



# Sisteme avansate de codare și compresie a datelor multimedia

## Curs 3 – Introducere în compresia informațiilor

Sl.Dr.Ing. Camelia FLOREA

Intelligent and multimodal image processing and analysis group (IMIPA),  
Communications Departament, ETTI, TUCN,

E-mail: [Camelia.Florea@com.utcluj.ro](mailto:Camelia.Florea@com.utcluj.ro); Phone: +4 0264 401285;

Address: C. Daicoviciu, 15, room 433, Cluj-Napoca, RO.



# Introducere

---

- **Compresia**

- reprezentare digitala compacta a informației multimedia
- cu aplicații în special:
  - în transmisia informației
  - stocarea datelor/imaginilor/secvențelor video
  - aplicații multimedia

- Fără reprezentarea comprimată

- multe dintre aplicațiile multimedia nu ar putea fi implementate practic.

# Introducere – concepte de bază

- **Compresia**

- reducerea cantității de date necesare pentru reprezentarea unei informații

$$C = \frac{b}{b'}$$

$$R = 1 - \frac{1}{C}$$

- **C – factorul de compresie**

- $b$  – număr biți reprezentare informație în forma originală

- $b'$  – număr biți reprezentare informație după compresie

- R - redundanța relativă

- **Exemplificare:**

- $C=10$  rezultă  $R=0.9$  adică 90% din  $b$  sunt date redundante

- **Simbol,  $s_i$**

- o sursă  $S$  de simboluri aleatoare  $s_1, s_2, \dots, s_N$ .
  - $S$  – semnalul - poate fi o imagine digitală
  - $s_i$  reprezintă una din cele  $N$  valori posibile pe care le poate lua un pixel din imagine

- **Informația,  $I(s_i)$**

= numărul de biți necesari pentru codarea simbolului  $s_i$

- dacă baza logaritmului = 2  $\Rightarrow$  informația se măsoară în *biți*.
- dacă baza logaritmului = 10  $\Rightarrow$  informația se măsoară în *digiti*.

$$I(s_i) = \log \frac{1}{p_i} = -\log p_i \rightarrow p$$

- **Entropia,  $H(S)$**

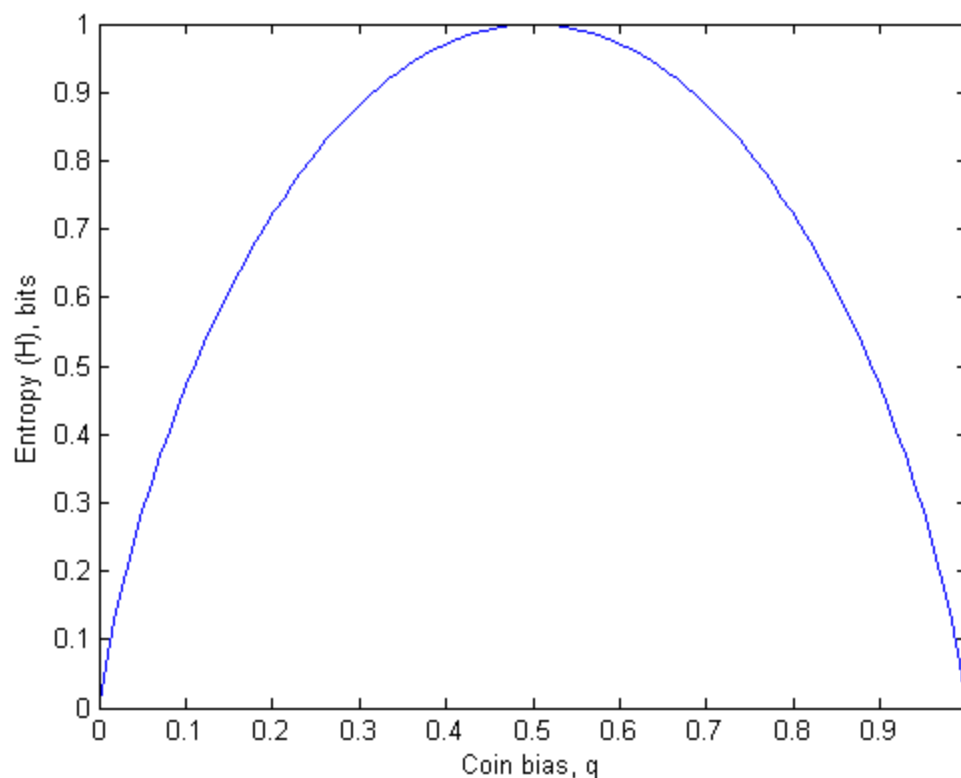
= informația medie pe simbol ( $\Leftrightarrow$  numărul mediu de biți pe simbol)

