

Εισαγωγή στην επιστήμη των υπολογιστών

Υπολογιστές και Δεδομένα

Κεφάλαιο 2ο

Αναπαράσταση Δεδομένων

Τύποι Δεδομένων

- ❖ Τα δεδομένα σήμερα συναντώνται σε διάφορες μορφές, στις οποίες περιλαμβάνονται αριθμοί, κείμενο, εικόνες, ήχος, και βίντεο.



Τύποι Δεδομένων

- ◆ Ένα πρόγραμμα μηχανικής χρησιμοποιεί τον υπολογιστή κυρίως για επεξεργασία αριθμών
- ◆ Ένα πρόγραμμα επεξεργασίας κειμένου χρησιμοποιεί τον υπολογιστή κυρίως για εργασίες με κείμενο
- ◆ Ένα πρόγραμμα επεξεργασίας εικόνων χρησιμοποιεί τον υπολογιστή για χειρισμό εικόνων
- ◆ Ένας υπολογιστής χειρίζεται επίσης δεδομένα ήχου
- ◆ Οι υπολογιστές χρησιμοποιούνται για τη δημιουργία των ειδικών εφέ που βλέπουμε στις ταινίες



Σημείωση:

Στη πληροφορική χρησιμοποιείται ο όρος “multimedia” για ορισμό πληροφοριών που περιέχουν αριθμούς, κείμενο, εικόνες, ήχο, και video.

Τα Δεδομένα στο Εσωτερικό του Υπολογιστή

- ◆ Όλοι οι τύποι δεδομένων του πραγματικού κόσμου μετατρέπονται σε μία ενιαία αναπαράσταση, όταν αποθηκεύονται στον υπολογιστή, και ξαναπαίρνουν την αρχική τους μορφή όταν βγαίνουν από τον υπολογιστή.
- ◆ Αυτή η καθολική μορφή ονομάζεται σχήμα μπιτ (**bit pattern**).

Μπιτ

- ❖ **Μπιτ** (**bit**, από τη φράση **binary digit**, που σημαίνει **δυναδικό ψηφίο**) είναι η μικρότερη μονάδα δεδομένων που μπορεί να αποθηκευτεί σε έναν υπολογιστή και μπορεί να πάρει είτε την τιμή 0 είτε την τιμή 1
- ❖ Ένα μπιτ αντιπροσωπεύει την κατάσταση μιας συσκευής η οποία μπορεί να πάρει μία από δύο δυνατές τιμές.

Μπιτ

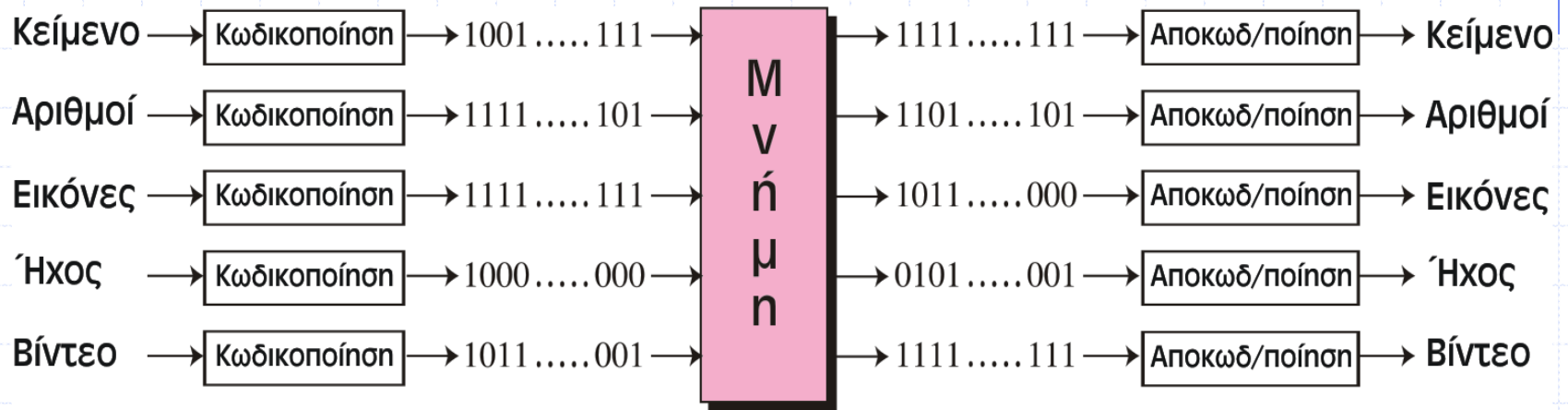
- ◆ Ένας ηλεκτρονικός διακόπτης αντιπροσωπεύει ένα μπιτ. Με άλλα λόγια, σε ένα διακόπτη μπορεί να αποθηκευτεί ένα μπιτ πληροφορίας
- ◆ Ένας διακόπτης μπορεί να είναι είτε αναμμένος κλειστός (on), είτε σβηστός ανοικτός (off). Η σύμβαση είναι να αναπαριστάται η κατάσταση "on" με 1 και η κατάσταση "off" με 0

