

Περί ήχου και βίντεο

Ο ήχος

Ο ήχος είναι μία μορφή ενέργειας. Προκαλείται από μια κρούση, μια νύξη, ή μια τριβή σε ένα σώμα (ονομάζεται πηγή), το οποίο ταλαντώνεται και οι παλμικές δονήσεις του δημιουργούν ηχητικά κύματα, που μεταφέρουν ενέργεια και μεταδίδονται σφαιρικά προς όλες τις κατευθύνσεις. Όταν φτάσουν και προσκρούσουν στη μεμβράνη του τυμπάνου του αυτιού προκαλούν ταλάντωση του και θέτουν σε ενέργεια τον μηχανισμό της ακοής. Από άποψη φυσικής, είναι ένα κύμα. Χρειάζεται κάποιο μέσο (στερεό, υγρό ή αέριο) για να μεταδοθεί (δεν μεταδίδεται στο κενό). Τα χαρακτηριστικά του ήχου είναι:

- Η **συχνότητα**, εκφράζει την ταχύτητα ταλάντωσης και μετράται σε κύκλους ανά δευτερόλεπτο (Hertz, Hz) και ορίζεται ως ο αριθμός των κύκλων (κυμάτων) που παράγει η πηγή του ήχου στο δευτερόλεπτο.
- Η **ένταση**, ορίζεται η ισχύς του ηχητικού κύματος ανά μονάδα επιφανείας (πόσο ισχυρή ή ασθενής είναι η ταλάντωση ενός σώματος). Μετριέται σε decibel (db) και καθορίζει πόσο δυνατά ή χαμηλά ακούγεται ένας ήχος. Εξαρτάται από τη δύναμη που παράγεται ο ήχος και από την απόσταση που έχουμε από την πηγή. Ένταση ήχου 120 db είναι το **όριο πόνου** στο ανθρώπινο αυτί.
- Η **διάρκεια**, ορίζει τον συνολικό χρόνο για τον οποίο ένας ήχος γίνεται αντιληπτός.

Εκτός από τα παραπάνω χαρακτηριστικά, υπάρχουν και ορισμένα χαρακτηριστικά του ήχου που έχουν να κάνουν με υποκειμενικά κριτήρια (**αντίληψη του ήχου**). Η **αντίληψη του ήχου** πραγματοποιείται μέσω της ακοής. Ο άνθρωπος αντιλαμβάνεται ήχους με συχνότητα μεταξύ 20 Hz και 20.000 Hz. Υποκειμενικά κριτήρια είναι τα ακόλουθα:

Η **ακουστότητα** (loudness), η οποία είναι η υποκειμενική ένταση του ήχου, που εξαρτάται από την ακουστική ικανότητα του ακροατή, τη διάρκεια του ήχου και τη συχνότητα. Η μονάδα μέτρησης της είναι το Phon.

Το **ύψος** (pitch), το οποίο αφορά το πως αντιλαμβανόμαστε τη συχνότητα του ήχου. Ήχοι ίδιας συχνότητας που παράγονται από διαφορετικά υλικά, μπορούν να έχουν διαφορετικό ύψος, να ακούγονται δηλαδή διαφορετικά. Η μονάδα μέτρησης του ύψους είναι η Mel (από το Melody). Ανάλογα με το ύψος τους οι ήχοι διακρίνονται σε:

- **οξείς** (υψηλής συχνότητας), τα λεγόμενα πρίμα (tremble) όπως ο ήχος του φλάουτου,
- **μέσους** (μεσαίας συχνότητας) και
- **βαρείς** (μικρής συχνότητας), τα λεγόμενα μπάσα (bass), όπως ο ήχος του τύμπανου.

Η **χροιά** (timbre) ή αλλιώς **ηχόχρωμα**, η οποία μας επιτρέπει να ξεχωρίσουμε δύο ήχους με το ίδιο ύψος και ένταση. Αφορά σύνθετους ήχους που έχουν μεν την ίδια συχνότητα, αλλά έχουν διαφορετικές υποσυχνότητες. Οι υποσυχνότητες αυτές ονομάζονται **αρμονικές** και είναι πολλαπλάσιες μιας βασικής συχνότητας. Επειδή οι αρμονικές σε κάθε όργανο είναι διαφορετικές, έχει ως αποτέλεσμα η ίδια νότα να ακούγεται με διαφορετική χροιά σε διαφορετικά όργανα, αν και η βασική συχνότητα της νότας είναι ίδια.

