

Υπολογιστικά συστήματα: ψηφιακά μέσα

Διδάσκων: Α. Ντελόπουλος

ΑΠΘ/ΤΗΜΜΥ/2003-4

17-Υπολογιστικά Συστήματα – Ψηφιακά Μέσα

Διδάσκων:
Α. Ντελόπουλος

Τα ψηφιακά μέσα είναι πληροφορίες που...

- είναι αντιληπτές από τις αισθήσεις μας
 - προς το παρόν: όραση, ακοή, αφή
 - στο μέλλον: όσφρηση και γεύση(;)
- μπορούν οι υπολογιστές να τις διαχειριστούν
 - σύλληψη, επεξεργασία, αποθήκευση-μετάδοση, αναπαραγωγή
- αναπαρίστανται από bits

ΑΠΘ/ΤΗΜΜΥ/2003-4

17-Υπολογιστικά Συστήματα – Ψηφιακά Μέσα

Διδάσκων:
Α. Ντελόπουλος

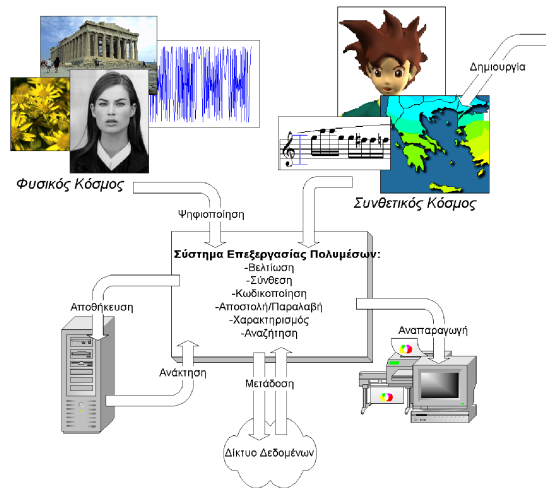
είδη ψηφιακών μέσων

- Κείμενο: Μη δομημένο (π.χ. ASCII text documents), Δομημένο (word processors, postscript, html, κ.λπ.), hypertext
- Ήχος: μουσική, φωνή, συνθετικός δομημένος ήχος (π.χ. MIDI)
 - Μονοκαναλικός – Μονοφωνικός
 - Πολυκαναλικός (Στερεοφωνικός, surround, κ.λπ.)
- Εικόνες: Φυσικές φωτογραφίες, συνθετικές εικόνες (2D γραφικά), Fax, στερεοσκοπικές και πανοραμικές εικόνες, ιατρικές εικόνες
 - Έγχρωμες
 - Αποχρώσεων του γκρι
 - Δυαδικές
- Βίντεο: Φυσικό βίντεο (Ταινίες, ειδήσεις, κ.λπ), 3D γραφικά

μερικές χρήσεις ψηφιακών μέσων...

- μετάδοση και λήψη ψηφιακής τηλεόρασης
 - βίντεο κατ' απαίτηση
- συστήματα τηλεφωνίας (σταθερής & κινητής)
- fax
- συστήματα διαχείρισης απεικονιστικών ιατρικών εξετάσεων
 - Medical Picture Archiving and Communication Systems – PACS
- παιχνιδομηχανές
- επεξεργασία και έκδοση εντύπων
 - Desk-top Publishing DTP
 - Pre-press
- ηλεκτρονική αρχειοθέτηση οπτικοακουστικού υλικού και ηλεκτρονικών εγγράφων

συστήματα διαχείρισης ψηφιακών μέσων



ΑΠΘ/ΤΗΜΜΥ/2003-4

17-Υπολογιστικά Συστήματα – Ψηφιακά Μέσα

Διδάσκων:
Α. Ντελόπουλος

ψηφιοποίηση

δειγματοληψία + κβαντισμός

ΑΠΘ/ΤΗΜΜΥ/2003-4

17-Υπολογιστικά Συστήματα – Ψηφιακά Μέσα

Διδάσκων:
Α. Ντελόπουλος

δειγματοληψία

- από ένα αναλογικό μέσο παράγεται μια ακολουθία δειγμάτων
 - μίας διάστασης: ήχος
 - δύο διαστάσεων: εικόνες
 - τριών διαστάσεων: βίντεο
- δείγμα: μέτρηση του σήματος σε μία συγκεκριμένη θέση
- θέση: στο χρόνο, στο χώρο, στο χώρο-χρόνο
- απώλεια πληροφορίας;
 - ίσως όχι ... αν δειγματοληπτούμε αρκετά συχνά

