

## ΒΕΣ 04 – Συμπίεση και Μετάδοση Πολυμέσων



# Συμπίεση Ήχου με βάση την Αντίληψη:

## Τα πρότυπα συμπίεσης MPEG-1 layer I, layer II, layer III

- \* Εισαγωγή
- Βασικό Διάγραμμα
- Ψυχοακουστική
- Κωδικοποίηση ήχου με το πρότυπο MPEG-1
- MPEG-1 Layer II
- MPEG-1 Layer III (mp3)

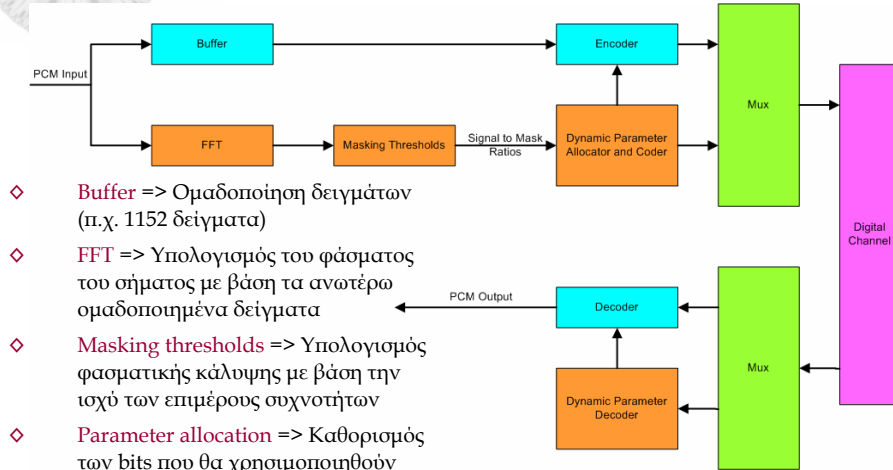
## Εισαγωγή



- ◇ Υπάρχουν πολλοί αλγόριθμοι κωδικοποίησης με βάση την αντίληψη οι κυριότεροι από τους οποίους βασίζονται στο πρότυπο MPEG (Moving Picture Expert Groups)
- ◇ Είναι αλγόριθμοι συμπίεσης με απώλειες και βασίζονται στη **λογική της μη μετάδοσης του τμήματος εκείνου ενός ηχητικού** (ή οπτικού) σήματος το οποίο δεν μπορεί να γίνει αντιληπτό από τα ανθρώπινα αισθητήρια όργανα
- ◇ Για την συμπίεση ήχου λαμβάνονται τρία βασικά χαρακτηριστικά του ακουστικού συστήματος του ανθρώπου:
  - ◇ Ακουστική ικανότητα των ανθρώπων στις διάφορες συχνότητες
  - ◇ Διακριτική ικανότητα αντίληψης συχνοτήτων από το αυτί (critical bands)
  - ◇ Φασματική κάλυψη (Spectral masking)

- ☑ Εισαγωγή
- ★ Βασικό Διάγραμμα
- ☐ Ψυχοακουστική
- ☐ Κωδικοποίηση ήχου με το πρότυπο MPEG-1
- ☐ MPEG-1 Layer II
- ☐ MPEG-1 Layer III (mp3)

## Κωδικοποίηση με βάση την αντίληψη: Βασικό διάγραμμα



- ◇ **Buffer** => Ομαδοποίηση δειγμάτων (π.χ. 1152 δείγματα)
- ◇ **FFT** => Υπολογισμός του φάσματος του σήματος με βάση τα ανωτέρω ομαδοποιημένα δείγματα
- ◇ **Masking thresholds** => Υπολογισμός φασματικής κάλυψης με βάση την ισχύ των επιμέρους συχνοτήτων
- ◇ **Parameter allocation** => Καθορισμός των bits που θα χρησιμοποιηθούν για την κωδικοποίηση των δειγμάτων (ανά ζώνη συχνοτήτων)

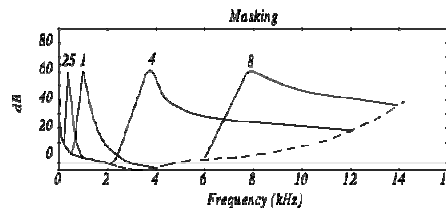
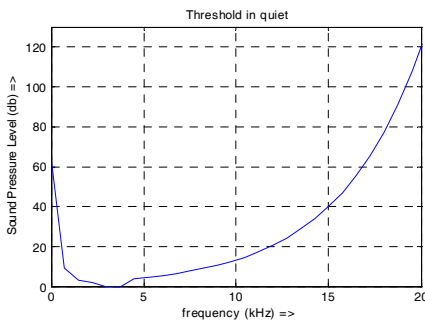
© 2006 Nicolas Tsapatsoulis

- ☑ Εισαγωγή
- ☑ Βασικό Διάγραμμα
- ★ Ψυχοακουστική
- ☐ Κωδικοποίηση ήχου με το πρότυπο MPEG-1
- ☐ MPEG-1 Layer II
- ☐ MPEG-1 Layer III (mp3)

