

La Télé-Vision : codage, compression et correction

Où tout se réduit à des 0 et à des 1.

{Anne.Rasse,Jean-Marc.Vincent,Benjamin.Wack}@univ-grenoble-alpes.fr
{Maryline.Bruel,Herve.Barbe,Simon.Billouet}@ac-grenoble.fr

Maison Pour la Science : informatique débranchée



2017

REPRÉSENTATION DE L'INFORMATION

1 Représentation numérique de l'information

- LE PROBLÈME : représentation numérique de l'information
- CODAGE DES NOMBRES
- CODAGE DE TEXTES

2 Codes compressifs

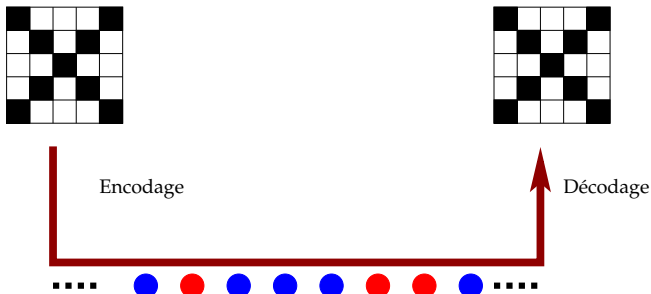
- COMPRESSER LES TEXTES
- COMPRESSER LES IMAGES

3 Codes correcteurs d'erreurs

- LE PROBLÈME : préserver l'information lors d'une transmission non fiable
- DÉBRANCHÉ : tour de magie
- CLÉ SIMPLE
- CODES LINÉAIRES

4 RÉFÉRENCES

INTRODUCTION AU PROBLÈME : TRANSMISSION D'UNE IMAGE SIMPLE



Démarche :

- ❶ représentation numérique d'une image (codage commun, ici pixels noir/blanc)
- ❷ protocole de communication (spécification)
 - algorithme de codage
 - algorithme de décodage
- ❸ test sur des exemples
- ❹ Preuve

UN MONDE NUMÉRIQUE

Numérique

- **Grand Robert 4. Techn.** Se dit de la représentation de données ou de grandeurs physiques sous forme de nombres (par oppos. à analogique) et par ext. de systèmes, dispositifs ou procédés employant ce mode de représentation. | Traitement numérique de l'information ; traitement numérique du son (→ Audionumérique ; D. A. T.), de l'image, du texte (→ Numériser). — Appareil photo, caméra, caméscope numérique. | Signal numérique : signal qui peut prendre en compte un nombre fini de valeurs discrètes. — Transmission numérique : transmission de signaux codés sous la forme de données binaires. | Télévision numérique. | Bouquet* numérique.
N. m. | Le numérique : l'ensemble des techniques de communication qui utilisent des signaux numériques, notamment dans la reproduction des images.
- **Trésor de la langue française informatisé A. MATH., INFORMAT.** Qui concerne des nombres, qui se présente sous la forme de nombres ou de chiffres, ou qui concerne des opérations sur des nombres. Méthode numérique ; donnée, échelle, évaluation, rapport, valeur numérique ; équation, expression numérique ; coefficient, exposant, notation numérique.

UN MONDE NUMÉRIQUE

Exemples

- ▶ image, son, video
- ▶ médecine
- ▶ automobile
- ▶ agriculture
- ▶ biologie
- ▶ ...

REPRÉSENTATION DE L'INFORMATION

1 Représentation numérique de l'information

- LE PROBLÈME : représentation numérique de l'information
- CODAGE DES NOMBRES
- CODAGE DE TEXTES

2 Codes compressifs

- COMPRESSER LES TEXTES
- COMPRESSER LES IMAGES

3 Codes correcteurs d'erreurs

- LE PROBLÈME : préserver l'information lors d'une transmission non fiable
- DÉBRANCHÉ : tour de magie
- CLÉ SIMPLE
- CODES LINÉAIRES

4 RÉFÉRENCES

CODAGE DES ENTIERS NATURELS

- ▶ Représentation du nombre (numération additive ou de position)
- ▶ choix d'une base (alphabet = ensemble de symboles)
- ▶ représentation *machine*
 - bit (*binary digit* atome d'information)
 - byte (mot machine élémentaire (de *bite*)) actuellement correspond à 8 bits
 - généralisation :
 - base octale (8),
 - base hexadécimale (16) {0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E, F}
- ▶ autres représentations (*redundant binary representation (RBR)*)

$$n = \sum_{i=0}^k d_i 2^i \text{ avec } d_i \in \{-1, 0, 1\}$$

Attention représentation finies de nombres

LES PUISSANCES DE 2

k	0	1	2	3	4	5	6	7	8	9	10	11	12
2^k	1	2	4	8	16	32	64	128	256	512	1024	2048	4096

$$2^{20} = 1048576$$

...