ΑΠΘ/ΤΗΜΜΥ/2003-4

17-Υπολογιστικά Συστήματα – Ψηφιακά Μέσα

Διδάσκων:

A. Ντελόπουλος

δειγματοληψία μονοδιάστατων σημάτων

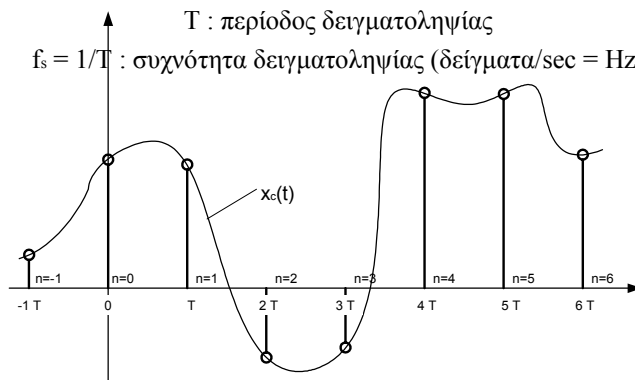
$x_c(t)$: αναλογικό σήμα ήχου

$x(n)$: δειγματοληπτημένο σήμα = ακολουθία δειγμάτων

$$x(n) = x_c(Tn), \quad n \text{ ακέραιος}$$

T : περίοδος δειγματοληψίας

$f_s = 1/T$: συχνότητα δειγματοληψίας (δείγματα/sec = Hz)



ΑΠΘ/ΤΗΜΜΥ/2003-4

17-Υπολογιστικά Συστήματα – Ψηφιακά Μέσα

Διδάσκων:

A. Ντελόπουλος

δειγματοληψία διδιάστατων σημάτων

$x_c(t_v, t_h)$: αναλογικό σήμα εικόνας

$x(n_v, n_h)$: δειγματοληπτημένο σήμα = 2Δ ακολουθία δειγμάτων

$$x(n_v, n_h) = x_c(T_v n_v, T_h n_h), \quad n_v, n_h \text{ ακέραιοι}$$

T_v : περίοδος δειγματοληψίας στην κατακόρυφη κατεύθυνση

T_h : περίοδος δειγματοληψίας στην οριζόντια κατεύθυνση

$f_{sv} = 1/T_v$: συχνότητα κατακόρυφης δειγματοληψίας
(γραμμές / μονάδα μήκους)

$f_{sh} = 1/T_h$: συχνότητα οριζόντιας δειγματοληψίας
(στήλες / μονάδα μήκους)

δειγματοληψία τρισδιάστατων σημάτων

$x_c(t_v, t_h, t)$: αναλογικό σήμα βίντεο

$x(n_v, n_h, n)$: δειγματοληπτημένο σήμα = 3Δ ακολουθία δειγμάτων

$$x(n_v, n_h, n) = x_c(T_v n_v, T_h n_h, T n), \quad n_v, n_h, n \text{ ακέραιοι}$$

T_v : περίοδος δειγματοληψίας στην κατακόρυφη κατεύθυνση

T_h : περίοδος δειγματοληψίας στην οριζόντια κατεύθυνση

T : περίοδος χρονικής δειγματοληψίας

$f_{sv} = 1/T_v$: συχνότητα κατακόρυφης δειγματοληψίας
(γραμμές / μονάδα μήκους)

$f_{sh} = 1/T_h$: συχνότητα οριζόντιας δειγματοληψίας
(στήλες / μονάδα μήκους)