$$H(S) = \sum_i p_i \log \frac{1}{p_i}$$

### Example: Entropy of a fair coin.

The coin emits symbols  $s_1 = \text{heads}$  and  $s_2 = \text{tails}$  with  $p_1 = p_2 = 1/2$ . Therefore, the entropy of this source is:

$$\begin{aligned} H(\text{coin}) &= -(1/2 \times \log_2 1/2 + 1/2 \times \log_2 1/2) = \\ &= -(1/2 \times -1 + 1/2 \times -1) = -(-1/2 - 1/2) = 1 \text{ bit.} \end{aligned}$$



# Exemplu

- 8 simboluri -  $s_1, s_2, s_3, s_4, s_5, s_6, s_7, s_8$

- $p_i = 1/8$

- entropia = 3

- nu putem găsi un algoritm de codare cu mai puțin de 3 biți/eșantion

- $p_1=p_2=3/12; p_3=p_4=p_5=p_6=p_7=p_8=1/12$

- entropia = 2.79

# Exemplu

- O imagine de 1024x1024 pixeli și 8 biți/pixel are entropia 5.3 biți/pixel.

- Care este raportul de compresie maxim care se poate obține?

$$C = \frac{b}{b'}$$

$$C = 8 / 5.3 = 1.5$$

- Care este redundanta relativă?

$$R = 1 - \frac{1}{C}$$

$$R = 1 - 1/1.5 = 0.33$$

=> 33% date redundante

# Redundanțe

- **Redundanța codării**

- simbolurile de codare sunt reprezentate pe mai mulți biți decât este necesar

- **Redundanța spațială, temporală și spectrală**

- Informația se repetă în timp/spațiu

- **Redundanța spectrale**

- Exploatează reprezentarea semnalului în domeniu frecvență

- **Informație nerelevantă**

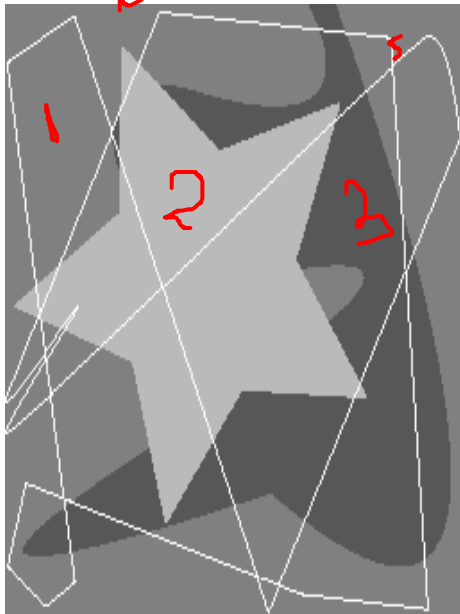
- din punct de vedere al sistemului vizual uman



