

**ΟΙΚΟΝΟΜΙΚΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΑΘΗΝΩΝ**



**ATHENS UNIVERSITY
OF ECONOMICS
AND BUSINESS**

Τεχνολογία Πολυμέσων

Ενότητα # 4: Ήχος

Διδάσκων: Γεώργιος Ξυλωμένος

Τμήμα: Πληροφορικής



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στα πλαίσια του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «**Ανοικτά Ακαδημαϊκά Μαθήματα στο Οικονομικό Πανεπιστήμιο Αθηνών**» έχει χρηματοδοτήσει μόνο τη αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



Άδειες Χρήσης

- Το παρόν εκπαιδευτικό υλικό υπόκειται σε άδειες χρήσης Creative Commons.
- Οι εικόνες προέρχονται από το βιβλίο «Τεχνολογία Πολυμέσων και Πολυμεσικές Επικοινωνίες», Γ.Β. Ξυλωμένος, Γ.Κ. Πολύζος, 1^η έκδοση, 2009, Εκδόσεις Κλειδάριθμος.



Σκοποί ενότητας

- Κατανόηση των βασικών χαρακτηριστικών του ήχου.
- Εισαγωγή στις βασικές τεχνικές ψηφιοποίησης και κβαντοποίησης σημάτων
- Κατανόηση της παλμοκωδικής διαμόρφωσης.
- Εισαγωγή στη συμβολική αναπαράσταση της μουσικής.

Περιεχόμενα ενότητας

- Χαρακτηριστικά του ήχου
- Ψηφιοποίηση με μετασχηματισμό
- Ψηφιοποίηση με δειγματοληψία
- Κβαντοποίηση δειγμάτων
- Παλμοκωδική διαμόρφωση
- Συμβολική αναπαράσταση

**ΟΙΚΟΝΟΜΙΚΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΑΘΗΝΩΝ**

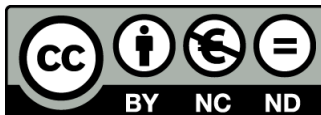


**ATHENS UNIVERSITY
OF ECONOMICS
AND BUSINESS**

Χαρακτηριστικά του ήχου

Μάθημα: Τεχνολογία Πολυμέσων, **Ενότητα # 4:** Ήχος

Διδάσκων: Γιώργος Ξυλωμένος, **Τμήμα:** Πληροφορικής



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



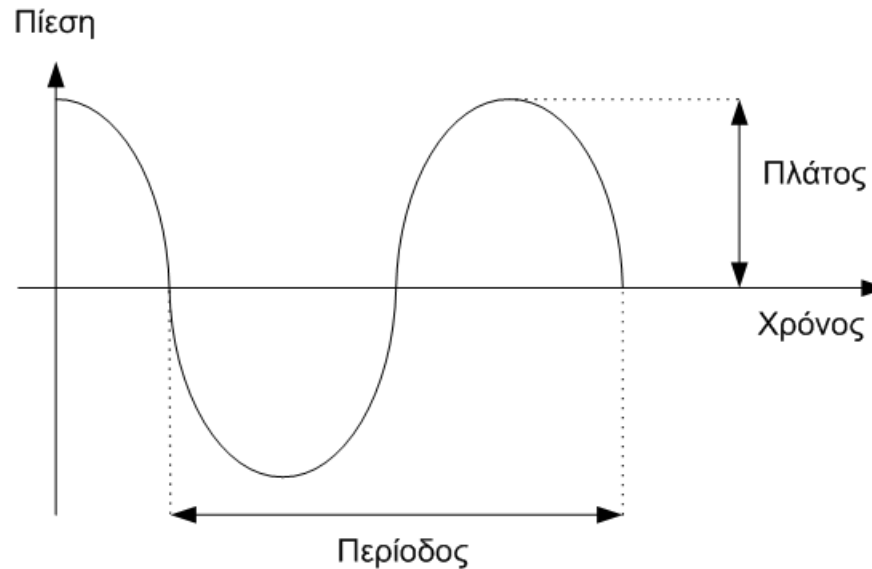
Τι είναι ο ήχος;