$f = 1/T$: συχνότητα χρονικής δειγματοληψίας
(καρέ / μονάδα χρόνου)

μπορούμε να πάμε προς τα πίσω;

Από τα δείγματα, $x(\mathbf{n})$, μπορούμε να ανακατασκευάσουμε το αρχικό σήμα $x_c(t)$ υπό δύο προϋποθέσεις:

1. Το αρχικό σήμα $x_c(t)$ έχει πεπερασμένο εύρος συχνοτήτων

= έχει μια μέγιστη συχνότητα μεταβολής f_{\max}

2. Η δειγματοληψία γίνεται αρκετά πυκνά:

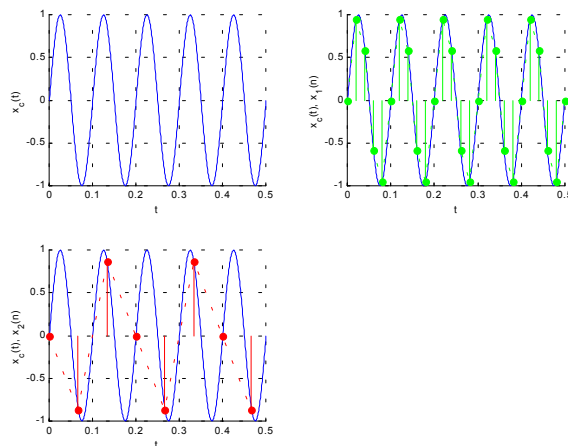
$$f_s > 2f_{\max}$$

(η συχνότητα $2f_{\max}$ λέγεται συχνότητα Nyquist)

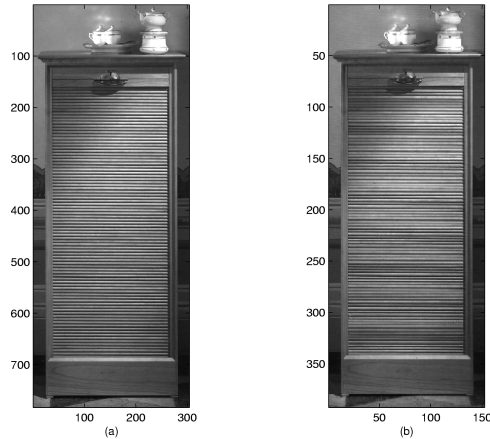
3. Αλλιώς έχουμε αναδίπλωση συχνότητας

= στο ανακατασκευασμένο σήμα εμφανίζονται συχνότητες που δεν υπάρχουν στο αρχικό

αναδίπλωση συχνότητας σε σήματα 1Δ



αναδίπλωση συχνότητας σε σήματα 2Δ (I)



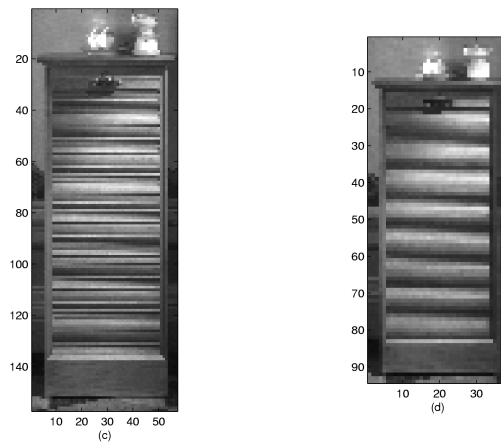
ΑΠΘ/ΤΗΜΜΥ/2003-4

17-Υπολογιστικά Συστήματα – Ψηφιακά Μέσα

Διδάσκων:

Α. Ντελόπουλος

αναδίπλωση συχνότητας σε σήματα 2Δ (II)



ΑΠΘ/ΤΗΜΜΥ/2003-4

17-Υπολογιστικά Συστήματα – Ψηφιακά Μέσα

Διδάσκων:

Α. Ντελόπουλος

κβαντισμός

- ανάγκη αναπαράσταση των δειγμάτων με πεπερασμένο αριθμό bits
 - με b bits $\rightarrow n=2^b$ διαφορετικά κβαντικά σύμβολα
- η περιοχή τιμών των δειγμάτων (πεδίο τιμών) διαμερίζεται σε κβαντικές ζώνες
- οι τιμές δειγμάτων της ίδιας ζώνης αντιστοιχίζονται σε ένα κβαντικό σύμβολο που αντιπροσωπεύει τη ζώνη
- αντιστοιχεί σε στρογγυλοποίηση
 - οποιαδήποτε τιμή στο εσωτερικό της ζώνης αντιστοιχίζεται στο ίδιο σύμβολο
- απώλεια πληροφορίας

κβαντιστής

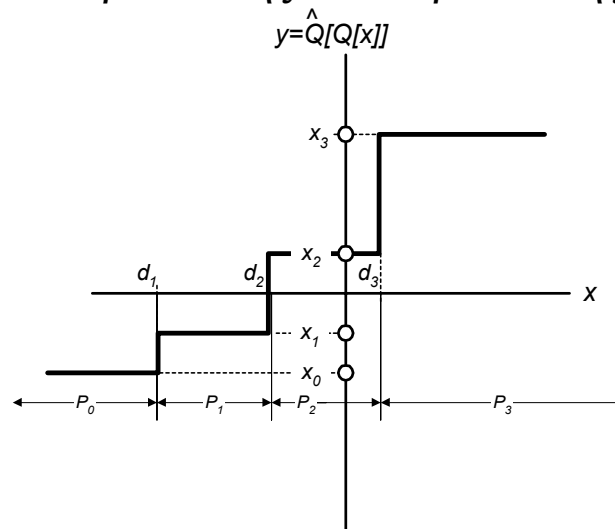
- Ο κβαντιστής $s=Q[x]$ με n στάθμες ορίζεται από:
 1. Τις ζώνες κβαντισμού p_i (=διαστήματα τιμών του x)
 - τα διαστήματα ορίζονται από τα άκρα τους, d_i :
 - $p_0 = (-\infty, d_1)$
 - $p_1 = [d_1, d_2)$
 - ...
 - $p_{n-1} = [d_{n-1}, \infty)$
 2. Τα σύμβολα s_0, \dots, s_{n-1}
- Η έξοδος του κβαντιστή είναι το σύμβολο s_i αν το x ανήκει στη ζώνη p_i
- Για n στάθμες χρειάζονται $b=\log_2(n)$ bits/σύμβολο
 - ισοδύναμα με b bits ορίζονται $n=2^b$ σύμβολα
 - π.χ. $n=8=2^3 \rightarrow b=3$ αλλά. Αναγκαστικά βέβαια και για $n=7 \rightarrow b=3$

αποκβαντιστής

- Ο αποκβαντιστής $y=Q_D [s]$ ορίζεται από τις στάθμες κβαντισμού x_0, \dots, x_{n-1}
- Τότε

$$Q_D [s_i] \rightarrow x_i$$

κβαντιστής - αποκβαντιστής



σφάλμα κβαντισμού

- η διαδικασία κβαντισμού-αποκβαντισμού εισάγει πάντα σφάλμα
 - αντίστοιχο του σφάλματος στρωγγυλοποίησης
- Όσο αυξάνει ο αριθμός n των συμβόλων (= ο αριθμός των σταθμών = ο αριθμός των ζωνών) το σφάλμα μειώνεται
- Γενικά το σφάλμα δίνεται από την έκφραση:

$$\varepsilon = x - Q_d[Q[x]]$$

Ψηφιακή Αναπαράσταση Μέσων

... τι ακριβώς περιγράφουν τα δείγματα;
... πόσα δείγματα χρησιμοποιούνται;
... πόσα bits ανά δείγμα χρησιμοποιούνται;

ψηφιακή αναπαράσταση ήχου (I)