Ο ψηφιακός ήχος

Ο φυσικός ήχος, όπως αναφέραμε και προηγουμένως, είναι μια κυματοφορφή (μια συνεχής αρμονική ημιτονοειδής καμπύλη) και μπορεί να καταγραφεί από συσκευές, όπως τα μικρόφωνα και να αποθηκευθεί σε κατάλληλα μέσα. Η χρησιμοποίησή τους σε εφαρμογές ηλεκτρονικών υπολογιστών προϋποθέτει τη μετατροπή του αναλογικού σήματος σε ψηφιακό, η οποία όπως και στην περίπτωση των εικόνων, χρειάζεται **δειγματοληψία** και **κβαντοποίηση**. Η χρήση του σε ιστοσελίδες και δικτυακές εφαρμογές, απαιτεί μείωση των αρχείων αποθήκευσης του, οπότε χρειάζονται **αλγόριθμοι κωδικοποίησης** και **συμπίεσης** τους.

Η **κωδικοποίηση ενός ψηφιακού ήχου** υλοποιείται με προγράμματα που είναι γνωστοί ως **codecs** (**κωδικοποιητές**), όπως για παράδειγμα το εργαλείο **MP3 Encoder**. Παίρνουν ως είσοδο ένα ψηφιακό ήχο και δημιουργούν ένα αντίγραφο του με κωδικοποίηση και συμπίεση του. Η αναπαραγωγή ενός αρχείου MP3 χρειάζεται αντίστοιχο αποκωδικοποιητή, γνωστό ως **mp3-player**.

Ψηφιακός ήχος μπορεί να παραχθεί και από κατάλληλες εφαρμογές ηλεκτρονικών υπολογιστών και αποθηκεύεται με το πρότυπο **Midi**

Ο **ρυθμός μετάδοσης** ενός ψηφιακού αρχείου στο διαδίκτυο (**bit rate**) εκφράζεται σε bits ανά sec (**bps**).

Μορφές αποθήκευσης αρχείων ήχου

Υπάρχουν διάφορες μορφές αποθήκευσης ήχου σε ψηφιακή μορφή, οι οποίες διακρίνονται σε κατηγορίες ανάλογα με το αν χρησιμοποιούν ή όχι συμπίεση και αν αυτή είναι απωλεστική ή μη.

Αυτές που δεν χρησιμοποιούν συμπίεση έχουν καλή απόδοση κατά την αναπαραγωγή των ψηφιακών αρχείων, καθόσον έχουν ελάχιστες απώλειες πληροφορίας. Δημιουργούνται όμως αρχεία με μεγάλο όγκο (περίπου 34 MB για κάθε ένα λεπτό στερεοφωνικού ήχου), που τα καθιστούν ακατάλληλα για ιστοσελίδες και εν γένει εφαρμογές διαδικτύου. Η αναλογική κυματομορφή για να μετατραπεί σε ψηφιακή δυαδική μορφή, χρειάζεται να γίνει **δειγματοληψία** και ο ήχος πρέπει να καταγραφεί σε συγκεκριμένα διαστήματα (**παλμοί**). Η μορφή του ψηφιακού αρχείου που δημιουργείται λέγεται PCM (Pulse Code Modulation) και είναι η πιο συνηθισμένη που χρησιμοποιείται σε CD και DVD.

Σε αυτή την κατηγορία ανήκουν οι ακόλουθες μορφές:

WAV (Waveform Audio), είναι ασυμπίεστα αρχεία με μεγάλο μέγεθος, που έχουν πολύ καλή ποιότητα ήχου και υποστηρίζονται από σχεδόν όλα τα προγράμματα επεξεργασίας ήχου. Αναπτύχθηκαν από MicroSoft και IBM. Χρησιμοποιούνται ευρέως σε εφαρμογές των Windows, ενώ στο διαδίκτυο σπάνια. Έχουν επέκταση **.wav**.