- Παράγεται από την ταλάντωση της ύλης
 - Δημιουργεί μεταβολές στην πίεση του αέρα
 - Η πίεση γίνεται αντιληπτή από το αυτί
- Μεταδίδεται με κυματοειδή μορφή
 - Μετάδοση και σε στερεά / υγρά
 - Δεν μπορεί να μεταδοθεί στο κενό
 - Το μέσο μετάδοσης είναι απαραίτητο

Κατηγοριοποίηση ήχων

- Ακουστοί ήχοι: 20 Hz - 20 kHz
 - Ακουστικά σήματα (acoustic signals)
 - Αντιληπτά από το ανθρώπινο αυτί
 - Κάτω από 20 Hz: υπόηχοι
 - Πάνω από 20 kHz: υπέρηχοι
 - Στην πράξη, κάθε άνθρωπος διαφέρει

Ηχητικές κυματομορφές (1 από 2)



- Κυματομορφή (waveform)
 - Αναπαράσταση των μεταβολών της πίεσης
 - Το πλάτος δείχνει την πίεση σε κάθε στιγμή

Ηχητικές κυματομορφές (2 από 2)

- Πλάτος (amplitude): ένταση του ήχου
 - Μετατόπιση του κύματος από το μέσο όρο
- Περίοδος κυματομορφής
 - Οι περιοδικοί ήχοι είναι γενικά πιο μουσικοί
- Συχνότητα: αντίστροφο της περιόδου
 - Κύκλοι ανά δευτερόλεπτο (Hertz, Hz)

Ψηφιοποίηση

- Ηχητική κυματομορφή: αναλογικό σήμα
 - Περίπλοκη μαθηματική συνάρτηση $g(t)$
 - Οποιαδήποτε τιμή σε οποιαδήποτε στιγμή
 - Στον υπολογιστή έχουμε περιορισμούς
 - Πεπερασμένο πλήθος τιμών
 - Πεπερασμένη ακρίβεια τιμών
- Ψηφιοποίηση
 - Μετατροπή αναλογικού σε ψηφιακό

**ΟΙΚΟΝΟΜΙΚΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΑΘΗΝΩΝ**

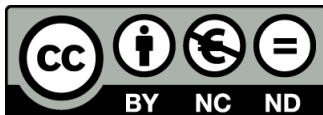


**ATHENS UNIVERSITY
OF ECONOMICS
AND BUSINESS**

Ψηφιοποίηση με μετασχηματισμό

Μάθημα: Τεχνολογία Πολυμέσων, **Ενότητα # 4:** Ήχος

Διδάσκων: Γιώργος Ξυλωμένος, **Τμήμα:** Πληροφορικής



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



Μετασχηματισμός Fourier (1 από 3)

- Θεώρημα Fourier
 - Η $g(t)$ αναλύεται σε απλές συναρτήσεις
 - Μπορεί να είναι άπειρες...
 - $f=1/T$: θεμελιώδης συχνότητα
 - a_n και b_n : αρμονικές (harmonics) του σήματος

$$g(t) = \frac{1}{2}c + \sum_{n=1}^{\infty} a_n \sin(2\pi nft) + \sum_{n=1}^{\infty} b_n \cos(2\pi nft)$$

Μετασχηματισμός Fourier (2 από 3)

$$a_n = \frac{2}{T} \int_0^T g(t) \sin(2\pi nft) dt$$

$$b_n = \frac{2}{T} \int_0^T g(t) \cos(2\pi nft) dt$$

$$c = \frac{2}{T} \int_0^T g(t) dt$$

- Υπολογισμός συντελεστών $g(t)$
 - Ανασύνθεση με χρήση των συντελεστών
 - Οι συντελεστές είναι μία σειρά αριθμών

