



# Συμπίεση Δεδομένων: Συμπίεση Ψηφιακού Βίντεο



Αλέξανδρος Ελευθεριάδης

Αναπ. Καθηγητής & Marie Curie Chair

Τμήμα Πληροφορικής και Τηλεπικοινωνιών

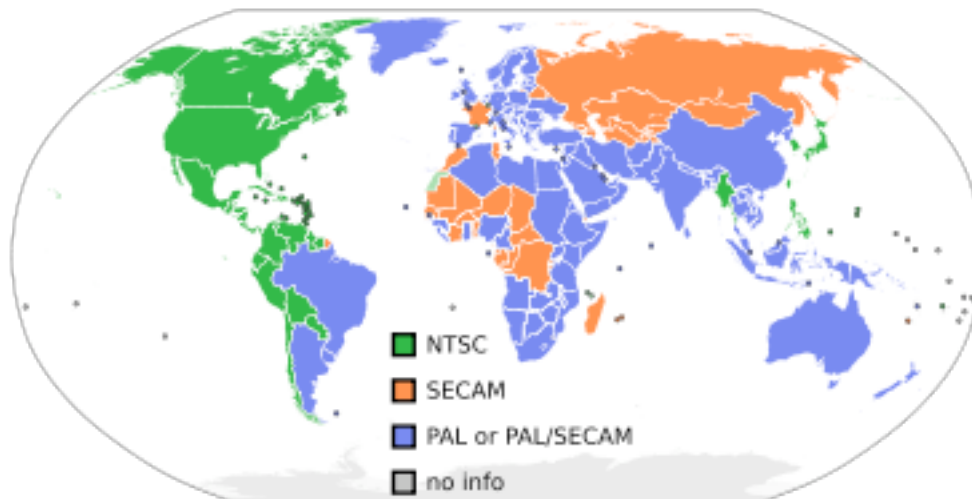
Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών

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

[Διάλεξη 1<sup>η</sup>]

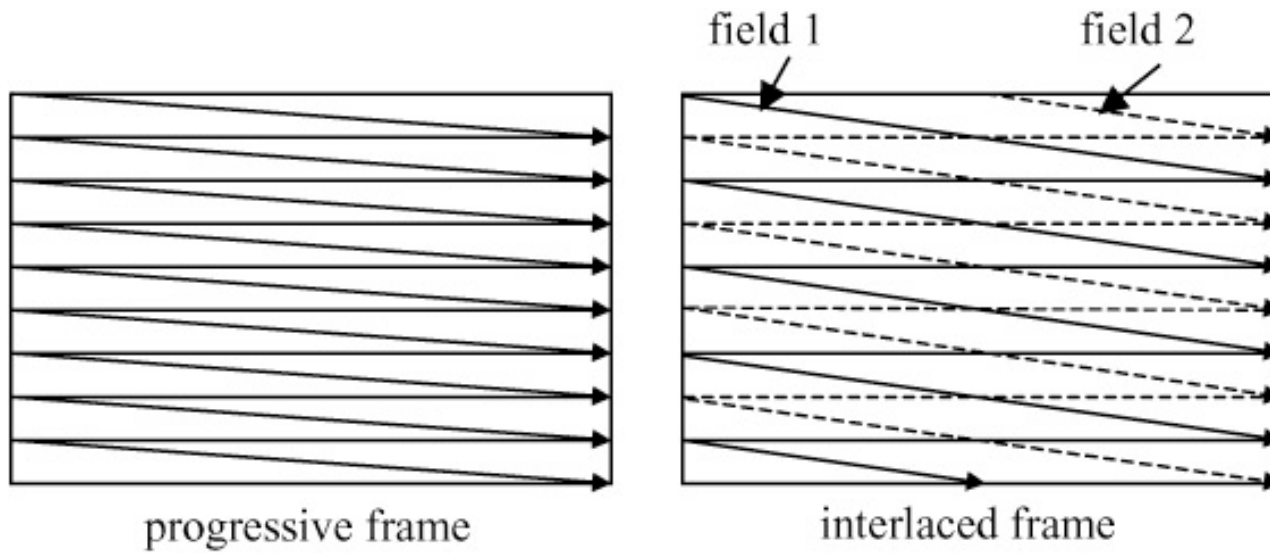
# Διεθνή Πρότυπα Αναλογικού Βίντεο

---



# Progressive/Interlaced

---



# Ασπρόμαυρο - Έγχρωμο Βίντεο

---

- Αναλογικό

$$Y = 0.299R' + 0.587G' + 0.114B'$$

$$U = -0.147R' - 0.289G' + 0.436B' = 0.492(B' - Y)$$

$$V = 0.615R' - 0.515G' - 0.100B' = 0.877(R' - Y)$$

- Ψηφιακό (CCIR-601)

