

VIDEO ΚΑΙ ΕΦΑΡΜΟΓΕΣ

- Υπάρχουσες εφαρμογές:
 - Αναπαραγωγή αποθηκευμένου οπτικοακουστικού υλικού (εκπαιδευτικές/ψυχαγωγικές π.χ. video on demand)
 - Οπτικοακουστική επικοινωνία πραγματικού χρόνου (ένας-προς-έναν π.χ. ενημέρωση διευθυντή από υπαλλήλους ή ένας-προς-πολλούς π.χ. σεμινάριο για υπαλλήλους ή φοιτητές)

ΣΥΛΛΗΨΗ VIDEO

- Υποσυστήματα γραφικών: σύνθετο σήμα RGB, σάρωση raster
- Τηλεοράσεις: σύνθετο σήμα YUV ή YIQ, πλεκτή (interlaced) σάρωση
- *Συμπέρασμα*: απαιτείται ειδικό υλικό για μετατροπή σήματος

ΣΥΛΛΗΨΗ VIDEO

- *Κάρτες υπέρθεσης video (video overlay boards):* παρουσίαση εικόνας σε τμήμα οθόνης, συγχρονισμός, μίξη με κείμενο, γραφικά (όχι ψηφιοποίηση)
- *Ψηφιοποιητές:* ψηφιοποιούν το σήμα PAL ή NTSC από video player, camera, δυνατότητα editing, παρουσίαση οπουδήποτε στην οθόνη
- *Κάρτες συμπίεσης:* DVI (Intel 1989) πρότυπο για εφαρμογές κινούμενης εικόνας, MJPEG (Motion JPEG), MPEG.

ΤΕΧΝΙΚΕΣ ΣΥΜΠΙΕΣΗΣ

- Χωρικό και χρονικό πλεόνασμα πληροφορίας
- Συμπύεση κινούμενης εικόνας: εκμετάλλευση χωρικού (π.χ. MJPEG) ή και χρονικού πλεονάσματος (π.χ. MPEG)
- MJPEG: ευκολία επέμβασης σε πλαίσια και ανθεκτικότητα σε λάθη
- MPEG: συσχέτιση πλαισίου με προηγούμενα (άρα λάθη), αλλά μεγαλύτεροι λόγοι συμπίεσης

Η ΑΝΑΓΚΗ ΓΙΑ ΨΗΦΙΑΚΗ ΣΥΜΠΙΕΣΗ

- Στόχος: οικονομία στο εύρος φάσματος
- Ασυμπίεστο τηλεοπτικό σήμα PAL: 5 MHz
- Ψηφιακή μετάδοση: 10 MHz (θεώρημα Nyquist)
- Τουλάχιστον 8 bits/δείγμα (256 επίπεδα)
- Έγχρωμη εικόνα: τρία κανάλια χρώματος
- Συνολικά: $3 \cdot 10 \cdot 8 = 240$ Mbits/sec
- Transfer rate από/προς επεξεργαστή (PCI bus): 133 Mbits/sec
- Transfer rate CD-ROM: 60 Mbits/sec
- Συμπέρασμα: αδύνατη αποθήκευση/αναπαραγωγή

Η ΑΝΑΓΚΗ ΓΙΑ ΨΗΦΙΑΚΗ ΣΥΜΠΙΕΣΗ

- Τυπική κινηματογραφική ταινία 90 min (5400 sec): 162 GB (μεγαλύτερο από συνήθη χωρητικότητα σκληρού δίσκου)
- CD-DA: 75 λεπτά ασυμπίεστου στερεοφωνικού ήχου στα 44.1 KHz με 16 bit/sample → 650 MB
- Αποθήκευση ψηφιακού video στο CD: 30 sec
- DVD (Digital Versatile Disk): 4.7 GB (θέλω 35 DVD!)