# Introducere – concepte de bază

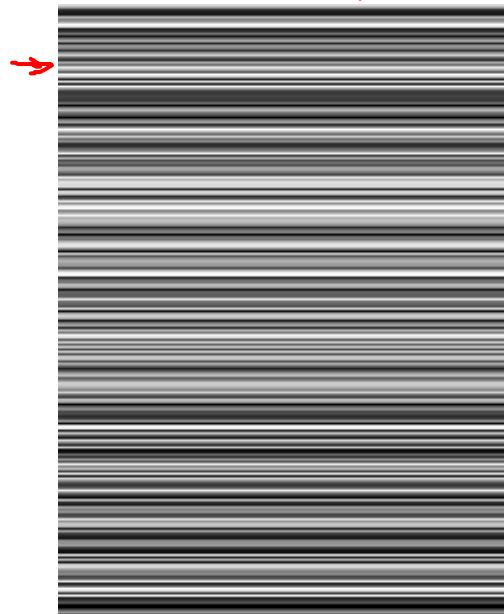
- Imagini 256x256 pixeli

8 bit?  $\rightarrow 2^8 = 255$   
2 bit

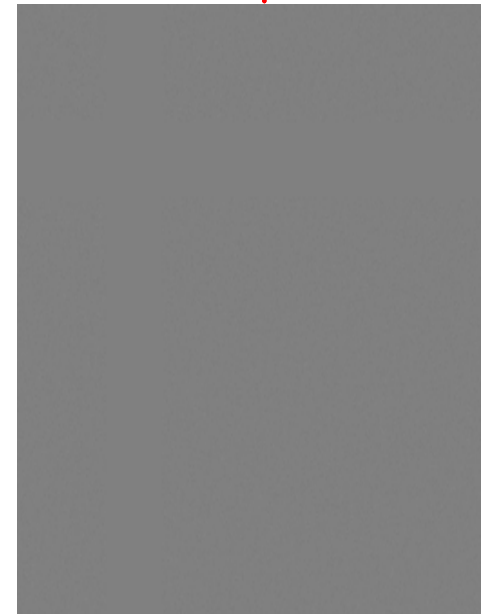


Redundanța codării

1 = ?



