



# Συμπίεση Δεδομένων: Συμπίεση Ψηφιακού Βίντεο



Αλέξανδρος Ελευθεριάδης

Αναπ. Καθηγητής & Marie Curie Chair

Τμήμα Πληροφορικής και Τηλεπικοινωνιών

Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών

eleft@di.uoa.gr, (210) 727-5210

[Διάλεξη 2<sup>η</sup>]

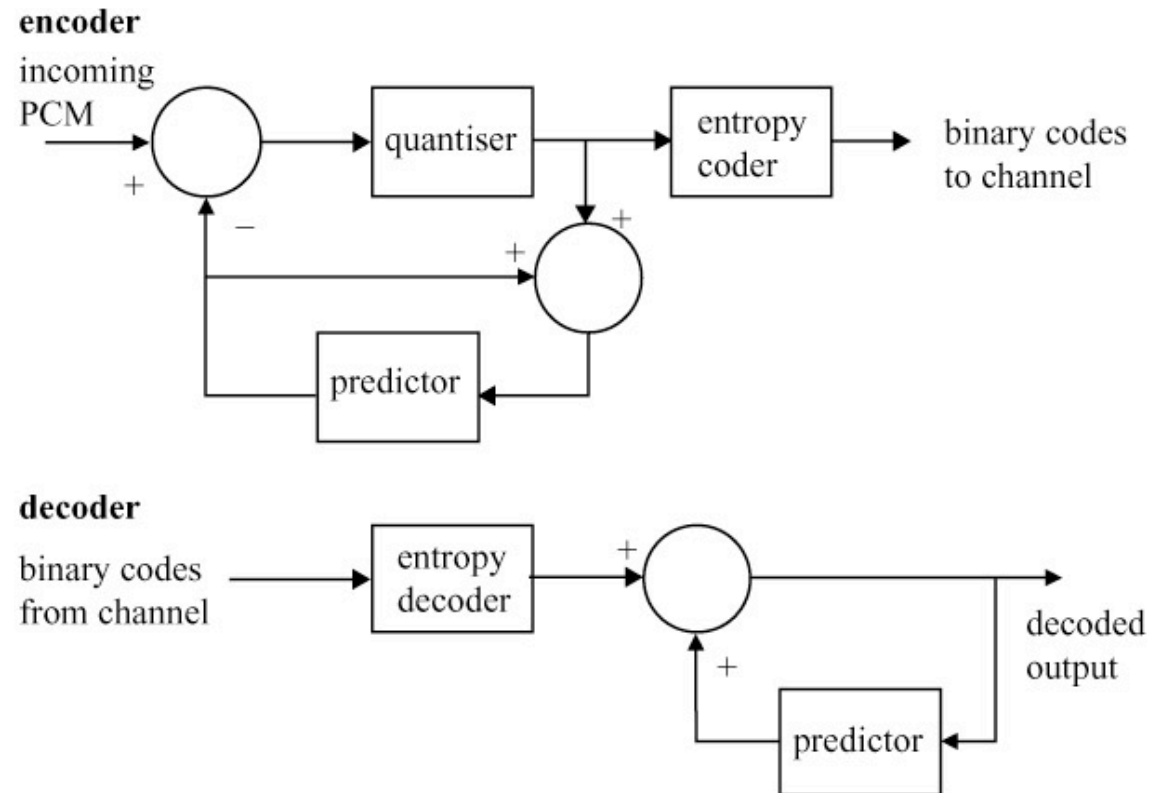
# Βασικές Αρχές Συμπίεσης

---

1. Ελαχιστοποίηση χωρικής πλεονασματικότητας (spatial redundancy) - π.χ. transform coding
2. Ελαχιστοποίηση χρονικής πλεονασματικότητας (temporal redundancy) - π.χ. motion compensated prediction
3. Κωδικοποίηση εντροπίας (entropy coding) - π.χ. Huffman coding

# DPCM

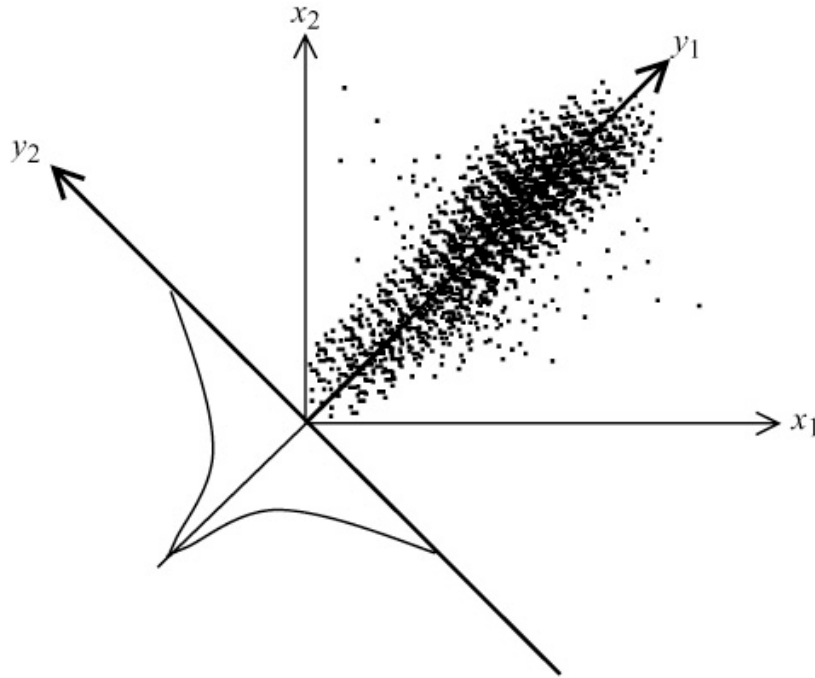
- Μεταξύ pixels αλλά και μεταξύ frames



# Transform Coding

---

- Παράδειγμα



$$[y_1, y_2] = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

# Discrete Cosine Transform

---

- Προσέγγιση του KLT
- Κατανομή σταθερών ώστε οι συντελεστές να είναι μεταξύ -2047 και +2047

$$f(x) = \sqrt{\frac{2}{N}} \sum_{u=0}^{N-1} C(u) F(u) \cos\left(\frac{\pi(2x+1)u}{2N}\right) \quad x = 0, 1, \dots, N-1$$

