

ψηφιακά Μέσα

10110011011000
10010011100011
1111000100110

06

**Ψηφιακός
Ήχος**

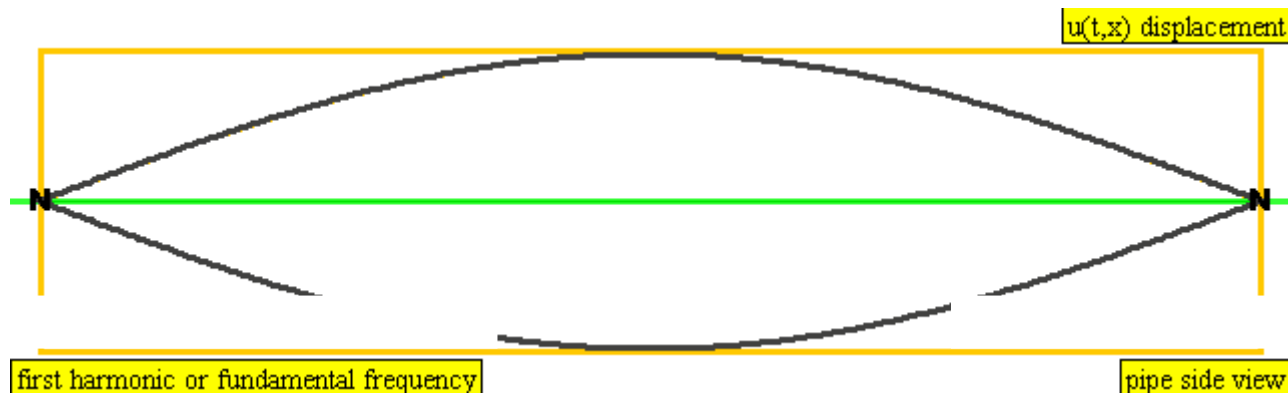
1ο Μέρος

Ψηφιακός Ήχος

Βασικές Έννοιες

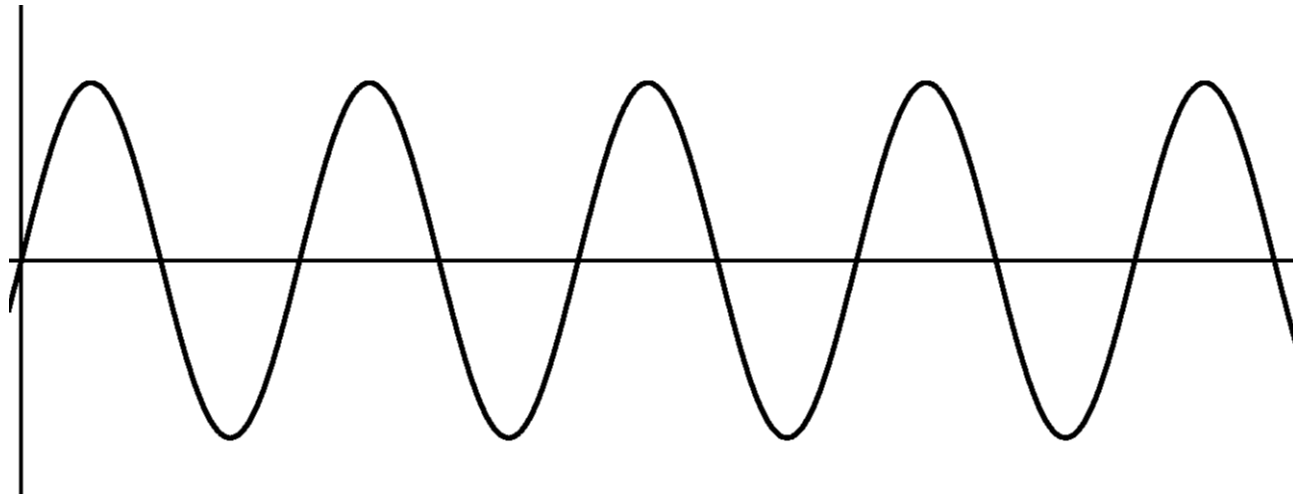
Ήχος

- παραγωγή ήχου: ταλάντωση της ύλης δημιουργεί μεταβολές στην πίεση του αέρα



Ήχος