AIFF (Audio Interchange File Format), δημιουργήθηκαν από την Apple, διατηρούν εξαιρετική την ποιότητα του ήχου για αυτό χρησιμοποιούνται από επαγγελματικά προγράμματα επεξεργασίας ήχου. Μειονέκτημα τους το μεγάλο μέγεθος αρχείου. Έχουν επέκταση **.aif**. Είναι ασυμπίεστα αρχεία, ενώ υπάρχει και μια συμπίεσμένη μορφή η **aifc**.

AU, δημιουργήθηκαν από τη Sun Microsystems για το λειτουργικό σύστημα Unix. Τα αρχεία έχουν προέκταση **.au** ή **.snd**. Το μέγεθος τους είναι μικρό, αλλά η ποιότητα είναι χαμηλή. Χρησιμοποιείται σε ιστοσελίδες για σύντομα ηχητικά εφε, όπου η ποιότητα δεν είναι σημαντική.

FLAC (Free Lossless Audio Codec) επιτυγχάνουν συμπίεση ως 60% χρησιμοποιώντας μη απωλεστική συμπίεση, οπότε η αποσυμπίεση τους γίνεται χωρίς σημαντική αλλοίωση του αρχικού ήχου. Επίσης,

περιέχουν και επιπλέον μεταδεδομένα, όπως για παράδειγμα τους στίχους ενός τραγουδιού (ένα τραγούδι 3-4 λεπτά θα έχει μέγεθος γύρω στα 30MB. Είναι ανοιχτό και δωρεάν φορμάτ.

WMA (Windows Media Audio) δημιουργήθηκαν από τη Microsoft, χρησιμοποιώντας απωλεστική και μη απωλεστική συμπίεση, ως 60% και μπορούν να χρησιμοποιηθούν και για streaming audio. Επειδή μπορούν να έχουν copyright, προτιμούνται σε εμπορικές χρήσεις. Έχουν επέκταση .wma.

MP3 είναι το πιο δημοφιλές πρότυπο απωλεστικής συμπίεσης ηχητικών δεδομένων. Έχουν επέκταση .mp3. Χρησιμοποιείται σε εφαρμογές λογισμικού, μπορούν να χρησιμοποιηθούν και για streaming audio και αναπαράγονται με mp3 players. Η κωδικοποίηση τους δεν είναι δωρεάν. Προσφέρουν μεγάλη συμπίεση, αλλά χαμηλή ποιότητα. Τα αρχεία MP3 έχουν ρυθμούς bit μεταξύ 32kbps και 320kbps (Ένα αρχείο MP3 διάρκειας 3 λεπτών, πχ. όσο ένα τραγούδι και bit rate 32kbps έχει μέγεθος 720kb). Ποιότητα ήχου κοντά σε αυτή του CD χρειάζεται bit rate 192kbps. Το πιο συνηθισμένο bit rate αρχείου mp3 είναι 128kbps, που διασφαλίζει μικρό μέγεθος και αρκετά καλή ποιότητα ήχου.

MPEG AAC (Advanced Audio Coding) αποτελούν εξέλιξη του MP3, προσφέροντας καλύτερη ποιότητα ήχου. Χρησιμοποιείται απωλεστική συμπίεση. Έχει μέγεθος αρχείου μικρότερο από αυτό του mp3 με το ίδιο bit rate, ιδιαίτερα για τιμές κάτω από 100kbps. Είναι το φορμάτ ήχου για ταχύτερες μεταφορές δεδομένων και είναι το πλέον κατάλληλο για χρήση στο διαδίκτυο. Χρησιμοποιείται επίσης σε συσκευές όπως iPhone, iPod, Playstation και κινητά με το λειτουργικό Android. Έχουν επέκταση mp4.

Ogg Vorbis είναι μια ελεύθερη εκδοχή μορφής συμπιεσμένων αρχείων αντίστοιχων των MP3 και AAC, ευρέως διαδεδομένα στους χρήστες Linux. Χρησιμοποιείται απωλεστική συμπίεση με δωρεάν κωδικοποίηση, ενώ έχουν ποιότητα καλύτερη των MP3. Έχουν επέκταση .ogg.