THE MAGIC TRICK

1	3	5	7	9	11	13	15
17	19	21	23	25	27	29	31
33	35	37	39	41	43	45	47
49	51	53	55	57	59	61	63

2	3	6	7	10	11	14	15
18	19	22	23	26	27	30	31
34	35	38	39	42	43	46	47
50	51	54	55	58	59	62	63

4	5	6	7	12	13	14	15
20	21	22	23	28	29	30	31
36	37	38	39	44	45	46	47
52	53	54	55	60	61	62	63

8	9	10	11	12	13	14	15
24	25	26	27	28	29	30	31
40	41	42	43	44	45	46	47
56	57	58	59	60	61	62	63

16	17	18	19	20	21	22	23
24	25	26	27	28	29	30	31
48	49	50	51	52	53	54	55
56	57	58	59	60	61	62	63

32	33	34	35	36	37	38	39
40	41	42	43	44	45	46	47
48	49	50	51	52	53	54	55
56	57	58	59	60	61	62	63

LES NOMBRES RÉELS

La norme IEEE définit la façon de coder un nombre réel sur 32 bits et définit trois composante

- ❶ Signe sur 1 bit
- ❷ Mantisse sur 23 bits
- ❸ Exposant sur 8 bits (biais = $2^7 - 1 = 127$)

La valeur d'un nombre est donnée par :

$$N = (-1)^s \left(1 + \sum_{i=1}^{23} m_i 2^{-i} \right) \cdot 2^e.$$

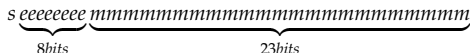
avec e calculé en supprimant le biais.

Conventions :

- ▶ Le 0 a deux codages différents : $0000 \dots 0$ ou $1000 \dots 0$
- ▶ NaN (not a number) est codé $0111 \dots 1$
- ▶ l'infini $+\infty$: $011111111000 \dots 0$ et $-\infty$: $111111111000 \dots 0$

EXEMPLE

Ainsi le codage d'un réel se fait sous la forme



Codage de la valeur 525, 5

- ▶ 525, 5 est positif donc le 1er bit s sera 0
- ▶ Sa représentation en base 2 est : 1000001101, 1 ;
- ▶ En normalisant on trouve 1, 0000011011.2⁹ ;
- ▶ On ajoute 127 à l'exposant qui vaut 9 ce qui donne 136, soit en base 2 : 10001000
- ▶ La mantisse est composée de la partie décimale de 525, 5 en base 2 normalisée, c. à.d 0000011011
- ▶ on complète la mantisse en ajoutant des 0 à droite pour qu'il occupe 23 bits, soit 01000100000000110110000000000000.

REPRÉSENTATION DE L'INFORMATION

1 Représentation numérique de l'information

- LE PROBLÈME : représentation numérique de l'information
- CODAGE DES NOMBRES
- CODAGE DE TEXTES

2 Codes compressifs

- COMPRESSER LES TEXTES
- COMPRESSER LES IMAGES

3 Codes correcteurs d'erreurs

- LE PROBLÈME : préserver l'information lors d'une transmission non fiable
- DÉBRANCHÉ : tour de magie
- CLÉ SIMPLE
- CODES LINÉAIRES

4 RÉFÉRENCES

BINAIRE, VOUS AVEZ DIT BINAIRE...

Description de chaque trigramme [[modifier](#) | [modifier le code](#)]

Trigramme	Sinogramme	Pinyin	unicode	Image naturelle	Qualités	Autres images
☰	乾	qián	☰ U+2630	le Ciel 天	Créativité, force, initiative	Le créateur , le cheval (bon, vieux, maigre, sauvage), le père, la tête, le rond, le prince, le jade, le métal, le froid le glace, le rouge sombre, un fruit...
☷	坤	kūn	☷ U+2637	La Terre 地	Disponibilité, adaptabilité, accueil, don de soi	Le réceptif , la vache, la mère, le ventre, une étoffe, un chaudron, l'économie, l'égalité, le veau avec la vache, un grand char, la multitude, le tronc, le sol noir parmi les autres...
☳	震	zhèn	☳ U+2633	Le Tonnerre 雷	Impulsion, mise en route, secousse	L'éveilleur , le dragon, le fils aîné, le pied, jaune sombre, une grande rue, un roseau ou un jonc...
☴	巽	xùn	☴ U+2634	Le Vent, le Bois 風	Pénétration, soumission, intériorisation	Le doux , le coq, la fille aînée, les cuisses, le corbeau, le travail, le blanc, le long, le haut, l'indécis...
☲	離	lí	☲ U+2632	Le Feu 火	Clarté, lucidité, vivacité, éclat	Ce qui s'attache , le faisan, la fille cadette, l'œil, le brillant, la cuirasse et le casque, la lance et les armes, la sécheresse, la tortue, le crabe, l'escargot, l'arbre desséché dans sa partie haute...
☵	坎	kǎn	☵ U+2635	L'Eau 水	Profondeur, endurance, peur	L'insondable , le porc, le fils cadet, l'oreille, les fosses, les pièges, l'arc et la flèche, le sang, le rouge, la lune, le bois ferme avec beaucoup de marques...
☶	艮	gèn	☶ U+2636	La Montagne 山	Rigueur, cohésion, calme, solidité	L'immobilisation , le chien, la main, le 3 ^e /le plus jeune fils, le chemin détourné, les pierres, les portes, les fruits, les semences, le bois ferme et noueux...
☴	兌	duì	☴ U+2631	Le Marais 澤	Apptitude à l'expression et à la communication, joie, légèreté	Le joyeux , le mouton, la 3 ^e /la plus jeune fille, la bouche (& la langue), la magicienne, écraser briser en morceaux, la voisine, le sol dur et dallé...