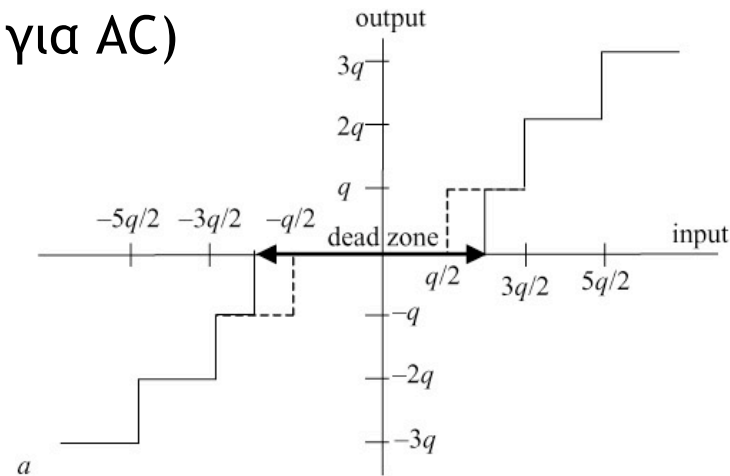
$$F(u) = \sqrt{\frac{2}{N}} C(u) \sum_{x=0}^{N-1} f(x) \cos\left(\frac{\pi(2x+1)u}{2N}\right) \quad u = 0, 1, \dots, N-1$$

$C(u) = \sqrt{1/2}$  αν  $u=0, 1$  διαφορετικά

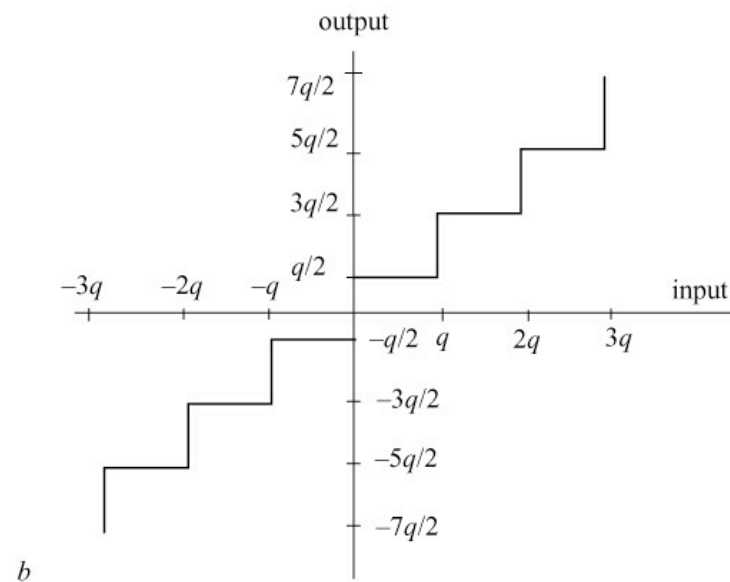
- Mismatch Control
  - Μη τυποποιημένη υλοποίηση συσσωρεύει σφάλμα στο δέκτη.
  - Προσθήκη ψευδοτυχαίου σήματος το κρατάει μέσα σε όρια

# Κβαντισμός (quantization)

- Με dead-zone (τυπικά για AC)



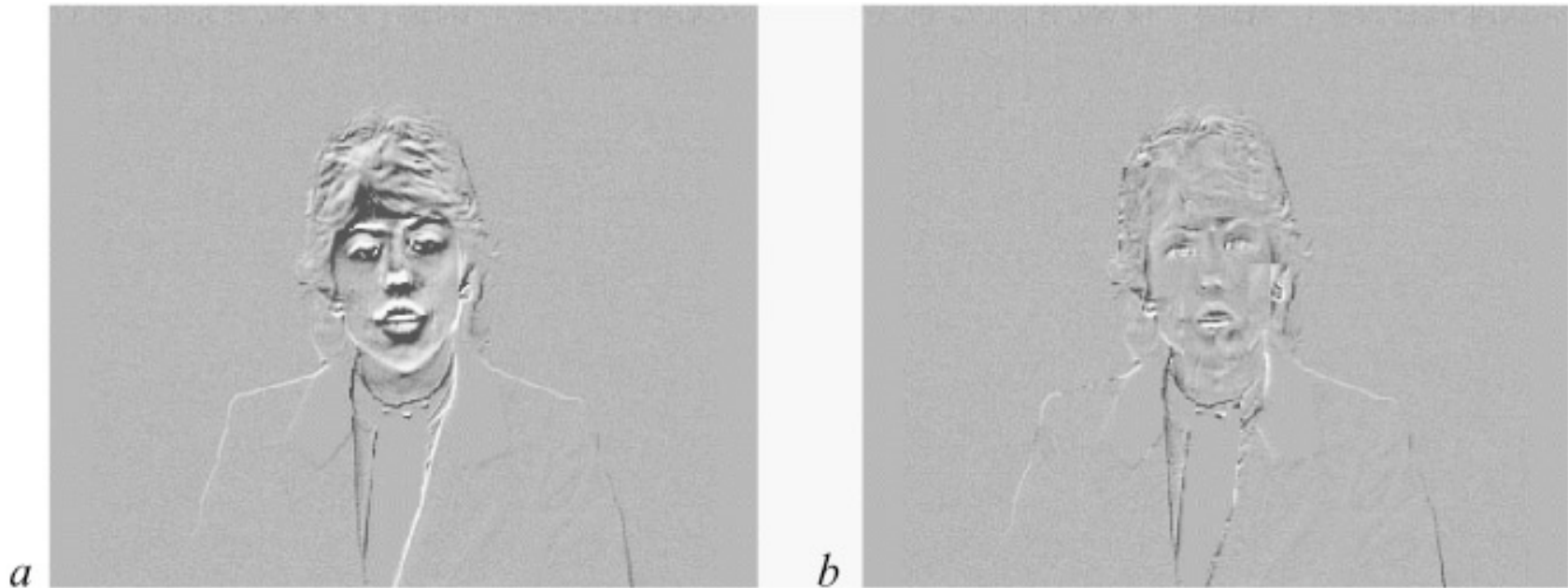
- Χωρίς (τυπικά για DC)



