

**ΟΙΚΟΝΟΜΙΚΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΑΘΗΝΩΝ**



**ATHENS UNIVERSITY
OF ECONOMICS
AND BUSINESS**

Τεχνολογία Πολυμέσων

Ενότητα # 10: Κωδικοποίηση ήχου

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



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



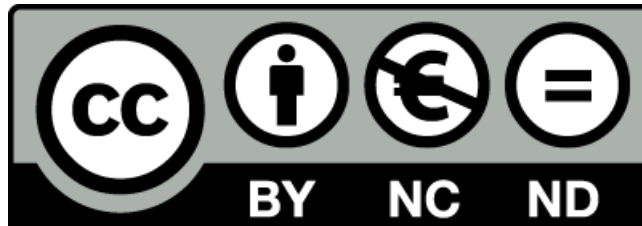
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

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 τεχνικών κωδικοποίησης καναλιού και πηγής, ειδικά για τη φωνή.
- Εισαγωγή στην αντιληπτική κωδικοποίηση και την εφαρμογή της στην κωδικοποίηση γενικών ήχων στα πρότυπα MPEG.

Περιεχόμενα ενότητας

- Κωδικοποίηση καναλιού φωνής
- Κωδικοποίηση πηγής φωνής
- Αντιληπτική κωδικοποίηση
- Κωδικοποίηση ήχου MPEG-1
- Κωδικοποίηση ήχου MPEG-2
- Κωδικοποίηση ήχου MPEG-4

**ΟΙΚΟΝΟΜΙΚΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΑΘΗΝΩΝ**



**ATHENS UNIVERSITY
OF ECONOMICS
AND BUSINESS**

Κωδικοποίηση καναλιού φωνής

Μάθημα: Τεχνολογία Πολυμέσων, **Ενότητα # 10:** Κωδικοποίηση ήχου

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

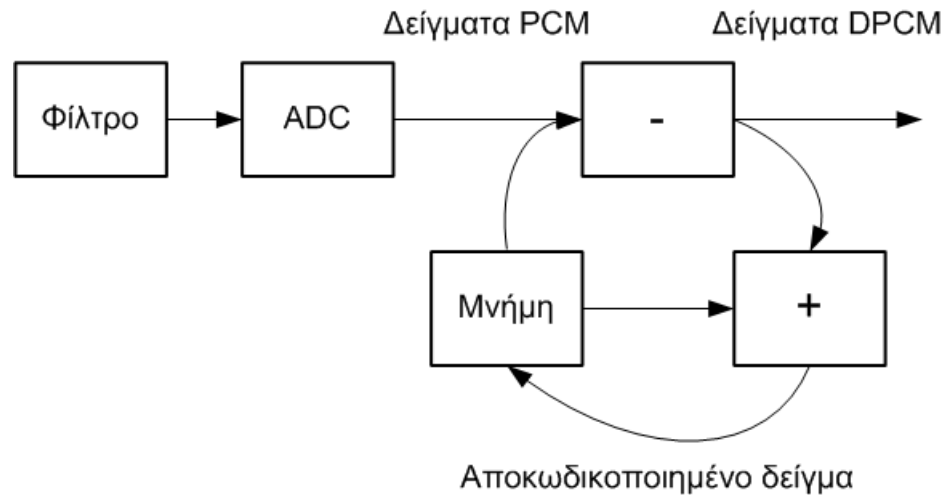


ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

PCM / G.711

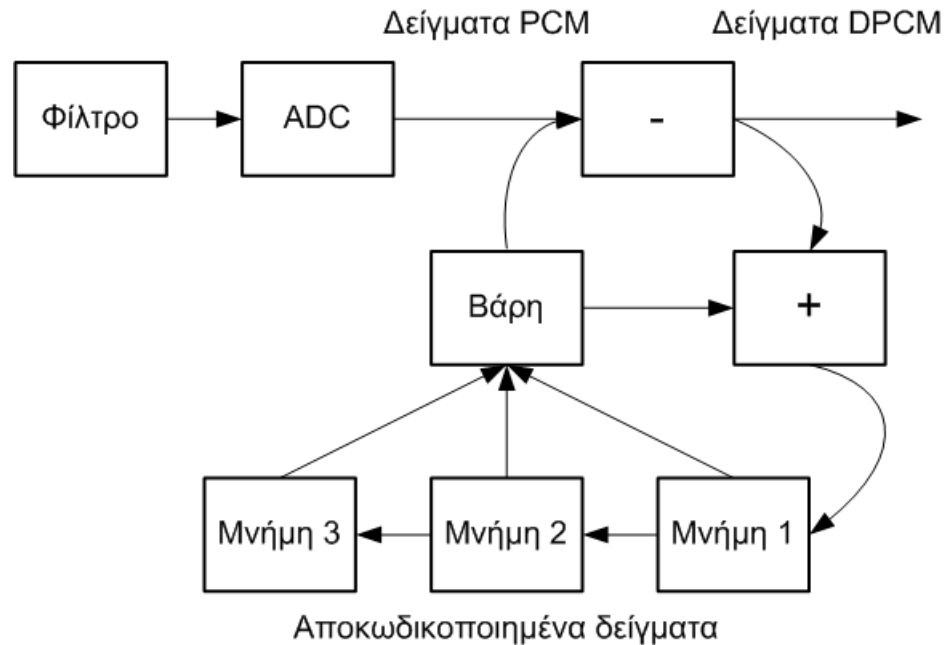
- Πρότυπο ITU G.711
 - Χρήση στην κλασική τηλεφωνία
 - Φιλτράρισμα συχνοτήτων (300-3400 Hz)
 - Λογαριθμική «συμπίεση» σήματος
 - Δειγματοληψία στα 8 KHz
 - Ρυθμός μετάδοσης 64 Kbps
- Βάση για πολλά άλλα πρότυπα
 - Κωδικοποιήσεις καναλιού

