

Ανώτατο Τεχνολογικό Εκπαιδευτικό Ίδρυμα Κρήτης
Παράρτημα Ρεθύμνου
Τμήμα Μουσικής Τεχνολογίας και Ακουστικής

Τίτλος Πτυχιακής Εργασίας

Εκτίμηση της διαφοράς αποτελεσματικότητας κωδικοποίησης του προτύπου MPEG-1 Layer III (MP3), για διαφορετικά είδη μουσικής, με τη χρήση listening tests (μέθοδος Semantic Differential).

Σπουδάστρια:

Αρόνη Ιωάννα (Α.Μ. 655)

Επιβλέπων καθηγητής:

κ. Παπαδάκης Νικόλαος

Ρέθυμνο, 2011

Περιεχόμενα

Περιεχόμενα	1
Περίληψη.....	5
Ευχαριστίες	5
Πρόλογος.....	6
ΚΕΦΑΛΑΙΟ 1 ^ο : ΑΝΘΡΩΠΙΝΟ ΣΥΣΤΗΜΑ ΑΚΟΗΣ ΚΑΙ ΨΥΧΟΑΚΟΥΣΤΙΚΗ	7
1.1 Ανθρώπινο Ακουστικό Σύστημα	7
1.1.1 Εξωτερικό Αυτί.....	8
1.1.2 Μέσο Αυτί.....	8
1.1.3 Εσωτερικό Αυτί.....	9
1.2 Ψυχοακουστική.....	10
1.3 Critical Bands.....	10
1.4 Κατώφλι Ακουστότητας.....	14
1.5 Masking.....	14
1.5.1 Simultaneous Masking.....	15
1.5.2 Temporal Masking	16
ΚΕΦΑΛΑΙΟ 2 ^ο : ΒΑΣΙΚΕΣ-ΕΙΣΑΓΩΓΙΚΕΣ ΕΝΝΟΙΕΣ ΓΙΑ ΤΗ ΣΥΜΠΙΕΣΗ ΣΗΜΑΤΩΝ.....	17
2.1 Απωλεστική Συμπίεση	17
2.2 Παλμοκωδική Διαμόρφωση (PCM).....	17
2.3 Δειγματοληψία	18
2.4 Κβαντοποίηση.....	18
2.5 Κωδικοποίηση.....	19
2.6 Κωδικοποίηση Εντροπίας.....	19
2.6.1 Στατιστική Κωδικοποίηση	20
2.7 Κωδικοποίηση Πηγής	21
2.7.1 Κωδικοποίηση Μετασχηματισμού.....	21
2.7.2 Διαφορική ή Προβλεπτική Κωδικοποίηση	22
ΚΕΦΑΛΑΙΟ 3 ^ο : ΠΡΟΤΥΠΟ MPEG-1 LAYER 3 (MP3).....	23
3.1 Μια Γενική Επισκόπηση του Προτύπου.....	23
3.1.1 Το Πρότυπο MPEG-1.....	23
3.1.2 Μείωση του Όγκου των Δεδομένων κατά 12	25
3.1.3 Ελευθερία Εκτέλεσης.....	25
3.1.4 Bitrate	26
3.1.5 Συχνότητα Δειγματοληψίας	27
3.1.6 Channel Modes.....	27

3.2 Ανατομία Ενός Αρχείου MP3	28
3.2.1 Μορφή των Frames	28
3.2.2 ID3	43
3.3 Κωδικοποίηση	43
3.3.1 Analysis Polyphase Filterbank	45
3.3.2 Modified Discrete Transform (MDCT)	46
3.3.3 FFT	48
3.3.4 Psychoacoustic Model	48
3.3.5 Nonuniform Quantization	55
3.3.6 Huffman Encoding	58
3.3.7 Coding of Side Information	58
3.3.8 Bitstream Formatting and CRC word generation	59
3.4 Αποκωδικοποίηση	59
3.4.1 Synchronization and Error Checking	60
3.4.2 Huffman Decoding and Huffman info Decoding	60
3.4.3 Scalefactor decoding	61
3.4.4 Requantization	62
3.4.5 Reordering	63
3.4.6 Stereo Processing	63
3.4.7 Alias Reduction	64
3.4.8 Inverse Modified Discrete Cosine Transform (IMDCT) and Overlapping	66
3.4.9 Frequency Inversion	66
3.4.10 Synthesis Polyphase Filterbank	67
3.5 Επίδραση του Αλγορίθμου στην Ποιότητα της Μουσικής	67
3.5.1 Φασματικά Χαρακτηριστικά	68
3.5.2 Δυναμικά Χαρακτηριστικά	68
3.5.3 Θόρυβος και Παραμόρφωση	68
ΚΕΦΑΛΑΙΟ 4 ^ο : ΜΕΘΟΔΟΣ SEMANTIC DIFFERENTIAL	70
4.1 Σύντομο Ιστορικό	71
4.2 Ανάλυση της Μεθόδου	72
4.2.1 Η Λογική της Μεθόδου	72
4.2.2 Δομή EPA	73
4.2.3 Επιλογή Αντικειμένων	74
4.2.4 Επιλογή Κλιμάκων	75
4.2.5 Χρήση Εφτά Σημείων	76
4.2.6 Αξιολόγηση της Μεθόδου	77
4.2.7 Διαδικασία Ανάλυσης	85
4.3 Πλεονεκτήματα της Μεθόδου	86

ΚΕΦΑΛΑΙΟ 5 ^ο : ΣΧΕΔΙΑΣΜΟΣ ΤΗΣ ΠΕΙΡΑΜΑΤΙΚΗΣ ΔΙΑΔΙΚΑΣΙΑΣ-ΜΕΘΟΔΟΛΟΓΙΑ	88
5.1 Ερευνητικά Ερωτήματα	88
5.2 Επιλογή Κλιμάκων.....	89
5.2.1 Αριθμός Κλιμάκων.....	89
5.2.2 Συμφωνία με τη Δομή ΕΡΑ.....	90
5.2.3 Φύση του Αντικειμένου	91
5.2.4 Ερμηνεία των κλιμάκων από τα υποκείμενα	91
5.3 Επιλογή Ακουσμάτων	92
5.3.1 Rock	92
5.3.2 Jazz.....	92
5.3.3 Ηλεκτρονική Μουσική.....	93
5.3.4 Κλασσική Μουσική (Ορχήστρα)	93
5.3.5 Solo Όργανο	93
5.4 Επιλογή Υποκειμένων.....	94
5.5 Εξοπλισμός και Λογισμικό	95
5.6 Πειραματική Διαδικασία.....	96
ΚΕΦΑΛΑΙΟ 6 ^ο : ΑΠΟΤΕΛΕΣΜΑΤΑ ΤΗΣ ΕΡΕΥΝΑΣ-ΣΥΓΚΡΙΣΕΙΣ.....	98
6.1 Rock	98
6.1.1 Rock με bitrate 96kbps.....	98
6.1.2 Rock με bitrate 160kbps.....	99
6.1.3 Rock με bitrate 320kbps.....	100
6.2 Jazz.....	101
6.2.1 Jazz με bitrate 96kbps	101
6.2.2 Jazz με bitrate 160kbps	101
6.2.3 Jazz με bitrate 320kbps	102
6.3 Ηλεκτρονική.....	103
6.3.1 Ηλεκτρονική με bitrate 96kbps	103
6.3.2 Ηλεκτρονική με bitrate 160kbps	104
6.3.3 Ηλεκτρονική με bitrate 320kbps	104
6.4 Κλασσική Μουσική (με ορχήστρα)	105
6.4.1 Κλασσική με bitrate 96kbps.....	105
6.4.2 Κλασσική με bitrate 160kbps.....	106
6.4.3 Κλασσική με bitrate 320kbps.....	106
6.5 Solo όργανο.....	107
6.5.1 Solo όργανο με bitrate 96kbps	107
6.5.2 Solo όργανο με bitrate 160kbps	108
6.5.3 Solo όργανο με bitrate 320kbps	108

6.6 Συγκρίσεις Αποτελεσμάτων.....	109
6.6.1 Rock	109
6.6.2 Jazz.....	111
6.6.3 Ηλεκτρονική μουσική	113
6.6.4 Κλασσική (με ορχήστρα)	115
6.6.5 Solo όργανο.....	117
6.6.6 Bitrate στα 96kbps.....	119
6.6.7 Bitrate στα 160kbps.....	121
6.6.8 Bitrate στα 320kbps.....	123
ΚΕΦΑΛΑΙΟ 7 ^ο : ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΠΕΡΙΟΡΙΣΜΟΙ ΤΗΣ ΕΡΕΥΝΑΣ.....	126
7.1 Συμπεράσματα	126
7.2 Περιορισμοί της Έρευνας	130
Βιβλιογραφία.....	132
Ελληνική	132
Ξενόγλωσση.....	132
ΠΑΡΑΡΤΗΜΑ Α	138
ΠΑΡΑΡΤΗΜΑ Β	146

Περίληψη

Το αντικείμενο μελέτης της παρούσας πτυχιακής εργασίας είναι η αποτελεσματικότητα της κωδικοποίησης του προτύπου MPEG-1 Layer III για πέντε διαφορετικά είδη μουσικής. Στο θεωρητικό της μέρος, αρχικά, παρουσιάζεται ο τρόπος λειτουργίας του ανθρώπινου ακουστικού συστήματος και αναφέρονται ορισμένα ψυχοακουστικά φαινόμενα τα οποία λαμβάνουν χώρα σε αυτό και σχετίζονται με τον τρόπο λειτουργίας του προτύπου. Επίσης, γίνεται εκτενής ανάλυση του αλγορίθμου του MP3 καθώς και της μεθοδολογίας με την οποία κατασκευάστηκαν τα listening tests. Στο πειραματικό μέρος, αναλύεται η κατασκευή των listening tests και ο τρόπος διεξαγωγής της πειραματικής διαδικασίας, καθώς και τα αποτελέσματά της.

Ευχαριστίες

Εδώ, θα ήθελα να ευχαριστήσω όλους τους ανθρώπους που βοήθησαν προκειμένου να ολοκληρωθεί η παρούσα πτυχιακή εργασία. Πρώτο απ' όλους, ευχαριστώ τον επιβλέποντα καθηγητή μου κ. Παπαδάκη Νίκο, για τη βοήθεια και συμπαράσταση που έδειξε καθ' όλη τη διάρκεια της εργασίας. Επίσης, ευχαριστώ θερμά την καθηγήτρια κ. Τζεδάκη Κατερίνα για την παραχώρηση του εργαστηρίου «Μουσικής Διάδρασης και Πολυφωνίας», προκειμένου να διεξαχθούν τα listening tests. Συνεχίζοντας, θα ήθελα να ευχαριστήσω την κ. Βαρβαντάκη Σταυρούλα, βιβλιοθηκονόμο του τμήματος μας, για τη βοήθειά της στην αναζήτηση βιβλιογραφίας. Τέλος, ευχαριστώ θερμά όλα τα άτομα τα οποία προσφέρθηκαν εθελοντικά να συμπεριληφθούν στο δείγμα της έρευνας.

Πρόλογος

Το πρότυπο MPEG-1 Layer 3 έχει κατακτήσει μια πολύ σημαντική θέση στην καθημερινότητα των ανθρώπων, μιας και είναι ο πιο διαδεδομένος αλγόριθμος συμπίεσης ηχητικών σημάτων. Λόγω των προτερημάτων του, όπως ο μικρός απαιτούμενος χώρος αποθήκευσης και η συμβατότητα του με τα περισσότερα προγράμματα μουσικής αναπαραγωγής, έχει σχεδόν αντικαταστήσει όλες τις υπόλοιπες μορφές μουσικών αρχείων, ενώ πάνω του έχει κατασκευαστεί μια ολόκληρη μουσική βιομηχανία. Δυστυχώς, παρά τα πλεονεκτήματά του, το πρότυπο, δεδομένο ότι είναι αλγόριθμος απωλεστικής συμπίεσης, μειονεκτεί σχετικά στην ποιότητα του αναπαραγόμενου ήχου. Λόγω των διαδικασιών που χρησιμοποιεί για τη μείωση του όγκου των δεδομένων, αναπόφευκτα, μειώνει τα ποιοτικά χαρακτηριστικά του συμπιεζόμενου ηχητικού σήματος.

Αυτή τη μείωση καλείται να μελετήσει η παρούσα πτυχιακή εργασία και λόγω της μεγάλης διάδοσης του προτύπου, η ενασχόληση με το συγκεκριμένο αντικείμενο θεωρήθηκε ωφέλιμη και ενδιαφέρουσα.

Η εργασία χωρίζεται σε δύο μέρη: το θεωρητικό και το ερευνητικό.

Στο θεωρητικό κομμάτι, αναφέρεται η λειτουργία του ακουστικού μας συστήματος και ορισμένα ψυχοακουστικά φαινόμενα τα οποία σχετίζονται με το πρότυπο (κεφάλαιο 1^ο) και αναλύονται ορισμένες βασικές έννοιες για τη συμπίεση σημάτων (κεφάλαιο 2^ο). Στη συνέχεια, αναλύεται διεξοδικά ο τρόπος λειτουργίας του προτύπου (κεφάλαιο 3^ο) και η μέθοδος Semantic Differential, σύμφωνα με την οποία κατασκευάστηκαν τα listening tests (κεφάλαιο 4^ο).

Στο πειραματικό μέρος, περιγράφεται η μεθοδολογία της έρευνας (κεφάλαιο 5^ο), τα αποτελέσματα που προέκυψαν (κεφάλαιο 6^ο), καθώς και τα συμπεράσματα και οι περιορισμοί της (κεφάλαιο 7^ο).

ΚΕΦΑΛΑΙΟ 1^ο: ΑΝΘΡΩΠΙΝΟ ΣΥΣΤΗΜΑ ΑΚΟΗΣ ΚΑΙ ΨΥΧΟΑΚΟΥΣΤΙΚΗ

Εισαγωγή 1^ο Κεφαλαίου

Στο κεφάλαιο αυτό γίνεται αναφορά στις βασικές λειτουργίες του ακουστικού μας συστήματος και ανάλυση των ψυχοακουστικών φαινομένων τα οποία σχετίζονται με το πρότυπο MPEG-1 Layer 3, τα οποία είναι το masking και τα critical bands καθώς και το κατώφλι ακουστότητας.

Σκοπός του κεφαλαίου αυτού είναι η εισαγωγή σε βασικές έννοιες και λειτουργίες της ανθρώπινης ακοής καθώς και η κατανόηση των φαινομένων που αξιοποιεί το πρότυπο προκειμένου να επιτύχει το επιθυμητό ακουστικό αποτέλεσμα με τη συμπίεση σημάτων ήχου.

1.1 Ανθρώπινο Ακουστικό Σύστημα

Η ανθρώπινη ακοή είναι η πιο ανεπτυγμένη μας αίσθηση και τα όργανα της έχουν αξιοθαύμαστα χαρακτηριστικά. Για παράδειγμα, το αυτί μας έχει πολύ μεγάλο δυναμικό εύρος μιας και το κατώφλι του πόνου για την ένταση των ήχων είναι στα 120dB_{SPL}, δηλαδή 1.000.000.000.000 φορές μεγαλύτερη ένταση από το κατώφλι της ακοής που βρίσκεται στα 0dB_{SPL}. Όσον αφορά τη συχνότητα, το αυτί μας αντιλαμβάνεται ένα εύρος συχνοτήτων το οποίο, για τους περισσότερους ανθρώπους είναι από τα 20Hz έως τα 20kHz.

Το ανθρώπινο ακουστικό σύστημα είναι ένας περίπλοκος μηχανισμός με συγκεκριμένα χαρακτηριστικά και ιδιότητες. Το σύστημα αυτό αντιλαμβάνεται τους ήχους του περιβάλλοντός μας, απορρίπτει ήχους που δε μας ενδιαφέρουν, προσαρμόζεται σε πολύ δυνατούς ήχους και έχει μηχανισμό αυτοπροστασίας για την αποφυγή βλάβης. Στο σημείο αυτό θα αναφερθεί η λειτουργία του ακουστικού μας συστήματος περιληπτικά

μιας και είναι απαραίτητο να γνωρίζουμε τις βασικές λειτουργίες προκειμένου να γίνουν κατανοητά τα ψυχοακουστικά φαινόμενα που θα αναλυθούν αργότερα.

Τα τμήματα που συνθέτουν το σύστημα ακοής μας είναι τρία: το εξωτερικό αυτί, το μέσο αυτί και το εσωτερικό αυτί. Καθένα από αυτά αποτελείται από διάφορα μικρά τμήματα τα οποία επιτελούν συγκεκριμένες διεργασίες.

1.1.1 Εξωτερικό Αυτί

Το εξωτερικό αυτί αποτελείται από το περύγιο, το ακουστικό κανάλι και το τύμπανο.

Το περύγιο είναι υπεύθυνο για τη συλλογή των ήχων από το περιβάλλον και συμβάλλει σημαντικά στον εντοπισμό της θέσης μιας ηχητικής πηγής. Λόγω της θέσης και του σχήματός του, διαφοροποιεί σημαντικά την ηχητική πληροφορία που εισέρχεται σε αυτό.

Το ακουστικό κανάλι έχει μήκος 3 εκατοστά και έχει την ικανότητα να αυξάνει την ηχηρότητα των συλλεγμένων ήχων. Είναι ουσιαστικά ένας φυσικός ενισχυτής με ικανότητα ενίσχυσης κοντά στα 10 με 15dB_{SPL}. Είναι επίσης υπεύθυνο για την προστασία του μέσου αυτιού από εξωτερικούς κινδύνους (εξωτερικά σώματα, αλλαγές στη θερμοκρασία κλπ).

1.1.2 Μέσο Αυτί

Το μέσο αυτί αποτελείται από το τύμπανο, τα οστά και τους μύες του μέσου αυτιού, την κοιλότητα και την ευσταχιανή σάλπιγγα (Παπαδάκης, 2009, 11).

Το τύμπανο είναι μια μεμβράνη η οποία έχει την ιδιότητα να αυξομειώνει την καμπυλότητα και τη σκληράδα της ανάλογα με την ένταση του ήχου που δέχεται. Η ικανότητα αυτή ονομάζεται aural reflex και προστατεύει το τύμπανο από την παραμόρφωση σε περίπτωση απότομης αύξησης της έντασης του ήχου.

Τα οστά του μέσου αυτιού είναι τρία: η σφύρα, ο άκμονας και ο αναβολέας. Ενώνονται μεταξύ τους με δύο μύες. Η σφύρα είναι ενωμένη και με το τύμπανο και ο αναβολέας με το οβάλ παράθυρο. Τα οστά αυτά είναι υπεύθυνα για τη μεταφορά της ενέργειας από την ταλάντωση του τυμπάνου στο οβάλ παράθυρο και από εκεί στον κοχλία. Οι μύες που

βρίσκονται ενδιάμεσα επιτελούν διάφορες λειτουργίες όπως ο έλεγχος της έντασης για πολύ δυνατούς ήχους χαμηλής συχνότητας, ελάττωση της αντίληψης των ήχων που παράγει το ανθρώπινο σώμα κ.α..

Η ευσταχιανή σάλπιγγα συνδέει το μέσο αυτί με το άνω μέρος του λαιμού και η βασική της λειτουργία είναι η εξίσωση της στατικής πίεσης του αέρα στο μέσο αυτί με την εξωτερική ατμοσφαιρική πίεση προκειμένου να λειτουργεί σωστά το τύμπανο. Μια ακόμα λειτουργία της είναι η απαγωγή υγρών σε περίπτωση μόλυνσης του αυτιού.

1.1.3 Εσωτερικό Αυτί

Το εσωτερικό αυτί αποτελείται από δύο μέρη: το μεμβρανώδη λαβύρινθο και τον οστέινο λαβύρινθο. Η πολυπλοκότητα της κατασκευής αυτού του συστήματος κάνει τη διάκριση των μερών του σχεδόν αδύνατη. Οι δυνατότητες του εσωτερικού αυτιού είναι πολλές και εκτείνονται από την αντίληψη των σωματικών κινήσεων μέχρι και την ικανότητα ανάλυσης κυματικών μετακινήσεων των ηχητικών πιέσεων.

Ο μεμβρανώδης λαβύρινθος περιλαμβάνει δύο όργανα: το αιθουσαίο και τον κοχλία. Το αιθουσαίο είναι υπεύθυνο για την αντίληψη της στάσης του σώματός στους τρεις άξονες του χώρου. Ο κοχλίας είναι ίσως το σημαντικότερο όργανο της ακοής μας και αποτελείται από πολλά μικρότερα τμήματα με συγκεκριμένες λειτουργίες. Τα σημαντικότερα από αυτά είναι η βασική μεμβράνη, τα τριχοειδή κύτταρα και το όργανο του Corti.

Το ηχητικό σήμα ταξιδεύει από το τύμπανο, μέσω των προαναφερθέντων οργάνων και καταλήγει στο ακουστικό νεύρο από όπου και στέλνεται στον εγκέφαλο. Η αντίληψη της συχνότητας του ήχου προέρχεται από τη συχνότητα στην οποία συντονίζεται κάθε φορά η βασική μεμβράνη. Καθώς τα κύματα «ταξιδεύουν» κατά μήκος της μεμβράνης προκαλούν ταλάντωση αυτής, της οποίας το πλάτος εξαρτάται από την ένταση του και την συχνότητα του ερεθίσματος. Συγκεκριμένα, οι περιοχές της βασικής μεμβράνης κοντά στην βάση του κοχλίου παρουσιάζουν μέγιστη διέγερση στις υψηλές συχνότητες, οι περιοχές κοντά στην κορυφή στις χαμηλές, ενώ οι μεσαίες περιοχές στις μεσαίες συχνότητες. Έτσι ένα ερέθισμα 10KHz θα διεγείρει κυρίως την βασική μεμβράνη κοντά

στην βάση του κοχλίου, των 2KHz στη μεσότητα του κοχλίου, και των 200Hz στην κορυφή του κοχλίου. Αυτή η κατανομή των συχνοτήτων πάνω στην βασική μεμβράνη είναι λογαριθμική. Χάριν σ' αυτές τις μηχανικές ιδιότητες της βασικής μεμβράνης, επιτυγχάνεται η διάκριση των συχνοτήτων.

1.2 Ψυχοακουστική

Η Ψυχοακουστική είναι ένας όρος ο οποίος περιλαμβάνει τη μελέτη κατασκευής και λειτουργίας του αυτιού, τις διαδρομές κίνησης του ήχου, την αντίληψή του καθώς και τις μεταξύ τους σχέσεις. Σαν ορισμό, λοιπόν, θα μπορούσαμε να πούμε πως η Ψυχοακουστική είναι ο τομέας της επιστήμης ο οποίος εξετάζει τη σύνδεση ανάμεσα στις φυσικές ιδιότητες των ακουστικών ερεθισμάτων και τις φυσιολογικές και ψυχολογικές αντιδράσεις που αυτά προκαλούν στον άνθρωπο (*Pohlmann, 2005, 316*). Μέσω αυτής εξάγεται το συμπέρασμα πως η σημασία που δίνεται στον ήχο από τους ανθρώπους προέρχεται αποκλειστικά και μόνο από την αντίληψη που έχουν οι άνθρωποι γι' αυτόν. Διαφορετικά, ο ήχος θα ήταν μια αδιάφορη και ασαφής ακαδημαϊκή έννοια.

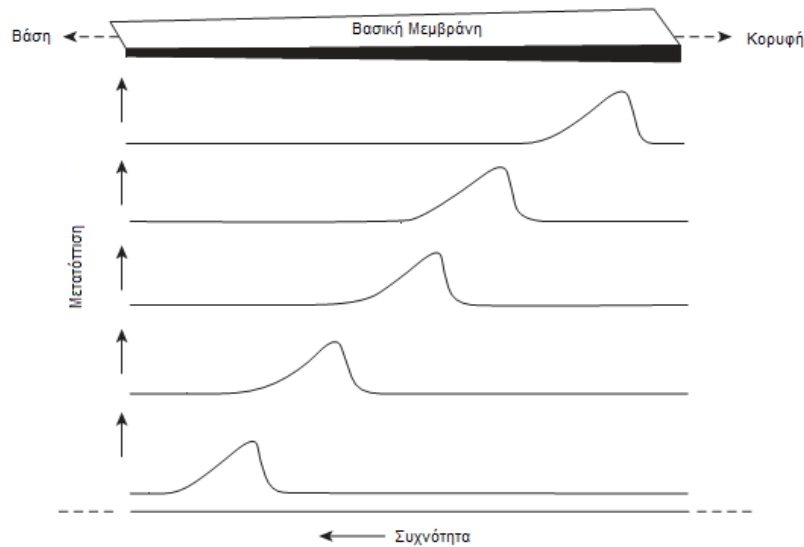
Ο τομέας της Ψυχοακουστικής δίνει απαντήσεις σε ερωτήματα που εγείρονται σχετικά με τον τρόπο που ακούμε, αποκαλύπτοντας τη μη γραμμικότητα του αυτιού μας και τα φαινόμενα που συνεπάγονται της ιδιότητας αυτής. Τα φαινόμενα αυτά είναι πολλά και ενδιαφέροντα, εδώ όμως θα εξετασθούν αυτά τα οποία το πρότυπο MPEG-1 Layer 3 (MP3) αξιοποιεί προκειμένου να συρρικνώσει το μέγεθος του ηχητικού σήματος με όσο το δυνατόν μικρότερες απώλειες ποιότητας. Έτσι, λοιπόν, θα αναλυθούν τα Critical Bands, το Κατώφλι Ακουστότητας και το Masking.

1.3 Critical Bands

Το αυτί μας δέχεται ανά πάσα στιγμή πολλά ταυτόχρονα ακουστικά ερεθίσματα από διαφορετικές ηχητικές πηγές. Πόσο εύκολα όμως μπορεί να ξεχωρίσει δύο συχνότητες,

ειδικά όταν αυτές είναι κοντινές; Και ποια είναι τα όρια μέσα στα οποία αντιλαμβανόμαστε τη διαφορά αυτή;

Κάθε φορά που ένα ακουστικό ερέθισμα φτάνει στο αυτί μας, η βασική μεμβράνη διεγείρεται σε συγκεκριμένα σημεία ανάλογα με τη συχνότητα του ακουστικού ερεθίσματος, όπως φαίνεται και στην Εικόνα 1.1.



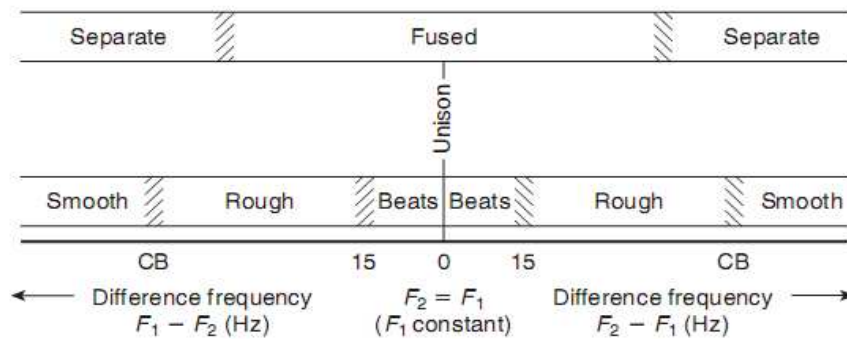
Εικόνα 1.1: Διέγερση της βασικής μεμβράνης ανάλογα με τη συχνότητα¹

Η διάκριση της συχνότητας ανάμεσα σε δύο ακουστικά ερεθίσματα με κοντινές συχνότητες εξαρτάται από τη θέση στην οποία βρίσκεται η μετατόπιση της βασικής μεμβράνης για κάθε ένα από τα ερεθίσματα. Εάν οι μετατοπίσεις είναι σε διαφορετικά σημεία, τότε η διάκριση είναι εφικτή. Επί παραδείγματι, υποθέτουμε πως έχουμε δύο ημιτονοειδή σήματα με ίδιο πλάτος και συχνότητες f_1 και f_2 που συνηχούν. Η f_1 παραμένει σταθερή ενώ η f_2 μεταβάλλεται σταδιακά, ξεκινώντας από την τιμή της f_1 . Όταν η f_2 έχει την ίδια τιμή με την f_1 τότε ακούγεται μια και μόνο συχνότητα. Όταν η f_2 αρχίζει να μεταβάλλεται προς τα πάνω ή προς τα κάτω, για ένα συγκεκριμένο εύρος τιμών, δημιουργείται διακρότημα. Το διακρότημα μας δίνει την αίσθηση της ύπαρξης

¹ Η εικόνα αυτή αποτελεί παράλλαγή αντίστοιχης εικόνας από: Howard, D. M. & Angus J. (2006). Acoustics and Psychoacoustics, 3rd edition. Focal Press. 73.

μιας συχνότητας αλλά κάπως θολής και μπερδεμένης. Όσο μεταβάλλεται περισσότερο η f_2 και περνάει τα όρια του διακροτήματος, το αυτί μας αντιλαμβάνεται δύο συχνότητες με την αίσθηση όμως της σκληράδας στον ήχο τους. Καθώς η διαφορά των f_1 και f_2 μεγαλώνει περισσότερο, η αίσθηση της σκληράδας εξαφανίζεται και το αυτί μας αντιλαμβάνεται πλέον δυο καθαρές διαφορετικές συχνότητες.

Για την πλειοψηφία των ανθρώπων, διακροτήματα σχηματίζονται όταν η διαφορά των δύο τόνων είναι μικρότερη από 12.5Hz και η αίσθηση της σκληράδας εμφανίζεται όταν η διαφορά των δύο συχνοτήτων ξεπεράσει τα 15Hz. Όσο περισσότερο αυξάνεται η συχνοτική διαφορά, η αίσθηση της σκληράδας εξαφανίζεται και ακούγονται άνετα δύο τόνοι διαφορετικής συχνότητας, όπως φαίνεται στην Εικόνα 1.2.



Εικόνα 1.2: Critical bands²

Το σημείο όπου η σκληράδα και η άνετη ακρόαση των ήχων διαχωρίζεται δεν έχει σταθερή τιμή αλλά διαφέρει από άνθρωπο σε άνθρωπο, ενώ ανάλογα με τη συχνότητα αυξομειώνεται και το εύρος της σκληρής περιοχής το οποίο υπολογίζεται σύμφωνα με την παρακάτω σχέση:

$$ERB = \{24.7 * [(4.37 * f_c) + 1]\} \quad (1.1)$$

όπου f_c είναι η κεντρική συχνότητα

και ERB (Equivalent Rectangular Bandwidth) το εύρος σε Hz.

² Η εικόνα αυτή προέρχεται από: Παπαδάκης Νίκος. (2009). Πανεπιστημιακές σημειώσεις για το μάθημα «Εισαγωγή στην Ψυχοακουστική». ΑΤΕΙ Κρήτης. 57.

Με την πάροδο της αίσθησης της σκληράδας σηματοδοτείται η έναρξη της επόμενης κρίσιμης περιοχής. Το ακουστικό μας φάσμα χωρίζεται σε 24 περιοχές για συχνότητες έως και 15kHz, ενώ υπάρχει και μια τελευταία, 25^η περιοχή, η οποία εκτείνεται από τα 15kHz ως τα 20kHz. Στον Πίνακα 1.1 φαίνεται ένα παράδειγμα διαχωρισμού του ακουστικού μας φάσματος σε κρίσιμες ζώνες.

Critical band number	Center frequency(Hz)	Critical band(Hz)	Lower cutoff frequency(Hz)	Upper cutoff frequency(Hz)
1	50	-	-	100
2	150	100	100	200
3	250	100	200	300
4	350	100	300	400
5	450	110	400	510
6	570	120	510	630
7	700	140	630	770
8	840	150	770	920
9	1000	160	920	1080
10	1170	190	1080	1270
11	1370	210	1270	1480
12	1600	240	1480	1720
13	1850	280	1720	2000
14	2150	320	2000	2320
15	2500	380	2320	2700
16	2900	450	2700	3150
17	3400	550	3150	3700
18	4000	700	3700	4400
19	4800	900	4400	5300
20	5800	1100	5300	6400
21	7000	1300	6400	7700
22	8500	1800	7700	9500
23	10500	2500	9500	12000
24	13500	3500	12000	15500
25	18775	6550	15500	22050

Πίνακας 1.1: Διαχωρισμός ακουστικού φάσματος σε κρίσιμες ζώνες³

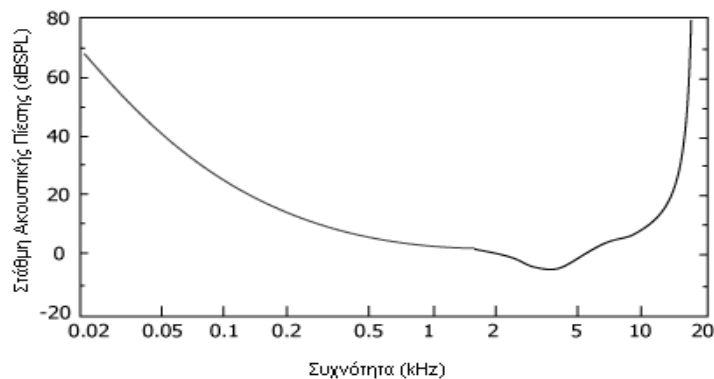
³ Ο πίνακας αυτός προέρχεται από: Pohlmann, K. C. (2005). Principles of digital audio. 5th edition. 322.

1.4 Κατώφλι Ακουστότητας

Το κατώφλι ακουστότητας κατατάσσεται στην κατηγορία των ορίων ακουστότητας του ανθρώπινου ακουστικού συστήματος και σχετίζεται με την ένταση του ήχου. Συγκεκριμένα, το κατώφλι ακουστότητας ενός ήχου ορίζεται ως η ελάχιστη αντιληπτή ένταση του ήχου αυτού, χωρίς την παρουσία άλλων ήχων (Παπαδάκης, 2009, 42).

Στο κατώφλι ακουστότητας αντιστοιχούν τα 0dB_{SPL} (20μPa), στη συχνότητα των 4kHz. Οποιοσδήποτε ήχος κάτω από το όριο αυτό δε γίνεται αντιληπτός. Είναι σημαντικό να σημειωθεί ότι το κατώφλι ακουστότητας δεν είναι σταθερό αλλά μεταβάλλεται ανάλογα με τη συχνότητα του ήχου, όπως φαίνεται στην Εικόνα 1.3.

Εύκολα γίνεται αντιληπτό ότι το αντί μας αντιλαμβάνεται πιο εύκολα τις συχνότητες που βρίσκονται στην περιοχή από 2 έως 5kHz ενώ απαιτούνται πολύ μεγαλύτερες εντάσεις για τις πολύ χαμηλές και πολύ υψηλές συχνοτικές περιοχές.



Εικόνα 1.3: Κατώφλι ακουστότητας συναρτήσει της συχνότητας

1.5 Masking

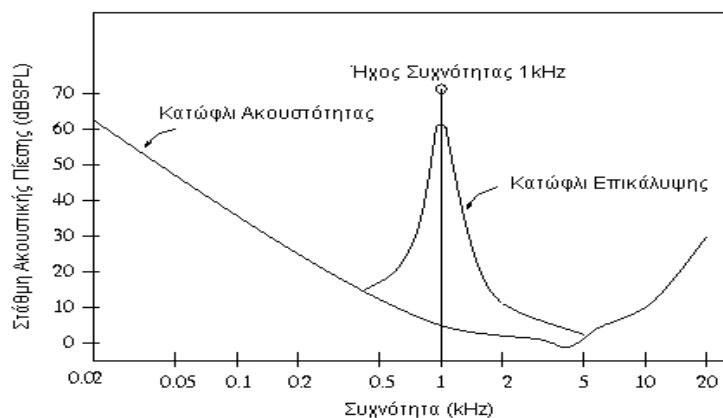
Όπως αναφέρθηκε νωρίτερα, μια ηχητική πληροφορία της οποίας η ένταση δεν ξεπερνά το κατώφλι της ακουστότητας δε γίνεται αντιληπτή από το σύστημά μας. Παρ'όλα αυτά, υπάρχουν περιπτώσεις που μεν ο ήχος έχει αρκετή ένταση αλλά και πάλι δε γίνεται αντιληπτός. Αυτό μπορεί να συμβεί αν ο ήχος αυτός επικαλυφθεί από κάποιον άλλο. Η

επικάλυψη αυτή ονομάζεται masking και λαμβάνει χώρα υπό ορισμένες προϋποθέσεις, που αφορούν το χρόνο και την τονικότητα των ήχων.

1.5.1 Simultaneous Masking

Η επικάλυψη αυτή λαμβάνει χώρα όταν ένας ήχος καθίσταται μη ακουστός λόγω κάποιου άλλου ήχου της ίδιας διάρκειας που συνηχεί με τον πρώτο. Ο ήχος που επικαλύπτεται ονομάζεται maskee και ο ήχος που επικαλύπτει ονομάζεται masker. Οι λόγοι για τους οποίους προκύπτει το φαινόμενο αυτό σχετίζονται με τη συχνότητα και το πλάτος των masker και maskee και συνδέεται άμεσα με τη λειτουργία της βασικής μας μεμβράνης. Σε γενικές γραμμές, ένας ήχος με πλάτος A και συχνότητα f επικαλύπτει έναν ήχο πλάτους $B < A$ και παραπλήσιας στην f συχνότητας. Έχει παρατηρηθεί ότι ήχοι με ψηλές συχνότητες επικαλύπτονται πολύ πιο εύκολα από ήχους με χαμηλές συχνότητες.

