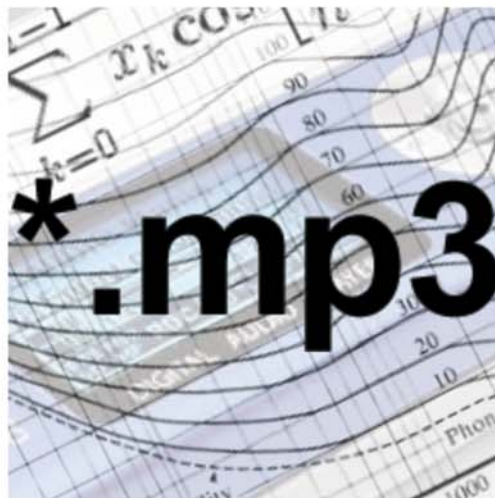


KnowHow Series

Τεχνολογίες
Συμπίεσης του ήχου:
Γνωρίζοντας το MP3
και άλλους Perceptual Codecs



Δημήτρης Σταματάκος, έκδοση 11/2005

ΜΕ ΤΗΝ ΥΠΟΣΤΗΡΙΞΗ:



COLOR SOUND MP3 PLAYER
iAUDIO

Μουσικός Πειρασμός!



Όλα τα iAudio ενσωματώνουν το σύστημα ψηφιακής επεξεργασίας ήχου της BBE.



Όλα τα iAudio υποστηρίζουν πολλαπλά format: MP3, WMA, ASF, WAV, OGG & FLAC (X5).



Όλα τα iAudio διαθέτουν ραδιόφωνο FM υψηλής ποιότητας με 24 προεπιλογές.



Όλα τα iAudio υποστηρίζουν εγγραφή φωνής, ραδιοφώνου & line με άμεση κωδικοποίηση σε MP3.



Όλα τα iAudio συνοδεύονται από ακουστικά υψηλής ποιότητας της Cresyn.



Όλα τα iAudio είναι πλήρως συμβατά με MAC & Linux OS.



**iAudio X5:
Το Απόλυτο
Gadget!**

MP3 player/encoder/Ραδιόφωνο FM. Αναπαραγωγή Video & φωτογραφιών (MPEG4, JPG). Έγχρωμη οθόνη LCD-TFT 160x128 pixels. Ισοσταθμιστής JetEffect. Δυνατότητα αναβάθμισης firmware. Μπαταρία λιθίου. Σύνδεση USB 2.0. Χωρητικότητα: 20MB / 30MB / 60MB



**iAudio 5:
Ήχος και Χρώμα!**

MP3 player/encoder/Ραδιόφωνο FM. Ισοσταθμιστής JetEffect. Led Backlight με 1000 χρώματα. Δυνατότητα αναβάθμισης firmware. 20 ώρες αναπαραγωγής με μία μόνο μπαταρία AAA. Σύνδεση USB 2.0. Χωρητικότητα: 256MB / 512MB / 1GB



**iAudio G3:
Η Μουσική Βίβλος!**

MP3 player/encoder/Ραδιόφωνο FM. Ισοσταθμιστής JetEffect. Δυνατότητα αναβάθμισης firmware. 50 ώρες αναπαραγωγής με μία μόνο μπαταρία AA. Σύνδεση USB 2.0. Χωρητικότητα: 256MB / 512MB / 1GB



**iAudio U2:
Ένα Μελωδικό
Κόσμημα!**

MP3 player/encoder/Ραδιόφωνο FM. Ισοσταθμιστής JetEffect. Δυνατότητα αναβάθμισης firmware. Μπαταρία λιθίου. Σύνδεση USB 2.0. Χωρητικότητα: 256MB / 512MB / 1GB

Εισαγωγή

Στις 14 Ιουλίου του 1995, δηλαδή πριν από δέκα χρόνια και κάτι μήνες, ένα από τα σημαντικότερα file extensions είδε το φως της ημέρας: Οι άνθρωποι του Fraunhofer IIS αποφάσισαν να δίνουν στα αρχεία που προέκυπταν από την κωδικοποίηση κατά MPEG-1 Layer III την κατάληξη .mp3. Η τεχνολογία που είχαν αναπτύξει έμελλε να αλλάξει τον τρόπο που ένα μεγάλο μέρος των ακροατών -και κυρίως οι νεώτεροι- βλέπουν την μουσική, να δημιουργήσει σημαντικούς κραδασμούς στο οικοδόμημα της δημιουργίας, της διακίνησης και της πώλησης δίσκων και να δώσει στους ηλεκτρονικούς υπολογιστές, τα δίκτυα και το διαδίκτυο έναν νέο λόγο ύπαρξης. Η διακίνηση της μουσικής μέσω συμπιεσμένων αρχείων από χρήστη σε χρήστη, η διαχείριση και η αναπαραγωγή τους από ένα νέο είδος προϊόντων που ονομάστηκαν portable media players και οι δυνατότητες που όλα αυτά προσέφεραν, είχαν ως βάση τους την δυνατότητα δημιουργίας αρχείων με μεγέθη αρκετά μικρά ώστε να μπορούν να αποθηκευτούν εύκολα, καθώς και της δημιουργίας συστημάτων που μπορούσαν να επεξεργαστούν τα αρχεία αυτά απαιτώντας την ελάχιστη δυνατή ροή δεδομένων. Μέσα σε ελάχιστο χρόνο, μία περιθωριακή τεχνολογία, αυτή της συμπίεσης των ηχητικών σημάτων, έγινε κυριολεκτικά κοινός τόπος και οι Perceptual codecs -ανεξαρτήτως με το ποιοί και για ποιούς λόγους διαφωνούν "επί της αρχής"- είναι ίσως η πλέον ενδιαφέρουσα και προκλητική ιδέα των τελευταίων δεκαετιών στον χώρο του audio...

Δημήτρης Σταματάκος, 11/2005
http://www.avmentor.gr/about/ds_bio.htm

Τεχνολογίες Συμπίεσης του Ήχου: Γνωρίζοντας το MP3 και άλλους Perceptual Codecs

Από την στιγμή που ένα ηχητικό σήμα ψηφιοποιείται, αποτελείται από μία σειρά ψηφιακών δεδομένων που το περιγράφουν πλήρως (ή έτσι ελπίζουμε...). Τα δεδομένα αυτά μπορούν να αποθηκευθούν σε διάφορα μέσα (CD, σκληρούς δίσκους, μνήμες, κ.λπ), να μεταδοθούν μέσα από ενσύρματες ή ασύρματες γραμμές μετάδοσης (ειδικές συνδέσεις μεταξύ συσκευών, τοπικά δίκτυα Ethernet, το διαδίκτυο) και να υποστούν επεξεργασία. Σε όλες αυτές τις εφαρμογές (οι οποίες συνιστούν το σημερινό περιβάλλον των πολυμέσων, μέσα στο οποίο ζούμε -τουλάχιστον σε ό,τι αφορά τον ήχο) αυτό που έχει μεγάλη σημασία είναι, κατ' αρχήν, ο όγκος της πληροφορίας ποσοτική έκφραση του οποίου είναι ο αριθμός των ψηφίων που απαιτείται για να περιγραφεί το σήμα και στην συνέχεια, ο ρυθμός ροής του σήματος που είναι αναγκαίος για την σωστή λειτουργία του συστήματος μεταφοράς και αποκωδικοποίησης.

Για παράδειγμα, τα σήματα που αποθηκεύονται στα CD, αποτελούνται από δύο κανάλια, κάθε ένα από τα οποία μεταφέρει 44.100 δείγματα το δευτερόλεπτο, με το κάθε δείγμα να απαιτεί 16 ψηφία για την περιγραφή του: $2 \times 44.100 \times 16 = 1.411.200 \text{ bits/s}$, ή, κατά προσέγγιση, 1.41Mbps

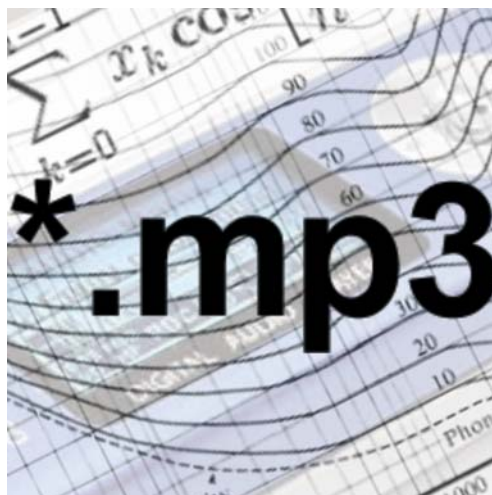
Αυτός είναι ο ρυθμός (bitrate) δεδομένων που θα πρέπει να υποστηρίζει κάθε ψηφιακό σύστημα που διαχειρίζεται σήματα κατάλληλα για CD. (την τυπική αυτή προδιαγραφή, την ονομάζουμε "CD Quality" και την συναντάμε αρκετά συχνά. Δεν πρέπει να συγχέεται με την έννοια "Ποιότητα αναπαραγωγής του CD" η οποία εξαρτάται βεβαίως από πολλούς άλλους παράγοντες. Προσέξτε επίσης ότι η μέτρηση του bitrate γίνεται πάντοτε σε bits και όχι σε bytes! 1byte=8bits). Στην πραγματικότητα, ο ρυθμός δεδομένων είναι πολύ μεγαλύτερος, επειδή μαζί με το σήμα κωδικοποιούνται και άλλες πληροφορίες. Αν θέλουμε να αποθηκεύσουμε ένα τέτοιο σήμα χρησιμοποιώντας την συμβατική τεχνολογία του προσωπικού υπολογιστή θα χρειαστούμε:

$2 \times 44.100 \times 2 (\text{bytes}) \times 60 (\text{δευτερόλεπτα}) = 10.584.000 \text{ bytes/min}$, ή, κατά προσέγγιση, 10.1MB χώρου για κάθε λεπτό σήματος. Ο χώρος, επομένως που απαιτείται για την αποθήκευση ψηφιακών ηχητικών δεδομένων είναι σημαντικός και σίγουρα

όχι εύκολο να εξασφαλισθεί στην πράξη, αν μιλούμε για τις ανάγκες μίας μικρής -έστω- δισκοθήκης ακόμη και με την σημερινή χαμηλή τιμή μνήμης και σκληρών δίσκων. Αλλά και η μεταφορά του σήματος δεν είναι εύκολη: Το μεγαλύτερο μέρος των συνδέσεων του διαδικτύου χρησιμοποιεί ταχύτητες αισθητά χαμηλότερες από 1Mbps και η μεταφορά τέτοιου όγκου δεδομένων απαιτεί μεγάλο χρόνο ή είναι αδύνατη αν απαιτείται streaming, δηλαδή ταυτόχρονη μεταφορά/χρήση.