Κωδικοποίηση καναλιού (1 από 5)



- Κωδικοποίηση DPCM
 - Κωδικοποίηση διαφορών αντί δειγμάτων
 - Μεταδίδουμε προσέγγιση της διαφοράς
 - Βάση: η προηγούμενη προσέγγιση
 - Προσοχή στο βρόχο αποκωδικοποίησης!

Κωδικοποίηση καναλιού (2 από 5)



- Κωδικοποίηση DPCM με γραμμική πρόβλεψη
 - Γραμμικός συνδυασμός παλιών προβλέψεων
 - Καλύτερες προσεγγίσεις με λιγότερα bit

Κωδικοποίηση καναλιού (3 από 5)

- Κωδικοποίηση ADPCM
 - Χρήση πολλών τιμών για πρόβλεψη
 - Μεταβολή βήματος κβαντοποίησης
- G.721: ποιότητα G.711 στα 32 kbps (DECT)
 - Χρήση 8 προηγούμενων τιμών για πρόβλεψη
 - G.723: αντίστοιχα για 24 και 40 kbps
- G.726: επεκτείνει τα G.721 και G.723
 - Υποστηρίζει ρυθμούς 16, 24, 32 και 40 Kbps

Κωδικοποίηση καναλιού (4 από 5)

- Πώς μεταβάλλεται το βήμα;
 - Το βήμα πολλαπλασιάζεται με παράγοντα μ
 - Παρακολουθούμε την τιμή των διαφορών
 - Κοντά στο 0: περισσότερη λεπτομέρεια
 - Μειώνουμε το μ
 - Μακριά από το 0: λιγότερη λεπτομέρεια
 - Αυξάνουμε το μ

Κωδικοποίηση καναλιού (5 από 5)

- G.722: 64 kbps για εύρος 7 KHz (HD Voice)
 - Χωρίζουμε τη φωνή σε δύο ζώνες συχνοτήτων
 - Κωδικοποίηση κάθε ζώνης με ADPCM
 - 0-3,5 KHz: κωδικοποίηση με 48 kbps
 - Αντίστοιχο με κλασική τηλεφωνία
 - 3,5-7 KHz: κωδικοποίηση με 16 kbps
 - Προσθήκη υψηλότερων συχνοτήτων
 - Πιο φυσική απόδοση της φωνής

**ΟΙΚΟΝΟΜΙΚΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΑΘΗΝΩΝ**



**ATHENS UNIVERSITY
OF ECONOMICS
AND BUSINESS**

Κωδικοποίηση πηγής φωνής

Μάθημα: Τεχνολογία Πολυμέσων, **Ενότητα # 10:** Κωδικοποίηση ήχου

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



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



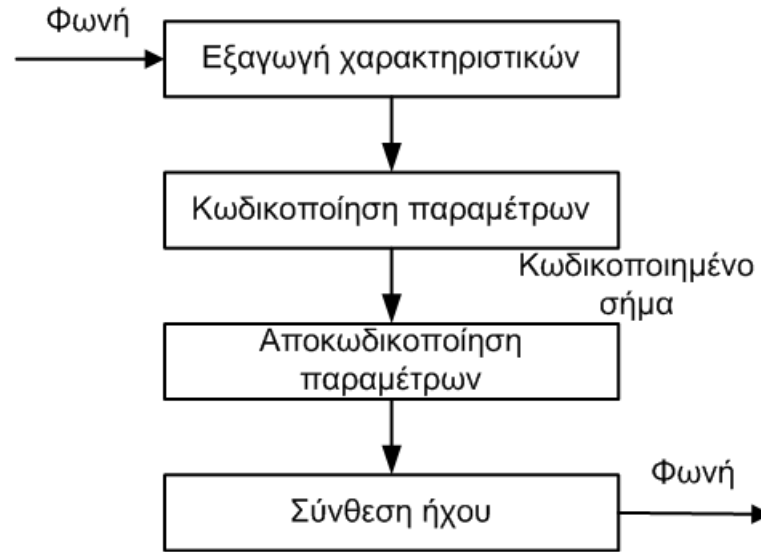
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