Yi-Jing *Traité canonique des mutations* (~ -1000 av. JC)



source wikipedia

CODAGE DE CARACTÈRES

Un codage historique : ASCII (American Standard Coding for Information Interchange)

- ▶ 128 caractères (jeu minimal) : 33 non imprimables, 94 imprimables, et l'espace (considéré comme invisible)
- ▶ Des limitations claires (US-ASCII), liées à son histoire
- ▶ Jusque fin 2007, l'encodage le plus répandu (UTF-8 depuis)
- ▶ 32 premiers caractères de contrôle (pour des machines comme les imprimantes)
 - Caractère 10 : LF (Line Feed), permet de faire avancer le papier dans une imprimante)
 - Caractère 8 : BS (backspace, retour arrière)
- ▶ Pas de description de la structure d'une page (langage de balisage)
- ▶ Beaucoup d'extensions et de propositions nationales pour utiliser le huitième bit et coder les caractères suivants
- ▶ Certaines extensions non compatibles avec le jeu minimal !

TABLE DES CARACTÈRES ASCII

ASCII TABLE

Decimal	Hex	Char	Decimal	Hex	Char	Decimal	Hex	Char	Decimal	Hex	Char
0	0	[NULL]	32	20	[SPACE]	64	40	@	96	60	`
1	1	[START OF HEADING]	33	21	!	65	41	A	97	61	a
2	2	[START OF TEXT]	34	22	"	66	42	B	98	62	b
3	3	[END OF TEXT]	35	23	#	67	43	C	99	63	c
4	4	[END OF TRANSMISSION]	36	24	\$	68	44	D	100	64	d
5	5	[ENQUIRY]	37	25	%	69	45	E	101	65	e
6	6	[ACKNOWLEDGE]	38	26	&	70	46	F	102	66	f
7	7	[BELL]	39	27	'	71	47	G	103	67	g
8	8	[BACKSPACE]	40	28	(72	48	H	104	68	h
9	9	[HORIZONTAL TAB]	41	29)	73	49	I	105	69	i
10	A	[LINE FEED]	42	2A	*	74	4A	J	106	6A	j
11	B	[VERTICAL TAB]	43	2B	+	75	4B	K	107	6B	k
12	C	[FORM FEED]	44	2C	,	76	4C	L	108	6C	l
13	D	[CARRIAGE RETURN]	45	2D	-	77	4D	M	109	6D	m
14	E	[SHIFT OUT]	46	2E	.	78	4E	N	110	6E	n
15	F	[SHIFT IN]	47	2F	/	79	4F	O	111	6F	o
16	10	[DATA LINK ESCAPE]	48	30	0	80	50	P	112	70	p
17	11	[DEVICE CONTROL 1]	49	31	1	81	51	Q	113	71	q
18	12	[DEVICE CONTROL 2]	50	32	2	82	52	R	114	72	r
19	13	[DEVICE CONTROL 3]	51	33	3	83	53	S	115	73	s
20	14	[DEVICE CONTROL 4]	52	34	4	84	54	T	116	74	t
21	15	[NEGATIVE ACKNOWLEDGE]	53	35	5	85	55	U	117	75	u
22	16	[SYNCHRONOUS IDLE]	54	36	6	86	56	V	118	76	v
23	17	[ENG OF TRANS. BLOCK]	55	37	7	87	57	W	119	77	w
24	18	[CANCEL]	56	38	8	88	58	X	120	78	x
25	19	[END OF MEDIUM]	57	39	9	89	59	Y	121	79	y
26	1A	[SUBSTITUTE]	58	3A	:	90	5A	Z	122	7A	z
27	1B	[ESCAPE]	59	3B	;	91	5B	[123	7B	{
28	1C	[FILE SEPARATOR]	60	3C	<	92	5C	\	124	7C	
29	1D	[GROUP SEPARATOR]	61	3D	=	93	5D]	125	7D	}
30	1E	[RECORD SEPARATOR]	62	3E	>	94	5E	^	126	7E	~
31	1F	[UNIT SEPARATOR]	63	3F	?	95	5F	_	127	7F	[DEL]

Source Wikipedia

REPRÉSENTATION DE L'INFORMATION

1 Représentation numérique de l'information

- LE PROBLÈME : représentation numérique de l'information
- CODAGE DES NOMBRES
- CODAGE DE TEXTES

2 Codes compressifs

- COMPRESSER LES TEXTES
- COMPRESSER LES IMAGES

3 Codes correcteurs d'erreurs

- LE PROBLÈME : préserver l'information lors d'une transmission non fiable
- DÉBRANCHÉ : tour de magie
- CLÉ SIMPLE
- CODES LINÉAIRES

4 RÉFÉRENCES

TRANSMETTRE UN TEXTE AVEC UN MINIMUM DE COMMUNICATIONS

Un texte c'est quoi ?

- Suite de symboles pris dans un alphabet restreint
- On veut compresser **sans perte** (sinon fichier corrompu, perte d'information, programme faux...)

TRANSMETTRE UN TEXTE AVEC UN MINIMUM DE COMMUNICATIONS

Un texte c'est quoi ?

- Suite de symboles pris dans un alphabet restreint
- On veut compresser **sans perte** (sinon fichier corrompu, perte d'information, programme faux...)

En fait pas limité au texte : si l'alphabet est celui des octets, on peut compresser n'importe quel fichier !

TRANSMETTRE UN TEXTE AVEC UN MINIMUM DE COMMUNICATIONS

Un texte c'est quoi ?

- ▶ Suite de symboles pris dans un alphabet restreint
- ▶ On veut compresser **sans perte** (sinon fichier corrompu, perte d'information, programme faux...)

En fait pas limité au texte : si l'alphabet est celui des octets, on peut compresser n'importe quel fichier !