## Ψυχοακουστική



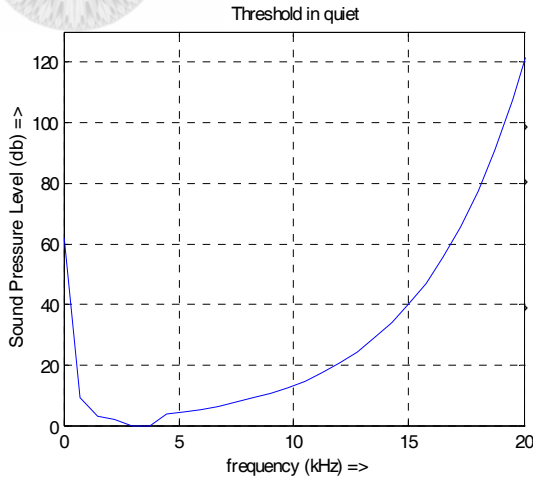
- ◇ Εξέταση των δυνατοτήτων του ανθρώπινου ακουστικού συστήματος:
  - ◇ Ακουστική ικανότητα σε σχέση με τη συχνότητα (threshold of hearing)
  - ◇ Διακριτική ικανότητα αυτιού σε σχέση με τη συχνότητα (critical bands)
  - ◇ Φασματική κάλυψη (spectral masking)
  - ◇ Χρονική κάλυψη (temporal masking)



© 2006 Nicolas Tsapatsoulis

- Εισαγωγή
- Βασικό Διάγραμμα
- Ψυχοακουστική
- Κωδικοποίηση ήχου με το πρότυπο MPEG-1
- MPEG-1 Layer II
- MPEG-1 Layer III (mp3)

## Ακουστική ικανότητα σε σχέση με τη συχνότητα



Το διάγραμμα του σχήματος δίνει την ακουστική ικανότητα των ανθρώπων σε σχέση με τη συχνότητα.

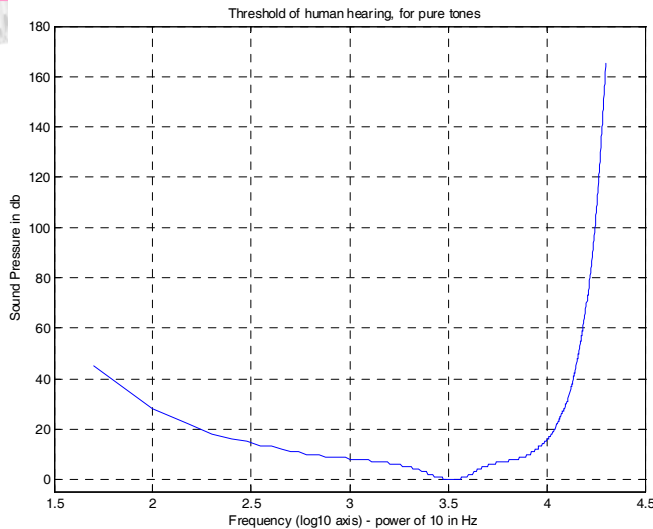
Μεγαλύτερη ευαισθησία (ακουστική ικανότητα) στις συχνότητες 2-4kHz

Συχνότητες με ισχύ μικρότερη από τη από τη ελάχιστη ακουστή μπορούν να απαλειφθούν κατά τη διαδικασία της κωδικοποίησης χωρίς να υπάρξει αντιληπτή διαφορά στην ποιότητα.

Για παράδειγμα μια συχνότητα 15kHz με ισχύ μικρότερη από 40 db δεν είναι ακουστή

- Εισαγωγή
- Βασικό Διάγραμμα
- Ψυχοακουστική
- Κωδικοποίηση ήχου με το πρότυπο MPEG-1
- MPEG-1 Layer II
- MPEG-1 Layer III (mp3)

## Ακουστική ικανότητα σε σχέση με τη συχνότητα (λογαριθμική κλίμακα)



- Εισαγωγή
- Βασικό Διάγραμμα
- Ψυχοακουστική
- Κωδικοποίηση ήχου με το πρότυπο MPEG-1
- MPEG-1 Layer II
- MPEG-1 Layer III (mp3)

## Διακριτική ικανότητα αυτιού σε σχέση με τη συχνότητα (Critical Bands)



- ◇ Το ανθρώπινο αυτί έχει περιορισμένη διακριτική ικανότητα. Μέσα σε μια περιορισμένη κριτική περιοχή εύρους ζώνης συγχέει (δεν μπορεί να διαχωρίσει) τις συχνότητες που ακούει
- ◇ Η διακριτική ικανότητα του αυτιού μπορεί να προσομοιωθεί με το άθροισμα μιας σειράς ζωνοπερατών φίλτρων με εύρος ζώνης:
  - ◇ Περίπου 100 Hz για συχνότητες <500 Hz
    - ◇ Κάθε περιοχή που αντιστοιχεί στο εύρος ζώνης ενός ζωνοπερατού φίλτρου ονομάζεται κρίσιμη περιοχή συχνοτήτων (critical band)
    - ◇ Η διακριτική ικανότητα του αυτιού, σε σχέση με τις συχνότητες, εντός της κρίσιμης περιοχής συχνοτήτων είναι σχετικά περιορισμένη ενώ αντίθετα το αυτί μπορεί πολύ εύκολα να διακρίνει συχνότητες που ερρίσκονται σε διαφορετικά critical bands
    - ◇ π.χ. μπορεί να διακρίνει τις συχνότητες 370 Hz και 430 Hz ανεξάρτητα την ισχύ που έχει καθεμία από αυτές (εφόσον φυσικά η ένταση τους ξεπερνά την ελάχιστη ακουστή τιμή) γιατί υπάγονται σε διαφορετικά critical bands (band # 4, band # 5) ενώ δεν μπορεί να διακρίνει τις συχνότητες 650 Hz και 740 Hz γιατί υπάγονται στο ίδιο critical bands (band # 7).