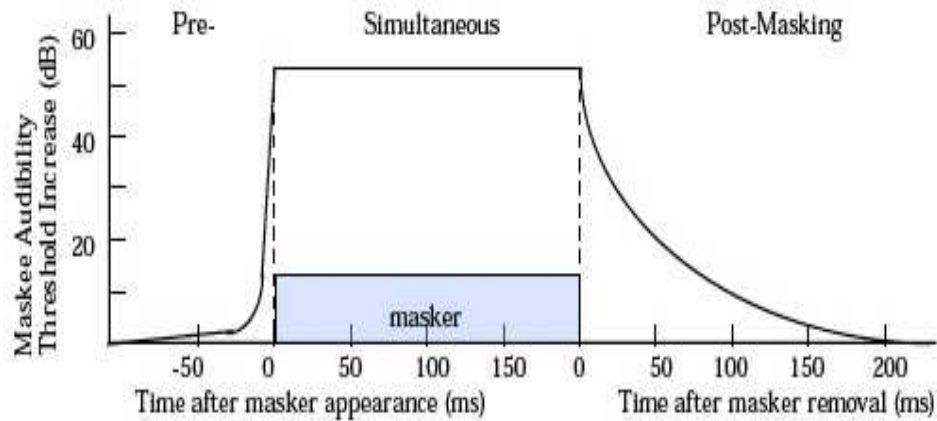
Στην Εικόνα 1.4 φαίνεται ένα παράδειγμα της δυνατότητας επικάλυψης ενός τόνου με συχνότητα 1kHz. Το κατώφλι ακουστότητας δείχνει τις απαιτούμενες εντάσεις για κάθε συχνότητα προκειμένου ο ήχος να γίνει αντιληπτός όταν αυτός δε συνηχεί με άλλους ήχους. Το κατώφλι επικάλυψης δείχνει τις εντάσεις που χρειάζονται οι ήχοι για να γίνουν αντιληπτοί όταν συνηχούν με έναν τόνο συχνότητας 1kHz. Από το παραπάνω παράδειγμα γίνεται αντιληπτό ότι οι παραπλήσιες στο masker συχνότητες χρειάζονται αρκετά μεγαλύτερη ένταση για να ακουστούν σε σχέση με το κατώφλι ακουστότητας.



Εικόνα 1.4: Περιοχή επικάλυψης για ήχο συχνότητας 1kHz

1.5.2 Temporal Masking

Η επικάλυψη αυτή προκύπτει όταν οι ήχοι ακούγονται διαδοχικά σε μικρή χρονική απόσταση. Χωρίζεται σε δύο υποκατηγορίες: το backward masking στο οποίο ο masker ακολουθεί το maskee και το forward masking στο οποίο ο masker προηγείται του maskee. Όσο κοντινότερη είναι η χρονική απόσταση των δύο τόνων, τόσο πιο αποτελεσματική είναι η επικάλυψη. Το forward masking λαμβάνει χώρα όταν η διαφορά μεταξύ των ηχητικών ερεθισμάτων είναι έως 200ms και εξαρτάται από την ένταση και τη διάρκεια του masker ενώ το backward masking συμβαίνει για διαφορά έως 50ms (Εικόνα 1.5).



Εικόνα 1.5: Temporal Masking⁴

⁴ Η εικόνα αυτή προέρχεται από: Jobert S. Jacaba. (2001). Audio compression using modified discrete cosine transform: The MP3 coding standard. Undergraduate research paper. 20.

ΚΕΦΑΛΑΙΟ 2ο: ΒΑΣΙΚΕΣ-ΕΙΣΑΓΩΓΙΚΕΣ ΕΝΝΟΙΕΣ ΓΙΑ ΤΗ ΣΥΜΠΙΕΣΗ ΣΗΜΑΤΩΝ

Εισαγωγή 2^ο Κεφαλαίου

Στο κεφάλαιο αυτό γίνεται αναφορά σε βασικές έννοιες σχετιζόμενες με τη συμπίεση ψηφιακών ηλεκτροακουστικών σημάτων όπως η Απωλεστική Συμπίεση, η Κωδικοποίηση Εντροπίας κ.α.

Σκοπός αυτού του κεφαλαίου είναι η κατανόηση των χρήσιμων αυτών εννοιών οι οποίες κρίνονται απαραίτητες για την ανάλυση των αρχών συμπίεσης του προτύπου MPEG-1 Layer III (MP3).

2.1 Απωλεστική Συμπίεση

Οι τεχνικές απωλεστικής συμπίεσης, όπως φαίνεται και από το όνομά τους, εμπεριέχουν απώλεια ορισμένων πληροφοριών σε σύγκριση με το αρχικό, ασυμπίεστο σήμα, οι οποίες δεν μπορούν να ανακτηθούν μετά τη συμπίεση. Τέτοιου είδους τεχνικές, όμως, είναι ευρέως αποδεκτές μιας και η απώλεια των πληροφοριών αντισταθμίζεται από το πλεονέκτημα του πολύ μικρού όγκου αυτών των συμπιεσμένων αρχείων.

Κατά την απωλεστική συμπίεση, αξιοποιούνται ορισμένα χαρακτηριστικά και ιδιότητες του ακουστικού μας συστήματος προκειμένου το τελικό προϊόν της συμπίεσης να ακούγεται όσο το δυνατόν πιο κοντά στο ασυμπίεστο σήμα.

2.2 Παλμοκωδική Διαμόρφωση (PCM)

Ο όρος Παλμοκωδική Διαμόρφωση αναφέρεται σε μια διαδικασία κατά την οποία αναλογικά σήματα audio μετατρέπονται σε ψηφιακά. Το αποτέλεσμα της διαδικασίας

αυτής είναι ψηφιακά δεδομένα τα οποία αναπαριστούν το αρχικό αναλογικό σήμα και συντελείται σε τρία στάδια:

- Τη δειγματοληψία
- Την κβαντοποίηση
- Την κωδικοποίηση

2.3 Δειγματοληψία

Η δειγματοληψία είναι έννοια που σχετίζεται με τη διαδικασία ψηφιοποίησης ενός αναλογικού σήματος. Για να είναι η δειγματοληψία ικανοποιητική για ένα σήμα, θα πρέπει να ισχύει ο παρακάτω περιορισμός:

«Για ένα αναλογικό σήμα με εύρος από 0 έως f Hz, όπου f η μέγιστη συχνότητα που περιλαμβάνεται στο σήμα, θα πρέπει να ληφθούν τουλάχιστον $2f$ δειγμάτων ανά δευτερόλεπτο».

Έτσι λοιπόν, για παράδειγμα, αν υπάρχει ένα αναλογικό σήμα του οποίου η μέγιστη συχνότητα είναι τα 22kHz, τότε η ελάχιστη, ικανοποιητική συχνότητα δειγματοληψίας θα πρέπει να είναι 44kHz.

2.4 Κβαντοποίηση

Όπως αναφέρθηκε παραπάνω, η δειγματοληψία αναφέρεται στη συχνότητα του προς ψηφιοποίηση αναλογικού σήματος. Η κβαντοποίηση είναι η αναπαράσταση του πλάτους του σήματος για το χρόνο κατά τον οποίο έγινε η δειγματοληψία. Είναι προφανές ότι αφού οι τιμές πλάτους ενός αναλογικού σήματος είναι θεωρητικά άπειρες, το ψηφιακά αναπαραστάμενο πλάτος θα έχει διαφορές από το αρχικό. Παρ' όλα αυτά, κατά τη

διαδικασία της κβαντοποίησης, το αρχικό σήμα προσεγγίζεται ικανοποιητικά, μιας και τα διαστήματα ανάμεσα στις τιμές είναι, χρονικά, πολύ κοντά μεταξύ τους.

2.5 Κωδικοποίηση

Οι διαδικασίες της δειγματοληψίας και της κβαντοποίησης δημιουργούν ένα ψηφιακό σήμα το οποίο όμως δεν είναι σε κατάλληλη μορφή για αποθήκευση σε κάποιο από τα μέσα που υπάρχουν. Για να αποθηκευτεί, λοιπόν η συγκεκριμένη πληροφορία, θα πρέπει να γίνει μια τελευταία διαδικασία, αυτή της κωδικοποίησης.

Για την κωδικοποίηση audio σημάτων, χρησιμοποιείται ο δυαδικός κώδικας (binary code), ο οποίος αποτελείται από τα σύμβολα «0» και «1». Ένας δυαδικός αριθμός (byte), μπορεί να περιλαμβάνει αρκετά ψηφία (bits), ανάλογα με την ανάλυση. Για παράδειγμα, στο CD Audio, ένα δείγμα αποτελείται από δύο λέξεις των 8 bits. Μέσω των δύο αυτών λέξεων, δίνονται οι πληροφορίες για τη συχνότητα και το πλάτος του κάθε δείγματος.

2.6 Κωδικοποίηση Εντροπίας

Η κωδικοποίηση εντροπίας αναφέρεται σε τεχνικές οι οποίες δε λαμβάνουν υπ' όψιν τους το είδος της πληροφορίας που πρόκειται να συμπιεστεί και την αντιμετωπίζουν ως μια απλή ακολουθία bits. Γι' αυτό το λόγο μπορεί να εφαρμοσθεί ανεξάρτητα από το είδος της πληροφορίας.

Οι τεχνικές κωδικοποίησης εντροπίας χωρίζονται σε δύο βασικές κατηγορίες:

- Περιορισμός των επαναλαμβανόμενων ακολουθιών (Suppression of repetitive sequences)
- Στατιστική Κωδικοποίηση (Statistical encoding)

2.6.1 Στατιστική Κωδικοποίηση

Η στατιστική κωδικοποίηση είναι μια μέθοδος που χρησιμοποιείται πολύ συχνά και η βασική της αρχή βρίσκεται στην ανεύρεση των πιο συχνά εμφανιζόμενων ακολουθιών χαρακτήρων και στην κωδικοποίησή τους με λιγότερα bits. Έτσι οι σπάνια εμφανιζόμενες ακολουθίες θα έχουν μεγαλύτερους κωδικούς ενώ οι συχνές μικρότερους.

Η μέθοδος αυτή απαιτεί την ύπαρξη λεξικού μέσα στο οποίο αποθηκεύονται οι ακολουθίες που αντιστοιχούν σε κάθε κωδικό για να μπορεί να πραγματοποιηθεί η συμπίεση. Η στατιστική επεξεργασία των δεδομένων παίζει πολύ σημαντικό ρόλο για την ανεύρεση των ακολουθιών που θα κωδικοποιηθούν με μικρούς κωδικούς. Στις πιο απλές περιπτώσεις το λεξικό είναι σταθερό ενώ στις πιο σύνθετες βρίσκεται κάθε φορά που γίνεται συμπίεση κάποιας ποσότητας δεδομένων.

Η στατιστική κωδικοποίηση μπορεί να έχει δύο μορφές: την αντικατάσταση προτύπων (pattern substitution) και την κωδικοποίηση Huffman (Huffman encoding).

- **Αντικατάσταση προτύπων**

Η μέθοδος αυτή χρησιμοποιείται αποκλειστικά για κείμενα. Συχνά εμφανιζόμενα πρότυπα (ακολουθίες χαρακτήρων, λέξεις κλπ) αντικαθίστανται με λίγους χαρακτήρες.

- **Κωδικοποίηση Huffman**

Η κωδικοποίηση Huffman αποτελεί μια γενίκευση της στατιστικής κωδικοποίησης. Για κάποιο συγκεκριμένο ρεύμα δεδομένων υπολογίζεται η συχνότητα εμφάνισης κάθε χαρακτήρα. Από αυτήν τη συχνότητα, ο αλγόριθμος Huffman υπολογίζει το ελάχιστο μήκος κωδικού που πρέπει να δοθεί σε κάθε χαρακτήρα και πραγματοποιεί τη βέλτιστη ανάθεση κωδικών οι οποίοι αποθηκεύονται στο λεξικό.

Η μέθοδος αυτή χρησιμοποιείται στη συμπίεση ήχου και εικόνας. Σε κάθε περίπτωση η διαδικασία πρέπει να αποθηκεύει το λεξικό για να είναι δυνατή η αποσυμπίεση.

2.7 Κωδικοποίηση Πηγής

Στην τεχνική αυτή, οι μετασχηματισμοί που υφίσταται το αρχικό σήμα εξαρτώνται άμεσα από τον τύπο του. Κατά την κωδικοποίηση πηγής γίνεται χρήση όλων των ιδιαίτερων σημασιολογικών χαρακτηριστικών που μεταφέρει το σήμα προκειμένου αυτά να περιγραφούν με πιο αποτελεσματικό τρόπο.

Οι τεχνικές αυτές παρέχουν μεγαλύτερα ποσοστά συμπίεσης σε σχέση με την κωδικοποίηση εντροπίας αλλά μειονεκτούν σε σταθερότητα γιατί το ποσοστό συμπίεσης εξαρτάται από το αντικείμενο το οποίο συμπιέζεται.

Διακρίνονται σε τρεις διαφορετικές κατηγορίες:

- Κωδικοποίηση μετασχηματισμού (transform encoding)
- Διαφορική ή προβλεπτική κωδικοποίηση (differential or predictive encoding)
- Διανυσματική κβαντοποίηση (vector quantization)

2.7.1 Κωδικοποίηση Μετασχηματισμού

Κατά την κωδικοποίηση μετασχηματισμού, το σήμα υφίσταται ένα μαθηματικό μετασχηματισμό από το αρχικό πεδίο του χρόνου ή του χώρου σε ένα αφηρημένο πεδίο το οποίο είναι πιο κατάλληλο για συμπίεση. Η διαδικασία αυτή είναι αντιστρεπτή, μιας και υπάρχει και ο αντίστροφος μετασχηματισμός ο οποίος θα επαναφέρει το σήμα στην αρχική μορφή του. Ένα χαρακτηριστικό παράδειγμα τέτοιου μετασχηματισμού είναι ο μετασχηματισμός Fourier μέσω του οποίου μια χρονική συνάρτηση $f(t)$ μπορεί να μετασχηματιστεί σε μια $g(\lambda)$ στο πεδίο των συχνοτήτων.

Η διαδικασία συμπίεσης με την εφαρμογή της τεχνικής αυτής έχει απώλειες μιας και μετά την εκτέλεση του μετασχηματισμού, παραμένουν μόνο οι σημαντικότεροι συντελεστές.

2.7.2 Διαφορική ή Προβλεπτική Κωδικοποίηση

Η μέθοδος αυτή στηρίζεται στο γεγονός ότι το μόνο που κωδικοποιείται είναι η διαφορά ανάμεσα στην πραγματική και την προβλεπόμενη τιμή. Η διαφορά αυτή ονομάζεται διαφορά πρόβλεψης ή παράγοντας λάθους.

Η διαφορική κωδικοποίηση είναι κατάλληλη για σήματα οι διαδοχικές τιμές των οποίων αναμένεται να διαφέρουν, αλλά όχι πάρα πολύ. Κατά συνέπεια, μπορεί να χρησιμοποιηθεί για συμπίεση ήχου.

ΚΕΦΑΛΑΙΟ 3^ο: ΠΡΟΤΥΠΟ MPEG-1 LAYER 3 (MP3)

Εισαγωγή 3^{ου} Κεφαλαίου

Στο κεφάλαιο αυτό εξετάζεται αναλυτικά η μέθοδος κατασκευής των αλγορίθμων MP3, ο τρόπος λειτουργίας τους καθώς και η εξέλιξη τους με το πέρασμα του χρόνου. Πιο συγκεκριμένα, στο κεφάλαιο αυτό θα αναφερθούν τα βασικά δομικά χαρακτηριστικά του προτύπου αυτού, ο τρόπος με τον οποίο φτάνουμε στο επιθυμητό αποτέλεσμα, τα προβλήματα τα οποία παρουσιάζονται κατά τη διαδικασία συμπίεσης καθώς και οι τρόποι αντιμετώπισης των προαναφερθέντων προβλημάτων.

Κύριος στόχος αυτού του κεφαλαίου είναι η μελέτη της λογικής που διέπει τους αλγορίθμους MP3 και η όσο το δυνατόν πληρέστερη ανάλυση της, προκειμένου ο αναγνώστης να κατανοήσει το πρότυπο αυτό καθώς και τη σημασία της ύπαρξής του.

3.1 Μια Γενική Επισκόπηση του Προτύπου

3.1.1 Το Πρότυπο MPEG-1

Ο ISO (International Organization for Standardization) είναι ένας διεθνής οργανισμός που στοχεύει στη διευκόλυνση της διεθνούς ανταλλαγής προϊόντων και υπηρεσιών θεσπίζοντας ορισμένα πρότυπα. Η ομάδα MPEG (Moving Picture Experts Group), συνεργαζόμενη με το ISO, κλήθηκε να ξεκινήσει την ανάπτυξη ενός κοινού προτύπου κωδικοποίησης και συμπίεσης εικόνων και ήχου καθώς και του συνδυασμού τους. Το πρότυπο αυτό έπρεπε να πληροί δύο σημαντικές προϋποθέσεις:

να είναι γενικό, δηλαδή κάθε αποκωδικοποιητής που θα το χρησιμοποιούσε να μπορεί να αποκωδικοποιήσει δεδομένα που κωδικοποιήθηκαν από οποιονδήποτε κωδικοποιητή

να γίνει όσο το δυνατόν καλύτερη διατήρηση της αρχικής ποιότητας του ήχου και της εικόνας.

Η ανάπτυξη του ξεκίνησε το 1988 και ολοκληρώθηκε το 1992, παίρνοντας το όνομα MPEG-1. Το πρότυπο αποτελείται από τρία διαφορετικά μέρη:

- Audio part
- Video part
- System part

Το τρίτο μέρος είναι μια περιγραφή του τρόπου με τον οποίο μπορεί να γίνει μετάδοση πολλαπλών σημάτων ήχου και εικόνας με μόνο ένα μέσο μετάδοσης. Με τη χρήση του προτύπου έγινε δυνατή η μετάδοση ήχου και εικόνας με bitrate που κυμαίνεται από 1 έως 2 Mbit/s.

Για το μέρος του ήχου, υπάρχουν τρία διαφορετικά επίπεδα (layers) συμπίεσης, που διαφέρουν αρκετά ως προς την πολυπλοκότητα: τα layer I, layer II και layer III. Η αυξανόμενη πολυπλοκότητα απαιτεί μικρότερο εύρος ζώνης (bandwidth), μιας και η συμπίεση γίνεται όλο και πιο αποτελεσματική. Στον Πίνακα 3.1 δίνονται οι ρυθμοί μετάδοσης που απαιτούνται για το κάθε layer προκειμένου να μεταδίδουν με ποιότητα CD.

Coding	Ratio	Required Bitrate
PCM CD Quality	1:1	1.4Mbps
Layer I	4:1	384kbps
Layer II	8:1	192kbps
Layer III	12:1	128kbps

Πίνακας 3.1: Απαιτούμενοι ρυθμοί μετάδοσης για ποιότητα CD

Το τρίτο επίπεδο συμπιέζει το αρχικό audio αρχείο δώδεκα φορές χωρίς κάποια ιδιαίτερα αντιληπτή διαφορά, καθιστώντας το έτσι το πιο αποτελεσματικό και πολύπλοκο επίπεδο από όλα. Αυτό το επίπεδο είναι που ονομάζεται MPEG-1 Layer III, ή πιο απλά MP3.

Η έρευνα στον τομέα της συμπίεσης συνεχίστηκε και κάποια χρόνια αργότερα βγήκαν κι άλλα πρότυπα, όπως το MPEG-2, MPEG-4 και άλλα τα οποία έχουν κι άλλα χαρακτηριστικά όπως η κωδικοποίηση σε μικρότερες συχνότητες δειγματοληψίας.

3.1.2 Μείωση του Όγκου των Δεδομένων κατά 12

Το MP3 κατατάσσεται στην κατηγορία της κωδικοποίησης με βάση την αντίληψη, μιας και εκμεταλλεύεται το ανθρώπινο ακουστικό σύστημα προκειμένου να αφαιρέσει πληροφορίες με σκοπό τη μείωση του όγκου των δεδομένων. Γίνεται εύκολα αντιληπτό το γεγονός ότι οι πληροφορίες που αφαιρούνται κατά την κωδικοποίηση είναι αδύνατο να ανακτηθούν κατά την αποσυμπίεση.

Το layer III, χωρίζει το ακουστικό φάσμα σε ζώνες που ονομάζονται scalefactor bands. Για κάθε scalefactor band υπολογίζεται ένα κατώτατο όριο επικάλυψης (masking threshold). Ανάλογα με το όριο, ορίζονται συγκεκριμένοι συντελεστές (scalefactors) για κάθε ζώνη με σκοπό τη μείωση του θορύβου κβαντισμού που προκαλείται από τον κβαντισμό των συχνοτικών γραμμών που περιέχονται σε κάθε ζώνη.

Όμως αυτή η διαδικασία δεν είναι αρκετή. Για περαιτέρω συμπίεση το layer III χρησιμοποιεί την κωδικοποίηση Huffman, διαδικασία η οποία θα αναλυθεί αργότερα.

3.1.3 Ελευθερία Εκτέλεσης

Οι προδιαγραφές του MP3 (ISO 11172-3) καθορίζουν τη δομή και τη μετάφραση της κωδικοποιημένης και αποκωδικοποιημένης ροής των δεδομένων. Σύμφωνα με αυτό, το παραγόμενο σήμα στην έξοδο ενός κωδικοποιητή αναγνωρίζεται από οποιονδήποτε αποκωδικοποιητή MP3 και αντίστροφα. Οι προδιαγραφές, όμως, δεν καθορίζουν τα ακριβή βήματα της κωδικοποίησης ενός ασυμπίεστου σήματος. Αυτό σημαίνει ότι οι κωδικοποιητές μπορεί να λειτουργούν με κάποιες διαφορές μεταξύ τους παράγοντας, όμως, συμμορφούμενα με τις προδιαγραφές δεδομένα. Ορισμένες από τις ενέργειες του κωδικοποιητή καθορίζονται μόνο από τον άνθρωπο που κατασκευάζει τον κωδικοποιητή, δίνοντας του έτσι μια σχετική ελευθερία κατά την εκτέλεση ενός κωδικοποιητή.

Δύο σημαντικά χαρακτηριστικά ενός κωδικοποιητή είναι η ταχύτητα και η ποιότητα. Δυστυχώς, η εκτέλεση που δίνεται από τις προδιαγραφές δεν εφαρμόζεται πάντα από τους πιο αποδοτικούς αλγόριθμους. Γι' αυτό, η ταχύτητα και η ποιότητα του αποτελέσματος διαφέρουν, συχνά πολύ, από τον έναν κωδικοποιητή στον άλλο.

Σε ότι αφορά την αποκωδικοποίηση, όλοι οι μετασχηματισμοί που απαιτούνται προκειμένου να παραχθούν δείγματα PCM είναι καθορισμένοι. Παρ' όλα αυτά, υπάρχουν ορισμένες λεπτομέρειες οι οποίες δεν καθορίζονται πλήρως.

Η ελευθερία που δίνεται από το πρότυπο πρέπει να αξιοποιείται καταλλήλως προκειμένου να βρίσκονται οι πιο αποδοτικές λύσεις σε θέματα ποιότητας αλλά και ταχύτητας.

3.1.4 Bitrate

Το bitrate είναι μια πληροφορία η οποία πρέπει να επιλεγεί πριν την κωδικοποίηση. Με αυτή πληροφορείται ο κωδικοποιητής για την ποσότητα των δεδομένων που επιτρέπεται να αποθηκευτούν για κάθε δευτερόλεπτο ασυμπιεστού σήματος. Έτσι, ο χρήστης έχει τη δυνατότητα να επιλέξει την ποιότητα του προς κωδικοποίηση σήματος. Το layer III ορίζει πως το bitrate επιτρέπεται να κυμαίνεται από 8kbit/s έως 320kbit/s. Η πιο διαδεδομένη χρησιμοποιούμενη τιμή είναι τα 128kb/s. Όσο μεγαλύτερη η τιμή του bitrate, τόσο καλύτερη η ποιότητα του παραγόμενου σήματος.

Αξίζει εδώ να σημειωθεί ότι για ένα στερεοφωνικό αρχείο με συγκεκριμένο bitrate, χωρίζεται στα δύο κανάλια δίνοντας μεγαλύτερη ποσότητα στο κανάλι με την περισσότερη πολυπλοκότητα.

Το πρότυπο καθορίζει δύο είδη bitrate, το Constant Bitrate (CBR) και το Variable Bitrate (VBR).

Όταν ο κωδικοποιητής χρησιμοποιεί το CBR, κάθε μέρος του αρχείου κωδικοποιείται με τον ίδιο αριθμό bits. Τα περισσότερα μουσικά αρχεία όμως έχουν μεταβαλλόμενη πολυπλοκότητα. Κάποια μέρη τους μπορεί να χρησιμοποιούν πολλά διαφορετικά μουσικά όργανα ή effects ενώ κάποια άλλα να είναι πιο απλά. Το πρόβλημα εδώ είναι εμφανές μιας και τα πιο πολύπλοκα μέρη, που λογικά θα απαιτούσαν μεγαλύτερο αριθμό bits για την κωδικοποίησή τους, καταλήγουν να έχουν τον ίδιο αριθμό bits με τα απλά μέρη, δημιουργώντας έτσι «άνιση» ποιότητα.

Το VBR αποτελεί λύση σε αυτό το πρόβλημα μιας και επιτρέπει αλλαγές στο bitrate ανάλογα με τις δυναμικές του σήματος. Σε αυτή την περίπτωση, η ποιότητα καθορίζεται από το χρήστη όχι ως μια σταθερή τιμή αλλά ως ένα ανώτατο επιτρεπόμενο όριο. Δυστυχώς όμως, το VBR μπορεί να προκαλέσει προβλήματα χρόνου σε ορισμένους αποκωδικοποιητές, όπως η μη σωστή ένδειξη διάρκειας του κομματιού.

3.1.5 Συχνότητα Δειγματοληψίας

Η ανάλυση ενός αρχείου ήχου εξαρτάται κυρίως από τη συχνότητα δειγματοληψίας, η οποία ορίζει πόσες φορές το δευτερόλεπτο αποθηκεύεται το σήμα. Ένα υψηλό bitrate δίνει μεγάλη ακρίβεια στην τιμή ενός δείγματος ενώ μια μεγάλη συχνότητα δειγματοληψίας δίνει τη δυνατότητα αποθήκευσης περισσότερων τιμών, παρέχοντας έτσι μεγάλο συχνοτικό φάσμα. Το πρότυπο ορίζει τις τιμές της συχνότητας δειγματοληψίας στα 32kHz, 44.1kHz και 48kHz.

3.1.6 Channel Modes

Υπάρχουν τέσσερα είδη καθορισμένα από το πρότυπο:

- Single Channel
- Dual Channel (τα κανάλια κωδικοποιούνται ανεξάρτητα το ένα από το άλλο και κάθε κανάλι χρησιμοποιεί ακριβώς το μισό από το καθορισμένο bitrate)
- Stereo
- Joint Stereo

Η περίπτωση του joint stereo παρουσιάζει ενδιαφέρον, μιας και στο mode αυτό χρησιμοποιείται η συσχέτιση των δύο καναλιών για τη βελτίωση της κωδικοποίησης. Υπάρχουν δύο τεχνικές εδώ: η Middle/Side Stereo (MS Stereo) και η Intensity Stereo αν και κάποιες φορές χρησιμοποιείται συνδυασμός των δύο τεχνικών.

Η MS είναι ιδιαίτερα χρήσιμη όταν τα δύο κανάλια σχετίζονται πολύ μεταξύ τους. Στην τεχνική αυτή, στο δεξί κανάλι μεταδίδεται το άθροισμα των L και R, ενώ στο αριστερό κανάλι μεταδίδεται η διαφορά των L και R. Αφού τα δύο κανάλια είναι αρκετά ίδια, τις

περισσότερες φορές σήμα που προκύπτει από την άθροιση περιέχει περισσότερες πληροφορίες από το σήμα της διαφοράς. Αυτό επιτρέπει πιο αποδοτική συμπίεση σε σύγκριση με την ανεξάρτητη μετάδοση των δύο καναλιών.

Στην Intensity Stereo οι υψηλότερες συχνοτικά υποζώνες κωδικοποιούνται σε ένα και μόνο σήμα με θέσεις έντασης που αντιστοιχούν στις scalefactor bands που αναφέρθηκαν προηγουμένως. Έτσι, οι πληροφορίες περιορίζονται σε ένα και μόνο κανάλι το οποίο και μεταδίδεται. Είναι φυσικό να υπάρχουν «ασυνέπειες» λόγω της μετάδοσης ενός σήματος και από τα δύο κανάλια, όμως εάν κρατηθούν σε χαμηλά επίπεδα τότε δε γίνονται αντιληπτές.

3.2 Ανατομία Ενός Αρχείου MP3

Όλα τα αρχεία MP3 διαιρούνται σε μικρότερα κομμάτια που ονομάζονται frames. Κάθε frame αποθηκεύει 1152 δείγματα ήχου, έχει διάρκεια 26ms και χωρίζεται σε δύο μικρότερα μέρη, τα granules, με 576 δείγματα ήχου το καθένα. Αφού, το bitrate καθορίζει το μέγεθος του κάθε δείγματος, η αύξηση του προκαλεί αύξηση στο μέγεθος του frame. Το μέγεθος εξαρτάται επίσης και από τη συχνότητα δειγματοληψίας, σύμφωνα με τον παρακάτω τύπο:

$$\frac{144 * \textit{bitrate}}{\textit{samplefrequency}} + \textit{Padding} \quad (\text{bytes}) \quad (2.1)$$

Ο όρος padding αναφέρεται σε ένα ειδικό bit το οποίο βρίσκεται στην αρχή του frame και χρησιμοποιείται σε ορισμένα frames, όπως θα δούμε παρακάτω.

3.2.1 Μορφή των Frames

Ένα frame αποτελείται από πέντε μέρη: την επικεφαλίδα (frame header), το CRC, το side information, τα κύρια δεδομένα (main data) και τα βοηθητικά δεδομένα (ancillary data), με τη σειρά που φαίνεται στο Σχήμα 3.1.

Header	CRC	Side Information	Main Data	Ancillary Data
--------	-----	------------------	-----------	----------------

Σχήμα 3.1: Μορφή των frames

3.2.1.1 Frame Header

Η επικεφαλίδα αποτελείται από 32 bits που περιέχουν μια λέξη συγχρονισμού και μία περιγραφή του frame. Η λέξη συγχρονισμού βρίσκεται στην αρχή του κάθε frame, επιτρέποντας έτσι στο δέκτη του MP3 να «κλειδώνει» το σήμα σε οποιοδήποτε σημείο της ροής των δεδομένων. Έτσι γίνεται δυνατή η μετάδοση οποιουδήποτε MP3 αρχείου. Το μόνο που χρειάζεται να κάνει ο δέκτης κατά τη διάρκεια της μετάδοσης ενός MP3 είναι να βρει τη λέξη συγχρονισμού. Ένα πρόβλημα που παρουσιάζεται σε αυτό το σημείο είναι οι χαλκευμένες λέξεις συγχρονισμού που μπορεί να εμφανίζονται σε άλλα σημεία ενός frame. Γι' αυτό το λόγο, ένας αποκωδικοποιητής πρέπει είτε να ελέγχει για έγκυρες λέξεις συγχρονισμού σε δύο συνεχόμενα frames είτε για έγκυρα δεδομένα στο μέρος side information (κάτι το οποίο είναι πρακτικά πιο δύσκολο).

Τα bits που αποτελούν την επικεφαλίδα χωρίζονται σε ομάδες κάθε μία από τις οποίες επιτελούν διαφορετικό έργο.

Sync			
ID	Layer	Prot. bit	
Bitrate			
Frequency	Pad. bit	Priv. bit	
Mode		Mode extention	
Copy	Home	Emphasis	

Εικόνα 3.1: Η επικεφαλίδα ενός frame

Sync (12 bits)

Αυτή είναι η λέξη συγχρονισμού που αναφέρθηκε παραπάνω. Είναι υποχρεωτικό να έχουν τιμή και τα 12 bits.

ID (1 bit)

Καθορίζει την έκδοση του MPEG. Εάν το bit έχει τιμή, σημαίνει ότι το frame έχει κωδικοποιηθεί σύμφωνα με το πρότυπο MPEG-1. Διαφορετικά, έχει κωδικοποιηθεί σύμφωνα με το πρότυπο MPEG-2.

Σε ορισμένες περιπτώσεις, η επικεφαλίδα έχει 11 bits, χαρίζοντας έτσι ένα επιπλέον bit στο ID. Όταν ισχύει αυτό, τα bits του ID μεταφράζονται σύμφωνα με τον Πίνακα 2.2 που φαίνεται παρακάτω.

00	MPEG-2.5
01	Reserved
10	MPEG-2
11	MPEG-1

Πίνακας 3.2: Χρήση των bits όταν υπάρχουν 2 bits στο ID

Layer (2 bits)

Καθορίζει το layer του frame, σύμφωνα με τον Πίνακα 3.3.

00	Reserved
01	Layer III
10	Layer II
11	Layer I

Πίνακας 3.3: Καθορισμός του layer

Protection bit (1 bit)

Εάν το protection bit έχει καθορισμένη τιμή, τότε το πεδίο CRC θα χρησιμοποιηθεί.

Bitrate (4 bits)

Αυτά τα 4 bits πληροφορούν τον αποκωδικοποιητή για το bitrate στο οποίο έχει κωδικοποιηθεί το frame. Αυτή η τιμή θα είναι ίδια για κάθε frame εάν το αρχείο έχει κωδικοποιηθεί με CBR. Στον Πίνακα 3.4 φαίνονται οι τιμές που αντιστοιχούν σε κάθε bitrate.

Bits	MPEG-1 layer I	MPEG-1 layer II	MPEG-1 layer III
0000	free*	free*	free*
0001	32	32	32
0010	64	48	40
0011	96	56	48
0100	128	64	56
0101	160	80	64
0110	192	96	80
0111	224	112	96
1000	256	128	112
1001	288	160	128
1010	320	192	160
1011	352	224	192
1100	384	256	224
1101	416	320	256
1110	448	384	320
1111	bad**	bad**	bad**

Πίνακας 3.4: Καθορισμός του bitrate

Frequency (2 bits)

Αυτά τα 2 bits δίνουν πληροφορίες για τη συχνότητα δειγματοληψίας, όπως φαίνεται στον Πίνακα 2.5.

Bits	MPEG-1	MPEG-2	MPEG-2.5
00	44100Hz	22050Hz	11025Hz
01	48000Hz	24000Hz	12000Hz
10	32000Hz	16000Hz	8000Hz
11	Reserved	Reserved	Reserved

Πίνακας 3.5: Καθορισμός συχνοτήτων δειγματοληψίας

Padding bit (1 bit)

Τα bit αυτό «γεμίζει» κενά τα οποία μπορεί να δημιουργηθούν στα frames. Για παράδειγμα, ένα κωδικοποιημένο αρχείο με bitrate στα 128kbit/s και με συχνότητα δειγματοληψίας 44100Hz, ενδέχεται να δημιουργήσει frames μήκους 417 bytes. Προκειμένου να χωρέσει όμως το bitrate των 128kbit/s, χρειάζονται frames μήκους 418 bytes. Το padding bit, λοιπόν, γεμίζει αυτό το κενό. Εάν έχει την τιμή «0», τότε το frame δεν χρησιμοποιεί padding bit. Εάν έχει την τιμή «1», τότε χρησιμοποιείται.

Private bit (1 bit)

Το bit αυτό παρέχει ορισμένες πληροφορίες στην εφαρμογή που χρησιμοποιεί το MP3.

Mode (2 bits)

Καθορίζουν το channel mode που χρησιμοποιείται, σύμφωνα με τον Πίνακα 3.6.

00	Stereo
01	Joint Stereo
10	Dual Channel
11	Single Channel

Πίνακας 3.6: Καθορισμός του channel mode

Mode Extension (2 bits)

Αυτά τα δύο bits χρησιμοποιούνται μόνο στο Joint Stereo και καθορίζουν τις χρησιμοποιούμενες μεθόδους. Το Joint Stereo μπορεί να διαφέρει από frame σε frame, ακόμα και να απενεργοποιείται. Οι μέθοδοι που χρησιμοποιούνται φαίνονται στον Πίνακα 3.7.

Bits	Intensity Stereo	MS Stereo
00	Off	Off
01	On	Off
10	Off	On
11	On	On

Πίνακας 3.7: Καθορισμός μεθόδων του Joint Stereo

Copyright bit (1 bit)

Εάν αυτό το bit έχει τιμή, τότε απαγορεύεται η αντιγραφή του περιεχομένου του αρχείου.

Home (Original bit) (1 bit)

Το bit αυτό, εάν είναι καθορισμένο, δείχνει πως το frame βρίσκεται στο πρωτότυπο αρχείο.

Emphasis (2 bits)

Η ένδειξη αυτή χρησιμοποιείται για να δείξει στον αποκωδικοποιητή ότι το αρχείο πρέπει να αλλάξει λόγω κάποιας προηγούμενης ενέργειας. Για παράδειγμα, να γίνει re-

equalization μετά από μια ενέργεια μείωσης θορύβου. Τα bits αυτά χρησιμοποιούνται πολύ σπάνια.

2.2.1.2 CRC (Cycling Redundancy Code)

Αυτό το πεδίο υπάρχει μόνο εάν το protection bit της επικεφαλίδας είναι καθορισμένο και επιτρέπει τον έλεγχο των πιο ευαίσθητων δεδομένων για λάθη κατά τη μετάδοση. Τα ευαίσθητα δεδομένα έχουν καθοριστεί από το πρότυπο και είναι τα bits 16 έως 31 της επικεφαλίδας και του side information. Εάν υπάρχουν σφάλματα στα bits αυτά, τότε καταστρέφεται ολόκληρο το frame, σε αντίθεση με σφάλματα σε άλλα bits που προκαλούν μερική καταστροφή του frame. Ένα κατεστραμμένο frame πρέπει είτε να μη γίνει ακουστό είτε να αντικατασταθεί από το προηγούμενο frame.