- διάδοση: οι μεταβολές στην πίεση μεταδίδονται στον αέρα ως κύμα

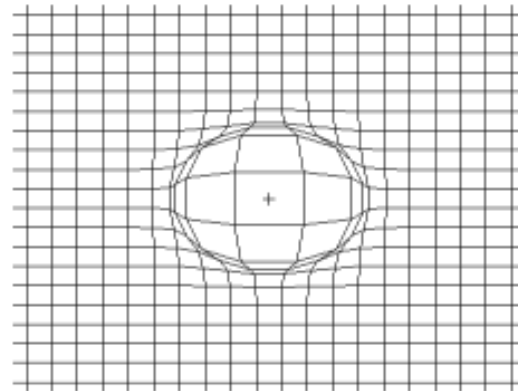
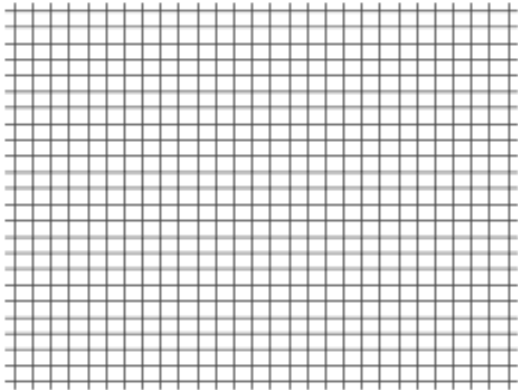


Ήχος

- απαιτείται μέσο μετάδοσης
- η ταχύτητα ως συνάρτηση του μέσου:
 - ταχύτερη μετάδοση σε μέσα με μεγαλύτερη πυκνότητα

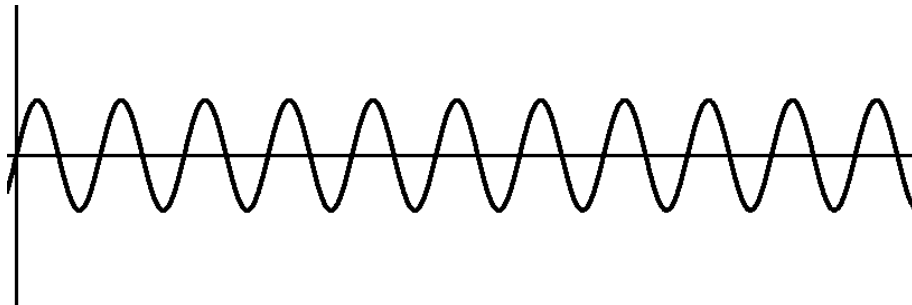
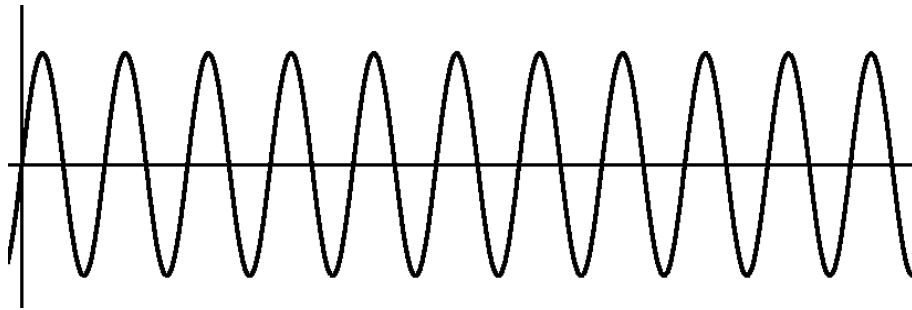
Ήχος