© 2006 Nicolas Tsapatsoulis

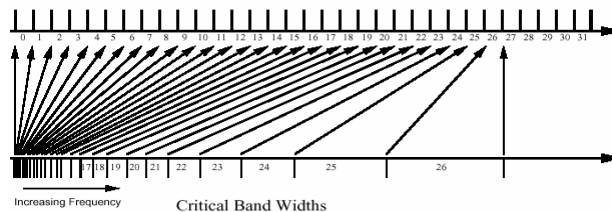
- Εισαγωγή
- Βασικό Διάγραμμα
- Ψυχοακουστική
- Κωδικοποίηση ήχου με το πρότυπο MPEG-1
- MPEG-1 Layer II
- MPEG-1 Layer III (mp3)

## Διακριτική ικανότητα αυτιού σε σχέση με τη συχνότητα (II)



- ◇ Περίπου 100 Hz για συχνότητες <500 Hz (συν.)
  - ◇ Το φαινόμενο της φασματικής κάλυψης είναι ισχυρότερο εντός ενός critical band
- ◇ Σύμφωνα με τη σχέση  $bw = 25 + 75 \cdot [1 + 1.4 \cdot (f/1000)^2]^{0.69}$ , (για συχνότητες  $f > 500$  Hz)
- ◇ Η διακριτική ικανότητα του αυτιού σε σχέση με τις συχνότητες οδήγησε στη χρήση φίλτρων ανάλυσης σε ζώνες (subband filters) στο πλαίσιο της κωδικοποίησης με βάση την αντίληψη

MPEG/Audio Filter Bank Bands



© 2006 Nicolas Tsapatsoulis

- Εισαγωγή
- Βασικό Διάγραμμα
- Ψυχοακουστική
- ★ Κωδικοποίηση ήχου με το πρότυπο MPEG-1
  - MPEG-1 Layer II
  - MPEG-1 Layer III (mp3)

## Κωδικοποίηση ήχου με το πρότυπο MPEG-1



- ◇ **MPEG-1 Layer I**
  - ◇ Κασέτα DCC της Philips
  - ◇ Συμπίεση 4:1
  - ◇ Ποιότητα μέτρια
  - ◇ Ρυθμός μετάδοσης 384 Kbps/κανάλι
- ◇ **MPEG-1 Layer II**
  - ◇ Ψηφιακό ραδιόφωνο DAB, VideoCD & DVD
  - ◇ Συμπίεση 6:1 & 8:1
  - ◇ Ποιότητα εφάμιλλη του CD
  - ◇ Ρυθμός μετάδοσης βελτιστοποιημένος στην περιοχή 192 ή 256 kbps / κανάλι
- ◇ **MPEG-1 Layer III (mp3)**
  - ◇ Αρχεία mp3 στο διαδίκτυο
  - ◇ Συμπίεση 10:1 & 12:1
  - ◇ Ποιότητα εφάμιλλη του CD
  - ◇ Ρυθμός μετάδοσης 128 Kbps για στερεοφωνικό ήχο (64 kbps/κανάλι)

© 2006 Nicolas Tsapatsoulis

- Εισαγωγή
- Βασικό Διάγραμμα
- Ψυχοακουστική
- ★ Κωδικοποίηση ήχου με το πρότυπο MPEG-1
  - MPEG-1 Layer II
  - MPEG-1 Layer III (mp3)

## Συνοπτική εικόνα

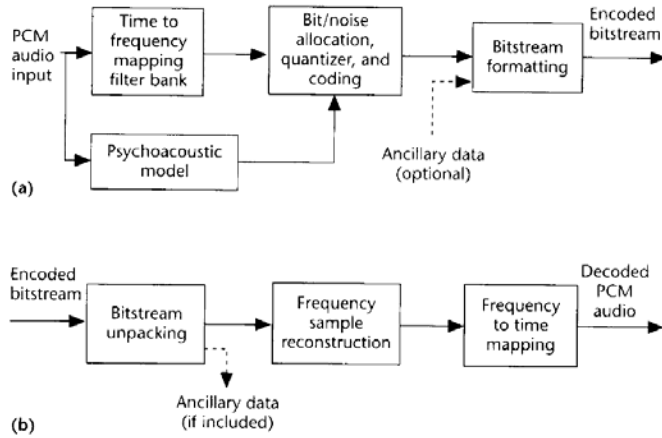


- ◇ **Βήματα κωδικοποίησης στο MPEG-1, layer I & II:**
  1. Χρήση 36 x 32 δείγματα = 1152 (samples)  
Με δειγματοληψία 48000 samples /sec αυτό αντιστοιχεί σε διάρκεια 24 ms
  2. Ανάλυση κάθε ομάδας 1152 δειγμάτων σε 32 ζώνες συχνοτήτων (για δειγματοληψία στα 48 kHz κάθε ζώνη έχει εύρος 750 Hz).
  3. Υπολογισμός του μέγιστου συντελεστή (scalefactor) σε κάθε ζώνη συχνοτήτων (ομάδα 36 samples)
  4. Υπολογισμός της κάλυψης (φασματικής και χρονικής) για κάθε ζώνη συχνοτήτων (χρησιμοποιώντας το σύνολο των 36 x 32 = 1152 δείγματα με χρήση του FFT)
  5. Αν η ισχύς σε κάποια ζώνη συχνοτήτων είναι μικρότερη από το κατώφλι κάλυψης τότε αυτή η ζώνη συχνοτήτων δεν κωδικοποιείται (0 bits)
  6. Σε διαφορετική περίπτωση υπολογίζουμε τον αριθμό των bits που απαιτούνται για την κωδικοποίηση της έντασης του σήματος που υπερβαίνει το κατώφλι κάλυψης (1 bit προσθέτει 6.02 db σηματοθροβικού λόγου)
  7. Μορφοποίηση του frame και μετάδοση

© 2006 Nicolas Tsapatsoulis

- ☑ Εισαγωγή
- ☑ Βασικό Διάγραμμα
- ☑ Ψυχοακουστική
- ★ Κωδικοποίηση ήχου με το πρότυπο MPEG-1
- ☐ MPEG-1 Layer II
- ☐ MPEG-1 Layer III (mp3)