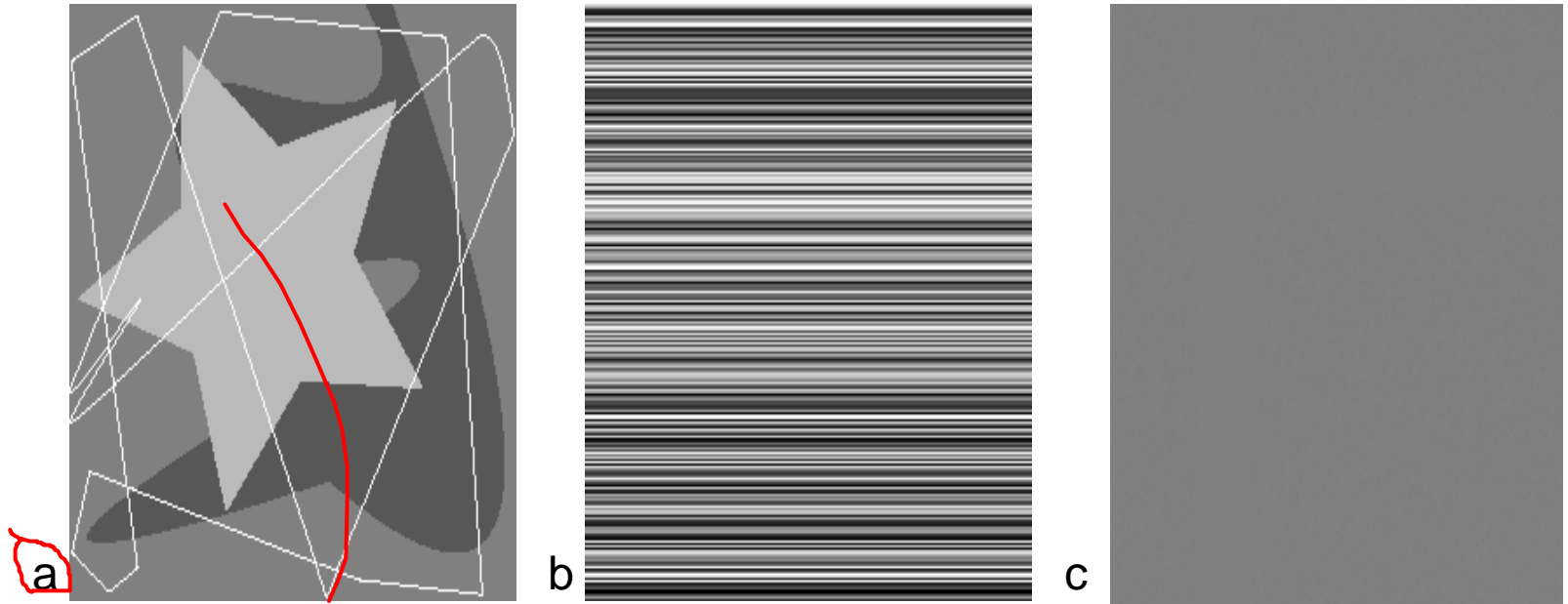
Redundanța spațială



Informație nerelevantă

# Imaginile de test ...

Imagini 256x256 pixeli



- Imaginea a – entropia = 1.6614 biți/pixel
- Imaginea b – entropia = 8 biți /pixel
- Imaginea c – entropia = 1.566 biți/ pixel
- Concluzie => Entropia și informația **NU sunt intuitive**

# Redundanța codării

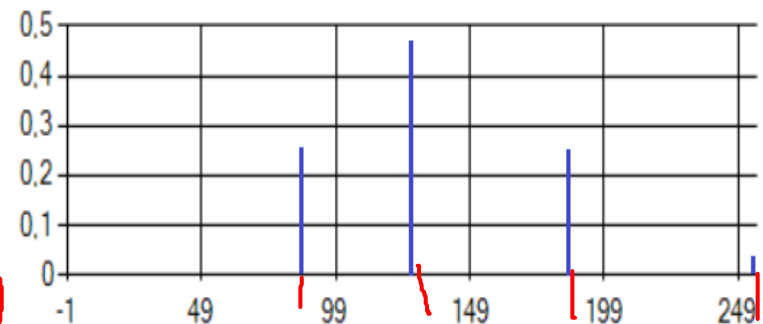
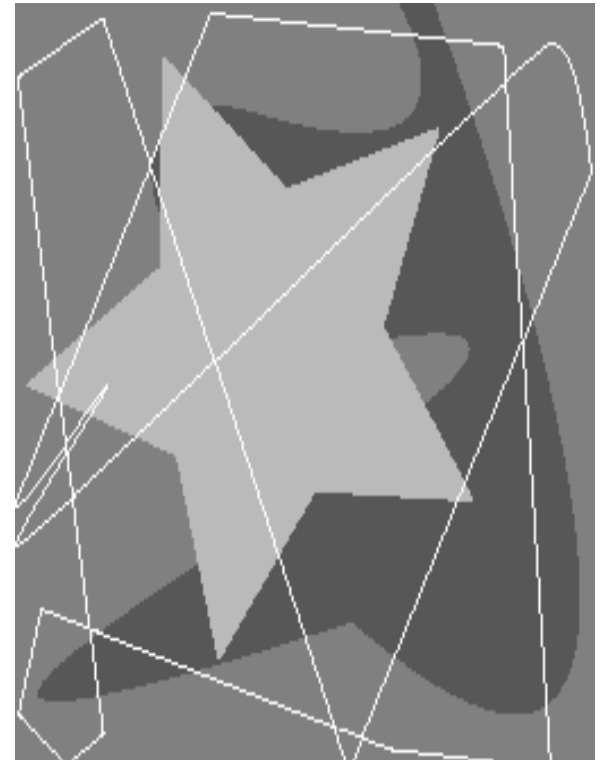
- Codarea optimă a informației pe un număr cât mai mic de biți
- Numărul mediu de biți, pe care se poate reprezenta informația:

$$L_{avg} = \sum_{k=0}^{L-1} l(r_k) p_r(r_k)$$

$r_k$  - nivelul de gri k în intervalul [0, L-1]