Κωδικοποίηση πηγής (1 από 8)



- Κωδικοποιητές φωνής (vocoders)
 - Χρήση μοντέλου ανθρώπινης φωνής
 - Εξαγωγή χαρακτηριστικών και μετάδοσή τους
 - Χαρακτηριστικά: παράμετροι μοντέλου φωνής

Κωδικοποίηση πηγής (2 από 8)

- Φωνήματα (phoneme): βασικοί ήχοι γλώσσας
- Φωνητικοί ήχοι (voiced)
 - Παράγονται από τις φωνητικές χορδές
 - Φωνήεντα και ορισμένα σύμφωνα (π.χ. a, b)
 - Έχουν περιοδική μορφή
- Μη φωνητικοί ήχοι (unvoiced/voiceless)
 - Τα υπόλοιπα σύμφωνα (π.χ. p)
 - Μοιάζουν με θόρυβο

Κωδικοποίηση πηγής (3 από 8)

- Σχηματιστικές συχνότητες (formants)
 - Συχνότητες με σημαντική ενέργεια (2 ή 3)
 - Εξαρτώνται από το φώνημα
 - Διαμόρφωση από λαιμό/στόμα/μύτη
 - Εντοπίζονται με φίλτρα συχνοτήτων
 - Αλλάζουν ανάλογα με το φώνημα
 - Αναλύουμε «πλαίσια» δειγμάτων
 - Αναγνωρίζουμε το αντίστοιχο φώνημα

Κωδικοποίηση πηγής (4 από 8)

- Γραμμική προβλεπτική κωδικοποίηση (LPC)
 - Ανάλυση ενός πλαισίου δειγμάτων
 - Δεν ταυτίζεται με τη διάρκεια του φωνήματος
 - Είναι ο ήχος είναι φωνητικός ή όχι;
 - Φωνητικοί ήχοι: γεννήτρια συχνοτήτων και φίλτρα
 - Μη φωνητικοί ήχοι: γεννήτρια θορύβου
 - Συνδυασμός νέων και παλιών παραμέτρων
 - Γραμμική πρόβλεψη από προηγούμενες παραμέτρους
 - Υπολογισμός νέων παραμέτρων προς αποστολή

Κωδικοποίηση πηγής (5 από 8)

- Γραμμική προβλεπτική κωδικοποίηση (LPC)
 - Αποκωδικοποίηση
 - Γραμμικός συνδυασμός παλιών παραμέτρων
 - Προσθήκη νέας πρόβλεψης
 - LPC-10: γραμμικός συνδυασμός 10 πλαισίων
 - Πλαίσιο: 180 δείγματα στα 8 kHz = 22,5 ms
 - Ρυθμοί μετάδοσης έως και 2,4 Kbps
 - Αναγνωρίσιμη αλλά «τεχνητή» φωνή

Κωδικοποίηση πηγής (6 από 8)

- LPC διεγχειρόμενη με κωδικούς (CELP)
 - Λειτουργεί σε δύο επίπεδα πρόβλεψης
 - Βραχυπρόθεσμη σε επίπεδο δειγμάτων (STP)
 - Μακροπρόθεσμη σε επίπεδο πλαισίων (LTP)
 - Χρήση βιβλιοθήκης έτοιμων προτύπων
 - Περιλαμβάνουν παραμέτρους κωδικοποιητή
 - Εντοπίζεται το καλύτερο ταίριασμα
 - Ουσιαστικά είναι κβαντοποίηση διανυσμάτων

Κωδικοποίηση πηγής (7 από 8)

- LPC διεγχειρόμενη με κωδικούς (CELP)
 - Προσθήκη προσαρμοσμένων προτύπων
 - Χρήση προηγούμενων προβλέψεων σε LTP
- G.723.1: 5,3 ή 6,3 kbps
 - Κατάλληλο για τηλεδιάσκεψη σε δίκτυα PSTN
 - Δειγματοληψία 8 kHz στα 16 bit, πλαίσια 30 ms
 - Διακρίνονται σε 4 υποπλαίσια για πρόβλεψη

Κωδικοποίηση πηγής (8 από 8)