Μη Απωλεστική - Απωλεστική κωδικοποίηση, Perceptual Codecs

Το τεχνικό πρόβλημα του χώρου και της ανεπάρκειας στην ταχύτητα μεταφοράς



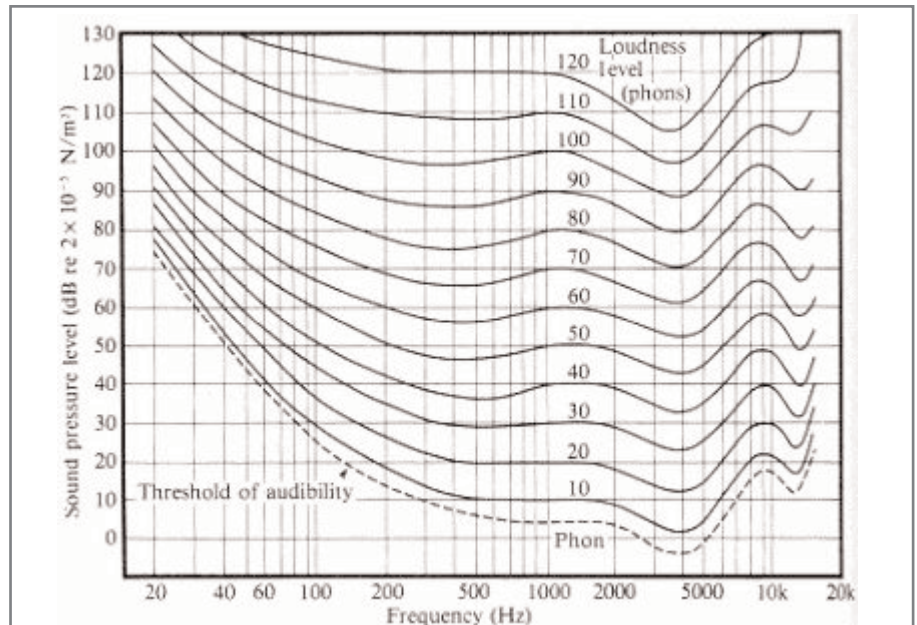
έρχονται να λύσουν οι τεχνολογίες συμπίεσης του ηχητικού σήματος. Οι τεχνολογίες αυτές χρησιμοποιούν διάφορες τεχνικές μείωσης του όγκου (bits ή bytes) και της ροής των δεδομένων (bitrate) που απαιτούνται για την κωδικοποίηση των ηχητικών σημάτων και χωρίζονται γενικώς σε δύο κατηγορίες: Τις μη-απωλεστικές (lossless) και τις απωλεστικές (lossy). Οι μη-απωλεστικές τεχνικές περιλαμβάνουν μεθόδους κωδικοποίησης οι οποίες μειώνουν μεν την ροή δεδομένων αλλά επιτυγχάνουν τον στόχο τους χωρίς να αλλοιώνουν τα δεδομένα αυτά. Αντίθετα, οι απωλεστικές τεχνικές αλλοιώνουν τα δεδομένα, απορρίπτοντας τα μέρη εκείνα τα οποία δεν είναι χρήσιμα -χρησιμοποιώντας

κάποιο συγκεκριμένο κριτήριο (irrelevance). Οι απωλεστικές τεχνικές είναι περισσότερο αποτελεσματικές, επιτυγχάνοντας συμπίεσεις που φθάνουν το 24:1 (μειώνουν, δηλαδή το bitrate στο 1/24, για παράδειγμα το 1.41Mbps γίνεται 60kbps) με καλή ποιότητα, ενώ οι μη απωλεστικές σπανίως ξεπερνούν το 4:1. Στην ορολογία του χώρου, το σύνολο των μηχανισμών που λαμβάνουν μέρος στην διαδικασία κωδικοποίησης-αποκωδικοποίησης ενός ψηφιακού σήματος, χρησιμοποιώντας έναν συγκεκριμένο αλγόριθμο (μια μαθηματική διαδικασία με σαφώς καθορισμένα βήματα, δηλαδή) ονομάζεται Codec (από την σύντμηση των λέξεων COder-DECoder). Σήμερα έχουμε στην διάθεσή μας μία ενδιαφέρουσα ποικιλία επιτυχημένων απωλεστικών codecs, με περισσότερο γνωστούς αυτούς της οικογένειας MP3 (MP3, MP3Plus, MP3 Surround), της οικογένειας AAC (MPEG-2 AAC, MPEG-4 AAC), τον ανοιχτού πηγαίου κώδικα και ελεύθερο δικαιωμάτων (open source/rights free) Ogg Vorbis, τον ιδιωτικό (proprietary) codec της Microsoft (WMA), ενώ συζητείται αρκετά και ο μη-απωλεστικός (και κατ'εξαιρέσει αναφερόμενος, εδώ λόγω της χρήσης του σε πολλά συστήματα) FLAC. Όλοι αυτοί οι codecs χρησιμοποιούν ως κριτήριο για την συμπίεση του σήματος την λειτουργία του ανθρώπινου συστήματος ακοής. Βασίζονται δηλαδή σε ένα ψυχοακουστικό μοντέλο αντίληψης των ήχων από τον άνθρωπο που τους επιτρέπει να αποφασίζουν "τί ακούγεται" και "τι δεν ακούγεται". Οι codecs που χρησιμοποιούν αυτή την προσέγγιση ονομάζονται Perceptual Codecs. Το ψυχοακουστικό μοντέλο αναπτύσσεται μεταξύ άλλων και από στατιστικές μετρήσεις οι οποίες οδηγούν στον προσδιορισμό μίας "μέσης συμπεριφοράς ακροατή". Για τον λόγο αυτό τα μοντέλα αυτά τα ονομάζουμε και "Υποκειμενικά Μοντέλα Ακοής".

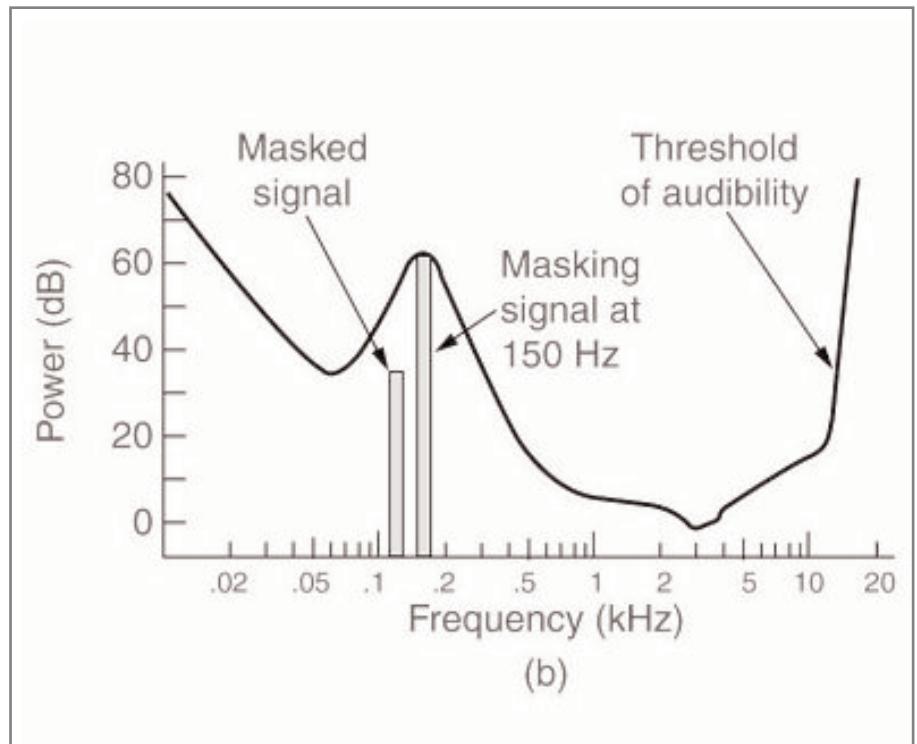
Οι πρώτες προσπάθειες. (PASC, ATRAC, MPEG-1 Layer II/III)

Η επίσημη ιστορία των Υποκειμενικών (Perceptual) κωδικοποιητών στην αγορά ξεκινά το 1993 όταν σχεδόν ταυτόχρονα οι Sony και Philips παρουσίασαν δύο προϊόντα

ψηφιακής εγγραφής (τον οπτικό δίσκο MiniDisc και την μαγνητική ταινία DCC) που χρησιμοποιούσαν απώλεστική συμπίεση. Οι PASC (Precision Adaptive Sub-band Coding) και ATRAC (Adaptive Transform Acoustic Coding) επιτύγχαναν συμπίεσεις μεταξύ του 4:1 και 5:1 και όπως είναι φυσικό είχαν μάλλον μέτρια απόδοση με τα σημερινά κριτήρια, ήταν ωστόσο εντυπωσιακοί, με βάση την εποχή που παρουσιάστηκαν, έχοντας ως βασικό στόχο σε εμπορικό επίπεδο την αντικατάσταση της αναλογικής κασέτας. Από τυπικής πλευράς, ανήκαν στην προδιαγραφή MPEG-1 Layer I και ο μόνος πρώτος έπαψε να υπάρχει, καθώς διεκόπη η παραγωγή των συσκευών DCC ο δε δεύτερος εξακολουθεί να χρησιμοποιείται από το MiniDisc και άλλα media players όντας πλέον στην τρίτη (και καλύτερη) εκδοχή του, ATRAC-3 και ATRAC3plus. Τους codecs PASC και ATRAC ακολούθησε η επόμενη γενιά της οποίας βασικός στόχος ήταν η μεταφορά ψηφιακού ήχου μέσα από τηλεοπτικά κανάλια για τις ανάγκες της τηλεόρασης. Το Layer I του MPEG-1 ακολούθησε το Layer II το οποίο έγινε γνωστό και ως MP2 (κατά πληροφορίες, μετά την έλευση του MP3...) αλλά και ως Musicam. Το Musicam, αποτελεί έναν μη-υποχρεωτικό codec για τα DVD-Video players και πιθανόν υπάρχουν ακόμη ταινίες με ήχο κωδικοποιημένο με αυτόν τον τρόπο καθώς και συσκευές που περιλαμβάνουν αποκωδικοποιητές Musicam. Οι προσπάθειες για μετάδοση ψηφιακού ήχου (DAB) είχαν ξεκινήσει στην Ευρώπη από το 1987 με το πρόγραμμα Eureka, στο οποίο συμμετείχαν ερευνητές από το Ινστιτούτο Fraunhofer (Fraunhofer IIS) με επικεφαλής τους Karlheinz Brandenburg και Dieter Seitzer. Το 1989 το Fraunhofer εξασφάλισε τα δικαιώματα για μία τεχνολογία συμπίεσης αναγνωρισμένη ως MPEG-1 Layer III (το Layer III, υποδηλώνει περισσότερο αποτελεσματική συμπίεση και προς τα πίσω συμβατότητα με τους codecs των Layers I&II). Το γεγονός ότι η πρόταση του Fraunhofer δεν χρησιμοποιήθηκε τελικώς για ψηφιακή μετάδοση θα μπορούσε να είναι μοιραίο για την συγκεκριμένη τεχνολογία, όμως οι Γερμανοί συνέχισαν την εξέλιξη της και το 1995 (14 Ιουλίου) αυτή απέκτησε μία περισσότερο εμπορική ονομασία: MP3. Η ιδέα του Brandenburg και των άλλων για χρήση του MP3 σε φορητές συσκευές, στον υπολογιστή και στο διαδίκτυο πήρε σάρκα και οστά την διετία 1997-98, όταν οι Tomislav Uzelac, Justin Frankel και Dmitry Boldyren κατασκεύασαν το λογισμικό για το πρώτο player (ο πρώτος) και την έκδοση για το λειτουργικό Windows οι δεύτεροι. Τα προϊόντα ονομάζονταν AMP MP3 Playback Engine και WINAMP αντίστοιχα. Η συνέχεια έκρυβε ταχέως την εξέλιξη: Τον Ιανουάριο του 2001 στην CES παρουσιάστηκε από την Coding Technologies η τεχνολογία MP3Pro, μία βελτιωμένη έκδοση του MP3, ενώ ήδη από το 1997 κυκλοφορούσε ένας διαφορετικός codec, συμβατός με τις προδιαγραφές MPEG-2, ο AAC, ο οποίος το 2003 μεταφέρθηκε και στην προδιαγραφή MPEG-4 και έγινε γνωστός όταν τον επέλεξε η Apple για το περιφημο iTunes. Ο χώρος



Η δέσημη καμπυλών Fletcher και Munson δείχνουν πώς μεταβάλλεται το κατώφλι ακουστότητας σε συνάρτηση με την συχνότητα και την ένταση (loudness) ενός ήχου.