2.2.1.3 Side Information

Το πεδίο αυτό αποτελείται από πληροφορίες που χρειάζονται για την αποκωδικοποίηση των κυρίως δεδομένων (main data). Το μέγεθός του εξαρτάται από το κωδικοποιημένο channel mode. Εάν το channel mode είναι single channel, τότε το μέγεθος του side information θα είναι 17 bytes. Σε κάθε άλλη περίπτωση θα είναι 32 bytes. Το πεδίο αυτό χωρίζεται σε πέντε διαφορετικά μέρη που φαίνονται στον Πίνακα 3.8 και αναλύονται παρακάτω. Να σημειωθεί ότι στις παρενθέσεις δίπλα στα μέρη δηλώνεται ο αριθμός των χρησιμοποιούμενων bits. Τα μέρη που έχουν μια τιμή έχουν και σταθερό μέγεθος. Τα μέρη με δύο τιμές έχουν μεταβλητό μέγεθος ανάλογα με τα modes. Η πρώτη τιμή ισχύει εάν χρησιμοποιείται mono mode ενώ η δεύτερη για όλες τις άλλες περιπτώσεις.

main_data_begin	private bits	scfsi	Side info gr. 0	Side info gr. 1
------------------------	---------------------	--------------	------------------------	------------------------

Πίνακας 3.8: Side information

main_data_begin (9 bits)

Η χρήση του layer III περιλαμβάνει μια τεχνική που ονομάζεται bit reservoir. Σύμφωνα με αυτή, ο εναπομένον ελεύθερος χώρος στην περιοχή των κυρίως δεδομένων ενός frame μπορεί να χρησιμοποιηθεί από διαδοχικά frames. Για να βρει ο αποκωδικοποιητής το σημείο που αρχίζουν τα κύρια δεδομένα ενός frame πρέπει να διαβάσει το main_data_begin. Η τιμή αυτή έχει συγκεκριμένη απόσταση από το πρώτο byte της

λέξης συγχρονισμού. Τα δεδομένα ενός frame μπορεί να βρίσκονται σε αρκετά προηγούμενα frames. Αξίζει να σημειωθεί ότι τα στατικά μέρη του frame, όπως η επικεφαλίδα, δε συμπεριλαμβάνονται στην απόσταση. Εάν η τιμή του `main_data_begin` ισούται με 0, τότε τα δεδομένα ξεκινούν αμέσως μετά από το side information.

private_bits (5 bits, 3 bits)

Τα bits αυτά ήταν ιδιωτικής χρήσεως και δεν χρησιμοποιούνται πλέον.

scfsi (4 bits, 8 bits)

Σε αυτό το πεδίο καθορίζεται αν θα χρησιμοποιηθούν οι ίδιοι συντελεστές κλίμακας και στα δύο granules ή όχι. Εδώ, οι scalefactor bands χωρίζονται σε τέσσερις κατηγορίες, όπως φαίνεται στον Πίνακα 3.9.

Group	Scalefactor Bands
0	0, 1, 2, 3, 4, 5
1	6, 7, 8, 9, 10
2	11, 12, 13, 14, 15
3	16, 17, 18, 19, 20

Πίνακας 3.9: Κατηγορίες scalefactor bands

Μεταδίδονται 4 bits για κάθε κανάλι, ένα για κάθε ζώνη. Εάν ένα bit που ανήκει σε μια ζώνη είναι μηδέν, οι συντελεστές της συγκεκριμένης ζώνης μεταδίδονται για κάθε granule ξεχωριστά, ενώ αν έχει τιμή «1» τότε ισχύει και για τα δύο granules. Αυτό σημαίνει ότι οι συντελεστές χρειάζεται να μεταδοθούν μόνο για το granule0 και το bit που μένει ελεύθερο στο granule1 χρησιμοποιείται αργότερα από την κωδικοποίηση Huffman.

Εάν σε κάποιο granule χρησιμοποιούνται short windows, τότε οι συντελεστές στέλνονται ξεχωριστά για κάθε granule στο κάθε κανάλι.

Side_info gr.0, Side_info gr.1

Τα δύο αυτά μέρη έχουν την ίδια δομή και αποτελούνται από αρκετά υπομέρη, όπως φαίνεται στον Πίνακα 3.10. Σε αυτά περιέχονται συγκεκριμένες πληροφορίες για κάθε granule.

part2 3 length	big values	global gain	scalefac compress
windows switching flag	block type	mixed block flag	table select
subblock gain	region0 count	region1 count	Preflag
scalefac scale	count1 table select		

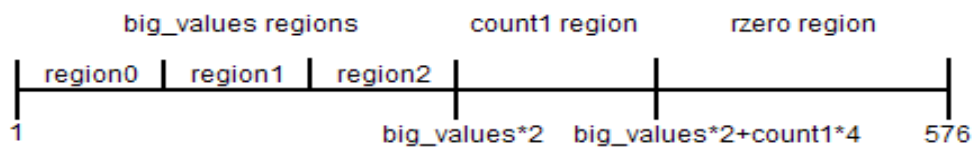
Πίνακας 3.10: Πεδία του side information για κάθε granule

part2_3_length (12 bits, 24 bits)

Δηλώνει τον αριθμό των bits που βρίσκονται στο main data για τους συντελεστές (part2) και για την κωδικοποίηση Huffman (part3). Για ένα μόνο κανάλι χρησιμοποιούνται 12 bits ενώ για stereo 24 bits. Αυτό το πεδίο μπορεί να χρησιμοποιηθεί για τον υπολογισμό της τοποθεσίας του επόμενου granule και των βοηθητικών πληροφοριών (ancillary data), εάν αυτές υπάρχουν.

big_values (9 bits, 18 bits)

Οι 576 συχνοτικές γραμμές του κάθε granule δε χρησιμοποιούν τον ίδιο πίνακα για την κωδικοποίησή τους. Οι συχνότητες αυτές έχουν εύρος από 0Hz έως τη συχνότητα Nyquist και διαιρούνται σε πέντε περιοχές, όπως φαίνεται στο Σχήμα 3.2. Ο σκοπός αυτού του διαχωρισμού είναι να μεγιστοποιήσει την απόδοση του κωδικοποιητή Huffman, χρησιμοποιώντας διαφορετικές τιμές για κάθε περιοχή.



Σχήμα 3.2: Περιοχές του συχνοτικού φάσματος

Ο διαχωρισμός γίνεται ανάλογα με τις μέγιστες κβαντισμένες τιμές, υποθέτοντας ότι οι τιμές σε υψηλότερες συχνότητες αναμένεται είτε να έχουν πολύ μικρότερα πλάτη είτε να μη χρειάζεται να κωδικοποιηθούν καθόλου.

global_gain (8 bits, 16 bits)

Το πεδίο αυτό καθορίζει το βήμα κβαντισμού, το οποίο χρησιμοποιείται από τον αποκωδικοποιητή, στο block του επανακβαντισμού.

scalefac_compress (4 bits, 8 bits)

Καθορίζει τον αριθμό των bits που απαιτούνται για τη μετάδοση των συντελεστών. Ένα granule χωρίζεται σε 12 ή 21 scalefactors bands, ανάλογα με τον τύπο παραθύρου που χρησιμοποιείται (long windows, short windows). Εάν χρησιμοποιούνται short windows, το granule χωρίζεται σε 12 ζώνες, ενώ αν υπάρχουν long windows χωρίζεται σε 21 ζώνες. Οι συντελεστές χωρίζονται και αυτοί σε δύο κατηγορίες για κάθε τύπο παραθύρου. Από 0 έως 10 και από 11 έως 20 για τα long windows και από 0 έως 6 και 7 έως 11 για τα short windows.

Η μεταβλητή scalefac_compress είναι ένας δείκτης για έναν καθορισμένο πίνακα τιμών, όπως φαίνεται στον Πίνακα 3.11. Τα slen1 και slen2 δηλώνουν τον αριθμό των bits που χρησιμοποιούνται από την πρώτη και τη δεύτερη κατηγορία των συντελεστών αντίστοιχα.

scalefac_compress	slen1	slen2
0	0	0
1	0	1
2	0	2
3	0	3
4	3	0
5	1	1
6	1	2
7	1	3
8	2	1
9	2	2
10	2	3
11	3	1
12	3	2

13	3	3
14	4	2
15	4	3

Πίνακας 3.11: Πίνακας τιμών για τη μεταβλητή scalefac_compress

windows_switching_flag (1 bit, 2 bits)

Εάν έχει τιμή, υποδηλώνει ότι χρησιμοποιείται κάποιος διαφορετικός τύπος παραθύρου από τον κανονικό. Επίσης, όλες οι εναπομένουσες τιμές που δε βρίσκονται στο region0 περιέχονται στο region1, κι έτσι δε χρησιμοποιείται το region2.

Τα πεδία block_type, mixed_blockflag και subblock_gain χρησιμοποιούνται μόνο όταν χρησιμοποιείται το windows_switching_flag.

block_type (2 bits, 4 bits)

Σε αυτό το πεδίο δηλώνεται ο τύπος παραθύρου που χρησιμοποιείται από το συγκεκριμένο granule, όπως φαίνεται στον Πίνακα 3.12. Η τιμή 00 είναι απαγορευμένη μιας και το block_type χρησιμοποιείται μόνο στην περίπτωση που είναι ενεργό το windows_switching_flag.

Block type	Window Type
00	forbidden
01	start
10	3 short windows
11	end

Πίνακας 3.12: Ορισμός του block_type

mixed_blockflag (1 bit, 2 bits)

Στο πεδίο αυτό δηλώνεται η ύπαρξη διαφορετικών παραθύρων για τις χαμηλές και τις υψηλές συχνότητες. Εάν το πεδίο έχει τιμή, οι δύο χαμηλότερες υποζώνες μετασχηματίζονται χρησιμοποιώντας κανονικό παράθυρο και οι υπόλοιπες 30 υποζώνες χρησιμοποιούν τον τύπο παραθύρου που καθορίζεται από το block_type.

table_select [(10 bits, 20 bits) ή (15 bits, 30 bits)]

Υπάρχουν 32 πιθανοί πίνακες Huffman που προβλέπονται από το πρότυπο. Η τιμή στο πεδίο αυτό καθορίζει ποιος πίνακας θα χρησιμοποιηθεί κατά την αποκωδικοποίηση και έχει μέγεθος 5 bits, δηλαδή 32 διαφορετικές τιμές για κάθε περιοχή, granule και κανάλι. Το table_select καθορίζει μόνο τον πίνακα που θα χρησιμοποιηθεί κατά την αποκωδικοποίηση του μέρους big_values. Ο καθορισμός του πίνακα εξαρτάται από τοπική μορφή του σήματος καθώς και από τη μέγιστη τιμή κβαντισμού που επιτρέπεται για τον κβαντισμό των 576 συχνοτικών γραμμών του granule.

Όπως αναφέρθηκε παραπάνω, όταν το windows_switching_flag είναι ενεργό, το region2 είναι κενό, οπότε μόνο δύο περιοχές κωδικοποιούνται. Αυτό δείχνει ότι στη μονοφωνική περίπτωση χρειάζονται $5*2*1 = 10$ bits ενώ στο stereo mode $5*2*2 = 20$ bits. Στην περίπτωση όμως που το windows_switching_flag είναι «0» τότε κωδικοποιούνται και οι τρεις περιοχές. Έτσι χρειάζονται $5*3*1 = 15$ bits για τη μονοφωνία και $5*3*2 = 30$ bits για τη στερεοφωνία.

subblock_gain (9 bits, 18 bits)

Αυτό το πεδίο είναι ενεργό όταν το windows_switching_flag έχει τιμή «1» και όταν το block_type = 10 (3 short windows) και δείχνει την απόσταση από το global_gain για κάθε short block.

region0_count (4 bits, 8 bits), region1_count (3 bits, 6 bits)

Στα πεδία αυτά περιέχονται πληροφορίες για τα όρια της κάθε περιοχής (region0 και region1) και προσαρμόζονται ανάλογα με τους scalefactors κάθε περιοχής.

preflag (1 bit, 2 bits)

Το preflag είναι μια συντόμευση για επιπλέον μετατροπές των υψηλών συχνοτικών περιοχών. Όταν είναι ενεργό, ορισμένες τιμές προστίθενται στους scalefactor, σύμφωνα

με τον Πίνακα 3.13 (το pretab δηλώνει την τιμή της αύξησης). Εάν χρησιμοποιούνται μικρά παράθυρα (short windows), το preflag δε χρησιμοποιείται ποτέ.

Scalefac Scale	Pretab
0	0
1	0
2	0
3	0
4	0
5	0
6	0
7	0
8	0
9	0
10	0
11	1
12	1
13	1
14	1
15	2
16	2
17	3
18	3
19	3
20	2

Πίνακας 3.13: Αύξηση των scalefactors όταν χρησιμοποιείται το preflag

scalefac_scale (1 bit, 2 bits)

Οι scalefactors κβαντίζονται λογαριθμικά με βήμα 2 ή $\sqrt{2}$ όπως φαίνεται στον Πίνακα 3.14.

scalefac_scale	step size
0	2
1	$\sqrt{2}$

Πίνακας 3.14: Βήμα κβαντισμού για τους scalefactors

count1table_select (1 bit, 2 bits)

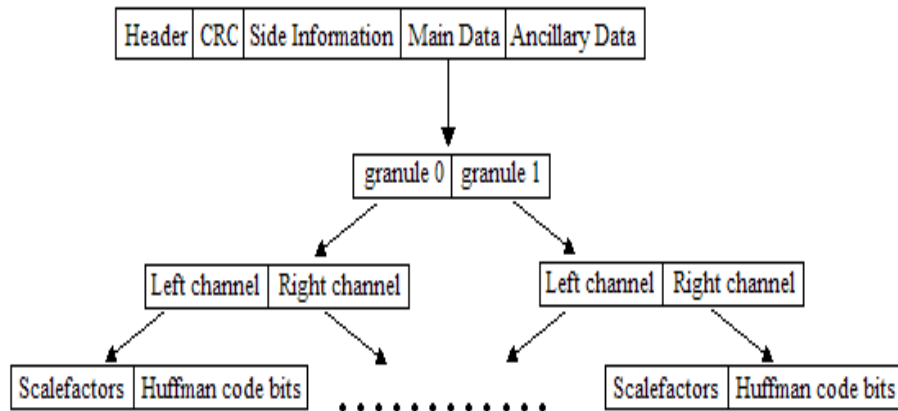
Στο πεδίο αυτό προσδιορίζεται ο πίνακας Huffman που θα χρησιμοποιηθεί στην περιοχή count1.

3.2.1.4 Main Data

Τα κυρίως δεδομένα (main data) ενός frame αποτελούνται από scalefactors, κωδικοποιημένα bits (με τη μέθοδο Huffman) και βοηθητικά δεδομένα (ancillary data).

- **Scalefactors**

Ο στόχος των scalefactors είναι η μείωση του θορύβου κβαντισμού. Εάν τα δείγματα σε μια συγκεκριμένη ζώνη χωριστούν σωστά τότε ο θόρυβος κβαντισμού θα επικαλυφθεί εντελώς. Για κάθε ζώνη μεταδίδεται ένας scalefactor. Το πεδίο scfsi δηλώνει εάν οι scalefactors είναι κοινοί και για τα δύο granules ή όχι. Τα bits που χρησιμοποιούνται για κάθε scalefactor καθορίζονται από το πεδίο scalefac_compress.



Σχήμα 3.3: Κατανομή των scalefactors σε granules και κανάλια

Η διαίρεση του ακουστικού φάσματος σε ζώνες είναι καθορισμένη για κάθε τύπο παραθύρου και συχνότητα δειγματοληψίας. Ένας τέτοιος πίνακας είναι ο Πίνακας 3.15, για μεγάλα παράθυρα και συχνότητα δειγματοληψίας 44.1kHz.

scalefactor band	Width	start index	end index
0	4	0	3
1	4	4	7
2	4	8	11
3	4	12	15
4	4	16	19
5	4	20	23
6	6	24	29

7	6	30	35
8	8	36	43
9	8	44	51
10	10	52	61
11	12	62	73
12	16	74	89
13	20	90	109
14	24	110	133
15	28	134	161
16	34	162	195
17	42	196	237
18	50	238	287
19	54	288	341
20	76	342	417

Πίνακας 3.15: Scalefactors για συχνότητα δειγματοληψίας 44.1kHz και long windows

- **Huffman code bits**

Αυτό το μέρος περιέχει τα bits της Huffman κωδικοποίησης. Οι πληροφορίες για την αποκωδικοποίησή τους βρίσκονται στις side information. Για τις τρεις περιοχές των big_values (region0, region1 και region2), η κωδικοποίηση γίνεται πάντα σε ζεύγη. Ένα παράδειγμα τέτοιας κωδικοποίησης δίνεται στον Πίνακα 3.16, όπου τα x και y είναι τα ζεύγη που κωδικοποιούνται, το hlen είναι το μήκος του κώδικα Huffman που θα χρησιμοποιηθεί και το hcod είναι ο ίδιος ο κώδικας. Για την περιοχή count1_values, η κωδικοποίηση γίνεται σε τετράδες αριθμών ενώ για την περιοχή rzero δε χρησιμοποιείται η κωδικοποίηση Huffman.

X	y	Hlen	hcod
0	0	1	1
0	1	3	010
0	2	6	001010
0	3	8	00010011
0	4	8	00010000
0	5	9	000001010
1	0	3	011
1	1	4	0011
1	2	6	000111
1	3	7	0001010
1	4	7	0000101

1	5	8	00000011
2	0	6	001011
2	1	5	00100
2	2	7	0001101
2	3	8	00010001
2	4	8	00001000
2	5	9	000000100
3	0	7	0001100
3	1	7	0001011
3	2	8	00010010
3	3	9	000001111
3	4	9	000001011
3	5	9	000000010
4	0	7	0000111
4	1	7	0000110
4	2	8	00001001
4	3	9	000001110
4	4	9	000000011
4	5	10	0000000001
5	0	8	00000110
5	1	8	00000100
5	2	9	000000101
5	3	10	0000000011
5	4	10	0000000010
5	5	10	0000000000

Πίνακας 3.16: Πίνακας κωδικοποίησης Huffman 7

Η κωδικοποίηση Huffman διαφέρει ανάλογα με τον τύπο παραθύρου που χρησιμοποιείται. Αν χρησιμοποιούνται μεγάλα παράθυρα, η κωδικοποίηση γίνεται ανάλογα με την αύξηση της συχνότητας.

3.2.1.5 Ancillary Data

Τα βοηθητικά δεδομένα είναι χρησιμοποιούνται προαιρετικά και δεν έχουν καθορισμένο αριθμό bits. Βρίσκονται πάντα μετά bits της κωδικοποίησης Huffman και μπορούν να έχουν έκταση μέχρι και το σημείο main_data_begin του επόμενου frame.

3.2.2 ID3

Αν και το MP3 κατάφερε να συμπιέσει αποτελεσματικά τα αρχεία ήχου χωρίς ιδιαίτερα ακουστές διαφορές, δεν είχε ποτέ τη δυνατότητα να αποθηκεύσει πληροφορίες κειμένου. Γι' αυτό το λόγο κατασκευάστηκε μια ετικέτα (tag) μήκους 128 bytes η οποία τοποθετείται στο τέλος του αρχείου. Η ετικέτα αυτή ονομάζεται ID3 και περιέχει πεδία για τις πληροφορίες του αρχείου, χωρισμένα σε κατηγορίες, όπως φαίνεται στον Πίνακα 3.17.

“TAG”	Title	Artist	Album	Year	Comment	“0”	Track	Genre
3 bytes	30 bytes	30 bytes	30 bytes	4 bytes	28 bytes		1 byte	1 byte

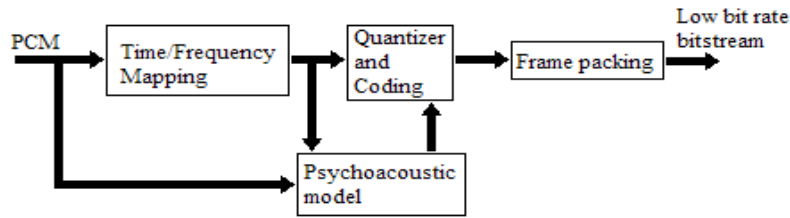
Πίνακας 3.17: Πληροφορίες του ID3 (ID3v1.1)

Στο πεδίο Genre η τιμή που χρησιμοποιείται λαμβάνεται από έναν προκαθορισμένο πίνακα. Το “0” χρησιμοποιείται για να γεμίσει το κενό που δημιουργείται.

Λόγω της μικρής χωρητικότητας σε πληροφορίες του ID3, αργότερα βγήκαν άλλες εκδόσεις (ID3v2) οι οποίες έκαναν εφικτή την καταχώρηση επιπλέον πληροφοριών. Η έκδοση που χρησιμοποιείται σήμερα είναι η ID3v2.4.

3.3 Κωδικοποίηση

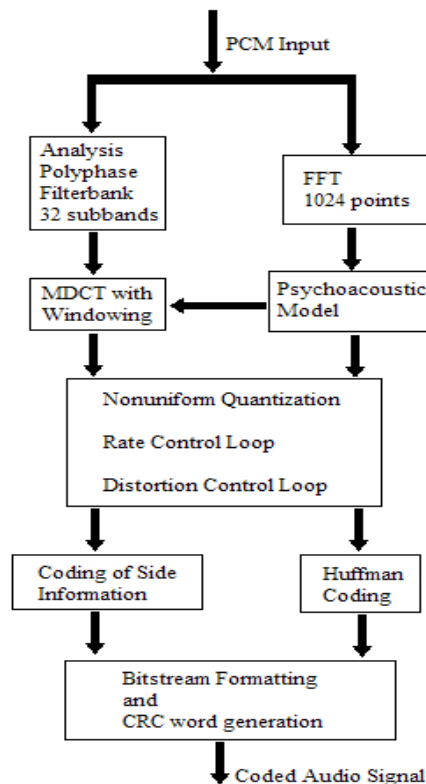
Οι κωδικοποιητές που λειτουργούν με τις αρχές της απωλεστικής συμπίεσης αναλύουν το συχνοτικό φάσμα και το πλάτος του εισερχόμενου σήματος και το συγκρίνουν με ένα τεχνητό μοντέλο της ανθρώπινης ακοής.



Σχήμα 3.4: Βασική δομή κωδικοποιητή

Η διαδικασία κωδικοποίησης του MPEG-1 Layer III είναι ιδιαίτερα πολύπλοκη. Όπως φαίνεται και στο Σχήμα 3.5, το εισερχόμενο σήμα περνά από πολλά στάδια επεξεργασίας προκειμένου να κωδικοποιηθεί καταλλήλως.

Στα στάδια αυτά περιλαμβάνεται ανάλυση της τράπεζας φίλτρων, μετασχηματισμός FFT, μετασχηματισμός MDCT, επεξεργασία μέσω ψυχοακουστικού μοντέλου, κωδικοποίηση Huffman κ.α, τα οποία θα αναλυθούν εκτενώς παρακάτω.



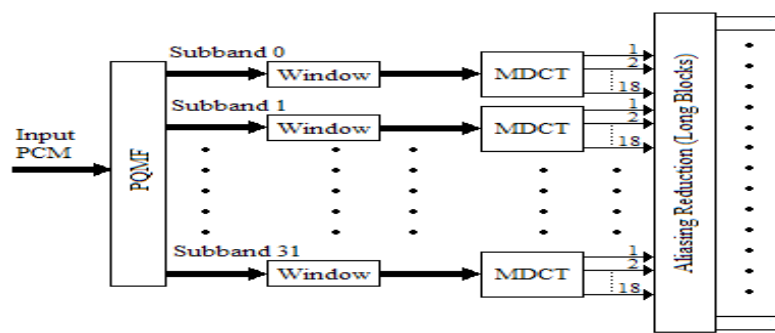
Σχήμα 3.5: Κωδικοποιητής MPEG-1 Layer III

3.3.1 Analysis Polyphase Filterbank

Τα πολυφασικά φίλτρα (PQMF), αποτελούνται από μια σειρά από ίσου μεγέθους φίλτρα με συσχετιζόμενες φάσεις. Μια πολυφασική τράπεζα φίλτρων θα μπορούσε εύκολα να κατασκευαστεί με ένα FIR φίλτρο το οποίο φιλτράρει το φάσμα και το αποδεκατίζει μέσω μικρότερων φίλτρων. Τα φίλτρα αυτά κατασκευάζονται τροποποιώντας ένα απλό lowpass FIR φίλτρο με διαφορετικές φάσεις.

Εδώ, μια ακολουθία από 1152 δείγματα PCM φιλτράρεται και χωρίζεται σε 32 ίσες συχνοτικές υποζώνες, το μήκος των οποίων εξαρτάται από τη συχνότητα Nyquist του αρχικού PCM σήματος. Εάν η PCM συχνότητα δειγματοληψίας είναι 44.1kHz τότε η συχνότητα Nyquist είναι 22.05kHz. Συνεπώς κάθε υποζώνη θα έχει έκταση $22050/32=689\text{Hz}$. Η πρώτη υποζώνη θα εκτείνεται από 0 - 688Hz, η δεύτερη από 689 – 1378Hz κλπ. Κάθε δείγμα μπορεί να περιέχει στοιχεία από ολόκληρο το φάσμα και ανάλογα με τα στοιχεία αυτά φιλτράρεται στην κατάλληλη υποζώνη. Αυτό σημαίνει ότι ο αριθμός των δειγμάτων έχει πλέον αυξηθεί κατά 32 φορές, μιας και κάθε υποζώνη αποθηκεύει ένα υπομέρος κάθε δείγματος.

Η έξοδος των PQMF οδηγεί στον MDCT μέσω του οποίου δίνεται η δυνατότητα μεγαλύτερης συχνοτικής ανάλυσης.



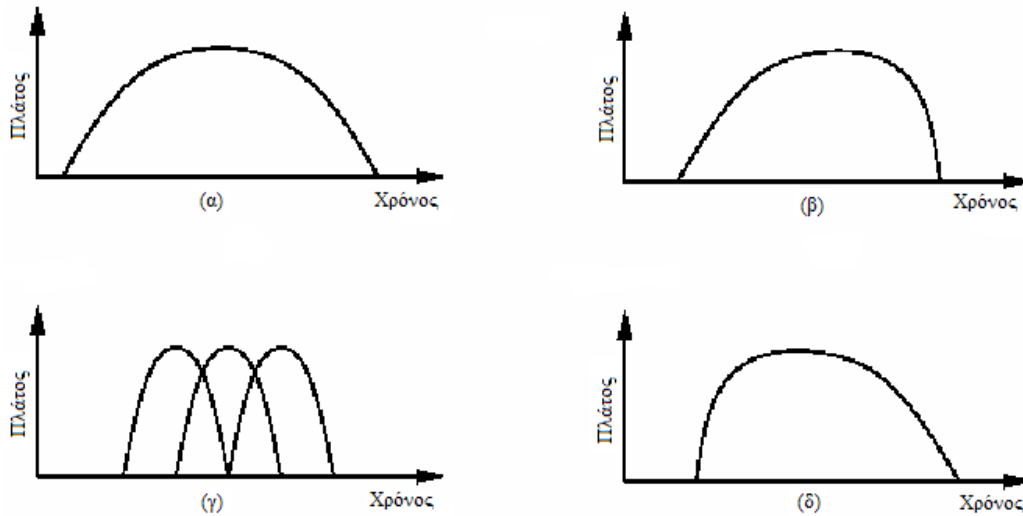
Σχήμα 3.6: Διαδικασία παραθυροποίησης και MDCT⁵

⁵ Το σχήμα αυτό αποτελεί παραλλαγή αντίστοιχου σχήματος από: Bosi, M & Goldberg R. E. (2003). Introduction to digital audio coding and standards. ISBN: 1-4020-7357-7. 274

3.3.2 Modified Discrete Transform (MDCT)

Όπως φαίνεται και στο Σχήμα 3.6, με την εφαρμογή του MDCT, οι συχνοτικές γραμμές πολλαπλασιάζονται με το 18, δίνοντας έτσι συνολικά $32 \cdot 18 = 576$ συχνοτικές γραμμές. Πριν την εφαρμογή του, όμως, το σήμα της κάθε υποζώνης πρέπει να «παραθυροποιηθεί».

Η διαδικασία αυτή γίνεται για να διορθωθούν ορισμένα ψεγάδια που προκαλούνται από τις άκρες του χρονικά πεπερασμένου μέρους του σήματος. Υπάρχουν τέσσερεις διαφορετικοί τύποι παραθύρων που ορίζονται από το πρότυπο, όπως φαίνονται στο Σχήμα 2.7. Ανάλογα με το βαθμό της στατικότητας, το ψυχοακουστικό μοντέλο καθορίζει τον τύπο παραθύρου που θα χρησιμοποιηθεί και προωθεί τις αντίστοιχες πληροφορίες.

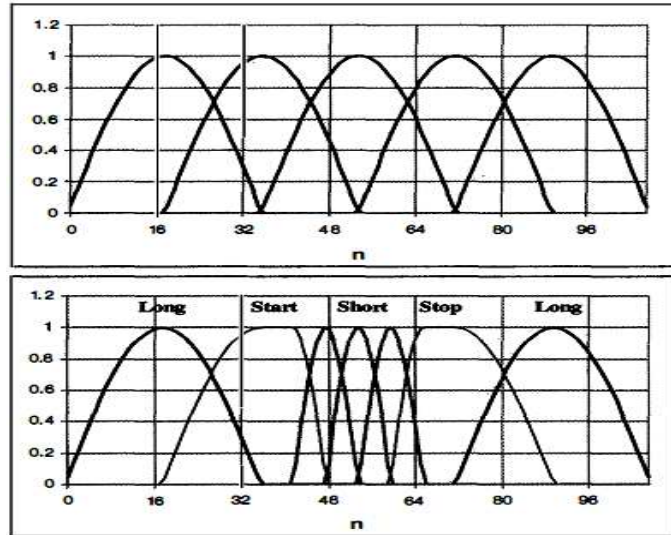


Σχήμα 3.7: Τύποι παραθύρων

(α) long window, (β) start window, (γ) short windows, (δ) stop window

Εάν το ψυχοακουστικό μοντέλο αποφασίσει ότι το frame που κωδικοποιείται εκείνη τη στιγμή έχει πολύ μικρές διαφορές από το προηγούμενο, τότε θα μπει ένα long window το οποίο θα αυξήσει τη φασματική ανάλυση που δίνεται από το MDCT. Εναλλακτικά, εάν

τα δύο frames παρουσιάζουν μεγάλες διαφορές μεταξύ τους, τότε θα μπου short windows. Αυτός ο τύπος αποτελείται από τρία short windows τα οποία επικαλύπτονται μεταξύ τους και θα αυξήσουν τη χρονική ανάλυση που δίνεται από το MDCT. Η μεγάλη χρονική ανάλυση είναι απαραίτητη για τη διόρθωση ορισμένων ατελειών. Προκειμένου να επιτευχθεί καλύτερη προσαρμογή όταν απαιτείται μεταβολή των παραθύρων, ορίζονται δύο άλλοι τύποι, τα start window και stop window.



Σχήμα 3.8: Τυπική αλληλουχία παραθύρων⁶

Ένα long window μετατρέπεται σε start window όταν ακολουθείται από short window. Όμοια, ένα long window γίνεται stop window όταν προηγείται αυτού ένα short window.

Οι διάφοροι τύποι παραθύρων προκύπτουν από τέσσερεις διαφορετικούς τύπους, οι οποίοι παρατίθενται στην Εικόνα 3.2.

⁶ Το σχήμα αυτό προέρχεται από: Bosi, M & Goldberg R. E. (2003). Introduction to digital audio coding and standards. ISBN: 1-4020-7357-7. 277.

$$w[n] = \sin\left(\frac{\pi}{36}\left(n + \frac{1}{2}\right)\right) \quad n = 0, \dots, 35 \quad (\text{long window}) \quad (2.2)$$

$$w[n] = \sin\left(\frac{\pi}{12}\left(n + \frac{1}{2}\right)\right) \quad n = 0, \dots, 11 \quad (\text{short window}) \quad (2.3)$$

$$w[n] = \begin{cases} \sin\left(\frac{\pi}{36}\left(n + \frac{1}{2}\right)\right) & n = 0, \dots, 17 \\ 1 & n = 18, \dots, 23 \quad (\text{start window}) \\ \sin\left(\frac{\pi}{12}\left(n - 18 + \frac{1}{2}\right)\right) & n = 24, \dots, 29 \\ 0 & n = 30, \dots, 35 \end{cases} \quad (2.4)$$

$$w[n] = \begin{cases} 0 & n = 0, \dots, 5 \\ \sin\left(\frac{\pi}{12}\left(n - 6 + \frac{1}{2}\right)\right) & n = 6, \dots, 11 \\ 1 & n = 12, \dots, 17 \quad (\text{stop window}) \\ \sin\left(\frac{\pi}{36}\left(n + \frac{1}{2}\right)\right) & n = 18, \dots, 35 \end{cases} \quad (2.5)$$

Εικόνα 3.2: Εξισώσεις υπολογισμού των παραθύρων

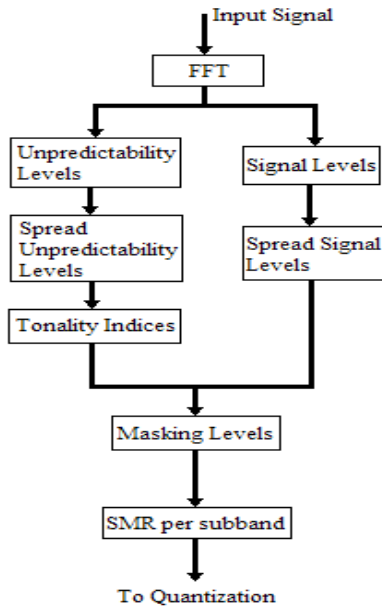
Το aliasing που προκύπτει από το PQMF αφαιρείται με σκοπό τη μείωση του όγκου πληροφοριών. Αυτό επιτυγχάνεται με μια σειρά υπολογισμών που προσθέτουν υποζώνες με τέτοιο τρόπο ώστε να προκαλούνται ακυρώσεις.

3.3.3 FFT

Παράλληλα με την επεξεργασία του σήματος από την τράπεζα φίλτρων γίνεται και ο μετασχηματισμός του σήματος στο πεδίο των συχνοτήτων μέσω ενός Μετασχηματισμού Fourier (FFT). Λαμβάνουν χώρα δύο μετασχηματισμοί για κάθε 1152 PCM δείγματα, ένας με 1024 σημεία και ένας με 256. Αυτό συμβαίνει με σκοπό την αύξηση της συχνοτικής ανάλυσης και των πληροφοριών των φασματικών αλλαγών με την πάροδο του χρόνου.

3.3.4 Psychoacoustic Model

Το ψυχοακουστικό μοντέλο που χρησιμοποιείται από το MP3 για την κωδικοποίηση ονομάζεται Μοντέλο 2 και είναι εξαιρετικά πολύπλοκο και λεπτομερές. Στο Σχήμα 3.9 φαίνεται το διάγραμμα του. Όπως φαίνεται εκεί, υπάρχουν δύο ταυτόχρονοι δρόμοι που ακολουθούνται: αυτός της ενέργειας επικάλυψης και αυτός του δείκτη τονικότητας.



Σχήμα 3.9: Διάγραμμα του ψυχοακουστικού μοντέλου

3.3.4.1 Στάδιο Ανάλυσης

Το μοντέλο εφαρμόζει FFT με παράθυρα Hanning στο ευθυγραμμισμένο χρονικά σήμα. Το μέγεθος των παραθύρων είναι 1024 σημεία. Εδώ πρέπει να σημειωθεί ότι στο Layer III εφαρμόζονται δύο FFT σε κάθε frame, ένας με μέγεθος 1024 σημείων και ένας με 256 σημεία. Η έξοδος της FFT ανάλυσης χρησιμοποιείται για τον υπολογισμό των καμπύλων επικάλυψης και σχετίζονται με τα signal-to-mask ratios των υποζωνών. Λόγω της χρήσης δύο FFT, υπολογίζονται και δύο signal-to-mask ratios. Το SMR που επιλέγεται για την κάθε υποζώνη είναι το μεγαλύτερο από τα δύο αυτά.

3.3.4.2 Υπολογισμός SPL

Οι συχνοτικές γραμμές ομαδοποιούνται σε μέρη που ονομάζονται “threshold calculation partitions”, των οποίων το μέγεθος είναι περίπου το 1/3 μιας critical band. Για κάθε μέρος επιλέγεται ένας masker SPL ο οποίος προκύπτει από το άθροισμα των πυκνωμάτων ενέργειας στο συγκεκριμένο μέρος.

Η συνολική επικαλυπτική ενέργεια για ένα frame υπολογίζεται με τη συνέλιξη μιας εξίσωσης με κάθε ένα από τους maskers που υπάρχουν στο σήμα. Αυτή η διαδικασία είναι αντίστοιχη με την εξάπλωση της ενέργειας κάθε masker σε όλες τις συχνότητες και με την πρόσθεση των σχετικών ενεργειών. Η εξίσωση που χρησιμοποιείται ονομάζεται “spreading function” και είναι μια παραλλαγή μιας εξίσωσης Schroeder.

3.3.4.3 Spreading Function

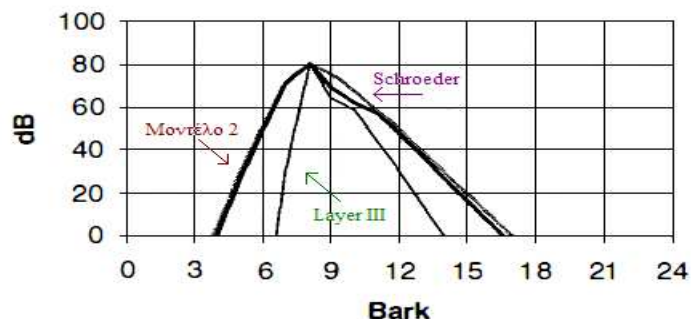
Η βασική spreading function που χρησιμοποιείται από το Μοντέλο 2 αυτό είναι η σχέση (2.6).

$$B(dz) = 15.8111389 + 7.5 * (1.05 * dz + 0.474) - 17.5 * \sqrt{1 + (1.05 * dz + 0.474)^2} + 8 * \text{MIN}(0, (1.05 * dz - 0.5)^2 - 2 * (1.05 * dz - 0.5)) \quad (2.6)$$

όπου το dz είναι η απόσταση σε bark ανάμεσα στον masker και τον maskee.