Trois grandes stratégies

- ▶ Idée principale : profiter des **fréquences** de chaque symbole (Huffman, cf activité de cet après-midi)
- ▶ Pour aller plus loin : changer d'**alphabet** (coder des groupes de 2 caractères consécutifs par exemple)
- ▶ Supprimer la **redundance** : dictionnaires (Lempel-Ziv, utilisé aussi dans les formats ZIP et GIF)

REPRÉSENTATION DE L'INFORMATION

1 Représentation numérique de l'information

- LE PROBLÈME : représentation numérique de l'information
- CODAGE DES NOMBRES
- CODAGE DE TEXTES

2 Codes compressifs

- COMPRESSER LES TEXTES
- COMPRESSER LES IMAGES

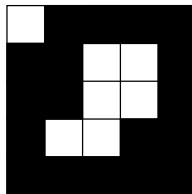
3 Codes correcteurs d'erreurs

- LE PROBLÈME : préserver l'information lors d'une transmission non fiable
- DÉBRANCHÉ : tour de magie
- CLÉ SIMPLE
- CODES LINÉAIRES

4 RÉFÉRENCES

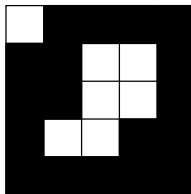
RUN LENGTH ENCODING

Constat : une image contient souvent de grandes zones uniformes (encore plus lorsque la palette est limitée)



RUN LENGTH ENCODING

Constat : une image contient souvent de grandes zones uniformes (encore plus lorsque la palette est limitée)

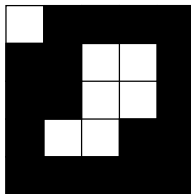


Principe du **Run Length Encoding** : au lieu de coder la suite des couleurs des pixels, on note successivement une couleur, puis le nombre d'occurrences de celle-ci, etc.

BNNNNNNBBNNBBNNBBNNNNNNNN \rightarrow 1B6N2B3N2B2N2B7N

RUN LENGTH ENCODING

Constat : une image contient souvent de grandes zones uniformes (encore plus lorsque la palette est limitée)



Principe du **Run Length Encoding** : au lieu de coder la suite des couleurs des pixels, on note successivement une couleur, puis le nombre d'occurrences de celle-ci, etc.

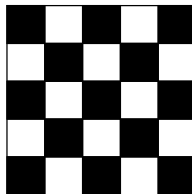
BNNNNNNBBNNBBNNBBNNNNNNN → 1B6N2B3N2B2N2B7N

Applications

- ▶ Utilisé dans les fax, le format BMP
- ▶ Adapté aux images monochromes (journal...)

RUN LENGTH ENCODING

Constat : une image contient souvent de grandes zones uniformes (encore plus lorsque la palette est limitée)



Principe du **Run Length Encoding** : au lieu de coder la suite des couleurs des pixels, on note successivement une couleur, puis le nombre d'occurrences de celle-ci, etc.

BNNNNNNBBNNBBNNBBNNNNNNN → 1B6N2B3N2B2N2B7N

Applications

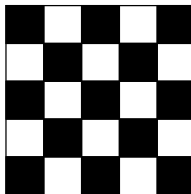
- ▶ Utilisé dans les fax, le format BMP
- ▶ Adapté aux images monochromes (journal...)

Propriétés et limitations

- ▶ Compression **sans perte**
- ▶ Sans risque... sauf de générer un fichier plus gros que l'original !

RUN LENGTH ENCODING

Constat : une image contient souvent de grandes zones uniformes (encore plus lorsque la palette est limitée)



Principe du **Run Length Encoding** : au lieu de coder la suite des couleurs des pixels, on note successivement une couleur, puis le nombre d'occurrences de celle-ci, etc.

BNNNNNNBBNNBBNNBBNNNNNNN \rightarrow 1B6N2B3N2B2N2B7N

Applications

- ▶ Utilisé dans les fax, le format BMP
- ▶ Adapté aux images monochromes (journal...)

Propriétés et limitations

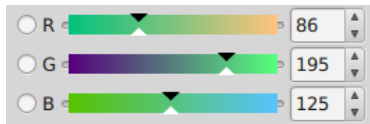
- ▶ Compression **sans perte**
- ▶ Sans risque... sauf de générer un fichier plus gros que l'original !

Version 2D : les quadrees

SOUS-ÉCHANTILLONNAGE

Codage des couleurs

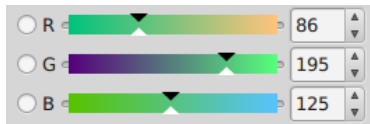
- En général 1 pixel = 3 valeurs (RVB par exemple) exprimées chacune sur 1 octet



SOUS-ÉCHANTILLONNAGE

Codage des couleurs

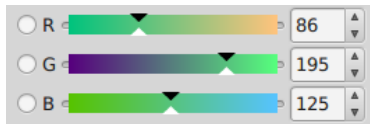
- ▶ En général 1 pixel = 3 valeurs (RVB par exemple) exprimées chacune sur 1 octet
- ▶ D'où $2^{24} \simeq 16$ millions de couleurs ;
l'œil humain n'en distingue que 300 000.



SOUS-ÉCHANTILLONNAGE

Codage des couleurs

- ▶ En général 1 pixel = 3 valeurs (RVB par exemple) exprimées chacune sur 1 octet
- ▶ D'où $2^{24} \simeq 16$ millions de couleurs ;
l'œil humain n'en distingue que 300 000.