- χαρακτηριστικά
 - 1Δ κύμα
 - διαμήκες κύμα



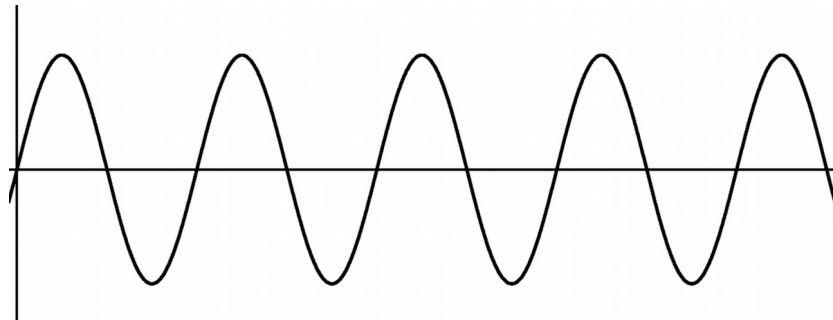
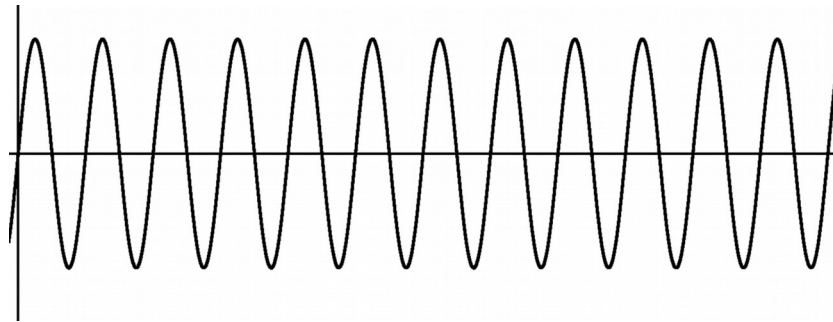
Ήχος

- χαρακτηριστικά
 - πλάτος



Ήχος

- χαρακτηριστικά
 - συχνότητα: ταλαντώσεις/δευτερόλεπτο (Hz)



Ήχος

- χαρακτηριστικά
 - ένταση
 - ανάλογη της ταλάντωσης
 - μικρή αλλαγή πίεσης: μικρή ένταση
 - μεγάλη αλλαγή πίεσης: μεγάλη ένταση

Ακοή

- εύρος ανθρώπινης ακουστικής αντίληψης
 - 20 Hz -20 KHz
 - ακουστικά σήματα
- ωστόσο
 - δεν υπάρχει η ίδια ευαισθησία σε όλο το εύρος
 - μεγαλύτερη ευαισθησία μεταξύ 2 KHz - 5 KHz

Ακοή

- Ακουστική αντίληψη
 - μεγάλο δυναμικό εύρος
 - οι ήχοι γίνονται αντιληπτοί σε **σχέση** με άλλους ήχους και όχι ως απόλυτες τιμές
 - π.χ.
 - κλάμα μωρού σε βιβλιοθήκη
 - κλάμα μωρού σε εργοστάσιο

Ακοή

- η αντίληψη της έντασης είναι υποκειμενική
- η ανθρώπινη ακοή είναι λογαριθμική
 - σύγκριση έντασης ήχων: λόγος αντί η διαφορά τους σε απόλυτες τιμές
 - μονάδα μέτρησης: decibel (db)
 - σχετική κλίμακα - διαφορά: $10 \log P/P_0$
 - P = ένταση ήχου που καταγράφηκε
 - P_0 = ένταση ήχου αναφοράς

Ακοή

- δυναμικό εύρος ακοής: 0 db – 140 db
 - 0 db: όριο ακοής (threshold of hearing) 20 μ Pa
 - 140 db: όριο πόνου
- δυναμική προσαρμογή αυτιού
 - αναλογία με μάτι
 - προσαρμοζόμενη κλίμακα $\sim$ 90 db