- κάθε δείγμα έχει τιμή αντίστοιχη προς
 - την πίεση που καταγράφει ένα μικρόφωνο σε μία συγκεκριμένη χρονική στιγμή ή ισοδύναμα
 - την ηλεκτρική τάση στην έξοδο του μικροφώνου
- ψηφιακός μονοκαναλικός ήχος
 - μονοδιάστατη ακολουθία βαθμωτών δειγμάτων
- πολυκαναλικός ήχος
 - μονοδιάστατη ακολουθία διανυσμάτων ή ισοδύναμα
 - πολλαπλές μονοδιάστατες ακολουθίες βαθμωτών δειγμάτων: μία για κάθε κανάλι

ψηφιακή αναπαράσταση ήχου (II)

- μονοκαναλικός ήχος:
 $\dots x(n) \quad x(n+1) \quad x(n+2) \dots$
- πολυκαναλικός ήχος με k κανάλια
 $\dots \underline{x}(n) \quad \underline{x}(n+1) \quad \underline{x}(n+2) \dots$ όπου $\underline{x}(n) = \begin{bmatrix} x_1(n) \\ x_2(n) \\ \cdot \\ \cdot \\ x_k(n) \end{bmatrix}$
 ή
 $\begin{bmatrix} \dots x_1(n) & x_1(n+1) & x_1(n+2) & \dots \\ \dots x_2(n) & x_2(n+1) & x_2(n+2) & \dots \\ \dots & & & \\ \dots x_k(n) & x_k(n+1) & x_k(n+2) & \dots \end{bmatrix}$

ψηφιακή αναπαράσταση ήχου (III)

Τύπος	Εύρος αναλογικών συχνοτήτων	Συχνότητα δειγματοληψίας	bits/δείγμα	bitrate
φωνή χαμηλής ποιότητας (π.χ. τηλέφωνο)	200-3200 Hz	8 KHz	16	128 Kbits/sec
φωνή υψηλής ποιότητας	50-7000 Hz	16 KHz	16	256 Kbits/sec
ήχος CD	20-20000 Hz	44,1 KHz	2 κανάλια x 16	1,41 Mbits/sec

ΑΠΘ/ΤΗΜΜΥ/2003-4

17-Υπολογιστικά Συστήματα – Ψηφιακά Μέσα

Διδάσκων:

A. Ντελόπουλος

ψηφιακή αναπαράσταση εικόνας (I)

- εικόνες αποχρώσεων του γκρι (“ασπρόμαυρες”)
 - κάθε δείγμα έχει τιμή αντίστοιχη προς τη φωτεινότητα στην αντίστοιχη θέση
 - αναπαρίστανται από έναν διδιάστατο πίνακα βαθμωτών τιμών

$$Y = \begin{bmatrix} y(1,1) & y(1,2) & \cdots & y(1,N) \\ y(2,1) & y(2,2) & \cdots & y(2,N) \\ \vdots & \vdots & \vdots & \vdots \\ y(M,1) & y(M,2) & \cdots & y(M,N) \end{bmatrix}$$

ΑΠΘ/ΤΗΜΜΥ/2003-4

17-Υπολογιστικά Συστήματα – Ψηφιακά Μέσα

Διδάσκων:

A. Ντελόπουλος

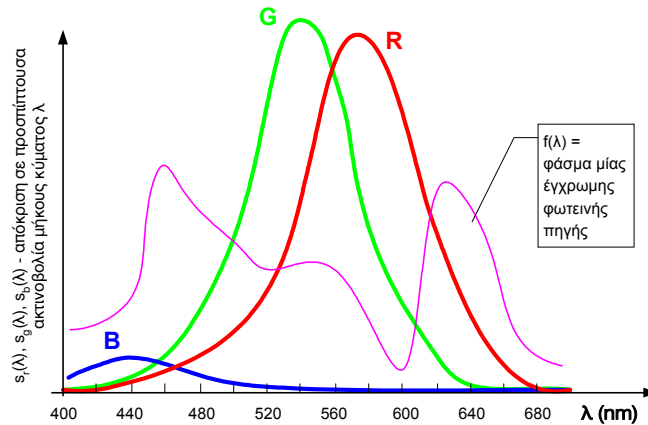
ψηφιακή αναπαράσταση εικόνας (II)