Σχήματα Μπιτ

- ◆ Για την αναπαράσταση των διαφόρων τύπων δεδομένων χρησιμοποιούμε σχήματα μπιτ (bit patterns), δηλαδή ακολουθίες ή όπως μερικές φορές λέγονται συμβολοσειρές μπιτ

1 0 0 0 1 0 1 0 1 1 1 1 1

Μνήμη και σχήματα Μπιτ



❖ Τα δεδομένα κωδικοποιούνται όταν εισέρχονται στον υπολογιστή και αποκωδικοποιούνται όταν παρουσιάζονται στο χρήστη

Μπάιτ

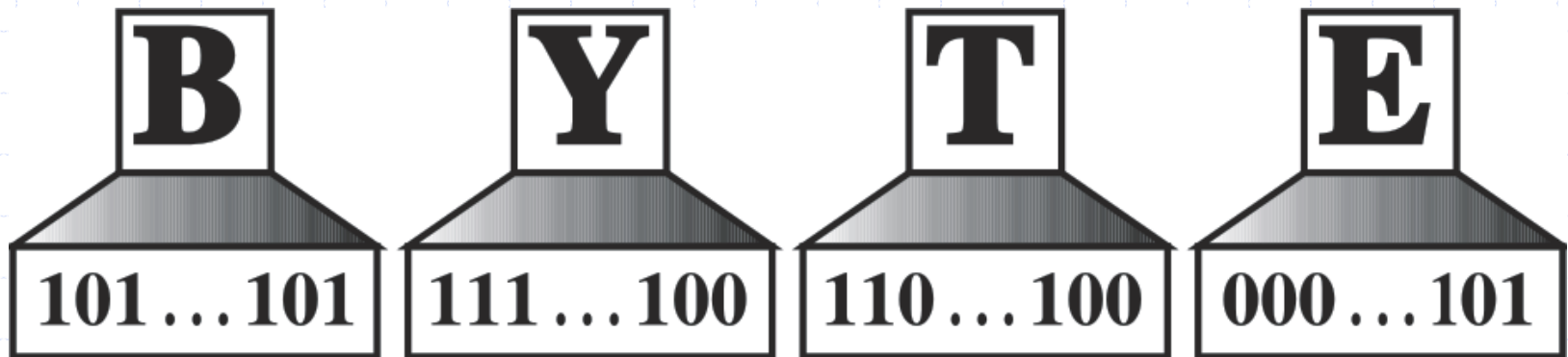
- ◆ Ένα σχήμα μπιτ με μήκος 8 μπιτ ονομάζεται **μπάιτ (byte)**.
- ◆ Ο όρος αυτός χρησιμοποιείται επίσης για τη μέτρηση του μεγέθους της μνήμης ή άλλων συσκευών αποθήκευσης.

Αναπαράσταση Δεδομένων

Κείμενο, Αριθμοί, Εικόνες, Ήχος,
Βίντεο

Κείμενο

- ◆ Σε οποιαδήποτε γλώσσα, ένα τμήμα κειμένου (text) δεν είναι παρά μια ακολουθία συμβόλων που χρησιμοποιούνται για να αναπαραστήσουν μια έννοια στη συγκεκριμένη γλώσσα
- ◆ Κάθε σύμβολο μπορεί να αναπαρασταθεί από ένα σχήμα μπιτ



Κείμενο

❖ Ο αριθμός των μπιτ που χρειάζεται ένα σχήμα μπιτ για να αναπαραστήσει ένα σύμβολο σε κάποια γλώσσα εξαρτάται από το πλήθος των συμβόλων που διαθέτει η γλώσσα.

❖ **Η σχέση αυτή είναι λογαριθμική**

$$(\log_2 \text{Αριθμός_Συμβόλων} = \text{Μήκος_Σχήματος_Μπιτ})$$

Κείμενο

- ◆ Ένα σχήμα μπιτ με μήκος 2 μπιτ μπορεί να πάρει τέσσερις διαφορετικές μορφές ($\log_2 4 = 2$): 00, 01, 10, και 11. Κάθε μία από αυτές τις μορφές αναπαριστά ένα σύμβολο. Με τον ίδιο τρόπο, ένα σχήμα μπιτ με μήκος 3 μπιτ μπορεί να πάρει οκτώ διαφορετικές μορφές: 000, 001, 010, 011, 100, 101, 110, και 111
- ◆ Για την αναπαράσταση των συμβόλων κειμένου έχουν σχεδιαστεί διάφορα σύνολα σχημάτων μπιτ. Κάθε σύνολο ονομάζεται **κώδικας**, και η διαδικασία της αναπαράστασης συμβόλων ονομάζεται **κωδικοποίηση**