Πρέπει να σημειωθεί ότι το Layer III χρησιμοποιεί αυτή την εξίσωση με μικρές τροποποιήσεις. Οι όροι (1.05*dz) της εξίσωσης αντικαθίστανται με (1.5*dz) εάν το dz είναι αρνητικό (συχνότητα masker < συχνότητα maskee) και με (3*dz) εάν το dz είναι θετικό (συχνότητα masker > συχνότητα maskee).

Στο Σχήμα 3.10 φαίνονται συγκριτικά οι τρεις spreading functions, η Schroeder, η εξίσωση του Μοντέλου 2 και η παραλλαγή που χρησιμοποιείται στο MP3.



Σχήμα 3.10: Σύγκριση των spreading functions⁷

Η spreading function είναι διαμορφώνεται έτσι ώστε να διατηρεί την ενέργεια στα δύο όρια του συχνοτικού φάσματος. Δηλαδή, ένα flat φάσμα το οποίο είναι πάνω από το θετικό συχνοτικό εύρος του FFT κάνει συνέλιξη με την $B(dz)$. Έτσι, η ενέργεια του flat φάσματος πέφτει στα όρια του θετικού συχνού εύρους λόγω έλλειψης φασματικής ενέργειας εκτός του εύρους. Η spreading function $B(dz)$ προσαρμόζεται έτσι ώστε να απομακρύνει αυτή την πτώση με το να διαιρεί την ενέργεια του σήματος μετά τη συνέλιξη διά την ενέργεια του flat φάσματος μετά τη συνέλιξη. Τέλος, η ενέργεια του τελικού σήματος μειώνεται σε κάθε μέρος με ρυθμό που εξαρτάται από την τονικότητα του φάσματος στο συγκεκριμένο μέρος.

3.3.4.4 Tonality Index

Ο δείκτης τονικότητας σε κάθε μέρος υπολογίζεται με βάση την προβλεψιμότητα του σήματος των συχνοτικών γραμμών των δύο προηγούμενων frames. Για κάθε frame m και για κάθε συχνοτική γραμμή k , το πλάτος $A_m[k]$ και η φάση $\phi_m[k]$ του σήματος υπολογίζονται με τις σχέσεις (2.7) και (2.8).

$$A'_m[k] = A_{(m-1)}[k] + \{A_{(m-1)}[k] - A_{(m-2)}[k]\} \quad (2.7)$$

$$\phi'_m[k] = \phi_{(m-1)}[k] + \{\phi_{(m-1)}[k] - \phi_{(m-2)}[k]\} \quad (2.8)$$

⁷ Το σχήμα αυτό αποτελεί παραλλαγή αντίστοιχου σχήματος από: Bosi, M & Goldberg R. E. (2003). Introduction to digital audio coding and standards. ISBN: 1-4020-7357-7. 189

όπου τα $A'_m[k]$ και $\phi'_m[k]$ είναι οι προβλεπόμενες τιμές.

Στη συνέχεια οι τιμές αυτές χρησιμοποιούνται για τον ορισμό του μέτρου μη προβλεψιμότητας (unpredictability measure), μέσα από τη σχέση (2.9).

$$c_m[k] = \frac{\sqrt{\{A'_m[k] \cos \phi'_m[k] - A_m[k] \cos \phi_m[k]\}^2 + \{A'_m[k] \sin \phi'_m[k] - A_m[k] \sin \phi_m[k]\}^2}}{A_m[k] + |A'_m[k]|} \quad (2.9)$$

όπου το $c_m[k]$ είναι ίσο με το μηδέν όταν η τιμή προβλέπεται ακριβώς και ίσο με ένα όταν η ενέργεια ενός από τα δύο σήματα (του προβλεπόμενου και του πραγματικού) έχει πολύ μεγάλη διαφορά από την ενέργεια του άλλου.

Το μέτρο της μη προβλεψιμότητας συνδυάζεται με την ενέργεια σε κάθε μέρος, εξάγοντας το μέτρο προβλεψιμότητας για κάθε μέρος (partition unpredictability measure). Στη συνέχεια πραγματοποιείται η συνέλιξη του μέτρου αυτού με τη spreading function. Το αποτέλεσμα αυτής της συνέλιξης ομαλοποιείται και συνδυάζεται με το δείκτη τονικότητας του οποίου οι τιμές κυμαίνονται ανάμεσα στο μηδέν και το ένα. Με βάση τη λογική του δείκτη, η μεγάλη μη προβλεψιμότητα τείνει στο μηδέν και η μικρή στο ένα. Ο δείκτης τονικότητας σε συγκεκριμένη συχνοτική περιοχή αναπαριστά την τονικότητα των κύριων maskers της περιοχής μιας και το μέτρο μη προβλεψιμότητας συνδυάζεται μέσω συνέλιξης με τη spreading function που καθορίζει την ενέργεια επικάλυψης της περιοχής αυτής.

3.3.4.5 Κατώφλι Επικάλυψης

Οι δείκτες τονικότητας που προκύπτουν στο προηγούμενο στάδιο χρησιμοποιούνται για τον υπολογισμό της μετατόπισης $\Delta(z)$ σε dB της ομαλοποιημένης ενέργειας του σήματος η οποία μετατρέπεται στο ολικό level επικάλυψης. Οι τιμές της $\Delta(z)$ έχουν γραμμική κατανομή με βάση το δείκτη τονικότητας από τα 6dB για μηδενική τονικότητα (ο masker είναι θόρυβος) έως τα 29dB για τονικότητα ίση με ένα (ο masker είναι ένας τόνος).

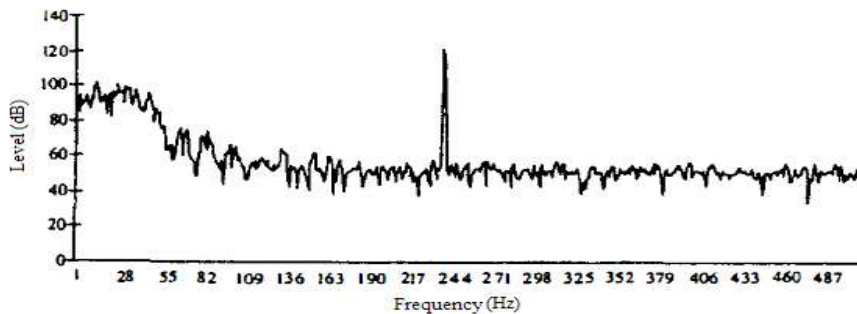
Για τη μείωση των pre-echoes, το πρότυπο συγκρίνει το κατώφλι επικάλυψης του τωρινού frame με τα κατώφλια των δύο προηγούμενων και το αντικαθιστά με τη

μικρότερη τιμή των δύο πριν το κατώφλι ακουστότητας. Αυτό εξυπηρετεί στη μείωση των κατωφλίων επικάλυψης που ακολουθούν πολύ μικρής έντασης σήματα.

3.3.4.6 Υπολογισμός Signal-to-mask Ratio (SMR)

Τα level των κατωφλίων επικάλυψης για κάθε μέρος συγκρίνονται με τα κατώφλια ησυχίας και χρησιμοποιείται το μεγαλύτερο από αυτά. Για κάθε μέρος η ενέργεια επικάλυψης χαρτογραφείται ομοιόμορφα σε κάθε συχνοτική γραμμή διαιρώντας τις τιμές των κατωφλίων επικάλυψης κάθε μέρους με τον αριθμό των συχνοτικών γραμμών του συγκεκριμένου μέρους. Τέλος, υπολογίζεται το SMR για κάθε υποζώνη συνδυάζοντας την πυκνότητα της ενέργειας του φάσματος και το κατώφλι επικάλυψης με τις ζώνες που περιλαμβάνουν τους scalefactors. Τα signal-to-mask ratios που προκύπτουν από αυτούς τους υπολογισμούς στέλνονται στη διαδικασία κατανομής για τον ορισμό των της κατανομής των bits σε κάθε υποζώνη.

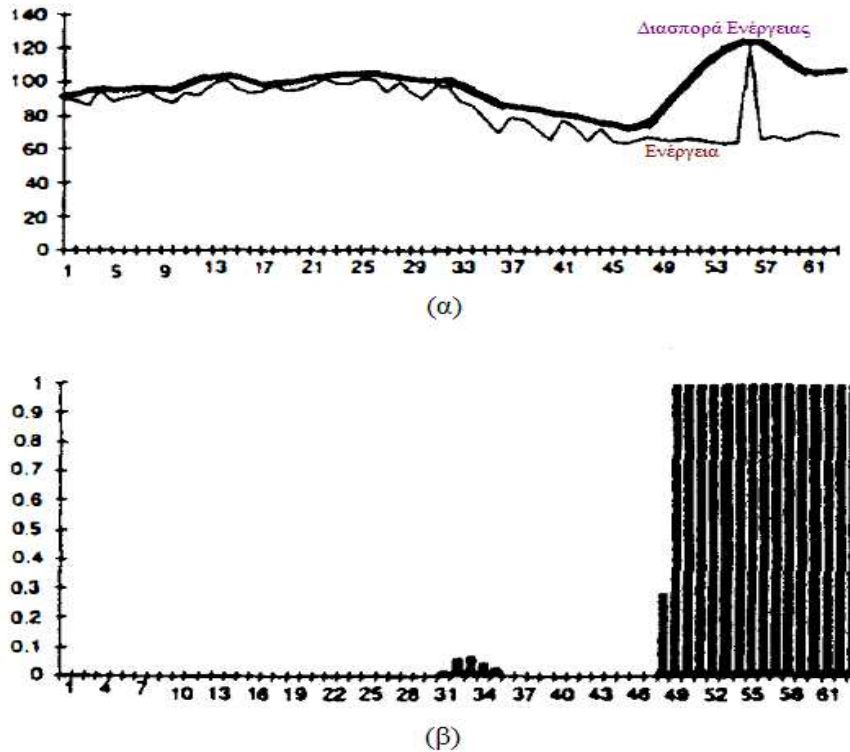
Ακολουθεί ένα παράδειγμα της λειτουργίας του ψυχοακουστικού μοντέλου. Στο Σχήμα 3.11 φαίνεται το συχνοτικό φάσμα ενός σήματος που αποτελείται από ένα τονικό στοιχείο το οποίο ξεχωρίζει από το υπόλοιπο φάσμα.



Σχήμα 3.11: Φάσμα σήματος⁸

⁸ Το σχήμα αυτό προέρχεται από: Bosi, M & Goldberg R. E. (2003). Introduction to digital audio coding and standards. ISBN: 1-4020-7357-7. 286

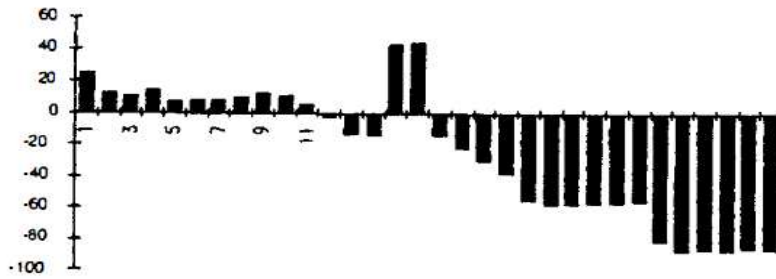
Στο Σχήμα 3.12 φαίνεται η ενέργεια του σήματος και η διασπορά της ενέργειας του (α) καθώς και ο δείκτης τονικότητας (β) για 62 μέρη, των οποίων το μέγεθος ισοδυναμεί με περίπου το 1/3 μιας critical band.



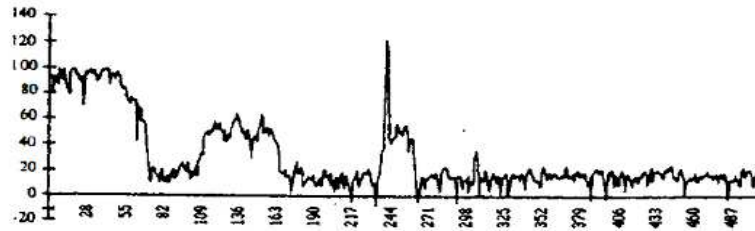
Σχήμα 3.12: (α) Ενέργεια και Διασπορά Ενέργειας (β) Δείκτης Τονικότητας⁹

Τέλος, στο Σχήμα 3.13 φαίνονται οι τιμές των signal-to-mask ratios για κάθε scalefactor ζώνη (α) καθώς και το τελικό κωδικοποιημένο σήμα με bitrate 64kbps (β).

⁹ Το σχήμα αυτό αποτελεί παραλλαγή αντίστοιχου σχήματος από: Bosi, M & Goldberg R. E. (2003). Introduction to digital audio coding and standards. ISBN: 1-4020-7357-7. 293



(α)



(β)

Σχήμα 3.13: (α) Signal-to-mask ratios (β) Κωδικοποιημένο σήμα στα 64kbps¹⁰

3.3.5 Nonuniform Quantization

Ο αλγόριθμος ελέγχου κατανομής που προτείνεται για τον κωδικοποιητή του Layer III χρησιμοποιεί δυναμικό κβαντισμό. Ο δυναμικός κβαντισμός περιέχει μια τεχνική που ονομάζεται κατανομή θορύβου (noise allocation) κατά την οποία ένας βρόγχος (loop) χρησιμοποιείται για τον υπολογισμό του βέλτιστου θορύβου κβαντισμού για κάθε υποζώνη. Αντί για την απευθείας κατανομή των bits από την τιμή του sound-noise-ratio (SNR), στην κατανομή θορύβου ο προσδιορισμός των bits γίνεται αυτόματα λόγω της στρατηγικής του noise allocation.

Για παράδειγμα, μπορεί να χρησιμοποιηθεί η μέθοδος analysis-by-synthesis για να υπολογιστεί ένα κβαντισμένο φάσμα το οποίο να ικανοποιεί τις απαιτήσεις θορύβου του ψυχοακουστικού μοντέλου. Χρησιμοποιούνται δύο βρόγχοι για την εύρεση δύο τιμών

¹⁰ Το σχήμα αυτό αποτελεί παραλλαγή αντίστοιχου σχήματος από: Bosi, M & Goldberg R. E. (2003). Introduction to digital audio coding and standards. ISBN: 1-4020-7357-7. 294

που χρησιμοποιούνται στην κατανομή: της τιμής *global gain* που καθορίζει το βήμα κβαντισμού και της τιμής των *scalefactors* που διαμορφώνουν την κατανομή του θορύβου για κάθε ζώνη. Για τη διαμόρφωση των *scalefactors*, τα *long windows* χωρίζονται σε 21 ζώνες και τα *short windows* σε 12 ζώνες. Ο διαχωρισμός αυτός σε ζώνες προσομοιάζει τις *critical bands* και ποικίλει ανάλογα με τη συχνότητα δειγματοληψίας.

Οι βρόγχοι που χρησιμοποιούνται ονομάζονται *rate control loop* και *distortion control loop*.

- **Rate control loop**

Ο βρόγχος αυτός καθορίζει το βήμα κβαντισμού και εκτελεί τον κβαντισμό στο πεδίο των συχνοτήτων. Επίσης, καθορίζει τη διαίρεση των *big_values* σε περιοχές, επιλέγει τον πίνακα Huffman που θα χρησιμοποιηθεί για κάθε περιοχή και υπολογίζει τα όρια τους.

Στην αρχή, τα δείγματα κβαντίζονται με αυξανόμενο βήμα κβαντισμού μέχρι οι κβαντισμένες τιμές να μπορούν να κωδικοποιηθούν με έναν μόνο πίνακα Huffman. Το μεγάλο βήμα κβαντισμού οδηγεί σε μικρές τιμές. Στη συνέχεια υπολογίζονται τα bits της κωδικοποίησης Huffman και συγκρίνονται με τον αριθμό των διαθέσιμων bits. Εάν το σύνολο των κωδικοποιημένων bits ξεπερνά τα διαθέσιμα bits, το βήμα κβαντισμού αυξάνεται περισσότερο και επαναλαμβάνεται όλη η διαδικασία μέχρι τα διαθέσιμα bits να είναι αρκετά.

Η μη γραμμικότητα εξασφαλίζεται με την ύψωση κάθε δείγματος εις την 3/4.

- **Distortion control loop**

Ο βρόγχος αυτός ελέγχει το θόρυβο κβαντισμού ο οποίος παράγεται από τον κβαντισμό μέσα στο *rate control loop*. Ο στόχος είναι να κρατηθεί ο θόρυβος κάτω από το κατώφλι επικάλυψης (*masking threshold*) για κάθε ζώνη. Το κατώφλι επικάλυψης είναι το όριο του επιτρεπτού θορύβου που ορίζεται από το ψυχοακουστικό μοντέλο.

Για τη διαμόρφωση του θορύβου, οι scalefactors κατανέμονται στις ζώνες που αντιστοιχεί στον κάθε ένα. Κατόπιν, υπολογίζεται ο θόρυβος. Η διαδικασία αυτή επαναλαμβάνεται μέχρι να μην υπάρχουν πλέον ζώνες των οποίων ο θόρυβος να ξεπερνά το masking threshold. Στις ζώνες που υπάρχει πολύς θόρυβος οι scalefactors αυξάνονται διαρκώς μέχρι να επιτευχθεί ο στόχος. Στο τέλος, ο εναπομένον θόρυβος δε θα γίνει αντιληπτός από το ανθρώπινο αφτί λόγω πολύ χαμηλής έντασης κι έτσι ο βρόγχος τελειώνει.

Υπάρχουν κάποιες περιπτώσεις στις οποίες οι βρόγχοι μπορεί να κάνουν άπειρες επαναλήψεις και η διαδικασία να μην ολοκληρώνεται ποτέ. Για την αποφυγή αυτής της κατάστασης, υπάρχουν ορισμένες προϋποθέσεις οι οποίες πρέπει να ελέγχονται:

Για την είσοδο των βρόγχων

- το μέγεθος των 576 φασματικών τιμών
- η επιτρεπόμενη παραμόρφωση των ζωνών
- τα διαθέσιμα bits για την κωδικοποίηση Huffman και την κωδικοποίηση των scalefactors
- ο αριθμός των bits και ο μέσος όρος των bits, όπως ορίζεται από την τιμή της ψυχοακουστικής εντροπίας του κάθε granule

Για την έξοδο των βρόγχων

- το μέγεθος των 576 κβαντισμένων τιμών
- οι scalefactors
- οι πληροφορίες για το βήμα κβαντισμού
- ο αριθμός των αχρησιμοποίητων bits για άλλη χρήση
- οι πληροφορίες που σχετίζονται με τον κώδικα Huffman (side information)
 - big_values
 - count1table_select
 - table_select

- region0_count και region1_count
- part2_3_length

3.3.6 Huffman Encoding

Οι κβαντισμένες τιμές φάσματος χωρίζονται σε τρεις περιοχές και κάθε περιοχή κωδικοποιείται με διαφορετικούς πίνακες Huffman.

Στις υψηλές συχνότητες, ο κωδικοποιητής φτιάχνει την περιοχή rzero, η οποία περιέχει μόνο μηδενικές τιμές. Το μέγεθος αυτή της περιοχής υπολογίζεται με βάση τα μεγέθη των δύο άλλων περιοχών και δεν χρειάζεται κωδικοποίηση. Ο μόνος περιορισμός που έχει είναι ότι πρέπει να περιέχει περιττό αριθμό μηδενικών τιμών επειδή οι άλλες δύο περιοχές περιέχουν άρτιο αριθμό τιμών.

Η δεύτερη περιοχή ονομάζεται count1 και περιλαμβάνει μια σειρά τιμών που αποτελούνται μόνο από τους αριθμούς -1, 0, 1. Σε αυτή την περιοχή η κωδικοποίηση γίνεται σε τετράδες τιμών και το μέγεθος της είναι πάντα πολλαπλάσιο του τέσσερα.

Η τρίτη και τελευταία περιοχή ονομάζεται big_values και καλύπτει τις υπόλοιπες φασματικές τιμές οι οποίες κωδικοποιούνται σε ζεύγη. Η περιοχή αυτή χωρίζεται σε τρία μέρη (region0, region1, region2), κάθε ένα από τα οποία κωδικοποιείται με διαφορετικό πίνακα Huffman.

Για την κωδικοποίηση χρησιμοποιείται ένα σύνολο από 16 πίνακες. Για κάθε τμήμα επιλέγεται ο πίνακας που ταιριάζει περισσότερο με τις στατιστικές του σήματος. Αυτός ο τρόπος επιλογής μεγιστοποιεί την αποτελεσματικότητα του κώδικα και μειώνει την πιθανότητα λαθών.

3.3.7 Coding of Side Information

Όλες οι παράμετροι που κατασκευάστηκαν από τον κωδικοποιητή συγκεντρώνονται σε αυτό το στάδιο με σκοπό την ευκολότερη ανακατασκευή του audio σήματος από τον

αποκωδικοποιητή. Οι παράμετροι αυτές κωδικοποιούνται και μπαίνουν στο side information του frame.

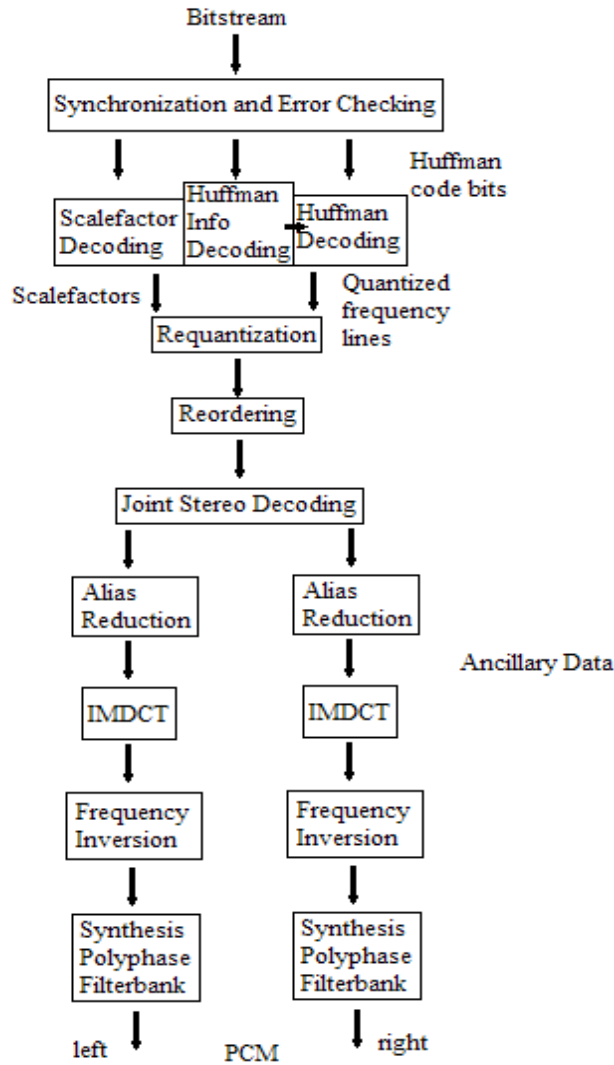
3.3.8 Bitstream Formatting and CRC word generation

Στο τελευταίο στάδιο παράγεται το οριστικό ρεύμα δεδομένων. Η επικεφαλίδα, το side information, το CRC και όλα τα επιμέρους στοιχεία συντίθενται και διαμορφώνουν τα frames. Κάθε ένα από αυτά τα frames αναπαριστούν 1152 κωδικοποιημένα PCM δείγματα.

3.4 Αποκωδικοποίηση

Το σημαντικότερο στοιχείο στη διαδικασία αποκωδικοποίησης είναι ο συγχρονισμός του αποκωδικοποιητή με το εισερχόμενο ρεύμα δεδομένων. Συγχρονισμός είναι η διαδικασία κατά την οποία ο αποκωδικοποιητής εντοπίζει τη θέση της πρώτης επικεφαλίδας καθώς και των επόμενων. Όταν πραγματοποιηθεί αυτό, ο αποκωδικοποιητής είναι σε θέση να γνωρίζει ολόκληρη τη ροή των δεδομένων και μπορεί να ξεκινήσει την αποκωδικοποίηση. Στο Σχήμα 3.14 φαίνεται το διάγραμμα ροής του αποκωδικοποιητή που χρησιμοποιείται από το MP3.

Στο πρώτο στάδιο (frame unpacking) περιλαμβάνεται η εύρεση της επικεφαλίδας, η αποκωδικοποίηση των side information, η αποκωδικοποίηση των scalefactors και των δεδομένων Huffman. Στο δεύτερο στάδιο (reconstruction) υπάρχει ο επανακβαντισμός και η ανακατάταξη του φάσματος. Στο τρίτο στάδιο (inverse mapping) γίνεται η επεξεργασία του joint stereo εάν υπάρχει, το alias reduction, η σύνθεση μέσω IMDCT (Inverse Modified Discrete Cosine Transform) και polyphase filterbank.



Σχήμα 3.14: Διάγραμμα αποκωδικοποιητή MPEG-1 Layer III

3.4.1 Synchronization and Error Checking

Σε αυτό το πεδίο εισέρχονται τα προς αποκωδικοποίηση δεδομένα. Ο αποκωδικοποιητής πρέπει να εντοπίσει την επικεφαλίδα κάθε frame. Χωρίς αυτή τη διαδικασία, η αποκωδικοποίηση δεν μπορεί να προχωρήσει.

3.4.2 Huffman Decoding and Huffman info Decoding

Η αποκωδικοποίηση των πληροφοριών Huffman γίνεται πρώτη διότι μέσω αυτών εντοπίζει ο αποκωδικοποιητής την αρχή των δεδομένων. Το μέρος αυτό εξασφαλίζει

επίσης ότι οι συχνοτικές γραμμές είναι 576. Σε περίπτωση που τα δεδομένα έχουν λιγότερες γραμμές, το μέρος αυτό συμπληρώνει τις κενές γραμμές με μηδενικά.

Όταν όλες οι παράμετροι έχουν διευκρινιστεί, τότε ξεκινά η αποκωδικοποίηση των δεδομένων Huffman. Πρώτα αποκωδικοποιούνται οι γραμμές που περιέχουν big values και μετά οι small values. Η αποκωδικοποίηση γίνεται μέσω προκαθορισμένων πινάκων που χρησιμοποιούνται από το πρότυπο και οι πληροφορίες γι' αυτούς βρίσκονται στο side information. Η αποκωδικοποίηση ολοκληρώνεται είτε μόλις όλα τα Huffman δεδομένα αποκωδικοποιηθούν είτε μόλις συμπληρωθούν οι 576 συχνοτικές γραμμές. Σε περίπτωση που περιέχονται περισσότερες από 576 γραμμές στα δεδομένα, οι εναπομένουσες διαγράφονται.

3.4.3 Scalefactor decoding

Κάθε granule περιέχει πρώτα τους scalefactors και μετά τα raw samples του Huffman. Πριν την ανάγνωση των scalefactors πρέπει πρώτα να διαπιστωθεί αν τα granules μοιράζονται κοινούς συντελεστές. Εάν ισχύει αυτό, τότε οι scalefactors μεταδίδονται μόνο από το πρώτο granule.

Στο ρεύμα δεδομένων βρίσκονται μόνο οι δείκτες των scalefactors. Αυτοί οι δείκτες αποθηκεύονται σε δύο σειρές και οι περισσότεροι από αυτούς επαναχρησιμοποιούνται από το δεύτερο granule. Η ανάγνωση των δεικτών αυτών γίνεται μέσω των slen1 και slen2, τα οποία αποκωδικοποιούνται από το scalefac_compress (Πίνακας 3.11).

Ο αριθμός των bits που χρησιμοποιούνται για την κωδικοποίηση των scalefactors ονομάζεται part2_length και υπολογίζεται σύμφωνα με τις παρακάτω σχέσεις:

Για block type = 0, 1, 3

$$part2_length = 11*slen1 + 10*slen2 \quad (2.10)$$

Για block type = 2 και mixed_block_flag = 0

$$part2_length = 18*slen1 + 18*slen2 \quad (2.11)$$

Για block type = 2 και mixed_block_flag = 1

$$part2_length = 17*slen1 + 18*slen2 \quad (2.12)$$

Όπως φαίνεται από τις σχέσεις (2.10), (2.11) και (2.12), ο αριθμός των bits εξαρτάται από το μέγεθος παραθύρου που χρησιμοποιείται κάθε φορά. Τα block type 0, 1, 3 αντιστοιχούν σε long windows ενώ το 2 σε short windows.

3.4.4 Requantization

Κβαντισμός ονομάζεται η διαδικασία κατά την οποία ένας πραγματικός αριθμός με σχεδόν απόλυτη ακρίβεια παρμένος από μια ομάδα πιθανών τιμών μετατρέπεται σε έναν ακέραιο αριθμό. Αυτό γίνεται κατά την κωδικοποίηση. Κατά την αποκωδικοποίηση, η διαδικασία του κβαντισμού αντιστρέφεται για τη δημιουργία των συχνοτικών γραμμών. Οι ακέραιοι αριθμοί των 576 συχνοτικών γραμμών που προκύπτουν από την κωδικοποίηση Huffman επανακβαντίζονται και χωρίζονται σε κλίμακες. Ο επανακβαντισμός γίνεται ξεχωριστά για τα short και long windows, σύμφωνα με τις σχέσεις (2.5) και (2.8) αντίστοιχα. Η διαίρεση σε κλίμακες γίνεται με τον πολλαπλασιασμό των τιμών αυτών με τους αντίστοιχους scalefactors και αποθηκεύονται.

Για short windows

$$xr[i] = sign(is[i]) * is[i]^{4/3} * 2^A * 2^B \quad (2.13)$$

όπου: το [i] είναι η αποκωδικοποιημένη τιμή Huffman που βρίσκεται στο δείκτη buffer, το xr[i] είναι η είσοδος στο synthesis filterbank και οι όροι A και B δίνονται από τις σχέσεις (2.14) και (2.15) αντίστοιχα.

$$A = 1/4*(global_gain[gr] - 210 - 8*subblock_gain>window][gr]) \quad (2.14)$$

$$B = -(scalefac_multiplier*scalefac_s[gr][ch][sfb][window]) \quad (2.15)$$

Για long windows

$$xr[i] = \text{sign}(is[i]) * is[i]^{4/3} * 2^C * 2^D \quad (2.16)$$

όπου οι όροι C και D δίνονται από τις σχέσεις (2.17) και (2.18) αντίστοιχα.

$$C = 1/4 * (\text{global_gain}[gr] - 210) \quad (2.17)$$

$$D = - (\text{scalefac_multiplier} * (\text{scalefac_l}[sfb][ch][gr][window] + \text{preflag}[gr] * \text{pretab}[sfb])) \quad (2.18)$$

Εάν η διαφορά μεταξύ δύο διαδοχικών frames είναι πολύ μικρή, τότε χρησιμοποιείται το long window. Διαφορετικά, όταν το σήμα της υποζώνης παρουσιάζει σημαντικές διαφορές μεταξύ δύο frames τότε χρησιμοποιείται το short window.

3.4.5 Reordering

Ο επαναπροσδιορισμός του φάσματος εξαρτάται από τον τύπο παραθύρου που έχει χρησιμοποιηθεί πριν τη διαδικασία IMDCT. Εάν χρησιμοποιείται short window τότε οι συχνотικές γραμμές θα κατασκευαστούν πρώτα με βάση τις υποζώνες, μετά με βάση το παράθυρο και τέλος με βάση τη συχνότητα. Με αυτή την κατάταξη αυξάνεται η αποτελεσματικότητα του κώδικα Huffman. Όταν χρησιμοποιούνται long windows οι γραμμές κατασκευάζονται πρώτα με βάση τις υποζώνες και μετά με βάση τη συχνότητα.

3.4.6 Stereo Processing

Σε αυτό το στάδιο τα επανακβαντισμένα δείγματα επεξεργάζονται για την απόδοση του τύπου stereo που θα χρησιμοποιηθεί. Οι λεπτομέρειες για τον τύπο αυτό λαμβάνονται από το mode_extension της επικεφαλίδας του frame.

Τα κανάλια του τυπικού stereo σήματος δεν είναι ανεξάρτητα και το joint stereo εκμεταλλεύεται αυτές τις ομοιότητες. Η επεξεργασία του joint stereo είναι πολύπλοκη επειδή η διαχείριση των long windows διαφέρει από αυτή των short. Ακόμη, τα granules ενδέχεται να περιέχουν ένα μείγμα από short και long windows και οι ζώνες σε κάθε granule μπορεί να συνδυάζονται με διαφορετικούς τύπους stereo. Υπάρχουν δύο τύποι stereo: το Mid/Side και το Intensity stereo.

3.4.6.1 Mid/Side stereo

Αυτός ο τύπος είναι διαθέσιμος μόνο για το Layer III του προτύπου. Εδώ, αντί να μεταδίδονται ξεχωριστά το left και right channel, μεταδίδεται ένας συνδυασμός αυτών, μέσω δύο νέων σημάτων, του mid σήματος και του side σήματος. Το mid σήμα $[M(i)]$ εξάγεται από την πρόσθεση των δύο καναλιών ενώ το side σήμα $[S(i)]$ εξάγεται από την αφαίρεση του δεξιού καναλιού από το αριστερό. Για να επανακατασκευαστούν τα σήματα των δύο καναλιών χρησιμοποιούνται οι σχέσεις (2.19) και (2.20).

Για το αριστερό κανάλι

$$L(i) = \frac{1}{\sqrt{2}} * [M(i) + S(i)] \quad (2.19)$$

Για το δεξί κανάλι

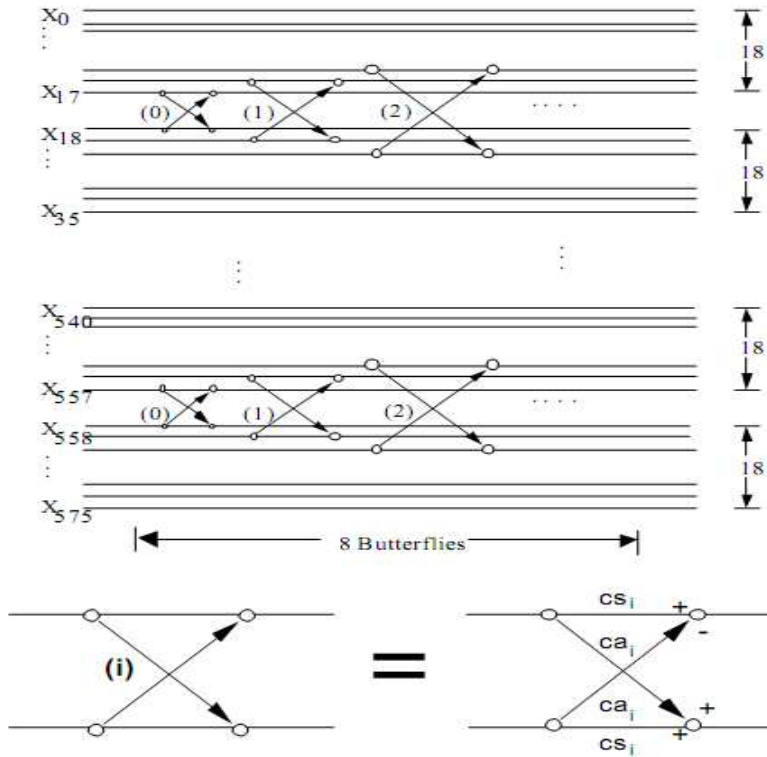
$$R(i) = \frac{1}{\sqrt{2}} * [M(i) - S(i)] \quad (2.20)$$

3.4.6.2 Intensity Stereo

Στον τύπο αυτό και τα δύο κανάλια μοιράζονται το ίδιο σήμα, όχι όμως και την ίδια ένταση του σήματος. Με τον τρόπο αυτό, ο ήχος φτάνει δυνατότερος στο αφτί που «βλέπει» στην πηγή σε σχέση με το άλλο αφτί. Για να γίνει αυτό, πρέπει να οριστεί το μέγεθος μέσω των scalefactors του δεξιού καναλιού και μια μεταβλητή που λέγεται `is_pos(sb)`. Αυτή η μεταβλητή μεταδίδεται στη θέση των scalefactors για το δεξί κανάλι.

3.4.7 Alias Reduction

Στο στάδιο του MDCT κατά την κωδικοποίηση, αναφέρθηκε η εφαρμογή της μείωσης του aliasing. Για να επανακατασκευαστεί σωστά το audio σήμα κατά την αποκωδικοποίηση πρέπει τα ψεγάδια που προκλήθηκαν από το φαινόμενο Alias να προστεθούν ξανά στο σήμα. Η προσθήκη αυτών γίνεται μέσω ορισμένων υπολογισμών που ονομάζονται butterfly calculations και γίνονται οχτώ για κάθε υποζώνη, όπως φαίνεται στο Σχήμα 3.15.



Σχήμα 3.15: Butterflies για τη μείωση alias

Η μείωση alias εφαρμόζεται μόνο στα μικρά παράθυρα. Η διαδικασία πραγματοποιείται με τη συγχώνευση κάποιων συχνοτικών γραμμών μέσω ορισμένων υπολογισμών. Για κάθε υπολογισμό απαιτούνται ορισμένοι συντελεστές. Οι υπολογισμοί γίνονται χρησιμοποιώντας τις σχέσεις (2.21) και (2.22), ενώ οι συντελεστές φαίνονται στον Πίνακα 3.18.

$$cs(i) = \frac{1}{\sqrt{(1+(c(i)))^2}} \quad (2.21)$$

$$ca(i) = \frac{c(i)}{\sqrt{(1+(c(i)))^2}} \quad (2.22)$$

I	0	1	2	3	4	5	6	7
c(i)	0.6	0.535	0.33	0.185	0.095	0.041	0.00142	0.0037

Πίνακας 3.18: Συντελεστές για τη μείωση alias

3.4.8 Inverse Modified Discrete Cosine Transform (IMDCT) and Overlapping

Ο IMDCT σε συνδυασμό με την τράπεζα φίλτρων σύνθεσης παράγει χρονικά δείγματα $x(i)$ από συχνοτικές γραμμές $X(k)$. Το “ n ” είναι ο αριθμός των «παραθυροποιημένων» δειγμάτων. Κατά την «παραθυροποίηση» δημιουργούνται τέσσερεις διαφορετικοί τύποι παραθύρων, οι πληροφορίες για τους οποίους βρίσκονται στο side information του κάθε frame. Για τα short windows $n=12$ και για τα long windows $n=36$. Υπολογίζεται σύμφωνα με τη σχέση 2.23.