ΠΡΟΤΥΠΟ ΣΥΜΠΙΕΣΗΣ MPEG

- MPEG (Moving Picture Experts Group): ομάδα ISO/IEC JTC1/SC29/WG11 (1988)
- Συντονιστής επιτροπής MPEG: Leonardo Chiariglione («πατέρας» του MPEG)
- MPEG: οικογένεια τυποποιήσεων (MPEG-1, MPEG-2, MPEG-4, MPEG-7, MPEG-21)

ΠΡΟΤΥΠΟ ΣΥΜΠΙΕΣΗΣ MPEG

- *MPEG-1 (Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbit/sec):*
 - ανάλυση 352×240 pixels (NTSC) ή 352×288 pixels (PAL)
 - ποιότητα VHS video
 - Αποθήκευση σε CD-ROM, Video-CD, CD-i
 - Εφαρμογές με ρυθμό μετάδοσης 4-5 Mbits/sec (όχι καλά αποτελέσματα)

ΠΡΟΤΥΠΟ ΣΥΜΠΙΕΣΗΣ MPEG

- *MPEG-2 (Generic coding of moving pictures and associated audio information):*
 - Ανάλυση 704×480 pixels (NTSC) ή 704×576 pixels (PAL) (broadcast quality)
 - Εικόνα πλεκτής σάρωσης (interlaced)
 - Ρυθμός μετάδοσης 3-10 Mbits/sec
 - Εφαρμογές στην καλωδιακή τηλεόραση (CableTV), στη δορυφορική (Direct Broadcasting Satellite TV), αποθήκευση ταινιών σε DVD.

ΠΡΟΤΥΠΟ ΣΥΜΠΙΕΣΗΣ MPEG

- *MPEG-4 (Coding of audio-visual objects):*
 - Audio-visual objects (AVO): οντότητες που απαρτίζουν την εικόνα (σχήματα, ήχοι κλπ.)
 - Ανάλυση εικόνας 176×144 pixels
 - Ρυθμός μετάδοσης 4.8-64 Kbits/sec (δίκτυα με μικρό bandwidth)
 - Εφαρμογές επικοινωνίας πολυμέσων (video-phone, video-conference, video e-mail, electronic news κλπ.)

ΠΡΟΤΥΠΟ ΣΥΜΠΙΕΣΗΣ MPEG

- *MPEG-7*: πρότυπο κωδικοποίησης με αναπαράσταση περιεχομένου (content representation) για αναζήτηση πληροφοριών σε εφαρμογές πολυμέσων
- CPB (Constrained Parameters Bitstream): ορίζει διαστάσεις MPEG σημάτων (επιτρέπονται και υψηλότερες αναλύσεις, χωρίς εγγυημένη αναπαραγωγή)
- MPEG-2 έως 1920×1080 pixels, MPEG-1 έως 4095×4095 pixels

ΠΡΟΤΥΠΟ ΣΥΜΠΙΕΣΗΣ MPEG

- MPEG-2 προσφέρει καλύτερη εικόνα και λόγους συμπίεσης από το MPEG-1 και είναι συμβατό (backwards compatible)
- MPEG-3 προοριζόταν για Τηλεόραση Υψηλής Ευκρίνειας (HDTV – High Definition TV) αλλά MPEG-4 κάλυψε τις απαιτήσεις

ΠΟΙΟΤΙΚΗ ΑΝΟΧΗ

- Δεχόμαστε ένα ποσοστό παραμόρφωσης
- Αναβαθμίζουμε το μέσο όρο ανάλυσης σκηνών
- Χρυσή τομή ανάμεσα σε ποιότητα και bandwidth
- Υποκειμενικές μετρήσεις που βασίζονται σε έμπειρους εκτιμητές
- Αποδεκτή συμπίεση: ανεκτές ατέλειες (artifacts) εικόνας
- Ήχος προσφέρεται για μεγαλύτερα ποσοστά συμπίεσης

ΠΛΕΟΝΑΣΜΟΣ

- *Χωρικός πλεονασμός (Spatial Redundancy):*
Γειτονικές περιοχές στο ίδιο frame είναι παρόμοιες.
- *Χρονικός πλεονασμός (Temporal Redundancy):*
Αντίστοιχες περιοχές σε διαδοχικά frames είναι παρόμοιες.
- Εκμετάλλευση χρονικού πλεονασμού στην αναλογική τηλεόραση παλιότερα και τώρα σε ψηφιακή μετάδοση

DCT Coding (Discrete Cosine Transform Coding)

- Διακριτός μετασχηματισμός συνημίτονου (DCT):

$$DCT(i, j) = \frac{1}{\sqrt{2N}} C(i)C(j) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} pixel(x, y) \cos\left[\frac{(2x+1)i\pi}{2N}\right] \cos\left[\frac{(2y+1)j\pi}{2N}\right]$$

$$\text{Όπου } C(x) = \begin{cases} 0.7071 & , x=0 \\ 1 & , x>0 \end{cases}$$

- Κάθε συντελεστής μεταφέρει πληροφορία που αντιστοιχεί σε τμήμα του φάσματος
- Ανθρώπινη όραση αντιλαμβάνεται περισσότερο χαμηλές συχνότητες