- έγχρωμες εικόνες
 - κάθε δείγμα είναι ένα διάνυσμα μήκους 3
 - οι συνιστώσες του διανύσματος αναπαριστούν την ένταση μιας χρωματικής συνιστώσας
 - στην απλούστερη περίπτωση οι τρεις χρωματικές συνιστώσες αντιστοιχούν στο κόκκινο (r), στο πράσινο (g) και στο μπλε (b)
 - άρα η αναπαράσταση αποτελείται από
 - τρεις πίνακες (R, G, B) με βαθμωτές τιμές ή ισοδύναμα
 - ένα πίνακα με διανυσματικά στοιχεία

ψηφιακή αναπαράσταση εικόνας (III) γιατί χρησιμοποιούμε 3 χρωματικές συνιστώσες;

- χρειάζεται να κατανοήσουμε τη φυσιολογία του ανθρώπινου ματιού
- τα κωνία του ματιού (φωτοευαίσθητα κύτταρα) χωρίζονται σε 3 οικογένειες
- κάθε μία από τις 3 είναι ευαίσθητη σε διαφορετική περιοχή του φωτεινού φάσματος
- για κάθε σημείο της εικόνας φτάνουν στον εγκέφαλο 3 τιμές από τα αντίστοιχα κωνία

ψηφιακή αναπαράσταση εικόνας (IV) η φασματική απόκριση των 3 οικογενειών κωνίων



ΑΠΘ/ΤΗΜΜΥ/2003-4

17-Υπολογιστικά Συστήματα – Ψηφιακά Μέσα

Διδάσκων:

A. Ντελόπουλος

ψηφιακή αναπαράσταση εικόνας (IV) η φασματική απόκριση των 3 οικογενειών κωνίων

- η πληροφορία που φτάνει στον εγκέφαλο από κάθε σημείο της εικόνας είναι:

$$\bar{c} = \begin{bmatrix} r \\ g \\ b \end{bmatrix} = \begin{bmatrix} \int s_r(\lambda) f(\lambda) d\lambda \\ \int s_g(\lambda) f(\lambda) d\lambda \\ \int s_b(\lambda) f(\lambda) d\lambda \end{bmatrix}$$

- τα $s_r(\lambda)$, $s_g(\lambda)$, $s_b(\lambda)$ και ένα παράδειγμα του $f(\lambda)$ φαίνονται στην προηγούμενη διαφάνεια
- στην ψηφιακή αναπαράσταση χρησιμοποιούμε ένα διάνυσμα $\bar{a} \propto \bar{c}$

ΑΠΘ/ΤΗΜΜΥ/2003-4

17-Υπολογιστικά Συστήματα – Ψηφιακά Μέσα

Διδάσκων:

A. Ντελόπουλος

ψηφιακή αναπαράσταση βίντεο

- βίντεο = ακολουθία (ακίνητων) εικόνων
- καρτέ ή πλαίσιο = κάθε μία από τις ακίνητες εικόνες
- συχνότητα χρονικής δειγματοληψίας:
 - περί τα 20 – 30 πλαίσια/sec (frames per second – fps)
 - σύστημα PAL (ευρωπαϊκή τηλεόραση) → 25 fps
 - σύστημα NTSC (αμερικανική τηλεόραση) → 30 fps
- διαστάσεις πλαισίου:
 - PAL → 576 γραμμές x 720 στήλες
 - NTSC → 486 γραμμές x 720 στήλες
 - High Definition TV → 720 γραμμές x 1280 στήλες
μέχρι → 1080 γραμμές x 1920 στήλες

Συμπύεση - Κωδικοποίηση

να μειώσουμε το πλήθος των bits που απαιτούνται για την αναπαράσταση των ψηφιακών μέσων

μία ώρα μουσικής ...