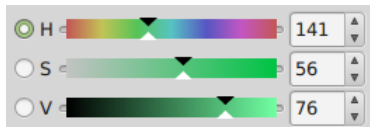
Compressions basées sur la couleur

- ▶ Sous-échantillonnage : assimiler des couleurs proches à une même couleur

SOUS-ÉCHANTILLONNAGE

Codage des couleurs

- ▶ En général 1 pixel = 3 valeurs (RVB par exemple) exprimées chacune sur 1 octet
- ▶ D'où $2^{24} \simeq 16$ millions de couleurs ;
l'œil humain n'en distingue que 300 000.



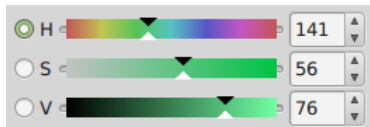
Compressions basées sur la couleur

- ▶ Sous-échantillonnage : assimiler des couleurs proches à une même couleur
- ▶ En fait on triche sur la chrominance (teinte), pas sur la luminance.
- ▶ On atténue aussi les variations de luminance très localisées

SOUS-ÉCHANTILLONNAGE

Codage des couleurs

- ▶ En général 1 pixel = 3 valeurs (RVB par exemple) exprimées chacune sur 1 octet
- ▶ D'où $2^{24} \simeq 16$ millions de couleurs ;
l'œil humain n'en distingue que 300 000.



Compressions basées sur la couleur

- ▶ Sous-échantillonnage : assimiler des couleurs proches à une même couleur
- ▶ En fait on triche sur la chrominance (teinte), pas sur la luminance.
- ▶ On atténue aussi les variations de luminance très localisées

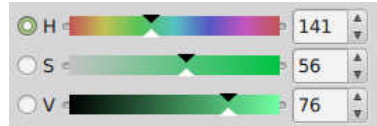
Applications

- ▶ utilisé en JPEG (en association avec d'autres techniques : RLE, Huffman)
- ▶ adapté aux photographies

SOUS-ÉCHANTILLONNAGE

Codage des couleurs

- ▶ En général 1 pixel = 3 valeurs (RVB par exemple) exprimées chacune sur 1 octet
- ▶ D'où $2^{24} \simeq 16$ millions de couleurs ;
l'œil humain n'en distingue que 300 000.



Compressions basées sur la couleur

- ▶ Sous-échantillonnage : assimiler des couleurs proches à une même couleur
- ▶ En fait on triche sur la chrominance (teinte), pas sur la luminance.
- ▶ On atténue aussi les variations de luminance très localisées

Applications

- ▶ utilisé en JPEG (en association avec d'autres techniques : RLE, Huffman)
- ▶ adapté aux photographies

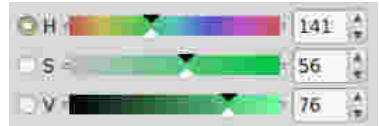
Propriétés et limitations

- ▶ risque de créer des artefacts autour des « bords »

SOUS-ÉCHANTILLONNAGE

Codage des couleurs

- ▶ En général 1 pixel = 3 valeurs (RVB par exemple) exprimées chacune sur 1 octet
- ▶ D'où $2^{24} \simeq 16$ millions de couleurs ;
l'œil humain n'en distingue que 300 000.



Compressions basées sur la couleur

- ▶ Sous-échantillonnage : assimiler des couleurs proches à une même couleur
- ▶ En fait on triche sur la chrominance (teinte), pas sur la luminance.
- ▶ On atténue aussi les variations de luminance très localisées

Applications

- ▶ utilisé en JPEG (en association avec d'autres techniques : RLE, Huffman)
- ▶ adapté aux photographies

Propriétés et limitations

- ▶ risque de créer des artefacts autour des « bords »
- ▶ crée des blocs visibles à l'œil nu si on cherche à trop compresser

INDEXATION DES COULEURS

Notion de palette

- Certaines images n'utilisent presque que des nuances d'une même teinte



INDEXATION DES COULEURS

Notion de palette

- Certaines images n'utilisent presque que des nuances d'une même teinte
- Dans d'autres à l'inverse c'est le contraste qui est important mais les nuances importent peu.



INDEXATION DES COULEURS

Notion de palette

- ▶ Certaines images n'utilisent presque que des nuances d'une même teinte
- ▶ Dans d'autres à l'inverse c'est le contraste qui est important mais les nuances importent peu.
- ▶ D'où l'idée de sélectionner un nombre limité de couleurs à utiliser (typiquement 256).



INDEXATION DES COULEURS

Notion de palette

- ▶ Certaines images n'utilisent presque que des nuances d'une même teinte
- ▶ Dans d'autres à l'inverse c'est le contraste qui est important mais les nuances importent peu.
- ▶ D'où l'idée de sélectionner un nombre limité de couleurs à utiliser (typiquement 256).



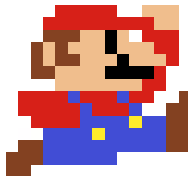
Une image indexée =

- ▶ 1 palette associant un octet à chaque couleur (décrite sur 3 octets)
- ▶ suivie de l'image elle-même mais avec 1 pixel = 1 octet

INDEXATION DES COULEURS

Notion de palette

- ▶ Certaines images n'utilisent presque que des nuances d'une même teinte
- ▶ Dans d'autres à l'inverse c'est le contraste qui est important mais les nuances importent peu.
- ▶ D'où l'idée de sélectionner un nombre limité de couleurs à utiliser (typiquement 256).