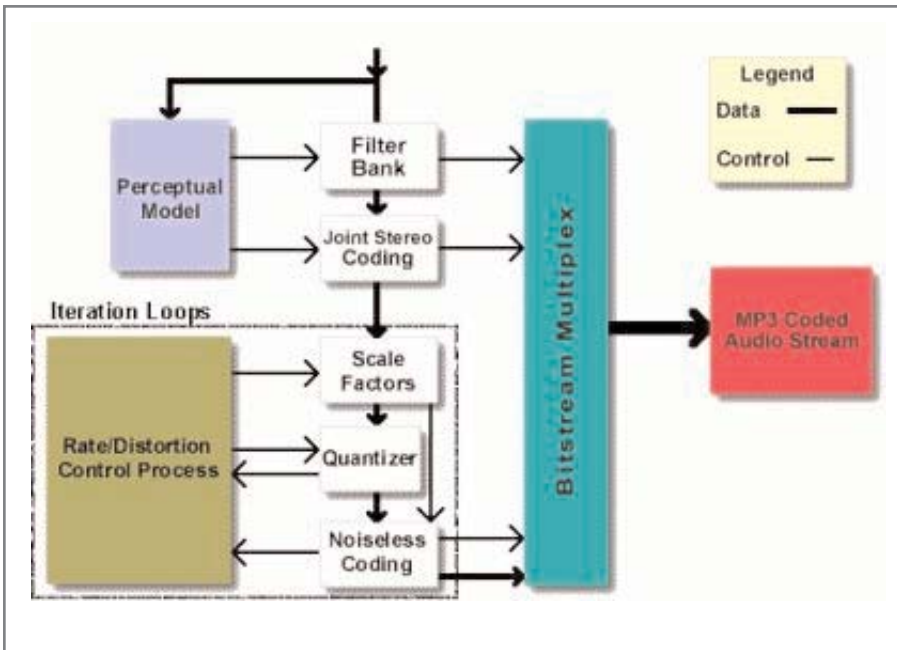


Το φαινόμενο της επικάλυψης: Ένας δυνατός ήχος, μεταβάλλει τοπικά την καμπύλη του κατωφλίου ακουστότητας, επικαλύπτοντας έναν γειτονικό αλλά αδύναμο ήχο.

της τεχνολογίας συμπίεσης ηχητικών σημάτων περιλαμβάνει -βεβαίως- και τα άκρα: Τον εντελώς "κλειστό" codec των Windows που είναι γνωστός ως WMA (Windows Media Audio) και οποίος βρίσκεται στην έκδοση 9 και τον εντελώς ελεύθερο codec Ogg Vorbis ο οποίος αποτελεί προϊόν συλλογικής ανάπτυξης (υπό το πνεύμα του open source) αλλά και ελεύθερης εμπορικής χρήσης, καθώς δεν απαιτείται πληρωμή δικαιωμάτων για την ενσωμάτωσή του σε συσκευές. Το 2004 το Fraunhofer IIS παρουσίασε μία πολυκαναλική έκδοση του codec MP3, το MP3 Surround. Το MP3 Surround επιτρέπει την κωδικοποίηση 5.1 καναλιών σε ένα stream το οποίο μπορεί να χρησιμοποιηθεί είτε από έναν συμβατικό decoder MP3,



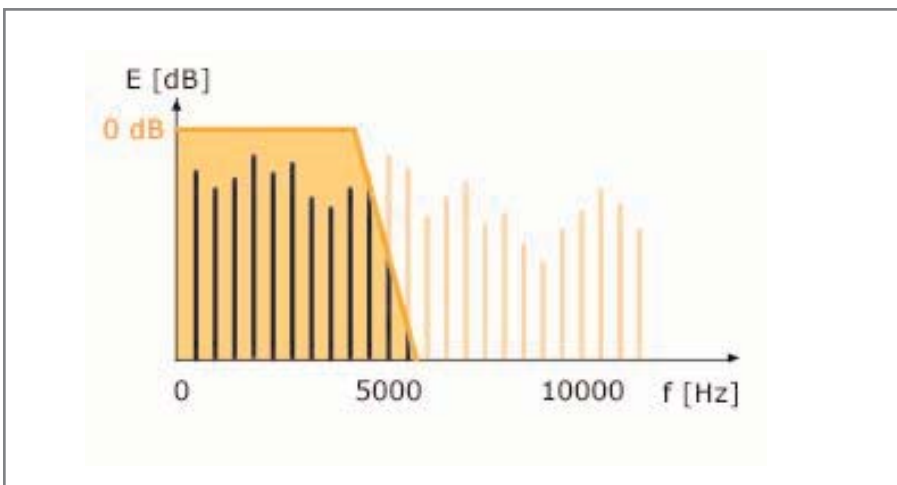
Ο Karlheinz Brandenburg του Fraunhofer Institute έθεσε τα θεμέλια και ανέπτυξε τον codec MPEG-1 Layer III που αργότερα έγινε γνωστός με το όνομα MP3



Το διάγραμμα βαθμίδων του κωδικοποιητή MP3 (πηγή: Fraunhofer IIS)

$$f_j = \sum_{k=0}^{2n-1} x_k \cos \left[\frac{\pi}{n} \left(j + \frac{1}{2} \right) \left(k + \frac{1}{2} + \frac{n}{2} \right) \right]$$

Αυτή η συνάρτηση περιγράφει τον μετασχηματισμό MDC. Ο μετασχηματισμός αυτός επιτρέπει την μετάβαση των υπολογισμών του codec από το πεδίο του χρόνου και τις τιμές του σήματος x στο πεδίο της συχνότητας αναπαριστώντας το σήμα με μία σειρά συντελεστές f.



Η τεχνική Spectral Band Replication (SBR) επιτρέπει την δημιουργία της ανώτερης περιοχής συχνοτήτων του φάσματος με την βοήθεια του κατώτερου φάσματος και μίας ομάδας παραμέτρων. (πηγή: Coding Technologies)

ώστε να διατηρείται η προς τα πίσω συμβατότητα είτε από έναν ειδικό decoder, με bitrate συγκρίσιμο αυτό του MP3. Οι codecs MP3 χρησιμοποιούν ως βασικό μοντέλο ακοής αυτό που βασίζεται στις καμπύλες κατωφλίου ακουστότητας (Minimal Audition Threshold).

MPEG-1 Layer III (MP3): Οι τεχνικές κωδικοποίησης κατά MP3 και MP3Plus

Η ανθρώπινη ακοή παρουσιάζει την μέγιστη ευαισθησία της στην περιοχή 1kHz-5kHz

και έξω από τα όρια αυτά μειώνεται δραστικά. Οι Fletcher και Munson κατασκεύασαν μία δέσμη από καμπύλες ελάχιστης ακουστότητας (οι οποίες περιγράφουν την στάθμη ηχητικής πίεσης που πρέπει να έχει ένας ήχος για να είναι μόλις ακουστός σε συνάρτηση με την συχνότητά του) όπου φαίνεται πώς η ευαισθησία της ακοής μεταβάλλεται όχι μόνο με την συχνότητα αλλά και με την ένταση (loudness). Όσο μικρότερη είναι η ένταση, τόσο μικρότερη η ευαισθησία της ακοής και τόσο μεγαλύτερη στάθμη απαιτείται για να γίνει ακουστή μία συγκεκριμένη συχνότητα. Οι καμπύλες Fletcher/Munson προσφέρουν ένα πρώτο κριτήριο για την συμπίεση: Οι

πληροφορίες κάτω από το κατώφλι ακουστότητας μπορούν να αποκοπούν από το σήμα ή, σκεπτόμενοι αντίστροφα, μπορούμε να αποθηκεύσουμε ανεπιθύμητα σήματα -όπως ο θόρυβος κβάντισης- σε περιοχές όπου το κατώφλι ακουστότητας είναι υψηλό κάτι που θα έχει αποτέλεσμα τα σήματα αυτά να μην γίνονται αντιληπτά. Ωστόσο πολύ μεγαλύτερο ενδιαφέρον έχει το φαινόμενο της επικάλυψης (masking): Με δεδομένους δύο γειτονικούς ήχους (δηλαδή ήχους με παραπλήσιες συχνότητες) ο ισχυρότερος αλλοιώνει τοπικά την καμπύλη κατωφλίου ακουστότητας επικαλύπτοντας τον ασθενέστερο ο οποίος πλέον δεν γίνεται αντιληπτός από τους ακροατές. Μπορεί να φανταστεί κανείς το φαινόμενο της επικάλυψης ως μία δυναμική, σε πραγματικό χρόνο, μεταβολή του κατωφλίου ακουστότητας σε σχέση με τους ήχους που φθάνουν στα αυτιά μας. Το μοντέλο που χρησιμοποιείται στον codec MP3 κάνει ακριβώς αυτό: Υπολογίζει σε κάθε στιγμή το φασματικό περιεχόμενο του σήματος, αποφασίζει ποιού ήχου επικαλύπτονται από ισχυρότερους και δεν τους κωδικοποιεί. Σε κάθε χρονική στιγμή, ο codec έχει στην διάθεσή του έναν αριθμό ψηφίων, ο οποίος εξαρτάται από τον βαθμό συμπίεσης που του έχει ζητηθεί (δηλαδή το επιτρεπόμενο bitrate). Καθώς η συμπίεση εξελίσσεται δεν χρησιμοποιούνται όλα τα ψηφία, ιδιαίτερα αν η στιγμιαία μορφή του σήματος είναι "εύκολα" συμπίεσιμη, έτσι είναι δυνατή η συντήρηση μίας δεξαμενής ψηφίων (bits reservoir) από την οποία αντλούνται περισσότερα ψηφία για σήματα που απαιτούν μεγαλύτερη ακρίβεια στην κωδικοποίηση.

