

ΗΧΟΣ ΚΑΙ ΕΦΑΡΜΟΓΕΣ

- Ήχος σε δεύτερο πλάνο σε εφαρμογές πολυμέσων (εστίαση σε βίντεο)
- Επικέντρωση σε ψυχαγωγικές ή εκπαιδευτικές εφαρμογές (π.χ. Information kiosks)
- Αφήγηση/σχολιασμός βοηθούν στη μετάδοση μηνύματος
- Γίνεται αντιληπτός χωρίς εστίαση προσοχής
- Εφαρμογές για μουσική ή για άτομα με προβλήματα όρασης
- Αναγνώριση και σύνθεση ομιλίας

ΜΟΥΣΙΚΗ ΚΑΙ ΥΠΟΛΟΓΙΣΤΕΣ

- Ήχος αποθηκεύεται ως ψηφιακό σήμα (συμπιεσμένο ή ασυμπίεστο)
- Αναλογία με διανυσματικά γραφικά: πρότυπο MIDI (Musical Instrument Digital Interface)
- Καθορίζει την κωδικοποίηση των τόνων και των οργάνων
- Δυνατότητα χρήσης 127 οργάνων και εφέ
- Καθορίζει πρότυπα επικοινωνίας οργάνων (π.χ. synthesizers) και υπολογιστή
- Συνήθως η σύνθεση ήχων γίνεται με FM σύνθεση (άλλοτε αποθηκεύονται δείγματα πραγματικών οργάνων σε ROM)

ΜΟΥΣΙΚΗ ΚΑΙ ΥΠΟΛΟΓΙΣΤΕΣ

- MIDI σε αναλογία με vector graphics:
 - Μεγαλύτερη ευελιξία στην επεξεργασία
 - Μικρότερος αποθηκευτικός χώρος
 - Υπολογιστικό κόστος για μετατροπή σε ακουστικό σήμα
 - Αποτέλεσμα εξαρτάται από συσκευή εξόδου

ΓΕΝΙΚΑ ΓΙΑ ΤΟ MIDI

- Το MIDI εμφανίστηκε το 1983 (συνεργασία Roland, Yamaha, Korg κλπ.)
- Επιτρέπει αλληλεπίδραση και συγχρονισμό ηλεκτρονικών συσκευών (υπολογιστές, συνθετητές, ελεγκτές συστημάτων φωτισμού, πολυκαναλικά μαγνητόφωνα, VCR)
- Επιτρέπει έλεγχο όλων των συσκευών από μια (master control device)
- Αναλογία με σύνδεση υπολογιστών μέσω modem

Η ΜΕΘΟΔΟΣ ΤΟΥ MIDI

- Το MIDI περιέχει πληροφορία για:
 - Ποια νότα παίχτηκε (πλήκτρο πιέστηκε)
 - Πότε παίχτηκε η νότα
 - Πόσο κράτησε
 - Με τι πίεση πατήθηκε
 - Τι ένταση είχε η νότα
 - Τι διαμόρφωση είχε η νότα
 - Τι όργανο παίχτηκε
- Οι εντολές MIDI αποτελούνται από ένα σύνολο bytes (το πρώτο είναι το status)

GENERAL MIDI

- **General MIDI:** πρότυπο που υιοθετήθηκε σχετικά πρόσφατα από κατασκευαστές
 - *Πρόβλημα:* ένα MIDI αρχείο ενός συνθέτη που μεταφέρεται σε άλλη συσκευή δεν αντιστοιχεί τα ίδια όργανα στο ίδιο patch (αριθμό οργάνου)
 - *Λύση:* το General MIDI καθορίζει κανόνες αντιστοίχισης (π.χ. Patch 0=Grand Piano, Patch 40=Violin)
 - Το ίδιο συμβαίνει πλέον με τις κάρτες ήχου

MIDI CHANNELS

- Channels (κανάλια):
 - Το General MIDI καθορίζει 16 κανάλια
 - Διαφορετικά όργανα αντιστοιχίζονται σε διαφορετικά κανάλια
 - Η πολυφωνία (πόσες νότες μπορούν να ακούγονται ταυτόχρονα) ποικίλει σε κάρτες ήχου και συνθετητές

MIDI TRACKS

- **Tracks:**
 - Το πλήθος των tracks εξαρτάται από το λογισμικό (συνήθως πάνω από 100)
 - Κάθε track περιέχει ένα σύνολο από events
 - Είναι δυνατόν δύο tracks με διαφορετικό περιεχόμενο να αντιστοιχηθούν στο ίδιο κανάλι (π.χ. αριστερό και δεξί χέρι στο πιάνο)