Μετασχηματισμός Fourier (3 από 3)

$$\sqrt{a_n^2 + b_n^2}$$

- Ενέργεια αρμονικής (RMS)
 - Δεν μειώνεται με το n
 - Οι αρμονικές μπορεί να είναι πάρα πολλές
- Προβλήματα με τους συντελεστές
 - Περιορισμένη ακρίβεια πράξεων
 - Περιορισμένη ακρίβεια αποθήκευσης
 - Περιορισμένο πλήθος συντελεστών

**ΟΙΚΟΝΟΜΙΚΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΑΘΗΝΩΝ**



**ATHENS UNIVERSITY
OF ECONOMICS
AND BUSINESS**

Ψηφιοποίηση με δειγματοληψία

Μάθημα: Τεχνολογία Πολυμέσων, **Ενότητα # 4:** Ήχος

Διδάσκων: Γιώργος Ξυλωμένος, **Τμήμα:** Πληροφορικής



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο

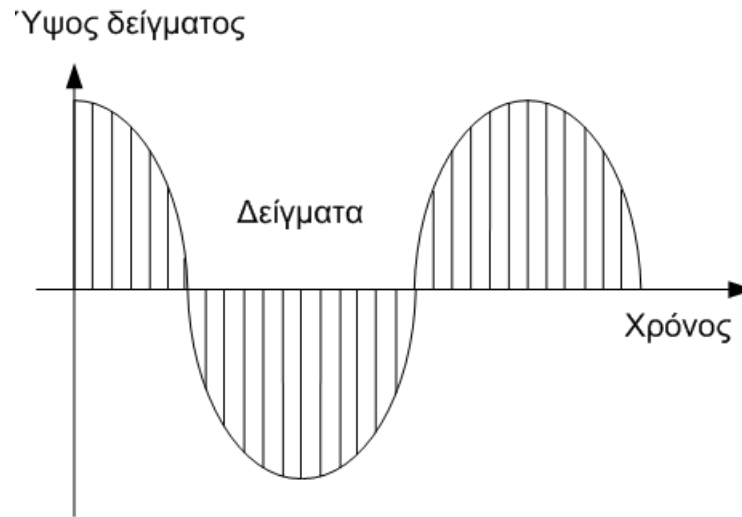


ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



Δειγματοληψία (1 από 2)

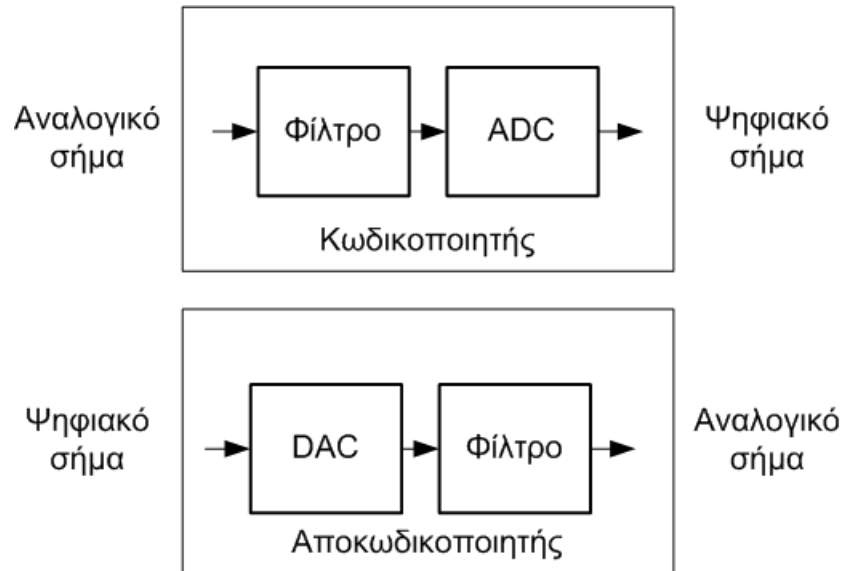


- Ψηφιοποίηση με δειγματοληψία
 - Μέτρηση πλάτους κυματομορφής
 - Κάθε μέτρηση ονομάζεται δείγμα (sample)
 - Διακριτή διάσταση χρόνου