Ο βασικός μηχανισμός του μοντέλου επικάλυψης, υποβοηθείται στην περίπτωση του MP3 και από δύο άλλους μηχανισμούς: Ανάλογα με την πολυπλοκότητα του προγράμματος και τις ποιοτικές απαιτήσεις του ο χρήστης μπορεί να επιλέξει την κωδικοποίηση "Joint Stereo" κατά την οποία, κωδικοποιούνται οι μεν χαμηλές συχνότητες ως ένα μονοφωνικό κανάλι (mid channel, L+R), οι δε υψηλές (οι οποίες φέρουν κατά τεκμήριο και τις πληροφορίες για την στερεοφωνική εικόνα) σε ένα δεύτερο κανάλι ως διαφορές μεταξύ των αρχικών καναλιών (side channel, L-R). Το δεύτερο εργαλείο για την μείωση του bitrate που χρησιμοποιεί ο codec MP3 είναι, τέλος, γνωστό ως εντροπική κωδικοποίηση (entropy coding) ένα μάλλον τρομακτικό όνομα για μία διαδικασία η οποία κωδικοποιεί τα ψηφία που προκύπτουν από τα προηγούμενα στάδια με βάση την πιθανότητα εμφάνισής τους (probability coding, κατ' αναλογία με τον γραπτό λόγο, όπου κάποιες λέξεις έχουν μεγαλύτερη πιθανότητα εμφάνισης σε σχέση με άλλες). Η εντροπική κωδικοποίηση είναι γνωστή και ως Huffman Coding (από τον David A. Huffman που την περιέγραψε το 1952) και είναι μια μη-απωλεστική διαδικασία. Από τα παραπάνω, γίνεται προφανές ότι το ψυχοακουστικό μοντέλο για την συμπίεση στην περίπτωση του MP3 "τρέχει" στο πεδίο της συχνότητας (γεγονός λογικό, αν αναλογιστεί κανείς ότι τόσο οι καμπύλες Fletcher/Munson όσο και η καμπύλη που περιγράφει το φαινόμενο της επικάλυψης

έχουν ως οριζόντιο άξονα την συχνότητα, κάτι που μπορείτε να δείτε και στα σχετικά σήματα). Ωστόσο, το ηχητικό σήμα εξελίσσεται στον χρόνο. Πώς μεταφερόμαστε από το πεδίο του χρόνου στο πεδίο της συχνότητας για να κάνουμε τους υπολογισμούς μας; Στην περίπτωση του MP3 (αλλά και των codecs MPEG-2 AAC, MPEG-4 AAC, Ogg Vorbis και AC-3 -ο τελευταίος χρησιμοποιείται από το Dolby Digital) χρησιμοποιείται μία πολύπλοκη μαθηματική διαδικασία κατά την οποία το υπό κωδικοποίηση μέρος του σήματος (block) υπόκειται σε επεξεργασία από μία τράπεζα φίλτρων (filter bank) η οποία χωρίζει το φάσμα σε 32 περιοχές και στην συνέχεια σε μετασχηματισμό MDC (Modified Discrete Cosine Transform). Το σύνολο της διαδικασίας βασίζεται στις αρχές του μετασχηματισμού Fourier, ο οποίος αναλύει ένα σήμα εξελισσόμενο στον χρόνο σε μία σειρά συνιστωσών στο πεδίο της συχνότητας.

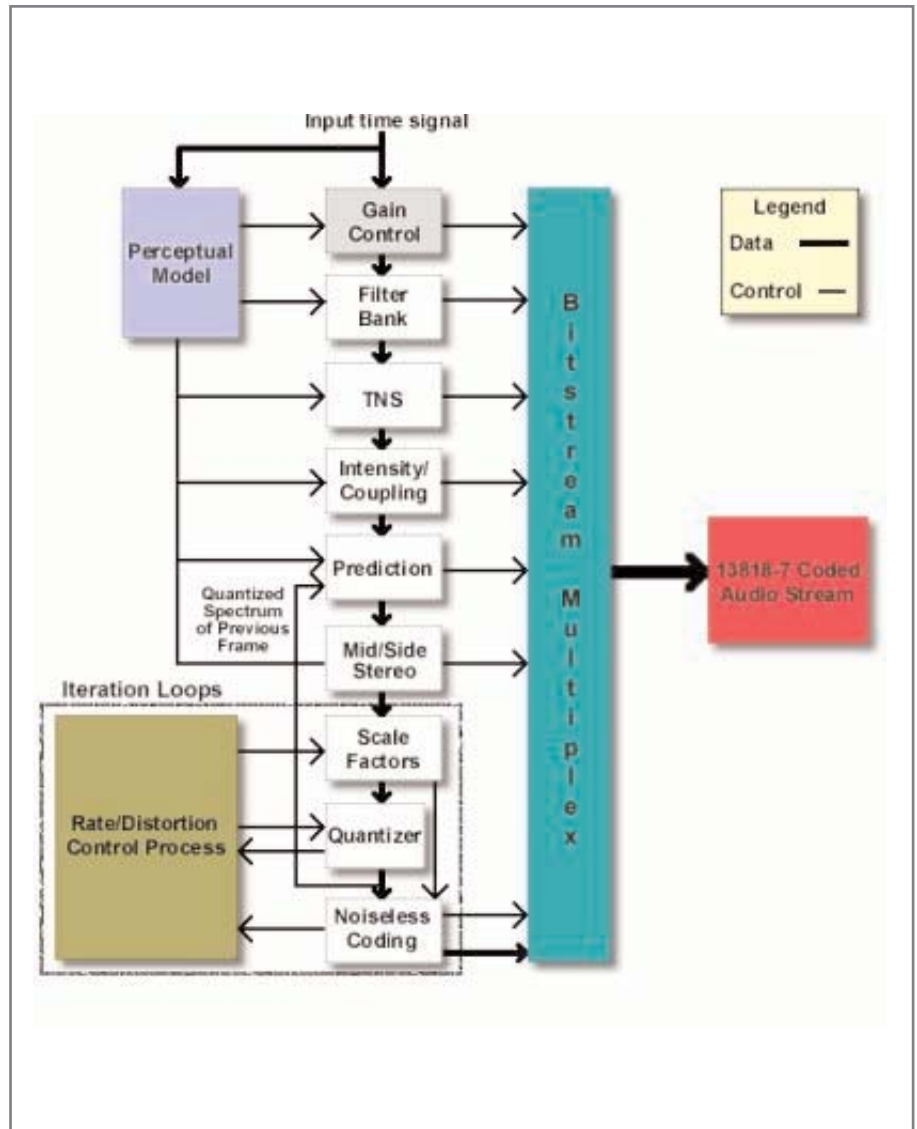
Στις παραπάνω τεχνικές, προστέθηκε, το 2001, μία ακόμη, με την ονομασία SBR (Spectral Band Replication, από την Coding Technologies), μετονομάζοντας τον codec που την υποστηρίζει σε MP3Pro. Ο τυπικός codec MP3 όταν λειτουργεί σε χαμηλά bitrates (μέχρι 96kbps, τα οποία απαιτούνται για streaming και μεταφορά δεδομένων από το διαδίκτυο) έχει σχετικά περιορισμένο εύρος συχνοτήτων που κυμαίνεται από τα 8-11kHz για bitrates 64kbps μέχρι τα 15kHz για bitrates 96kbps, μία επίδοση που είναι γνωστή και ως "FM Quality" επειδή το άνω όριο του φάσματος (15kHz) συμπίπτει με το θεωρητικό άνω όριο των ραδιοφωνικών εκπομπών με διαμόρφωση FM. Η τεχνική SBR χρησιμοποιεί την μορφή του φάσματος στις χαμηλές και τις μεσαίες συχνότητες καθώς και πληροφορίες για την μορφή του φάσματος στις υψηλές που έχουν κωδικοποιηθεί μαζί με το σήμα για να δημιουργήσει εκ του μηδενός την περιοχή που λείπει. Η τεχνική αυτή είναι αποτελεσματική αλλά απαιτεί πολύ μεγαλύτερη ισχύ επεξεργασίας: Για ένα σήμα με bitrate 64kbps (με πραγματικό εύρος 8kHz), χρησιμοποιεί τα 4kbps για την μεταφορά των παραμέτρων που περιγράφουν το φάσμα στις υψηλές συχνότητες και επιτρέπει την δημιουργία της περιοχής από τα 8kHz μέχρι τα 16kHz, απαιτώντας όμως τριπλάσια ισχύ από τον επεξεργαστή.

Άλλοι Codecs: MPEG-2 AAC, MPEG-4 AAC, WMA, Ogg Vorbis, FLAC

Εκτός του MP3, ο οποίος είναι ίσως ο πλέον γνωστός, υπάρχουν αρκετοί ακόμη codecs του οποίους αξίζει να γνωρίζετε κανείς:

AAC

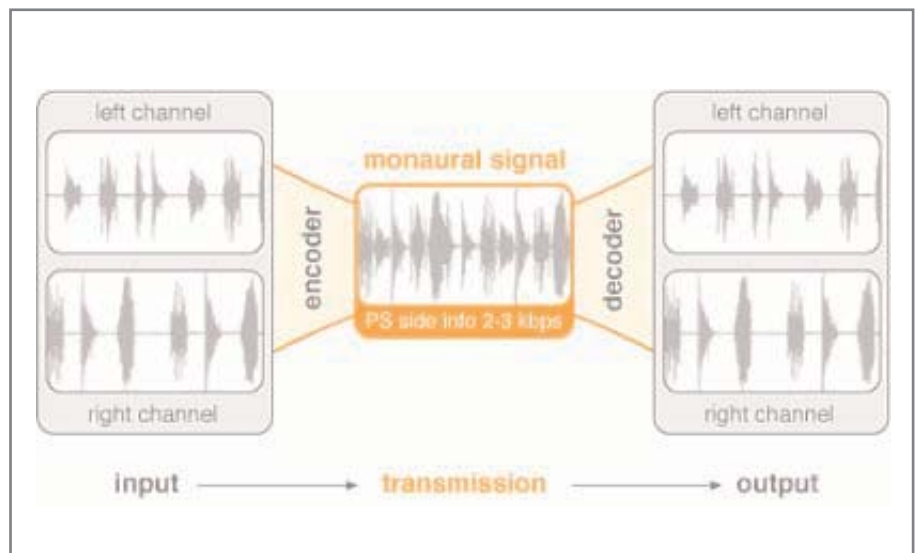
Ο Codec AAC (Advanced Audio Codec) παρουσιάστηκε για πρώτη φορά το 1997 και χρησιμοποιεί, όπως και ο MP3, το ψυχοακουστικό μοντέλο της επικάλυψης. Στον τομέα της επεξεργασίας χρησιμοποιεί επίσης τον MDCT αλλά χωρίς την τράπεζα φίλτρων, καθώς επίσης και την