Για $i = 0$ έως $n-1$

$$x(i) = \sum_{k=0}^{(n/2)-1} X(k) \cos\left(\frac{\pi}{2n}\left(2i+1+\frac{n}{2}\right)(2k+1)\right) \quad (2.23)$$

Για $n=36$ ο IMDCT δέχεται στην είσοδό του 18 συχνοτικές γραμμές και παράγει 36 δείγματα. Αυτά τα δείγματα πολλαπλασιάζονται με ένα παράθυρο 36 σημείων πριν περάσουν στο επόμενο στάδιο της αποκωδικοποίησης.

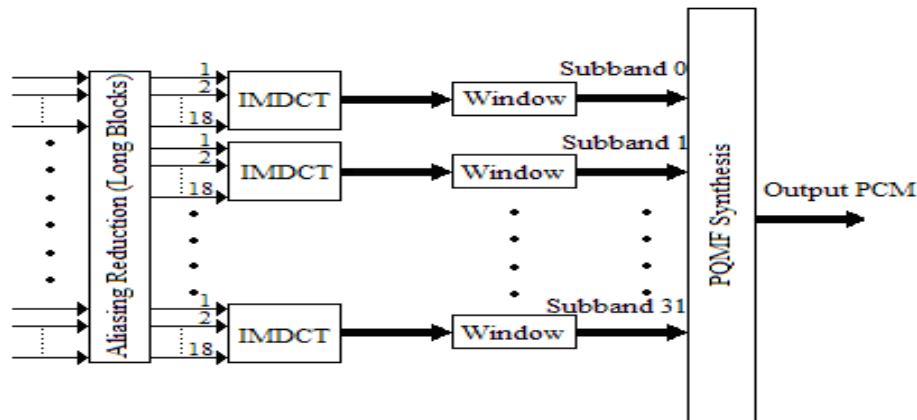
Η παραγωγή 36 δειγμάτων από 18 συχνοτικές γραμμές υποδεικνύει ότι μόνο τα 18 δείγματα είναι μοναδικά. Ο IMDCT, λοιπόν, χρησιμοποιεί επικάλυψη σε ποσοστό 50%. Οι 36 τιμές που προέρχονται από την παραθυροποίηση χωρίζονται σε δύο ομάδες. Η πρώτη ομάδα (18 τιμές) του block που αποκωδικοποιείται επικαλύπτεται με τη δεύτερη ομάδα του αμέσως προηγούμενου block. Η επικάλυψη γίνεται με την πρόσθεση των τιμών της πρώτης ομάδας του δεύτερου block με αντίστοιχες τιμές της δεύτερης ομάδας του πρώτου block.

3.4.9 Frequency Inversion

Όταν η επικάλυψη ολοκληρωθεί, παράγονται 18 δείγματα για κάθε μια από τις 32 πολυφασικές υποζώνες. Προτού γίνει η επεξεργασία μέσω της σύνθεσης της πολυφασικής τράπεζας φίλτρων, κάθε περιττό δείγμα κάθε περιττής υποζώνης πολλαπλασιάζεται με το -1 για να αντισταθμίσει την αναστροφή της συχνότητας που μπορεί να γίνει κατά τη σύνθεση.

3.4.10 Synthesis Polyphase Filterbank

Το τελικό στάδιο της αποκωδικοποίησης περιλαμβάνει το μετασχηματισμό των 18 χρονικών δειγμάτων κάθε μιας από τις 32 υποζώνες σε 18 blocks των 32 PCM δειγμάτων, όπως φαίνεται και στο Σχήμα 3.16.



Σχήμα 3.16: Δομή της τράπεζας σύνθεσης¹¹

3.5 Επίδραση του Αλγορίθμου στην Ποιότητα της Μουσικής

Ο τρόπος λειτουργίας του αλγορίθμου, όπως αναλύθηκε παραπάνω, επηρεάζει αρκετά τα ποιοτικά χαρακτηριστικά του συμπιεζόμενου ήχου. Οι αλλοιώσεις που προκύπτουν είναι αισθητές, κυρίως σε περίπλοκα ηχητικά σήματα, όπως είναι τα μουσικά κομμάτια. Ο αλγόριθμος επιδρά σε αρκετά διαφορετικά μουσικά χαρακτηριστικά, τα οποία μπορούν να χωριστούν σε τρεις κατηγορίες:

- Φασματικά χαρακτηριστικά
- Δυναμικά χαρακτηριστικά
- Θόρυβος και παραμόρφωση

¹¹ Το σχήμα αυτό αποτελεί παραλλαγή αντίστοιχου σχήματος από: Bosi, M & Goldberg R. E. (2003). Introduction to digital audio coding and standards. ISBN: 1-4020-7357-7. 275

3.5.1 Φασματικά Χαρακτηριστικά

Λόγω της εκμετάλλευσης των ψυχοακουστικών φαινομένων, όπως το masking, αλλά και συγκεκριμένων τεχνικών κωδικοποίησης της στερεοφωνίας, όπως αυτή του Intensity Stereo, το πρότυπο MP3 κατά τη συμπίεση του ηχητικού σήματος αλλοιώνει το συχνοτικό φάσμα του αρχικού σήματος. Αυτή η αλλοίωση γίνεται συχνά αντιληπτή από τον ακροατή, με τις αλλοιώσεις να είναι περισσότερο εμφανείς στις ψηλές συχνότητες. Φυσικά, σημαντικό ρόλο παίζει το bit rate με το οποίο γίνεται η συμπίεση. Όσο μικρότερο είναι, τόσο πιο εμφανείς είναι οι αλλοιώσεις. Στις χαμηλές συχνότητες, παρατηρούνται ορισμένες αλλοιώσεις λόγω του διαχωρισμού του ακουστικού φάσματος από τον αλγόριθμο σε ζώνες. Ενώ το ανθρώπινο αυτί διαχωρίζει το ακουστικό φάσμα σε περιοχές μεταβλητού μεγέθους, ανάλογα με τη συχνοτική περιοχή, ο αλγόριθμος του MP3 το διαχωρίζει σε ζώνες σταθερού μεγέθους. Στις χαμηλότερες συχνότητες, από 20 έως 700 Hz, ενώ το ανθρώπινο αυτί έχει 7 διαφορετικές critical bands, το MP3 έχει μόνο μια. Τα ψυχοακουστικά φαινόμενα που λαμβάνουν χώρα σε αυτή την περιοχή, όπως το masking, είναι πολύ διαφορετικά από αυτά που συμβαίνουν στον ανθρώπινο διαχωρισμό του φάσματος.

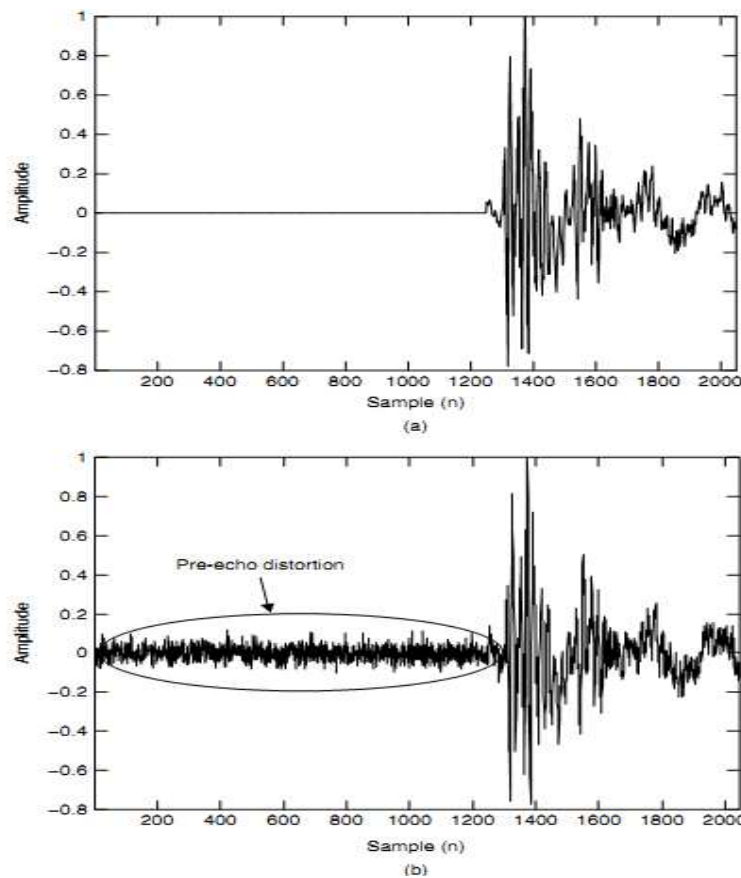
3.5.2 Δυναμικά Χαρακτηριστικά

Στον τομέα της δυναμικής, το πρότυπο MP3, λόγω της συμπίεσης την οποία πραγματοποιεί, μειώνει το δυναμικό εύρος του συμπιεζόμενου ηχητικού σήματος. Αυτό έχει σαν αποτέλεσμα την εμφανή μείωση των αλλαγών της δυναμικής ενός μουσικού κομματιού κάνοντας τες ιδιαίτερα αντιληπτές σε είδη μουσικής στα οποία η αυξομείωση της δυναμικής παίζει πρωτεύοντα ρόλο, όπως η κλασική μουσική.

3.5.3 Θόρυβος και Παραμόρφωση

Κατά τη συμπίεση, το πρότυπο MP3 χρησιμοποιεί ορισμένες διαδικασίες, όπως τα πολυφασικά φίλτρα, οι οποίες εισάγουν θόρυβο στο παραγόμενο αρχείο. Ένα από τα μεγαλύτερα προβλήματα που παρουσιάζονται, είναι το φαινόμενο pre-echo, το οποίο,

όπως φαίνεται και από το όνομά του, είναι η εμφάνιση ηχούς, η οποία προηγείται του σήματος, όπως φαίνεται και στην Εικόνα 5.1.



Εικόνα 5.1: (a) Αρχικό σήμα (b) Σήμα μετά την εισαγωγή του pre-echo¹²

Για την αποφυγή του, το πρότυπο MP3 χρησιμοποιεί μετασχηματισμούς Fourier οι οποίοι δημιουργούν short windows. Παρ' όλα αυτά, σε ορισμένους αλγορίθμους, το pre-echo δεν αντιμετωπίζεται αποτελεσματικά, με αποτέλεσμα την εισαγωγή θορύβου στο σήμα.

Εκτός του φαινομένου pre-echo, αρκετές φορές εισάγονται και άλλοι τύποι παραμόρφωσης, όπως clicks και pops, ενώ σε ορισμένες περιπτώσεις αλγορίθμων, ο ακροατής μπορεί να αντιληφθεί ακόμα και μικρά ηχητικά κενά κατά τη διάρκεια του ηχητικού σήματος.

¹² Η εικόνα αυτή προέρχεται από: Spanias, A. & Painter, T. & Atti, V. (2007). Audio Signal Processing and Coding. John Wiley & Sons, Inc.. 183.

ΚΕΦΑΛΑΙΟ 4^ο: ΜΕΘΟΔΟΣ SEMANTIC DIFFERENTIAL

Εισαγωγή 4^ο Κεφαλαίου

Στο κεφάλαιο αυτό γίνεται η ανάλυση της μεθόδου έρευνας Semantic Differential η οποία χρησιμοποιείται στο πρακτικό μέρος για τη βαθμολόγηση των ΜΡ3. Αναλύεται η λογική που διέπει τη μέθοδο, η δομή της, τα κριτήρια επιλογής της και η αξιολόγηση της καθώς επίσης και ορισμένα προβλήματα που παρουσιάζονται κατά τη χρήση της.

Σκοπός του κεφαλαίου αυτού είναι η βαθιά και πλήρης κατανόηση της μεθοδολογίας της Semantic Differential με σκοπό την ορθή και έγκυρη χρήση της κατά το πρακτικό μέρος.

Η έρευνα για θέματα που άπτονται της υποκειμενικότητας και της κρίσης του κάθε ανθρώπου είναι μια δύσκολη διαδικασία. Υπάρχουν πολλοί παράγοντες οι οποίοι μπορούν να επηρεάσουν και να αλλοιώσουν τα αποτελέσματα, καθιστώντας τα άκυρα και μη ρεαλιστικά. Για να θεωρείται μια μεθοδολογία έρευνας σωστή και έγκυρη, είτε πρόκειται για μαθησιακή μέθοδο, είτε για ψυχολογική, είτε για μέθοδο αντίληψης πρέπει να πληροί ορισμένα αντικειμενικά κριτήρια, τα βασικότερα εκ των οποίων είναι τα παρακάτω:

- **Αντικειμενικότητα:** Από τη μέθοδο θα πρέπει να εξάγονται αποτελέσματα τα οποία να είναι επαληθεύσιμα και ανεξάρτητα από την ιδιοσυγκρασία και την κρίση του ερευνητή.
- **Αξιοπιστία:** Σε περίπτωση επανάληψης ακριβώς της ίδιας διαδικασίας κάτω από τις ίδιες συνθήκες θα πρέπει να εξάγονται τα ίδια συμπεράσματα, με μικρά περιθώρια αποκλίσεων.
- **Εγκυρότητα:** Τα δεδομένα που θα προκύψουν πρέπει να είναι παραπλήσια με τα δεδομένα που θα εξάγονταν για το ίδιο θέμα αλλά με τη χρήση διαφορετικής μεθοδολογίας.

- **Εναισθησία:** Η μέθοδος θα πρέπει να αποδίδει σωστά τις διαφοροποιήσεις ανάλογα με το θέμα το οποίο μελετάται.
- **Συγκρισιμότητα:** Η μέθοδος πρέπει να είναι εφαρμόσιμη σε μεγάλο εύρος θεμάτων και σε πολλά θεματικά πεδία και να δίνεται η δυνατότητα σύγκρισης των αποτελεσμάτων ανάμεσα σε διαφορετικές ομάδες ατόμων και σε μεμονωμένα άτομα.
- **Χρηστικότητα:** Η μέθοδος πρέπει να αποδίδει πληροφορίες σχετικές με τα θεωρητικά ή πρακτικά ζητήματα τα οποία εξετάζει με τρόπο αποδοτικό. Για παράδειγμα, δεν θα πρέπει η μέθοδος να είναι κατασκευασμένη με τέτοιο τρόπο ώστε να κουράζει τα υποκείμενα ούτε και να καθοδηγεί με τον οποιοδήποτε τρόπο τη σκέψη τους.

Η τήρηση αυτών των προϋποθέσεων είναι απαραίτητη για τη σωστή ερευνητική μεθοδολογία και για την ορθή εξαγωγή των αποτελεσμάτων. Η επιλογή της μεθόδου πρέπει πάντα να ακολουθεί αυτά τα κριτήρια για να θεωρηθεί έγκυρη.

4.1 Σύντομο Ιστορικό

Η μέθοδος Semantic Differential είναι ένα όργανο μέτρησης των αντιδράσεων των ανθρώπων με τη χρήση αριθμημένων κλιμάκων οι οποίες στα δύο άκρα τους έχουν δύο αντίθετες έννοιες (διπολικές κλίμακες).

Το 1957 οι Osgood, Suci και Tannenbaum στα πλαίσια της προσπάθειας τους να βρουν έναν αντικειμενικό τρόπο μέτρησης αντιδράσεων και συμπεριφορών συγκέντρωσαν όλες τις μελέτες που είχαν γίνει νωρίτερα και συνέγραψαν το πρώτο βιβλίο για τη μέθοδο αυτή.

Αργότερα, έγιναν πολλές έρευνες γύρω από τη μέθοδο αυτή και εξετάστηκαν λεπτομερώς όλα τα στοιχεία της. Ενδεικτικά, οι Green και Goldfried το 1965¹³ εξέτασαν

¹³ Green, R.F., & Goldfried, M. R. (1965). On the bipolarity of semantic space. *Psychological Monographs* 79 (6, Whole No.601).

το ενδεχόμενο χρήσης μονοπολικών κλίμακων, ο Messick το 1957¹⁴ αποκάλυψε ότι το κέντρο της κλίμακας δεν είναι ακριβώς το μηδέν, ενώ τα μετρικά χαρακτηριστικά των ποσοτικών επιρρηματικών προσδιορισμών που χρησιμοποιούνται στις κλίμακες μελετήθηκαν από πολλούς όπως ο Cliff (1959)¹⁵ και ο Howe (1962¹⁶, 1966¹⁷). Ένας από τους σημαντικότερους ερευνητές που συνέβαλε στην εξέλιξη της μεθόδου είναι ο David R. Heise, ο οποίος μάλιστα το 1967 δημοσίευσε έναν οδηγό για τα χαρακτηριστικά 1000 αγγλικών λέξεων που χρησιμοποιούνται πολύ στη μέθοδο¹⁸.

Η μέθοδος Semantic Differential είναι πολύ διαδεδομένη και χρησιμοποιείται σήμερα σε πολλά πεδία, από την ψυχολογία μέχρι και τις έρευνες για καταναλωτικά αγαθά. Μια τέτοια μέθοδος, λοιπόν, με τόσο ευρύ φάσμα πεδίων και τόσο καλά προσαρμόσιμη στις ανάγκες κάθε ερευνητή, είναι μια σωστή και λογική επιλογή για κάθε είδους μελέτη.

4.2 Ανάλυση της Μεθόδου

Η ανάλυση της μεθόδου Semantic Differential θα γίνει με βάση τη λογική την οποία ακολουθεί, τη δομή που πρέπει να έχουν τα μετρητικά όργανα που χρησιμοποιούνται (ερωτηματολόγια), την αξιολόγησή της καθώς και ορισμένα προβλήματα που μπορεί κανείς να συναντήσει κατά την εφαρμογή της.

4.2.1 Η Λογική της Μεθόδου

Η γενική λογική της σημασιολογικής διαφοράς (semantic differential), στηρίζεται στη λογική του σημασιολογικού διαστήματος. Κάθε κλίμακα αυτής της λογικής ορίζεται από ένα ζεύγος πολικών (αντίθετων) επιθετικών προσδιορισμών X-Y και υποτίθεται πως

¹⁴ Messick, S. J. (1957). Metric properties of the semantic differential. *Educational and Psychological Measurement*, 17, 200-206.

¹⁵ Cliff, N. (1959). Adverbs as multipliers. *Psychological Review*, 66, 27-44.

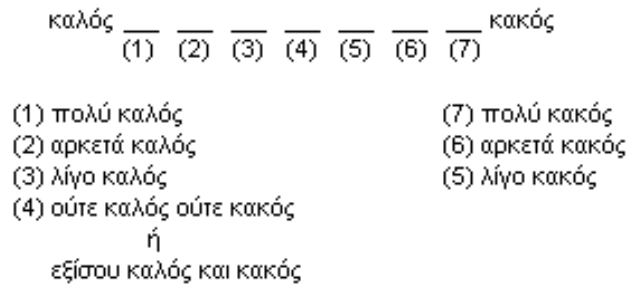
¹⁶ Howe, E. S. (1962). Probabilistic adverbial qualification of adjectives. *Journal of Verbal Learning and Verbal Behavior*, 225-242.

¹⁷ Howe, E. S. (1966). Associative structure of quantifiers. *Journal of Verbal Learning and Verbal Behavior*, 5, 156-162.

¹⁸ Heise, D. R. (1959). Semantic differential profiles for 1,000 most frequent English words. *Psychological Monographs*, 79 (8, Whole No. 601).

αναπαριστά μια ευθεία γραμμή πάνω στην οποία τοποθετούνται σημεία ίσης απόστασης μεταξύ τους.

Τα σημεία αυτά, χωρίζουν την κλίμακα σε ισόποσα τμήματα πάνω στα οποία τοποθετούνται οι απαντήσεις των υποκειμένων που συμμετέχουν σε τέτοιου είδους έρευνες. Σύμφωνα με τους ανθρώπους που κατασκεύασαν και μελέτησαν τη μέθοδο αυτή, τα τμήματα αυτά μπορούν να χαρακτηριστούν με επιρρηματικούς προσδιορισμούς, ανάλογα με την απόσταση που απέχουν από τα δύο άκρα της κλίμακας. Οι προσδιορισμοί αυτοί είναι συγκεκριμένοι: «πολύ», «αρκετά» και «λίγο» και τοποθετούνται γύρω από το ουδέτερο σημείο, όπως φαίνεται στο Σχήμα 4.1.



Σχήμα 4.1: Παράδειγμα διπολικής κλίμακας

4.2.2 Δομή EPA

Ένα από τα χαρακτηριστικότερα στοιχεία της μεθόδου είναι η κατηγοριοποίηση των απαντήσεων σε τρεις βασικές κατηγορίες. Μια σειρά μελετών έχει πραγματοποιηθεί για τον καθορισμό των διαστάσεων των διπολικών κλιμάκων. Η πιο σημαντική ήταν αυτή του Roget (Thesaurus), στην οποία επελέγησαν 76 ζεύγη επιθέτων για τον προσδιορισμό 20 διαφορετικών αντικειμένων από 100 φοιτητές. Οι συσχετίσεις ανάμεσα στα αποτελέσματα των διαφορετικών κλιμάκων υπολογίστηκαν και κατηγοριοποιήθηκαν. Η δομή EPA ήταν εμφανής σε αυτή την ανάλυση καθώς και σε προγενέστερες μελέτες. Αν και σε ορισμένες περιπτώσεις προέκυψαν περισσότερες διαστάσεις, εν τέλει οι ερευνητές κατέληξαν στο συμπέρασμα ότι η δομή αυτή είναι η βασική και η ουσιώδης και πως οι υπόλοιπες διαστάσεις είναι, λίγο-πολύ, υποκατηγορίες των βασικών.

Η δομή EPA αποτελείται από τρεις βασικές κατηγορίες χαρακτηρισμών. Οι κατηγορίες αυτές ονομάζονται factors και είναι οι evaluation, potency και activity (EPA). Οι

περισσότεροι επιθετικοί προσδιορισμοί που χρησιμοποιούνται στη μέθοδο Semantic Differential κατατάσσονται σε μια από τις κατηγορίες αυτές. Δυστυχώς, δεν υπάρχουν σαφή ορισμοί μέσω των οποίων μπορεί κανείς να κατατάξει κάποιο επίθετο σε μια από τις κατηγορίες. Παρ' όλα αυτά, λόγω των ονομάτων των κατηγοριών αυτών, τα επίθετα κατατάσσονται σχετικά εύκολα με βάση τη σχετικότητά τους με μια από τις τρεις κατηγορίες. Επί παραδείγματι, το ζεύγος «αργός-γρήγορος» θα μπορούσε να βρίσκεται μόνο στην κατηγορία activity, μιας και το νόημα του προδίδει κίνηση και δραστηριότητα.

Για να γίνει καλύτερα κατανοητή η κατηγοριοποίηση που γίνεται στα επίθετα, θα μπορούσε κανείς να πει γενικά ότι:

- Τα επίθετα που ανήκουν στην κατηγορία evaluation είναι αυτά μέσω των οποίων τα υποκείμενα αξιολογούν γενικά θετικά ή αρνητικά κάποιο αντικείμενο. Μερικά χαρακτηριστικά ζεύγη είναι τα «καλός-κακός», «ωραίος-άσχημος», «γλυκός-πικρός» κλπ.
- Τα επίθετα που ανήκουν στην κατηγορία potency σχετίζονται συνήθως με μεγέθη, δυνατότητες και το μέτρο δύναμης ενός αντικειμένου. Χαρακτηριστικά παραδείγματα είναι τα «μεγάλος-μικρός», «δυνατός-αδύναμος», «βαθύς-ρηχός» κλπ.
- Τα επίθετα που ανήκουν στην κατηγορία activity έχουν να κάνουν κυρίως με την κινητικότητα ενός αντικειμένου. Τα πιο χαρακτηριστικά είναι τα «αργός-γρήγορος», «ζωντανός-νεκρός», «θορυβώδης-ήσυχος» κλπ.

4.2.3 Επιλογή Αντικειμένων

Στην παρούσα εργασία, υπάρχει το αντικείμενο έρευνας είναι η ποιότητα ορισμένων ακουσμάτων που ανήκουν σε διαφορετικά είδη μουσικής με την εφαρμογή του αλγορίθμου συμπίεσης του MP3.

Η επιλογή των κομματιών έγινε με τρόπο που να διασφαλίζει τη σχέση τους με το κάθε είδος μουσικής (κάθε κομμάτι να περιέχει όλα τα χαρακτηριστικά που χρειάζεται για να

κατηγοριοποιηθεί σε συγκεκριμένο είδος) αλλά και την οικειότητα των υποκειμένων με αυτά (ιδιαίτερα δημοφιλή κομμάτια).

4.2.4 Επιλογή Κλιμάκων

Οι ιδανικές κλίμακες θα ήταν αυτές οι οποίες θα ανήκαν ξεκάθαρα και χωρίς καμία αμφιβολία σε έναν από τους συντελεστές που περιγράφηκαν παραπάνω. Εάν υπήρχαν, θα μπορούσε να χρησιμοποιηθεί μια κλίμακα από κάθε κατηγορία και τα αποτελέσματα θα ήταν έγκυρα. Πρακτικά, όμως, λίγα είναι αυτά τα ζεύγη επιθέτων τα οποία μπορούν να τοποθετηθούν με απόλυτη σιγουριά σε μία μόνο κατηγορία. Γι' αυτό το λόγο, πρέπει να χρησιμοποιηθούν αρκετές κλίμακες για την αναπαράσταση κάθε κατηγορίας. Τα κριτήρια για την επιλογή των κλιμάκων αυτών είναι τέσσερα:

1. Η ισοκατανομή των συντελεστών. Για την επίτευξή της, συνήθως χρησιμοποιείται ο ίδιος αριθμός κλιμάκων για κάθε συντελεστή. Για παράδειγμα, θα μπορούσαν να χρησιμοποιηθούν τρεις κλίμακες του παράγοντα evaluation, τρεις του potency και τρεις του activity. Αυτό όμως δεν είναι απόλυτο και ούτε συμβαίνει πάντα μιας και δεν έχουν όλες οι κλίμακες την ίδια βαρύτητα. Ορισμένες κλίμακες είναι πολύ πιο στενά συνδεδεμένες με έναν παράγοντα απ' ότι κάποιες άλλες καθιστώντας τις πιο δυνατές. Επί παραδείγματι, μια πολύ αντιπροσωπευτική κλίμακα ενός παράγοντα θα χρειαζόταν δύο κλίμακες από τον καθένα από τους άλλους δύο παράγοντες προκειμένου να υπάρχει ισόποση κατανομή.
2. Οι κλίμακες που επιλέγονται πρέπει να έχουν άμεση σχέση με το προς βαθμολόγηση αντικείμενο. Όταν χρησιμοποιούνται κλίμακες οι οποίες δεν έχουν μεγάλη σχέση με τα αντικείμενα, παρατηρείται μεγάλη συγκέντρωση απαντήσεων γύρω από το ουδέτερο τμήμα της κλίμακας. Αυτό έχει σαν αποτέλεσμα τη μεγάλη μείωση των πληροφοριών ακόμα και με τη χρήση μεγάλου αριθμού κλιμάκων.
3. Ακόμα ένα κριτήριο για την επιλογή είναι η σημασιολογική σταθερότητα που έχει η κάθε κλίμακα. Η σημασιολογική σταθερότητα έχει να κάνει με το κατά

πόσον μια κλίμακα έχει την ίδια σημασία για τα υποκείμενα όταν χρησιμοποιείται για διαφορετικά αντικείμενα. Για παράδειγμα, το ζεύγος «μεγάλος-μικρός» έχει διαφορετική σημασία εάν χρησιμοποιηθεί για να χαρακτηρίσει τη λέξη «έντομο» και διαφορετική για να χαρακτηρίσει τη λέξη «πονοκέφαλος». Για να χρησιμοποιηθεί μια κλίμακα, η σημασία της θα πρέπει να γίνεται απόλυτα κατανοητή από τα υποκείμενα χωρίς περιθώρια παρερμηνείας.

4. Το τελευταίο κριτήριο είναι η γραμμικότητα των κλιμάκων. Οι κλίμακες οι οποίες χρησιμοποιούνται πρέπει να είναι γραμμικές και τα επίθετα τα οποία χρησιμοποιούνται πρέπει να είναι απόλυτα αντίθετα μεταξύ τους.

Πρέπει εδώ να σημειωθεί ότι η επιλογή των κλιμάκων δεν εξαντλείται μόνο σε αυτά τα τέσσερα κριτήρια. Υπάρχουν αρκετές έρευνες οι οποίες έχουν γίνει αποδεκτές και δεν πληρούν ορισμένες από αυτές τις προϋποθέσεις. Αυτό συμβαίνει διότι το πεδίο στο οποίο κινείται η μέθοδος έχει μεγάλη δόση υποκειμενικότητας. Ο γενικός κανόνας λέει ότι οι βασικές κατηγορίες συντελεστών είναι τρεις και τα κριτήρια επιλογής των κλιμάκων είναι τέσσερα. Όμως οι περιπτώσεις στις οποίες αυτά δεν τηρούνται δε συνεπάγονται ότι οι έρευνες δεν είναι έγκυρες.

4.2.5 Χρήση Εφτά Σημείων

Μέσω ενός μεγάλου αριθμού πειραμάτων και ερευνών έχει προκύψει ότι ο χωρισμός της κλίμακας σε επτά τμήματα είναι ο πιο εύχρηστος από όλους. Όλα τα σημεία έχουν την ίδια πιθανότητα να επιλεγθούν μιας και είναι ξεκάθαρο ότι αντιπροσωπεύουν επτά διακριτές διαβαθμίσεις. Σε πειράματα που έγιναν με εννέα διαβαθμίσεις της κλίμακας παρατηρήθηκε το γεγονός ότι τα υποκείμενα απέφευγαν να επιλέξουν ορισμένες από τις διαβαθμίσεις της κλίμακας (κυρίως αυτές που βρίσκονταν γύρω από το ουδέτερο σημείο). Επίσης, σε κλίμακες που είχαν μόνο πέντε διαβαθμίσεις, παρατηρήθηκε ο εκνευρισμός των υποκειμένων για το γεγονός ότι μπορούσαν να επιλέξουν μόνο ανάμεσα στα επιρρήματα «πολύ», «λίγο» και την ουδέτερη θέση (απουσίαζε ο χαρακτηρισμός «αρκετά»).

4.2.6 Αξιολόγηση της Μεθόδου

Στην αρχή του κεφαλαίου αναφέρθηκαν τα βασικά κριτήρια σύμφωνα με τα οποία μια μέθοδος θεωρείται έγκυρη. Σε αυτό το σημείο, λοιπόν, θα αξιολογηθεί η μέθοδος Semantic Differential με βάση τα προαναφερθέντα.

4.2.6.1 Αντικειμενικότητα

Μία μέθοδος θεωρείται αντικειμενική όταν οι πράξεις και τα μέσα με τα οποία καταλήγει κανείς στα αποτελέσματα είναι σαφή και επαληθεύσιμα.

Οι διαδικασίες μέτρησης που ακολουθούνται από τη μέθοδο Semantic Differential δίνουν σαφείς και επαληθεύσιμες. Τα μέσα με τα οποία κάνεις εξάγει τα αποτελέσματα, από τα αριθμημένα κενά της κλίμακας μέχρι και τον τρόπο με τον οποίο επιλέγεται η δομή των ερωτηματολογίων, είναι αντικειμενικά. Εάν τα ίδια ερωτηματολόγια δοθούν σε δύο διαφορετικούς ερευνητές και αυτοί ακολουθήσουν τους ίδιους κανόνες για την εξαγωγή των αποτελεσμάτων, τότε θα καταλήξουν στα ίδια συμπεράσματα. Θα μπορούσε κάποιος να πει ότι τα δεδομένα τα οποία συλλέγονται μέσω τέτοιων μεθόδων είναι υποκειμενικά και ότι το μόνο που καταφέρνει κανείς είναι να εκφράσει αντικειμενικά υποκειμενικές καταστάσεις. Αυτό είναι αλήθεια, αλλά δεν είναι απαραίτητα κακό μιας και η αντικειμενικότητα της μεθόδου αφορά τον ερευνητή και όχι τους ανθρώπους που συμμετέχουν στην έρευνα. Οι διαδικασίες που ακολουθούνται δεν αφήνουν περιθώρια αλλοίωσης των αποτελεσμάτων λόγω της υποκειμενικότητας του ερευνητή και αυτή είναι η ουσία της αντικειμενικότητας.

4.2.6.2 Αξιοπιστία

Η αξιοπιστία ενός οργάνου μέτρησης είναι ο βαθμός στον οποίο μπορούν να αναπαραχθούν τα ίδια αποτελέσματα όταν επαναλαμβάνεται η διαδικασία μέτρησης στα ίδια υποκείμενα. Το βασικό σκορ που λαμβάνεται από τη μέθοδο είναι ένα ψηφίο (από το 1 ως το 7, ή από το -3 ως το +3) το οποίο αντιπροσωπεύει την επιλογή του υποκειμένου σε μια συγκεκριμένη κλίμακα ενός συγκεκριμένου θέματος. Η αναπαραγωγή των βασικών αυτών σκορ ονομάζεται αξιοπιστία αντικειμένου (item reliability).

Τα σκορ αυτά συνήθως αντιπροσωπεύουν το μέσο όρο των factors. Για παράδειγμα, τις πέντε απαντήσεις για τον evaluation, τις πέντε του rotency και τις πέντε του activity. Η αναπαραγωγή των τιμών αυτών ονομάζεται αξιοπιστία του σκορ των συντελεστών (factor-score reliability).

- **Item Reliability**

Η συμβατική έννοια της αξιοπιστίας σε ψυχολογικές και εκπαιδευτικές έρευνες εστιάζει στο πόσο συνεπή είναι τα άτομα που συμμετέχουν σε διαδοχικές εφαρμογές του οργάνου μέτρησης. Ένας πιο γενικός ορισμός της αξιοπιστίας, όπως αυτός προκύπτει από τις επιστήμες, είναι η προαναφερθείσα αναπαραγωγή των σκορ κατόπιν επαναλαμβανόμενων μετρήσεων.

Απόλυτη και ιδανική αξιοπιστία υπάρχει μόνο όταν, πραγματοποιώντας δύο ίδιες μετρήσεις, βρεθούν απόλυτα ίδια αποτελέσματα. Αυτό σημαίνει ότι η οποιαδήποτε απόκλιση των αποτελεσμάτων μεταξύ τους εισάγει ένα βαθμό αναξιοπιστίας. Επειδή όμως στην πράξη δεν είναι δυνατόν να αναπαραχθούν ακριβώς τα ίδια αποτελέσματα, εισάγεται ο βαθμός της αξιοπιστίας, σύμφωνα με τον οποίο κρίνεται η μέθοδος. Έτσι λοιπόν, η μέτρηση θεωρείται αξιόπιστη όταν κατά την επανάληψη των μετρήσεων διαπιστώνονται διαφορές εντός ορισμένων ορίων.

Μια σειρά μελετών έχει γίνει γύρω από την αξιοπιστία, προκειμένου να ικανοποιηθούν αυτές οι προδιαγραφές. Οι μελέτες αυτές εστίασαν στο πρόβλημα με διαφορετικούς τρόπους κάθε φορά. Τρεις από τις σημαντικότερες αναφέρονται παρακάτω:

1. Κοινή κατανομή των σκορ της πρώτης μέτρησης και της επανάληψή της

Με το συγκεκριμένο τρόπο προσπάθησαν να προσεγγίσουν το πρόβλημα ο Solomon¹⁹ και ο Kellog Wilson. Σύμφωνα με αυτούς, οι απαντήσεις που δίνονται για το ίδιο αντικείμενο τις δύο φορές που επαναλαμβάνεται η μέτρηση μπορούν να σχεδιαστούν σε έναν πίνακα. Οι γραμμές του πίνακα ορίζονται από τις επτά εναλλακτικές απαντήσεις της πρώτης μέτρησης και οι στήλες από τις επτά εναλλακτικές της

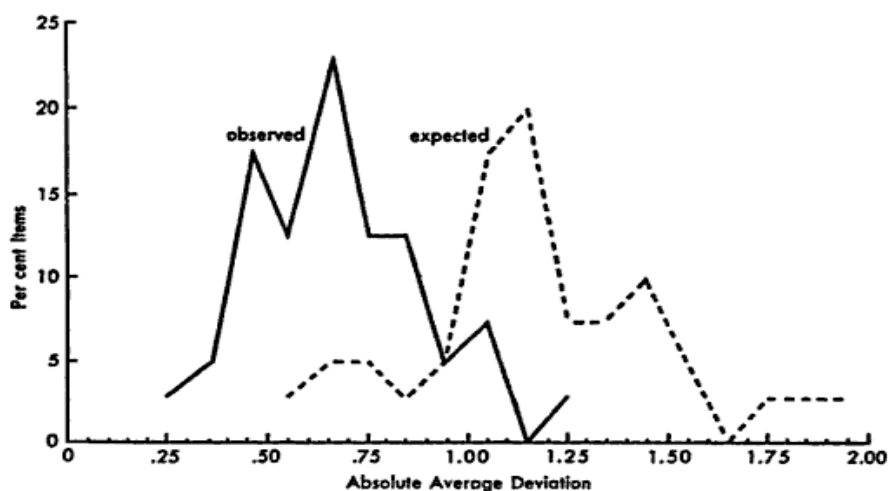
¹⁹ Solomon, L. N. (1954). A factorial study of complex auditory stimuli (passive sonar sounds). Unpublished doctor's dissertation, University of Illinois, 66-68, 128.