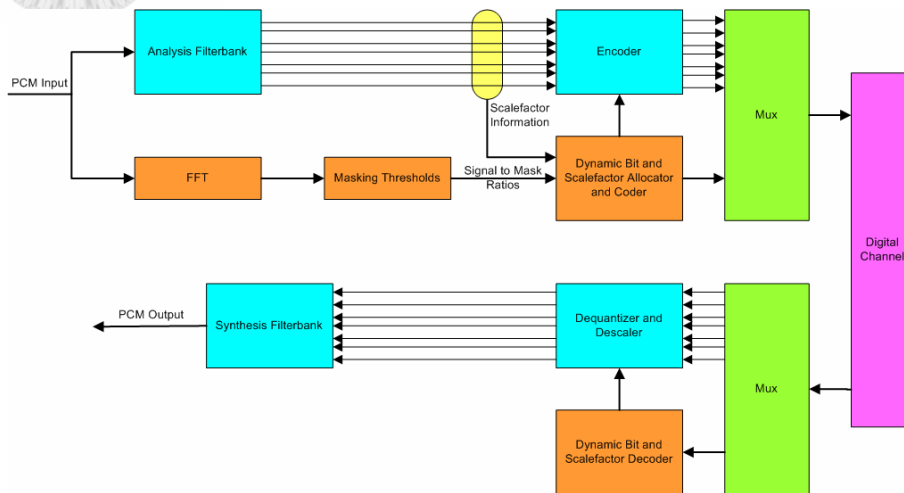
## Συνοπτική εικόνα (II)



© 2006 Nicolas Tsapatsoulis

- ☑ Εισαγωγή
- ☑ Βασικό Διάγραμμα
- ☑ Ψυχοακουστική
- ☑ Κωδικοποίηση ήχου με το πρότυπο MPEG-1
- ★ MPEG-1 Layer II
- ☐ MPEG-1 Layer III (mp3)

## MPEG-1 Layer I και II



© 2006 Nicolas Tsapatsoulis

- Εισαγωγή
- Βασικό Διάγραμμα
- Ψυχοακουστική
- Κωδικοποίηση ήχου με το πρότυπο MPEG-1
- ★ MPEG-1 Layer II
- MPEG-1 Layer III (mp3)

## Το MPEG-1 Audio frame



- ◇ Για κάθε 1152 δείγματα σχηματίζεται ένα πλαίσιο από bits που περιλαμβάνει τα ακόλουθα:
  - ◇ Επικεφαλίδα (header)
  - ◇ CRC (Cyclic Redundancy Code) 0-16 bits, προαιρετικό για έλεγχο σφαλμάτων
  - ◇ Επιμερισμός bits (Bit Allocation): Δηλώνει τον αριθμό των bits που χρησιμοποιούνται για την κωδικοποίηση των δειγμάτων σε κάθε μία από τις 32 ζώνες. Κάθε δείγμα μπορεί να κωδικοποιηθεί σε 0-15 bits άρα χρειάζονται 4 bits για να μας δώσουν αυτή την πληροφορία
  - ◇ Scale factors: Για τις ζώνες συχνοτήτων που δεν έχουν μηδενικό bit allocation η μέγιστη τιμή κάθε μιας από τις 32 ζώνες (στο σύνολο των 36 δειγμάτων) κωδικοποιείται με 6 bits.
  - ◇ Δείγματα (Samples): κωδικοποίηση των 1152 (36 x 32) δειγμάτων με βάση τον αριθμό των bits που υποδηλώνονται στο πεδίο bit allocation.
  - ◇ Συμπληρωματικά δεδομένα (Ancillary data): προαιρετικά

Header (32)	CRC (0,16)	Bit allocation (128-256)	Scale factors (0-384)	Samples	Ancillary data
----------------	---------------	--------------------------------	--------------------------	---------	-------------------

atsoulis

- Εισαγωγή
- Βασικό Διάγραμμα
- Ψυχοακουστική
- Κωδικοποίηση ήχου με το πρότυπο MPEG-1
- ★ MPEG-1 Layer II
- MPEG-1 Layer III (mp3)

## Επικεφαλίδα (Header)



- ◇ Η επικεφαλίδα αποτελείται από 32 bits και τα εξής πεδία:
  - ◇ Κώδικας συγχρονισμού 16 bits (SYNC = 1111111111111111)
  - ◇ Πληροφορίες καναλιού 2 bits (SI = Stereo Information)
    - ◇ Monophonic => Code 00
    - ◇ Dual Monophonic => Code 01
    - ◇ Stereo => Code 10
    - ◇ Joint Stereo => Code 11
  - ◇ Πληροφορίες Bit Rate 4 bit (BR = Bit Rate)
    - ◇ 32 kbps => Code 0001
    - ◇ 46 kbps => Code 0010
    - ◇ ... kbps => Code .....
    - ◇ 448 kbps => Code 1110
  - ◇ Συχνότητα δειγματοληψίας 2 bits (SR = Sampling Rate)
    - ◇ 32 kHz => Code 00
    - ◇ 44.1 kHz => Code 01
    - ◇ 48 kHz => Code 10
    - ◇ other => Code 11 (η συχνότητα ορίζεται στα επόμενα 8 bits)

© 2006 Nicolas Tsapatsoulis

- Εισαγωγή
- Βασικό Διάγραμμα
- Ψυχοακουστική
- Κωδικοποίηση ήχου με το πρότυπο MPEG-1
- ★ MPEG-1 Layer II
- MPEG-1 Layer III (mp3)