$$Y = 0.257R' + 0.504G' + 0.098B' + 16$$

$$C_b = -0.148R' - 0.291G' + 0.439B' + 128$$

$$C_r = 0.439R' - 0.368G' - 0.071B' + 128$$

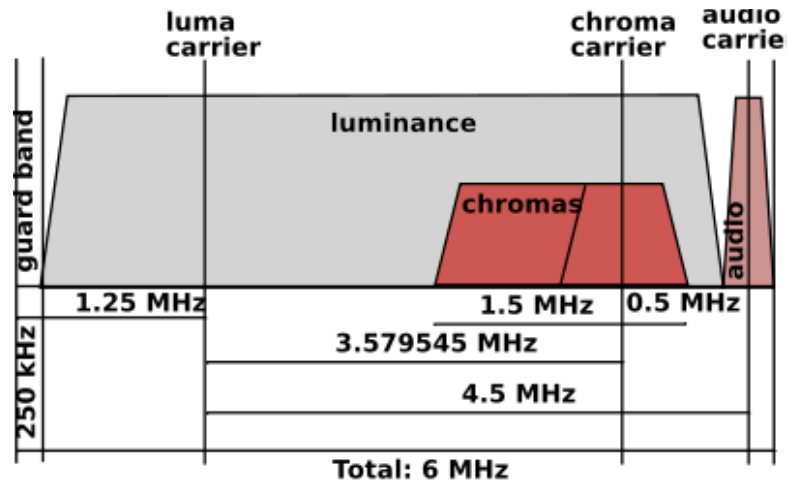
# Πρότυπα

---

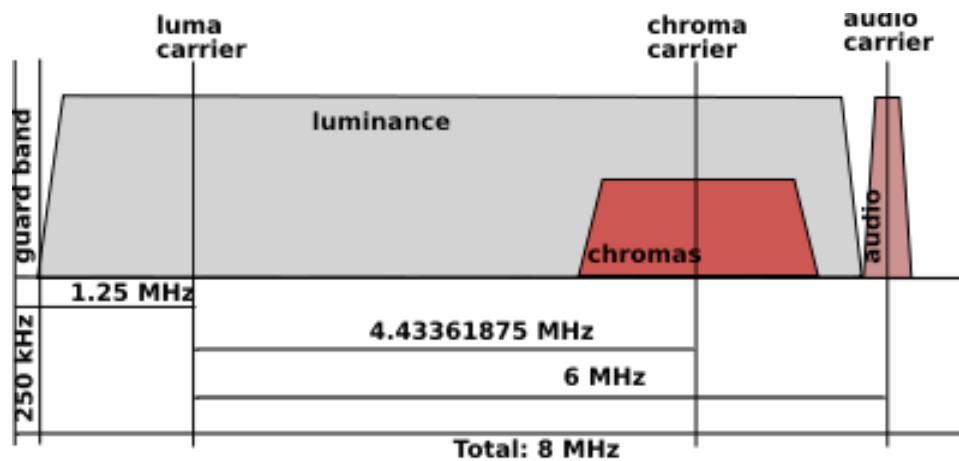
- NTSC - National Television Standards Committee
  - 1941 μονόχρωμο, 525 lines, 30 fps, 2:1 interlaced, 4:3 aspect ratio
  - 1950 έγχρωμο, 29.97 fps
- PAL - Phase Alternating Line
  - 1963 έγχρωμο, 625 lines, 25 fps, 2:1 interlaced, 4:3 aspect ratio

# Διαμόρφωση

- NTSC



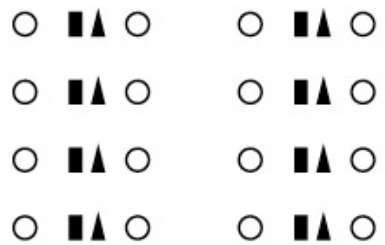
- PAL



# Ψηφιακό Βίντεο

- CCIR-601 / ITU-R BT.601
  - 720x525@60 (active 480)
  - 720x625@50 (active 576)
  - 2:1 interlace, 4:2:2
- Chroma Formats:

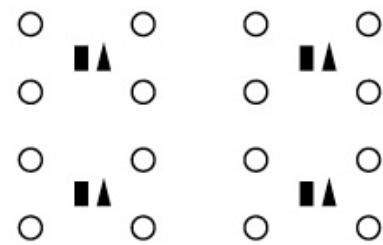
4:2:2 CCIR-601



4:2:2

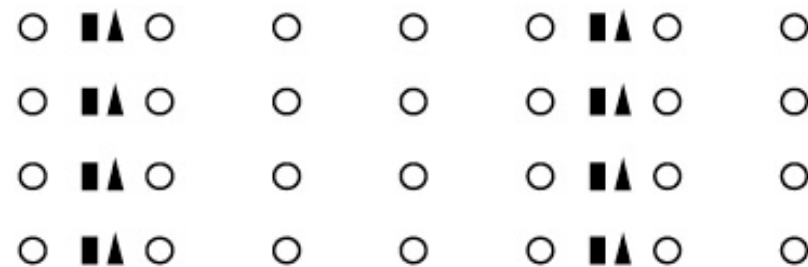
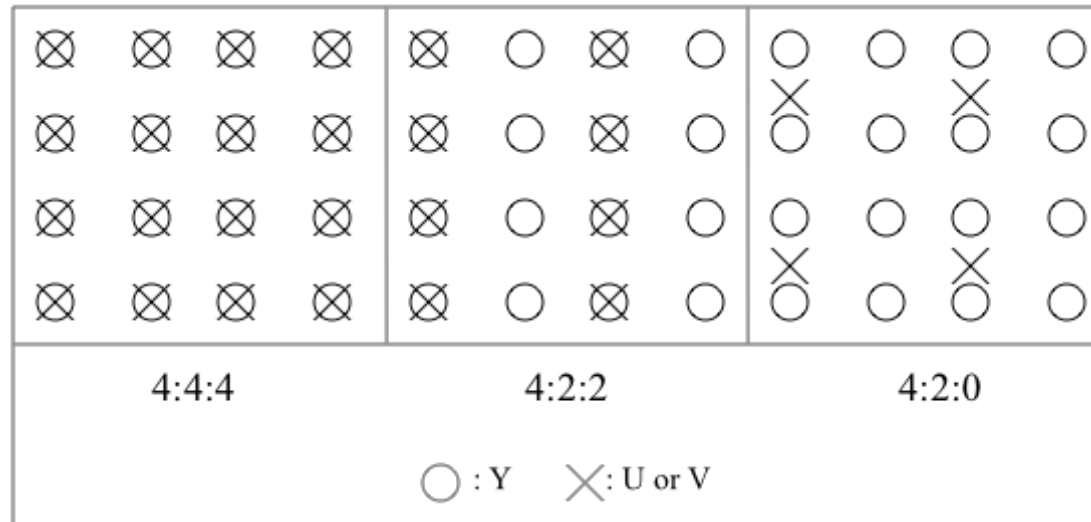
○ Y block   ■ C<sub>b</sub> block   ▲ C<sub>r</sub> block

4:2:0 SIF



4:2:0

# Άλλα Chroma Formats



4:1:1

○ Y block    ■  $C_b$  block    ▲  $C_r$  block



# SIF - Source Interchange Format

---

- Ευρώπη: 360x288@25 (μισές γραμμές των active)
- Β. Αμερική: 360x240@30
- Μετατροπή από CCIR-601 απαιτεί φιλτράρισμα (non-standard):

- Luminance

$$[-29 \ 0 \ 88 \ 138 \ 88 \ 0 \ -29]//256$$

- Chrominance (σημ.: ζυγός αριθμός συντελεστών)

$$[1 \ 3 \ 3 \ 1]//8$$

# CIF - Common Interchange Format

---

- Εφαρμογές videoconferencing
- Μισό resolution σε σχέση με CCIR-601, παίρνοντας το μέγιστο αριθμό γραμμών (625) και μέγιστο ρυθμό εικόνων (30): 360x288@30 (29.97)
- Υποδιαιρέσεις: QSIF, QCIF

# HDTV Formats

---

Video Format Supported	Native Resolution (W×H)	Pixels (Advertised Megapixels)	Aspect Ratio (X:Y)	
			Image	Pixel
720p 1280×720	1024×768 XGA	786,432 (0.8)	16:9	4:3
	1280×720	921,600 (0.9)	16:9	1:1
	1366×768 WXGA	1,049,088 (1.0)	683:384 (Approx 16:9)	1:1 Approx
1080i 1920×1080	1280×1080	1,382,400 (1.4)	32:27 (Approx 16:9)	3:2
1080p 1920×1080	1920×1080	2,073,600 (2.1)	16:9	1:1
2160p 3840×2160	3840×2160	8,294,400 (8.3)	16:9	1:1

# Formats H/Y

---

- Formats από υπολογιστές:
  - VGA 640x480
  - SVGA 800x600
  - XGA 1024x768
  - SXGA 1280x1024
  - UXGA 1600x1200
- Πρόθεμα 'W' για widescreen (π.χ. WXGA 1366x768)
- Συχνότητα 60 Hz, 75 Hz, κτλ., progressive

# Film

---

- 24 εικόνες/sec
- Δειγματοληψία μόνο χρονικά
- Μετατροπή σε NTSC/PAL δεν είναι απλή (telecine).

