Size matters: The Importance of Image Compression

Fernando Gastón

Research Technician @ CRM

Data Days 2024

CENTRE DE RECERCA MATEMÀTICA

EXCELENCIA MARÍA DE MAEZTU 2022 - 2025

Recordatori: el codi binari de 8 bits (1 byte) ens permet representar nombre entre 0 i 255 per tal d'emmagatzemar-los a un ordinador

	Codi binari
0	00000000
1	00000001
2	00000010
3	00000011
4	00000100
5	00000101
	…
254	11111110
255	11111111

229	229	217	217	240	240	240	240
229	229	217	217	240	240	240	240
229	217	217	217	217	240	240	240
217	217	217	217	217	217	240	240
217	217	179	179	217	217	194	194
217	179	179	179	194	194	194	194
217	179	179	179	179	194	194	194
217	217	194	194	194	194	194	194

→

1 byte	1 byte	1 byte	1 byte	1 byte	1 byte	1 byte	1 byte
1 byte	1 byte	1 byte	1 byte	1 byte	1 byte	1 byte	1 byte
1 byte	1 byte	1 byte	1 byte	1 byte	1 byte	1 byte	1 byte
1 byte	1 byte	1 byte	1 byte	1 byte	1 byte	1 byte	1 byte
1 byte	1 byte	1 byte	1 byte	1 byte	1 byte	1 byte	1 byte
1 byte	1 byte	1 byte	1 byte	1 byte	1 byte	1 byte	1 byte
1 byte	1 byte	1 byte	1 byte	1 byte	1 byte	1 byte	1 byte
1 byte	1 byte	1 byte	1 byte	1 byte	1 byte	1 byte	1 byte

Per què és important comprimir?

Un pixel és 1byte

Un frame d’un video 1080p60 (1920x1080 pxls) són 2 073 600 bytes=2MB!

Un segon del vídeo (60 frames) són 120MB!!

Un video de youtube d’una hora (3600s) són 432GB!!!

CENTRE DE RECERCA MATEMÀTICA

EXCELENCIA MARÍA DE MAEZTU 2022 - 2025

La entropia

una imatge són mil paraules?

Distribució de probabilitat

Recordem: La distribució de probabilitat d'una variable aleatòria que prèn valors (símbols) en un conjunt Ω és una funció $P : \Omega \to \mathbb{R}_{\geq 0}$ tal que:

$$\sum_{x \in \Omega} P(x) = 1$$

➔ El llançament d'un dau:

$$\Omega = \{1, 2, 3, 4, 5, 6\} \qquad P(x) = \tfrac{1}{6} \quad \forall x \in \Omega$$

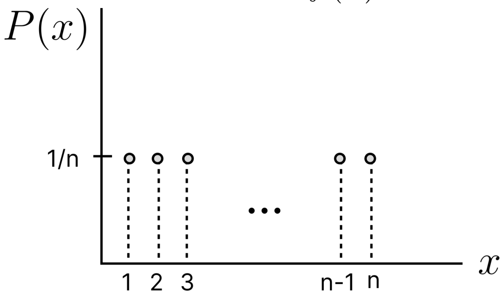
➔ Pel d'una moneda:

$$\Omega = \{\text{Cara}, \text{Creu}\} \qquad P(x) = \tfrac{1}{2} \quad \forall x \in \Omega$$

Algunes distribucions ben conegudes...

- ➔ Uniforme: tenim n valors equiprobables (llançament d'un dau)
- ➔ Bernouilli: un 0 o 1 amb probabilitat 1-p i p (llançament d'una moneda, p=0.5)
- ➔ Geomètrica: quantes tirades d'una Bernouilli em calen per aconseguir un 1

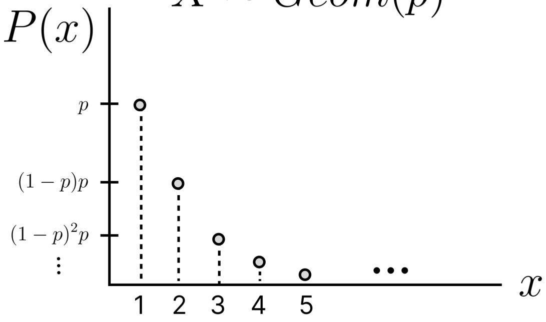
$$X \sim Unif(n)$$



$$P(x) = 1/n$$

$$\Omega = \{1, 2, \ldots, n\}$$

$$X \sim Geom(p)$$



$$P(x) = (1-p)^{x-1}p$$

$$\Omega = \mathbb{N}$$

Si em diuen quin valor prendrà una variable aleatòria, quanta informació m'estan donant?