- 1 h x
- 3600 sec/h x
- 44100 δείγματα/sec x
- 16 bits/δείγμα x
- 2 κανάλια = 5.080.320.000 bits
≈ 4.7 Gigabits
≈ 600 Mbytes

δηλαδή όσο η χωρητικότητα 1 CD

ένα δευτερόλεπτο φωνής

- 1 sec x
- 8000 δείγματα/sec x
- 16 bits/δείγμα x
- 1 κανάλι = 128000 bits ≈ 128 kbits
- → άρα απαιτούνται 128 kbps ενώ οι τηλεφωνικές συνδέσεις παρέχουν από 9,6 kbps έως το πολύ 56 kbps.

μία σελίδα A4 στο fax ...

- υπόθεση:
 - οριζόντια συχνότητα δειγματοληψίας $f_h = 150$ dots per inch (dpi)
 - κατακόρυφη συχνότητα δειγματοληψίας $f_v = 150$ lines per inch (lpi)
 - διαστάσεις A4: 8,27 inches x 11,69 inches
 - ρυθμός μετάδοσης (bitrate) = 9600 bps (bits per second)
- → $8,27 \times 11,69 \times 150 \times 150$ δείγματα x 1 bit/δείγμα = 2.175.216 bits \approx 2 Mbits \approx 0,25 Mbytes
- → \approx **227 sec** για μετάδοση πάνω από σύνδεση 9600 bps
- → \approx 16 σελίδες την ώρα !!!

ένα δευτερόλεπτο ψηφιακού βίντεο PAL

- 1 sec x
- 25 πλαίσια / sec x
- 576 x 720 δείγματα / πλαίσιο x
- 3 χρωματικές συνιστώσες / δείγμα x
- 8 bits / χρωματική συνιστώσα
= 248.832.000 bits \approx **237 Mbits** \approx 30 Mbytes
- Για σύγκριση η χωρητικότητα (bandwidth):
 - του πανεπιστημιακού δικτύου δεδομένων: 100 Mbits/sec
 - των γραμμών ATM: \sim 155 Mbits/sec
 - των δορυφορικών καναλιών: \sim 32 Mbits/sec
 - των συνδέσεων DSL: 1,5 – 8 Mbits/sec
- Ένα DVD των 4 Gbytes θα χωρούσε \sim 133 sec βίντεο!

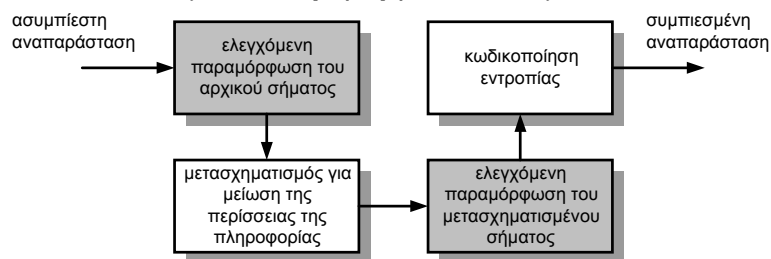
άρα χρειάζεται συμπίεση της αναπαράστασης

ΑΠΘ/ΤΗΜΜΥ/2003-4

17-Υπολογιστικά Συστήματα – Ψηφιακά Μέσα

Διδάσκων:
Α. Ντελόπουλος

συμπίεση ψηφιακών μέσων



- συμπίεση με απώλειες: από τη συμπιεσμένη αναπαράσταση μπορεί να ανακατασκευαστεί ένα παραμορφωμένο αντίγραφο της αρχικής (ασυμπίεστης)
- συμπίεση χωρίς απώλειες: είναι εφικτή η τέλεια ανακατασκευή
 - λείπουν τα γκρίζα blocks του διαγράμματος
 - λιγότερη συμπίεση

ΑΠΘ/ΤΗΜΜΥ/2003-4

17-Υπολογιστικά Συστήματα – Ψηφιακά Μέσα

Διδάσκων:
Α. Ντελόπουλος

ελεγχόμενη παραμόρφωση του αρχικού σήματος

- υποδειματοληψία
 - μείωση διάστασης της εικόνας
 - μετάδοση άλλοτε των άρτιων και άλλοτε των περιττών γραμμών στο βίντεο
 - κ.λπ.
- απάλειψη υψηλών συχνοτήτων
- μείωση του αριθμού των κβαντικών επιπέδων
- κ.λπ.