Une image indexée =

- ▶ 1 palette associant un octet à chaque couleur (décrite sur 3 octets)
- ▶ suivie de l'image elle-même mais avec 1 pixel = 1 octet

Applications

- ▶ utilisé en GIF ou en PNG
- ▶ adapté aux images de synthèse, aux graphiques, aux dessins

INDEXATION DES COULEURS

Notion de palette

- ▶ Certaines images n'utilisent presque que des nuances d'une même teinte
- ▶ Dans d'autres à l'inverse c'est le contraste qui est important mais les nuances importent peu.
- ▶ D'où l'idée de sélectionner un nombre limité de couleurs à utiliser (typiquement 256).



Une image indexée =

- ▶ 1 palette associant un octet à chaque couleur (décrite sur 3 octets)
- ▶ suivie de l'image elle-même mais avec 1 pixel = 1 octet

Applications

- ▶ utilisé en GIF ou en PNG
- ▶ adapté aux images de synthèse, aux graphiques, aux dessins

Propriétés et limitations

- ▶ Bon taux de compression mais seulement si l'image est assez grande
- ▶ Risque de créer de grands aplats uniformes artificiels

REPRÉSENTATION DE L'INFORMATION

1 Représentation numérique de l'information

- LE PROBLÈME : représentation numérique de l'information
- CODAGE DES NOMBRES
- CODAGE DE TEXTES

2 Codes compressifs

- COMPRESSER LES TEXTES
- COMPRESSER LES IMAGES

3 Codes correcteurs d'erreurs

- LE PROBLÈME : préserver l'information lors d'une transmission non fiable
- DÉBRANCHÉ : tour de magie
- CLÉ SIMPLE
- CODES LINÉAIRES

4 RÉFÉRENCES

EXEMPLES DE TAUX D'ERREUR

Types d'erreurs

- ▶ altération : modification d'un ou plusieurs bits
- ▶ insertion / suppression d'un ou plusieurs bits

Ordres de grandeur de taux d'erreurs

ligne	taux d'erreur
Disquette	10^{-9} : à 5 Mo/s, 3 bits erronés par minute
CD-ROM optique	10^{-5} : 7ko erronés sur un CD de 700 Mo
DAT audio	10^{-5} : à 48 kHz, deux erreurs par seconde
Mémoires à semi-conducteurs	$< 10^{-9}$
Liaison téléphonique	entre 10^{-4} et 10^{-7}
Télécommande infrarouge	10^{-12}
Communication par fibre optique	10^{-9}
Satellite	10^{-6} (Voyager), 10^{-11} (TDMA)
ADSL	10^{-3} à 10^{-9}
Réseau informatique	10^{-12}

TAB. 4.1: Ordre de grandeur du taux d'erreurs.

REPRÉSENTATION DE L'INFORMATION

1 Représentation numérique de l'information

- LE PROBLÈME : représentation numérique de l'information
- CODAGE DES NOMBRES
- CODAGE DE TEXTES

2 Codes compressifs

- COMPRESSER LES TEXTES
- COMPRESSER LES IMAGES

3 Codes correcteurs d'erreurs

- LE PROBLÈME : préserver l'information lors d'une transmission non fiable
- DÉBRANCHÉ : tour de magie
- CLÉ SIMPLE
- CODES LINÉAIRES

4 RÉFÉRENCES

INFORMATIQUE DÉBRANCHÉE

Activité à 3 : un magicien, un assistant et un participant

Sur une grille de 16 cartes disposées aléatoirement, le participant va en retourner une.
À l'aide d'une petite manipulation de l'assistant, le magicien sera capable de retrouver laquelle à coup sûr, **sans avoir vu la grille initiale**.

INFORMATIQUE DÉBRANCHÉE

Activité à 3 : un magicien, un assistant et un participant

Sur une grille de 16 cartes disposées aléatoirement, le participant va en retourner une. À l'aide d'une petite manipulation de l'assistant, le magicien sera capable de retrouver laquelle à coup sûr, **sans avoir vu la grille initiale**.

Plus fort encore, l'assistant va effectuer sa manipulation **avant que le participant ne choisisse sa carte**.

INFORMATIQUE DÉBRANCHÉE

Activité à 3 : un magicien, un assistant et un participant

Sur une grille de 16 cartes disposées aléatoirement, le participant va en retourner une. À l'aide d'une petite manipulation de l'assistant, le magicien sera capable de retrouver laquelle à coup sûr, **sans avoir vu la grille initiale**.

Plus fort encore, l'assistant va effectuer sa manipulation **avant que le participant ne choisisse sa carte**.

C'est de l'informatique

- ▶ La disposition initiale des cartes est un message à transmettre (qui peut être **presque** quelconque)
- ▶ Le magicien est le récepteur
- ▶ L'action du participant représente une erreur de transmission sur un canal non fiable (pas de perte ni d'ajout, 1 bit changé sur 16 au maximum)

INFORMATIQUE DÉBRANCHÉE

Activité à 3 : un magicien, un assistant et un participant

Sur une grille de 16 cartes disposées aléatoirement, le participant va en retourner une. À l'aide d'une petite manipulation de l'assistant, le magicien sera capable de retrouver laquelle à coup sûr, **sans avoir vu la grille initiale**.

Plus fort encore, l'assistant va effectuer sa manipulation **avant que le participant ne choisisse sa carte**.