Si em diuen quin valor prendrà una variable aleatòria, quanta informació m’estan donant?

Depèn de la seva distribució de probabilitat!

Per exemple, si al Sahara em diuen que demà no plourà (quan sabem que quasi mai plou), no m’estan donant gaire informació…

Com formalitzem matemàticament aquesta intuició?

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

CENTRE DE RECERCA MATEMÀTICA

EXCELENCIA MARÍA DE MAEZTU 2022 - 2025

L’entropia de Shannon

L’entropia ens diu la “informació” que esperem obtenir d’una variable aleatòria $X = (\Omega, P)$:

$$H(X) = -\sum_{x \in \Omega} P(x) \log P(x)$$

per exemple:

X	$\mathrm{H}(X)$
$\mathrm{Bernouilli}(p)$	$-p\log(p) - (1-p)\log(1-p)$
$\mathrm{Unif}(n)$	$\log\ (n)$
$\mathrm{Geom}(p)$	$\dfrac{-p\log(p) - (1-p)\log(1-p)}{p}$

CENTRE DE RECERCA MATEMÀTICA

EXCELENCIA MARÍA DE MAEZTU 2022 - 2025

Codificar els valors de la variable aleatòria?

Volem trobar un codi que representi els símbols de Ω amb una seqüència de 0s i 1s de la forma més curta possible.

Per exemple: pels possibles valors d'un dau

Ω	Codi
1	00000001
2	00000010
3	00000011
4	00000100
5	00000101
6	00000110

Ω	Codi
1	001
2	010
3	011
4	100
5	101
6	110

Ω	Codi
1	1
2	10
3	11
4	100
5	101
6	110

1011 = (2 1 1) o (5 1)
11 = (3) o (1 1)

NO es pot decodificar!

Codificar els valors de la variable aleatòria?

Volem trobar un codi que representi els símbols de Ω amb una seqüència de 0s i 1s de la forma més curta possible.

Per exemple: pels possibles valors d'un dau

Ω	Codi
1	00000001
2	00000010
3	00000011
4	00000100
5	00000101
6	00000110

Ω	Codi
1	001
2	010
3	011
4	100
5	101
6	110

Ω	Codi
1	1
2	01
3	001
4	0001
5	00001
6	000001

Ω	Codi
1	00
2	01
3	100
4	1010
5	1011
6	11

Codificar els valors de la variable aleatòria?

Volem trobar un codi que representi els símbols de Ω amb una seqüència de 0s i 1s de la forma més curta possible.

Per exemple: pels possibles valors d'un dau

Ω	Codi
1	
2	
3	8b/simbol
4	
5	
6	

Ω	Codi
1	
2	
3	3b/simbol
4	
5	
6	

Ω	Codi
1	
2	
3	3.5b/simbol
4	
5	
6	

Ω	Codi
1	
2	
3	2.8b/simbol
4	
5	
6	

El teorema de l'entropia de Shannon

Informalment: L'entropia d'una distribució dona un límit inferior per el nombre mitjà de bits que cal per codificar una seqüència i.i.d. de símbols ***sense pèrdua d'informació***.

Pel llançament d'un dau l'entropia és ~ 2.5 per tant no podem fer servir menys de 2.5 bits per símbol!