ΟΜΙΛΙΑ ΚΑΙ ΥΠΟΛΟΓΙΣΤΕΣ

- Ορισμένες εφαρμογές: ομιλία σημαντικότερη
- Λογισμικό, κάρτες: Apple, Microsoft, Creative Labs
- Microsoft Windows Sound System: κάρτα ήχου 16 bit, μικρόφωνο, ακουστικά, λογισμικό
- Λογισμικό για αναγνώριση και σύνθεση φωνής, εισαγωγή ηχητικών αντικειμένων σε εφαρμογές
- IBM Personal Dictation System (OS/2): 70-100 λέξεις/λεπτό (χειρισμός συστήματος)
- Αναγνώριση ομιλίας: 30000 λέξεις σε εκπαιδευμένο σύστημα, 400 λέξεις (συνεχής λόγος, άγνωστος ομιλητής)

ΣΥΜΠΙΕΣΗ ΗΧΟΥ

- Ήχος δυσκολότερος από εικόνα: ακοή πιο ευαίσθητη
- Ανάπτυξη πολλών τεχνικών (ενδιαφέρον για τηλεφωνία)
- Στη μουσική πιο δύσκολα (μεγάλο εύρος ζώνης)
- Επιλογή κατάλληλης συχνότητας δειγματοληψίας και μήκους λέξης
- Περαιτέρω συμπίεση με αφαίρεση τμημάτων σιωπής
- Χρήση μη-γραμμικής PCM (λογαριθμική ή μ -law, διαφορική PCM, προσαρμοστική διαφορική PCM)

ΣΥΜΠΙΕΣΗ ΗΧΟΥ

Συχνότητα Δειγματοληψίας (kHz)	Κβαντοποίηση (bits)	Τεχνική Κωδικοποίησης	Ποιότητα
44.1	16	PCM	Hi-Fi
37.8	8	ADPCM	FM μετάδοση (μουσική)
18.9	8	ADPCM	AM μετάδοση (ομιλία)
8	8	PCM	Τηλεφωνική

ΟΠΤΙΚΟΑΚΟΥΣΤΙΚΗ ΤΗΛΕΦΩΝΙΑ

- Οικογένεια προτύπων H.320 (ITU-TS) για teleconferencing και videophony
- Υποστηρίζει τηλεπικοινωνιακά δίκτυα αλλά και υπολογιστών
- Περιλαμβάνει πρότυπα ITU-G για συμπίεση ήχου

Πρότυπο	Περιγραφή
G.711	PCM συχνοτήτων φωνής
G.722	ADPCM με συχνότητα 7kHz και απαιτούμενο εύρος ζώνης 64Kbits/s
G.728	CELP κωδικοποίηση που απαιτεί 16Kbits/s

MP3 - **ΙΣΤΟΡΙΑ**

- 1987, Fraunhofer Institut και Erlangen University: project σε perceptual audio coding
- Προέκυψε το ISO-MPEG Audio Layer-3
- Χωρίς μείωση δεδομένων, ένα sec μουσικής stereo σε ποιότητα CD απαιτεί 1.4 Mbit
- MPEG κωδικοποίηση: συμπίεση με παράγοντα 12 χωρίς απώλειες ποιότητας
- Παράγοντας 24 διατηρεί ικανοποιητική ποιότητα
- Perceptual coding: αντίληψη των ηχητικών κυμάτων από το ανθρώπινο αυτί

MP3 - **ΙΣΤΟΡΙΑ**

1:4	με Layer 1 (αντιστοιχεί σε 384 kbps για ένα στερεοφωνικό σήμα)
1:6...1:8	με Layer 2 (αντιστοιχεί σε 256...192 kbps για ένα στερεοφωνικό σήμα)
1:10...1:12	με Layer 3 (αντιστοιχεί σε 128...112 kbps για ένα στερεοφωνικό σήμα)

MP3 - **ΙΣΤΟΡΙΑ**

- Εκμετάλλευση εφέ στερεοφωνίας και περιορισμός ηχητικού εύρους ζώνης
- MPEG Layer-3 είναι το πιο ισχυρό
- Για δοσμένο επίπεδο ποιότητας ήχου απαιτεί το χαμηλότερο ρυθμό μετάδοσης, ή
- για δοσμένο ρυθμό μετάδοσης πετυχαίνει την καλύτερη ποιότητα ήχου