$p_r(r_k)$  probabilitatea de apariție a nivelului de gri  $r_k$

$l(r_k)$  - numărul de biți necesari pentru reprezentarea nivelului  $r_k$



# Redundanța codării - exemplu

Nivel gri $r_k$	Probabilitatea de apariție $p_r(r_k)$	Code 1	Număr biți reprezentare code 1 $l_1(r_k)$	Code 2	Număr biți reprezentare code 2 $l_2(r_k)$
87	0.25	01010111	8	01	2
128	0.47	10000000	8	1	1
186	0.25	11000100	8	000	3
255	0.03	11111111	8	001	3

- să calculăm  $L_{avg}$  pentru codul 1 și 2
- asignarea codurilor scurte la simbolurile cu probabilitate mare de apariție

=> **coduri de lungime variabilă**

$$L_{avg} = 8 \times 0.25 + 8 \times 0.47 + 8 \times 0.25 + 8 \times 0.03 = 8$$

$$2 \cdot 0.25 + 1 \cdot 0.47 + 3 \cdot 0.25 + 3 \cdot 0.03 = 1.81$$

# Redundanța codării

Numărul total de biți necesari pentru reprezentarea unei imagini

$$b' = M \times N \times L_{avg} = 256 \times 256 \times 1.81$$

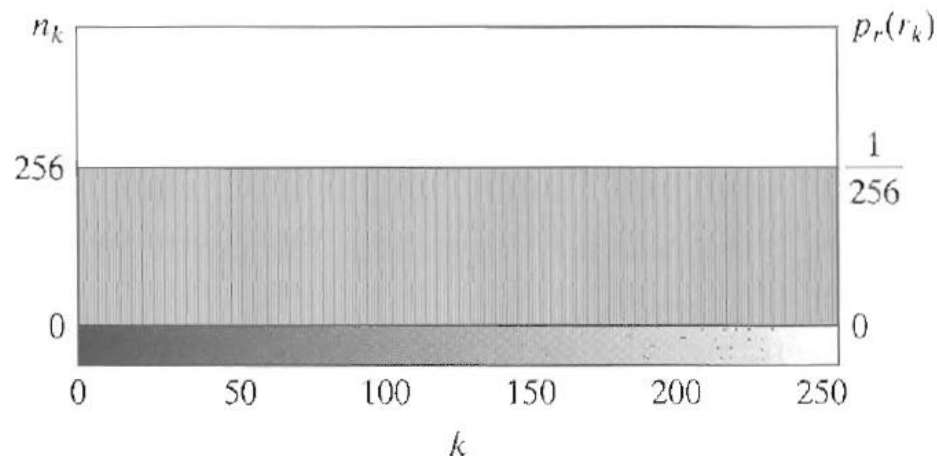
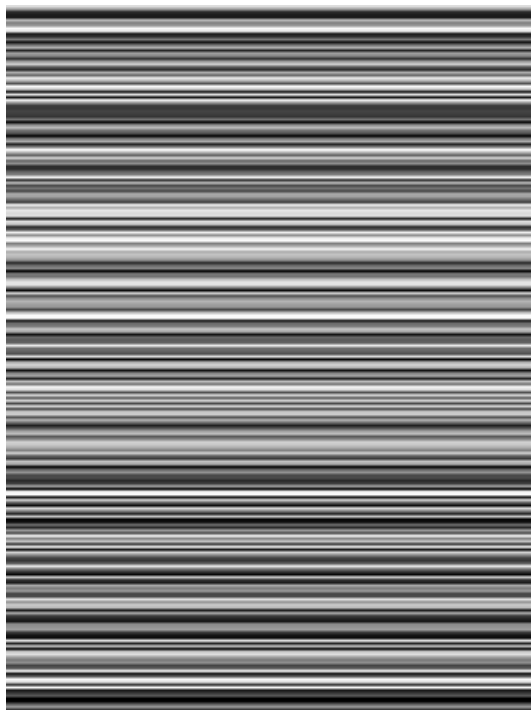
$$b = M \times N \times L_{avg} = 256 \times 256 \times 8$$

$$C = \frac{b}{b'} = 4.42$$

$$R = 1 - \frac{1}{C} = 1 - \frac{1}{4.42} = 0.774 \text{ adică } 77.4\% \text{ este informație redundanță}$$