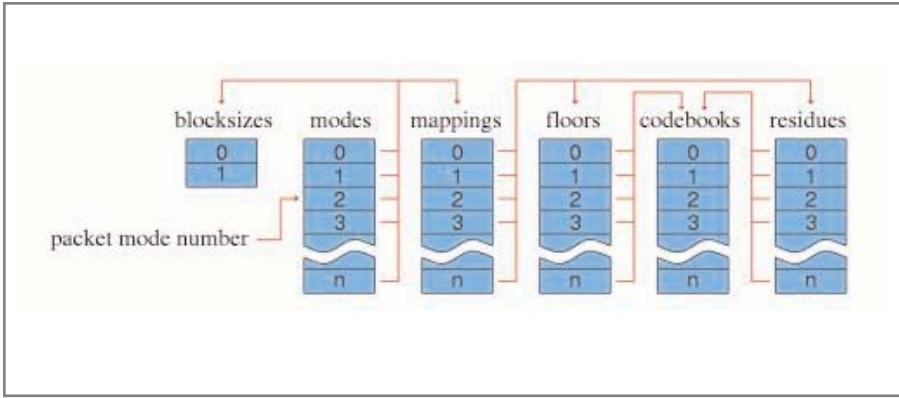
Το διάγραμμα βαθμίδων του κωδικοποιητή AAC (πηγή: Fraunhofer IIS)

κωδικοποίηση Huffman. Οι διαφορές του με τον MP3 βρίσκονται σε δύο διαφορετικά εργαλεία: Το Linear Predictive Coding, μια μαθηματική διαδικασία πρόβλεψης της εξέλιξης του σήματος στο πεδίο της συχνότητας, εξαγωγής ενός σφάλματος (στα σημεία που η πρόβλεψη δεν συμφωνεί με τις πραγματικές τιμές) και κωδικοποίησης

του σφάλματος αυτού, τεχνική η οποία αρχικά χρησιμοποιήθηκε σε συστήματα μετάδοσης φωνής όπως η κινητή τηλεφωνία GSM, και την Χρονικά Μεταβαλλόμενη Μορφοποίηση Θορύβου (Temporal Noise Shaping, TNS) η οποία με βάση την μορφή του σήματος στο πεδίο της συχνότητας, μεταβάλλει την κατανομή του θορύβου



Ο τρόπος λειτουργίας της τεχνικής Parametric Stereo (πηγή: Coding Technologies)



Το διάγραμμα βαθμίδων του κωδικοποιητή Ogg Vorbis. Η κωδικοποίηση του φάσματος βάσης και του απομείνου φάσματος, καθώς επίσης και η εντροπική κωδικοποίηση χρησιμοποιούν παραμέτρους που μεταφέρονται μαζί με το σήμα (codebooks)

$$\hat{x}(n) = - \sum_{i=1}^P a_i x(n - i)$$

Στην τεχνική LPC (Linear Prediction Coding) η τιμή του σήματος μπορεί να προβλεφθεί με βάση τις τιμές p προηγούμενων δειγμάτων...

$$e(n) = x(n) - \hat{x}(n)$$

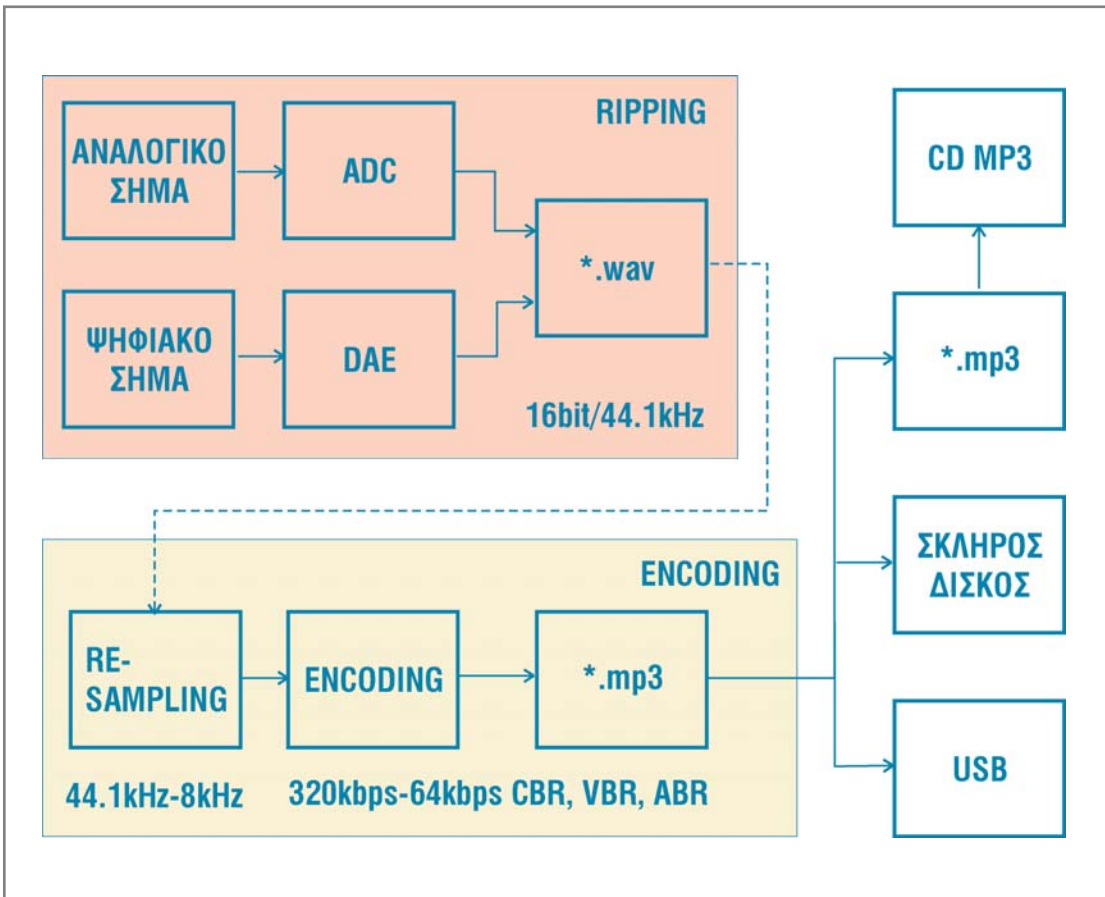
...και η διαφορά της τιμής που προβλέπεται με την πραγματική τιμή του σήματος $e(n)$ είναι το σφάλμα το οποίο και κωδικοποιείται.

κβάντισης που προκύπτει από την κωδικοποίηση με τον βέλτιστο τρόπο, έτσι ώστε αυτή να παραμένει πάντοτε κάτω από την στάθμη του κατωφλίου ακουστότητας, όπως αυτό ορίζεται κάθε στιγμή με βάση τις καμπύλες Fletcher/Munson. Ο codec AAC προσφέρει καλύτερη απόδοση από τον MP3, (96kbps AAC αντιστοιχούν στην ποιότητα των 128kbps MP3) μπορεί να δεχτεί προς κωδικοποίηση σήματα μέχρι 96kHz, και προσφέρει καλύτερη διαχείριση των σημάτων στην περίπτωση που επιλεγεί η τεχνική Joint Stereo.

Το 2003 παρουσιάστηκε η έκδοση του AAC συμβατή με τις προδιαγραφές MPEG-4. Ο Codec έγινε γνωστός ως MPEG-4 AAC και λίγο αργότερα διάσημος, όταν η Apple αποφάσισε να τον χρησιμοποιήσει για το iTunes. Η πλέον διαδεδομένη έκδοσή του είναι η HE AAC (High Efficiency AAC) πιο γνωστή και με το εμπορικό όνομα aacPlus (της Coding Technologies). Η έκδοση v1 του codec aacPlus είναι ένας συνδυασμός AAC και SBR ενώ η έκδοση v2 περιλαμβάνει και την τεχνική Parametric Stereo η οποία μεταφέρει κωδικοποιημένη την βασική πληροφορία του σήματος σε ένα κανάλι και μία σειρά από παραμέτρους που περιγράφουν την στερεοφωνική εικόνα σε ένα πλευρικό κανάλι με bitrate 3kbps (side channel). Οι παράμετροι του PS οι οποίες αποθηκεύονται στο κανάλι αυτό είναι η Panorama (PAN) η οποία αφορά στις διαφορές μεταξύ των δύο καναλιών (L,R) και η Stereo Ambience (SA) η οποία αφορά στις αντιλήψεις της στερεοφωνικής εικόνας. Μία άλλη έκδοσή του MPEG-2 AAC είναι το AAC SSR (Scalable Sample Rate) το οποίο ανέπτυξε η Sony, πλησιάζει αρκετά το ATRAC-3 και έχει γίνει αποδεκτό από την προδιαγραφή MPEG-4. Το AAC SSR που χρησιμοποιεί τον συνδυασμό τράπεζας φίλτρων τεσσάρων περιοχών και MDCT, έχει την φήμη ότι είναι καλό σε χαμηλά bitrates (που φθάνουν τα 24kbps).

WMA

Με δεδομένη την τάση μεταφοράς, αποθήκευσης και χρήσης ακουστικού υλικού υπό την μορφή απωλεστικών συμπιεσμένων αρχείων μέσω ηλεκτρονικών υπολογιστών, ο μεγαλύτερος κατασκευαστής λειτουργικών συστημάτων, η Microsoft, δεν μπορούσε, βεβαίως να μείνει έξω από το παιχνίδι. Ο codec WMA (Windows Media Audio) είναι ο proprietary codec των Windows ο οποίος χρησιμοποιείται από την αντίστοιχη εφαρμογή (Windows Media Player) και αρχικώς προσέφερε παρόμοιες δυνατότητες με



Το διάγραμμα βαθμίδων που περιγράφει την διαδικασία Ripping/Encoding

τον MP3, αλλά σε πολλές περιπτώσεις με καλύτερη ποιότητα. Ως προϊόν της Microsoft δίνει ιδιαίτερη έμφαση στην ψηφιακή διαχείριση των πνευματικών δικαιωμάτων (DRM) και συνοδεύεται από τις γνωστές μονομανίες και ιδιαιτερότητες των προϊόντων της εταιρίας -που συνήθως εκνευρίζουν του σκληροπυρηνικούς χρήστες... Με την έκδοσή 9.1 είναι συμβατός με την προδιαγραφή MPEG-4 προσφέροντας την δυνατότητα και για μη-απωλεστική συμπίεση καθώς και για πολυκαναλικό ήχο.