DCT Coding (Discrete Cosine Transform Coding)

- Αντίστροφος διακριτός μετασχηματισμός συνημίτονου (IDCT):

$$pixel(x, y) = \frac{1}{\sqrt{2N}} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} C(i)C(j) DCT(i, j) \cos\left[\frac{(2x+1)i\pi}{2N}\right] \cos\left[\frac{(2y+1)j\pi}{2N}\right]$$

- Λαμβάνεται σχεδόν ανέπαφη η αρχική πληροφορία (εκτός από σφάλματα στρογγυλοποίησης)

ΚΒΑΝΤΟΠΟΙΗΣΗ

- Απαλασσόμαστε από σημαντικό μέρος της πληροφορίας
- Μετατροπή σήματος άπειρων τιμών σε σήμα ορισμένων διακριτών τιμών
- *Παράδειγμα:* κβαντοποίηση εικόνας εκατομμυρίων χρωμάτων σε εικόνα 256 τιμών (πρότυπο JPEG)

ΚΒΑΝΤΟΠΟΙΗΣΗ

- *Παράδειγμα:*
 - 45 στο δυαδικό 101101 (6 bits)
 - Με 4 bits είναι 1011 (δηλ. 11)
 - Με 3 bits είναι 101 (δηλ. 5) κτλ.
- Συμπέρασμα: εισάγεται σφάλμα ανάλογο του πλήθους bits που απορρίπτονται (lossy compression)
- Διαίρεση τιμών με σταθερές τιμές (uniform quantization) ή με τιμές από πίνακα (quantization table) όπως στο MPEG (intra-frame και inter-frame)

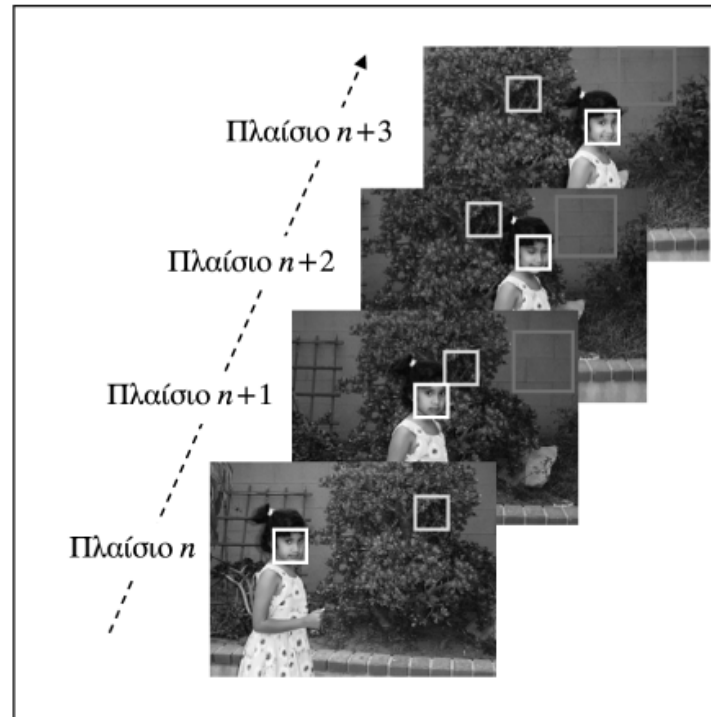
ΤΜΗΜΑΤΙΚΗ ΠΡΟΒΛΕΨΗ ΚΙΝΗΣΗΣ

(BLOCK MOTION COMPENSATION)

- Αποτελεσματική περιγραφή εικόνων με κίνηση
- Δια-πλαισιακή (inter-frame) κωδικοποίηση
- Περιγραφή με τμήματα που παραμένουν ίδια και διανύσματα που δείχνουν μετακίνηση άλλων τμημάτων
- *Προβλήματα:* τμήματα αλλάζουν θέση αλλά και χρώμα, εμφανίζονται νέα τμήματα
- MPEG: καθορίζεται η αναπαράσταση με διανύσματα κίνησης (motion vectors) αλλά δεν καθορίζεται η μέθοδος εύρεσής τους