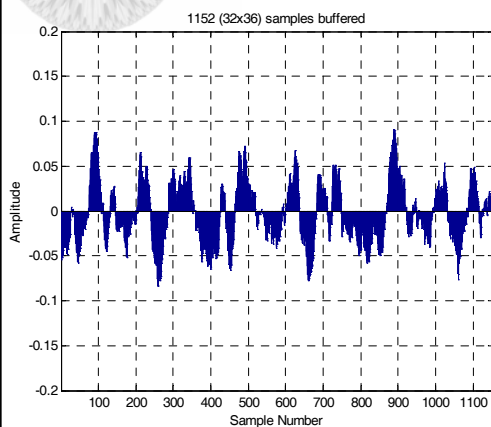
## Επιμερισμός bits (bit allocation)



- ◇ Το σήμα εισόδου διαχωρίζεται σε ομάδες 1152 δειγμάτων
- ◇ Κάθε ομάδα 1152 αναλύεται σε 32 ζώνες συχνοτήτων
- ◇ Υπολογίζεται η ακουστική ικανότητα σε κάθε μια από τις 32 ζώνες συχνοτήτων
- ◇ Υπολογίζεται η φασματική κάλυψη εξαιτίας ισχυρών συχνοτήτων (με τη βοήθεια του μετασχηματισμού FFT)
- ◇ Υπολογίζεται ποιες ζώνες συχνοτήτων θα είναι ακουστές μετά από το συνδυασμό ακουστικής ικανότητας και φασματικής κάλυψης
- ◇ Υπολογίζεται ο αριθμός των bits με τον οποίο θα κωδικοποιηθούν τα δείγματα εντός κάθε ζώνης συχνοτήτων

- Εισαγωγή
- Βασικό Διάγραμμα
- Ψυχοακουστική
- Κωδικοποίηση ήχου με το πρότυπο MPEG-1
- ★ MPEG-1 Layer II
- MPEG-1 Layer III (mp3)

## Δείγματα σήματος εισόδου



Το διάγραμμα του σχήματος δίνει τις τιμές των πρώτων 1152 δειγμάτων ενός σήματος μουσικής.

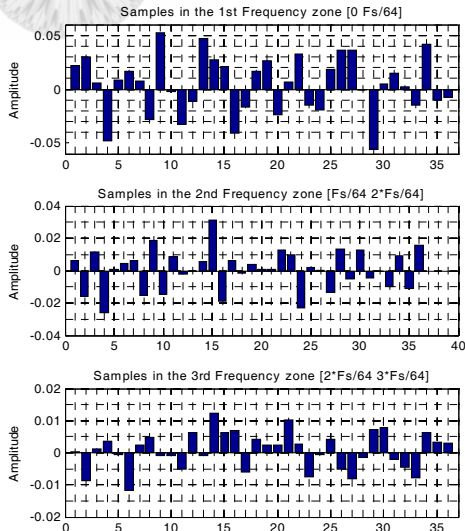
Τα δείγματα αυτά θα πρέπει να κωδικοποιηθούν με κάποια bits το καθένα. Στο τμήμα του σήματος που αντιπροσωπεύεται από τα δείγματα του σχήματος υπάρχουν συχνότητες που δεν είναι ακουστές (ακουστική ικανότητα) και συχνότητες που είναι ισχυρές και καλύπτουν γειτονικές τους (φασματική κάλυψη).

Τα δύο αυτά φαινόμενα λειτουργούν αθροιστικά (για παράδειγμα η ακουστική ικανότητα στη συχνότητα 8kHz απαιτεί ένταση τουλάχιστον 20 dB για να είναι ακουστή. Αν εξαιτίας της παρουσίας μιας ισχυρής συχνότητας στα 6kHz έχουμε φασματική κάλυψη ύψους 15 dB στη συχνότητα 8kHz τότε συνολικά η συχνότητα 8kHz θα είναι ακουστή αν έχει ένταση μεγαλύτερη από 35 dB.



- Εισαγωγή
- Βασικό Διάγραμμα
- Ψυχοακουστική
- Κωδικοποίηση ήχου με το πρότυπο MPEG-1
- ★ MPEG-1 Layer II
- MPEG-1 Layer III (mp3)

## Διαχωρισμός δειγμάτων σε 32 ζώνες συχνότητας

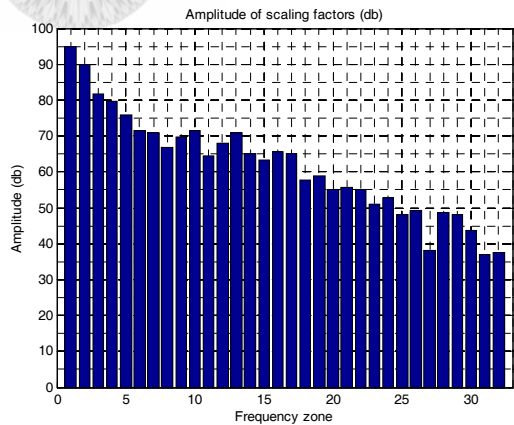


Το διάγραμμα του σχήματος δίνει τις 'τιμές' δειγμάτων που αντιστοιχούν στις 3 πρώτες ζώνες συχνότητας.