Ogg Vorbis

Ο codec Ogg Vorbis έχει αναπτυχθεί γύρω από το πρότυπο αρχείων Ogg (του xiph.org) και βασίζεται στην open source εφαρμογή απωλεστικής συμπίεσης με την ονομασία Vorbis. (η περιέργη αυτή σύνθετη ονομασία αντανακλά τη δομή του συστήματος: Format αρχείου [Ogg] - κωδικοποίηση [Vorbis], υπάρχει και η εκδοχή Ogg Flac, μη απωλεστικού codec). Ο Ogg Vorbis χρησιμοποιεί MDCT για τον μετασχηματισμό του σήματος από το πεδίο του χρόνου στο πεδίο της συχνότητας και μία εναλλακτική μέθοδο επεξεργασίας του φάσματος, κατά την οποία, κωδικοποιείται με λίγα ψηφία το φάσμα βάσης του οποίου η κατανομή είναι σχετικώς ομαλή (floor encoding) και με περισσότερα ψηφία το απομένον φάσμα του οποίου η δομή και η χρονική εξέλιξη είναι περισσότερο πολύπλοκη (residue encoding). Ο Ogg Vorbis χρησιμοποιεί επίσης εντροπική κωδικοποίηση και υποστηρίζει μεταβλητό bitrate ανάλογα με τις ανάγκες του σήματος. Το πλεονέκτημα του συγκεκριμένου codec είναι ότι προσφέρει καλή ποιότητα και είναι ελεύθερος δικαιωμάτων, μπορεί δηλαδή να χρησιμοποιηθεί από κατασκευαστές συσκευών και χρήστες χωρίς κόστος, ενώ το γεγονός ότι όλοι έχουν πρόσβαση στον πηγαίο κώδικα επιτρέπει την συλλογική του ανάπτυξη.

MP3 Surround

Το MP3 Surround παρουσιάστηκε από το Fraunhofer IIS το 2004 και αποτελεί μία επέκταση του MP3. Βασίζεται στην τεχνολογία Binaural Cue Coding της Agere η οποία επιτρέπει την μείξη σημάτων από πολλά κανάλια σε δύο (downmix) για την δημιουργία ενός σήματος συμβατού με τον απλό MP3 codec ενώ κωδικοποιεί μία σειρά από παραμέτρους που περιγράφουν πλήρως το ηχητικό πεδίο surround (side channel), τις οποίες επίσης θέτει στην διάθεση του codec ο οποίος, αν είναι συμβατός με MP3 Surround, μπορεί να τις αξιοποιήσει και να δημιουργήσει και πάλι το πολυκαναλικό σήμα. Οι παράμετροι αυτοί είναι οι χρονικές διαφορές μεταξύ των καναλιών (Inter-Channel Time Difference -ICTD), οι διαφορές στάθμης μεταξύ των καναλιών (Inter-Channel Level Difference -ICLD), και η συσχέτιση μεταξύ των καναλιών (Inter-Channel Coherence).

Τελευταίος, Αλλά όχι Εσχατός: Ο FLAC

Ο FLAC (Free Lossless Audio Codec) είναι ίσως ο σημαντικότερος μη απωλεστικός

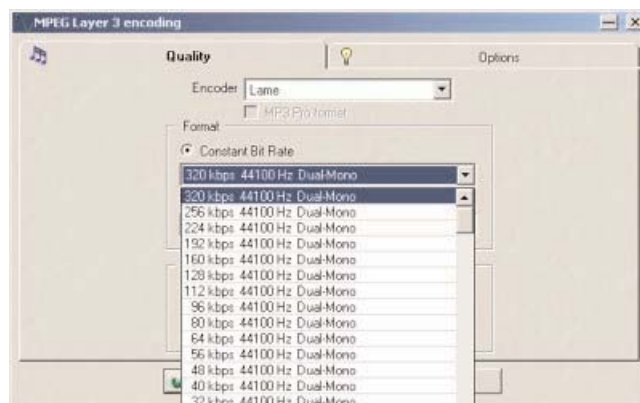
codec στον οποίο έχει εύκολη πρόσβαση τόσο ο χρήστης όσο και οι κατασκευαστές. Είναι open source, και δίνει ιδιαίτερη έμφαση στην ενσωμάτωση μεταδεδομένων (metadata, δηλαδή πληροφορίες που αφορούν στο περιεχόμενο του υπό κωδικοποίηση σήματος και όχι το σήμα καθ'αυτό, ώστε να είναι στην συνέχεια εύκολη η αρχειοθέτηση, η αναζήτηση και η ανάκτησή του, γεγονός που τον κάνει ιδιαίτερα συμπαθή codec στους σχεδιαστές μεγάλων audio servers). Ο FLAC χρησιμοποιεί μεταβλητό μήκος block (της ομάδας των δεδομένων που επεξεργάζεται σε κάθε χρονική στιγμή) από 16 έως 65535 δείγματα, μία προχωρημένη τεχνική διαχείρισης των καναλιών που προσομοιάζει το Joint Stereo αλλά μπορεί να λειτουργήσει σε τέσσερις ρυθμούς (Mid/Side, Left/Side, Right/Side, Independent, ανάλογα με την μορφή και τις ανάγκες του σήματος), καθώς επίσης και μία πολύ πιο ευέλικτη τεχνική πρόβλεψης σε σχέση με τον AAC η οποία περιλαμβάνει μηδενική πρόβλεψη/συμπίεση (Verbatim, όταν δεν υπάρχει σήμα), σταθερή πρόβλεψη (Constant, για σταθερές τιμές σήματος, όπως το dc), σταθερή LPC (παρόμοια με αυτήν που χρησιμοποιείται από τον AAC) και LPC με την χρήση ψηφιακού φίλτρου FIR 32ης τάξης. Όπως και στην περίπτωση του AAC ο codec κωδικοποιεί το σφάλμα (error, residue) που απομένει όταν από τις πραγματικές τιμές του σήματος αφαιρέσουμε αυτές που προκύπτουν από την πρόβλεψη. Αφού ο τρόπος της πρόβλεψης είναι σταθερός και γνωστός από το player, το μόνο που χρειάζεται είναι να γνωρίζουμε τις διαφορές αυτές.



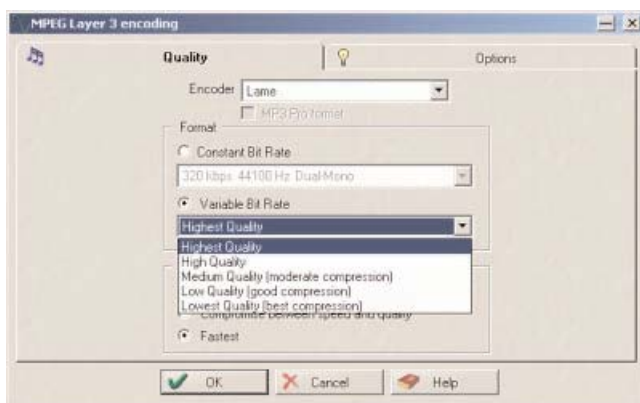
Ένα πρόγραμμα κωδικοποίησης έχει συνήθως ένα μενού επιλογών που μοιάζει (αλλά δεν είναι απαραίτητα όμοιο) με αυτό.



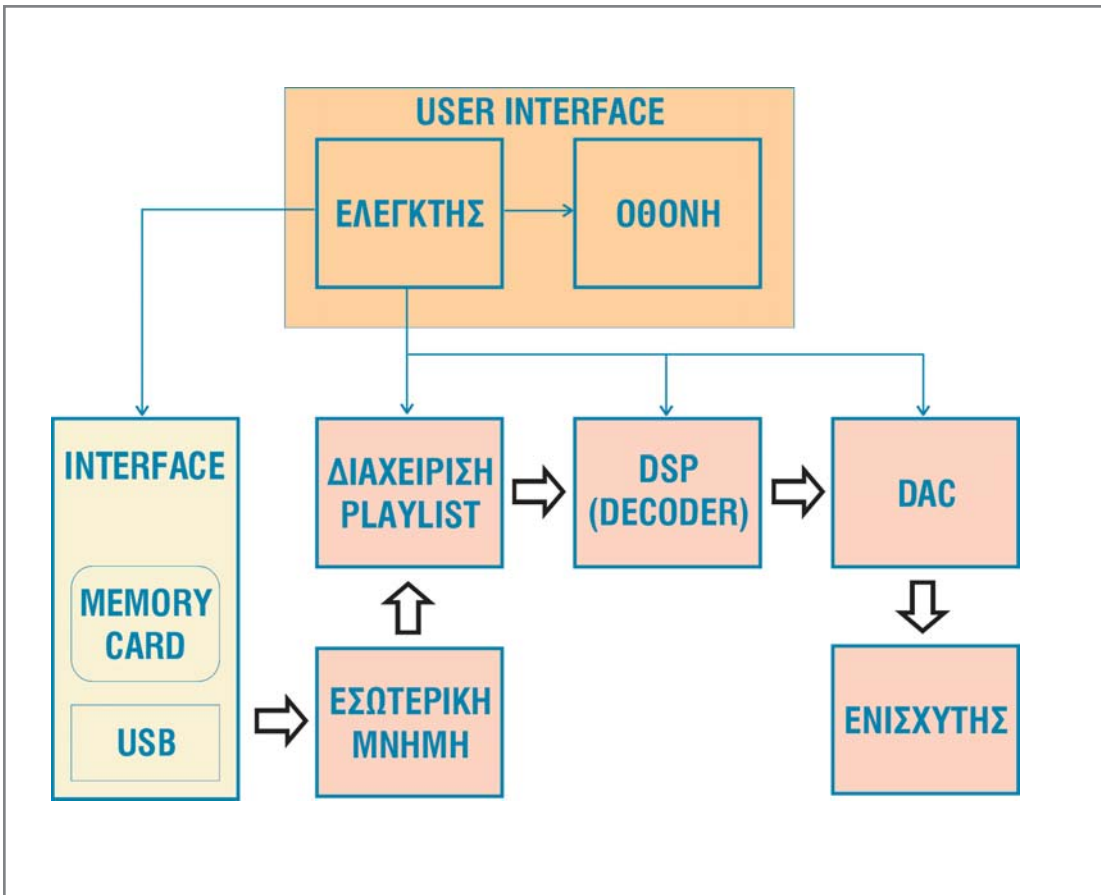
Η πρώτη επιλογή είναι ο codec που θα χρησιμοποιηθεί (εδώ υπάρχει επιλογή μεταξύ LAME και Fraunhofer)



Το επόμενο βήμα είναι η επιλογή του bitrate



Αν υπάρχει η δυνατότητα, ο χρήστης μπορεί να επιλέξει διάφορες στρατηγικές διαχείρισης του bitrate.