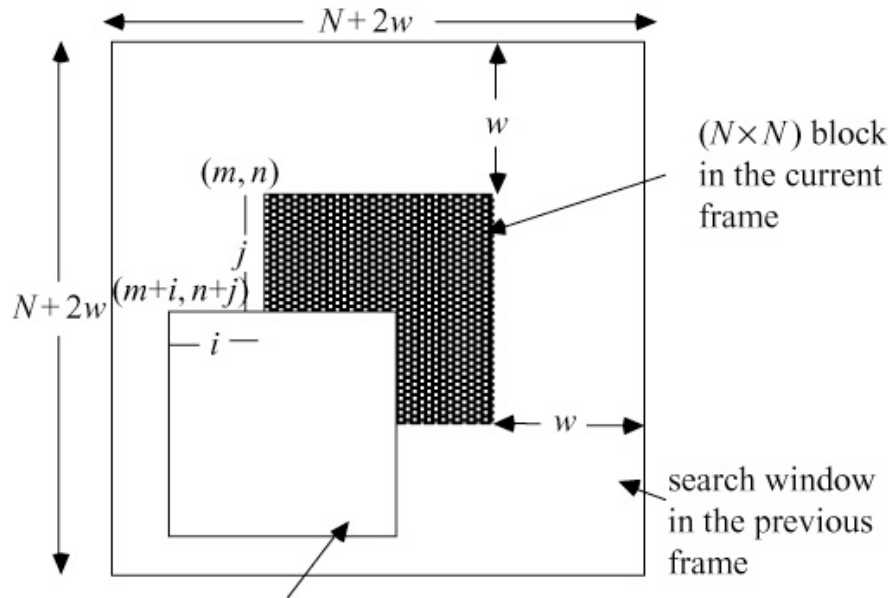
# Motion Estimation

---

- Διαφορά διαδοχικών εικόνων χωρίς (a) και με (b) εκτίμηση κίνησης



# Μοντέλο ΜΕ



( $N \times N$ ) block under the search in the previous frame, shifted by  $i, j$

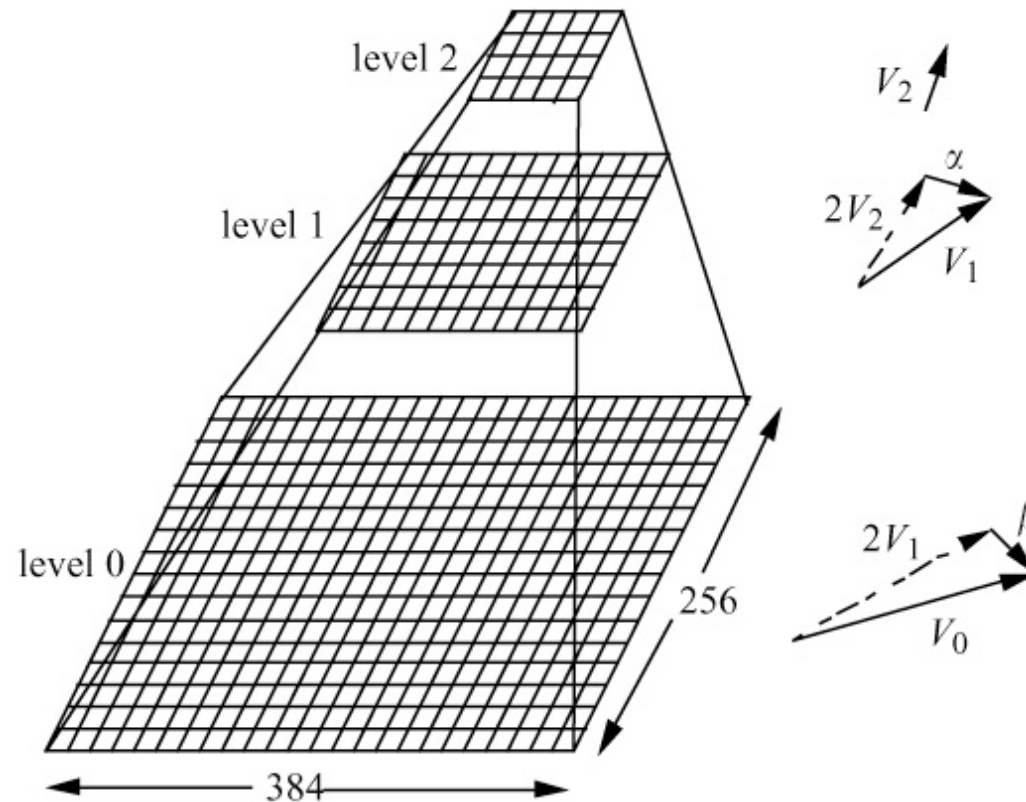
- **MSE** 
$$M(i, j) = \frac{1}{N^2} \sum_{m=1}^N \sum_{n=1}^N (f(m, n) - g(m + i, n + j))^2, \quad -w \leq i, j \leq w$$
- **MAE** 
$$M(i, j) = \frac{1}{N^2} \sum_{m=1}^N \sum_{n=1}^N |f(m, n) - g(m + i, n + j)|, \quad -w \leq i, j \leq w$$