Κείμενο

Πλήθος συμβόλων	Μήκος σχήματος μππ
2	1
4	2
8	3
16	4
...	...
128	7
256	8
...	...
65.536	16

Κείμενο

◆ ASCII

◆ Το Αμερικανικό Ινστιτούτο Εθνικών Προτύπων (American National Standard Institute, ANSI) ανέπτυξε έναν κώδικα τον οποίο ονόμασε Αμερικανικό Πρότυπο Κώδικα για την Ανταλλαγή Πληροφοριών (American Standard Code for Information Interchange, ASCII).

◆ Ο κώδικας αυτός χρησιμοποιεί 7 μπιτ για κάθε σύμβολο, δηλαδή με τον κώδικα ASCII μπορούν να οριστούν 128 ($= 2^7$) διαφορετικά σύμβολα

B	Y	T	E
1000010	1011001	1010100	1000101

01001000

H

01100101

e

01101100

l

01101100

l

01101111

o

00101110

.

Κείμενο

- ◆ Ο κώδικας ASCII χρησιμοποιεί ένα σχήμα 7 μπιτ με εύρος από 0000000 μέχρι 1111111.
- ◆ Το πρώτο σχήμα (0000000) αναπαριστά τον κενό χαρακτήρα (την απουσία χαρακτήρα).
- ◆ Το τελευταίο σχήμα (1111111) αναπαριστά το χαρακτήρα διαγραφής.
- ◆ Υπάρχουν 31 χαρακτήρες ελέγχου (μη εκτυπώσιμοι).
- ◆ Οι αριθμητικοί χαρακτήρες (0 έως 9) είναι τοποθετημένοι πριν από τα γράμματα.
- ◆ Υπάρχουν πολλοί ειδικοί εκτυπώσιμοι χαρακτήρες.
- ◆ Τα κεφαλαία γράμματα (A...Z, στην αγγλική γλώσσα) είναι πριν από τα πεζά γράμματα (a...z).
- ◆ Οι κεφαλαίοι και οι πεζοί χαρακτήρες ξεχωρίζουν μεταξύ τους από ένα μόνο μπιτ. Για παράδειγμα, το σχήμα του A είναι το 1000001 ενώ το σχήμα του a είναι το 1100001. Η μόνη διαφορά βρίσκεται στο έκτο μπιτ από τα δεξιά.
- ◆ Μεταξύ των κεφαλαίων και των πεζών γραμμάτων υπάρχουν έξι ειδικοί χαρακτήρες.

Κείμενο

- ◆ EBCDIC
- ◆ Στις αρχές της εποχής των υπολογιστών, η IBM ανέπτυξε έναν κώδικα τον οποίο ονόμασε Επεκτεταμένο Κώδικα Ανταλλαγής Δυαδικά Κωδικοποιημένων Δεκαδικών (Extended Binary Coded Decimal Interchange Code, EBCDIC).
- ◆ Ο κώδικας αυτός χρησιμοποιούσε σχήματα 8 μπιτ, δηλαδή μπορούσε να αναπαραστήσει μέχρι 256 σύμβολα. Χρησιμοποιείται αποκλειστικά σε μεγάλα συστήματα (mainframes) της IBM.

Κείμενο

◆ Unicode

- ◆ Ένας συνασπισμός από κατασκευαστές υλικού και λογισμικού σχεδίασε έναν κώδικα με όνομα **Unicode**, ο οποίος χρησιμοποιεί σχήματα 16 μπιτ και μπορεί να αναπαραστήσει μέχρι 65.536 (2^{16}) σύμβολα.
- ◆ Ο κώδικας αυτός είναι χωρισμένος σε διάφορες ενότητες, κάθε μία από τις οποίες αντιστοιχεί στα σύμβολα μιας διαφορετικής γλώσσας.
- ◆ Έτσι αναπαριστά σύμβολα που ανήκουν σε άλλες γλώσσες εκτός των Αγγλικών