CENTRE DE RECERCA MATEMÀTICA

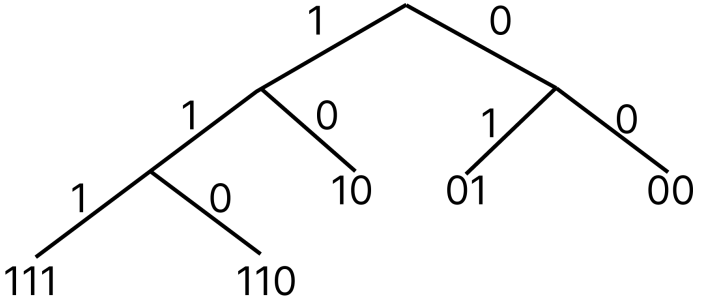
EXCELENCIA MARÍA DE MAEZTU 2022 - 2025

Codis de Huffman

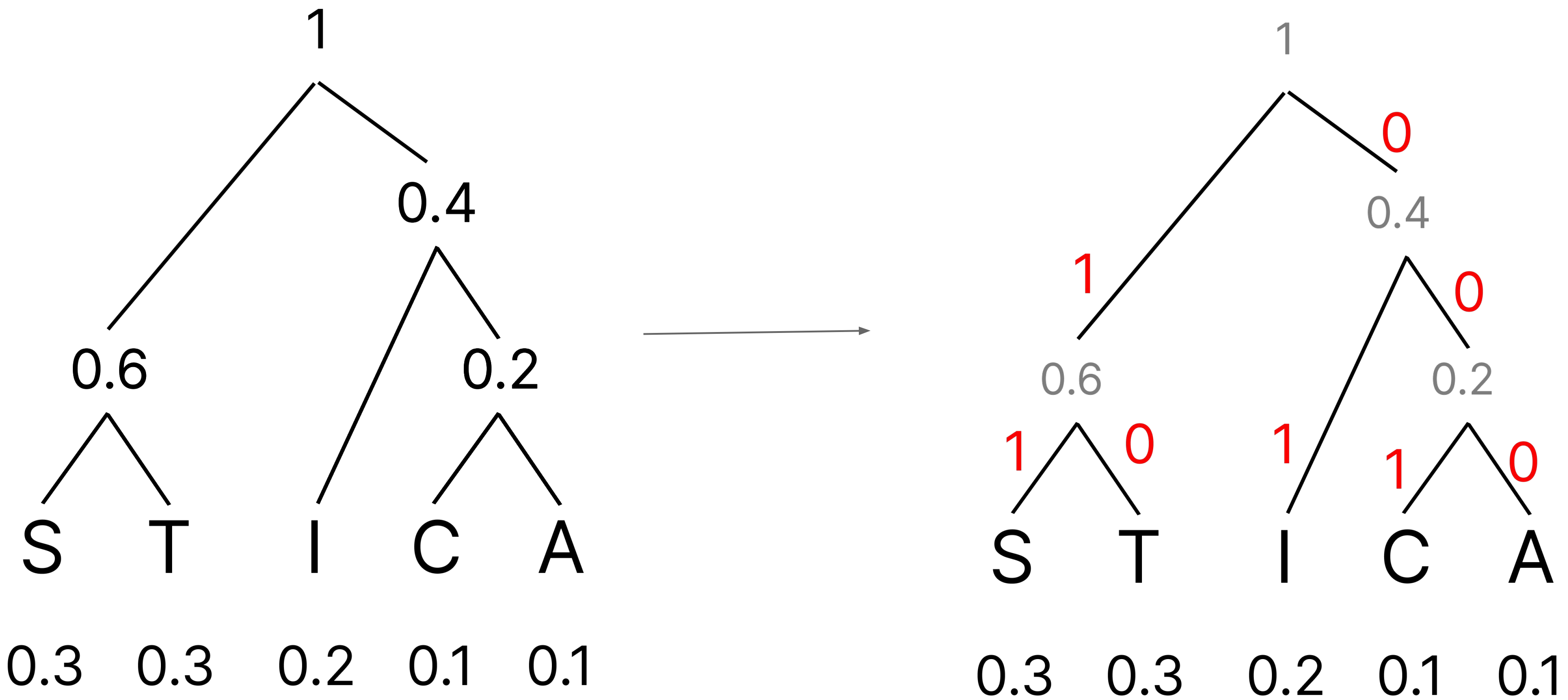
L'algorisme proposat per Huffman dona un codi òptim* (el més proper a l'entropia) per codificar símbol a símbol.

* (quan les probabilitats de la distribució són de la forma 1/2^k)