- ◇ Κάθε ζώνη συχνότητας καλύπτει ένα διάστημα  $F_s/64$  ( $F_s$  = Sampling frequency)
- ◇ Το δείγμα με την ισχυρότερη ένταση σε κάθε ζώνη αποτελεί το scaling factor της ζώνης αυτής
- ◇ Στο παράδειγμα του σχήματος έχουμε:
  - ◇ Ζώνη 1 => scaling factor το δείγμα 29
  - ◇ Ζώνη 2 => scaling factor το δείγμα 15
  - ◇ Ζώνη 3 => scaling factor το δείγμα 14

- Εισαγωγή
- Βασικό Διάγραμμα
- Ψυχοακουστική
- Κωδικοποίηση ήχου με το πρότυπο MPEG-1
- ★ MPEG-1 Layer II
- MPEG-1 Layer III (mp3)

## Ένταση των scaling factors (σε db)

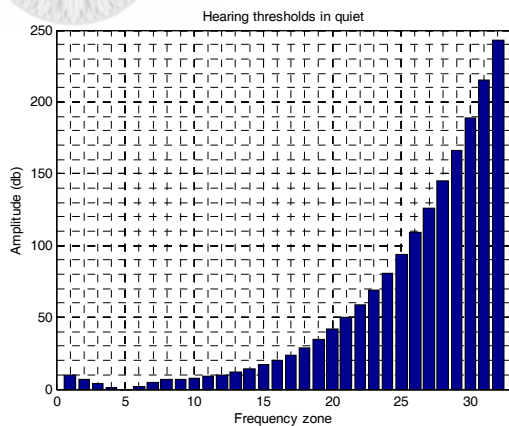


Το διάγραμμα του σχήματος δίνει τις 'τιμές' των scaling factors σε db.

- ◇ Οι τιμές αυτές θα συγκριθούν με τα επίπεδα ακουστικής ικανότητας + φασματικής κάλυψης.
- ◇ Για όσες ζώνες τα scaling factors έχουν ένταση μεγαλύτερη από το άθροισμα των ανωτέρω επιπέδων τα δείγματα στις ζώνες αυτές θα μεταδοθούν ενώ στις υπόλοιπες θα αγνοηθούν.

- Εισαγωγή
- Βασικό Διάγραμμα
- Ψυχοακουστική
- Κωδικοποίηση ήχου με το πρότυπο MPEG-1
- ★ MPEG-1 Layer II
- MPEG-1 Layer III (mp3)

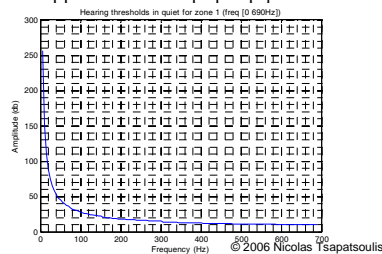
## Ακουστική ικανότητα ανά ζώνη συχνότητας



Το διάγραμμα του σχήματος δίνει την ακουστική ικανότητα ανά ζώνη συχνότητας.

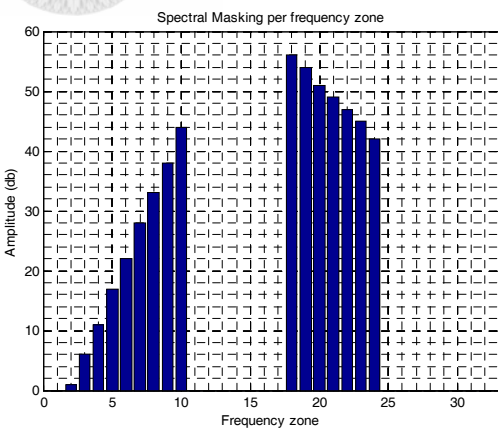
Τα επίπεδα ακουστικής ικανότητας λαμβάνονται ως οι ελάχιστες τιμές σε κάθε ζώνη.

Για παράδειγμα η ακουστική ικανότητα στη ζώνη 1 (συχνοτήτες 0-690 Hz) είναι καλύτερη στη συχνότητα 690 Hz. Επομένως το επίπεδο ακουστικότητας λαμβάνεται σε αυτή την τιμή



- Εισαγωγή
- Βασικό Διάγραμμα
- Ψυχοακουστική
- Κωδικοποίηση ήχου με το πρότυπο MPEG-1
- ★ MPEG-1 Layer II
- MPEG-1 Layer III (mp3)

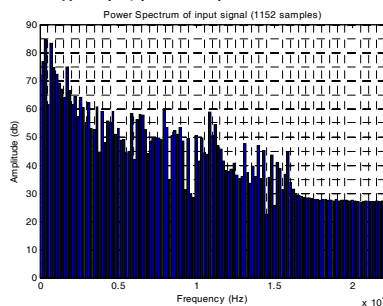
## Υπολογισμός Φασματικής Κάλυψης



Το διάγραμμα του σχήματος δίνει την φασματική κάλυψη ανά ζώνη συχνότητας με βάση τα 1152 δείγματα.

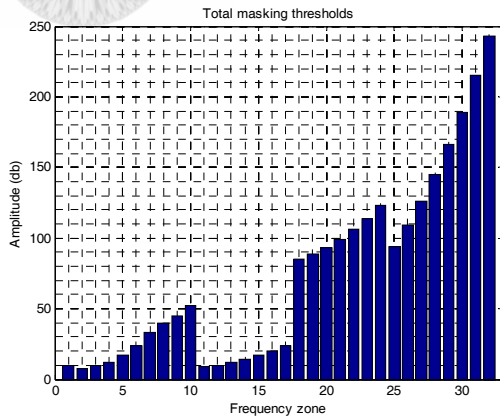
◇ Ο υπολογισμός της φασματικής κάλυψης έγινε με τη βοήθεια του μετασχηματισμού FFT 1024 σημείων.

◇ Φασματική κάλυψη δίνουν μόνο οι συχνότητες με ένταση πάνω από 60 db



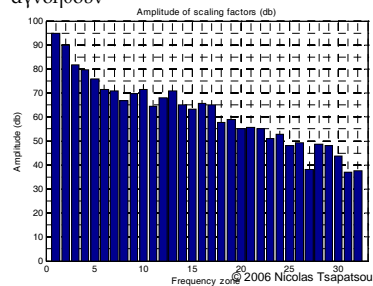
- Εισαγωγή
- Βασικό Διάγραμμα
- Ψυχοακουστική
- Κωδικοποίηση ήχου με το πρότυπο MPEG-1
- ★ MPEG-1 Layer II
- MPEG-1 Layer III (mp3)

## Συνολική κάλυψη



Το διάγραμμα του σχήματος δίνει το συνολικό κατώφλι ακουστότητας (ακουστική ικανότητα + φασματική κάλυψη) ανά ζώνη συχνοτήτων με βάση τα 1152 δείγματα.