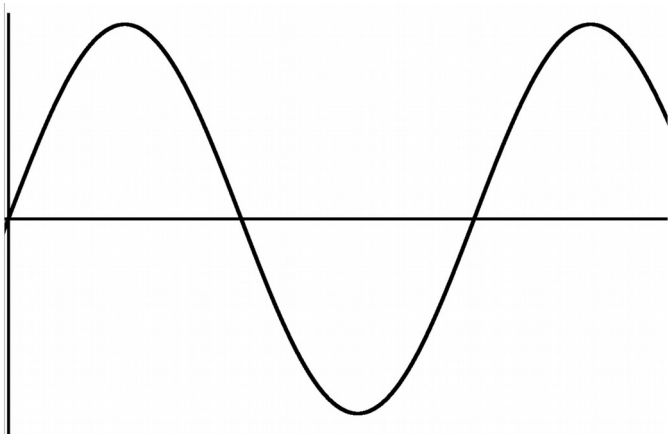
Ψηφιοποίηση Ήχου

Ψηφιοποίηση

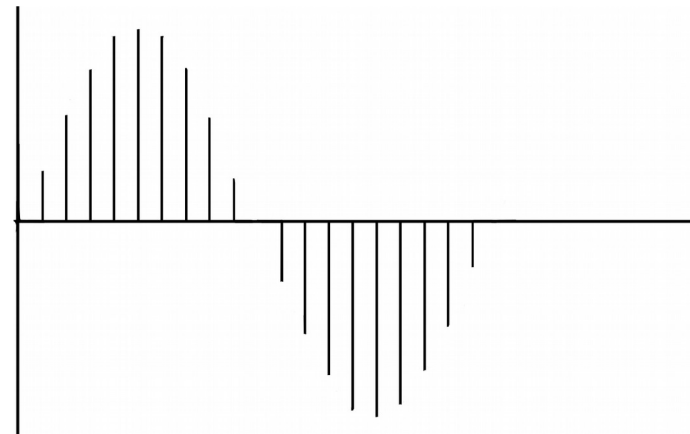
- Στάδια ψηφιοποίησης
 - δειγματοληψία
 - κβαντοποίηση
 - κωδικοποίηση