Δειγματοληψία (2 από 2)

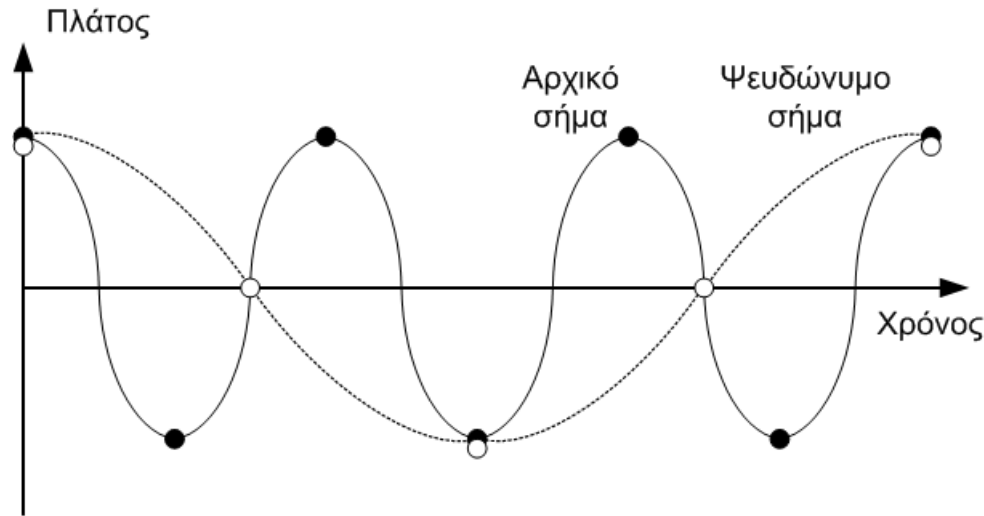
- Ρυθμός δειγματοληψίας (sampling rate)
 - Μετριέται σε Hz (δείγματα / δευτερόλεπτο)
 - 44,1 kHz στα CD, 8 KHz στην τηλεφωνία
- Θεώρημα δειγματοληψίας του Nyquist
 - Έστω σήμα με μέγιστη συχνότητα F
 - Ρυθμός δειγματοληψίας τουλάχιστον $2F$
 - Αποφεύγεται η απώλεια πληροφορίας

Μετατροπείς (1 από 3)



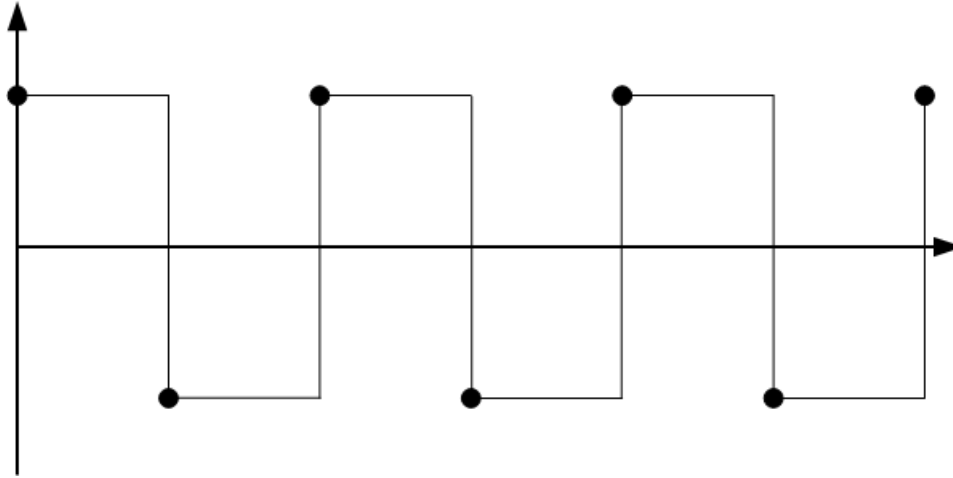
- Διατάξεις μετατροπής
 - ADC: Μετατροπή αναλογικού σε ψηφιακό
 - DAC: Μετατροπή ψηφιακού σε αναλογικό
 - Φίλτρα πριν το ADC και μετά το DAC