Κείμενο

◆ ISO

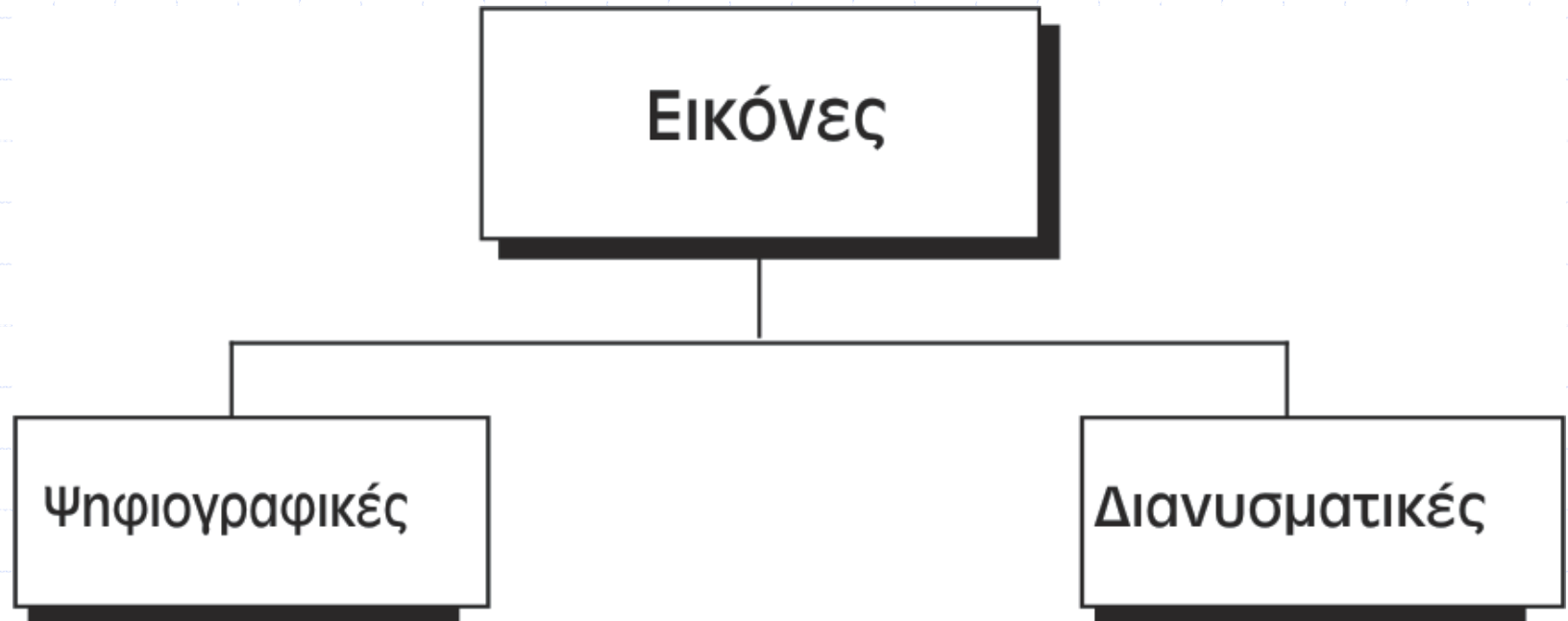
◆ Ο Διεθνής Οργανισμός Προτυποποίησης (International Organization for Standardization), γνωστός ως ISO, έχει σχεδιάσει έναν κώδικα που χρησιμοποιεί σχήματα 32 μπιτ.

◆ Ο κώδικας αυτός μπορεί να αναπαραστήσει μέχρι 4.294.967.296 (2^{32}) σύμβολα, αριθμός που είναι σαφώς επαρκής για την αναπαράσταση οποιουδήποτε γνωστού συμβόλου στον κόσμο σήμερα.

Αριθμοί

- ◆ Σε έναν υπολογιστή οι αριθμοί αναπαρίστανται με τη χρήση του **δυναδικού συστήματος** (binary system).
- ◆ Στο σύστημα αυτό κάθε αριθμός αναπαρίσταται με ένα σχήμα μπιτ (μια σειρά από μηδενικά και άσσους), χωρίς όμως να χρησιμοποιείται κάποιος κώδικας όπως ο ASCII