Ψηφιοποίηση

- Δειγματοληψία



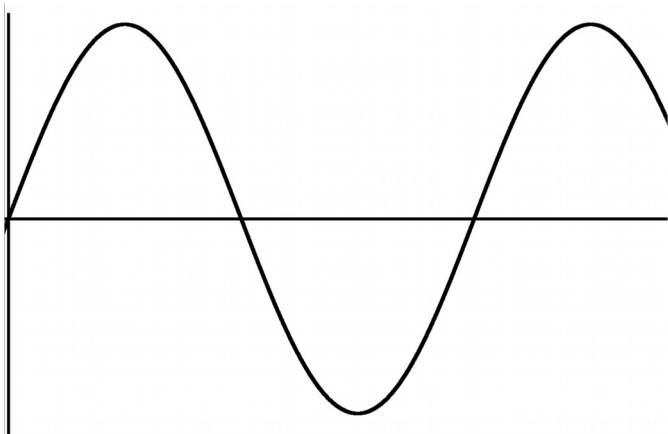
αναλογικό σήμα



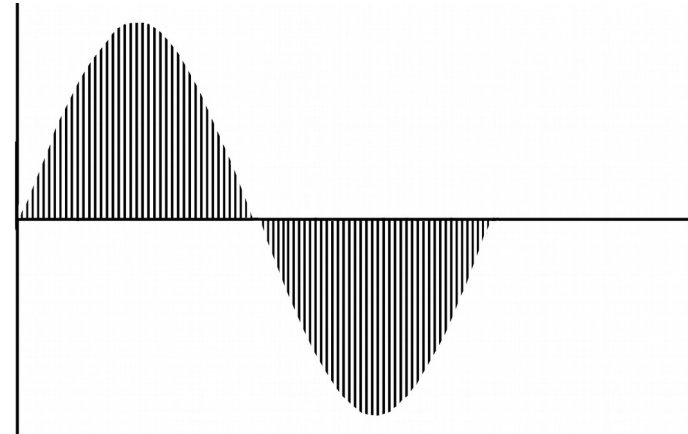
Ψηφιακό σήμα

Ψηφιοποίηση

- Δειγματοληψία



αναλογικό σήμα



Ψηφιακό σήμα

Ψηφιοποίηση

- **Δειγματοληψία**

- δείγμα Fur Elise – L. Van Beethoven

- 44100 Hz 

- 22050 Hz 

- 11025 Hz 

- 8000 Hz 

- 6000 Hz 

- 4000 Hz 

- 3000 Hz 

Ψηφιοποίηση

- **Δειγματοληψία**

- δείγμα Spring – A. Vivaldi

- 44100 Hz 

- 22050 Hz 

- 11025 Hz 

- 8000 Hz 

- 6000 Hz 

- 4000 Hz 

- 3000 Hz 

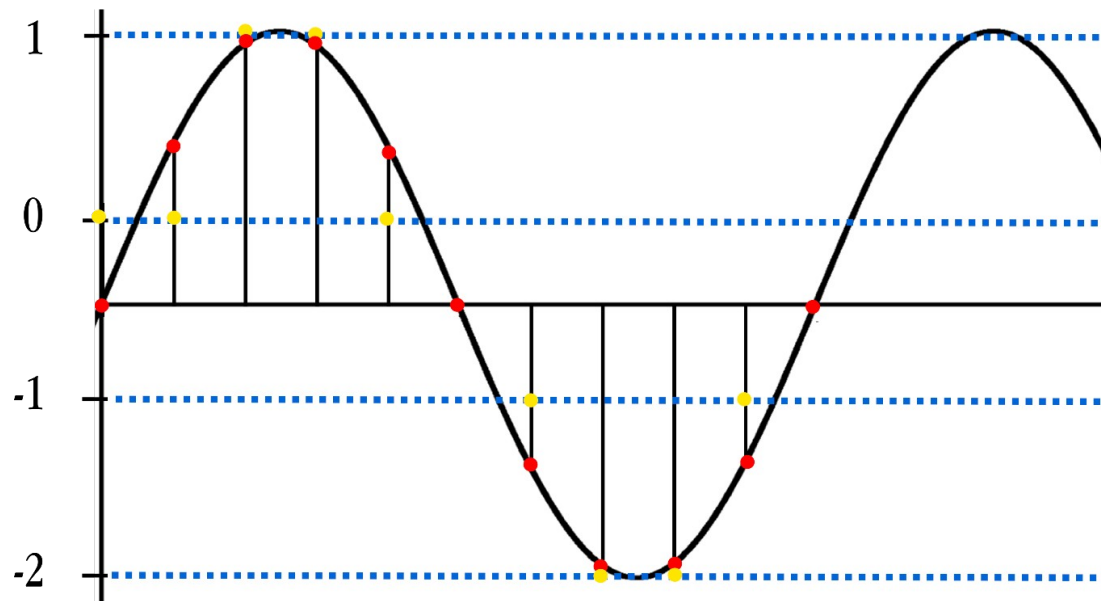
Ψηφιοποίηση

- **Κβαντοποίηση**

- σε κάθε δείγμα εκχωρείται μια τιμή από αυτές που επιτρέπει η κλίμακα
- κάθε δείγμα παίρνει ως ψηφιακή τιμή αυτή που είναι πλησιέστερη προς την αρχική αναλογική του τιμή
- μέγεθος δείγματος (sampling size)

Ψηφιοποίηση

- Κβαντοποίηση



Ψηφιοποίηση

- Μέγεθος δείγματος
 - δείγμα Fur Elise – L. Van Beethoven
 - 256 kbps
 - 192 kbps
 - 128 kbps
 - 96 kbps 
 - 64 kbps 
 - 32 kbps 
 - 16 kbps 
 - 8 kbps 

Ψηφιοποίηση

- Μέγεθος δείγματος
 - δείγμα Spring – A. Vivaldi
 - 256 kbps 
 - 192 kbps 
 - 128 kbps 
 - 96 kbps 
 - 64 kbps 
 - 32 kbps 
 - 16 kbps 
 - 8 kbps 

Ψηφιοποίηση

- **Κωδικοποίηση**

- δημιουργία ψηφιακού σήματος με την αντιστοίχιση τιμών κβάντωσης των δειγμάτων σε ψηφιακούς κωδικούς
- το ψηφιακό σήμα απαρτίζεται από μια ακολουθία από bit

Ψηφιοποίηση

- Κωδικοποίηση
 - χρήση Παλμοκωδικής διαμόρφωσης
 - χρήση άλλης τεχνικής

Παλμοκωδική Διαμόρφωση

- Παλμοκωδική διαμόρφωση (PCM)
 - πιο διαδεδομένος τρόπος ψηφιοποίησης σήματος γιατί είναι απλός και ασυμπίεστος
- Είδη PCM
 - γραμμική (Linear PCM)
 - διαφορική (Differential PCM)
 - λογαριθμική (logarithmic A-Law/μ-Law)