- G.728: 16 kbps
 - Κατάλληλο για τηλεδιάσκεψη σε ISDN
 - Χαμηλή καθυστέρηση σε σχέση με απλό CELP
- G.729: 8 kbps
 - Χρήση στην κινητή τηλεφωνία
 - Το G.729a χρησιμοποιείται στο GSM
 - Πλαίσιο 10 ms για χαμηλότερη καθυστέρηση
 - Προστασία από απώλειες παραμέτρων

**ΟΙΚΟΝΟΜΙΚΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΑΘΗΝΩΝ**



**ATHENS UNIVERSITY
OF ECONOMICS
AND BUSINESS**

Αντιληπτική κωδικοποίηση

Μάθημα: Τεχνολογία Πολυμέσων, **Ενότητα # 10:** Κωδικοποίηση ήχου

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

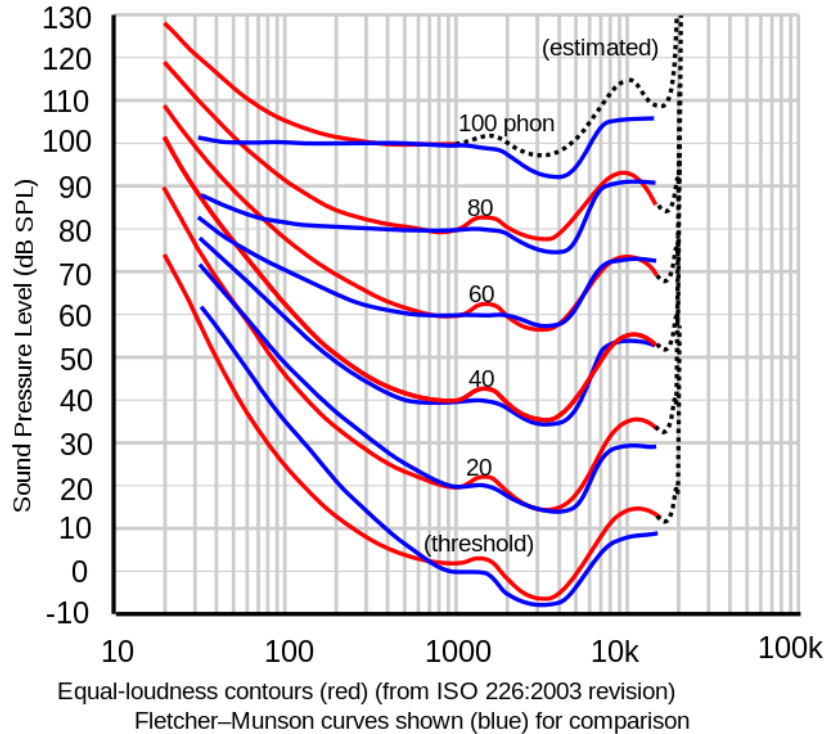
Κωδικοποίηση γενικών ήχων

- Η φωνή είναι ειδική περίπτωση ήχου
 - Συγκεκριμένα χαρακτηριστικά και μοντέλα
 - Ομαλή μεταβολή, συγκεκριμένοι ήχοι
- Πώς να κωδικοποιήσουμε τυχαίο ήχο;
 - Για παράδειγμα, μουσική
 - Η κωδικοποίηση πηγής δεν είναι εφικτή
 - Δεν γνωρίζουμε τα χαρακτηριστικά της πηγής
 - Γνωρίζουμε όμως τα χαρακτηριστικά της ακοής

Λειτουργία της ακοής (1 από 2)

- Η ακοή εξαρτάται από τη συχνότητα
 - Έστω μία συγκεκριμένη ένταση σήματος
 - Εκφράζεται σε dB: ένταση πίεσης ήχου (SPL)
 - Η αίσθηση έντασης δεν είναι σταθερή
 - phon: αντίληψη έντασης από άνθρωπο
 - 1 phon = 1 dB στο 1kHz
- Καμπύλες Fletcher-Munson
 - Κάθε καμπύλη έχει σταθερό phon

Λειτουργία της ακοής (2 από 2)



- Δυναμικό εύρος της ακοής
 - Μεγιστοποιείται στην περιοχή 2-5 kHz

Ψυχοακουστικό μοντέλο

- Αντιληπτική κωδικοποίηση
 - Χρήση ψυχοακουστικού μοντέλου της ακοής
 - Εντοπισμός μη αντιληπτών τμημάτων του ήχου
 - Κωδικοποίησή τους με μικρότερη ακρίβεια
- Δύο βασικές τεχνικές
 - Απόκρυψη συχνοτήτων
 - Απόκρυψη χρόνου

Απόκρυψη ήχων (1 από 3)