ΕΙΚΟΝΕΣ

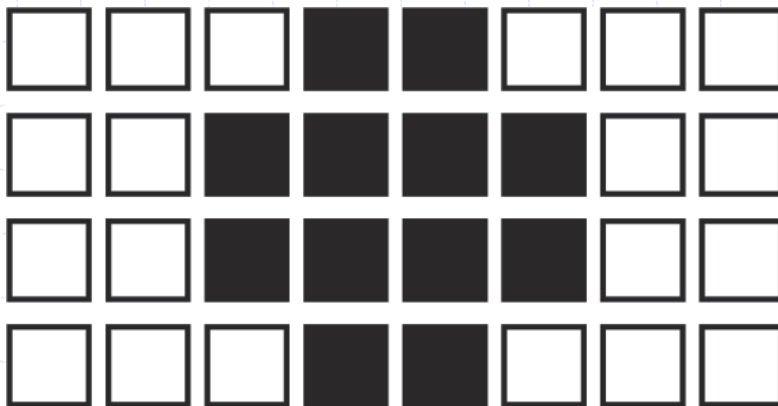


Εικόνες

- ❖ Οι ψηφιογραφικές εικόνες διαιρούνται σε μια διάταξη **πίξελ** (pixel, από τη φράση picture element, δηλαδή που σημαίνει **στοιχείο εικόνας** ή **εικονοστοιχείο**), δηλαδή σε μικρές κουκκίδες.
- ❖ Το μέγεθος ενός πίξελ εξαρτάται από τον παράγοντα που ονομάζουμε *ανάλυση* (resolution).

ΕΙΚΟΝΕΣ

- Ψηφιογραφική μέθοδος αναπαράστασης εικόνων σε μια ασπρόμαυρη εικόνα



Εικόνα

0 0 0 1 1 0 0 0

0 0 1 1 1 1 0 0

0 0 1 1 1 1 0 0

0 0 0 1 1 0 0 0

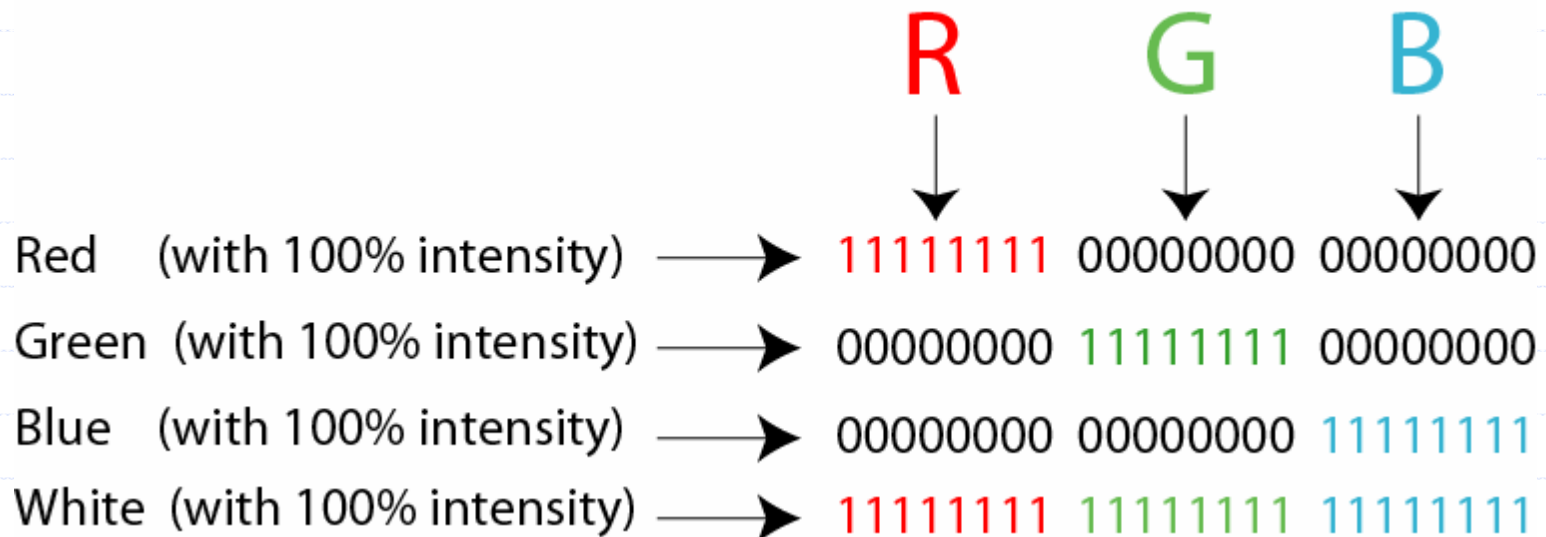
Αναπαράσταση πίνακα

00011000 00111100 00111100 00011000

Γραμμική αναπαράσταση

ΕΙΚΟΝΕΣ

- ◆ Αναπαράσταση έγχρωμων εικόνων
- ◆ Κάθε πίξελ διαθέτει τρία σχήματα μπιτ: ένα για την αναπαράσταση της έντασης του κόκκινου χρώματος, ένα του πράσινου και ένα του μπλε