# Redundanța spațială, temporală, spectrală

- Toate intensitățile au probabilități egale



- Intensitatea fiecărei linii este selectată aleator – pixelii sunt independenți pe verticală
- Pixelii dintr-o linie sunt identici – maxim corelați, depedenți unul de altul pe orizontală

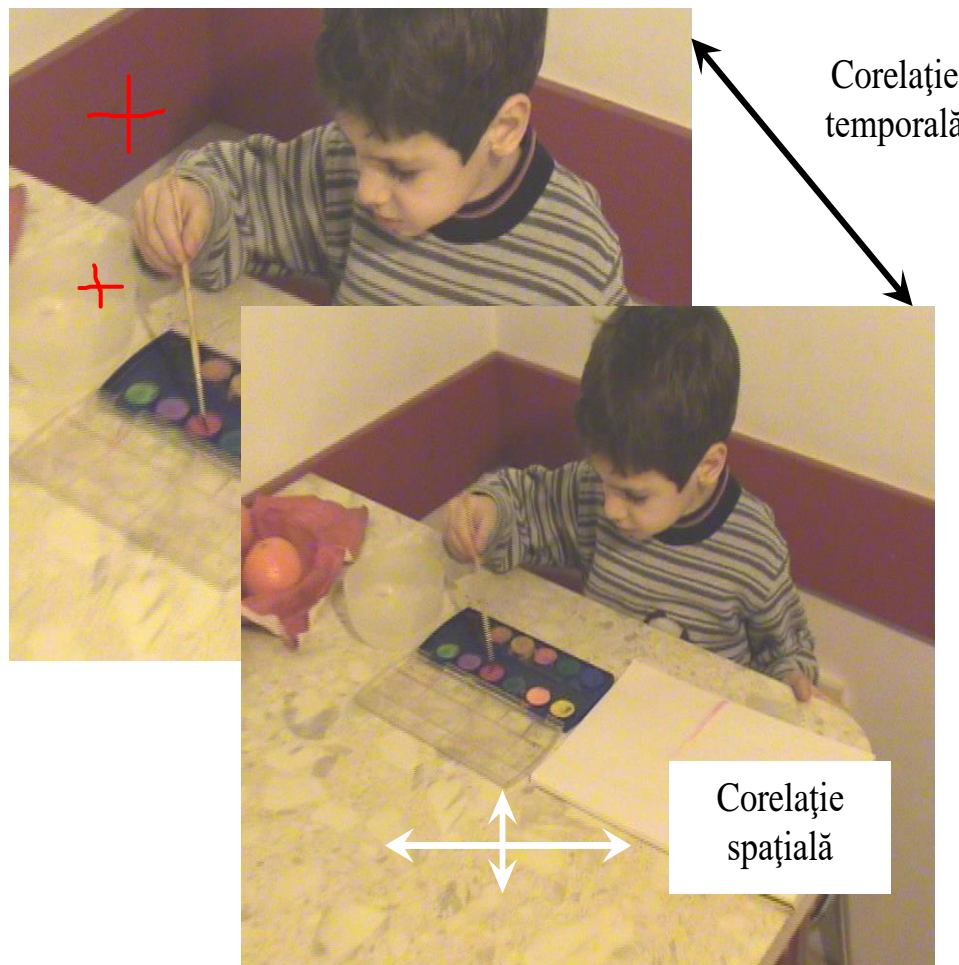
# Redundanța spațială, temporală

- Observații
  - NU se pot folosi codurile de lungime variabilă – cele 256 simboluri rk au aceeași probabilitate
  - Se poate coda lungimea curselor
  - Pereche de codare
    - Start nouă intensitate
    - Lungimea intensității
  - Să vedem un exemplu.....
    - Fiecare linie se înlocuiește cu
      - Valoare intensitate
      - Lungimea – care pentru exemplul nostru este 256