δεύτερης. Ένα τέτοιο σχήμα δείχνει τις αποκλίσεις ανάμεσα στις κλίμακες των δύο τεστ. Εάν υπάρχει «απόλυτη αξιοπιστία» όλα τα σημεία των υποκειμένων θα πέσουν στην κύρια διαγώνιο του πίνακα. Σύμφωνα λοιπόν με τη μέθοδο αυτή, όσο μεγαλώνει η διασπορά των σημείων γύρω από τη διαγώνιο, τόσο μικρότερη γίνεται η αξιοπιστία.

Η μελέτη αυτή χαρακτηρίστηκε ως γενική και κατά κάποιο τρόπο αβάσιμη μιας και προκύπτουν νέοι παράγοντες οι οποίοι έπρεπε να εξεταστούν για να βγάλει κανείς σωστά συμπεράσματα. Έτσι λοιπόν, κρίθηκε αναγκαία η ανάπτυξη μιας διαδικασίας μέτρησης των σφαλμάτων, η οποία αναφέρεται παρακάτω.

2. Μέτρηση Σφαλμάτων

Όταν μια μέτρηση επαναλαμβάνεται, είναι αναμενόμενο να υπάρχουν κάποιες διαφορές ανάμεσα στις μετρήσεις. Όσο μικρότερες είναι οι αποκλίσεις των τιμών, τόσο πιο αξιόπιστο θεωρείται το όργανο μέτρησης. Στην Εικόνα 4.1, φαίνεται ο μέσος όρος σφαλμάτων μιας έρευνας. Σε αυτή την εικόνα φαίνονται οι αναμενόμενες διαφορές στα αποτελέσματα της δεύτερης επανάληψης και οι διαφορές οι οποίες τελικά παρατηρήθηκαν. Η έρευνα αυτή αφορούσε την αξιολόγηση σαράντα διαφορετικών αντικειμένων και πραγματοποιήθηκε από τον Kellogg Wilson.



Εικόνα 4.1: Διαφορά αποτελεσμάτων επί τοις εκατό ανάμεσα σε δύο μετρήσεις²⁰

²⁰ Η εικόνα αυτή προέρχεται από: Osgood, C. E. & Suci, G. J. & Tannenbaum, P. H. (1957). The measurement of meaning. University of Illinois. 129.

Ο μέσος όρος των σφαλμάτων μέτρησης είναι πάντα μικρότερος από μία μονάδα (μονάδα της κλίμακας) και συγκεκριμένα για τις κλίμακες που σχετίζονται με τον παράγοντα evaluation είναι περίπου μισή μονάδα. Αυτό σημαίνει ότι οι συμμετέχοντες είναι ακριβείς και σταθεροί στις απαντήσεις τους. Η απόκλιση της μίας μονάδας η οποία παρατηρείται είναι αρκετά ικανοποιητική μιας και όπως αναφέρθηκε πριν, σε πρακτικό επίπεδο, οι απαντήσεις δε γίνεται να ταυτίζονται. Ακόμα όμως και αυτός ο τρόπος δεν παρέχει κάποιο όριο μέσα στο οποίο μπορεί κανείς να είναι σίγουρος για την αξιοπιστία της μεθόδου του.

3. Όρια Πιθανοτήτων

Η τελευταία αυτή προσπάθεια αξιολόγησης είναι και αυτή που χρησιμοποιείται ως σήμερα μιας και παρέχει στον ερευνητή σαφή όρια λαθών με τον καλύτερο δυνατό τρόπο.

Μέσω των μελετών κυρίως της Bopp²¹, προέκυψε το συμπέρασμα ότι για οποιοδήποτε τύπο κλίμακας (είτε ανήκει στο evaluation, είτε στο potency είτε στο activity), η διαφορά των απαντήσεων ανάμεσα στην πρώτη και τη δεύτερη μέτρηση είναι ικανοποιητική όταν δεν ξεπερνά τις δύο μονάδες κλίμακας. Εφαρμόζοντας λοιπόν τη μέθοδο Semantic Differential παρατήρησαν ότι δεν παρατηρούνται αποκλίσεις από αυτόν τον κανόνα (με εξαίρεση ορισμένες περιπτώσεις μελετών που έγιναν σε νοητικά ασταθείς ανθρώπους). Συνεπώς, η αξιοπιστία της μεθόδου κρίθηκε πολύ καλή, πράγμα που καθιστά τη μέθοδο αυτή έγκυρο όργανο μέτρησης.

Να σημειωθεί ότι οι μελέτες που έγιναν αφορούσαν εκατοντάδες διαφορετικά αντικείμενα από ονόματα διασημοτήτων μέχρι τον καιρό. Συνεπώς, καλύφθηκε ένα πολύ ευρύ φάσμα θεμάτων στα οποία η μέθοδος ανταποκρίθηκε πολύ καλά.

• Factor-Score Reliability

Ο παράγοντας αυτός έχει να κάνει με την αξιολόγηση των factors που χρησιμοποιούνται σε κατηγορίες. Έτσι, εδώ, δεν εξετάζεται πλέον η αξιολόγηση κάθε κλίμακας ξεχωριστά

²¹ Bopp, Joan. (1955). A quantitative semantic analysis of word association in schizophrenia. Unpublished doctor's dissertation, University of Illinois, 131, 170, 223-26, 227.

αλλά ολόκληρη η κατηγορία στην οποία υπάγονται κάποιες κλίμακες. Επειδή η μέθοδος Semantic Differential είναι χτισμένη με βάση τους παράγοντες αυτούς, η αξιολόγησή τους θεωρείται πολύ σημαντική διαδικασία.

Όταν κατηγοριοποιούνται κάποιες κλίμακες, είναι λογικό ορισμένα τυχαία σφάλματα να εξαλείφονται, σε σχέση με τη μελέτη κάθε κλίμακας ξεχωριστά. Η λογική αυτή ισχύει και στη συγκεκριμένη περίπτωση. Μέσω ορισμένων μελετών έχει διαπιστωθεί ότι τα περιθώρια αποκλίσεων είναι μικρότερα σε σχέση με το item reliability και μάλιστα ότι για κάθε παράγοντα ισχύουν διαφορετικά όρια. Για τον παράγοντα evaluation η απόκλιση είναι αποδεκτή μέχρι και μια μονάδα, για τον παράγοντα potency μέχρι και μιάμιση μονάδα ενώ για τον παράγοντα activity μέχρι και 1.33.

Και σε αυτόν τον τομέα η μέθοδος Semantic Differential ανταποκρίθηκε καλά μιας και στις μελέτες που έγιναν για την αξιοπιστία της δεν ξεπεράστηκαν τα όρια που αναφέρονται παραπάνω.

4.2.6.3 Εγκυρότητα

Ένα όργανο μέτρησης θεωρείται έγκυρο όταν μετρά αυτό το οποίο πρέπει να μετρηθεί. Με άλλα λόγια, η εγκυρότητα ενός οργάνου εξαρτάται από τη συσχέτιση ανάμεσα στις απαντήσεις που εξάγονται από αυτό και τις απαντήσεις τις οποίες θα περίμενε κανείς να δοθούν για το συγκεκριμένο αντικείμενο έρευνας.

Στην περίπτωση της Semantic Differential, εισάγεται η έννοια της face validity, μιας και το όργανο αυτό είναι σχεδιασμένο να μετρά υποκειμενικά πράγματα για τα οποία κανείς δεν μπορεί να περιμένει να δοθούν απόλυτα σαφείς απαντήσεις.

Όταν ένα όργανο θεωρείται ότι διαθέτει face validity, τότε αυτό σημαίνει ότι οι παρατηρήσεις που γίνονται μέσω αυτού θα μπορούσαν να γίνουν από τους περισσότερους παρατηρητές, χωρίς τη χρήση του οργάνου. Δηλαδή, οι παρατηρήσεις θα μπορούσαν να εξαχθούν και με κοινή λογική, από την πλειοψηφία των ανθρώπων.

Μέσω των μελετών που κατά καιρούς έχουν γίνει γύρω από την εγκυρότητα της Semantic Differential, δεν έχουν βρεθεί λόγοι για την αμφισβήτηση της εγκυρότητάς της.

Καλό θα ήταν, όμως, να αναφερθούν τα σημεία στα οποία πρέπει κανείς να εστιάσει για να ελέγξει την εγκυρότητα μιας μεθόδου.

- **Εγκυρότητα των factors**

Όταν οι συσχετίσεις ανάμεσα σε αρκετές κλίμακες αναλύονται, τότε προκύπτουν ορισμένα βασικά σκορ για κάθε factor (evaluation, potency και activity). Αυτά τα σκορ, όταν προκύπτουν με επαναλαμβανόμενες μετρήσεις, μπορεί κανείς να υποθέσει ότι αντιπροσωπεύουν τις κύριες διαστάσεις με τις οποίες οι άνθρωποι κρίνουν τα πράγματα αυθόρμητα και φυσικά. Η υπόθεση αυτή έχει επαληθευθεί από μελέτες (Rowan, 1954²²), σύμφωνα με τις οποίες έχει διαπιστωθεί ότι τα υποκείμενα κρίνουν με τον ίδιο τρόπο τα αντικείμενα, είτε χρησιμοποιείται κάποιο όργανο μέτρησης είτε όχι.

- **Εγκυρότητα ορισμένων υποθέσεων των κλιμάκων**

Η χρήση της μεθόδου περιλαμβάνει ορισμένες υποθέσεις για τις μεμονωμένες κλίμακες από τις οποίες αποτελείται. Όταν, για παράδειγμα, ένας ακέραιος αριθμός χρησιμοποιείται σε μια θέση της κλίμακας, τότε προκύπτει η υπόθεση ότι τα διαστήματα ανάμεσα στις θέσεις της κλίμακας είναι ίσα. Επίσης, σε διάφορες τεχνικές ανάλυσης των αποτελεσμάτων περιλαμβάνονται υποθέσεις οι οποίες σχετίζονται με τη γραμμικότητα των κλιμάκων, τη θέση «μηδέν» και το κέντρο της κλίμακας. Εάν οι υποθέσεις αυτές δεν επαληθεύονται για κάθε χρησιμοποιούμενη κλίμακα, τότε τα αποτελέσματα που θα εξαχθούν θα είναι αναληθή και διαστρεβλωμένα.

Μέσω των ψυχομετρικών μεθόδων που εφάρμοσαν οι Gulliksen (1954)²³ και Saffir (1937)²⁴, επαληθεύτηκαν οι υποθέσεις αυτές. Υπήρξαν, όμως, ορισμένες

²² Rowan, T. C. (1954). Some developments in multidimensional scaling applied to semantic relationships. Unpublished doctor's dissertation, University of Illinois. 143-46.

²³ Gulliksen, H. (1954). A least squares solution for successive intervals assuming unequal standard deviations. *Psychometrika*, 19, 117-139. 146.

²⁴ Saffir, M. (1937). A comparative study of scales constructed by three psychophysical methods. *Psychometrika*, 2, 179-198. 146.

κλίμακες οι οποίες δεν ανταποκρίθηκαν καλά στα τεστ αυτά. Έτσι, θεωρούνται σχετικά ακατάλληλες για χρήση στη μέθοδο αυτή και καλό είναι να αποφεύγονται.

4.2.6.4 Ευαισθησία

Η ευαισθησία ενός οργάνου σχετίζεται άμεσα με την εγκυρότητα και την αξιοπιστία. Για να διαπιστώσει κανείς εάν το μετρητικό του όργανο έχει ευαισθησία θα πρέπει να επιλέξει σει παραπλήσιων, αλλά σχετικά διαφορετικών, λέξεων και μέσω τεστ να φανεί ότι οι διακρίσεις που γίνονται μέσω του οργάνου μπορούν να γίνουν και από απλούς χρήστες της γλώσσας. Επίσης, άλλος τρόπος για να διαπιστωθεί η ευαισθησία ενός οργάνου είναι να γίνει προσπάθεια διαφοροποίησης συνώνυμων λέξεων. Οι μελέτες που έχουν διεξαχθεί έχουν δείξει ότι η Semantic Differential είναι μια μέθοδος κατασκευής μετρητικών οργάνων η οποία παρέχει μεγάλο βαθμό ευαισθησίας, καθιστώντας την, έτσι, έγκυρη και αποτελεσματική.

4.2.6.5 Συγκρισιμότητα

Η συγκρισιμότητα είναι μια «υποκατηγορία» της εγκυρότητας και σχετίζεται με το εύρος των περιστάσεων στο οποίο μπορεί χρησιμοποιηθεί ένα όργανο μέτρησης. Στην περίπτωση της Semantic Differential, οι περιστάσεις που προαναφέρθηκαν έχουν να κάνουν με τα αντικείμενα και τα υποκείμενα που μπορούν να χρησιμοποιηθούν για τη μέτρηση. Ορισμένα ερωτήματα τα οποία ανακύπτουν σχετίζονται με την ηλικία των υποκειμένων, το φύλο τους, τη νοητική τους κατάσταση καθώς και με τα αποτελέσματα που προκύπτουν για τους factors στα διάφορα αντικείμενα τα οποία μπορούν να μετρηθούν. Ένα από τα πιο σημαντικά ερωτήματα αφορά την περίπτωση στην οποία εφαρμόζονται τα ίδια τεστ σε ανθρώπους διαφορετικών πολιτισμικών κόσμων και διαφορετικών εθνικοτήτων.

Τα στοιχεία που θα αναφερθούν παρακάτω έχουν να κάνουν με τη συγκρισιμότητα μεταξύ των αντικειμένων και μεταξύ των υποκειμένων.

- **Συγκρισιμότητα μεταξύ των υποκειμένων**

Αδιαμφισβήτητα, η μεγαλύτερη πρόκληση για τη συγκρισιμότητα μεταξύ των υποκειμένων έχει να κάνει με τις συγκρίσεις ανθρώπων που χρησιμοποιούν διαφορετικές γλώσσες και προέρχονται από διαφορετικές κουλτούρες. Το ερώτημα που ανακύπτει από αυτό είναι το εξής: όταν γίνεται μετάφραση των επιθέτων που χρησιμοποιούνται στις κλίμακες από την αγγλική γλώσσα σε άλλες, προκύπτουν τα ίδια συμπεράσματα και τα ίδια σκορ των factors με αυτά της αγγλικής;

Σε διάφορες έρευνες που έχουν γίνει γύρω από το θέμα, με κυριότερες αυτές των Kumata και Schramm (1956)²⁵, έχει προκύψει το συμπέρασμα ότι η μέθοδος Semantic Differential μπορεί να χρησιμοποιηθεί σε ένα μεγάλο εύρος γλωσσών. Συγκεκριμένα, σε έρευνες που διεξήχθησαν από τους Osgood (1964)²⁶ και Jakobovits (1966)²⁷, διαπιστώθηκε ότι η μέθοδος λειτουργεί με τον ίδιο τρόπο σε 15 γλώσσες (Αμερικάνικα, Αραβικά, Γερμανικά, Ελληνικά, Ισπανικά, Φινλανδικά, Γαλλικά, Ιταλικά, Σουηδικά, Κροατικά, Ιαπωνικά, Ινδικά, Καναδικά και σε δύο κινέζικες διαλέκτους (Cantonese, Flemish)).

- **Συγκρισιμότητα μεταξύ των αντικειμένων**

Ιδανική και τέλεια περίπτωση συγκρισιμότητας σε αυτό το θέμα θα υπήρχε εάν μεμονωμένες κλίμακες διατηρούσαν το ίδιο ακριβώς νόημα, ανεξάρτητα από το προς κρίση αντικείμενο. Παρ' όλα αυτά, επειδή πρακτικά δεν μπορεί να επιτευχθεί η ιδανική περίπτωση, η προϋπόθεση θα μπορούσε απλά να μετατραπεί στο γεγονός ότι οι ίδιοι factors θα πρέπει να επανεμφανίζονται για κάθε αντικείμενο, αν και οι κλίμακες που χρησιμοποιούνται μπορούν να αλλάζουν κάθε φορά.

²⁵ Kumata, H., & W. Schramm. (1956). A pilot study of cross-cultural meaning. *Publ. Opin. Quart.*, 20, 229-238. 171-176, 222.

²⁶ Osgood, C. E. (1964). Semantic differential technique in the comparative study of cultures. *American Anthropologist*, 66(3, Pt. 2), 171-200.

²⁷ Jakobovits, L. A. (1966). Comparative psycholinguistics in the study of cultures. *International Journal of Psychology*, 1, 15-37.

Έχει διαπιστωθεί ότι υπάρχει μεγάλος βαθμός αλληλεπίδρασης ανάμεσα στο αντικείμενο και τις κλίμακες. Αυτό σημαίνει ότι το νόημα των επιθέτων που χρησιμοποιούνται στις κλίμακες αλλάζει αρκετά ανάλογα με το αντικείμενο το οποίο κρίνεται. Φυσικά, όπως αναφέρθηκε νωρίτερα, αυτή η διαπίστωση δε χρήζει ιδιαίτερης σημασίας μιας και κατατάσσεται στην κατηγορία της ιδανικής περίπτωσης η οποία δεν υφίσταται πρακτικά. Το πρακτικό συμπέρασμα που προκύπτει από τις έρευνες είναι το γεγονός ότι οι factors έχουν μια επανάληψη στην εμφάνισή τους για κάθε αντικείμενο. Το συμπέρασμα αυτό παρέχει την πληροφορία ότι στον τομέα της συγκρισιμότητας μεταξύ αντικειμένων, η μέθοδος είναι έγκυρη.

4.2.6.6 Χρηστικότητα

Η χρηστικότητα της μεθόδου Semantic Differential είναι κυρίως θέμα του κάθε ερευνητή που τη χρησιμοποιεί. Ιδιαίτερη προσοχή κατά την εφαρμογή της πρέπει να δοθεί στο γεγονός ότι δεν θα πρέπει με κανέναν τρόπο το όργανο μέτρησης να επηρεάζει την κρίση των υποκειμένων που συμμετέχουν στην έρευνα. Αυτό, πρακτικά, σημαίνει ότι τα ερωτηματολόγια και η διαδικασία θα πρέπει να είναι κατασκευασμένα με τέτοιο τρόπο ώστε να αφήνουν το υποκείμενο ανεπηρέαστο να επιλέξει την απάντηση που εκείνο θεωρεί σωστή. Όταν τηρείται αυτό, τα αποτελέσματα της έρευνας θεωρούνται χρηστικά, μιας και θα έχει μετρηθεί η πραγματική αντίδραση και άποψη των υποκειμένων για τα βαθμολογούμενα αντικείμενα.

4.2.7 Διαδικασία Ανάλυσης

Για την ανάλυση των αποτελεσμάτων μιας έρευνας με τη χρήση της μεθόδου Semantic Differential ακολουθούνται συγκεκριμένα βήματα.

- Αρχικά, πρέπει να συνδυαστούν οι απαντήσεις των κλιμάκων και να χωριστούν σε κατηγορίες, ανάλογα με τους factors στους οποίους κατατάσσονται. Αυτό περιλαμβάνει την αρίθμηση των θέσεων των κλιμάκων, από το -3 ως το +3 (συνήθως το -3 βρίσκεται στον αρνητικό πόλο και το +3 στο θετικό. Για παράδειγμα, στο ζεύγος «καλός-κακός», το -3 βρίσκεται στο «κακός» ενώ το +3 στο «καλός»).

Πραγματοποιώντας αυτή τη διαδικασία, εξάγονται οι βαθμολογίες για κάθε factor ξεχωριστά. Έτσι, ο κάθε factor καταλήγει να έχει έναν αριθμό για κάθε βαθμολογούμενο αντικείμενο, για κάθε υποκείμενο.

Προσοχή θα πρέπει να δοθεί στο γεγονός ότι ορισμένες κλίμακες δεν έχουν την ίδια βαρύτητα με κάποιες άλλες. Έτσι σε περίπτωση που χρησιμοποιηθούν άνισες κλίμακες, ο υπολογισμός των σκορ αυτών θα πρέπει να γίνει ακολουθώντας μια πιο προσεκτική και λεπτομερή διαδικασία.

- Στη συνέχεια πρέπει να υπολογιστούν οι ομαδικοί βαθμοί για κάθε factor, σε κάθε αντικείμενο. Για να γίνει αυτό, πρέπει να υπολογιστεί ο μέσος όρος όλων των υποκειμένων για έναν factor.
- Τέλος, γίνεται η ανάλυση των παραγόντων και η εξαγωγή συμπερασμάτων για τα αντικείμενα που εξετάστηκαν, μέσω μαθηματικών υπολογισμών.

4.3 Πλεονεκτήματα της Μεθόδου

Η μέθοδος Semantic Differential επελέγη, κατόπιν μελέτης άλλων μεθόδων υποκειμενικής αξιολόγησης, διότι τα πλεονεκτήματα που συγκεντρώνει δε βρίσκονται σε άλλες μεθόδους:

- Η χρήση των διπολικών κλιμάκων, τα διαστήματα των οποίων είναι ισοκατανομημένα, είναι ένα χαρακτηριστικό το οποίο βοηθά τα υποκείμενα να εκφράσουν όσο ακριβέστερα γίνεται την άποψη τους για κάποιο αντικείμενο. Επίσης, οι διπολικές κλίμακες της συγκεκριμένης μεθόδου έχουν και το «ουδέτερο» ενσωματωμένο σε αυτές, το οποίο περιλαμβάνει και τις απαντήσεις «δεν ξέρω, δε γνωρίζω», στοιχείο σημαντικό για τα υποκείμενα, καθώς χωρίς αυτό, κάποιες από τις απαντήσεις που θα δίνονταν, δεν θα αντιπροσώπευαν την πραγματική άποψη του κάθε υποκειμένου.
- Με σωστή επιλογή διπολικών κλιμάκων, μπορεί να εξετασθεί ένα πολύ μεγάλο φάσμα χαρακτηριστικών, δίνοντας έτσι το πλεονέκτημα της σφαιρικής μελέτης

ενός αντικειμένου. Αυτό, για την παρούσα εργασία, είναι ιδιαίτερα σημαντικό, μιας και ο αλγόριθμος του MP3 επηρεάζει πολλά από τα ποιοτικά χαρακτηριστικά του ήχου και της μουσικής, και πρέπει να υπάρχει η δυνατότητα εξέτασης όσο περισσότερων εξ' αυτών γίνεται.

- Οι κλίμακες δε χρειάζεται να είναι υπερβολικές σε αριθμό (τρεις για κάθε παράγοντα αρκούν για την εξαγωγή σωστών αποτελεσμάτων). Αυτό, καθιστά τη μέθοδο σχετικά ξεκούραστη για τα υποκείμενα που συμμετέχουν στην έρευνα, ιδίως στην παρούσα εργασία, όπου οι πολύωρες ακροάσεις μπορεί να γίνουν ιδιαίτερα κουραστικές.
- Η κατηγοριοποίηση των διπολικών κλιμάκων σε παράγοντες, που παρέχει η συγκεκριμένη μέθοδος, είναι εύκολη στην εφαρμογή της και εξασφαλίζει την εξαγωγή σωστών και έγκυρων συμπερασμάτων με απλές μαθηματικές μεθόδους. Ακόμη, λόγω των παραγόντων αυτών, τυχόν δειγματοληπτικά λάθη τα οποία εντοπίζονται σε μεμονωμένες κλίμακες και δεν οφείλονται σε κακό σχεδιασμό, μπορούν να εκμηδενιστούν.
- Είναι μια μέθοδος η οποία δεν απαιτεί ιδιαίτερες γνώσεις από τα υποκείμενα τα οποία συμμετέχουν στην έρευνα, καθώς δε χρειάζεται κανένας ιδιαίτερος εξοπλισμός για την αξιολόγηση των αντικειμένων.
- Ο αριθμός των υποκειμένων που απαιτούνται για την εξαγωγή γενικών συμπερασμάτων για ένα αντικείμενο είναι σχετικά μικρός ($N \geq 30$), όταν το δείγμα είναι πιθανοτικό.

ΚΕΦΑΛΑΙΟ 5^ο: ΣΧΕΔΙΑΣΜΟΣ ΤΗΣ ΠΕΙΡΑΜΑΤΙΚΗΣ ΔΙΑΔΙΚΑΣΙΑΣ-ΜΕΘΟΔΟΛΟΓΙΑ

Εισαγωγή 5^{ου} Κεφαλαίου

Στο κεφάλαιο αυτό, αναλύεται ο τρόπος με τον οποίο σχεδιάστηκε και διεξήχθη η πειραματική διαδικασία. Στη συνέχεια, αναφέρονται τα ερωτήματα στα οποία εστιάζεται το ενδιαφέρον της παρούσας εργασίας.

Σκοπός αυτού του κεφαλαίου είναι η κατανόηση της λογικής με την οποία σχεδιάστηκε η πειραματική διαδικασία καθώς και τα ερωτήματα τα οποία καλείται να απαντήσει.

Για τη διεξαγωγή της πειραματικής διαδικασίας, κατασκευάστηκαν listening tests, σύμφωνα με τη μέθοδο Semantic Differential, η οποία αναλύθηκε στο προηγούμενο κεφάλαιο.

Τα tests αυτά είναι ερωτηματολόγια τα οποία αποτελούνται από διπολικές κλίμακες οι οποίες έχουν στα άκρα τους αντίθετους προσδιορισμούς. Για την κατασκευή τους, ελήφθησαν υπ' όψιν ορισμένες παράμετροι, οι οποίες είναι ιδιαίτερα σημαντικές για την εγκυρότητα της έρευνας.

5.1 Ερευνητικά Ερωτήματα

Τα ερωτήματα στα οποία καλείται η παρούσα πτυχιακή να απαντήσει αφορούν την ποιότητα του ηχητικού σήματος που παράγεται με τη συμπίεση του, μέσω του προτύπου MPEG-1 Layer III. Συγκεκριμένα, τα ερωτήματα αυτά αφορούν τόσο στο συχνοτικό φάσμα του παραγόμενου ήχου, όσο και την εισαγωγή θορύβου και τον τομέα των δυναμικών και είναι τα ακόλουθα:

- Η εισαγωγή του θορύβου γίνεται αντιληπτή από τους ακροατές; Κι αν ναι, από ποιο bitrate και κάτω και σε ποια είδη μουσικής;

- Η αποκοπή συχνοτήτων γίνεται αντιληπτή από τους ακροατές; Κι αν ναι, ποιες συχνότητες παρουσιάζουν μεγαλύτερη έλλειψη, σε ποια μουσικά είδη και σε ποιά bitrate υπάρχει σημαντική διαφορά;
- Η αυξομείωση των δυναμικών επηρεάζεται από τη συμπίεση του προτύπου; Κι αν ναι, σε ποια είδη μουσικής γίνεται περισσότερο αντιληπτή από τους ακροατές και σε ποια bitrate;

5.2 Επιλογή Κλιμάκων

Ο αριθμός των κλιμάκων καθώς και οι προσδιορισμοί που χρησιμοποιούνται στα listening tests επελέγησαν με βάση τρία βασικά κριτήρια: τη συμφωνία τους με τη δομή EPA που είναι απαραίτητη για τη χρήση της μεθόδου Semantic Differential, τη φύση των αντικειμένων που εξετάζονται (ποιοτικά χαρακτηριστικά) και την κοινή ερμηνεία τους από τα υποκείμενα που έλαβαν μέρος στην έρευνα.

Η βαθμολόγηση των κλιμάκων γίνεται από το -3 έως το 3. Το σημείο -3 αντιπροσωπεύει τη χειρίστη βαθμολογία (π.χ. Πολύ Χειρότερο) και το 3 την άριστη (π.χ. Πολύ Καλύτερο). Η επιλογή αυτή έγινε με σκοπό την ευκολότερη κατανόηση της θετικής και της αρνητικής άποψης των υποκειμένων για κάθε ένα από τα ποιοτικά χαρακτηριστικά των ακουσμάτων.

5.2.1 Αριθμός Κλιμάκων

Πολλές μελέτες έχουν γίνει γύρω από τον αριθμό των κλιμάκων που πρέπει να χρησιμοποιούνται σε έρευνες που γίνονται με βάση τη μέθοδο Semantic Differential, με κυριότερες αυτές των DiVesta και Dick (1966)²⁸ και Norman (1959)²⁹. Τα αποτελέσματά τους έδειξαν ότι για κάθε παράγοντα θα πρέπει να χρησιμοποιούνται τρεις με τέσσερις κλίμακες, ενώ παραδόξως, η προσθήκη περισσότερων κλιμάκων, δε βοηθά σχεδόν καθόλου στην εξαγωγή πιο ακριβών συμπερασμάτων. Έτσι λοιπόν, η επιλογή

²⁸ DiVesta, F. J., & Dick, W. (1966). The test-retest reliability of children's ratings on the semantic differential. *Educational and Psychological Measurement*, 26, 605-616.

²⁹ Norman, W. (1959). Stability characteristics of the semantic differential. *American Journal of Psychology*, 72, 581-584.

εννέα συνολικά κλιμάκων (τρεις για evaluation, τρεις για potency και τρεις για activity), συμβαδίζει με τα ευρήματα αυτά κι έτσι χρησιμοποιήθηκε στην παρούσα εργασία.

Εδώ, θα πρέπει να σημειωθεί ότι η επιλογή του αριθμού των κλιμάκων συμβαδίζει και με πιο πρακτικά κριτήρια, όπως αυτό της υπομονής των υποκειμένων. Σε μια έρευνα με πάρα πολλές κλίμακες αξιολόγησης, τα υποκείμενα θα μπορούσαν εύκολα να κουραστούν και οι απαντήσεις που θα δίνονταν δεν θα αντιπροσώπευαν την πραγματική τους άποψη.

5.2.2 Συμφωνία με τη Δομή EPA

Τα listening tests συμβαδίζουν απόλυτα με τη λογική της δομής EPA. Ισχύουν, λοιπόν, όλες οι προϋποθέσεις και οι περιορισμοί που αναφέρθηκαν στο προηγούμενο κεφάλαιο.

Στα listening tests, οι προσδιορισμοί που χρησιμοποιούνται για κάθε factor φαίνονται στον Πίνακα 5.1.

Evaluation	Potency	Activity
Χειρότερο – Καλύτερο	Περισσότερος Θόρυβος – Λιγότερος Θόρυβος	Πιο Φτωχό Φάσμα – Πιο Πλούσιο Φάσμα
Πιο Ενοχλητικό – Πιο Απολαυστικό	Πιο Δύσκολος Διαχωρισμός Οργάνων – Πιο εύκολος Διαχωρισμός Οργάνων	Λιγότερο Ζεστό – Περισσότερο Ζεστό
Λιγότερο Δυνατό – Περισσότερο Δυνατό	Περισσότερη Παραμόρφωση – Λιγότερη Παραμόρφωση	Λιγότερο Λαμπρό – Περισσότερο Λαμπρό

Πίνακας 5.1: Αντιστοιχία κλιμάκων και παραγόντων των listening tests

Ένα χαρακτηριστικό στην έρευνα αυτή που διαφέρει από τις περισσότερες έρευνες που έχουν γίνει με τη μέθοδο αυτή, είναι η χρήση του συγκριτικού βαθμού στους προσδιορισμούς. Το ερωτηματολόγιο σχεδιάστηκε κατ' αυτόν τον τρόπο μιας και η αξιολόγηση των αντικειμένων προκύπτει από τη σύγκριση των αρχείων WAV με τα αντίστοιχα αρχεία MP3. Τα υποκείμενα καλούνται να αξιολογήσουν τα αρχεία MP3 με σημείο αναφοράς το αντίστοιχο αρχείο WAV. Συνεπώς, κατά το σχεδιασμό των ερωτηματολογίων, πρέπει να αποτυπώνεται η σύγκριση αυτή, μιας και μόνο έτσι είναι εφικτή η εξαγωγή συμπερασμάτων.

5.2.3 Φύση του Αντικειμένου

Η αξιολόγηση της συμπίεσης σημάτων ήχου είναι μια διαδικασία η οποία έχει ερευνηθεί διεξοδικά σε αντικειμενικό επίπεδο. Με τη χρήση συγκεκριμένων λογισμικών, οι ερευνητές έχουν αξιολογήσει διάφορα πρότυπα συμπίεσης. Οι αξιολογήσεις, όμως, αυτές δεν περιέχουν τον ανθρώπινο παράγοντα. Δεν έχει ληφθεί, δηλαδή, υπ' όψιν η κρίση των ακροατών οι οποίοι χρησιμοποιούν στην καθημερινότητά τους ηχητικά αρχεία που έχουν υποστεί συμπίεση. Αυτό είναι και το κυρίαρχο στοιχείο στις έρευνες με τη χρήση αυτής της μεθόδου. Οι κλίμακες που επιλέγονται αλλά και οι ομαδοποίηση αυτών σε διάφορους παράγοντες δίνουν μια πολύ σαφή εικόνα της υποκειμενικής άποψης των ακροατών.

Στον τομέα της υποκειμενικής αξιολόγησης ηχητικών σημάτων, έχουν κατασκευαστεί ορισμένα πρότυπα. Έχουν επίσης γίνει πολλές μελέτες για την αξιολόγηση των ήχων που παράγονται από συγκεκριμένες ηχητικές πηγές, όπως αεροσκάφη, κλιματιστικά, ηλεκτρικές σκούπες κλπ. Επίσης, έχουν γίνει μελέτες σχετικές με την αξιολόγηση των αλγορίθμων συμπίεσης και συγκεκριμένα του mp3 σε διάφορα είδη μουσικής με τη μέθοδο AB-Blind Test³⁰. Παρ' όλα αυτά, η μέθοδος Semantic Differential παρουσιάζει ορισμένα πλεονεκτήματα τα οποία δεν υπάρχουν στις προαναφερθείσες έρευνες. Η ομαδοποίηση των κλιμάκων είναι το σημαντικότερο εξ' αυτών μιας και μέσω αυτής αφ' ενός μπορεί κανείς να εξάγει εύκολα γενικευμένα συμπεράσματα, αφ' ετέρου, ελαχιστοποιούνται τυχόν ερευνητικά σφάλματα.

Τα ποιοτικά χαρακτηριστικά που ερευνώνται στην παρούσα εργασία, συνδέονται άμεσα με τη διαδικασία συμπίεσης η οποία ακολουθείται από το πρότυπο και είναι τα φασματικά και δυναμικά χαρακτηριστικά των μουσικών κομματιών, καθώς επίσης και ο θόρυβος ή η παραμόρφωση που μπορεί να εισαχθεί κατά τη συμπίεση.

5.2.4 Ερμηνεία των κλιμάκων από τα υποκείμενα

Οι έρευνες οι οποίες γίνονται με σκοπό τη μέτρηση υποκειμενικών χαρακτηριστικών, ενέχουν πάντα κινδύνους για την παρερμηνεία των ερωτήσεων από τα υποκείμενα. Στην

³⁰ Pras, A., Zimmerman, R., Levitin, D., Guastavino, C. (2009). Subjective evaluation of mp3 compression for different musical genres. AES 127th Convention, New York, NY, USA.

παρούσα εργασία, οι κλίμακες που έχουν επιλεγεί για την αξιολόγηση των MP3, δεν μπορούν εύκολα να παρερμηνευθούν. Παρ' όλα αυτά, κατά τη διεξαγωγή της πειραματικής διαδικασίας, τα υποκείμενα έλαβαν όλες τις απαραίτητες οδηγίες και επεξηγήσεις για τον αποκλεισμό παρερμηνειών. Συνεπώς, όλα τα υποκείμενα κατά τη διάρκεια της διαδικασίας, είχαν σαφή εικόνα για τη σημασία των κλιμάκων.

5.3 Επιλογή Ακουσμάτων

Η έρευνα αυτή γίνεται με σκοπό την αξιολόγηση του προτύπου MP3 ανάλογα με το είδος μουσικής το οποίο χρησιμοποιείται. Είναι, λοιπόν, λογικό να δοθεί μεγάλη βαρύτητα στην επιλογή των ακουσμάτων, μιας και κάθε είδος χαρακτηρίζεται από συγκεκριμένα πράγματα. Τα είδη που εξετάζονται στην παρούσα εργασία είναι: rock, jazz, ηλεκτρονική μουσική, κλασική μουσική (με ορχήστρα) και solo όργανο.

5.3.1 Rock

Η μουσική rock χαρακτηρίζεται από έντονο ρυθμό και από ευδιάκριτη, χαρακτηριστική μελωδία φωνητικών η οποία συνοδεύεται συνήθως από ηλεκτρικές κιθάρες, ηλεκτρικό μπάσο και drums, χωρίς ιδιαίτερες αυξομειώσεις στη δυναμική. Πολλές φορές χρησιμοποιούνται και πληκτροφόρα όργανα. Ένα χαρακτηριστικό παράδειγμα rock μουσικής το οποίο χρησιμοποιείται στην παρούσα εργασία είναι το κομμάτι “Poison” του Alice Cooper.

5.3.2 Jazz

Η jazz είναι ένα είδος μουσικής το οποίο χαρακτηρίζεται από περίπλοκες μελωδικές γραμμές, αυξομειώσεις των δυναμικών και χρήση κρουστών, πνευστών, πληκτροφόρων και εγχόρδων οργάνων. Μερικά χαρακτηριστικά όργανα που χρησιμοποιούνται είναι το μπάσο, η drums, η τρομπέτα, το σαξόφωνο, η κιθάρα, το πιάνο κ.α. Ένα χαρακτηριστικό δείγμα αυτής της μουσικής, που χρησιμοποιείται στην παρούσα εργασία είναι το “Blue Train” του John Coltrane.