- Απόκρυψη συχνοτήτων
 - Ισχυρά σήματα: μειώνουν το δυναμικό εύρος
 - Στις γειτονικές συχνότητες ανεβαίνει το κάτω όριο
 - Ιδιαίτερα προς τις υψηλότερες συχνότητες
 - Η επίδραση εξαρτάται από τη συχνότητα
 - Διάκριση του φάσματος σε κρίσιμες περιοχές
 - Το αυτί δεν διακρίνει συχνότητες μέσα σε περιοχές
 - Το εύρος της περιοχής μεγαλώνει με τη συχνότητα

Απόκρυψη ήχων (2 από 3)

- Απόκρυψη χρόνου
 - Ισχυρά σήματα: αποκρύπτουν ασθενή για λίγο
 - Στις γειτονικές συχνότητες ανεβαίνει το κάτω όριο
 - Εξασθένηση φαινομένου με το χρόνο
 - Η επίδραση εξαρτάται και από τη συχνότητα
 - Καμπύλη απόκρυψης δύο διαστάσεων
 - Απόκρυψη στο χρόνο
 - Απόκρυψη στις συχνότητες

Απόκρυψη ήχων (3 από 3)

- Αξιοποίηση των αποκρύψεων
 - Ομαδοποιούμε μία σειρά δειγμάτων (πλαίσιο)
 - Αναλύουμε το σήμα σε περιοχές συχνοτήτων
 - Σε κάθε περιοχή εντοπίζουμε τα ισχυρά σήματα
 - Υπολογίζουμε επίδραση στα ασθενή σήματα
 - Σε κάθε περιοχή έχουμε διαφορετική επίδραση
 - Απομονώνουμε τα ασθενή σήματα
 - Κωδικοποίηση με μικρότερη ακρίβεια

**ΟΙΚΟΝΟΜΙΚΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΑΘΗΝΩΝ**



**ATHENS UNIVERSITY
OF ECONOMICS
AND BUSINESS**

Κωδικοποίηση ήχου MPEG-1

Μάθημα: Τεχνολογία Πολυμέσων, **Ενότητα # 10:** Κωδικοποίηση ήχου
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

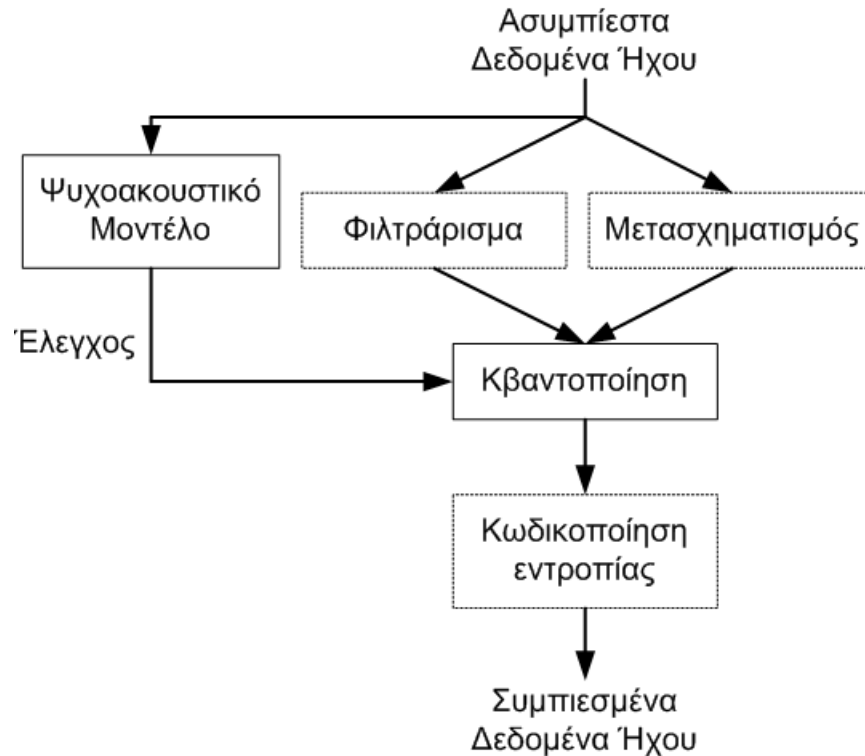


ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

Κωδικοποίηση MPEG-1 (1 από 7)

- MPEG Audio Layer 1, 2, 3
 - Τυποποιήθηκε στα πλαίσια του MPEG-1
 - Μέρος γενικότερης ομάδας προτύπων
- Τρία επίπεδα, προς τα πίσω συμβατότητα
 - Πιο δημοφιλές το επίπεδο 3 (MP3)
 - Αύξηση πολυπλοκότητας και καθυστέρησης
- Ψηφιοποίηση σήματος
 - 48, 44,1 ή 32 KHz με δείγματα 16 bit ανά κανάλι