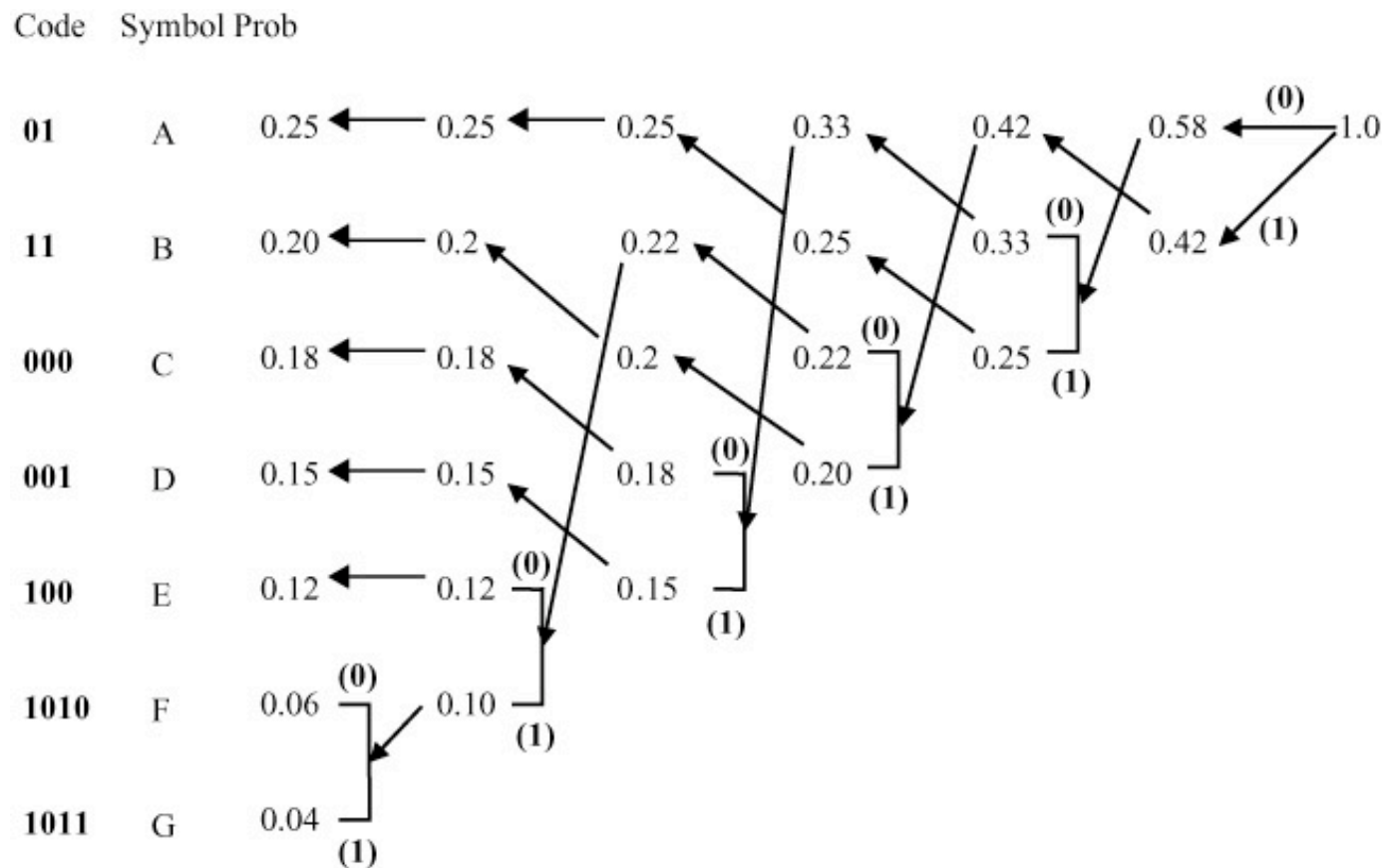
# Ιεραρχική Εκτίμηση (Hierarchical ME)

- Ιδιαίτερα χρήσιμη σε scalable codecs



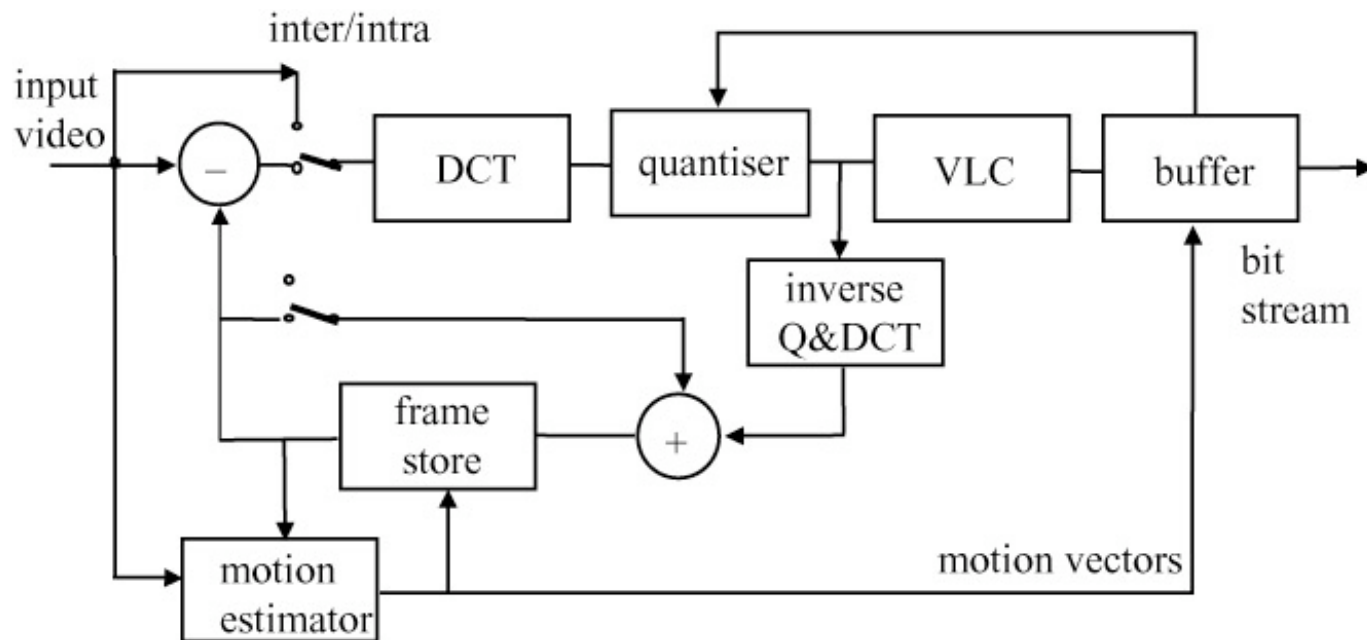
# Κωδικοποίηση Εντροπίας

- Εντροπία  $H(x) = - \sum_{i=1}^n p_i \log_2 p_i$
- Huffman & Arithmetic Coding
- Παράδειγμα Huffman:



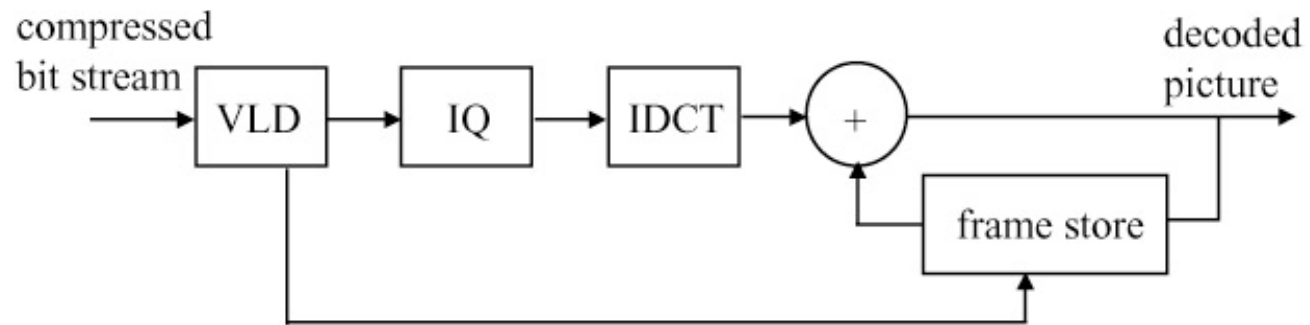
# Γενική Αρχιτεκτονική Κωδικοποιητή Βίντεο