Μετατροπείς (2 από 3)



- Γιατί φίλτρο πριν τον κωδικοποιητή;
 - Ψευδώνυμα σήματα (alias signals)
 - Μαύρες κουκκίδες: επαρκής δειγματοληψία
 - Λευκές κουκκίδες: ανεπαρκής δειγματοληψία
 - Το φίλτρο αποκόπτει συχνότητες άνω των f Hz

Μετατροπείς (3 από 3)



- Γιατί φίλτρο μετά τον αποκωδικοποιητή;
 - Τα δείγματα αναπαράγονται περιοδικά
 - Η έξοδος έχει τετραγωνική κυματομορφή
 - Το φίλτρο αποκόπτει συχνότητες πάνω από f Hz
 - Το παραμορφωμένο σήμα είναι πιο φυσικό

**ΟΙΚΟΝΟΜΙΚΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΑΘΗΝΩΝ**



**ATHENS UNIVERSITY
OF ECONOMICS
AND BUSINESS**

Κβαντοποίηση δειγμάτων

Μάθημα: Τεχνολογία Πολυμέσων, **Ενότητα # 4:** Ήχος

Διδάσκων: Γιώργος Ξυλωμένος, **Τμήμα:** Πληροφορικής



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



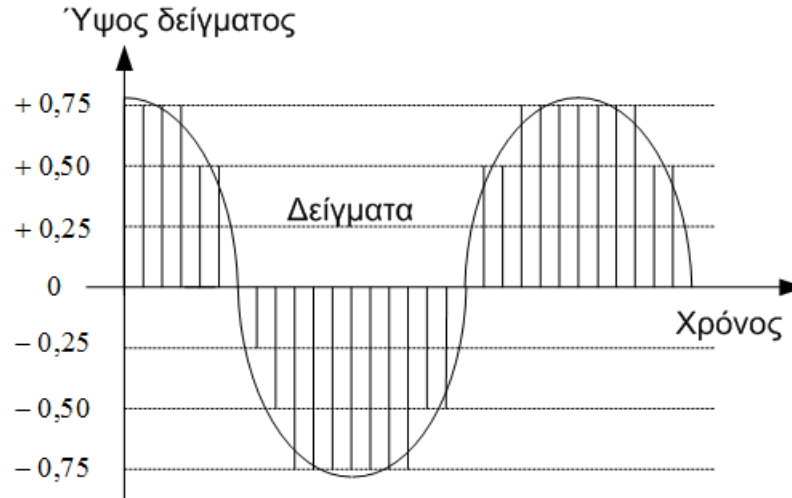
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

Κβαντοποίηση (1 από 2)



- Κβαντοποίηση (quantization)
 - Αναπαράσταση συνεχών με διακριτές τιμές
 - Κάθε δείγμα προσεγγίζεται με μία τιμή
 - Διακριτή διάσταση πλάτους

Κβαντοποίηση (2 από 2)

- Επηρεάζει την ποιότητα
 - 16 bits: 65536 τιμές, 8 bits: 256 τιμές
 - Λιγότερες τιμές, μεγαλύτερο σφάλμα
 - Σφάλμα κβαντοποίησης (quantization error)
- Πόσα επίπεδα χρειαζόμαστε;
 - Δεν υπάρχει γενικό θεώρημα
 - Θέτουμε άνω όριο στο σφάλμα
 - Με πόσα bit επιτυγχάνεται το όριο;