όλα τα προϊόντα είναι λίγοι και συγκεκριμένα τέσσερις: Οι codecs του Fraunhofer IIS (FastEnc, MP3Enc, L3Enc), ο open source codec LAME (παρά το γεγονός ότι το όνομά του σημαίνει ακριβώς το αντίθετο...! Lame Ain't a Mp3 Encoder) ο οποίος βασίζεται στο ψυχοακουστικό μοντέλο GPSYCHO, ο BladeEnc του Tord Jansson (ο οποίος έχει σταματήσει επισήμως να διανέμεται επειδή παραβιάζει κάποια δικαιώματα του Fraunhofer -όχι ότι δεν μπορείτε να τον βρείτε πάντως...!) και τέλος ο codec της Xing, ο πρώτος που επέτρεψε την κωδικοποίηση MP3 με μεταβλητή ροή δεδομένων (VBR). Η ποιότητα της κωδικοποίησης σχετίζεται άμεσα με τον τρόπο που λειτουργεί ο codec, οπότε είναι σαφές ότι γίνεται πολλή συζήτηση για το ποιός είναι καλύτερος και μπορείτε να βρείτε ολόκληρα fora γύρω από το θέμα αυτό. Μία περιήγηση

Το διάγραμμα βαθμίδων ενός mp3 audio player

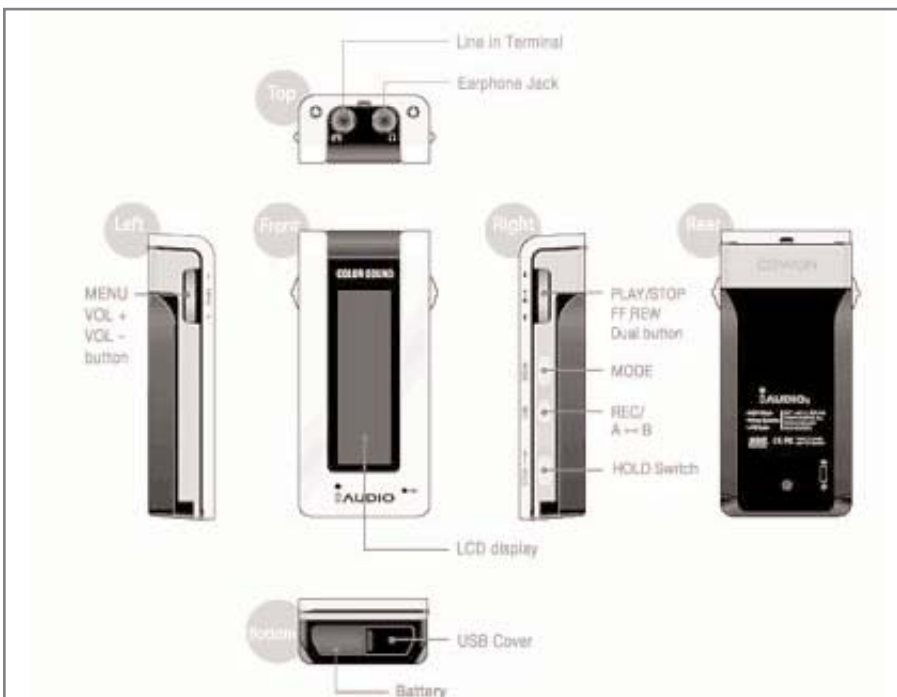
**Στην Πράξη:
Κωδικοποίηση/Δημιουργία
αρχείων MP3
(Ripping/Encoding)**

Με εξαίρεση κάποιες λεπτομέρειες οι οποίες γίνονται σαφείς σε κάθε περίπτωση, όλοι οι Perceptual codecs λειτουργούν ως προς την χρήση με τον ίδιο τρόπο. Έτσι, αν και τα

όσα πρακτικά αναφέρονται αφορούν τον MP3, τα περισσότερα ισχύουν και για τον AAC, και για τον Ogg Vorbis. Οσον αφορά τον MP3, αξίζει να αναφέρουμε κατ' αρχήν το εξής: Παρά τον τεράστιο αριθμό προϊόντων που κυκλοφορούν (software players, επιτραπέζια και φορητά players, λογισμικό επεξεργασίας και δημιουργίας αρχείων mp3 κ.λπ) στην πράξη οι πραγματικοί codecs, δηλαδή οι πυρήνες επάνω στους οποίους βασίζονται

θα δείξει ότι ο LAME χαιρεί της εκτίμησης όλων, οι Fraunhofer θεωρούνται καλοί αλλά αργοί (κάτι σαν τα γερμανικά αυτοκίνητα ας πούμε...) με εξαίρεση τον νεότερο FastEnc, ο BladeEnc είναι για τους φανατικούς που θα αντέξουν το command line interface του, και ο Xing είναι μάλλον ο ταχύτερος (γι' αυτό και τον προτιμούν πολλά consumer προϊόντα). Το ποιόν θα διαλέξει κανείς είναι, βεβαίως προσωπικό θέμα...

Η δημιουργία ενός αρχείου mp3 (η κατάληξη .mp3 υποδηλώνει αρχείο mp3, προφανώς...) ξεκινά από την διαδικασία ripping. Το ripping επιτρέπει την δημιουργία ενός αρχείου wav (η native-εγγενής μορφή αρχείων ήχου σε υπολογιστή που δεν περιλαμβάνει κανενός είδους συμπίεση) το οποίο αποθηκεύεται στον υπολογιστή, από το πρωτογενές υλικό. Αν το υλικό αυτό είναι αναλογικό (βινύλιο, ταινία ή όποιο άλλο σήμα εμφανίζεται σε εξόδους line κάποιας πηγής) θα πρέπει πρώτα να περάσει από έναν μετατροπέα A/D του οποίου η ποιότητα είναι επίσης καθοριστική (οπότε δώστε σημασία την κάρτα ήχου του υπολογιστή σας...). Αν το υλικό είναι αποθηκευμένο σε CD (η πλέον συνηθής περίπτωση) τότε μπορούμε να χρησιμοποιήσουμε την διαδικασία Digital Audio Extraction (DAE) και να μεταφέρουμε το ψηφιακό περιεχόμενο του δίσκου απ' ευθείας σε αρχείο wav. Κάθε ripper που σέβεται τον εαυτό του υποστηρίζει DAE, και το ίδιο συμβαίνει για κάθε CD-ROM drive. Ωστόσο, αν έχετε κάποιο πολύ παλιό drive θα πρέπει να βεβαιωθείτε. Είναι στατιστικώς πιθανόν αλλά όχι βέβαιο ότι θα το υποστηρίζει. Το αρχείο wav που προκύπτει από τον ripper έχει προδιαγραφές που εξαρτώνται από την διαδικασία: Η πλέον συνηθισμένες



Τα μέρη ενός mp3 audio player (πηγή: iAudio)

ονομάζονται συλλήβδην "CD Quality", έχουν δηλαδή συχνότητα δειγματοληψίας 44.1kHz, και μήκος λέξης 16bit. (χωρίς να αποκλείονται και άλλες επιλογές, όπως τα 48kHz ή τα 24bit). Αυτό το αρχείο είναι η πηγή των δεδομένων που τοποθετείται στην είσοδο του encoder. Ο encoder πραγματοποιεί resampling (με βάση τις οδηγίες που του δίνουμε), κωδικοποιεί το σήμα με βάση το μοντέλο της απωλεστικής συμπίεσης και δημιουργεί ένα αρχείο mp3. Το αρχείο αυτό μπορεί να αποθηκευτεί τοπικά (και να αναπαράγεται μέσω του σχετικού player, της κάρτας ήχου και των ηχείων του υπολογιστή) να μεταφερθεί σε κάποια εξωτερική συσκευή μέσω USB (συνήθως, χωρίς να αποκλείεται το Ethernet ή οι ασύρματες ζεύξεις), να μετατραπεί και πάλι σε wav με απώτερο στόχο την εγγραφή του σε CD που είναι συμβατό με απλά CD players (κακή ιδέα...) ή να εγγραφεί απ' ευθείας σε CD ώστε να χρησιμοποιηθεί από συσκευές που είναι συμβατές με CD MP3 (πολλά DVD-Video players και ακόμη περισσότερα CD αυτοκινήτου, προσωπικά CD και φορητά ηχοσυστήματα). Οι διαδικασίες που περιγράφονται ως ripping και encoding μπορεί πράγματι να λάβουν χώρα ξεχωριστά αλλά είναι δυνατόν, και αρκετά συνηθισμένο, μία και μόνη εφαρμογή να τις πραγματοποιεί αυτόματα. Μιλάμε τότε για ripper/encoder. Σε κάθε περίπτωση ο χρήστης έχει να κάνει μία σειρά από επιλογές: Κατ' αρχήν, θα πρέπει να επιλέξει τον codec εφόσον υπάρχει αυτή η δυνατότητα. Στην συνέχεια, το sample rate (η συχνότητα δειγματοληψίας) το οποίο είναι διαφορετικό από το bitrate και ίσως περισσότερο καθοριστικό: Με εξαίρεση τις εφαρμογές streaming ή την ύπαρξη μίας πολύ κακής σύνδεσης στο διαδίκτυο, δεν υπάρχει κανένας λόγος να επιλεγεί άλλο sample rate από τα 44.1kHz/16bit. Χαμηλότερες επιλογές περιορίζουν την ποιότητα σημαντικά. Η επόμενη παράμετρος είναι η ροή των δεδομένων, το bitrate. Είναι προφανές ότι όσο μεγαλύτερη ροή δεδομένων επιλέξει κανείς, τόσο μικρότερο λόγο συμπίεσης επιβάλλει, οπότε η ποιότητα είναι καλύτερη, εις βάρος του μεγέθους του αρχείου και της ταχύτητας μεταφοράς. Για "τοπική" χρήση, διαλέξτε 320kbps, για μεταφορά σε δίσκους (οπότε θέλετε πολλά τράκς) διαλέξτε κάτι

Συχνά ο κατασκευαστής μπορεί να προσφέρει ειδική βάση (cradle) που παίζει και τον ρόλο του φορηστού.



μεταξύ 128-320kbps) και για μεταφορά μέσω του διαδικτύου (από σύνδεση χαμηλής ταχύτητας) μπορείτε να πέσετε στα 64kbps γνωρίζοντας ότι γίνονται παραχωρήσεις στην ποιότητα. Αν ο encoder το επιτρέπει, μπορείτε να διαλέξετε την διαχείριση των δύο καναλιών (Joint Stereo -προφανώς για ποιοτικά αποτελέσματα καλό είναι το αποφύγετε), και την ταχύτητα κωδικοποίησης, συνήθως ως συμβιβασμό μεταξύ μικρού χρόνου και καλής ποιότητας. Εκτός και αν έχετε πραγματικούς λόγους να βιάζεσθε η πιο σοφή επιλογή εδώ είναι το slow. Από κάποια χρονική στιγμή και μετά, οι codecs MP3 υποστηρίζουν τρεις στρατηγικές διαχείρισης του Bitrate: Το σταθερό (Constant Bitrate -CBR) όπου η ροή των δεδομένων είναι σταθερή στην προρυθμισμένη τιμή, το μεταβλητό (Variable Bitrate -VBR) όπου η ροή των δεδομένων βελτιστοποιείται με βάση την μορφή του σήματος αλλά δεν ξεπερνά την προρυθμισμένη τιμή και το μέσο (Average Bitrate -ABR) όπου ο codec έχει την δυνατότητα να ταλαντώνεται γύρω από μία μέση τιμή που είναι η προρυθμισμένη. Οι δύο τελευταίες στρατηγικές επιτρέπουν την δημιουργία αρχείων μικρότερων από αυτά που θεωρητικά επιβάλλει το σταθερό bitrate και ταυτόχρονα επιτρέπουν στον codec να δώσει βαρύτητα σε δύσκολα σήματα, κάνοντας "οικονομία" σε απλούστερα και ίσως είναι μία καλή επιλογή αν ήδη έχει γίνει ένας συμβιβασμός προς μία χαμηλότερη τιμή του bitrate.