Κωδικοποίηση MPEG-1 (2 από 7)



- Συμπίεση με βάση ψυχοακουστικό μοντέλο
 - Ελέγχει την κωδικοποίηση με βάση την αντίληψη

Κωδικοποίηση MPEG-1 (3 από 7)

- Βασική κωδικοποίηση (Layer 1)
 - Επεξεργασία ομάδων 384 υπο-δειγμάτων
 - 32 υπο-δείγματα ανά πακέτο
 - 12 πακέτα δειγμάτων ανά ομάδα
 - Διάκριση σε 32 ζώνες με φίλτρα
 - Ζώνες ίσου πλάτους με επικάλυψη
 - Δεν μεγαλώνουν όπως οι κρίσιμες περιοχές
 - Εξετάζονται 12 δείγματα ανά ζώνη

Κωδικοποίηση MPEG-1 (4 από 7)

- Ψυχοακουστικό μοντέλο
 - Εντοπισμός μέγιστης τιμής σε κάθε ζώνη
 - Εκτίμηση σημασίας κάθε ζώνης/συντελεστή
 - Μετασχηματισμός Fourier 1024 σημείων
 - Κατανομή bit ανάλογα με τη σημασία
- Κβαντοποίηση
 - Γραμμική κβαντοποίηση των συντελεστών
 - Χρήση παράγοντα κλιμάκωσης κβαντοποίησης
 - Στόχος: έξοδος με σταθερό ρυθμό bit

Κωδικοποίηση MPEG-1 (5 από 7)

- Ενδιάμεση κωδικοποίηση (Layer 2)
 - Τρία τεμάχια κάθε φορά (1152 δείγματα)
 - Αύξηση καθυστέρησης αλλά και ακρίβειας
 - Αξιοποίηση και της απόκρυψης χρόνου
- Προηγμένη κωδικοποίηση (Layer 3)
 - Ζώνες άνισου μήκους (όπως οι κρίσιμες περιοχές)
 - Επιπλέον μετασχηματισμός MDCT
 - Μεγαλύτερη ακρίβεια στην απόκρυψη θορύβου
 - Μη γραμμική κβαντοποίηση

Κωδικοποίηση MPEG-1 (6 από 7)

- Κωδικοποίηση εντροπίας
 - MP1/2: απλή PCM
 - MP3: Huffman σε ζεύγη συντελεστών
 - Επιλογή πίνακα Huffman ανάλογα με την είσοδο
- Διπλός βρόχος προσαρμογής (MP3)
 - Εσωτερικός: ανάλογα με κωδικοποίηση εντροπίας
 - Ρυθμίζει βήμα κβαντοποίησης ώστε να πετύχουμε στόχο
 - Εξωτερικός: ανάλογα με θόρυβο σε κάθε ζώνη
 - Ρυθμίζει παράγοντες κβαντοποίησης ανά ζώνη

Κωδικοποίηση MPEG-1 (7 από 7)

- Τελική κωδικοποίηση
 - Επίπεδα 1 και 2: σταθερός ρυθμός bit
 - Επίπεδο 3: προαιρετικά μεταβλητός ρυθμός bit
 - Αλλαγή σε κάθε πλαίσιο (frame) ήχου
- Ρυθμός δεδομένων: τουλάχιστον 32 Kbps
 - Επίπεδο 1: Μέχρι 448 Kbps
 - Επίπεδο 2: Μέχρι 384 Kbps
 - Επίπεδο 3: Μέχρι 320 Kbps
 - Η ποιότητα ανά ρυθμό εξαρτάται από το επίπεδο

Στερεοφωνία (1 από 3)

- Στερεοφωνικός ήχος
 - Δύο κανάλια ήχου για μεγαλύτερη πιστότητα
 - Μικρόφωνα / ηχεία σε διαφορετικά σημεία
 - Ο άνθρωπος τα αντιλαμβάνεται με δύο τρόπους
 - Διαφορές στο χρονισμό των καναλιών
 - Διαφορές στην ένταση των καναλιών
- Κωδικοποίηση στέρεο στο MPEG-1
 - Ανεξάρτητη ή κοινή (joint)

Στερεοφωνία (2 από 3)

- Κωδικοποίηση έντασης (intensity)
 - Στις χαμηλότερες συχνότητες
 - Αντιλαμβανόμαστε κυρίως διαφορές στο χρονισμό
 - Στις ψηλότερες συχνότητες
 - Αντιλαμβανόμαστε κυρίως διαφορές στην ένταση
 - Συνένωση δεξιού/αριστερού καναλιού
 - Προσθήκη πληροφορίας για τις διαφορές εντάσεων