5.3.3 Ηλεκτρονική Μουσική

Η ηλεκτρονική μουσική είναι ένα είδος μουσικής το οποίο έχει ως κύριο χαρακτηριστικό την αποκλειστική χρήση ηλεκτρονικών ήχων για τη σύνθεση των κομματιών. Ουσιαστικά, δηλαδή, τα φυσικά μουσικά όργανα αντικαθίστανται από υπολογιστές, samplers και άλλες ηλεκτρονικές συσκευές. Ένα χαρακτηριστικό δείγμα της μουσικής αυτής, το οποίο χρησιμοποιείται και στην παρούσα εργασία, είναι το κομμάτι “Traffic” του DJ Tiesto.

5.3.4 Κλασική Μουσική (Ορχήστρα)

Η κλασική μουσική είναι ένα περίπλοκο είδος το οποίο χαρακτηρίζεται από πληθώρα μουσικών οργάνων, έγχορδα, πνευστά και κρουστά, από περίτεχνες μελωδίες και έντονες αυξομειώσεις δυναμικών. Ένα αντιπροσωπευτικό παράδειγμα του συγκεκριμένου είδους είναι η 3^η συμφωνία (Ηρωική) του Ludwig van Beethoven, το οποίο χρησιμοποιείται στην παρούσα εργασία.

5.3.5 Solo Όργανο

Αν και δεν αποτελεί συγκεκριμένο είδος μουσικής, ένα άκουσμα το οποίο περιέχει μόνο ένα όργανο έχει ιδιαίτερη σημασία για την εργασία αυτή μιας και το φαινόμενο masking είναι σχεδόν ανύπαρκτο σε τέτοιου είδους ακούσματα. Μιας και το πρότυπο MP3 είναι σχεδιασμένο έτσι ώστε να εκμεταλλεύεται τέτοιου είδους ακουστικά φαινόμενα, παρουσιάζει μεγάλο ενδιαφέρον η λειτουργία του σε ακούσματα που δεν έχουν πολλά τέτοια φαινόμενα. Στην παρούσα εργασία χρησιμοποιείται το κομμάτι “Moonlight Sonata” του Ludwig van Beethoven, παιγμένο από ένα πιάνο.

Τα αποσπάσματα των παραπάνω μουσικών κομματιών τα οποία επελέχθησαν είναι διάρκειας από 8 έως 12 δευτερόλεπτα και περιέχουν μεγάλο συχνοτικό φάσμα (ανάλογα με το μουσικό είδος) καθώς και τις απαραίτητες αυξομειώσεις δυναμικών, όπου αυτές είναι χαρακτηριστικές του κάθε μουσικού είδους. Η συγκεκριμένη χρονική διάρκεια επελέγη καθώς επιτρέπει στα υποκείμενα να συγκρατούν εύκολα στο μυαλό τους το εκάστοτε κομμάτι έτσι ώστε να μπορούν να κάνουν τη σύγκριση ανάμεσα στο ασυμπιεσμένο και συμπιεσμένο σήμα.

5.4 Επιλογή Υποκειμένων

Η επιλογή των υποκειμένων είναι μια διαδικασία ουσιώδης και πολύ σημαντική σε έρευνες τέτοιας φύσεως. Η διαδικασία αυτή γίνεται με βάση τρεις βασικούς παράγοντες επιλογής: τον αριθμό των υποκειμένων, την ηλικία και τη σχέση τους με το αντικείμενο.

Ο αριθμός των υποκειμένων που απαιτείται να λαμβάνουν μέρος σε έρευνες με τη χρήση της συγκεκριμένης μεθόδου είναι 30, σύμφωνα με τον Heise³¹. Έτσι, λοιπόν, στην παρούσα έρευνα συμμετέχουν 30 υποκείμενα και από τα δύο φύλα (15 άντρες, 15 γυναίκες) με εύρος ηλικιών από 18 έως 29 έτη.

Βρέθηκαν ορισμένες μελέτες, όπως αυτή των Hansen και Weber³², στις οποίες έχει χρησιμοποιηθεί μικρότερος αριθμός υποκειμένων. Παρ' όλα αυτά, προτιμήθηκε η συμμετοχή 30 ατόμων, προκειμένου να εξασφαλιστεί η εγκυρότητα της έρευνας.

Οι ηλικίες των υποκειμένων, όπως αναφέρθηκε και παραπάνω, είναι από 18 έως 29 ετών. Η επιλογή αυτή έγινε με βάση διάφορες μελέτες οι οποίες σχετίζονται με την απώλεια της ακοής λόγω ηλικίας, και συγκεκριμένα συχνοτικών απωλειών, όπως αυτή του Takeda³³. Έτσι, προτιμήθηκε η ηλικιακή ομάδα κάτω των 30 ετών για να μην υπάρχει κίνδυνος αλλοίωσης των αποτελεσμάτων λόγω προβλημάτων ακοής.

Ένας τρίτος παράγοντας που καθόρισε την επιλογή των υποκειμένων είναι η σχέση τους με το εξεταζόμενο αντικείμενο. Οι ακροατές που επελέγησαν, γνωρίζουν και χρησιμοποιούν σε καθημερινή βάση κομμάτια MP3, χωρίς όμως να διαθέτουν ιδιαίτερα εξειδικευμένες γνώσεις για το αντικείμενο (τρόπο κατασκευής και λειτουργίας του αλγορίθμου). Κατ' αυτό τον τρόπο, τα υποκείμενα δεν είναι υποψιασμένα και προκατειλημμένα για τις αλλοιώσεις που μπορεί να επιφέρει η συμπίεση. Έτσι, διασφαλίζεται η εγκυρότητα των απαντήσεων τους και το γεγονός ότι η αξιολόγηση θα

³¹ Heise, D. R. (1969). Some methodological issues in semantic differential. *Psychological Bulletin*, 72 6, 406-422.

³² Hansen, H., Weber, R. (2009). Semantic evaluations of noise with tonal components in Japan, France and Germany-A cross-cultural comparison. *Journal of acoustic society of America*. 125, 2, 850-862

³³ Takeda, S. et al (1992). Age variation in the upper limit of hearing. *European journal of applied physiology*. 65, 403-408

γίνει αυθόρμητα, με μοναδικό γνώμονα την ακουστική τους εμπειρία και όχι τις τεχνικές τους γνώσεις. Επίσης, τα υποκείμενα των ακροάσεων έχουν μερικές μουσικές και τεχνικές γνώσεις για τον ήχο, έτσι ώστε να είναι σε θέση να αντιληφθούν έννοιες όπως το φάσμα ενός ήχου ή τις δυναμικές ενός κομματιού.

Το δείγμα που χρησιμοποιήθηκε, όμως, είναι μη πιθανοτικό και, επομένως, όχι τυχαίο. Συνεπώς τα αποτελέσματα που παρατίθενται σε επόμενο κεφάλαιο είναι δύσκολο να γενικευτούν στον γενικό πληθυσμό. Το μη πιθανοτικό δείγμα επελέγη διότι κρίθηκε απαραίτητη η μερική τεχνική γνώση των υποκειμένων προκειμένου να κατανοήσουν πλήρως το ερωτηματολόγιο.

5.5 Εξοπλισμός και Λογισμικό

Για την προετοιμασία της πειραματικής διαδικασίας, χρησιμοποιήθηκαν:

- Το λογισμικό Wavelab v.5.01b, για την κατασκευή των αρχείων MP3 από τα πρωτότυπα WAV. Τα αρχεία κατασκευάστηκαν σε bitrates 96, 160 και 320kbps, με τον αλγόριθμο κωδικοποίησης Lame Encoder (ο οποίος έχει χρησιμοποιηθεί σε παραπλήσιες έρευνες³⁴), σε Intensity Stereo Mode και με συχνότητα δειγματοληψίας 44kHz. Ο αλγόριθμος αυτός επελέγη λόγω απόλυτης συμβατότητας με τον αλγόριθμο αποκωδικοποίησης.
- Τα ακουστικά κλειστού τύπου Sennheiser HD-215, με συχνοτική απόκριση 12 με 22.000Hz και συνολική αρμονική παραμόρφωση THD<0,2%, για τον έλεγχο των κωδικοποιημένων αρχείων.

Για την ακρόαση των κομματιών και την αξιολόγηση τους από τα υποκείμενα, χρησιμοποιήθηκαν:

- Τα ακουστικά κλειστού τύπου Sennheiser HD-215, που αναφέρθηκαν παραπάνω.

³⁴ Pras, A., Zimmerman, R., Levitin, D., Guastavino, C. (2009). Subjective evaluation of mp3 compression for different musical genres. AES 127th Convention, New York, NY, USA.

- Το λογισμικό Winamp v.5.541 για την αναπαραγωγή των κομματιών. Το λογισμικό αυτό επελέγη διότι χρησιμοποιεί αλγόριθμο αποκωδικοποίησης ο οποίος είναι απόλυτα συμβατός με τον αλγόριθμο κωδικοποίησης που χρησιμοποιεί το Wavelab.
- Έντυπα ερωτηματολόγια που δόθηκαν στα υποκείμενα και περιείχαν τις ερωτήσεις αξιολόγησης καθώς και οδηγίες για την ορθή συμπλήρωσή τους.

Ολόκληρο το ερωτηματολόγιο παρατίθεται στο Παράρτημα Α και η σειρά και μορφή των κομματιών στο Παράρτημα Β.

Οι ακροάσεις πραγματοποιήθηκαν στο Γ.Ε.Ι. Ρεθύμνου και συγκεκριμένα στις αίθουσες των εργαστηρίων Μουσικής Διάδρασης και Πολυφωνίας και Πρωτοκόλλων Μουσικής Επικοινωνίας. Οι αίθουσες αυτές επελέγησαν λόγω καλής ηχομόνωσης και πολύ μικρού θορύβου βάθους, έτσι ώστε τα υποκείμενα να μην ενοχλούνται από τους εσωτερικούς και εξωτερικούς θορύβους.

5.6 Πειραματική Διαδικασία

Οι ακροάσεις έγιναν ατομικά (ένα υποκείμενο κάθε φορά) και διεξήχθησαν με τον ακόλουθο τρόπο, ο οποίος επαναλήφθηκε για κάθε υποκείμενο:

- Κάθε υποκείμενο εισέρχεται στην αίθουσα και καθόταν μπροστά στο laptop. Στην επιφάνεια εργασίας βρισκόταν ήδη ανοιχτό το Winamp με τα προς ακρόαση κομμάτια φορτωμένα στη λίστα αναπαραγωγής του.
- Το υποκείμενο έπαιρνε το ερωτηματολόγιο, διάβαζε τις οδηγίες συμπλήρωσης και γίνονταν όλες οι απαραίτητες διευκρινίσεις για τις κλίμακες αξιολόγησης.
- Στη συνέχεια, προσαρμοζόταν η ένταση αναπαραγωγής των κομματιών για το κάθε υποκείμενο, με βάση τα κομμάτια που είχαν τη μεγαλύτερη και τη μικρότερη συνολική ένταση.
- Κατόπιν, το υποκείμενο προχωρούσε στην αξιολόγηση των ακουσμάτων. Οι επαναλήψεις των ακουσμάτων δεν είχαν περιορισμό στον αριθμό, όπως επίσης και ο χρόνος που θα αφιέρωνε το υποκείμενο στο κάθε κομμάτι. Ο μέσος όρος

συμπλήρωσης του ερωτηματολογίου ήταν τα 50 λεπτά, με το γρηγορότερο υποκείμενο να τα αξιολογεί σε περίπου 40 λεπτά και το πιο αργό σε περίπου 80 λεπτά.

Κατά τη διαδικασία αξιολόγησης, το κάθε υποκείμενο αξιολογούσε ένα άκουσμα με βάση ένα άλλο. Για παράδειγμα, αξιολογούσε το Κομμάτι 1 έχοντας σα μέτρο σύγκρισης το Κομμάτι 2. Τα ζεύγη των κομματιών αποτελούνταν από το αρχικό WAV και ένα MP3 σε συγκεκριμένο bitrate. Συνολικά αξιολογήθηκαν 15 ακούσματα, πέντε διαφορετικά είδη μουσικής σε τρεις διαφορετικές “εκδοχές”, κωδικοποιημένα κάθε φορά σε διαφορετικά bitrates.

Η επιλογή των κομματιών γινόταν κάνοντας διπλό click με το ποντίκι πάνω στο κομμάτι που ήθελε το υποκείμενο να ακούσει. Οι επιλογές “forward” και “rewind” γίνονταν με το δεξί και αριστερό βέλος του πληκτρολογίου αντίστοιχα, ενώ τα “play” και “pause” με το πλήκτρο C του πληκτρολογίου.

ΚΕΦΑΛΑΙΟ 6^ο: ΑΠΟΤΕΛΕΣΜΑΤΑ ΤΗΣ ΕΡΕΥΝΑΣ- ΣΥΓΚΡΙΣΕΙΣ

Εισαγωγή 6^{ου} Κεφαλαίου

Στο κεφάλαιο αυτό, παρουσιάζονται τα αποτελέσματα της έρευνας με τη βοήθεια πινάκων, γραφημάτων και ορισμένων στατιστικών δεικτών. Επίσης, γίνονται συγκρίσεις, αρχικά για κάθε είδος μουσικής ξεχωριστά, ανάλογα με το bitrate και στη συνέχεια συγκρίνονται τα μουσικά είδη μεταξύ τους έτσι ώστε να διαπιστωθεί η ποιοτική απόδοση των αρχείων MP3.

Η παρουσίαση των αποτελεσμάτων θα ξεκινήσει με την παράθεση πινάκων, για κάθε είδος μουσικής και για κάθε bitrate ξεχωριστά. Στη συνέχεια, μέσω γραφημάτων και πινάκων θα συγκριθούν τα διαφορετικά bitrates για κάθε είδος μουσικής. Τέλος, θα γίνει σύγκριση ανάμεσα στα είδη μουσικής, εξετάζοντας τα ανά bitrate.

Υπενθυμίζεται ότι η αρίθμηση των κλιμάκων ξεκινά από το -3 και φτάνει ως το +3.

6.1 Rock

Ξεκινώντας με το είδος της rock, θα παρουσιαστούν τα αποτελέσματα για κάθε bitrate ξεχωριστά, τόσο σε επίπεδο κλιμάκων όσο και σε επίπεδο factor scores³⁵.

6.1.1 Rock με bitrate 96kbps

Στον Πίνακα 6.1, φαίνονται τα αποτελέσματα του πειράματος που εξήχθησαν από τις απαντήσεις των υποκειμένων, για κάθε κλίμακα ξεχωριστά, για το είδος Rock στα 96kbps. Στον πίνακα αυτό, αναφέρεται το μέγεθος του δείγματος (N) και ο μέσος όρος των απαντήσεων.

³⁵ Τα factor scores υπολογίζονται βρίσκοντας το μέσο όρο των κλιμάκων που εντάσσονται σε κάθε factor, μιας και οι κλίμακες είναι ισοδύναμες. Έτσι, για παράδειγμα, ο factor “evaluation”, βρίσκεται από το μέσο όρο των κλιμάκων «Χειρότερο – Καλύτερο», «Πιο Ενοχλητικό - Πιο Απολαυστικό» και «Λιγότερο Δυνατό – Περισσότερο Δυνατό». Αντίστοιχα, υπολογίζονται και οι άλλοι δύο factors, σύμφωνα με την κατηγοριοποίηση των κλιμάκων που φαίνεται στον Πίνακα 5.1 του προηγούμενου κεφαλαίου.

Κλίμακες	N	Μέσος Όρος ³⁶
Χειρότερο - Καλύτερο	30	-2,2
Πιο Φτωχό Φάσμα - Πιο Πλούσιο Φάσμα	30	-1,8
Περισσότερος Θόρυβος - Λιγότερος Θόρυβος	30	-0,4
Πιο Ενοχλητικό - Πιο Απολαυστικό	30	-2
Λιγότερο Δυνατό - Περισσότερο Δυνατό	30	-1
Πιο Δύσκολος Διαχωρισμός Οργάνων - Πιο Εύκολος Διαχωρισμός Οργάνων	30	-1,8
Λιγότερο Ζεστό - Περισσότερο Ζεστό	30	-1,6
Περισσότερη Παραμόρφωση - Λιγότερη Παραμόρφωση	30	-1
Λιγότερο Λαμπρό - Περισσότερο Λαμπρό	30	-1,2

Πίνακας 6.1: Αποτελέσματα για το είδος Rock στα 96kbps

Σε επίπεδο factor scores, τα αποτελέσματα που εξήχθησαν από τις απαντήσεις των υποκειμένων φαίνονται στον Πίνακα 6.2.

Factor	Factor Scores
Evaluation	-1,73
Potency	-1,07
Activity	-1,53

Πίνακας 6.2: Factor scores για το είδος Rock στα 96kbps

6.1.2 Rock με bitrate 160kbps

Στον Πίνακα 6.3, φαίνονται τα αποτελέσματα του πειράματος που εξήχθησαν από τις απαντήσεις των υποκειμένων, για κάθε κλίμακα ξεχωριστά, για το είδος Rock στα 160kbps.

Κλίμακες	N	Μέσος Όρος
Χειρότερο - Καλύτερο	30	-1,4
Πιο Φτωχό Φάσμα - Πιο Πλούσιο Φάσμα	30	-1,6
Περισσότερος Θόρυβος - Λιγότερος Θόρυβος	30	-0,4
Πιο Ενοχλητικό - Πιο Απολαυστικό	30	-0,6
Λιγότερο Δυνατό - Περισσότερο Δυνατό	30	-0,4
Πιο Δύσκολος Διαχωρισμός Οργάνων - Πιο Εύκολος Διαχωρισμός Οργάνων	30	-0,8
Λιγότερο Ζεστό - Περισσότερο Ζεστό	30	-1

³⁶ Οι αρνητικές τιμές στο μέσο όρο δηλώνουν ότι στις συγκεκριμένες κλίμακες, το MP3 αξιολογήθηκε αρνητικά σε σύγκριση με το αντίστοιχο WAV.

Περισσότερη Παραμόρφωση - Λιγότερη Παραμόρφωση	30	-0,6
Λιγότερο Λαμπρό - Περισσότερο Λαμπρό	30	-1

Πίνακας 6.3: Αποτελέσματα για το είδος Rock στα 160kbps

Σε επίπεδο factor scores, τα αποτελέσματα από τις απαντήσεις των υποκειμένων φαίνονται παρακάτω, στον Πίνακα 6.4.

Factor	Factor Scores
Evaluation	-0,8
Potency	-0,6
Activity	-1,2

Πίνακας 6.4: Factor scores για το είδος Rock στα 160kbps

6.1.3 Rock με bitrate 320kbps

Στον Πίνακα 6.5, φαίνονται τα αποτελέσματα του πειράματος που εξήχθησαν από τις απαντήσεις των υποκειμένων, για κάθε κλίμακα ξεχωριστά, για το είδος Rock στα 320kbps.

Κλίμακες	N	Μέσος Όρος
Χειρότερο - Καλύτερο	30	-0,2
Πιο Φτωχό Φάσμα - Πιο Πλούσιο Φάσμα	30	-0,6
Περισσότερος Θόρυβος - Λιγότερος Θόρυβος	30	-0,2
Πιο Ενοχλητικό - Πιο Απολαυστικό	30	0,2
Λιγότερο Δυνατό - Περισσότερο Δυνατό	30	-0,6
Πιο Δύσκολος Διαχωρισμός Οργάνων - Πιο Εύκολος Διαχωρισμός Οργάνων	30	-0,2
Λιγότερο Ζεστό - Περισσότερο Ζεστό	30	-0,2
Περισσότερη Παραμόρφωση - Λιγότερη Παραμόρφωση	30	-0,2
Λιγότερο Λαμπρό - Περισσότερο Λαμπρό	30	-0,6

Πίνακας 6.5: Αποτελέσματα για το είδος Rock στα 320kbps

Τα factor scores για το MP3 στα 320kbps, φαίνονται στο Πίνακα 6.6.

Factors	Factor Scores
Evaluation	-0,2
Potency	-0,2
Activity	-0,47

Πίνακας 6.6: Factor scores για το είδος Rock στα 320kbps

6.2 Jazz

Σε αυτό το σημείο θα παρουσιαστούν τα αποτελέσματα για το είδος της jazz για κάθε bitrate ξεχωριστά, για τις κλίμακες και για τα factor scores.

6.2.1 Jazz με bitrate 96kbps

Στον Πίνακα 6.7, φαίνονται τα αποτελέσματα του πειράματος, για κάθε κλίμακα ξεχωριστά, για το είδος Jazz σε bitrate 96kbps.

Κλίμακες	N	Μέσος Όρος
Χειρότερο - Καλύτερο	30	-2,6
Πιο Φτωχό Φάσμα - Πιο Πλούσιο Φάσμα	30	-2
Περισσότερος Θόρυβος - Λιγότερος Θόρυβος	30	-0,6
Πιο Ενοχλητικό - Πιο Απολαυστικό	30	-2,6
Λιγότερο Δυνατό - Περισσότερο Δυνατό	30	-0,8
Πιο Δύσκολος Διαχωρισμός Οργάνων - Πιο Εύκολος Διαχωρισμός Οργάνων	30	-1,4
Λιγότερο Ζεστό - Περισσότερο Ζεστό	30	-2
Περισσότερη Παραμόρφωση - Λιγότερη Παραμόρφωση	30	-2,2
Λιγότερο Λαμπρό - Περισσότερο Λαμπρό	30	-1,2

Πίνακας 6.7: Αποτελέσματα για το είδος Jazz στα 96kbps

Τα factor scores που αντιστοιχούν στο είδος Jazz με bitrate 96kbps, φαίνονται στον Πίνακα 6.8.

Factors	Factor Scores
Evaluation	-2
Potency	-1,4
Activity	-1,73

Πίνακας 6.8: Factor scores για το είδος Jazz στα 96kbps

6.2.2 Jazz με bitrate 160kbps

Στον Πίνακα 6.9, παρουσιάζονται τα αποτελέσματα του πειράματος, για κάθε κλίμακα ξεχωριστά, για το είδος Jazz σε bitrate 160kbps.

Κλίμακες	N	Μέσος Όρος
Χειρότερο - Καλύτερο	30	-1,2
Πιο Φτωχό Φάσμα - Πιο Πλούσιο Φάσμα	30	-1,2
Περισσότερος Θόρυβος - Λιγότερος Θόρυβος	30	-1,2
Πιο Ενοχλητικό - Πιο Απολαυστικό	30	-1

Λιγότερο Δυνατό - Περισσότερο Δυνατό	30	-0,4
Πιο Δύσκολος Διαχωρισμός Οργάνων - Πιο Εύκολος Διαχωρισμός Οργάνων	30	-0,4
Λιγότερο Ζεστό - Περισσότερο Ζεστό	30	-1,4
Περισσότερη Παραμόρφωση - Λιγότερη Παραμόρφωση	30	-0,8
Λιγότερο Λαμπρό - Περισσότερο Λαμπρό	30	-1,2

Πίνακας 6.9: Αποτελέσματα για το είδος Jazz στα 160kbps

Τα factor scores που αντιστοιχούν στο είδος Jazz με bitrate 160kbps, φαίνονται στον Πίνακα 6.10.

Factors	Factor Scores
Evaluation	-0,87
Potency	-0,8
Activity	-1,27

Πίνακας 6.10: Factor scores για το είδος Jazz στα 160kbps

6.2.3 Jazz με bitrate 320kbps

Στον Πίνακα 6.11, παρουσιάζονται τα αποτελέσματα του πειράματος, για κάθε κλίμακα ξεχωριστά, για το είδος Jazz σε bitrate 320kbps.

Κλίμακες	N	Μέσος Όρος
Χειρότερο - Καλύτερο	30	-0,2
Πιο Φτωχό Φάσμα - Πιο Πλούσιο Φάσμα	30	-0,2
Περισσότερος Θόρυβος - Λιγότερος Θόρυβος	30	0
Πιο Ενοχλητικό - Πιο Απολαυστικό	30	-0,2
Λιγότερο Δυνατό - Περισσότερο Δυνατό	30	-0,2
Πιο Δύσκολος Διαχωρισμός Οργάνων - Πιο Εύκολος Διαχωρισμός Οργάνων	30	-0,4
Λιγότερο Ζεστό - Περισσότερο Ζεστό	30	-0,4
Περισσότερη Παραμόρφωση - Λιγότερη Παραμόρφωση	30	-0,2
Λιγότερο Λαμπρό - Περισσότερο Λαμπρό	30	0

Πίνακας 6.11: Αποτελέσματα για το είδος Jazz στα 320kbps

Τα factor scores που αντιστοιχούν στο είδος Jazz με bitrate 320kbps, φαίνονται στον Πίνακα 6.12.

Factors	Factor Scores
Evaluation	-0,2
Potency	-0,2
Activity	-0,2

Πίνακας 6.12: Factor scores για το είδος Jazz στα 320kbps

6.3 Ηλεκτρονική

Σε αυτό το σημείο θα παρουσιαστούν τα αποτελέσματα για το είδος της ηλεκτρονικής μουσικής για κάθε bitrate ξεχωριστά, για τις κλίμακες και για τα factor scores.

6.3.1 Ηλεκτρονική με bitrate 96kbps

Στον Πίνακα 6.13, παρουσιάζονται τα αποτελέσματα του πειράματος, για κάθε κλίμακα ξεχωριστά, για το είδος της Ηλεκτρονικής σε bitrate 96kbps.

Κλίμακες	N	Μέσος Όρος
Χειρότερο - Καλύτερο	30	-2,4
Πιο Φτωχό Φάσμα - Πιο Πλούσιο Φάσμα	30	-2
Περισσότερος Θόρυβος - Λιγότερος Θόρυβος	30	-1,4
Πιο Ενοχλητικό - Πιο Απολαυστικό	30	-1,8
Λιγότερο Δυνατό - Περισσότερο Δυνατό	30	-0,8
Πιο Δύσκολος Διαχωρισμός Οργάνων - Πιο Εύκολος Διαχωρισμός Οργάνων	30	-1
Λιγότερο Ζεστό - Περισσότερο Ζεστό	30	-1,6
Περισσότερη Παραμόρφωση - Λιγότερη Παραμόρφωση	30	-1,6
Λιγότερο Λαμπρό - Περισσότερο Λαμπρό	30	-1,6

Πίνακας 6.13: Αποτελέσματα για το είδος Ηλεκτρονική στα 96kbps

Τα factor scores που αντιστοιχούν στο είδος της Ηλεκτρονικής μουσικής με bitrate στα 96kbps, φαίνονται στον Πίνακα 6.14.

Factors	Factor Scores
Evaluation	-1,67
Potency	-1,33
Activity	-1,73

Πίνακας 6.14: Factor scores για το είδος Ηλεκτρονική στα 96kbps

6.3.2 Ηλεκτρονική με bitrate 160kbps

Στον Πίνακα 6.15, παρουσιάζονται τα αποτελέσματα του πειράματος, για κάθε κλίμακα ξεχωριστά, για το είδος της Ηλεκτρονικής σε bitrate 160kbps.

Κλίμακες	N	Μέσος Όρος
Χειρότερο - Καλύτερο	30	-1,4
Πιο Φτωχό Φάσμα - Πιο Πλούσιο Φάσμα	30	-1,2
Περισσότερος Θόρυβος - Λιγότερος Θόρυβος	30	-0,2
Πιο Ενοχλητικό - Πιο Απολαυστικό	30	-1,2
Λιγότερο Δυνατό - Περισσότερο Δυνατό	30	-0,8
Πιο Δύσκολος Διαχωρισμός Οργάνων - Πιο Εύκολος Διαχωρισμός Οργάνων	30	-1
Λιγότερο Ζεστό - Περισσότερο Ζεστό	30	-1,2
Περισσότερη Παραμόρφωση - Λιγότερη Παραμόρφωση	30	-0,8
Λιγότερο Λαμπρό - Περισσότερο Λαμπρό	30	-1

Πίνακας 6.15: Αποτελέσματα για το είδος Ηλεκτρονική στα 160kbps

Τα factor scores που αντιστοιχούν στο είδος της Ηλεκτρονικής μουσικής με bitrate στα 96kbps, φαίνονται στον Πίνακα 6.16.

Factors	Factor Scores
Evaluation	-1,13
Potency	-0,67
Activity	-1,13

Πίνακας 6.16: Factor scores για το είδος Ηλεκτρονική στα 160kbps

6.3.3 Ηλεκτρονική με bitrate 320kbps

Στον Πίνακα 6.17, παρουσιάζονται τα αποτελέσματα του πειράματος, για κάθε κλίμακα ξεχωριστά, για το είδος της Ηλεκτρονικής σε bitrate 320kbps.

Κλίμακες	N	Μέσος Όρος
Χειρότερο - Καλύτερο	30	-0,6
Πιο Φτωχό Φάσμα - Πιο Πλούσιο Φάσμα	30	-0,6
Περισσότερος Θόρυβος - Λιγότερος Θόρυβος	30	-0,4
Πιο Ενοχλητικό - Πιο Απολαυστικό	30	-0,2
Λιγότερο Δυνατό - Περισσότερο Δυνατό	30	-0,2
Πιο Δύσκολος Διαχωρισμός Οργάνων - Πιο Εύκολος Διαχωρισμός Οργάνων	30	-0,2
Λιγότερο Ζεστό - Περισσότερο Ζεστό	30	-0,6
Περισσότερη Παραμόρφωση - Λιγότερη Παραμόρφωση	30	-0,2

Λιγότερο Λαμπρό - Περισσότερο Λαμπρό	30	-0,4
--------------------------------------	----	------

Πίνακας 6.17: Αποτελέσματα για το είδος Ηλεκτρονική στα 320kbps

Τα factor scores που αντιστοιχούν στο είδος της Ηλεκτρονικής μουσικής με bitrate στα 320kbps, φαίνονται στον Πίνακα 6.18.

Factors	Factor Scores
Evaluation	-0,47
Potency	-0,27
Activity	-0,53

Πίνακας 6.18: Factor scores για το είδος Ηλεκτρονική στα 320kbps

6.4 Κλασσική Μουσική (με ορχήστρα)

Εδώ παρουσιάζονται τα αποτελέσματα για το είδος της κλασσικής μουσικής (με ορχήστρα) για κάθε bitrate ξεχωριστά, για τις κλίμακες και για τα factor scores.

6.4.1 Κλασσική με bitrate 96kbps

Στον Πίνακα 6.19, παρουσιάζονται τα αποτελέσματα του πειράματος, για κάθε κλίμακα ξεχωριστά, για το είδος της Κλασσικής σε bitrate 96kbps.

Κλίμακες	N	Μέσος Όρος
Χειρότερο - Καλύτερο	30	-2,4
Πιο Φτωχό Φάσμα - Πιο Πλούσιο Φάσμα	30	-1,8
Περισσότερος Θόρυβος - Λιγότερος Θόρυβος	30	-1,2
Πιο Ενοχλητικό - Πιο Απολαυστικό	30	-2
Λιγότερο Δυνατό - Περισσότερο Δυνατό	30	-1,4
Πιο Δύσκολος Διαχωρισμός Οργάνων - Πιο Εύκολος Διαχωρισμός Οργάνων	30	-1,2
Λιγότερο Ζεστό - Περισσότερο Ζεστό	30	-1,8
Περισσότερη Παραμόρφωση - Λιγότερη Παραμόρφωση	30	-0,8
Λιγότερο Λαμπρό - Περισσότερο Λαμπρό	30	-1,6

Πίνακας 6.19: Αποτελέσματα για το είδος Κλασσική στα 96kbps

Τα factor scores που αντιστοιχούν στο είδος της Κλασσικής μουσικής με bitrate στα 96kbps, φαίνονται στον Πίνακα 6.20.

Factors	Factor Scores
Evaluation	-1,93
Potency	-1,07

Activity	-1,73
----------	-------

Πίνακας 6.20: Factor scores για το είδος Κλασσική στα 96kbps

6.4.2 Κλασσική με bitrate 160kbps

Στον Πίνακα 6.21, παρουσιάζονται τα αποτελέσματα του πειράματος, για κάθε κλίμακα ξεχωριστά, για το είδος της Κλασσικής σε bitrate 160kbps.

Κλίμακες	N	Μέσος Όρος
Χειρότερο - Καλύτερο	30	-0,8
Πιο Φτωχό Φάσμα - Πιο Πλούσιο Φάσμα	30	-0,8
Περισσότερος Θόρυβος - Λιγότερος Θόρυβος	30	0
Πιο Ενοχλητικό - Πιο Απολαυστικό	30	-0,6
Λιγότερο Δυνατό - Περισσότερο Δυνατό	30	-0,6
Πιο Δύσκολος Διαχωρισμός Οργάνων - Πιο Εύκολος Διαχωρισμός Οργάνων	30	-0,6
Λιγότερο Ζεστό - Περισσότερο Ζεστό	30	-1
Περισσότερη Παραμόρφωση - Λιγότερη Παραμόρφωση	30	-0,2
Λιγότερο Λαμπρό - Περισσότερο Λαμπρό	30	-0,8

Πίνακας 6.21: Αποτελέσματα για το είδος Κλασσική στα 160kbps

Τα factor scores που αντιστοιχούν στο είδος της Κλασσικής μουσικής με bitrate στα 160kbps, φαίνονται στον Πίνακα 6.22.

Factors	Factor Scores
Evaluation	-0,67
Potency	-0,27
Activity	-0,87

Πίνακας 6.22: Factor scores για το είδος Κλασσική στα 160kbps

6.4.3 Κλασσική με bitrate 320kbps

Στον Πίνακα 6.23, παρουσιάζονται τα αποτελέσματα του πειράματος, για κάθε κλίμακα ξεχωριστά, για το είδος της Κλασσικής σε bitrate 320kbps.

Κλίμακες	N	Μέσος Όρος
Χειρότερο - Καλύτερο	30	-0,4
Πιο Φτωχό Φάσμα - Πιο Πλούσιο Φάσμα	30	-0,4
Περισσότερος Θόρυβος - Λιγότερος Θόρυβος	30	0
Πιο Ενοχλητικό - Πιο Απολαυστικό	30	-0,8
Λιγότερο Δυνατό - Περισσότερο Δυνατό	30	-0,4
Πιο Δύσκολος Διαχωρισμός Οργάνων - Πιο Εύκολος Διαχωρισμός Οργάνων	30	-0,2

Λιγότερο Ζεστό - Περισσότερο Ζεστό	30	-0,4
Περισσότερη Παραμόρφωση - Λιγότερη Παραμόρφωση	30	-0,2
Λιγότερο Λαμπρό - Περισσότερο Λαμπρό	30	-0,8

Πίνακας 6.23: Αποτελέσματα για το είδος Κλασσική στα 320kbps

Τα factor scores που αντιστοιχούν στο είδος της Κλασσικής μουσικής με bitrate στα 320kbps, φαίνονται στον Πίνακα 6.24.

Factors	Factor Scores
Evaluation	-0,53
Potency	-0,13
Activity	-0,53

Πίνακας 6.24: Factor scores για το είδος Κλασσική στα 320kbps

6.5 Solo όργανο

Εδώ παρουσιάζονται τα αποτελέσματα για το είδος του solo οργάνου για κάθε bitrate ξεχωριστά, για τις κλίμακες και για τα factor scores.

6.5.1 Solo όργανο με bitrate 96kbps

Στον Πίνακα 6.25, παρουσιάζονται τα αποτελέσματα του πειράματος, για κάθε κλίμακα ξεχωριστά, για το είδος του solo οργάνου σε bitrate 96kbps.

Κλίμακες	N	Μέσος Όρος
Χειρότερο - Καλύτερο	30	-0,8
Πιο Φτωχό Φάσμα - Πιο Πλούσιο Φάσμα	30	-0,6
Περισσότερος Θόρυβος - Λιγότερος Θόρυβος	30	-0,6
Πιο Ενοχλητικό - Πιο Απολαυστικό	30	-0,4
Λιγότερο Δυνατό - Περισσότερο Δυνατό	30	-0,6
Πιο Δύσκολος Διαχωρισμός Οργάνων - Πιο Εύκολος Διαχωρισμός Οργάνων	30	-0,6
Λιγότερο Ζεστό - Περισσότερο Ζεστό	30	-0,8
Περισσότερη Παραμόρφωση - Λιγότερη Παραμόρφωση	30	-0,6
Λιγότερο Λαμπρό - Περισσότερο Λαμπρό	30	-1,2

Πίνακας 6.25: Αποτελέσματα για το είδος solo όργανο στα 96kbps

Τα factor scores που αντιστοιχούν στο είδος solo όργανο με bitrate στα 96kbps, φαίνονται στον Πίνακα 6.26.

Factors	Factor Scores
Evaluation	-0,6
Potency	-0,6
Activity	-0,87

Πίνακας 6.26: Factor scores για το είδος solo όργανο στα 96kbps

6.5.2 Solo όργανο με bitrate 160kbps

Στον Πίνακα 6.27, παρουσιάζονται τα αποτελέσματα του πειράματος, για κάθε κλίμακα ξεχωριστά, για το είδος του solo οργάνου σε bitrate 160kbps.

Κλίμακες	N	Μέσος Όρος
Χειρότερο - Καλύτερο	30	-0,4
Πιο Φτωχό Φάσμα - Πιο Πλούσιο Φάσμα	30	-0,6
Περισσότερος Θόρυβος - Λιγότερος Θόρυβος	30	-0,4
Πιο Ενοχλητικό - Πιο Απολαυστικό	30	-0,2
Λιγότερο Δυνατό - Περισσότερο Δυνατό	30	-0,2
Πιο Δύσκολος Διαχωρισμός Οργάνων - Πιο Εύκολος Διαχωρισμός Οργάνων	30	-0,2
Λιγότερο Ζεστό - Περισσότερο Ζεστό	30	-0,6
Περισσότερη Παραμόρφωση - Λιγότερη Παραμόρφωση	30	0
Λιγότερο Λαμπρό - Περισσότερο Λαμπρό	30	-0,6

Πίνακας 6.27: Αποτελέσματα για το είδος solo όργανο στα 160kbps

Τα factor scores που αντιστοιχούν στο είδος του solo οργάνου με bitrate στα 160kbps, παρουσιάζονται στον Πίνακα 6.28.