## Παράδειγμα Film-to-NTSC: 3:2 pull-down

- “3:2” - Πρόσθεσε ένα field σε κάθε δεύτερο frame ( $24 \times 2.5 = 60$ )
- “pull-down” - Αφού το NTSC έχει 59.94 fields/sec, επιβράδυνε τον ήχο κατά 0.1% για συγχρονισμό

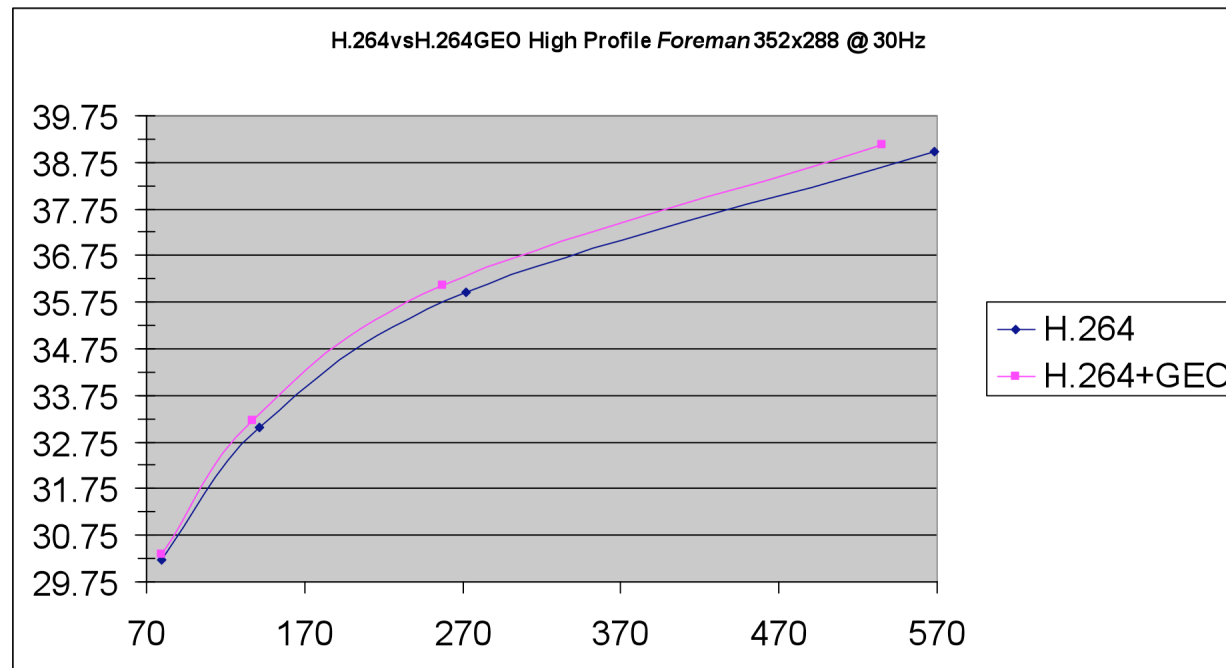
# Μέτρηση Ποιότητας σε Βίντεο

---

- Υποκειμενική - Subjective
  - Double-Stimulus Impairment Scale
  - Double-Stimulus Continuous Quality Scale
  - Single-Stimulus Continuous Quality Scale
  - Double-Stimulus Unknown Reference (2 γύροι, πρώτα για να βρεθεί το reference)
  - κτλ.MOS, Mean Opinion Score
- Αντικειμενική - Objective
  - PSNR 
$$\text{PSNR} = 10 \log_{10} \left[ \frac{255^2}{(1/N) \sum_i \sum_j (Y_{ref}(i, j) - Y_{prc}(i, j))^2} \right]$$
  - Μέση τιμή σε όλη την ακολουθία εικόνων (mean PSNR)
  - Συνήθως μόνο σε luma (Y-PSNR)

# Καμπύλες R-D

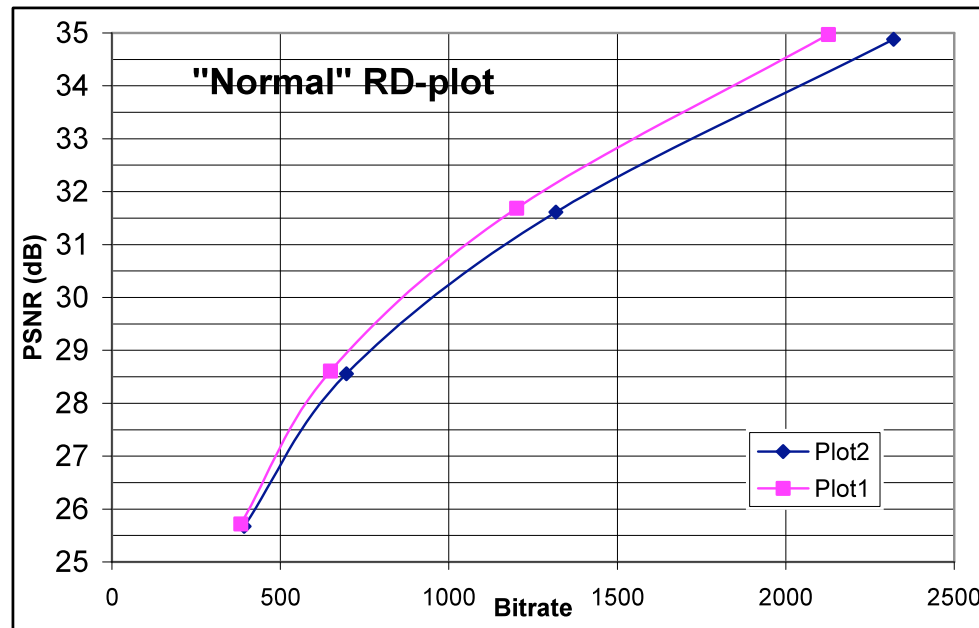
- Παράδειγμα σύγκρισης



# Bjontegaard Delta (2001)

- Ιδέα: κάνε 4 πειράματα, προσάρμοσε δύο καμπύλες, και υπολόγισε το ολοκλήρωμα της διαφοράς => ένας αριθμός σύγκρισης για όλα τα bit rates

$$\text{SNR} = (a + b \cdot \text{bit} + c \cdot \text{bit}^2) / (\text{bit} + d)$$