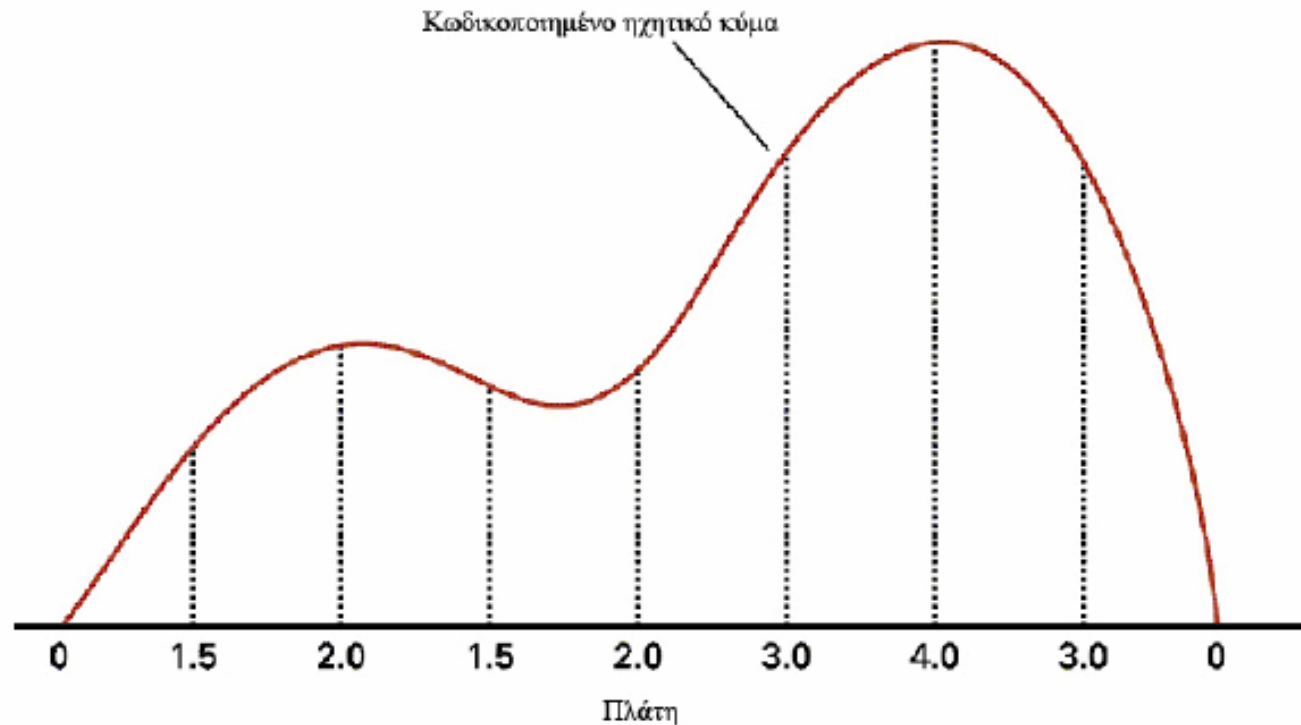
Εικόνες

- ◆ Κατά τη μέθοδο των διανυσματικών γραφικών δεν αποθηκεύονται τα σχήματα μπιτ αλλά η εικόνα αναλύεται σε ένα συνδυασμό από καμπύλες και ευθύγραμμα τμήματα που αναπαρίστανται από ένα μαθηματικό τύπο.
- ◆ Έτσι μπορεί να αλλάζει το μέγεθος της εικόνας χωρίς καμία χρωματική απώλεια αφού η εικόνα κάθε φορά επανασχεδιάζεται.

Ήχος (Audio)