Περί Μεταδεδομένων (Metadata), Tagging και ID3v2

Ο όρος metadata ή μεταδεδομένα στην ελληνική γλώσσα, αφορά γενικώς την παροχή πληροφοριών για ένα πακέτο πληροφοριών. Ενα απλό παράδειγμα μεταδεδομένων είναι η συλλογή με κάρτες σε μία βιβλιοθήκη (το index της δηλαδή) οι οποίες περιλαμβάνουν -για κάθε βιβλίο- το όνομα του συγγραφέα, τον τίτλο, την χρονολογία, την θεματική κατηγορία, κ.λπ. Στην περίπτωση των αρχείων MP3 αλλά και κάθε άλλου αρχείου που προέρχεται από έναν ripper/encoder, ανεξαρτήτως του αλγορίθμου που χρησιμοποιείται, τα μεταδεδομένα αφορούν κατ' αρχήν τον καλλιτέχνη, τον τίτλο και την χρονολογία του κάθε μουσικού έργου, και στην συνέχεια είναι δυνατόν να περιλαμβάνουν το είδος, τον τίτλο κάθε τράκ, ακόμη και την φωτογραφία του εξωφύλλου ή και σχόλια που εισάγει ο ίδιος ο χρήστης. Όλες αυτές οι



Το user interface είναι συνήθως απλό: Ένας ή περισσότεροι διακόπτες και μία οθόνη που πληροφορεί τον χρήστη για τις διάφορες λειτουργίες και το track που αναπαράγεται την συγκεκριμένη στιγμή.



Η συνηθέστερη επιλογή για την επικοινωνία ενός player με τον υπολογιστή είναι μέσω USB...



Ο διακόπτης "Hold" είναι συνηθισμένος στα περισσότερα προσωπικά players. Δεν επιτρέπει τους χειρισμούς από τυχαία πίεση των διάφορων διακοπών.



Κάποια players διαθέτουν και είσοδο line για την on-the-fly κωδικοποίηση και αποθήκευση αναλογικών σημάτων.

πληροφορίες περιλαμβάνονται σε μία εικονική ετικέτα ή tag η οποία συνοδεύει το κάθε αρχείο. Η χρησιμότητα του Tagging, της δημιουργίας δηλαδή metadata σε αρχεία μουσικής είναι εξαιρετικά σημαντική, διότι τα tags αποτελούν τον μοναδικό αξιόπιστο τρόπο πρόσβασης στο υλικό, το οποίο -σε αντίθεση με τους πραγματικούς δίσκους- δεν έχει "εξώφυλλο", "σχήμα" ή συγκεκριμένη θέση στην δισκοθήκη αλλά απλώς υπάρχει ως ένα filename "κάπου" μέσα σε ένα σύστημα αποθήκευσης. Τα σωστά tags εξασφαλίζουν εύκολη πρόσβαση με κάποιο από τα πολλά προγράμματα διαχείρισης που υπάρχουν και την δυνατότητα αναπαραγωγής με βάση το είδος, τον καλλιτέχνη ή την χρονολογία (με την δημιουργία μίας απλής αναζήτησης, έχοντας ως "κλειδί" το αντίστοιχο μεταδεδωμένο). Η ιστορία των tags ξεκίνησε το 1996, όταν ο Eric Kemp χρησιμοποίησε για πρώτη φορά τον όρο ID3 ως συντομογραφία για την φράση "IDentify an MP3". Η πρώτη έκδοση μίας προδιαγραφής για το Tagging ήταν η έκδοση 1.1 της αρχικής ιδέας του Kemp, ID3v1.1 η οποία για πρώτη φορά περιλάμβανε την δυνατότητα εισαγωγής των τράκς ενός αρχείου, ενώ η σημερινή έκδοση, ID3v2, περιλαμβάνει περισσότερα πεδία και δυνατότητα αποθήκευσης μεγάλου αριθμού χαρακτήρων. Θεωρητικά, το σωστό ripping/encoding απαιτεί την δημιουργία ενός λεπτομερούς tag, κυρίως επειδή πολύ γρήγορα μία συλλογή εικονικών τράκς σε έναν σκληρό δίσκο (ή



Το logo της ημεπίστης προδιαγραφής ID3v2 για το tagging των αρχείων ήχου που δημιουργούνται από codecs. Αν και η ιδέα ξεκίνησε από τον Eric Kemp το 1996 και αφορούσε μόνο αρχεία mp3, όλα τα αρχεία που δημιουργούνται από τους δημοφιλείς encoders την υποστηρίζουν.

ένα "αχανές" media player) δημιουργεί πολύ σοβαρά προβλήματα διαχείρισης -με απλά λόγια, ξεχνάτε πού έχετε τί... Αυτό μπορεί να σημαίνει αρκετή δουλειά, υπό την προϋπόθεση ότι λογισμικό σας επιτρέπει το editing των σχετικών πεδίων, αλλά συνήθως, η εισαγωγή των σχετικών

πληροφοριών μπορεί να γίνει αυτόματα αν έχετε πρόσβαση στο διαδίκτυο, όπου υπάρχουν αρκετές βάσεις δεδομένων με όλα αυτά τα στοιχεία, τα οποία "κατεβαίνουν" αυτόματα όταν το λογισμικό σας τα ζητήσει.

Στην Πράξη: Αναπαραγωγή αρχείων MP3 (Playback)

Η αναπαραγωγή αρχείων mp3 όσο αυτά βρίσκονται αποθηκευμένα στον υπολογιστή είναι απλή. Συνήθως τα αρχεία με κατάληξη .mp3 μέσω των "File Associations" "σηκώνουν" με διπλό κλικ το media player με τον κατάλληλο codec και δεν απαιτείται κάποια άλλη προσπάθεια από

την πλευρά του χρήστη. Αν τα αρχεία βρίσκονται στον υπολογιστή αλλά πρέπει να φορτωθούν σε κάποια εξωτερική συσκευή player ή μνήμη, αυτό συνήθως γίνεται χωρίς ειδικό λογισμικό, αφού οι μνήμες και όλο και πιο συχνά τα players φαίνονται από την πλευρά του υπολογιστή ως "removable storage device" όταν συνδέονται στο USB. Από την πλευρά του player τα πράγματα δεν είναι, επίσης, σύνθετα: Ένα απλό user interface (μερικοί διακόπτες και ένα display με άλλα λόγια) δίνει πρόσβαση στην δημιουργία και διαχείριση ενός καταλόγου (playlist) με τα περιεχόμενα του player (είτε αυτά βρίσκονται σε εξωτερική είτε σε εσωτερική μνήμη). Η διαχείριση αυτή μπορεί να είναι απλή ή περισσότερο πολύπλοκη (κατάταξη κατά μουσικό είδος, κ.λπ - με βάση τα tags) ανάλογα με τις δυνατότητες της συσκευής. Από την στιγμή που ο χρήστης επιλέξει το τράκ και πατήσει το play τα δεδομένα του αντίστοιχου αρχείου mp3 οδηγούνται στον αποκωδικοποιητή, στον μετατροπέα d/a και στον ενισχυτή ακουστικών για τα περαιτέρω. Ανάλογα με την ισχύ του dsp μπορεί να προσφέρονται διάφορες δυνατότητες όπως ισοσταθμιστής, ρυθμιστικά τονικότητας ή επεξεργασία των χαμηλών συχνοτήτων. Πολλά players διαθέτουν και αναλογικές εισόδους. Στην περίπτωση αυτή ο ψηφιακός επεξεργαστής διαθέτει μετατροπέα A/D και MP3 encoder.

Για περισσότερο διάβασμα:

Links

<http://www.iis.fraunhofer.de>
<http://www.codingtechnologies.com/>
<http://www.vorbis.com/>
<http://flac.sourceforge.net/>
<http://lame.sourceforge.net/>
<http://www.id3.org>

Άλλες πηγές:

1. Discrete-Time Signal Processing, Oppenheim/Schafer, Prentice Hall, 1989
2. Convergence in Broadcastand Communications Media, Watkinson, Focal Press, 2001
3. High Quality Audio Coding, (AES 17th International Conference), 1999
4. The MPEG Handbook, Watkinson, Focal Press, 2001

COLOR SOUND MP3 PLAYER
iAUDIO

Μουσικός Πειρασμός!



Όλα τα iAudio ενσωματώνουν το σύστημα ψηφιακής επεξεργασίας ήχου της BBE.



Όλα τα iAudio υποστηρίζουν πολλαπλά format: MP3, WMA, ASF, WAV, OGG & FLAC (X5).



Όλα τα iAudio διαθέτουν ραδιόφωνο FM υψηλής ποιότητας με 24 προεπιλογές.



Όλα τα iAudio υποστηρίζουν εγγραφή φωνής, ραδιοφώνου & line με άμεση κωδικοποίηση σε MP3.



Όλα τα iAudio συνοδεύονται από ακουστικά υψηλής ποιότητας της Cresyn.



Όλα τα iAudio είναι πλήρως συμβατά με MAC & Linux OS.



**iAudio X5:
Το Απόλυτο
Gadget!**

MP3 player/encoder/Ραδιόφωνο FM. Αναπαραγωγή Video & φωτογραφιών (MPEG4, JPG). Έγχρωμη οθόνη LCD-TFT 160x128 pixel. Ισοσταθμιστής JetEffect. Δυνατότητα αναβάθμισης firmware. Μπαταρία λιθίου. Σύνδεση USB 2.0. Χωρητικότητα: 20MB / 30MB / 60MB



**iAudio 5:
Ήχος και Χρώμα!**

MP3 player/encoder/Ραδιόφωνο FM. Ισοσταθμιστής JetEffect. Led Backlight με 1000 χρώματα. Δυνατότητα αναβάθμισης firmware. 20 ώρες αναπαραγωγής με μία μόνο μπαταρία AAA. Σύνδεση USB 2.0. Χωρητικότητα: 256MB / 512MB / 1GB



**iAudio G3:
Η Μουσική Βίβλος!**

MP3 player/encoder/Ραδιόφωνο FM. Ισοσταθμιστής JetEffect. Δυνατότητα αναβάθμισης firmware. 50 ώρες αναπαραγωγής με μία μόνο μπαταρία AA. Σύνδεση USB 2.0. Χωρητικότητα: 256MB / 512MB / 1GB



**iAudio U2:
Ένα Μελωδικό
Κόσμημα!**

MP3 player/encoder/Ραδιόφωνο FM. Ισοσταθμιστής JetEffect. Δυνατότητα αναβάθμισης firmware. Μπαταρία λιθίου. Σύνδεση USB 2.0. Χωρητικότητα: 256MB / 512MB / 1GB



avmentor

URL: <http://www.avmentor.gr>, ©Ακρίες Εκδόσεις 2005