(10256), (30)<sup>256</sup>

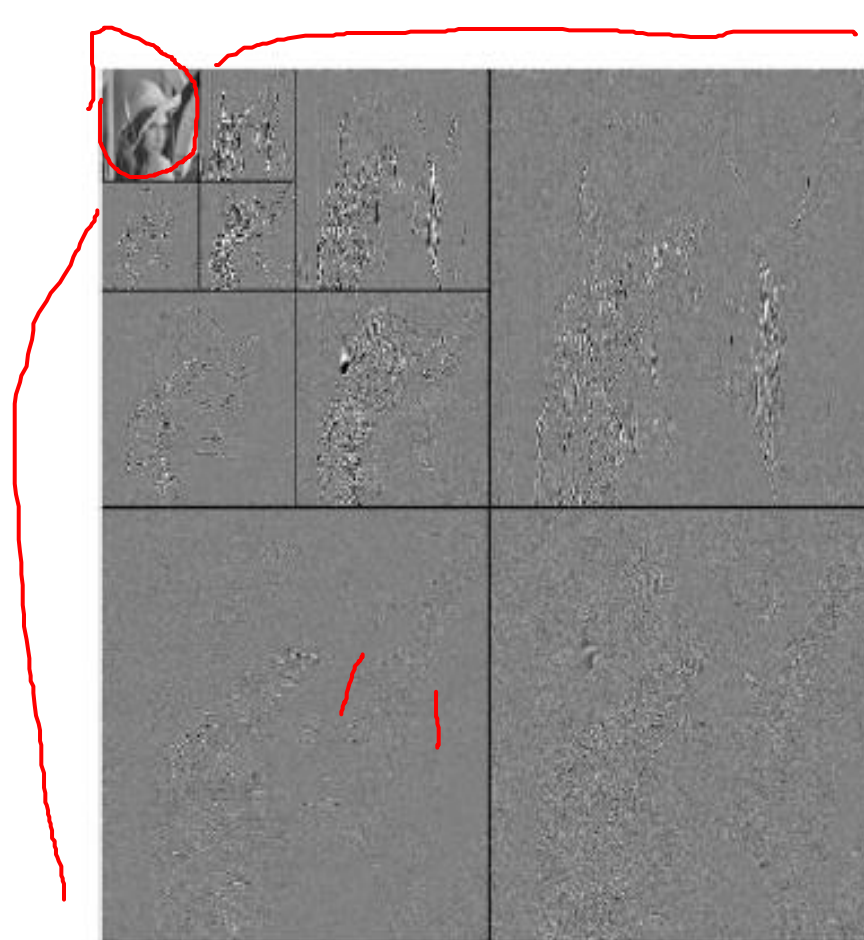
# Redundanța spațială, temporală, spectrală