- Interframe predictive encoder
- Καλύπτει όλες τις γενιές κωδικοποιητών βίντεο!
- Παρατηρήστε σύνδεση buffer-quantizer (rate control)

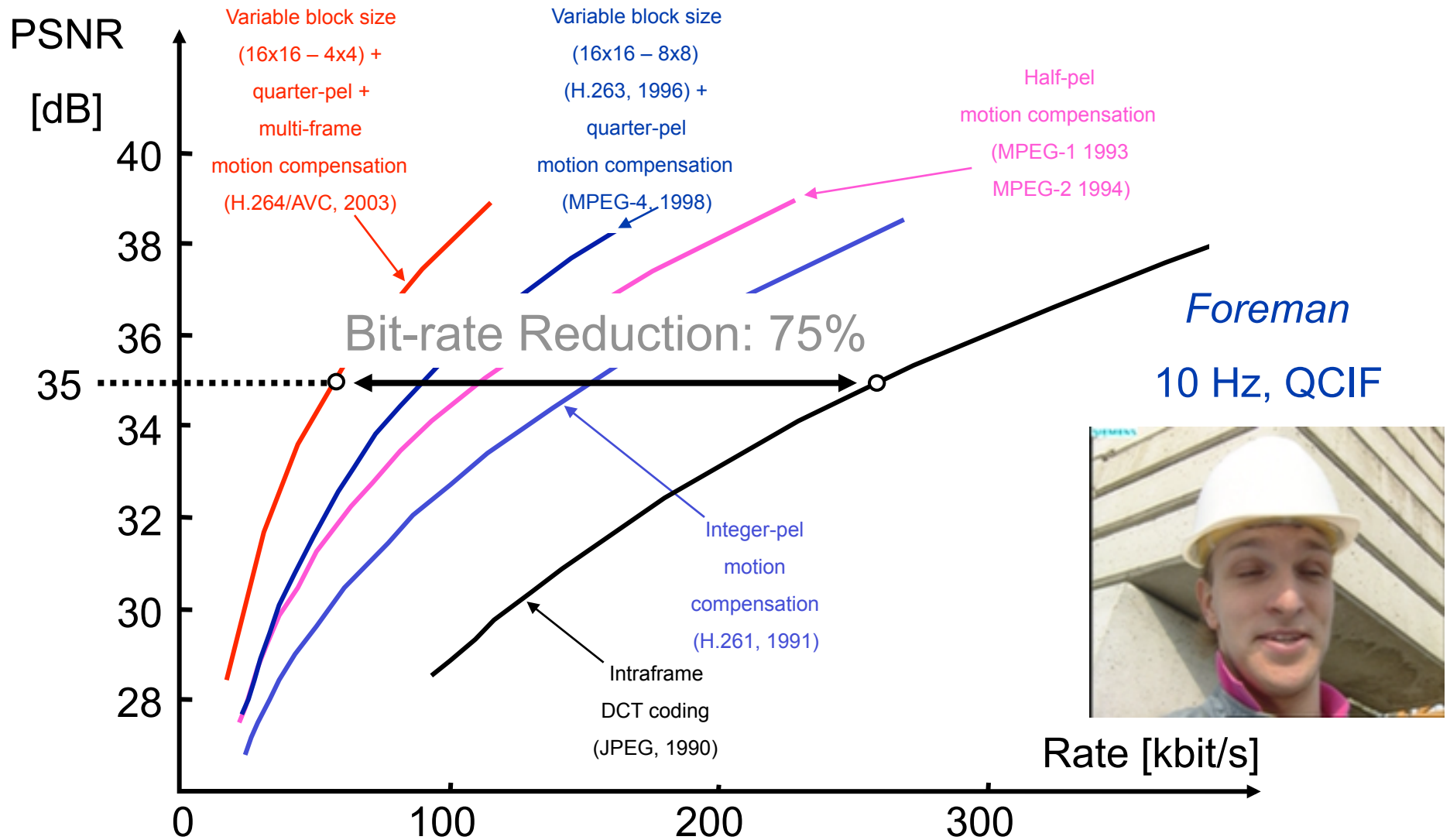


# Αποκωδικοποιητής

---



# Επιτεύγματα στην Κωδικοποίηση Βίντεο



# Ύλη (Ghanbari)

---

- Κεφάλαια 1-3
- Κεφάλαιο 4 - εκτός
- Κεφάλαιο 5
  - JPEG – Baseline
  - JPEG2000 - εκτός