μετασχηματισμός για μείωση της περίσσειας της πληροφορίας

- περίσσεια πληροφορίας: αποθηκεύουμε / μεταδίδουμε την ίδια πληροφορία πολλές φορές
- συνήθως γειτονικά δείγματα είναι συσχετισμένα
 - η τιμή ενός δείγματος μπορεί να προβλεφθεί από τα γειτονικά του
- μείωση περίσσειας = αποσυσχέτιση
- τεχνικές:
 - πρόβλεψη και αφαίρεση : π.χ. Differential Pulse Code Modulation (DPCM)
 - μετασχηματισμός: π.χ., Διακριτού Συνημιτόνου (DCT), Karhunen Loeve (KLT)
 - εκτίμηση και αντιστάθμιση κίνησης

ελεγχόμενη παραμόρφωση του μετασχηματισμένου σήματος

- κβαντισμός του μετασχηματισμένου σήματος με ελεγχόμενη ακρίβεια
 - οδηγεί σε απώλεια πληροφορίας

κωδικοποίηση εντροπίας (I)

- εντροπία (H):
 - έστω ένας μηχανισμός παραγωγής n διαφορετικών συμβόλων $\{s_0, s_1, \dots, s_{n-1}\}$
 - όπως ο κβαντιστής του προηγούμενου σταδίου
 - ο μηχανισμός αυτός χαρακτηρίζεται ως πηγή συμβόλων
 - αν p_i η πιθανότητα εμφάνισης του συμβόλου s_i

$$H = -\sum_{i=0}^{n-1} p_i \log_2(p_i)$$

κωδικοποίηση εντροπίας (II)

- Έστω ότι επιλέγουμε τις κωδικές δυαδικές λέξεις $\{w_0, w_1, \dots, w_{n-1}\}$ για να αναπαραστήσουμε τα σύμβολα της πηγής
 - το μήκος, $|w_i|$, της κάθε λέξης μπορεί να είναι διαφορετικό
 - π.χ. $s_0 \rightarrow w_0 = 1101 \rightarrow |w_0| = 4$ bits
 $\rightarrow w_1 = 101 \rightarrow |w_1| = 3$ bits, κ.λπ.

- το μέσο μήκος λέξης είναι
$$L = \sum_{i=0}^{n-1} p_i |w_i|$$

κωδικοποίηση εντροπίας (III)

- Θεώρημα: η εντροπία είναι το κάτω όριο του μέσου μήκους λέξης = δεν μπορούμε να πετύχουμε καλύτερη συμπίεση απότι μας δηλώνει η εντροπία

$$L \geq H$$

- κωδικοποίηση εντροπίας: επιλέγουμε τις κωδικές λέξεις $\{w_0, w_1, \dots, w_{n-1}\}$ έτσι ώστε το μέσο μήκος λέξης L να προσεγγίζει όσο γίνεται περισσότερο την εντροπία H .

πρότυπα συμπίεσης

- ήχος:
 - Differential Pulse Code Modulation (DPCM)
 - MPEG layer III (mp3)
- εικόνα:
 - πρότυπα του Joint Pictures Expert Group (JPEG)
 - με απώλειες → συμπίεση 10 προς 1
 - χωρίς απώλειες → συμπίεση 2-3 προς 1
- βίντεο:
 - πρότυπα της επιτροπής Motion Pictures Expert Group
 - MPEG 1 → ~1.5 Mbits/sec → εφαρμογές πολυμέσων/internet
 - MPEG 2 → 2 έως 60 Mbits/sec → ψηφιακή TV, DVD, ψηφιακός κινηματογράφος
 - H.261 και H.263 → 64, 128, p x 64 Kbits/sec → videoconference