- Corelația spațială – ambele direcții
- Corelație temporală
  - Cadre succesive

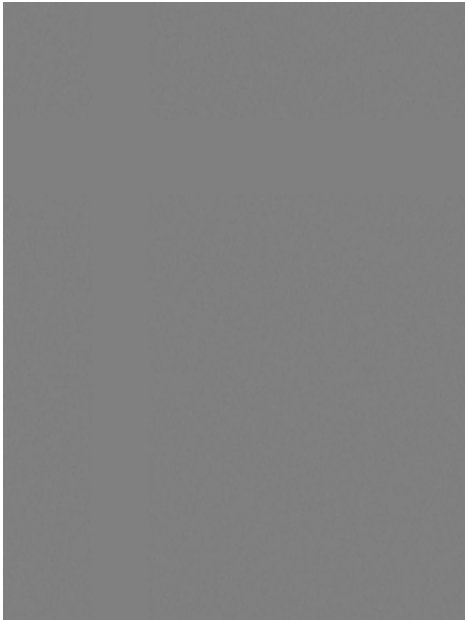




# Redundanța spectrală



# Informație irelevantă



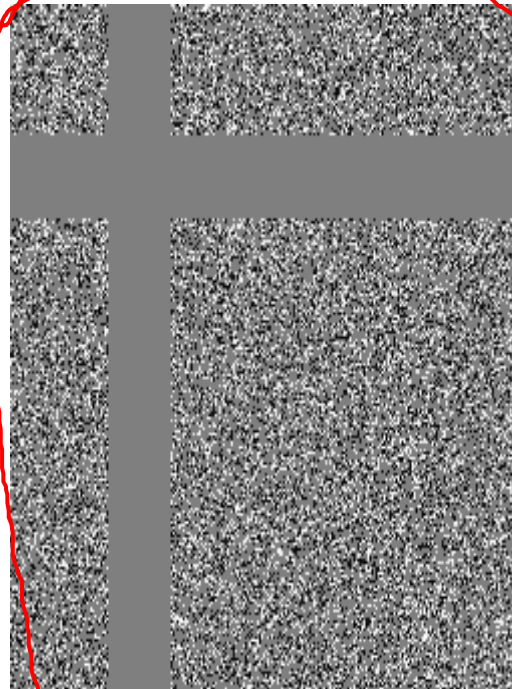
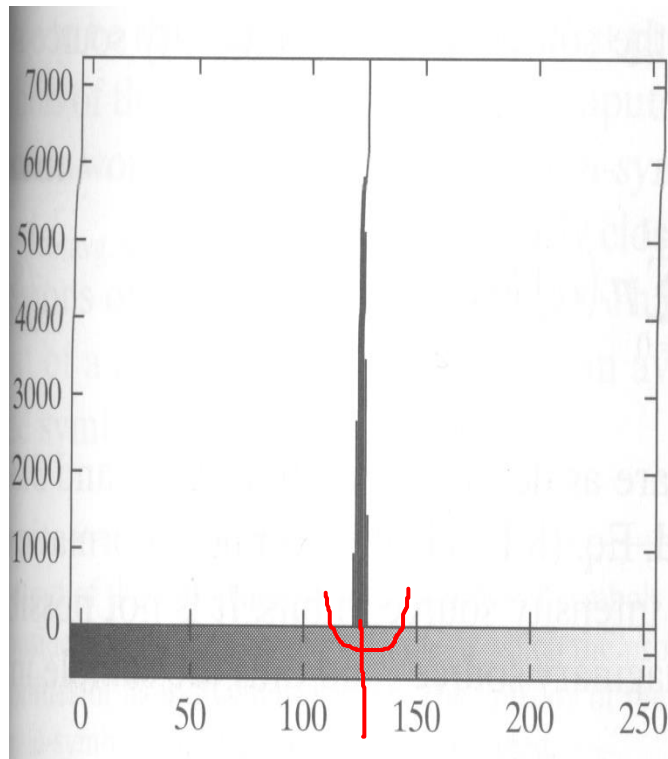
- Reprezentarea intensității medii (intensitățile sunt foarte apropiate și ochiul nu le poate distinge)
- Un singur octet
- Compresie

$$C = \frac{b}{b'} = \frac{256 \times 256 \times 8}{8} = 65536$$

- La decodare imaginea refăcută nu va avea de suferit foarte mult

# Informație irelevantă

- Totuși imaginea dacă se face egalizare nu are intensitate uniformă
- Imaginea initiala nivele de gri – 125-131, după egalizare apar informații relevante și pentru sistemul vizual
- Depinde de tipul aplicației... de ex. imaginile medicale!!!





# Criterii de calitate

- **Obiective**

- Funcții matematice
- Evaluare simplă

$$MAD(i, j) = \frac{1}{mn} \sum_i^n \sum_j^m |F(i, j) - G(i, j)|$$

**MAD** - medie absolută a diferenței

- **Subiective**

- Scală de evaluare:
  - excelent, bună, acceptabilă, satisfăcătoare, inferioară, inacceptabilă

$$MSE(U, \hat{U}) = \frac{1}{M \cdot N} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} (u(m, n) - \hat{u}(m, n))^2.$$

PSNR



# Criterii de calitate - exemplu

- 1-original, 2-decomprimată, 3-generată artif.
- ATENTIE – 1, 3 prin criteriul obiectiv RSZ da o valoare f.bună, DAR prin criteriul subiectiv NU este aceeași imagine!!!