# JPEG

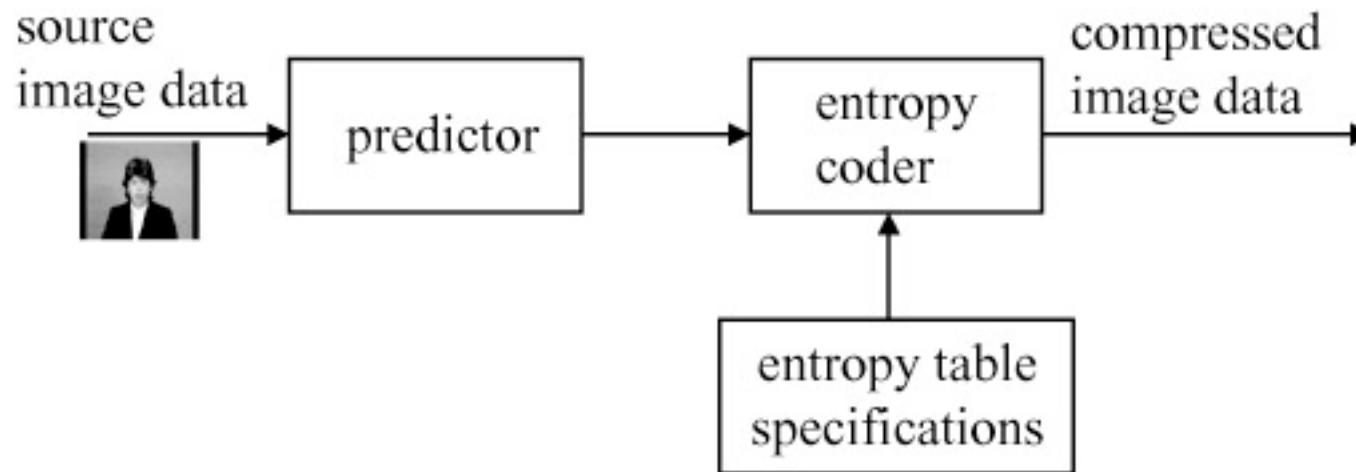
---

- Joint (ITU+ISO) Photographic Experts Group
- JPEG 1994 (ISO 10918-1) - .jpg
  - Coding
  - File format (JIF - Annex B, JFIF)
- JPEG2000 2000 (ISO 15444-1) - .jp2



# JPEG - Lossless

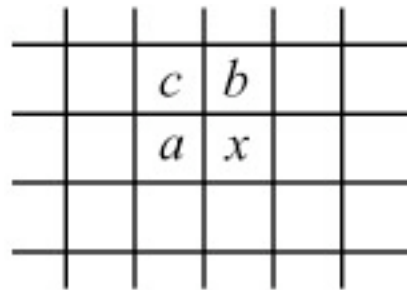
---



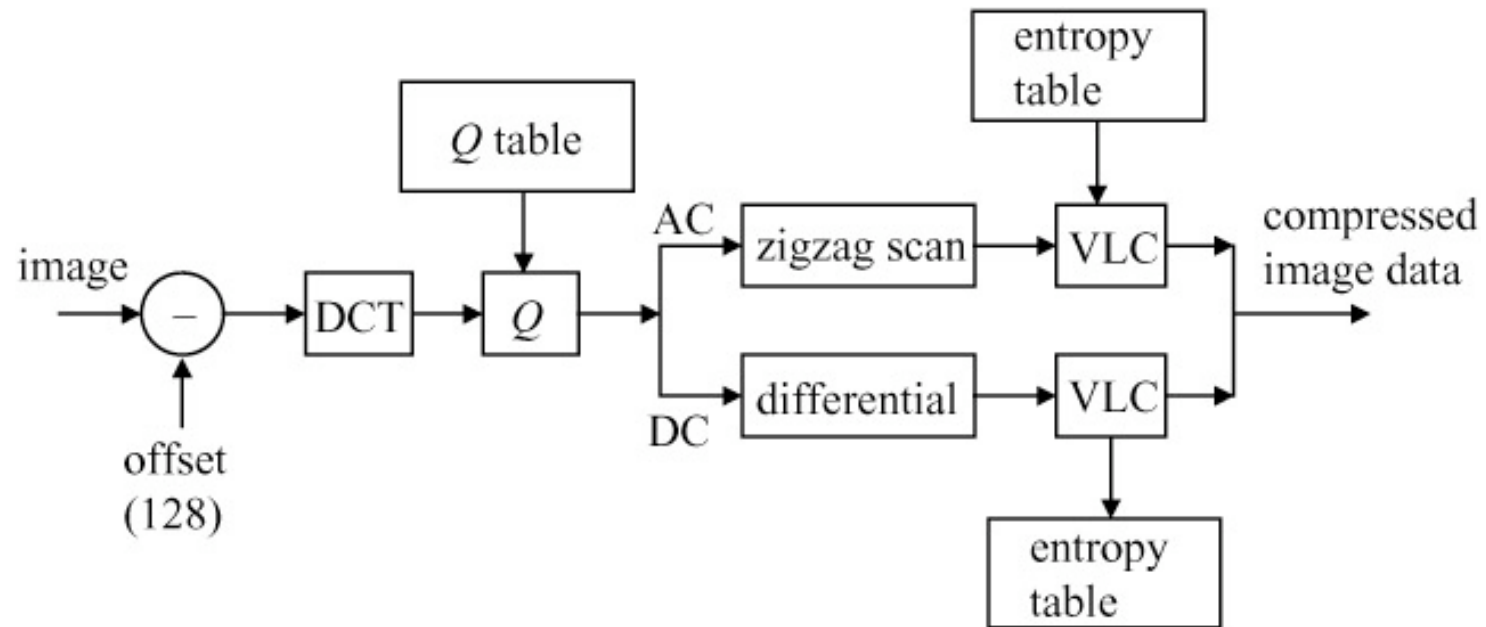
# JPEG - Lossless

---

- 3 sample prediction neighborhood



# Baseline Sequential JPEG (Lossy)



# Πίνακας Luminance Q

---

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	56	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

$$F^q(u, v) = \left[ \frac{F(u, v)}{Q(u, v)} \right]$$

$$F^Q(u, v) = F^q(u, v) \times Q(u, v)$$

# Q Factor

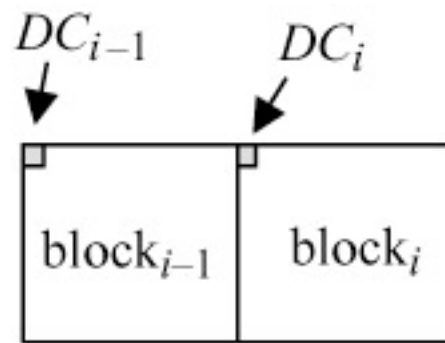
---

- Το  $\alpha$  πολλαπλασιάζει τα στοιχεία του πίνακα

$$\alpha = \frac{50}{q\_JPEG} \quad \text{if } 1 \leq q\_JPEG \leq 50$$
$$\alpha = 2 - \frac{2 \times q\_JPEG}{100} \quad \text{if } 50 \leq q\_JPEG \leq 99$$