RealAudio είναι ένα φορμάτ που πρωτοεμφανίστηκε το 1995, υποστήριζε streaming audio και έγινε γρήγορα αρκετά δημοφιλής. Οι κωδικοποιητές που χρησιμοποιούνται είναι απωλεστικοί και επιτυγχάνουν καλή συμπίεση, αλλά χαμηλή ποιότητα ήχου.

Λογισμικό και εργαλεία επεξεργασίας για ψηφιακό ήχο

Το λογισμικό ήχου που έχει αναπτυχθεί αφορά σε λειτουργίες, όπως: εγγραφή, αναπαραγωγή, μετατροπή μεταξύ διαφόρων μορφών ήχου, ραδιοφωνική μετάδοσης μέσω διαδικτύου και τέλος. Τα εργαλεία επεξεργασίας ήχου παρέχουν δυνατότητες όπως:

- ηχογράφηση και μίξη ήχου,
- απομάκρυνση θορύβου,
- αποκοπή ή εισαγωγή τμημάτων ήχου σε συγκεκριμένα σημεία,
- μετατροπή από μία μορφή σε άλλη,
- χρήση ειδικών εφέ.

Το πιο δημοφιλές δωρεάν λογισμικό είναι το **Audacity** (<http://audacity.sourceforge.net/>). Είναι διαθέσιμο για όλα τα διαδεδομένα λειτουργικά συστήματα, δηλαδή Windows, Mac OS και Linux.

Ψηφιακό βίντεο

Ένα κομμάτι βίντεο (**video clip**) αποτελείται από διαδοχικές εικόνες που προβάλλονται με τέτοια ταχύτητα ώστε τα αντικείμενα να εμφανίζονται ως κινούμενα. Οι διαδοχικές εικόνες, που έχουν μικρές διαφορές, ονομάζονται πλαίσια ή καρέ (**frames**). Ο αριθμός των πλαισίων που προβάλλονται ανά δευτερόλεπτο μετρίεται σε frames per second (**fps**). Είναι σημαντικός, γιατί καθορίζει την ομαλότητα της εικόνας και το μέγεθος του αρχείου. Η ομαλή εμφάνιση ενός βίντεο χρειάζεται 25 fps ή 30 fps. Προβάλλονται με κατάλληλο πρόγραμμα τον VLC player.

Τα βασικά χαρακτηριστικά του ψηφιακού βίντεο είναι τα ακόλουθα:

Η **ανάλυση ενός βίντεο (video resolution)**, η οποία είναι η διάσταση της εικόνας σε pixels. Στην αναλογική τηλεόραση ήταν 768×484 ή 768×576. Ανάλυση 1920×1080 είναι γνωστή ως Full High Definition ή **Full HD**. Συνήθεις αναλύσεις βίντεο που προβάλλονται μέσω του διαδικτύου, π.χ. στο YouTube, είναι οι 360p (480×360), 480p (640×480) και 720p (1280×720). Η ανάλυση **UHD** (Ultra High Definition) ή **4K**, αναφέρεται σε αναλύσεις υψηλότερες από Full HD. Ως UHD ανάλυση θεωρείται η 2160p (3840×2160). Ανάλυση 4K×2K (4096×2160 pixels), χρησιμοποιείται από βίντεο προβολείς και επαγγελματικές φωτογραφικές μηχανές. Το **8K UHD**, αφορά ανάλυση 7680×4320.

Ο **λόγος πλευρών (aspect ratio)**, της οριζόντιας διάστασης προς την κάθετη διάσταση. Στις τηλεοράσεις παλαιά ήταν 4:3 επειδή το τηλεοπτικό σήμα είχε λόγο 4:3 και γι' αυτό είχαν τετράγωνη εμφάνιση. Στο ψηφιακό βίντεο ο λόγος είναι 16:9 και η εμφάνιση είναι μακρόστενη.