Σφάλμα κβαντοποίησης

- Υπολογισμός σφάλματος κβαντοποίησης
 - Έστω σήμα εισόδου με πλάτος $-V$ έως $+V$
 - Έστω n bit για την κβαντοποίηση
 - Πλάτος διαστήματος $q = 2V / 2^n$
 - Υποθέτουμε γραμμική κβαντοποίηση
 - Μέγιστο σφάλμα κβαντοποίησης
 - $q/2 = V / 2^n$ (το μισό του διαστήματος)

Επίπεδα κβαντοποίησης (1 από 2)

- Δυναμικό εύρος εισόδου
 - Το ελάχιστο αντιληπτό σήμα έχει πλάτος v
 - Δυναμικό εύρος: $20 \log_{10}(V/v)$ dB
- Πρώτη προσέγγιση: χρήση δυναμικού εύρους
 - Σφάλμα $<$ ελάχιστο αντιληπτό σήμα
 - $V / 2^n < v$ σημαίνει $\log_{10}(V/v) < n \log_{10}2 = 0,3 n$
 - Σε όρους δυναμικού εύρους: $20 \log_{10}(V/v) < 6 n$
 - Αν το δυναμικό εύρος είναι 40 dB τότε $n \geq 7$

Επίπεδα κβαντοποίησης (2 από 2)

- Δεύτερη προσέγγιση: χρήση SNR
 - Ισχύς σήματος / ισχύς θορύβου $> x$ dB
 - SNR: $10 \log_{10}(V/(q/2))^2 = 20 \log_{10}(V/(q/2))$
 - Ισχύς ανάλογη με το τετράγωνο του πλάτους
 - Με αντικατάσταση έχουμε $20 \log_{10}(V/(V / 2^n))$
 - Τελικά λοιπόν $20 \log_{10}2^n = 6 n$
 - Έστω ότι θέλουμε SNR > 40 dB
 - Για να ισχύει $6 n > 40$ dB πρέπει $n \geq 7$

**ΟΙΚΟΝΟΜΙΚΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΑΘΗΝΩΝ**

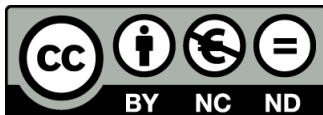


**ATHENS UNIVERSITY
OF ECONOMICS
AND BUSINESS**

Παλμοκωδική διαμόρφωση

Μάθημα: Τεχνολογία Πολυμέσων, **Ενότητα # 4:** Ήχος

Διδάσκων: Γιώργος Ξυλωμένος, **Τμήμα:** Πληροφορικής



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



Τύποι PCM (1 από 4)

- Παλμοκωδική διαμόρφωση (PCM)
 - Απλή περιοδική δειγματοληψία
 - Χωρίς μνήμη: κάθε δείγμα είναι ανεξάρτητο
 - Αναπαραγωγή από οποιοδήποτε δείγμα
- Γραμμική (linear) κβαντοποίηση
 - Ίσα διαστήματα τιμών δειγμάτων
 - Μία τιμή ανά διάστημα
 - Απλή στην υλοποίηση

Τύποι PCM (2 από 4)

- Μουσικό CD: πρότυπο CD-DA
 - Γραμμική κβαντοποίηση: ίδια ακρίβεια παντού
 - Μεγαλύτερη πιστότητα για τυχαίος ήχους
 - Εύρος συχνοτήτων 20 kHz
 - Δειγματοληψία 44,1 kHz
 - Κβαντοποίηση με τιμές των 16 bit
 - Ρυθμός μετάδοσης 1,411 Mbps
 - Στερεοφωνικός ήχος (δύο κανάλια)