MP3 – ΠΟΙΟΤΗΤΑ ΗΧΟΥ

Ποιότητα ήχου	Εύρος ζώνης	Τρόπος Λειτουργίας	Ρυθμός μετάδοσης	Αναλογία μείωσης
Ήχος Τηλεφώνου	2.5 kHz	mono	8 kbps	96:1
Καλύτερα από βραχεία	4.5 kHz	mono	16 kbps	48:1
Καλύτερα από ραδιόφωνο AM	7.5 kHz	mono	32 kbps	24:1
Παρόμοια με ραδιόφωνο FM	11 kHz	stereo	56...64 kbps	26...24:1
Σχεδόν-CD	15 kHz	stereo	96 kbps	16:1
CD	>15 kHz	stereo	112...128 kbps	14...12:1

MP3 - ΠΟΙΟΤΗΤΑ ΗΧΟΥ

- Σε όλα τα διεθνή τεστ ακρόασης, το MP3 έδειξε ανώτερη απόδοση
- Διατηρεί την αρχική ποιότητα ήχου με μείωση δεδομένων 1:12 (περίπου 64 kbps ανά κανάλι ήχου)
- Αν εύρος ζώνης εφαρμογής ≈ 10 kHz \rightarrow ικανοποιητική ποιότητα με 1:24 μείωση
- ITU-R συνιστά MP3 για εφαρμογές ευρείας μετάδοσης στα 60 kbit/s ανά κανάλι

MP3 SURROUND

- Το νέο format MP3 Surround προσφέρει πολυκαναλικό ήχο (MPEG-2)
- Είναι πλήρως συμβατό προς τα πίσω με το stereo MP3
- Υποστηρίζει κάθε διάταξη surround (όπως το 5.1 surround)
- Ο ρυθμός δεδομένων που απαιτεί συγκρίνεται με του stereo MP3
- Χρήσεις: ραδιοφωνικοί σταθμοί στο internet, music download shops

MP3 - ΛΕΠΤΟΜΕΡΕΙΕΣ

- **Filter Bank** (συστοιχία φίλτρων):
 - υβριδική (πολυφασική συστοιχία φίλτρων και MDCT)
- **Perceptual Model** (μοντέλο αντίληψης):
 - βασικά καθορίζει ποιότητα υλοποίησης κωδικοποιητή
 - Χρησιμοποιεί χωριστή συστοιχία φίλτρων ή
 - Συνδυάζει υπολογισμό τιμών ενέργειας με την κυρίως συστοιχία

MP3 - ΛΕΠΤΟΜΕΡΕΙΕΣ

- **Joint Stereo** (ενιαία στερεοφωνία):
 - Εκμεταλλεύεται γεγονός ότι στέρεο κανάλια έχουν κατά πολύ την ίδια πληροφορία
 - Μειώνεται το bitrate βάσει των πλεονασμών και ελλείψεων σχέσης των καναλιών
 - Χρησιμοποιείται όπου διατίθεται χαμηλό bitrate αλλά απαιτείται στερεοφωνικό σήμα

MP3 - ΛΕΠΤΟΜΕΡΕΙΕΣ

- **Κβαντισμός και κωδικοποίηση:**
 - Κβαντισμός με κβαντιστή νόμου ισχύος (μεγάλες τιμές κωδικοποιούνται με μικρότερη ακρίβεια)
 - Κωδικοποίηση κβαντισμένων τιμών κατά Huffman (lossless, noiseless)
 - Διαδικασία εύρεσης μέγιστης ενίσχυσης και παράγοντα κλίμακας για δοσμένο μπλοκ, bitrate και έξοδο perceptual model:
 - με δύο εμφωλιασμένους βρόχους επαναλήψεων

MP3 - ΛΕΠΤΟΜΕΡΕΙΕΣ

- **Εσωτερικός βρόχος (βρόχος ρυθμού):**
 - Τα codebooks του Huffman αναθέτουν μικρότερους κωδικούς σε συχνότερες τιμές
 - Αν πλήθος bits μετά από coding > πλήθος bits για coding ενός μπλοκ → ρύθμιση ολικού κέρδους
 - Μεγαλώνω το βήμα κβαντισμού (άρα μικρότερες κβαντισμένες τιμές)
 - Επαναλαμβάνω μέχρι η απαίτηση bits για Huffman να είναι αρκετά μικρή