Στερεοφωνία (3 από 3)

- Κεντρική-πλευρική κωδικοποίηση (mid-side)
 - Το κεντρικό είναι το άθροισμα δεξιού/αριστερού
 - $M=L+R$
 - Το πλευρικό κανάλι είναι η διαφορά τους
 - $S=L-R$
 - Μπορεί να κωδικοποιηθεί με λιγότερα bit
 - Μη απωλεστικός μετασχηματισμός
 - $L=(M+S)/2, R=(M-S)/2$

Ροή δεδομένων MPEG-1 (1 από 2)

- Ροή δεδομένων MPEG-1
 - Αν είναι αρχείο, μπορεί να έχει ιδιαίτερη κεφαλίδα
 - Εξαρτάται από το μορφότυπο, όχι από το πρότυπο
 - Η ροή διακρίνεται σε πλαίσια (frames) ήχου
 - Διάρκεια 24 ms στα 48 KHz
 - Κάθε πλαίσιο έχει μία κεφαλίδα
 - Επιτρέπει την άμεση έναρξη της αποκωδικοποίησης
 - Λέξη χρονισμού: έλεγχος για περιοδική εμφάνιση
 - Μπορεί να εμφανίζεται και στα δεδομένα

Ροή δεδομένων MPEG-1 (2 από 2)

- Ροή δεδομένων MPEG-1
 - Ρυθμός bit: επιτρέπει αλλαγή ανά πλαίσιο
 - Δειγματοληψία: επιτρέπει αλλαγή ανά πλαίσιο
 - Επίπεδο: 1, 2, 3 ή παραλλαγές
 - Τρόπος κωδικοποίησης
 - Στέρεο, κοινό στέρεο, κ.λπ.
 - Bit προστασίας: χρησιμοποιούνται ελάχιστα

**ΟΙΚΟΝΟΜΙΚΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΑΘΗΝΩΝ**



**ATHENS UNIVERSITY
OF ECONOMICS
AND BUSINESS**

Κωδικοποίηση ήχου MPEG-2

Μάθημα: Τεχνολογία Πολυμέσων, **Ενότητα # 10:** Κωδικοποίηση ήχου
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

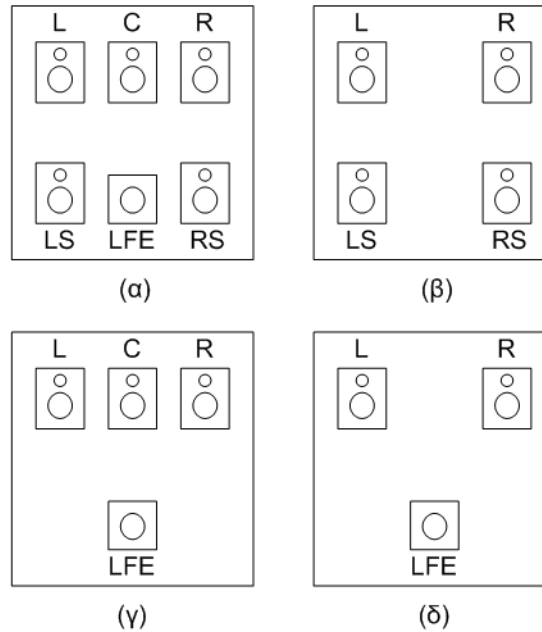


ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

Επεκτάσεις MPEG-2 (1 από 2)

- Πολυκαναλικός ήχος MPEG-2
 - Μέχρι πέντε κανάλια πλήρους εύρους
 - Κεντρικό, εμπρός, περιφερειακά
 - Κανάλι βελτίωσης μπάσων (LFE)
 - Συχνότητες 15-120 Hz
 - Επιτρέπονται διάφοροι συνδυασμοί
 - Πολύγλωσση ομιλία και σχολιασμός
 - Κινηματογραφικός ήχος 5.1

Επεκτάσεις MPEG-2 (2 από 2)



- Άλλες προσθήκες του MPEG-2
 - Ρυθμοί δειγματοληψίας 8-96 kHz
 - Αποδοτικό με 64 Kbps ανά κανάλι

Προβλήματα MPEG (1 από 2)

- Προβλήματα ψηφιακής κωδικοποίησης ήχου
 - Διαφέρουν από αυτά του αναλογικού ήχου
- Απώλεια ποιότητας
 - Σε συγκεκριμένες περιοχές συχνοτήτων
 - Όχι όπως η αναλογική αρμονική παραμόρφωση
 - Μπορεί να αλλάζει σε κάθε διάστημα κωδικοποίησης
 - Εξαφάνιση συχνοτήτων λόγω χαμηλού ρυθμού bit
 - Αναγκαστικός μηδενισμός ορισμένων συντελεστών