Παλμοκωδική Διαμόρφωση

- Γραμμική PCM

- κάθε δείγμα είναι ανεξάρτητο
- κωδικοποιείται κάθε δείγμα
- τα δείγματα ισοκατανέμονται σε όλο το εύρος τιμών σήματος
- δεν υπάρχει συμπίεση

Παλμοκωδική Διαμόρφωση

- **Διαφορική PCM (DCPM)**
 - δεν κωδικοποιείται κάθε δείγμα
 - χρήση πρόβλεψης για κωδικοποίηση της διαφοράς μεταξύ παρατηρούμενης και προβλεπόμενης τιμής
 - πλεονασμός καθώς τα διαδοχικά δείγματα παρουσιάζουν υψηλή συνάφεια
 - 4 bit ανά δείγμα αντί 16 bit

Παλμοκωδική Διαμόρφωση

- **Λογαριθμική**

- τα υψηλά πλάτη είναι λιγότερο αντιληπτά οπότε μπορούν να αναπαρασταθούν με μικρότερη ακρίβεια από ότι τα χαμηλά πλάτη
- μη γραμμική αντιστοίχιση τιμών: περισσότερες τιμές στα χαμηλά πλάτη και λιγότερες στα υψηλά
- χρήση 8 bit για κωδικοποίηση δυναμικού εύρους 14 bit γραμμικής PCM

CDDA

- **CDDA** (Compact Disc Digital Audio)
 - το στάνταρ για τα CD ήχου
 - χαρακτηριστικά
 - δειγματοληψία: 44100 Khz
 - αριθμός bit ανά δείγμα: 16 (2 byte)
 - κανάλια: 2 (στέρεο)
 - ρυθμός μετάδοσης δεδομένων (bitrate)
 - Δειγματοληψία x Αριθμός bit x Κανάλια
 - ρυθμός μετάδοσης δεδομένων = 1411.1 kbps

CDDA

- Μνήμη
 - μεγάλες απαιτήσεις σε μνήμη για ψηφιακό ήχο ποιότητας CD
 - ένα δευτερόλεπτο απαιτεί 172,27 KB αποθηκευτικού χώρου
 - ένα λεπτό ψηφιακού ήχου απαιτεί 10,09 MB αποθηκευτικού χώρου

DVDA

- **DVD Audio** (Digital Versatile Disk Audio)
 - στάνταρτ για ψηφιακό ήχο υψηλής πιστότητας
 - χαρακτηριστικά
 - δειγματοληψία: 44.100 Hz, 48.000 Hz, 88.200 Hz, 96.000 Hz, 176.400 Hz, 192.000 Hz
 - αριθμός bit: 16, 20, 24 bit
 - κανάλια: 2 (στέρεο) – 5.1 (full surround)

DVDA

- Μνήμη
 - οι απαιτήσεις σε μνήμη για τον ψηφιακό ήχο ποιότητας DVD είναι πολύ μεγαλύτερες σε σχέση με τις αντίστοιχες του CD ήχου
 - 1 δευτερόλεπτο DVD audio απαιτεί 1,09 MB μνήμης
 - 1 λεπτό ήχου πιστότητας DVD απαιτεί 65,91 MB μνήμης

Κωδικοποίηση

Συμπίεση

- κωδικοποίηση ήχου
 - χρήση PCM
 - δειγματοληψία
 - πλήθος bit ανά δείγμα
 - χρήση άλλων μεθόδων για να κωδικοποιηθεί ο αρχικός αριθμός bit με μικρότερο αριθμό

Συμπίεση

- κωδικοποίηση ήχου
 - κωδικοποίηση πηγής για φωνή
 - αντιληπτική κωδικοποίηση ήχου

Συμπύεση

- αντιληπτική κωδικοποίηση ήχου
 - εκμετάλλευση χαρακτηριστικών ακουστικής αντίληψης για αποφυγή κωδικοποίησης και μετάδοσης ήχων που δεν γίνονται αντιληπτοί
 - μέγιστο δυναμικό εύρος ακοής
 - 1-5 KHz
 - μειώνεται σε μικρότερες και μεγαλύτερες συχνότητες