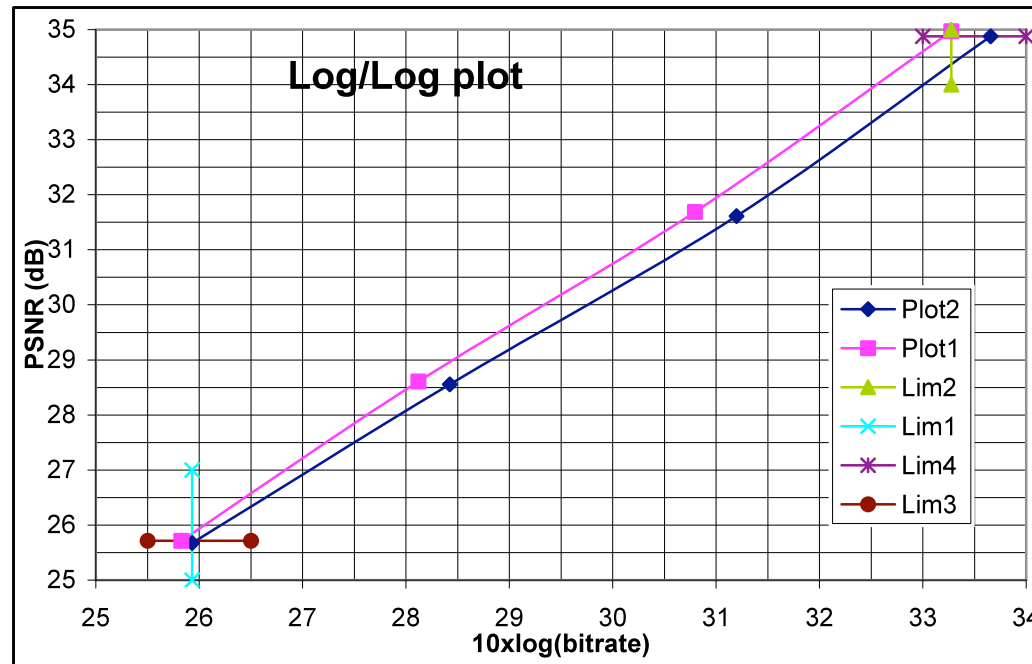
Προβλήματα:

- πόλοι σε μερικές περιπτώσεις
- υπερίσχυση υψηλών bit rates



# Bjontegaard Delta (2)

- Λύση: χρήση λογαριθμικής κλίμακας για το bit rate



- Απόσταση μεταξύ καθέτων γραμμών  $10^{0.05} = 1.122$  (ή 12.2%).
- Χρήση απλούστερου πολυωνύμου για παρεμβολή:

$$\text{SNR} = a + b \cdot \text{bit} + c \cdot \text{bit}^2 + d \cdot \text{bit}^3$$