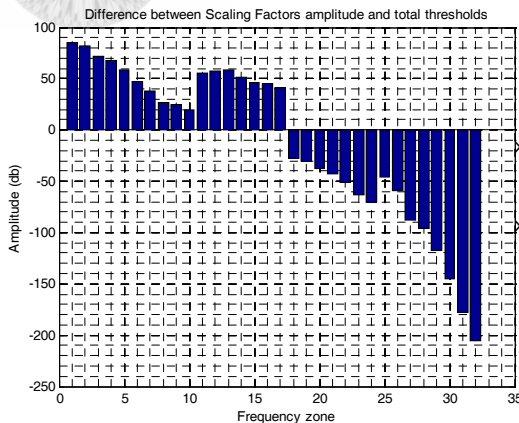
Οι ζώνες συχνοτήτων των οποίων η ένταση των scaling factors υπερβαίνει το συνολικό κατώφλι ακουστότητας θα κωδικοποιηθούν ενώ τα δείγματα στις υπόλοιπες θα αγνοηθούν



© 2006 Nicolas Tsapatsoulis

- Εισαγωγή
- Βασικό Διάγραμμα
- Ψυχοακουστική
- Κωδικοποίηση ήχου με το πρότυπο MPEG-1
- ★ MPEG-1 Layer II
- MPEG-1 Layer III (mp3)

## Ζώνες που κωδικοποιούνται



Οι ζώνες συχνοτήτων των οποίων η ένταση των scaling factors υπερβαίνει το συνολικό κατώφλι ακουστότητας θα κωδικοποιηθούν ενώ τα δείγματα στις υπόλοιπες θα αγνοηθούν

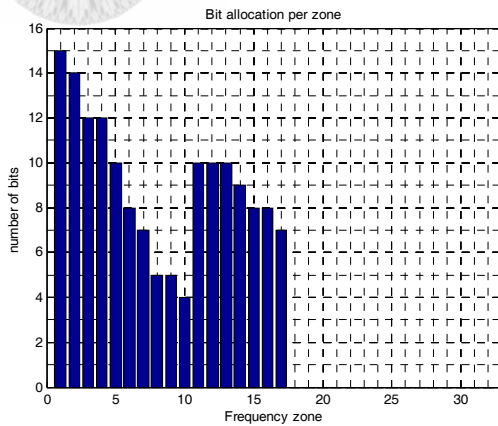
Από το διπλανό σχήμα φαίνεται ότι θα κωδικοποιηθούν τα δείγματα στις ζώνες συχνοτήτων 1-17.

Τα bits που θα διατεθούν για την κωδικοποίηση των δειγμάτων σε κάθε ζώνη υπολογίζονται διαγράφοντας τις θετικές τιμές του διαγράμματος με το 6 και στρογγυλοποιώντας προς την μεγαλύτερη ακέραια τιμή

© 2006 Nicolas Tsapatsoulis

- Εισαγωγή
- Βασικό Διάγραμμα
- Ψυχοακουστική
- Κωδικοποίηση ήχου με το πρότυπο MPEG-1
- ★ MPEG-1 Layer II
- MPEG-1 Layer III (mp3)

## Επιμερισμός bit ανά ζώνη συχνοτήτων



Στο σχήμα δίνεται ο αριθμός των bits με τα οποία θα κωδικοποιηθούν τα δείγματα σε κάθε ζώνη συχνοτήτων

Τα δείγματα σε 15 ζώνες δεν μεταδίδονται

Με περισσότερα bits κωδικοποιούνται τα δείγματα στη ζώνη 1 (15 bits) και με λιγότερα τα δείγματα στη ζώνη 10 (4 bits)

Συνολικά για την κωδικοποίηση όλων των δειγμάτων χρειάζονται:

$$36 \cdot (15+14+12+12+\dots+7) = 5544 \text{ bits}$$

- Εισαγωγή
- Βασικό Διάγραμμα
- Ψυχοακουστική
- Κωδικοποίηση ήχου με το πρότυπο MPEG-1
- ★ MPEG-1 Layer II
- MPEG-1 Layer III (mp3)