Συμπύεση

- ψυχοακουστικά μοντέλα
 - τεχνικές προσαρμοσμένες στην ανθρώπινη ακοή
- διαδικασία
 - ανάλυση σήματος για εντοπισμό των μη αντιληπτών τμημάτων
 - εφόσον είναι μη αντιληπτά δεν υπάρχει λόγος κωδικοποίησης τους

Συμπίεση

- **απόκρυψη συχνοτήτων (frequency masking):**
 - ένα ασθενές σήμα
 - γίνεται αντιληπτό όταν υφίσταται μόνο του
 - δεν γίνεται αντιληπτό όταν υπάρχει ένα ισχυρό σήμα σε γειτονική συχνότητα
 - ανάλυση σήματος για εντοπισμό ισχυρών τμημάτων που αποκρύπτουν γειτονικά

Συμπύεση

- **απόκρυψη χρόνου** (temporal masking)
 - ένα ισχυρό σήμα αποκρύπτει τα πιο ασθενή σήματα στην ίδια περιοχή για κάποιο χρονικό διάστημα πριν και μετά την εμφάνιση του
 - ανάλυση σήματος για εντοπισμό ισχυρών τμημάτων και υπολογισμό των ασθενών σημάτων που θα αποκρύπτονται στο ίδιο χρονικό διάστημα

MPEG-1

- Το πρότυπο MPEG-1 περιλαμβάνει 3 επίπεδα κωδικοποίησης ήχου (layers)
- 3ο επίπεδο: MPEG-1 Audio Layer 3 ή MP3
 - χαρακτηριστικά
 - υποστηρίζει δύο κανάλια (στέρεο)
 - μπορεί να πετύχει πολύ υψηλά επίπεδα συμπίεσης 10:1 -12:1

MPEG-1

- χαρακτηριστικά
 - ρυθμός δειγματοληψίας
 - 32 KHz
 - 44,1 KHz
 - 48 KHz
 - ροή δεδομένων
 - ελάχιστο: 32 Kbps
 - μέγιστο: 448 kbps (επίπεδο 1), 384 (επίπεδο 2), 320 (επίπεδο 3)

MPEG-1

- CDDA vs. MPEG-1 Audio Layer 3
 - η αντίληψη ποιότητας εξαρτάται από
 - είδος μουσικής
 - εκπαίδευση ακροατή

MPEG-1

- CDDA vs. MPEG-1 Audio Layer 3
 - για κάποια είδη μουσικής, μη εκπαιδευμένοι ακροατές δεν μπορούν να διακρίνουν αντιληπτικά μεταξύ ποιότητας CD και MPEG-1 Layer 3 με ροή δεδομένων 128 Kbps
 - ανεξαρτήτως είδους μουσικής και εκπαίδευσης ακροατών, το MPEG-1 Layer 3 με ροή δεδομένων > 192 Kbps δεν μπορεί να διακριθεί από ποιότητα CD

MPEG-2

- χαρακτηριστικά
 - επιτρέπει ρυθμούς δειγματοληψίας μειωμένους κατά 50%
 - 16 KHz
 - 22,05 KHz
 - 24 KHz

MPEG-2

- χαρακτηριστικά
 - πολυκάναλο ήχο
 - 5 κανάλια πλήρους εύρους ζώνης (L,C,R, LS, RS)
 - 1 κανάλι βελτίωσης χαμηλών συχνοτήτων

MPEG-4

- το part-3 αφορά ήχο και επεκτείνει το Advanced Audio Coding (AAC) του MPEG-2 part 7
- χαρακτηριστικά
 - αποσκοπεί στην καλύτερη ποιότητα ήχου για δεδομένη ροή δεδομένων
 - επιτρέπει το συνδυασμό φυσικού και συνθετικού ήχου
 - χρησιμοποιεί αντικείμενα ήχου (audio objects) για να ορίσει ηχητικά σήματα διαφόρων τύπων

MPEG-4

- χαρακτηριστικά
 - ρυθμός δειγματοληψίας: 8 - 96 kHz
 - κανάλια: μέχρι 48
 - ροή δεδομένων ελεγχόμενη από χρήστη
 - καλύτερη ποιότητα σε σχέση με το MP3 για την ίδια ροή δεδομένων
 - στάνταρτ μορμά ήχου για υπηρεσίες (π.χ. YouTube), συσκευές (π.χ. τηλέφωνα, ταμπλέτες κτλ) και εκπομπή (π.χ. HDTV)