MP3 - ΛΕΠΤΟΜΕΡΕΙΕΣ

- **Εξωτερικός βρόχος (βρόχος ελέγχου θορύβου/παραμόρφωσης):**
 - Παράγοντες κλίμακας εφαρμόζονται σε κάθε μπάντα παραγόντων κλίμακας
 - Αν ο θόρυβος κβαντισμού υπερβαίνει κατώφλι συγκάλυψης (επιτρεπόμενο θόρυβο) ρυθμίζω scalefactor
 - Για κάθε νέο scalefactor τρέχει το εσωτερικό loop (γιατί για μικρότερο θόρυβο θέλω περισσότερα βήματα κβαντισμού → μεγαλύτερο bitrate)
 - Επαναλαμβάνω μέχρι ο τρέχων θόρυβος να είναι κάτω από το κατώφλι για κάθε μπάντα

MP3 - ΛΕΠΤΟΜΕΡΕΙΕΣ

- **Επέκταση Surround:**

- Πολλά κανάλια ήχου στην είσοδο συνδυάζονται σε ένα στερεοφωνικό σήμα εξόδου με μια διαδικασία μίξης
- Παράλληλα, εξέχοντα διακαναλικά cues εξάγονται από κανάλια εισόδου και κωδικοποιούνται συμπαγώς ως παράπλευρη πληροφορία surround
- Το ολικό σήμα είναι ένα stereo MP3 που κωδικοποιείται και μεταδίδεται μαζί με την παράπλευρη πληροφορία στο δέκτη

MP3 - ΛΕΠΤΟΜΕΡΕΙΕΣ

- **Επέκταση Surround (συνέχεια):**
 - Στο δέκτη αποκωδικοποιούνται τα MP3 δεδομένα
 - Δημιουργείται ένα πολυκαναλικό σήμα εξόδου από το ολικό σήμα και τα cues ανασυνθέτοντας τα κανάλια του
 - Τα διακαναλικά cues είναι η διακαναλική διαφορά χρόνου, η διακαναλική διαφορά επιπέδου, η διακαναλική συνέπεια κλπ.

AC3 (Audio Code 3)

- Αναπτύχθηκε στα εργαστήρια της Dolby
- Χρησιμοποιείται στην κινηματογραφική βιομηχανία
- Υιοθετήθηκε και στα DVD και στην HDTV στην Αμερική (η Ευρώπη υιοθέτησε το MPEG-2)
- Συμπυκνώνει 6 κανάλια σε χώρο λιγότερο από ένα μόνο κανάλι σε CD
- Δημιουργείται αληθινή αίσθηση χώρου

AC3 (Audio Code 3)

- 6 κανάλια: δεξί, αριστερό, κεντρικό, δύο surround και ένα για χαμηλές συχνότητες (low frequency effects channel)
- Χρήση subwoofer για καλύτερη ακρόαση του τελευταίου καναλιού
- 5 κανάλια πλήρους εύρους (3Hz-20KHz)
- Έκτο κανάλι συχνότητες 3Hz-120Hz
- Υποστηρίζονται sampling rates 32, 44.1, 48 KHz

AC3 (Audio Code 3)

- Επιτρέπεται επικάλυψη ήχου σε κανάλι με θόρυβο, από ήχο άλλου καναλιού
- Ρυθμοί δεδομένων: 32-640 kbps/channel
- Χρησιμοποιείται από το Real Audio (Real Media)

AAC (Advanced Audio Codec)

- Δίνει υψηλής ποιότητας ήχο στα 64 kbps ανά κανάλι
- Κωδικοποιεί έως και 48 κανάλια ήχου και έως 16 κανάλια χαμηλών συχνοτήτων
- Υποστηρίζει πολλές γλώσσες παράλληλα
- Υποστηρίζει φωνή πάνω από προγράμματα ήχου
- Έχει τρεις τρόπους λειτουργίας: κύρια, χαμηλής πολυπλοκότητας και κλιμακούμενης συχνότητας δειγματοληψίας

AAC (Advanced Audio Codec)

- **Κύρια:** σε εφαρμογές χωρίς περιορισμούς σε υπολογιστική ισχύ και μνήμη
- **Χαμηλής πολυπλοκότητας:** εφαρμογές όπου υπάρχουν περιορισμοί ισχύος και μνήμης
- **Κλιμακούμενης συχνότητας δειγματοληψίας:** για εφαρμογές με ελάχιστες απαιτήσεις σε ισχύ και μνήμη
- Δίνει καλύτερη ποιότητα ήχου από κάθε άλλο coding με το μισό μόλις bitrate
- Είναι κομμάτι του MPEG-4