ΤΜΗΜΑΤΙΚΗ ΠΡΟΒΛΕΨΗ ΚΙΝΗΣΗΣ (BLOCK MOTION COMPENSATION)

Σχήμα 8-2 Χρονικός πλεονασμός στο βίντεο. Βλέπετε τέσσερα πλαίσια μίας ακολουθίας βίντεο. Στο προσκήνιο κινείται ένα παιδί, ενώ στο παρασκήνιο τα αντικείμενα είναι στατικά. Επίσης, κινείται και η κάμερα. Μολονότι κάθε πλαίσιο είναι διαφορετικό, υπάρχουν περιοχές στο παρασκήνιο και στο προσκήνιο που δεν μεταβάλλονται ή που μοιάζουν πολύ μεταξύ τους.



(Εικόνα από «Συστήματα Πολυμέσων», Havaladar & Medioni)

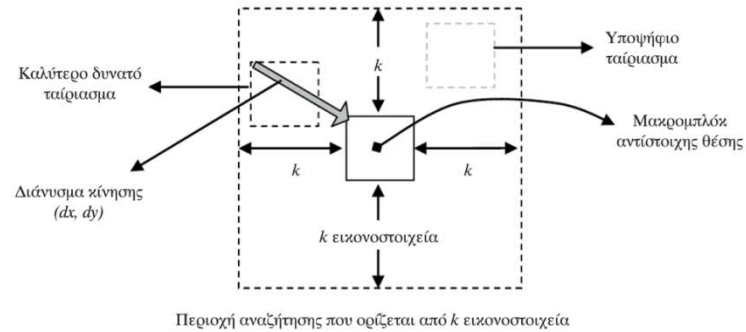
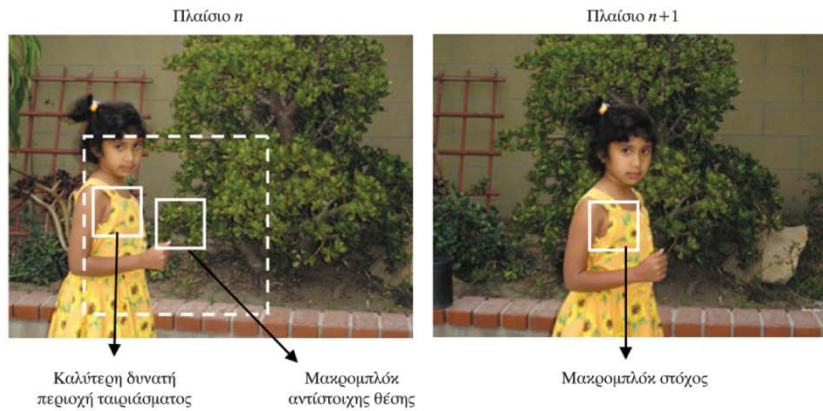
ΤΜΗΜΑΤΙΚΗ ΠΡΟΒΛΕΨΗ ΚΙΝΗΣΗΣ

(BLOCK MOTION COMPENSATION)

- Συνήθως χρησιμοποιείται το απόλυτο λάθος (Absolute Error):

$$AE(d_x, d_y) = \sum_{i=0}^{15} \sum_{j=0}^{15} \left| f(i, j) - g(i - d_x, j - d_y) \right|$$

- $f(i, j)$ το τρέχον μπλοκ και $g(i, j)$ το macroblock αναφοράς (περιοχή αναζήτησης)
- Το macroblock που δίνει ελάχιστο σφάλμα αντιστοιχεί στο διάνυσμα κίνησης (d_x, d_y)
- Πλήρης αναζήτηση (full search): Απλή αλλά υπολογιστικά πολύπλοκη



Μακρομπλόκ σε υπέρθεση με περιοχή αναζήτησης στην αντίστοιχη θέση



Μακρομπλόκ σε υπέρθεση με υποψήφια θέση



Μακρομπλόκ με υπέρθεση στη θέση καλύτερου ταυρίσματος που δείχνει το διάνυσμα κίνησης

Σχήμα 8-7

(Εικόνα από «Συστήματα Πολυμέσων», Havalдар & Medioni)