## Υπολογισμός βαθμού συμπίεσης

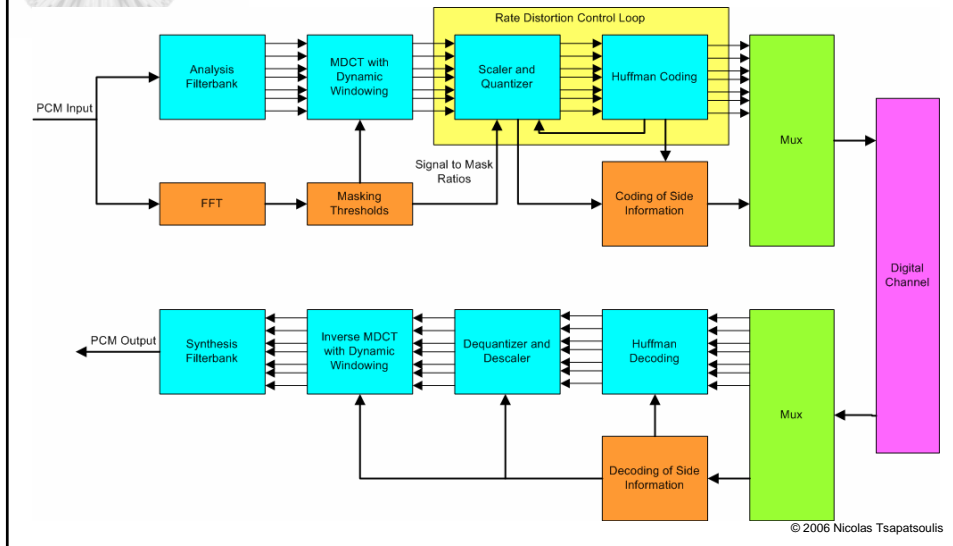


- ◇ Στο παράδειγμα μας έχουμε  $36 \times 32 = 1152$  δείγματα
  - ◇ Σε ασυμπίεστη μορφή απαιτούνται 2304 bytes (16 bits / δείγμα)
  - ◇ Με βάση τον αλγόριθμο που περιγράψαμε έχουμε:
    - ◇ 32 bits επικεφαλίδα
    - ◇ 128 bits για καθορισμό του bit allocation
    - ◇  $17 \times 6 = 102$  bits για κωδικοποίηση των 17 scaling factors που αντιστοιχούν στις ζώνες συχνοτήτων οι οποίες είναι ακουστές
    - ◇  $154 \times 36 = 5544$  bits για κωδικοποίηση των δειγμάτων (17 μη μηδενικές ομάδες δειγμάτων με αριθμό bits όπως υπολογίστηκε νωρίτερα)
  - ◇ Σύνολο: 5806 bits => 726 bytes
  - ◇ Συμπίεση περίπου 3.17:1

Header (32)	CRC (0,16)	Bit allocation (128-256)	Scale factors (0-384)	Samples	Ancillary data
-------------	------------	--------------------------	-----------------------	---------	----------------

# MPEG-1 Layer III (mp3)

- ☑ Εισαγωγή
- ☑ Βασικό Διάγραμμα
- ☑ Ψυχοακουστική
- ☑ Κωδικοποίηση ήχου με το πρότυπο MPEG-1
- ☑ MPEG-1 Layer II
- ★ MPEG-1 Layer III (mp3)



## Διαφορές του MPEG-1, Layer III σε σχέση με το MPEG-1, Layer II

- ☑ Εισαγωγή
- ☑ Βασικό Διάγραμμα
- ☑ Ψυχοακουστική
- ☑ Κωδικοποίηση ήχου με το πρότυπο MPEG-1
- ☑ MPEG-1 Layer II
- ★ MPEG-1 Layer III (mp3)



- ◇ Οι βασικές διαφορές του MPEG-1, Layer III (mp3) σε σχέση με το MPEG-1, Layer II είναι:
  - ◇ Η χρήση του MDCT (Modified Discrete Cosine Transform) για υποδιαίρεση του φάσματος συχνοτήτων σε 576 ζώνες αντί για 32 με στόχο την καλύτερη εκμετάλλευση της φασματικής κάλυψης και τις διακριτικής ικανότητας
    - ◇ Χρησιμοποιούνται 18 συντελεστές MDCT για κάθε μια από τις 32 ζώνες
    - ◇ Τελικά δημιουργούνται συνολικά  $32 \times 18 = 576$  ζώνες στο φάσμα συχνοτήτων
  - ◇ Δυναμικός επιμερισμός bits ανά ζώνη συχνοτήτων ώστε να μπορεί να επιτευχθεί το ζητούμενο από το χρήστη bit rate
  - ◇ Χρήση κωδικοποίησης Huffman επί της κωδικοποίησης των δειγμάτων (codewords) για να επιτευχθεί περαιτέρω συμπίεση.

- Εισαγωγή
- Βασικό Διάγραμμα
- Ψυχοακουστική
- Κωδικοποίηση ήχου με το πρότυπο MPEG-1
- MPEG-1 Layer II
- ★ MPEG-1 Layer III (mp3)

## Χαρακτηριστικά του MPEG-1, Layer III



- ◇ Το μέγεθος και η ποιότητα των αρχείων ήχου καθορίζονται από τον ρυθμό μετάδοσης των δεδομένων (bit rate).
  - ◇ Το πρότυπο MP3 υποστηρίζει αρκετούς προκαθορισμένους ρυθμούς μετάδοσης δεδομένων.
  - ◇ Τα αρχεία με μεγαλύτερο bit rate θα ακούγονται καλύτερα από αυτά με μικρότερο bit rate.

Συχρότητα	Chan-nels	Bit rate	Συμπίεση	Ποιότητα
2.5 kHz	Mono	8 kbps	96:1	Τηλεφωνική
4.5 kHz	Mono	16 kbps	48:1	Βραχέα (SW)
11 kHz	Stereo	56..64 kbps	26 .. 24:1	FM
15 kHz	Stereo	96 kbps	16:1	CD like
>15 kHz	Stereo	112..128 kbps	14..12:1	CD

© 2006 Nicolas Tsapatsoulis

- Εισαγωγή
- Βασικό Διάγραμμα
- Ψυχοακουστική
- Κωδικοποίηση ήχου με το πρότυπο MPEG-1
- MPEG-1 Layer II
- ★ MPEG-1 Layer III (mp3)

## Χαρακτηριστικά του MPEG-1, Layer III (II)



- ◇ Συμπίεση αρχείων ήχου μπορεί να φτάσει και το 85%
- ◇ Με ρυθμό μετάδοσης (bit rate) άνω των 128 Kbps η διαφορά από το πρωτότυπο CD δεν γίνεται αντιληπτή.
- ◇ Το αποτέλεσμα είναι να έχουμε μουσικά αρχεία 4-6 λεπτών που κανονικά θα καταλάμβαναν χώρο 40-70 MB, να πάνουν τώρα χώρο μόλις 3-7MB.

© 2006 Nicolas Tsapatsoulis