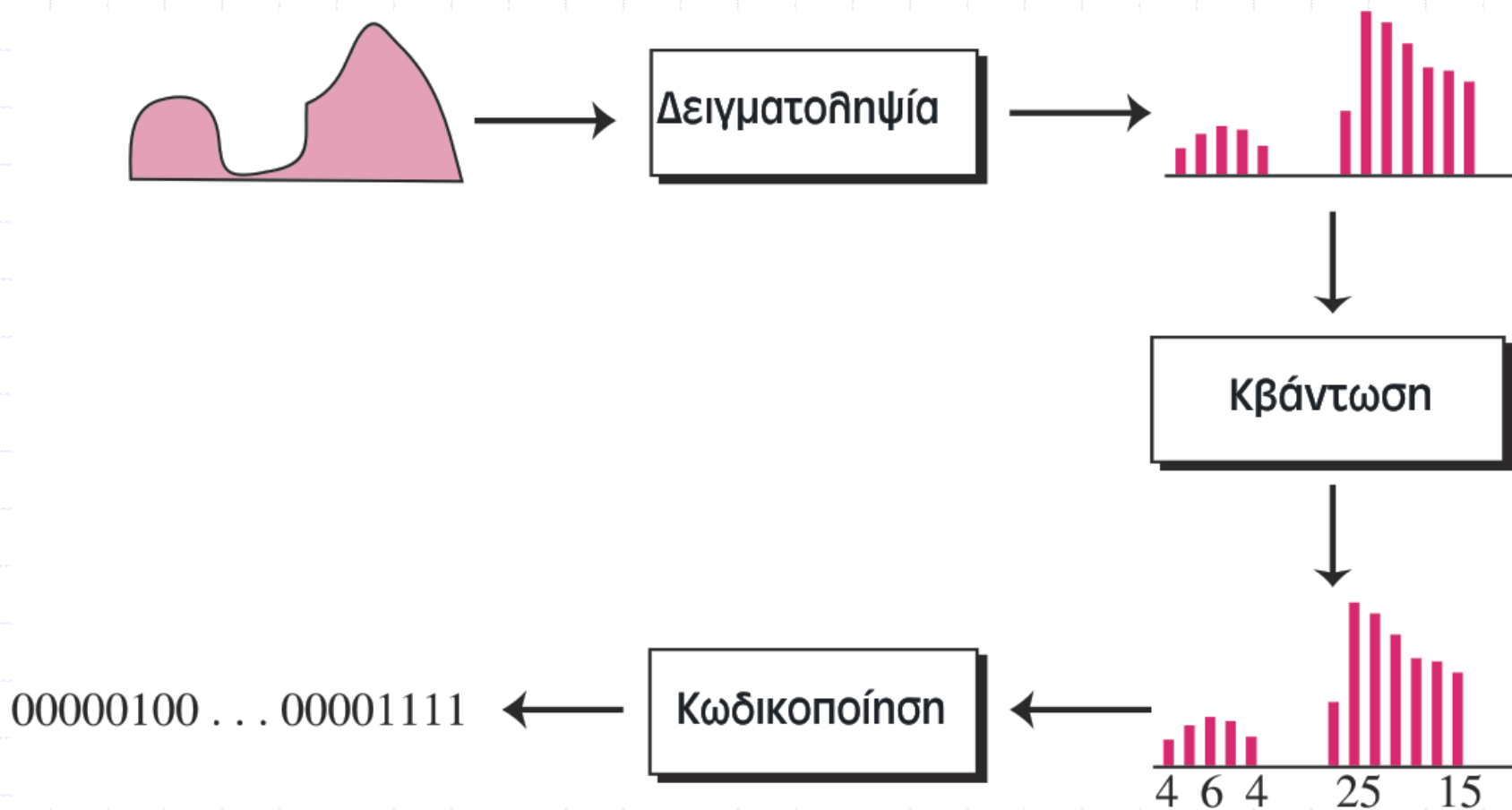
- ❖ Ο ήχος είναι εκ φύσεως **αναλογική** πληροφορία. Είναι συνεχής (αναλογικός), όχι διακριτός (ψηφιακός)
- ❖ Η βασική ιδέα είναι η μετατροπή του ήχου σε **ψηφιακά** δεδομένα τα οποία αποθηκεύονται με τη μορφή σχημάτων μπιτ

Ήχος



Για καλή ποιότητα αναπαραγωγής του ήχου στα CD απαιτούνται 44100 δείγματα το δευτερόλεπτο.

Ήχος (Audio)



Βίντεο

- ❖ Ο όρος **βίντεο** αναφέρεται στην αναπαράσταση εικόνων (ονομάζονται καρτέ) με το πέρασμα του χρόνου
- ❖ Μια ταινία είναι μια ακολουθία καρτέ τα οποία προβάλλονται το ένα μετά το άλλο ώστε να δημιουργήσουν την εντύπωση της κίνησης.
- ❖ Έτσι, αν γνωρίζουμε πώς να αποθηκεύσουμε μια εικόνα στον υπολογιστή, γνωρίζουμε επίσης και πώς να αποθηκεύσουμε βίντεο

Δεκαεξαδικός Συμβολισμός

- ❖ Ο δεκαεξαδικός συμβολισμός βασίζεται στον αριθμό 16.
- ❖ Αυτό σημαίνει ότι υπάρχουν 16 σύμβολα (δεκαεξαδικά ψηφία): 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E, και F
- ❖ Ένα σχήμα τεσσάρων μπιτ μπορεί να αναπαρασταθεί από ένα δεκαεξαδικό ψηφίο, και το αντίστροφο

Δεκαεξαδικός Συμβολισμός

Σχήμα Μπιτ	Δεκαεξαδικό ψηφίο	Σχήμα Μπιτ	Δεκαεξαδικό ψηφίο
0000	0	1000	8
0001	1	1001	9
0010	2	1010	A
0011	3	1011	B
0100	4	1100	C
0101	5	1101	D
0110	6	1110	E
0111	7	1111	F