Κωδικοποιητές Ήχου

- Διαδεδομένοι κωδικοποιητές ήχου (audio codecs)
 - MP3
 - AAC
 - Vorbis
 - FLAC
 - WMA

Φορμά

Φορμά Ήχου

- Διαδεδομένα φορμά ήχου (audio file formats)
 - AIFF
 - AAC
 - FLAC
 - MP3
 - OGG
 - RA
 - WAV
 - WMA

Φορμά Ήχου

- Διαδεδομένα φορμά ήχου
 - **AIFF** (Audio Interchange File Format)
 - αναπτύχθηκε από την Apple
 - χωρίς συμπίεση
 - επεκτάσεις αρχείων: *. aiff, *.aif
 - **AAC** (Advanced Audio Coding)
 - πρότυπο ήχου MPEG-2/4
 - απωλεστική συμπίεση
 - επεκτάσεις αρχείων: *.m4a, *.m4r, *m4p *.mp4, *.3gp
*.aac

Φορμά Ήχου

- Διαδεδομένα φορμά ήχου
 - **FLAC** (Free Lossless Audio Coding)
 - μη απωλεστική συμπίεση
 - ανοικτό φορμά, χωρίς πατέντες
 - επεκτάσεις: *.flac
 - **MP3** (MPEG-1 Audio Layer 3)
 - πρότυπο ήχου MPEG-1
 - απωλεστική συμπίεση
 - επεκτάσεις: *.mp3

Φορμά Ήχου

- Διαδεδομένα φορμά ήχου
 - **OGG**
 - αναπτύχθηκε από το ίδρυμα Xiph.org
 - ανοικτό φορμά, χωρίς πατέντες
 - επεκτάσεις: *.ogg, *.oga
 - **RA/RM** (RealAudio)
 - αναπτύχθηκε από την RealNetworks
 - ιδιόκτητο φορμά
 - επεκτάσεις: *.ra, *.ram

Φορμά Ήχου

- Διαδεδομένα φορμά ήχου
 - **WAV** (Waveform Audio File Format)
 - αναπτύχθηκε από την M\$
 - χωρίς συμπίεση
 - επεκτάσεις: *.wav, *.wave
 - **WMA** (Windows Media Audio)
 - αναπτύχθηκε από την M\$
 - μη απωλεστική & απωλεστική συμπίεση (WMA9)
 - επεκτάσεις: *.wma

Φορμά Ήχου

- Συμβατότητα
 - δεν είναι όλα τα φορμά ήχου (container formats) συμβατά με όλους τους κωδικοποιητές ήχου (audio codecs)
 - παράδειγμα
 - OGG: συμβατό με Vorbis αλλά όχι με MP3
 - MP4: συμβατό με AAC αλλά όχι με Vorbis

MIDI

- συμβολική αναπαράσταση μουσικής
 - πρότυπο ψηφιακής διεπαφής μουσικών οργάνων - **MIDI** (Musical Instrument Digital Interface)
 - ψηφιακό αντίστοιχο παρτιτούρας
 - διασύνδεση μουσικών οργάνων σε υπολογιστή για τον έλεγχο της παραγωγής μουσικών ήχων από υπολογιστή

MIDI

- χαρακτηριστικά
 - καταγραφή γεγονότων (events)
 - για κάθε νότα που παράγεται αποθηκεύονται
 - θεμελιώδης συχνότητα (νότα)
 - αρμονικές (ηχόχρωμα)
 - δυναμική (πλάτος ταλάντωσης)
 - διάρκεια (αρχή-τέλος ήχου)
 - παραδείγματα
 - fur Elise 
 - Sonata_en_Do-Mayor-K6 
 - Polonaise_in_F_major 

Μνήμη

- Πόση μνήμη απαιτείται για την αποθήκευση ενός ήχου διάρκειας 2' 30" στις ακόλουθες περιπτώσεις;
 - Wav (PCM)
 - MP3 με ροή δεδομένων 196 kbps