Τύποι PCM (3 από 4)

- Λογαριθμική (logarithmic) κβαντοποίηση
 - Λογαριθμίζουμε και μετά κβαντοποιούμε
 - Η ίδια η κβαντοποίηση γίνεται γραμμικά
 - Αντιστρέφουμε στον αποκωδικοποιητή
 - Μεγαλύτερη ακρίβεια σε ένα άκρο τιμών
 - Γιατί να δώσουμε έμφαση στο ένα άκρο;
 - Είτε το σήμα έχει περισσότερη πληροφορία εκεί
 - Είτε η αντίληψή μας είναι πιο οξεία εκεί

Τύποι PCM (4 από 4)

- Φωνητική τηλεφωνία: πρότυπο ITU G.711
 - Λογαριθμική: ακρίβεια στα χαμηλά πλάτη
 - Χρήση compressor/expander ή compander
 - Τροποποίηση πριν/μετά τη δειγματοληψία
 - A-law (Ευρώπη), μ-law (Αμερική / Ιαπωνία)
 - Εύρος 3.1-3.5 kHz, δειγματοληψία 8 kHz
 - Τιμές των 8 bit (ή 7 bit)
 - Ισοδύναμα με 12-14 bit σε γραμμική κβαντοποίηση
 - Ρυθμός μετάδοσης 64 Kbps

**ΟΙΚΟΝΟΜΙΚΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΑΘΗΝΩΝ**



**ATHENS UNIVERSITY
OF ECONOMICS
AND BUSINESS**

Συμβολική αναπαράσταση

Μάθημα: Τεχνολογία Πολυμέσων, **Ενότητα # 4:** Ήχος

Διδάσκων: Γιώργος Ξυλωμένος, **Τμήμα:** Πληροφορικής



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



Πρότυπο MIDI (1 από 3)

- Συμβολική αναπαράσταση: MIDI
 - Προδιαγραφές υλικού διασύνδεσης
 - Βύσματα, ηλεκτρικά σήματα
 - Προδιαγραφές μορφοποίησης δεδομένων
- Τα μηνύματα MIDI περιγράφουν γεγονότα
 - Ενέργειες που εκτελεί ένας μουσικός
 - Πάτημα και απελευθέρωση πλήκτρων
 - Συχνότητα και εφφέ των ήχων

Πρότυπο MIDI (2 από 3)

- 16 κανάλια γεγονότων
 - Ελέγχουν ένα ή περισσότερα όργανα
 - Μονοφωνικά ή πολυφωνικά όργανα
 - General MIDI: 128 τυποποιημένα όργανα
- Σύνθεση ήχου
 - Χρήση γεννήτριας συχνοτήτων (FM)
 - Τροποποίηση ήχου με φίλτρα
 - Αναπαραγωγή ψηφιοποιημένων ήχων (wavetable)
 - Αποθηκευμένα δείγματα

Πρότυπο MIDI (3 από 3)

- Πλεονεκτήματα και μειονεκτήματα MIDI
 - Οικονομική αναπαράσταση
 - Δυνατότητα συμβολικής επεξεργασίας
 - Αποτέλεσμα ανάλογο με το υλικό
 - Διαφέρει σημαντικά από συσκευή σε συσκευή
 - Κατάλληλο μόνο για μουσικά όργανα
 - Βασίζεται στη δομή συγκεκριμένων ήχων
 - Δεν είναι κατάλληλο π.χ. για ανθρώπινη φωνή

**ΟΙΚΟΝΟΜΙΚΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΑΘΗΝΩΝ**



**ATHENS UNIVERSITY
OF ECONOMICS
AND BUSINESS**

Τέλος Ενότητας #4

Μάθημα: Τεχνολογία Πολυμέσων, **Ενότητα # 4:** Ήχος
Διδάσκων: Γιώργος Ξυλωμένος, **Τμήμα:** Πληροφορικής



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