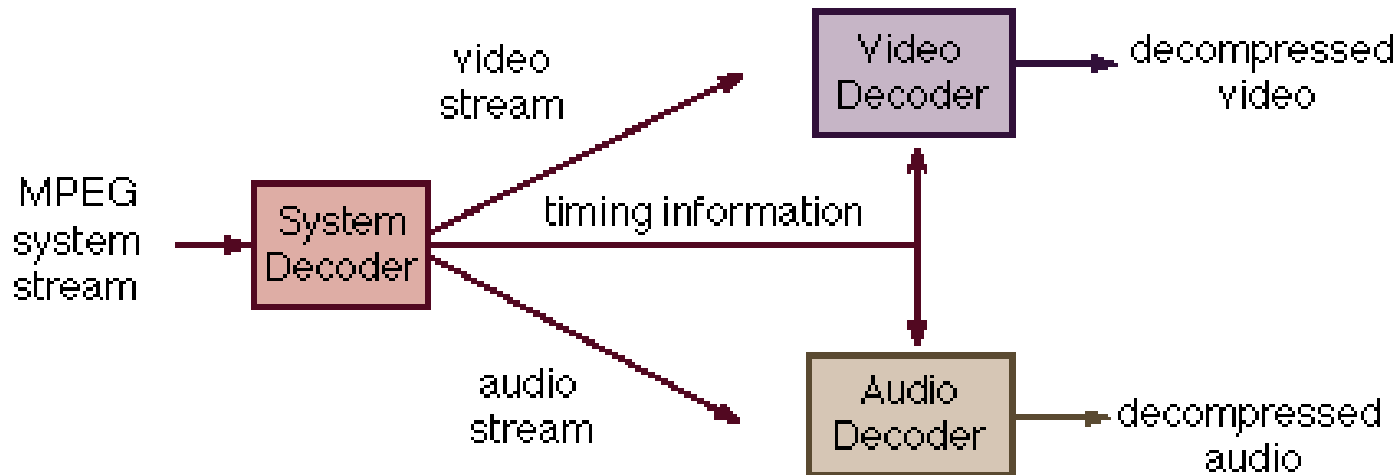
ΤΜΗΜΑΤΙΚΗ ΠΡΟΒΛΕΨΗ ΚΙΝΗΣΗΣ

(BLOCK MOTION COMPENSATION)

- Αναζήτηση τριών βημάτων (TSS – Three Step Search):
 - Υπολογίζεται το ΑΕ στο κέντρο και σε οκτώ περιοχές της περιοχής αναζήτησης (μεγέθους 32×32 pixels)
 - Περιοχή με ελάχιστο ΑΕ γίνεται κέντρο νέας αναζήτησης (μισού μεγέθους)
 - Επανάληψη διαδικασίας τρεις φορές (three steps)

ΑΝΑΤΟΜΙΑ ΣΗΜΑΤΟΣ MPEG

- Τρία επίπεδα ρεύματος (stream) MPEG:



ΑΝΑΤΟΜΙΑ ΣΗΜΑΤΟΣ MPEG

- **Επίπεδο system:** πληροφορίες σχετικά με συγχρονισμό, ροή σήματος, σημεία αναφοράς για τυχαία προσπέλαση και για διαχωρισμό video και audio και συγχρονισμένη απεικόνιση
- **Επίπεδα video και audio:** κωδικοποιημένη εικόνα και ήχος αντίστοιχα
- Κωδικοποίηση επιπέδων ταυτόχρονη είτε ξεχωριστή
- Συνένωση δεδομένων με **πολυπλεξία (multiplexing)** (real time είτε όχι)