C'est de l'informatique

- ▶ La disposition initiale des cartes est un message à transmettre (qui peut être **presque** quelconque)
- ▶ Le magicien est le récepteur
- ▶ L'action du participant représente une erreur de transmission sur un canal non fiable (pas de perte ni d'ajout, 1 bit changé sur 16 au maximum)

Correction d'erreur

Tous les messages ne sont pas valides ; en particulier :

- ▶ Celui préparé par l'assistant est valide
- ▶ Tout message dans lequel on a retourné un bit est invalide
- ▶ Et surtout, chaque message invalide ne peut correspondre qu'à **un** message valide

HYPERCUBES

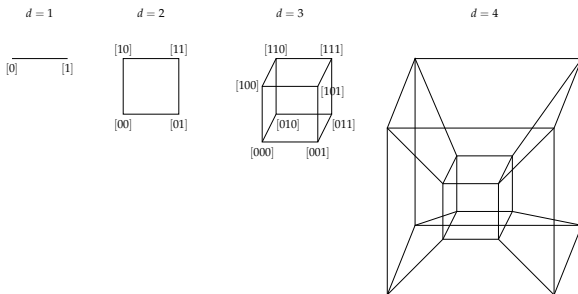
Espace des vecteurs de bits

Exercice : dessiner les hypercubes de dimension $d = 1, 2, 3, 4, \dots$

HYPERCUBES

Espace des vecteurs de bits

Exercice : dessiner les hypercubes de dimension $d = 1, 2, 3, 4, \dots$



Distance de Hamming

- Nombre de bits différents entre 2 vecteurs de bits. Calculer la taille de la boule à distance 1, 2, 3, ... d'un vecteur x de n bits.
- Exercice : calculer la **borne de Hamming**, capacité d'un code de longueur n à corriger 1 bit erroné.

REPRÉSENTATION DE L'INFORMATION

1 Représentation numérique de l'information

- LE PROBLÈME : représentation numérique de l'information
- CODAGE DES NOMBRES
- CODAGE DE TEXTES

2 Codes compressifs

- COMPRESSER LES TEXTES
- COMPRESSER LES IMAGES

3 Codes correcteurs d'erreurs

- LE PROBLÈME : préserver l'information lors d'une transmission non fiable
- DÉBRANCHÉ : tour de magie
- CLÉ SIMPLE
- CODES LINÉAIRES

4 RÉFÉRENCES

CONTRÔLE DE PARITÉ

Une technique de base pour construire un code détecteur

- ❶ Découper le message en mots de 7 bits $m = [x_0, \dots, x_6]$
- ❷ Ajouter aux mots leur parité : $f(m) = [x_0, \dots, x_6, p]$

Le nombre de 1 dans le mot est soit pair ($p = 0$) soit impair ($p = 1$)

$$p \stackrel{\text{def}}{=} \sum_{i=0}^6 x_i \text{ modulo } 2.$$

Standard n5 du Comité Consultatif International Télégraphique et Téléphonique (CCITT 5)

le plus populaire et celui utilisé par exemple aux USA.

Lettre	Codage de base sur 7 bits	Mot de code avec bit de parité
a	1000 001	1000 001 0
e	1010 001	1010 001 1
u	0110 101	0110 101 0

Détecte un nombre impair d'erreurs

CODES ACTUELS

Numéro de Sécurité Sociale

Un numéro de sécurité sociale est un nombre de $n = 15$ chiffres : un numéro d'identification K sur $k = 13$ chiffres suivi de la clé C de $r = 2$ chiffres calculée pour que $K + C$ soit un multiple de 97.

- ❶ Quelle est la clé du numéro de sécurité sociale 2.63.05.38.516.305 ?
- ❷ Quel est le rendement de ce code ?
- ❸ Combien d'erreurs de chiffres, la clé du numéro de sécurité sociale permet-elle de détecter ?

Formule de Luhn pour les cartes bancaires

Carte bancaire : 4 nombres de 4 chiffres

- ▶ Pour les chiffres de rang pair (le premier chiffre est de rang 0) on double ce chiffre modulo 9
- ▶ On additionne ces chiffres aux chiffres de rang impair.

Le résultat doit être divisible par 10

Exercice : vérifier sur votre carte bancaire, comment calculer la clé (dernier chiffre).

International Standard Book Number (ISBN)

Exercice : chercher la clé.

PARITÉ LONGITUDINALE ET TRANSVERSALE

Encore un tour de cartes :

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

Exercice : Calculer le rendement de ce type de code, calculer sa capacité de détection, de correction, calculer la probabilité d'avoir une erreur non détectée.

PARITÉ LONGITUDINALE ET TRANSVERSALE

Encore un tour de cartes :

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

Exercice : Calculer le rendement de ce type de code, calculer sa capacité de détection, de correction, calculer la probabilité d'avoir une erreur non détectée.

PARITÉ LONGITUDINALE ET TRANSVERSALE

Encore un tour de cartes :

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

Exercice : Calculer le rendement de ce type de code, calculer sa capacité de détection, de correction, calculer la probabilité d'avoir une erreur non détectée.

PARITÉ LONGITUDINALE ET TRANSVERSALE

Encore un tour de cartes :

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

Exercice : Calculer le rendement de ce type de code, calculer sa capacité de détection, de correction, calculer la probabilité d'avoir une erreur non détectée.