Δεκαεξαδικός Συμβολισμός

- ❖ Ο δεκαεξαδικός συμβολισμός χρησιμοποιεί δύο μορφές.
 - Στην πρώτη μορφή προστίθεται πριν από τα ψηφία ένα πεζό (ή κεφαλαίο) x, ως ένδειξη ότι η αναπαράσταση είναι σε δεκαεξαδική μορφή (xA34)
 - Στην άλλη μορφή, η βάση του αριθμού (16) υποδεικνύεται υποδηλώνεται με τη μορφή δείκτη μετά τον αριθμό(A34₁₆)



Δεκαεξαδικό

Δεκαεξαδικός Συμβολισμός

◆ Βρείτε το δεκαεξαδικό ισοδύναμο του σχήματος μπιτ 11001100010

◆ Λύση

Κάθε ομάδα των τεσσάρων μπιτ μεταφράζεται σε ένα δεκαεξαδικό ψηφίο

- 1100 C
- 1110 E
- 0010 2

Άρα το δεκαεξαδικό ισοδύναμο είναι xCE2

Δεκαεξαδικός Συμβολισμός

◆ Βρείτε το δεκαεξαδικό ισοδύναμο του σχήματος μπιτ 000011100010

◆ Λύση

Κάθε ομάδα των τεσσάρων μπιτ μεταφράζεται σε ένα δεκαεξαδικό ψηφίο

- 0000 0
- 1110 E
- 0010 2

Άρα το δεκαεξαδικό ισοδύναμο είναι x0E2

Δεκαεξαδικός Συμβολισμός

◆ Ποιο είναι το σχήμα μπιτ του x24C;

◆ Λύση

Γράφουμε στη θέση του κάθε δεκαεξαδικού ψηφίου το ισοδύναμο σχήμα μπιτ

- 2 0010
- 4 0100
- C 1100

Άρα το ισοδύναμο σχήμα μπιτ είναι
001001001100

Οκταδικός Συμβολισμός

- ❖ Ο **οκταδικός συμβολισμός** βασίζεται στον αριθμό 8.
- ❖ Αυτό σημαίνει ότι υπάρχουν οκτώ σύμβολα (οκταδικά ψηφία): 0, 1, 2, 3, 4, 5, 6, και 7
- ❖ Κάθε οκταδικό ψηφίο μπορεί να αναπαραστήσει 3 μπιτ, και 3 μπιτ μπορούν να αναπαρασταθούν από ένα οκταδικό ψηφίο

Οκταδικός Συμβολισμός

Σχήμα Μπιτ	Οκταδικό ψηφίο	Σχήμα Μπιτ	Οκταδικό ψηφίο
000	0	100	4
001	1	101	5
010	2	110	6
011	3	111	7

Οκταδικός Συμβολισμός

- ❖ Ο οκταδικός συμβολισμός χρησιμοποιεί δύο μορφές.
 - Στην πρώτη μορφή προστίθεται πριν από τα ψηφία το 0 (μηδέν) ή το πεζό ο, ως ένδειξη ότι η αναπαράσταση είναι οκταδική (ο634).
 - Με την άλλη μορφή υποδεικνύεται υποδηλώνεται η βάση του αριθμού (8) με τη μορφή δείκτη (634_8)

1	1 1 1	1 1 0	0 1 1	1 0 0	1 0 0
1	7	6	3	4	4

Οκταδικό

Οκταδικός Συμβολισμός

◆ Βρείτε το οκταδικό ισοδύναμο του σχήματος
101110010

◆ Λύση

Κάθε ομάδα των τριών μπιτ μεταφράζεται σε
ένα οκταδικό ψηφίο

- 101 5
- 110 6
- 010 2

Άρα το οκταδικό ισοδύναμο είναι 0562 ή 0562 ή
 562_8

Οκταδικός Συμβολισμός

◆ Βρείτε το οκταδικό ισοδύναμο του σχήματος 1100010

◆ Λύση

Κάθε ομάδα των τριών μπιτ μεταφράζεται σε ένα οκταδικό ψηφίο

Πρέπει όμως να προσθέσουμε δύο επιπλέον μηδενικά στα αριστερά ώστε ο συνολικός αριθμός των μπιτ να διαιρείται ακριβώς με το 3. Έτσι παίρνουμε το σχήμα 001100010

■ 001	1
■ 100	4
■ 010	2

Άρα το οκταδικό ισοδύναμο είναι 0142 ή 0142 ή 142_8

Οκταδικός Συμβολισμός

◆ Ποιο είναι το σχήμα μπιτ του 24_8 ;

◆ Λύση

Γράφουμε στη θέση του κάθε οκταδικού ψηφίου το ισοδύναμο σχήμα μπιτ

- 2 010
- 4 100

Άρα το ισοδύναμο σχήμα μπιτ είναι
010100