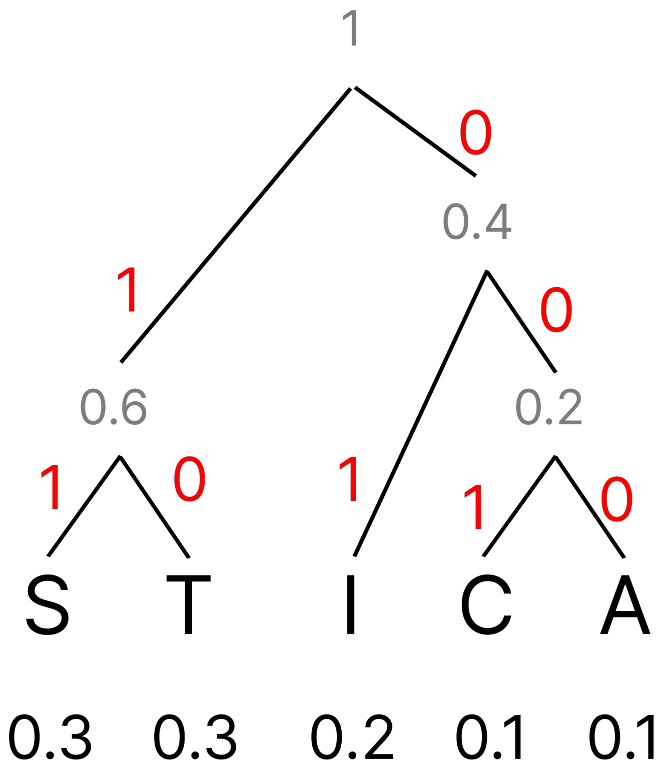
Es basa en generar un arbre, amb els símbols a les fulles, on s'assignarà a cada símbol un codi en funció del recorregut de l'arrel fins a la seva fulla.