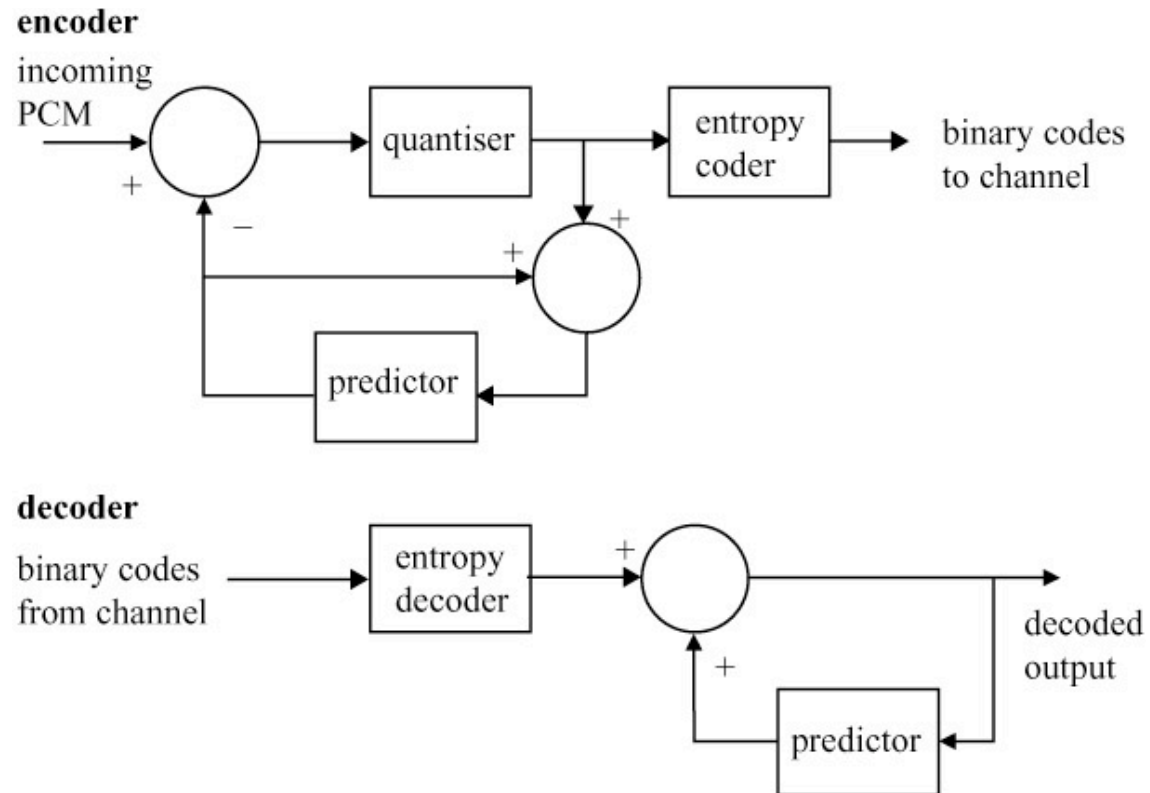
# Βασικές Αρχές Συμπίεσης

---

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

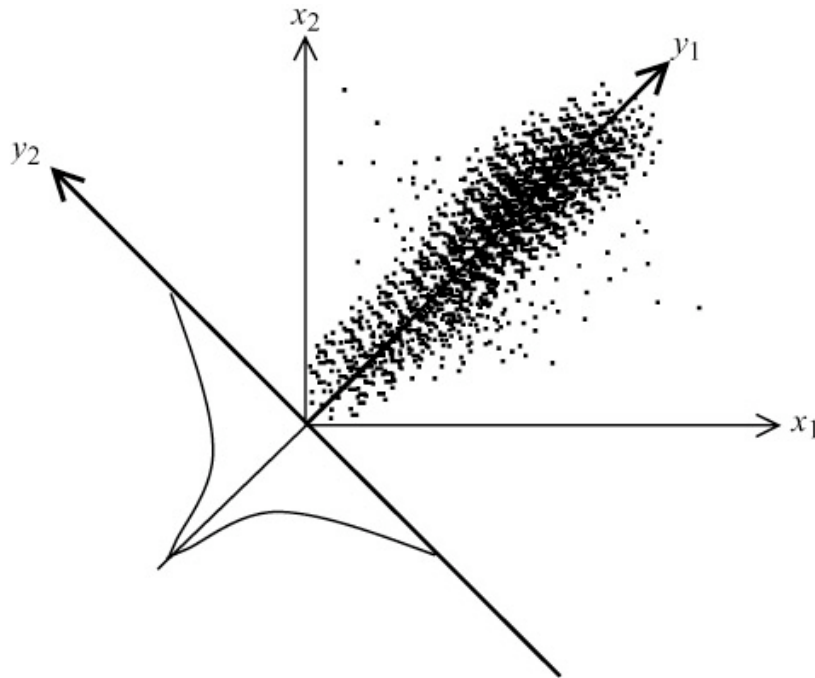
# DPCM

- 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

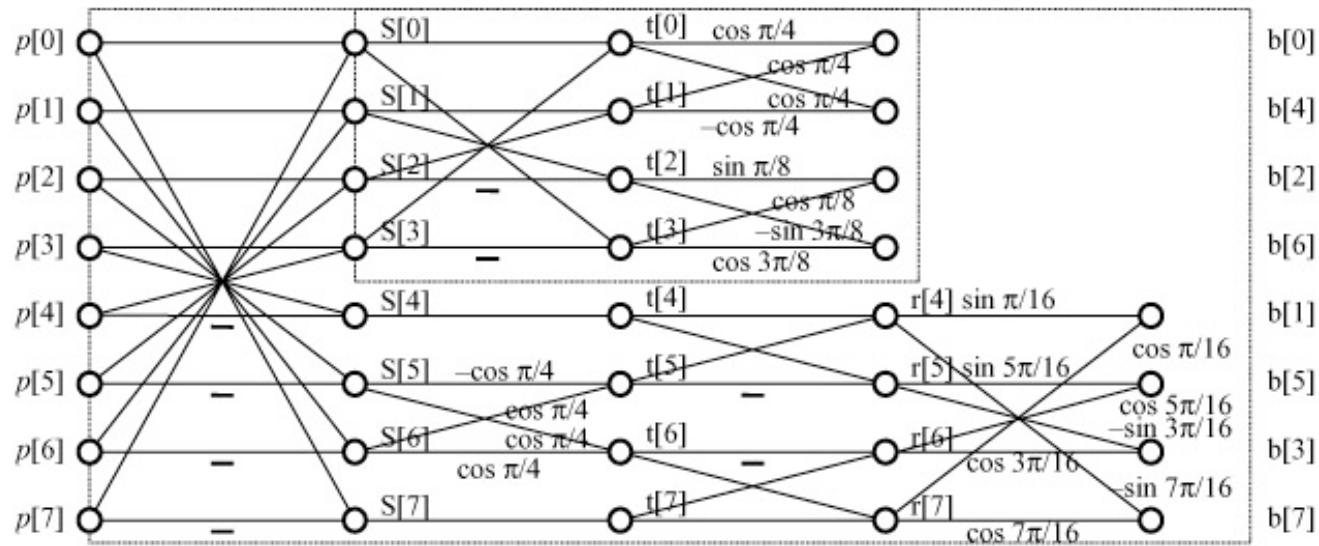
$$f(x) = \sqrt{\frac{2}{N}} \sum_{u=0}^{N-1} C(u) F(u) \cos\left(\frac{\pi(2x+1)u}{2N}\right) \quad x = 0, 1, \dots, N-1$$

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

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

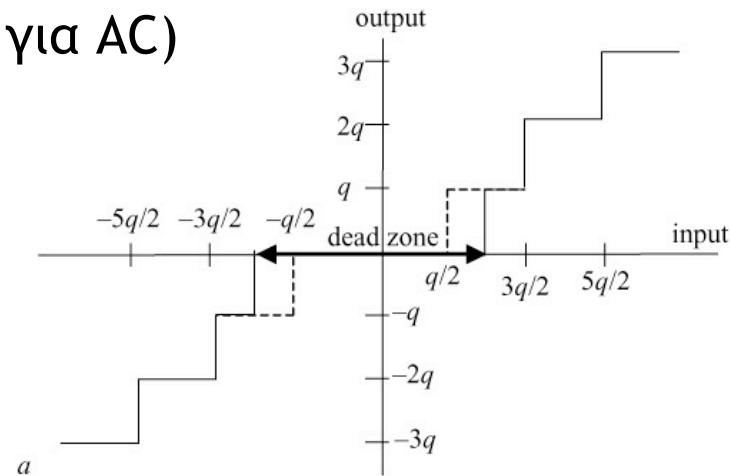
- Mismatch Control
  - Μη τυποποιημένη υλοποίηση συσσωρεύει σφάλμα στο δέκτη.
  - Προσθήκη ψευδοτυχαίου σήματος το κρατάει μέσα σε όρια