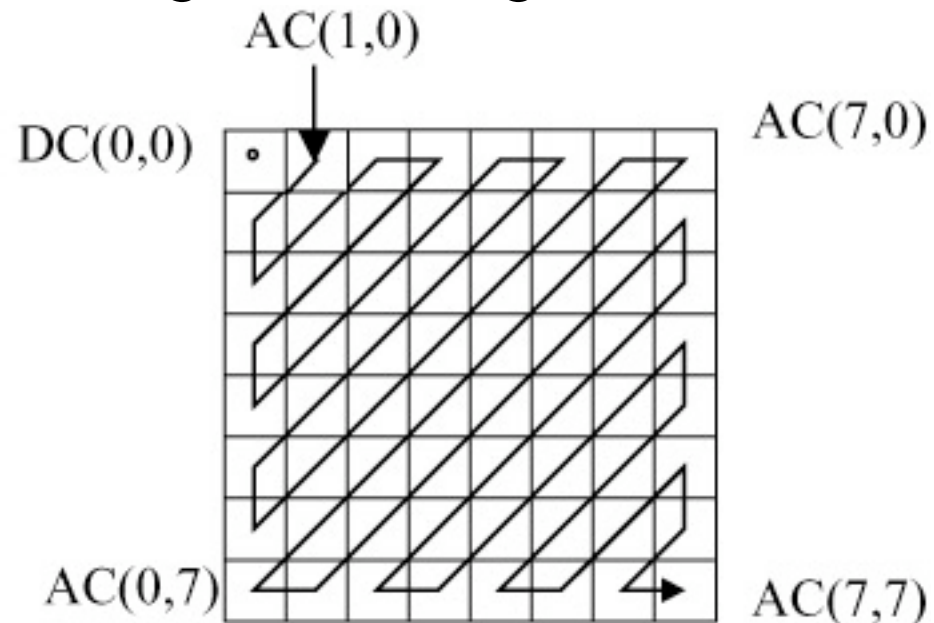
# Κωδικοποίηση DCT

- DC - πρόβλεψη
- AC - zig-zag scan + run-length encoding



$$\text{DIFF} = DC_i - DC_{i-1}$$

differential encoding of DC



zigzag scanning of AC

# Κωδικοποίηση Συντελεστών DC

---

- $\text{DIFF} = \text{Range}(\text{CAT}) + \text{Τιμή}$

<u>CAT</u>	<u>Range</u>
0	-
1	-1, 1
2	-3, -2, 2, 3
3	-7, ..., -4, 4, ..., 7
...	
11	-2047, ..., -1024, 1024, ..., 2047

- Κωδικοποίησης τελικών bits τιμής
  - Αν  $\text{DIFF} > 0$ , ξεκινάς από το πρώτο bit με τιμή 1
  - Αν  $\text{DIFF} < 0$ , ξεκινάς από το πρώτο bit του  $\text{DIFF}-1$  με τιμή 0

# Παράδειγμα

---

- $\text{DIFF} = 6$  (0...0110)
  - $\text{DIFF} > 0$
  - $\text{CAT} = 3$  ( $4 \leq \text{DIFF} \leq 7$ ), bits = 100 (Πίνακας B.1 βιβλίου)
  - Τελικά bits 110
  - Τελικός κώδικας = 100110
- $\text{DIFF} = -3$ 
  - $\text{DIFF} < 0$
  - $\text{CAT} = 2$  ( $-3 \leq \text{DIFF} \leq -2$ ), bits = 011 (Πίνακας B.1 βιβλίου)
  - $\text{DIFF}-1 = -4$  ή 1....100
  - Τελικά bits 00
  - Τελικός κώδικας 01100



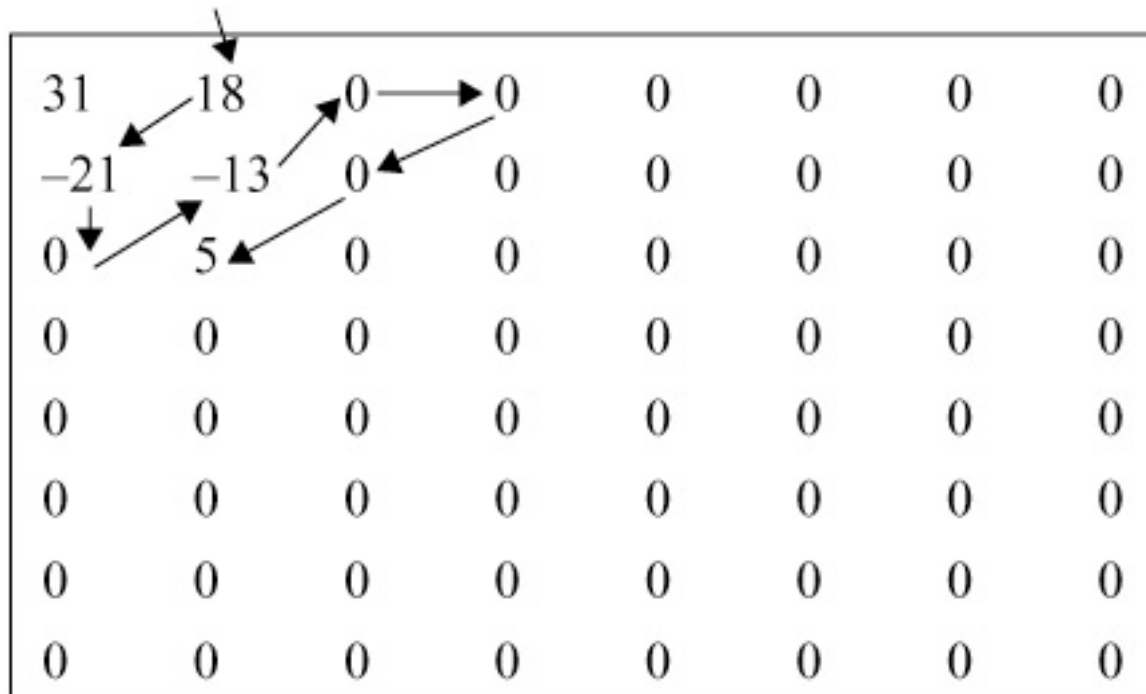
# Κωδικοποίηση Συντελεστών AC

---

- Κωδικοποίηση (RUN, CAT) ή (RUN, SIZE)
- Max run = 15, για μεγαλύτερα runs επαναλαμβάνεις το σύμβολο (15, 0)
- EOB = (0, 0) = τέλος σειράς AC
- Κωδικοποίηση υπόλοιπων bits όπως DC