Ο **ρυθμός μεταφοράς bit (bit rate)**, αφορά στα δεδομένα (bits) που περιέχονται σε ένα δευτερόλεπτο ενός βίντεο και εκφράζεται σε bit ανά δευτερόλεπτο (bits per second ή bit/s ή **bps**). Μεγαλύτερος ρυθμός μετάδοσης σημαίνει και καλύτερη ποιότητα βίντεο, γιατί μπορούν να μεταδοθούν περισσότερα δεδομένα, άρα βίντεο μεγαλύτερης ανάλυσης. Επειδή το μέγεθος ενός ψηφιακού βίντεο είναι αρκετά μεγάλο και απαγορευτικό για μετάδοση μέσω διαδικτύου, έχουν αναπτυχθεί διάφορες τεχνικές **συμπίεσης (compression)**. Οι αλγόριθμοι συμπίεσης έχουν διάφορα χαρακτηριστικά, όπως ο λόγος συμπίεσης, η τεχνική συμπίεσης και η πολυπλοκότητα. Το μέγεθος ενός ασυμπίεστου βίντεο εξαρτάται από την ανάλυση, το βάθος χρώματος, τον αριθμό των πλαισίων ανά δευτερόλεπτο και τη χρονική διάρκεια.

οριζόντια ανάλυση × κάθετη ανάλυση × βάθος χρώματος (bits) × ρυθμός πλαισίων × διάρκεια (sec)

Παράδειγμα 1 λεπτό (60 sec) έγχρωμου βίντεο με πραγματικό βάθος χρώματος 24 bit, ανάλυση 320x200 και 30 fps καταλαμβάνει αποθηκευτικό χώρο: $320 \times 200 \times 24 \times 30 \times 60 = 2.764.800.000$ bits ≈ 329 MB. Οπότε, το μέγεθος αποθήκευσης ασυμπίεστου βίντεο είναι απαγορευτικό. Γι' αυτό τον λόγο εφαρμόζονται διάφορες τεχνικές συμπίεσης, που μειώνουν αρκετά το μέγεθος του αρχείου. Η συμπίεση των δεδομένων εικόνας βίντεο είναι απαραίτητη, γιατί διαφορετικά η αποθήκευση και διανομή του βίντεο μέσω υπολογιστή θα ήταν ανέφικτη. Γνωστός απωλεστικός codec για βίντεο είναι ο MPEG-4, που μπορεί να πετύχει συμπίεση από 20 έως 200 φορές σε σχέση με το αρχικό μέγεθος.

Τα πιο γνωστά μορφότυπα αρχείων βίντεο.

AVI (Audio Video Interleave), έχουν προέκταση .avi και δημιουργήθηκαν από τη Microsoft αρχές της δεκαετίας του '90 για τα Windows. Πλεονέκτημα τους ότι μπορούν να αναπαραχθούν σχεδόν σε κάθε

λειτουργικό σύστημα και να χρησιμοποιηθούν σε εφαρμογές. Μειονέκτημα ότι το μέγεθος των αρχείων είναι μεγάλο.

QTFF (QuickTime File Format), έχουν προέκταση .qt ή .mov. και δημιουργήθηκαν από την Apple. Πρόκειται για streaming format, δηλαδή γίνεται σταδιακή μεταφορά ήχου στον υπολογιστή από το διαδίκτυο και αναπαραγωγή των τμημάτων που έρχονται. Η αναπαραγωγή τους γίνεται από όλους τους players και από το iTunes με πολύ καλή ποιότητα εικόνας.

MP4 είναι τύπος αρχείων στα οποία τα δεδομένα βίντεο συμπιέζονται με τον codec MPEG-4. Έχουν προέκταση .mp4. Χρησιμοποιείται ευρέως στο YouTube και σε άλλες πολυμεσικές διαδικτυακές εφαρμογές. Έχει δυνατότητα τμηματικής αναπαραγωγής και προτιμάται για streaming video. Αναπαράγεται σε υπολογιστές, κινητά τηλέφωνα, σύγχρονες τηλεοράσεις, έξυπνα ρολόγια κ.λπ.

Λογισμικό επεξεργασίας βίντεο

Το Avidemux είναι μια εφαρμογή δωρεάν και ανοιχτού κώδικα για την επεξεργασία βίντεο. Υποστηρίζει την εφαρμογή φίλτρων και ειδικών εφέ, την αλλαγή μεγέθους, την περιστροφή και τον υποτιτλισμό. Επιτρέπει την κωδικοποίηση με διαφορετικούς codecs και την εισαγωγή ήχου στο βίντεο ή την εξαγωγή του ήχου από αυτό.