# Fast DCT

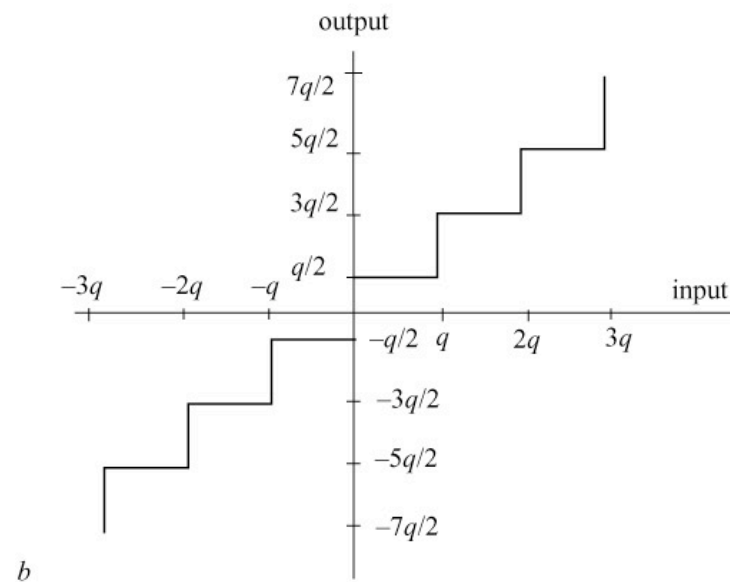


# Κβαντισμός (quantization)

- Με dead-zone (τυπικά για AC)



- Χωρίς (τυπικά για DC)



# Κβάντιση

---

- Quantization “index”

$$I(u, v) = \left\lfloor \frac{F(u, v)}{q} \right\rfloor$$

- Αντιστροφή

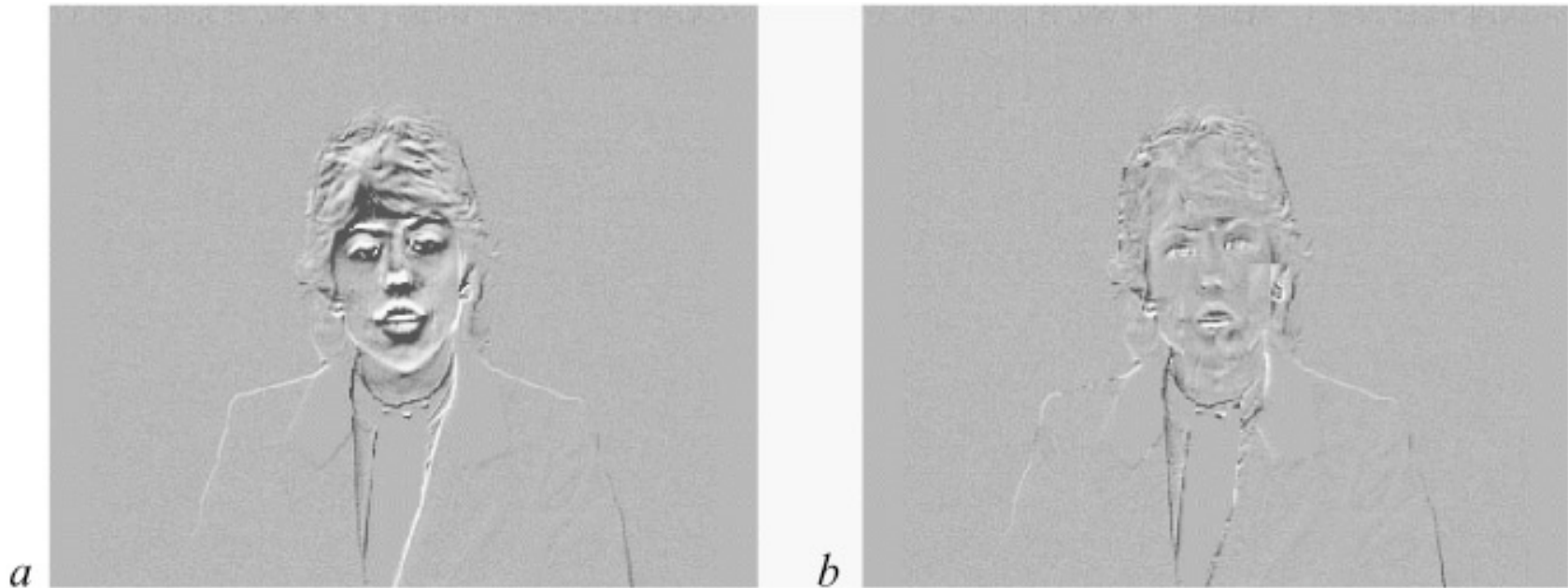
$$F^q(u, v) = \left\{ I(u, v) \pm \frac{1}{2} \right\} \times q$$