Codis de Huffman: STATISTICS



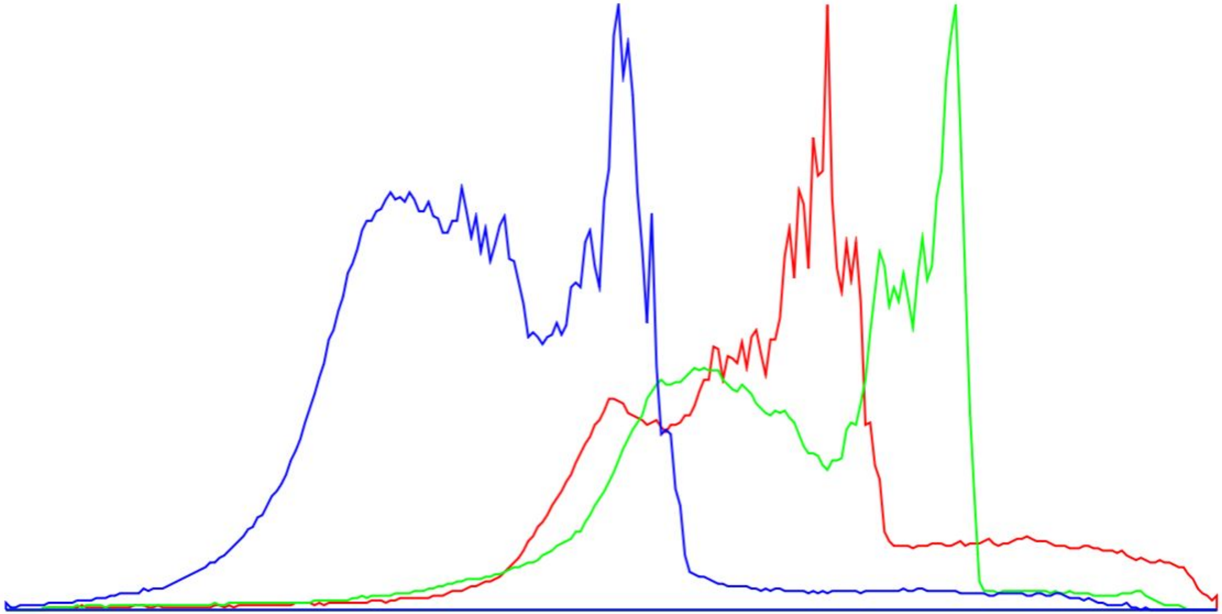
Codis de Huffman: STATISTICS

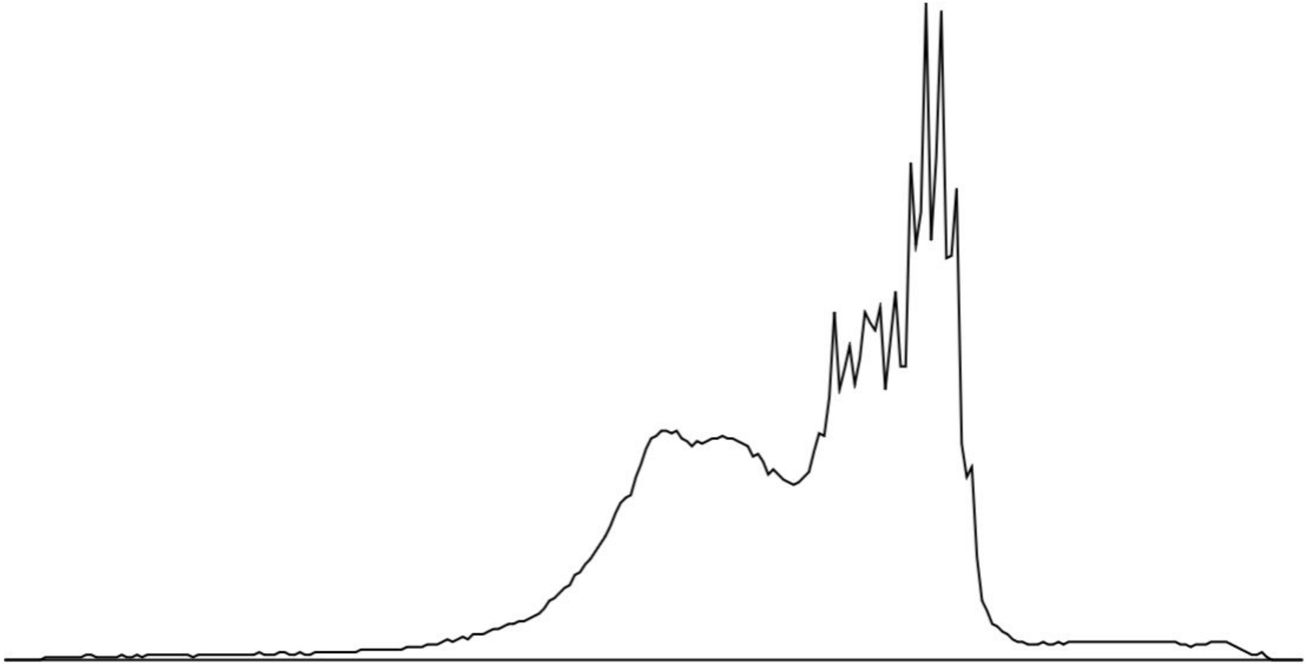


Ω	S	T	A	I	C
	11	10	000	01	001

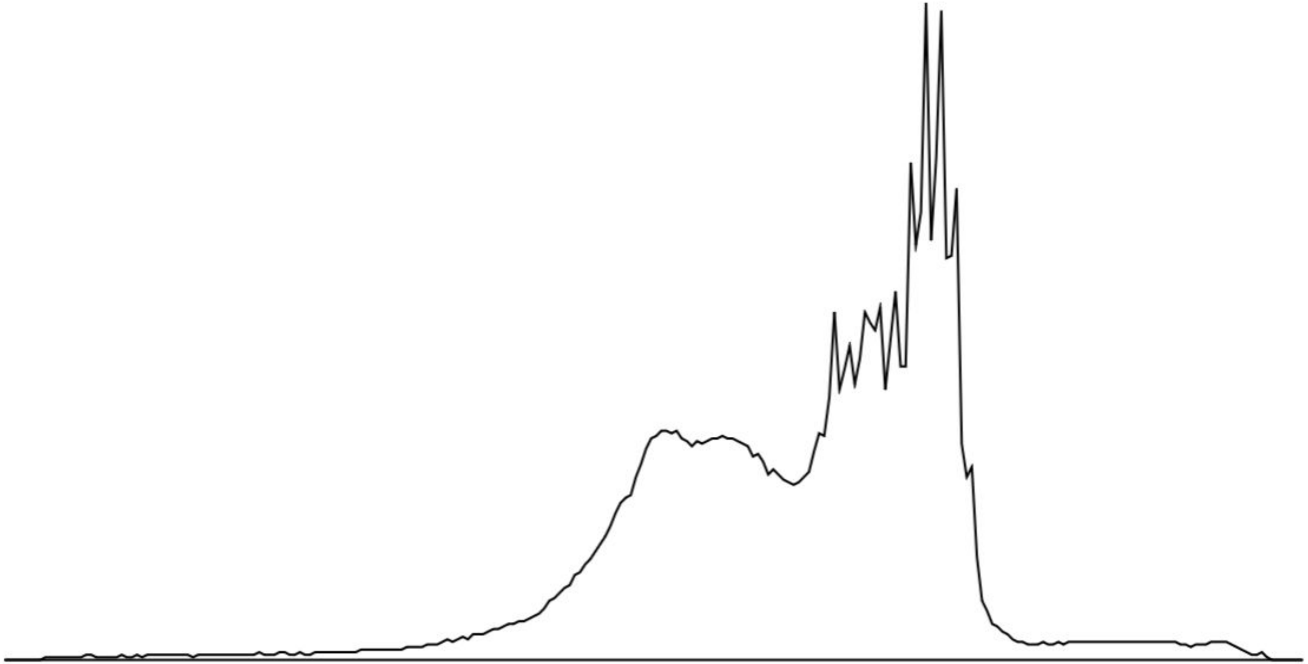
…i això té a veure amb les nostres imatges?