REPRÉSENTATION DE L'INFORMATION

1 Représentation numérique de l'information

- LE PROBLÈME : représentation numérique de l'information
- CODAGE DES NOMBRES
- CODAGE DE TEXTES

2 Codes compressifs

- COMPRESSER LES TEXTES
- COMPRESSER LES IMAGES

3 Codes correcteurs d'erreurs

- LE PROBLÈME : préserver l'information lors d'une transmission non fiable
- DÉBRANCHÉ : tour de magie
- CLÉ SIMPLE
- CODES LINÉAIRES

4 RÉFÉRENCES

CODE DE HAMMING

Code 1-correcteur à nombre de bits ajoutés minimal, $\delta = 3$

Idée : ajouter un contrôle de parité pour chaque puissance de 2 : b_1, b_2, b_4, \dots
 \Rightarrow localisation de l'erreur

Exemples

- ▶ Code de Hamming(4,7) : compléter $x_1x_2x_3x_4$ avec $x_2 + x_3 + x_4$, $x_1 + x_3 + x_4$, $x_1 + x_2 + x_4$
 \Rightarrow rendement $\simeq 57\%$
- ▶ Minitel(136,120) = Hamming(128,120)
1-correcteur avec un ajout de 8 bits toujours à 0, pour les grosses erreurs.
 \Rightarrow taux d'erreur = 10^{-4} ; rendement $\simeq 88\%$; bits ajoutés = 16

RETOUR SUR LE PREMIER TOUR DE MAGIE

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

Dans un tableau valide :

RETOUR SUR LE PREMIER TOUR DE MAGIE

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

Dans un tableau valide :

- La somme des colonnes 1 et 2 est paire

RETOUR SUR LE PREMIER TOUR DE MAGIE

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

Dans un tableau valide :

- La somme des colonnes 1 et 2 est paire
- La somme des colonnes 2 et 3 est paire

RETOUR SUR LE PREMIER TOUR DE MAGIE

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

Dans un tableau valide :

- La somme des colonnes 1 et 2 est paire
- La somme des colonnes 2 et 3 est paire
- La somme des lignes 1 et 2 est paire

RETOUR SUR LE PREMIER TOUR DE MAGIE

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

Dans un tableau valide :

- La somme des colonnes 1 et 2 est paire
- La somme des colonnes 2 et 3 est paire
- La somme des lignes 1 et 2 est paire
- La somme des lignes 2 et 3 est paire

RETOUR SUR LE PREMIER TOUR DE MAGIE

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

Dans un tableau valide :

- La somme des colonnes 1 et 2 est paire
- La somme des colonnes 2 et 3 est paire
- La somme des lignes 1 et 2 est paire
- La somme des lignes 2 et 3 est paire

Correction

- Les 2 premiers critères permettent d'identifier la colonne erronée
- Les 2 derniers identifient la ligne

RETOUR SUR LE PREMIER TOUR DE MAGIE

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

Dans un tableau valide :

- ▶ La somme des colonnes 1 et 2 est paire
- ▶ La somme des colonnes 2 et 3 est paire
- ▶ La somme des lignes 1 et 2 est paire
- ▶ La somme des lignes 2 et 3 est paire

Correction

- ▶ Les 2 premiers critères permettent d'identifier la colonne erronée
- ▶ Les 2 derniers identifient la ligne

Tout tableau erroné est à distance exactement 1 d'un (et un seul) tableau valide.

REPRÉSENTATION DE L'INFORMATION

1 Représentation numérique de l'information

- LE PROBLÈME : représentation numérique de l'information
- CODAGE DES NOMBRES
- CODAGE DE TEXTES

2 Codes compressifs

- COMPRESSER LES TEXTES
- COMPRESSER LES IMAGES

3 Codes correcteurs d'erreurs

- LE PROBLÈME : préserver l'information lors d'une transmission non fiable
- DÉBRANCHÉ : tour de magie
- CLÉ SIMPLE
- CODES LINÉAIRES

4 RÉFÉRENCES

CODAGE

- ▶ Le site web [Interstice](#) est une mine de renseignements, parfois très détaillés.
Voir en particulier :
 - [Interstice : Tout a un reflet numérique](#)
 - [Nom de code : binaire](#)
- ▶ *Introduction à la science informatique : Pour les enseignants de la discipline en lycée.* Gilles Dowek. et al CRDP Académie de Paris, 2011. Chapitre 1 Représentation numérique de l'information
- ▶ **Introduction aux sciences de l'information**, Jean-Yves Le Boudec, Patrick Thiran et Rüdiger Urbanke, Presses polytechniques et universitaires romandes, 2015
- ▶ *Informatique et sciences du numérique Spécialité ISN en terminale S. Avec des exercices corrigés et des idées de projets.* Gilles Dowek. et al Eyrolles 2012 Chapitres 7-8
- ▶ **Histoires d'algorithmes** Jean-Luc Chabert et al. Belin (2010)

CORRECTION

- ▶ Exposé de Jean-Guillaume Dumas à MidiSciences [Codes détecteurs et correcteurs d'erreurs](#)
- ▶ **Introduction aux sciences de l'information**, Jean-Yves Le Boudec, Patrick Thiran et Rüdiger Urbanke, Presses polytechniques et universitaires romandes, 2015
- ▶ **Théorie des codes**, J-G. Dumas, J-L. Roch, E. Tannier et S. Varrette, Dunod 2007, [Site](#)
- ▶ **Introduction aux codes correcteurs** Pierre Csillag, Ellipses 1990
- ▶ **L'information : L'histoire - La théorie - Le déluge** James Gleick, Cassini 2015