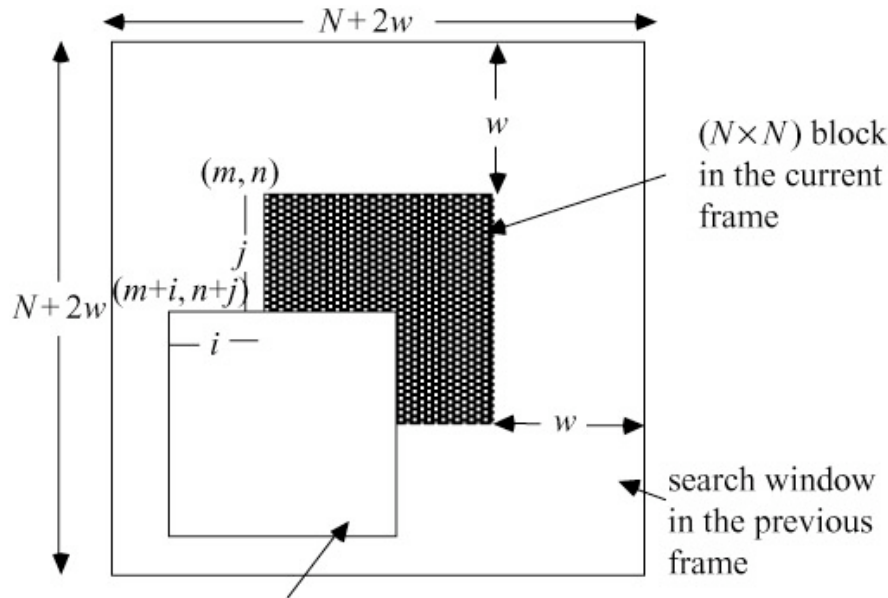
# Motion Estimation

---

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



# Μοντέλο ΜΕ

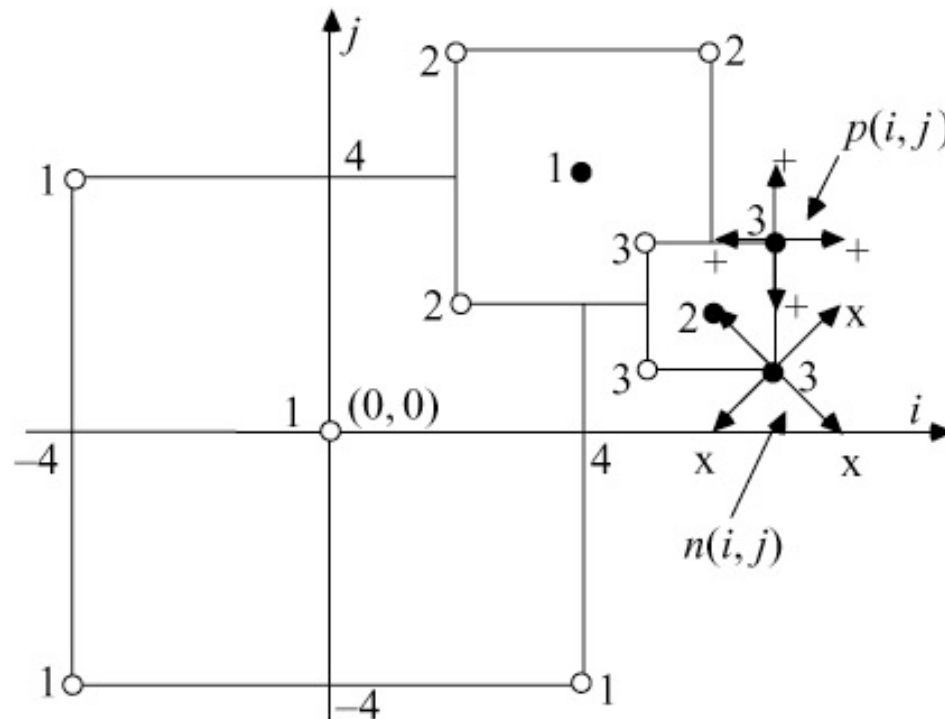


( $N \times N$ ) block under the search in the previous frame, shifted by  $i, j$

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

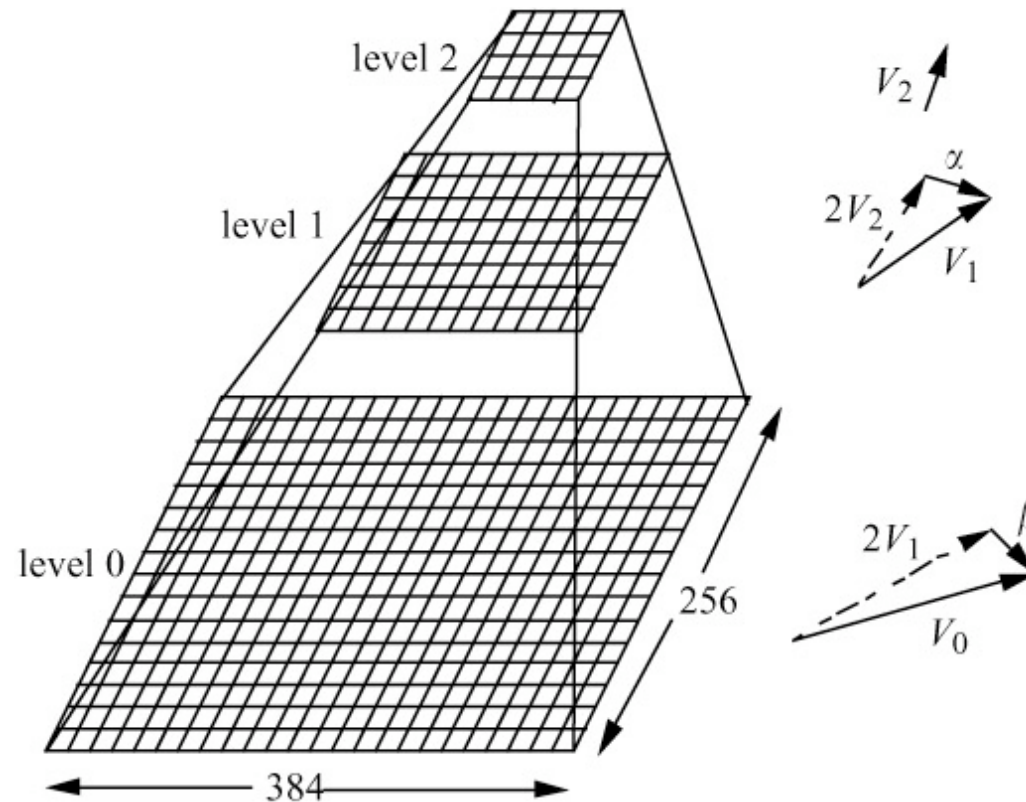
# Τεχνικές ΜΕ

- Full-search, πολυπλοκότητα  $(2w+1)^2$  - 50-65 % συνολικής υπολογιστικής πολυπλοκότητας τυπικού κωδικοποιητή
- Παράδειγμα ταχύ αλγορίθμου - Cross-Search Algorithm (CSA), πολυπλοκότητα  $5+4\log_2 w$