Podem prendre els valors dels pixels de la nostra imatge com una variable aleatòria, calcular la seva **entropia** i mirar d'aplicar **Huffman**!





H(gos)=7.8 bits/symbol

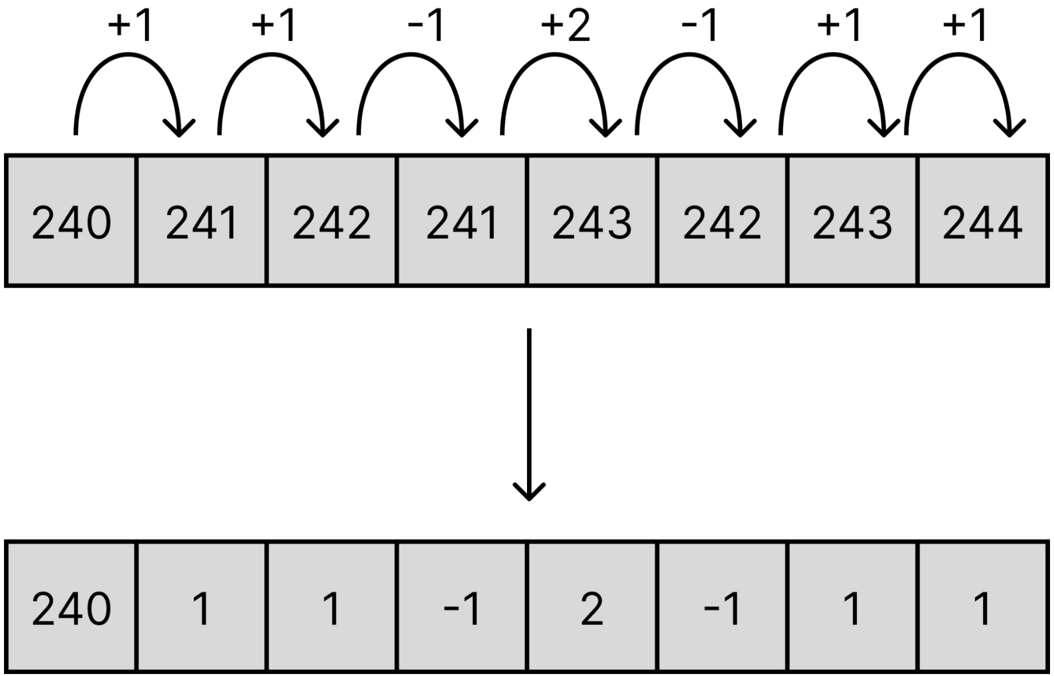


H(gos)=7.8 bits/pixel

CENTRE DE RECERCA MATEMÀTICA

EXCELENCIA MARÍA DE MAEZTU 2022 - 2025

Predictive coding



Podem fer-ho millor d'alguna forma?

[*https://github.com/fergascod/Image-Compression-Workshop*](https://github.com/fergascod/Image-Compression-Workshop)

Taller

- Puja la teva pròpia imatge!
  - Intenta aconseguir una imatge molt compressible i una altra que no ho sigui tant i mira com canvia la entropia
- Prova altres distribucions per les imatges de “soroll” (soroll Gaussià vs. Uniforme)
- Intenta trobar altres funcions per fer predictive coding (e.g. diferencies en y, mitjana dels pixels anteriors…)
- Prova idees per comprimir video (e.g. diferències entre frames)

CENTRE DE RECERCA MATEMÀTICA

EXCELENCIA MARÍA DE MAEZTU 2022 - 2025