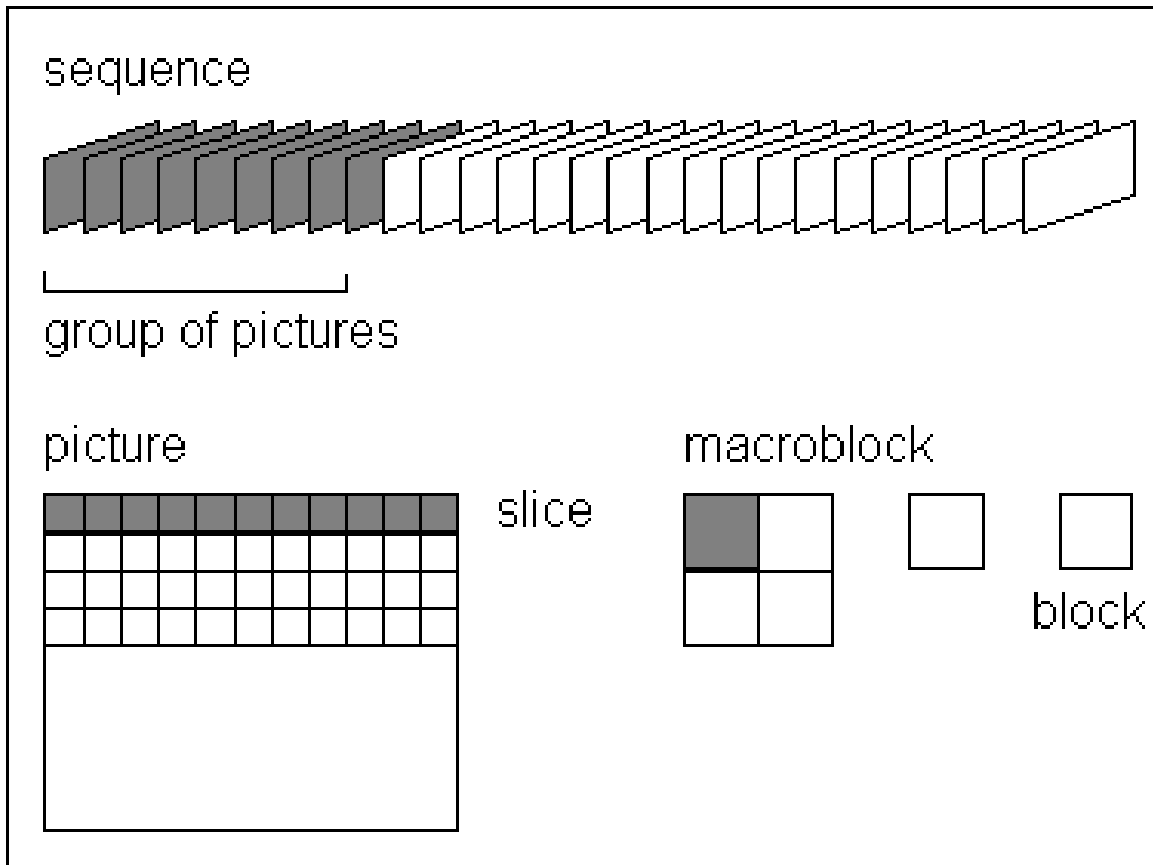
ΑΝΑΤΟΜΙΑ ΣΗΜΑΤΟΣ MPEG

- MPEG coding/decoding γίνεται με hardware ή software
- Λύση software: φθηνή αλλά χειρότερη ποιότητα εικόνας και απαίτηση υπολογιστικής ισχύος
- Λύση hardware: ακριβότερη, αλλά αυτόνομα συστήματα (encoders/decoders) και καλύτερη ποιότητα
- Κωδικοποιήσεις όπως Cinepak, Indeo, MJPEG στηρίζονται στο JPEG.

ΑΝΑΤΟΜΙΑ ΣΗΜΑΤΟΣ MPEG

- Κωδικοποίηση MPEG επεκτείνει βασικές αρχές JPEG παραμετροποιώντας την κίνηση (μεγαλύτερα ποσοστά συμπίεσης)
- Τα πλαίσια (frames) χωρίζονται σε τμήματα (slices)
- Τα slices περιέχουν macroblocks (και αυτά με τη σειρά τους blocks μεγέθους 8×8 pixels)
- Τα slices περιορίζουν διάδοση σφαλμάτων (error propagation) - βοηθούν στην απόκρυψη σφαλμάτων (error concealment)

ANATOMIA ΣΗΜΑΤΟΣ MPEG



ΑΝΑΤΟΜΙΑ ΣΗΜΑΤΟΣ MPEG

- Δειγματοληψία χρωματικών συνιστωσών αναλογικού σήματος τηλεόρασης (Cr και Cb) στα 6,75 MHz
- Δειγματοληψία φωτεινότητας (Y) στα 13,5 MHz (διπλάσια συχνότητα)
- Και τα δύο πολλαπλάσια του 2,25 MHz (ΕΚΠ ρυθμού σάρωσης γραμμών σε PAL και NTSC)
- Στην ψηφιακή τους μορφή για κάθε macroblock τρεις πίνακες (ένας 16×16 και δύο 8×8 pixels)
- Στους πίνακες αυτούς εφαρμόζεται ο DCT, οι άλλες διαδικασίες συμπίεσης και η πρόβλεψη κίνησης

ΑΝΑΤΟΜΙΑ ΣΗΜΑΤΟΣ MPEG