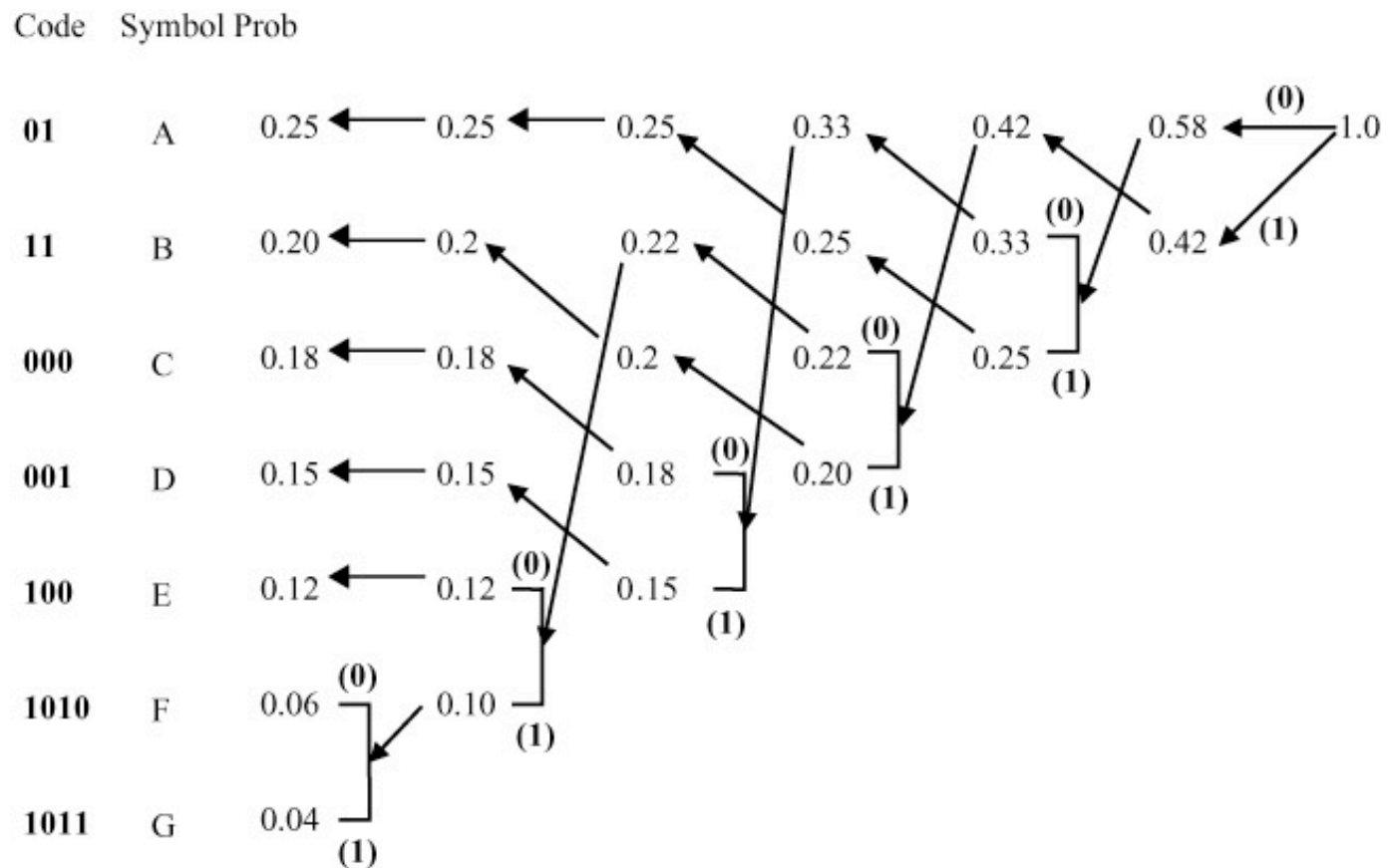
# Ιεραρχική Εκτίμηση (Hierarchical ME)

- Ιδιαίτερα χρήσιμη σε scalable codecs



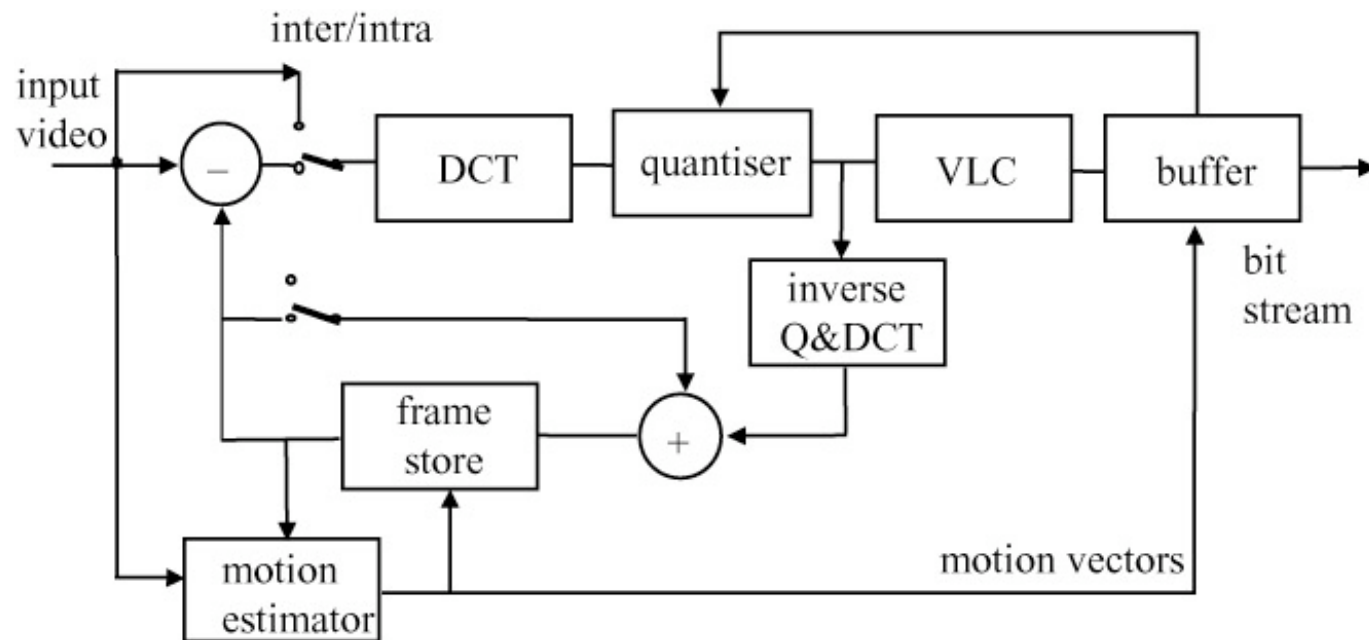
# Κωδικοποίηση Εντροπίας

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



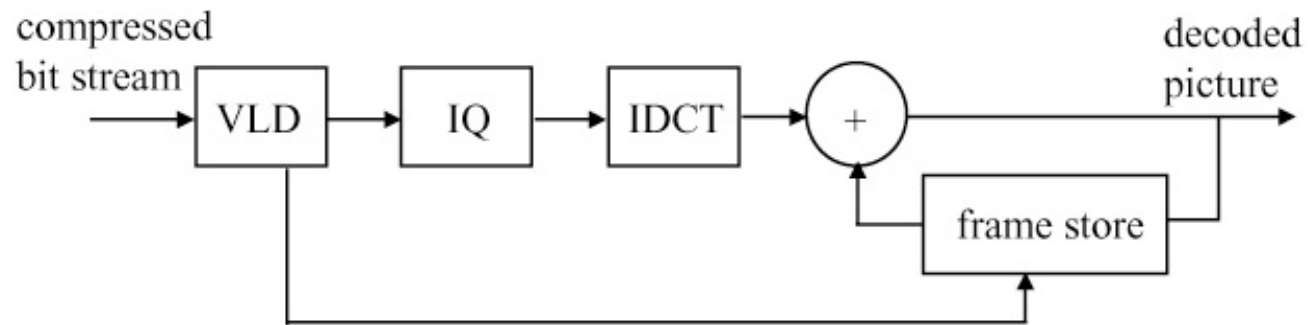
# Γενική Αρχιτεκτονική Κωδικοποιητή Βίντεο

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



# Αποκωδικοποιητής

---



# Επιτεύγματα στην Κωδικοποίηση Βίντεο