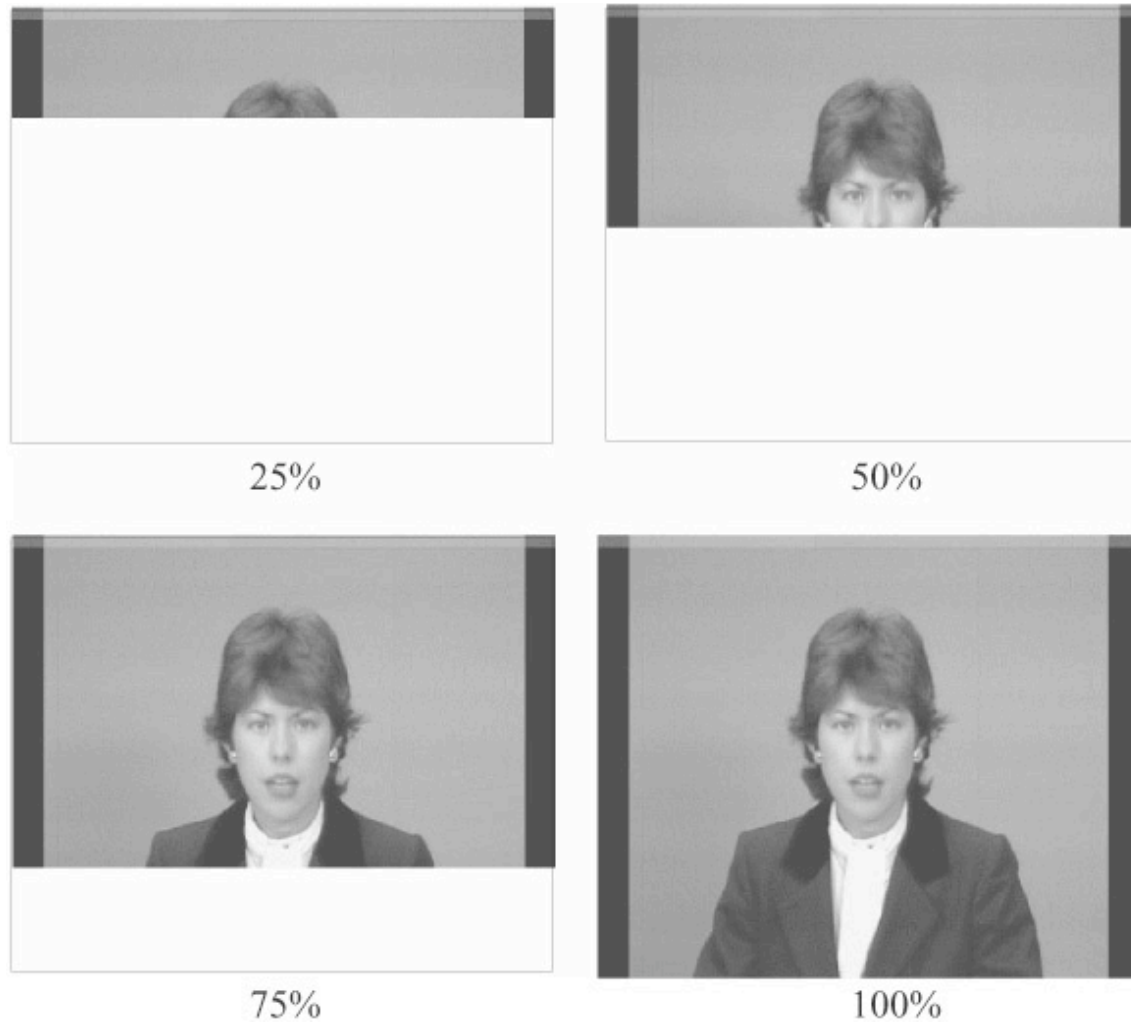
# Παράδειγμα

---



# Sequential JPEG

---



# Progressive JPEG

---

- Bitplane coding



0



0,1,8



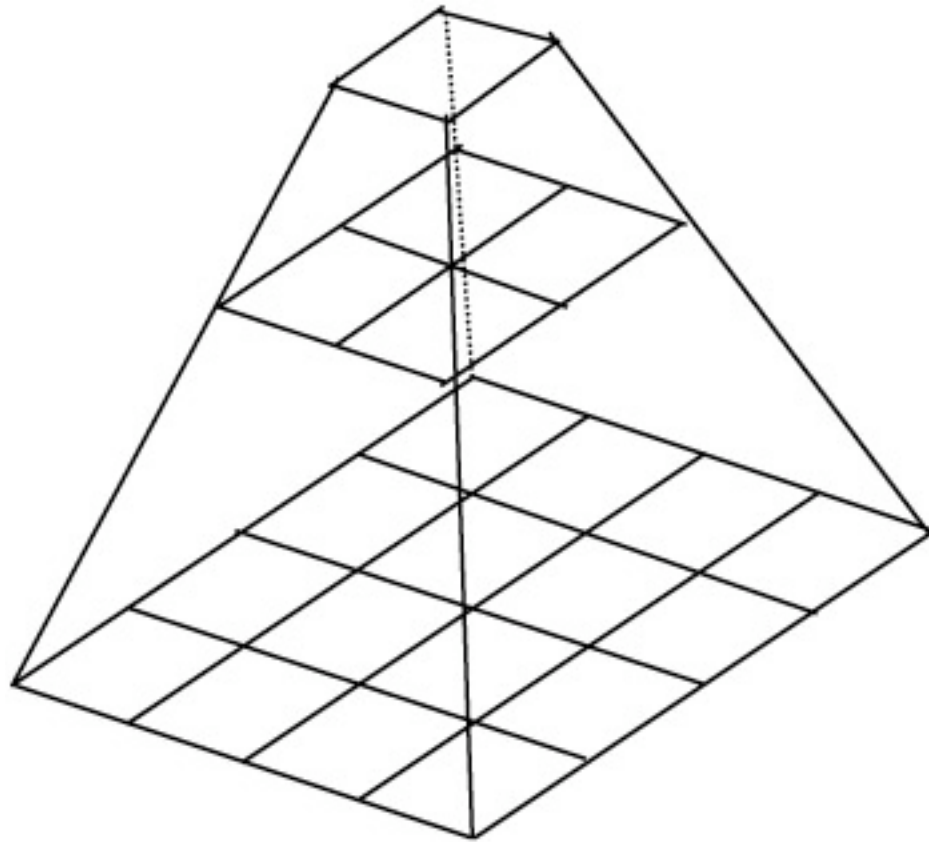
0,1,8,16,9,2,3,10,17,24



all

# Hierarchical Multiresolution Encoding

---



# Hierarchical Multiresolution Encoder