Προβλήματα MPEG (2 από 2)

- Προ-ηχώ (pre-echo)
 - Απότομη αλλαγή στο διάστημα κωδικοποίησης
 - Παράγει θόρυβο λόγω ανεπαρκούς ρυθμού bit
 - Απλώνεται σε όλο το διάστημα κωδικοποίησης
- Διπλή ομιλία (double speak)
 - Διαφορά περιόδου σήματος και κωδικοποίησης
 - Η ομιλία παρουσιάζει περιοδικότητα
 - Αλλοιώνεται από την κωδικοποίηση

MPEG-2 AAC (1 από 4)

- MPEG-2 Advanced Audio Coding (AAC)
 - Νέος κωδικοποιητής για το MPEG-2
 - Πιο αποδοτικός από το MPEG-1
 - Μείωση ρυθμού bit κατά 30% για ίδια ποιότητα
 - Δεν διατηρεί προς τα πίσω συμβατότητα
 - Ίδια βασική δομή αλλά με πολλές βελτιώσεις
 - Βασικός κωδικοποιητής ήχου και για το MPEG-4

MPEG-2 AAC (2 από 4)

- Βελτιώσεις κωδικοποίησης
 - 256 ή 2048 δείγματα
 - Χρησιμοποιεί μόνο MDCT, όχι φίλτρα
 - 256 δείγματα: μικρή καθυστέρηση
 - 2048 δείγματα: μικρότερος ρυθμός μετάδοσης
 - Οργάνωση συντελεστών σε 49 ζώνες
 - Παρόμοιες με τις κρίσιμες περιοχές
 - Πρόβλεψη συντελεστών σε κάθε συχνότητα

MPEG-2 AAC (3 από 4)

- Βελτιώσεις κωδικοποίησης
 - Βελτιωμένη κοινή στερεοφωνική κωδικοποίηση
 - Huffman σε τετράδες συντελεστών
- Βελτιώσεις ποιότητας
 - Μειωμένη προ-ηχώ (pre-echo)
 - Οφείλεται στη μικρότερη περίοδο κωδικοποίησης
 - Χρονική μορφοποίηση θορύβου (TNS)
 - Αποφυγή του φαινομένου της διπλής ομιλίας

MPEG-2 AAC (4 από 4)

- Ροή δεδομένων AAC
 - Audio Data Interchange Format (ADIF)
 - Όλες οι πληροφορίες είναι σε μία κεφαλίδα
 - Αποκωδικοποίηση μόνο από αρχή του αρχείου
 - Audio Data Transport Stream (ADTS)
 - Κεφαλίδες ανά πλαίσιο ήχου
 - Παρόμοιο με το MPEG-1
 - Επιτρέπει όμως και πλαίσια μεταβλητού μεγέθους
 - Επίπεδο 4 στην κεφαλίδα του πλαισίου (MP4!)

**ΟΙΚΟΝΟΜΙΚΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΑΘΗΝΩΝ**



**ATHENS UNIVERSITY
OF ECONOMICS
AND BUSINESS**

Κωδικοποίηση ήχου MPEG-4

Μάθημα: Τεχνολογία Πολυμέσων, **Ενότητα # 10:** Κωδικοποίηση ήχου
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

Εργαλεία ήχου MPEG-4 (1 από 2)

- Μέθοδοι κωδικοποίησης φυσικού ήχου
 - 2-6 Kbps: κωδικοποίηση LPC
 - 6-24 Kbps: κωδικοποίηση CELP
 - 24-64 Kbps: κωδικοποίηση AAC
- Κωδικοποιητής κειμένου σε φωνή (TTS)
 - Ρυθμοί δεδομένων 200 bps έως 1,2 Kbps
 - Απλό κείμενο ή κείμενο με προσωδία

Εργαλεία ήχου MPEG-4 (2 από 2)

- Μουσική σύνθεση με βάση παρτιτούρα
 - Ρυθμός μετάδοσης 2-3 Kbps
 - Ορχήστρα που αποτελείται από όργανα
 - Οδηγίες προς τα όργανα αυτά
 - Τράπεζα ήχων και σύνολο ηχητικών εφφέ
- Επεκτάσεις του MPEG-2 AAC
 - HD AAC: μη απωλεστική κωδικοποίηση
 - HE AAC: χαμηλότερος ρυθμός bit

**ΟΙΚΟΝΟΜΙΚΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΑΘΗΝΩΝ**



**ATHENS UNIVERSITY
OF ECONOMICS
AND BUSINESS**

Τέλος Ενότητας #10

Μάθημα: Τεχνολογία Πολυμέσων, **Ενότητα # 10:** Κωδικοποίηση ήχου
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