Factors	Factor Scores
Evaluation	-0,27
Potency	-0,2
Activity	-0,6

Πίνακας 6.28: Factor scores για το είδος solo όργανο στα 160kbps

6.5.3 Solo όργανο με bitrate 320kbps

Στον Πίνακα 6.29, παρουσιάζονται τα αποτελέσματα του πειράματος, για κάθε κλίμακα ξεχωριστά, για το είδος solo όργανο σε bitrate 320kbps.

Κλίμακες	N	Μέσος Όρος
Χειρότερο – Καλύτερο	30	0,2
Πιο Φτωχό Φάσμα - Πιο Πλούσιο Φάσμα	30	0
Περισσότερος Θόρυβος - Λιγότερος Θόρυβος	30	0

Πιο Ενοχλητικό - Πιο Απολαυστικό	30	0
Λιγότερο Δυνατό - Περισσότερο Δυνατό	30	-0,4
Πιο Δύσκολος Διαχωρισμός Οργάνων - Πιο Εύκολος Διαχωρισμός Οργάνων	30	0
Λιγότερο Ζεστό - Περισσότερο Ζεστό	30	-0,4
Περισσότερη Παραμόρφωση - Λιγότερη Παραμόρφωση	30	0
Λιγότερο Λαμπρό - Περισσότερο Λαμπρό	30	0

Πίνακας 6.29: Αποτελέσματα για το είδος solo όργανο στα 320kbps

Τα factor scores που αντιστοιχούν στο είδος του solo οργάνου με bitrate στα 320kbps, φαίνονται στον Πίνακα 6.30.

Factors	Factor Scores
Evaluation	-0,07
Potency	0
Activity	-0,13

Πίνακας 6.30: Factor scores για το είδος solo όργανο στα 320kbps

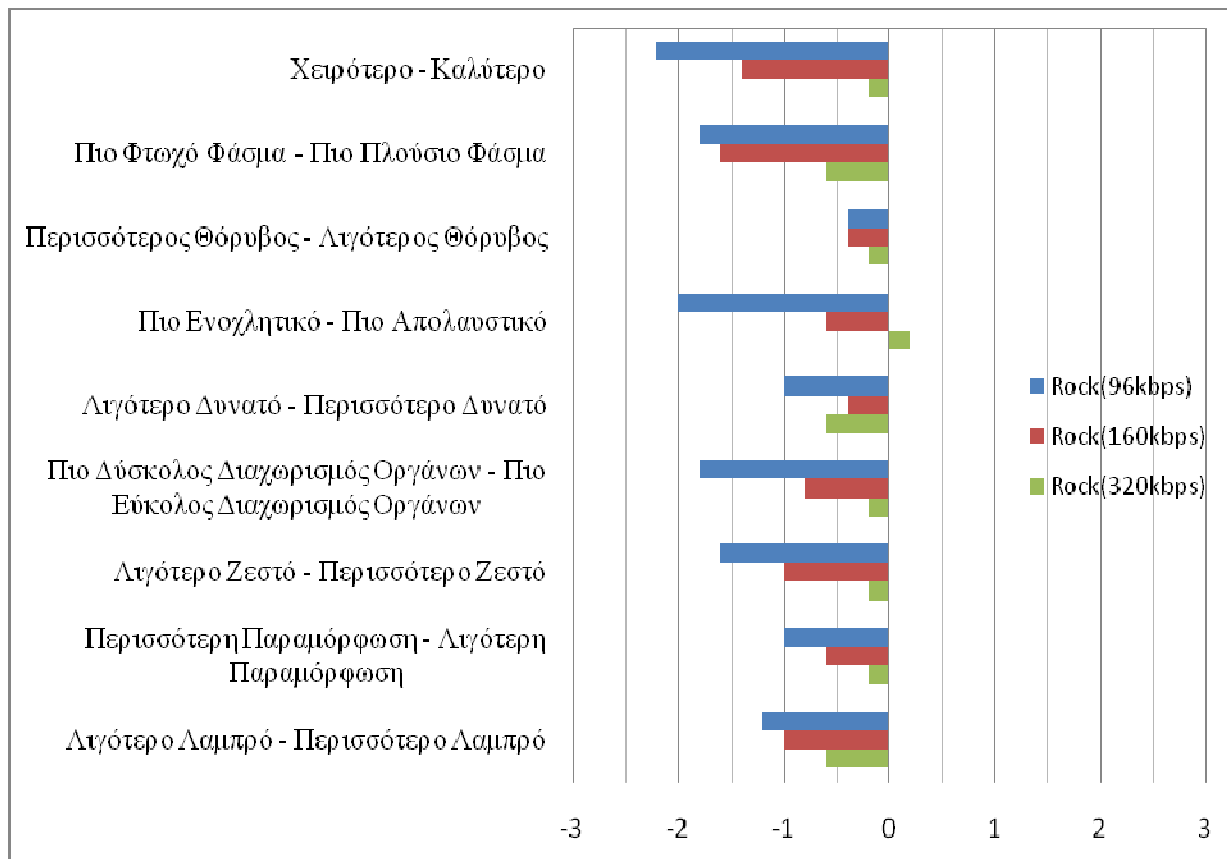
6.6 Συγκρίσεις Αποτελεσμάτων

Στο σημείο αυτό θα συγκριθούν τα αποτελέσματα που βρέθηκαν για κάθε είδος μουσικής και για κάθε bitrate ξεχωριστά, μέσω γραφημάτων. Μέσω των γραφικών αυτών παραστάσεων, γίνεται απόλυτα κατανοητή η ποιοτική απόδοση του αλγορίθμου ανάλογα με το είδος της μουσικής και το bitrate της κωδικοποίησης. Αρχικά θα συγκριθούν το κάθε είδος ανάλογα με το bitrate και στη συνέχεια, τα διαφορετικά είδη μουσικής μεταξύ τους για κάθε bitrate.

6.6.1 Rock

Στο Γράφημα 6.1, συγκρίνονται τα τρία διαφορετικά bitrates για το είδος μουσικής Rock, σε επίπεδο κλιμάκων. Όπως φαίνεται καθαρά, στα 96kbps, η απόδοση του MP3 δεν είναι καθόλου καλή, με τις διαφορές να είναι ιδιαίτερα εμφανείς σε όλες τις κλίμακες. Στο bitrate των 160kbps, η βελτίωση είναι ορατή, σε όλες τις κλίμακες, χωρίς όμως η ποιότητα της συμπίεσης να πλησιάζει αρκετά την ποιότητα του πρωτοτύπου WAV. Τέλος, στο bitrate των 320kbps, οι διαφορές ανάμεσα στο συμπιεσμένο και ασυμπιεστο αρχείο είναι πολύ μικρές, με τη βαθμολογία των κλιμάκων να μην πέφτει κάτω από το -

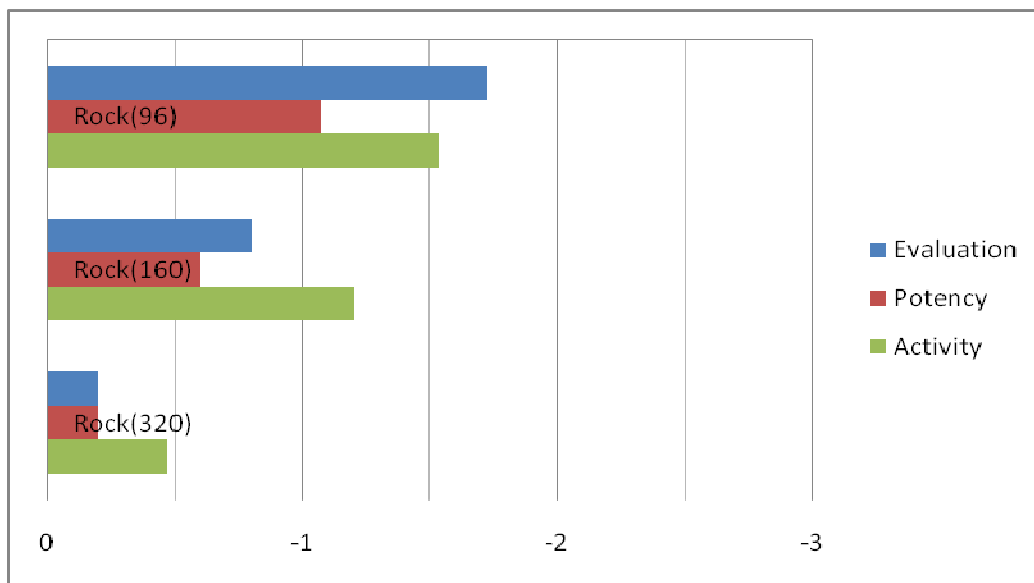
0,6. Έτσι, γίνεται ξεκάθαρο πως στο μεγαλύτερο bitrate, οι διαφορές είναι δυσδιάκριτες και οι ακροατές δεν είναι σε θέση να τις εντοπίσουν εύκολα.



Γράφημα 6.1: Σύγκριση των διαφορετικών bitrates για το μουσικό είδος rock

Στο Γράφημα 6.2, γίνεται και πάλι η σύγκριση των bitrates για τη rock, σε επίπεδο factor scores. Όπως φαίνεται, στα 96kbps, παρουσιάζεται αρκετά μεγάλο πρόβλημα και στους τρεις factors, με μεγαλύτερο αυτό του παράγοντα Evaluation. Ο factor αυτός αντιπροσωπεύει κυρίως τη γενική εικόνα και αισθητική του κομματιού. Το μικρότερο πρόβλημα φαίνεται να υπάρχει στον παράγοντα Potency, χωρίς όμως αυτό να σημαίνει ότι δεν εντοπίζονται διαφορές. Στα 160kbps, η βελτίωση φαίνεται από το γεγονός ότι στον παράγοντα Evaluation, το πρόβλημα έχει μειωθεί σε σχεδόν στο μισό. Στον παράγοντα Potency, η μείωση του προβλήματος είναι και πάλι σχεδόν στο μισό, ενώ στον παράγοντα Activity, υπάρχει μείωση κατά 1/3. Τέλος, στα 320kbps, και οι τρεις παράγοντες βρίσκονται κάτω από το -0,5, δείχνοντας έτσι ότι οι διαφορές ανάμεσα στο

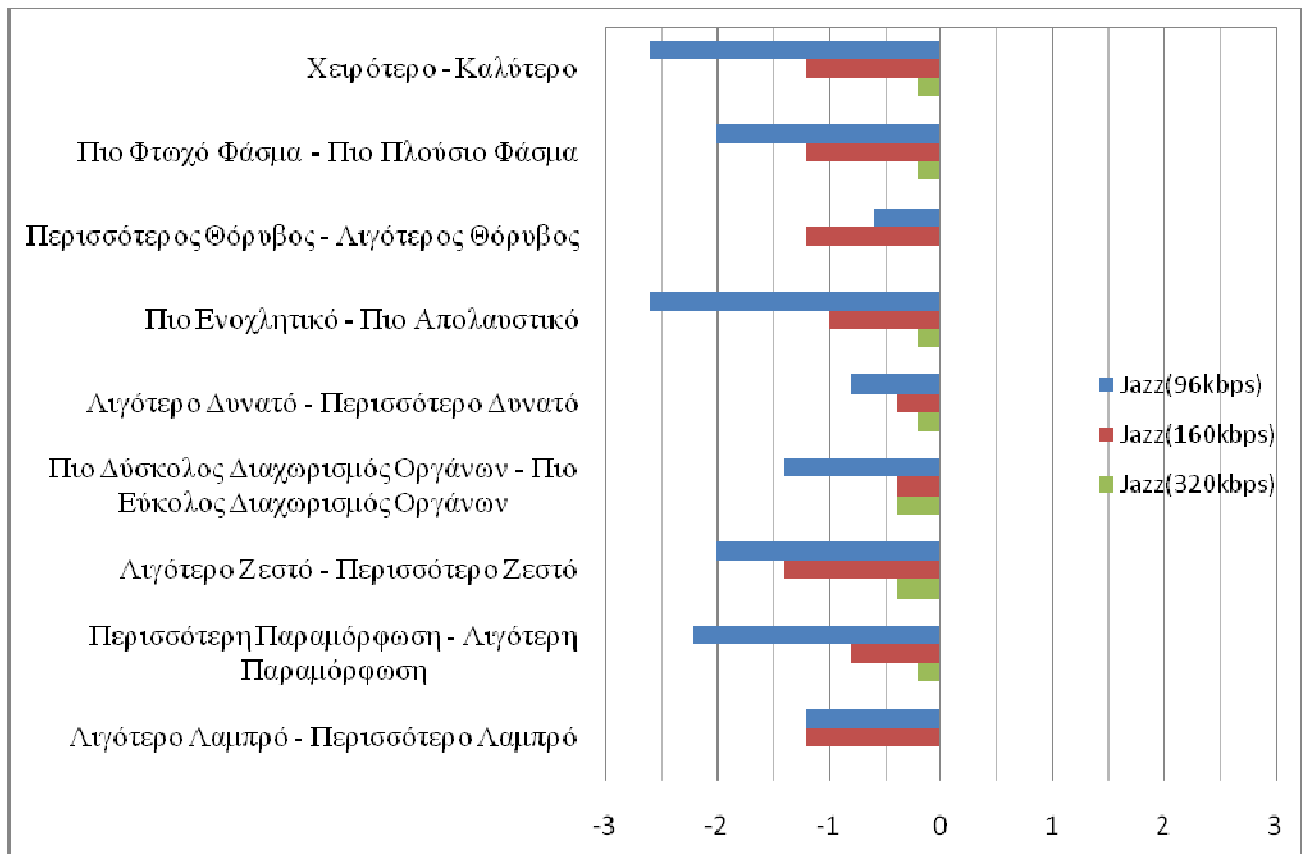
αρχείο MP3 και το πρωτότυπο αρχείο, είναι δυσδιάκριτες σε επίπεδο φάσματος και σχεδόν ανύπαρκτες σε επίπεδο δυναμικής και συνολικής αισθητικής.



Γράφημα 6.2: Σύγκριση των factor scores για διαφορετικά bitrates στο είδος rock

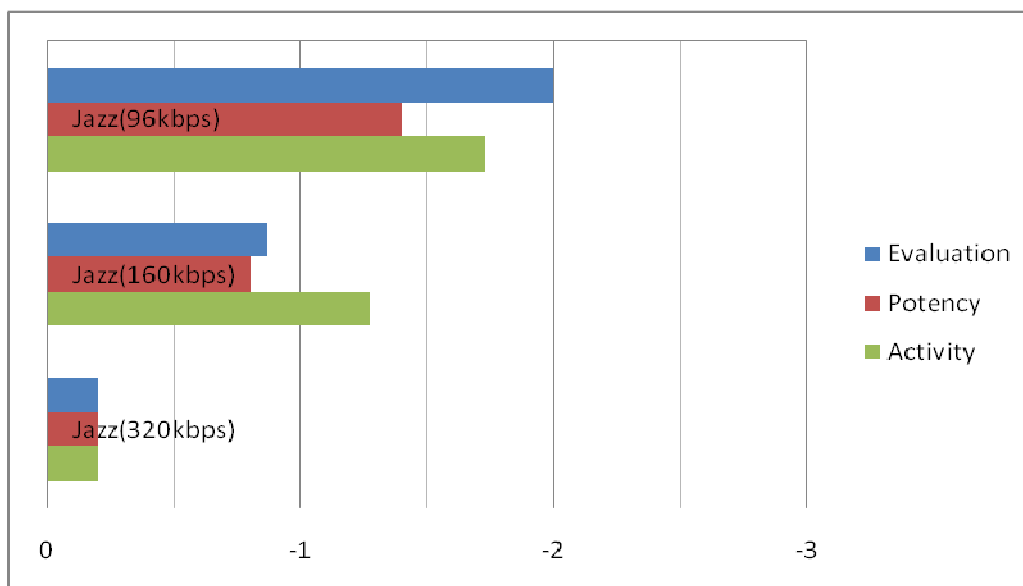
6.6.2 Jazz

Η σύγκριση των bitrates για τη Jazz παρατίθεται στο Γράφημα 6.3, σε επίπεδο ξεχωριστών κλιμάκων. Στο bitrate των 96kbps, η ποιότητα του αλγορίθμου είναι πάρα πολύ χαμηλή, με ορισμένες κλίμακες να φτάνουν σε βαθμολογία μέχρι και στο -2,6. Περνώντας στα 160kbps, η κατάσταση βελτιώνεται αισθητά, μιας και οι βαθμολογίες στις κλίμακες δεν πέφτουν κάτω από το -1,4. Παρ' όλα αυτά, ούτε και στο συγκεκριμένο bitrate αποδίδει αρκετά καλά ο αλγόριθμος, με περισσότερες από τις μισές κλίμακες να έχουν βαθμολογίες μικρότερες του -1. Στα 320kbps, η ποιότητα του αλγορίθμου έχει βελτιωθεί πάρα πολύ, με τις βαθμολογίες να κυμαίνονται από το -0,4 έως το 0, δείχνοντας έτσι πως σε αυτό το bitrate οι διαφορές από το πρωτότυπο κομμάτι είναι μηδαμινές.



Γράφημα 6.3: Σύγκριση των διαφορετικών bitrates για το μουσικό είδος jazz

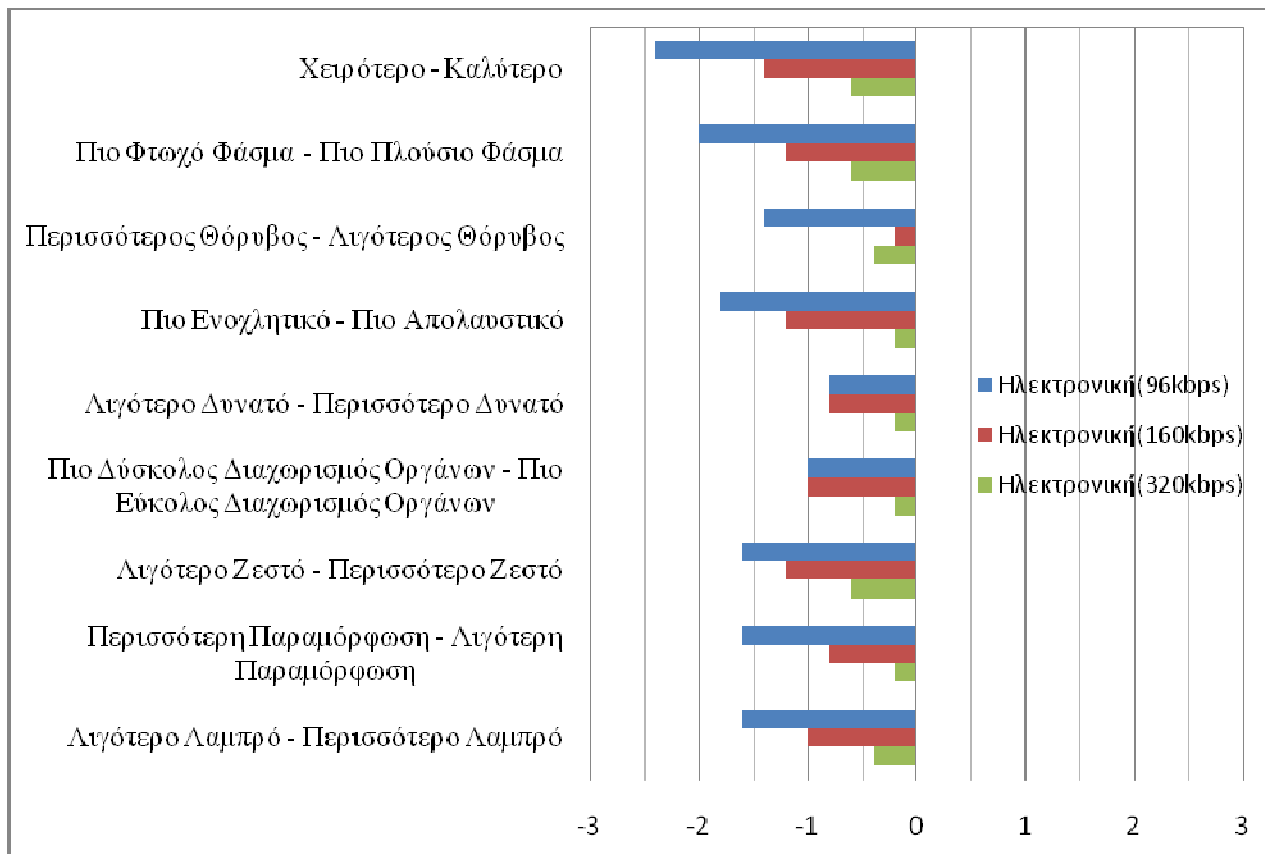
Προχωρώντας στο Γράφημα 6.4, γίνεται και πάλι η σύγκριση των bitrates για τη jazz, σε επίπεδο factor scores. Στα 96kbps, το πρόβλημα είναι ολοφάνερο, και στους τρεις factors, με μεγαλύτερο αυτό του παράγοντα Evaluation και μικρότερο αυτό του παράγοντα Potency. Κοιτάζοντας τα factor scores για το bitrate των 160kbps, υπάρχει μια μεγάλη βελτίωση με το score του παράγοντα Evaluation να έχει βελτιωθεί κατά 1,13 μονάδες, το score του παράγοντα Potency κατά 0,6 και του παράγοντα Activity κατά 0,46. Αν και οι βελτιώσεις στους παράγοντες είναι σχετικά μεγάλες, ιδίως αυτή του παράγοντα Evaluation, οι διαφορές ανάμεσα στο συμπίεσμένο και το πρωτότυπο αρχείο εξακολουθούν να είναι υπαρκτές. Τέλος, στο bitrate των 320kbps, υπάρχει μια μεγάλη βελτίωση σε σχέση με τα προηγούμενα bitrates. Τα scores έχουν φτάσει στο -0,2 και στους τρεις παράγοντες, δείχνοντας έτσι ότι η διαφορά του MP3 σε αυτό το bitrate είναι σχεδόν ανύπαρκτη σε σχέση με το πρωτότυπο.



Γράφημα 6.4: Σύγκριση των factor scores για διαφορετικά bitrates στο είδος jazz

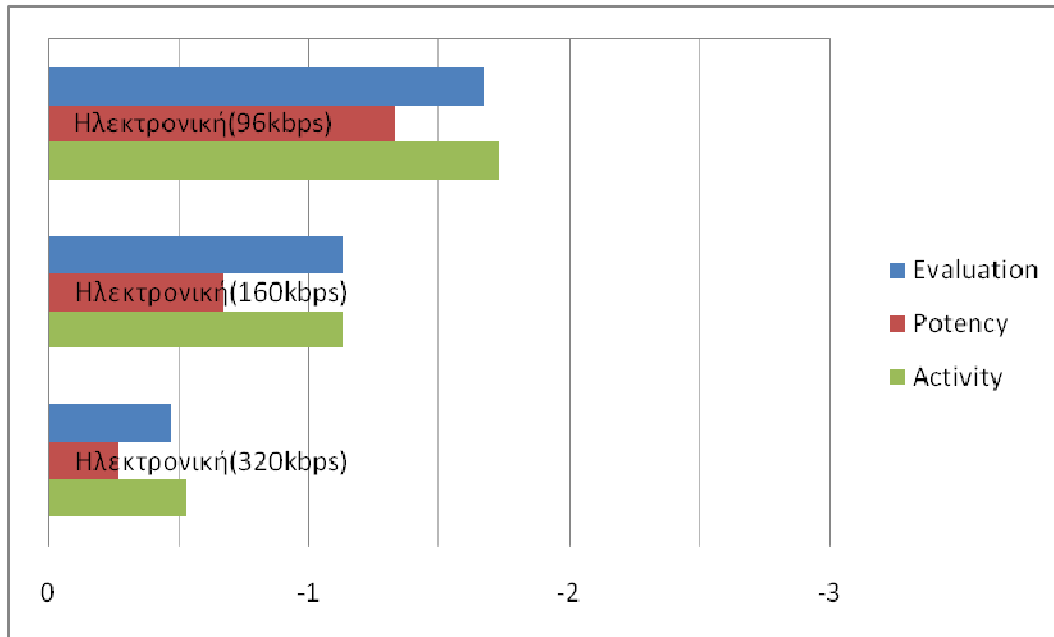
6.6.3 Ηλεκτρονική μουσική

Στο Γράφημα 6.5 αναπαρίσταται η σύγκριση των bitrates για την ηλεκτρονική μουσική, με βάση τις κλίμακες. Όπως φαίνεται, το MP3 με bitrate στα 96kbps, έχει αρκετά κακή απόδοση, με τη χειρότερη κλίμακα να βαθμολογείται με -2,4(Χειρότερο – Καλύτερο) και την καλύτερη με -0,8 (Λιγότερο Δυνατό – Περισσότερο Δυνατό). Στο bitrate των 160kbps, η βελτίωση που έχει επέλθει στη βαθμολογία των κλιμάκων είναι ορατή και μεγάλη, χωρίς όμως να μπορεί κανείς να πει ότι ο αλγόριθμος πλησιάζει αρκετά σε ποιότητα το πρωτότυπο WAV. Τέλος, στο bitrate των 320kbps, οι βαθμολογίες βελτιώνονται ακόμα περισσότερο, με τη χαμηλότερη να βρίσκεται στο -0,6 και την υψηλότερη στο -0,2, δείχνοντας έτσι τη δυσκολία με την οποία εντοπίζει κανείς διαφορές ανάμεσα στον αλγόριθμο στο συγκεκριμένο bitrate και το ασυμπίεστο σήμα.



Γράφημα 6.5: Σύγκριση των διαφορετικών bitrates για το μουσικό είδος ηλεκτρονική

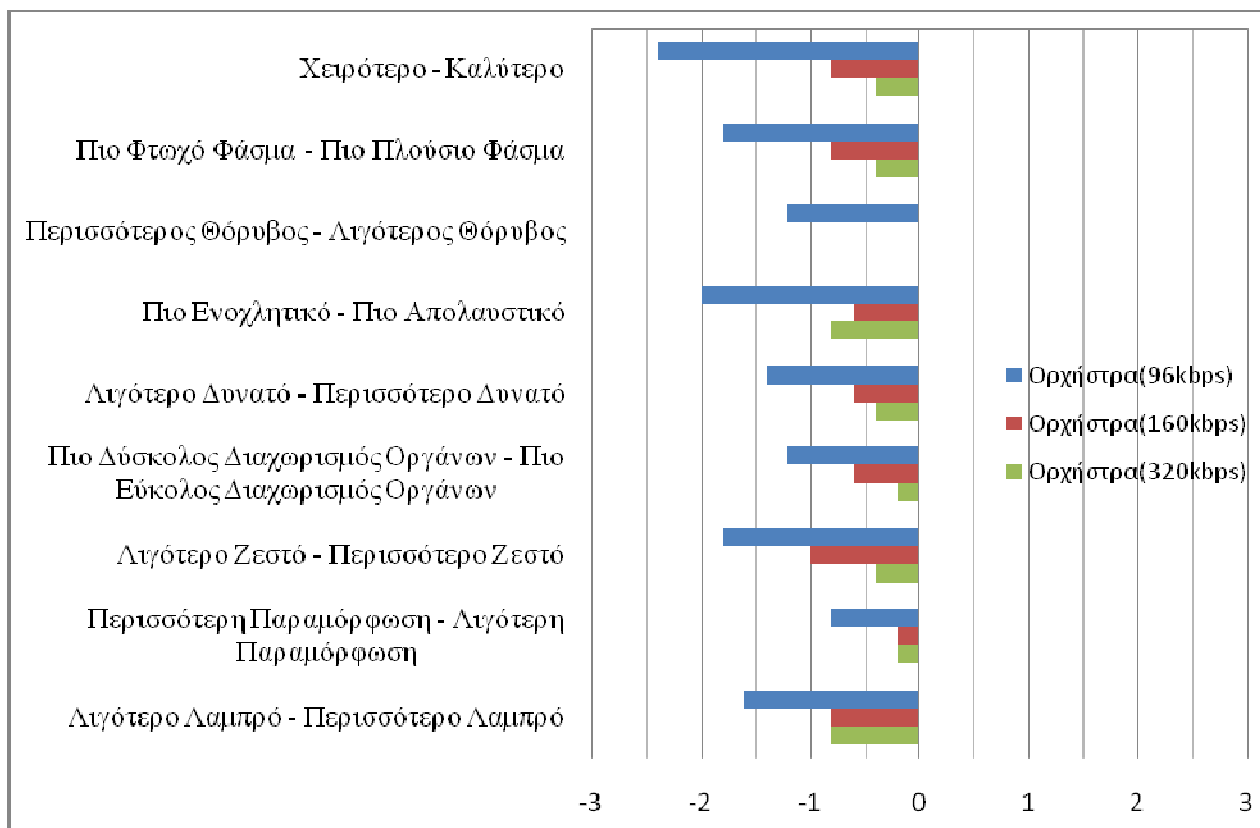
Περνώντας στο Γράφημα 6.6, φαίνεται η σύγκριση των factor scores για τα τρία διαφορετικά bitrates. Στα 96kbps, το μεγαλύτερο πρόβλημα εμφανίζεται στον παράγοντα Activity ενώ το μικρότερο στον παράγοντα Potency. Οι βαθμολογίες, όμως, και στους τρεις παράγοντες είναι αρκετά χαμηλές, πέφτοντας κάτω από το -1,3 και στις τρεις περιπτώσεις. Στο bitrate των 160kbps, η βελτίωση των παραγόντων αγγίζει το 66% στον παράγοντα Potency, το 53% στον παράγοντα Evaluation και το 60% στον παράγοντα Activity, με τις βαθμολογίες όμως και των τριών να βρίσκονται σε αρκετά χαμηλά επίπεδα. Κοιτώντας τα factor scores του bitrate των 320kbps, οι βαθμολογίες έχουν ανέβει πάρα πολύ, με μεγαλύτερη αυτή του παράγοντα Potency στο -0,27 και μικρότερη αυτή του παράγοντα Activity στο -0,53.



Γράφημα 6.6: Σύγκριση των factor scores για διαφορετικά bitrates στο είδος ηλεκτρονική

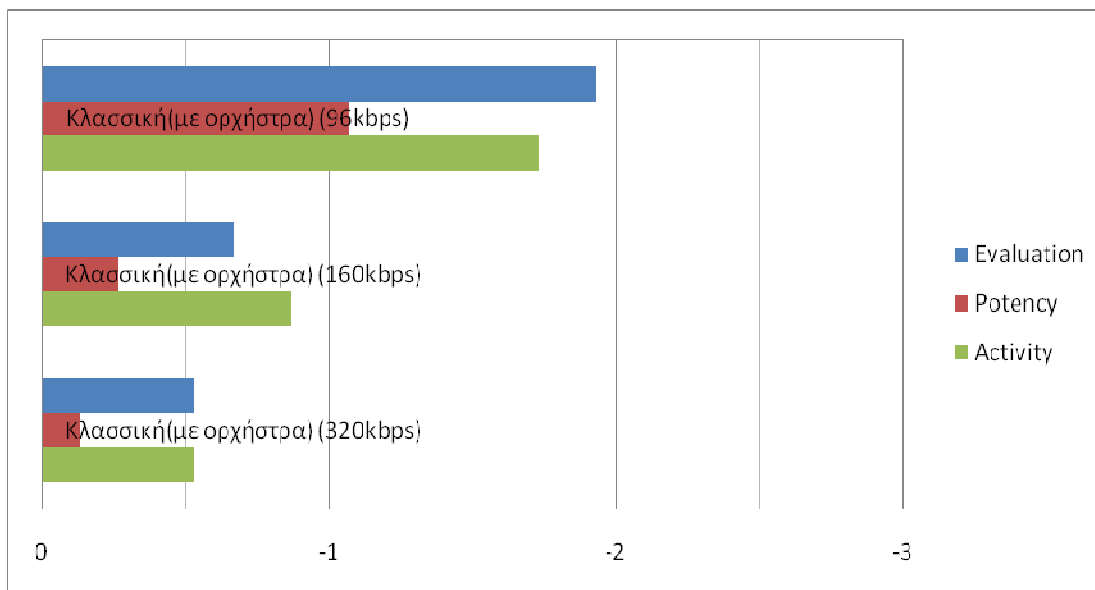
6.6.4 Κλασσική (με ορχήστρα)

Συνεχίζοντας με την κλασσική μουσική, στο Γράφημα 6.7, φαίνεται η σύγκριση ανάμεσα στα διαφορετικά bitrates, σε επίπεδο κλιμάκων. Κοιτάζοντας το γράφημα, γίνεται αμέσως αντιληπτό ότι στα 96kbps το MP3 παρουσιάζει ιδιαίτερα χαμηλές βαθμολογίες, οι οποίες κυμαίνονται από το -2,4 έως το -0,8, δείχνοντας έτσι την πολύ κακή ποιότητα ήχου στο συγκεκριμένο bitrate. Περνώντας στα 160kbps φαίνεται μια μεγάλη βελτίωση, με τις βαθμολογίες, όμως, να βρίσκονται ακόμα σε αρκετά χαμηλό επίπεδο. Στα 320kbps, η ποιότητα φαίνεται να βελτιώνεται περαιτέρω, με κάποιες, όμως, από τις κλίμακες να εξακολουθούν να έχουν αρκετά χαμηλές βαθμολογίες.



Γράφημα 6.7: Σύγκριση των διαφορετικών bitrates για το μουσικό είδος κλασσική (με ορχήστρα)

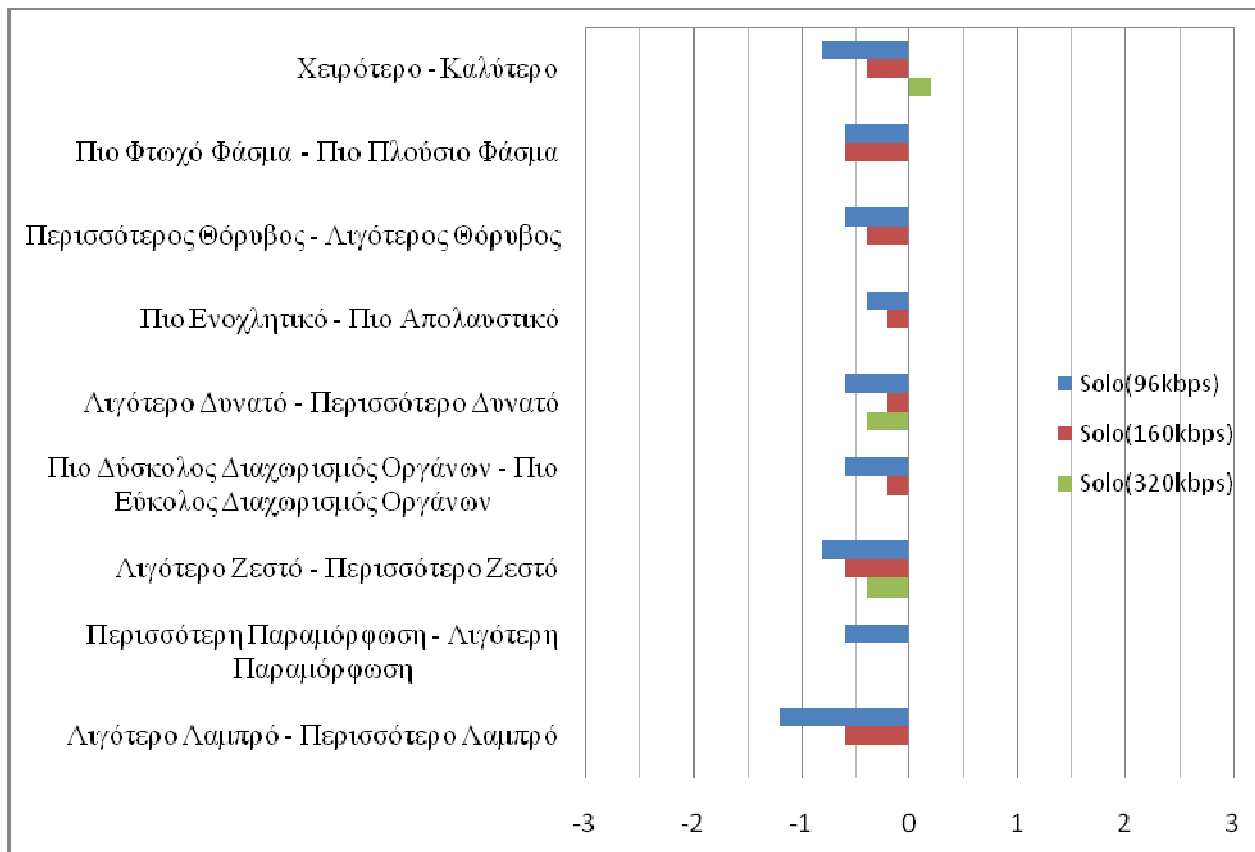
Περνώντας στα factor scores, όπως αυτά φαίνονται στο Γράφημα 6.8, στα 96kbps ο παράγοντας Evaluation είναι αυτός με το μεγαλύτερο πρόβλημα, με τον παράγοντα Activity να ακολουθεί σχεδόν οριακά. Οι τιμές και των δύο δείχνουν ότι το MP3 στο bitrate αυτό, παρουσιάζει πολύ μεγάλες αποκλίσεις από το πρωτότυπο, καθιστώντας το ακατάλληλο για το συγκεκριμένο είδος μουσικής. Στον παράγοντα Potency, υπάρχει σαφώς πρόβλημα, όχι όμως τόσο μεγάλο δύο προηγούμενους, αλλά ούτε και μικρό. Στα 160kbps, υπάρχει σαφέστατη βελτίωση στην ποιότητα του αλγορίθμου, χωρίς όμως αυτή να πλησιάζει αρκετά στα επίπεδα του ασυμπιεστού σήματος. Τέλος, στο bitrate των 320kbps, δεν υπάρχουν πολύ μεγάλες διαφορές ανάμεσα στα factor scores των 320kbps και των 160kbps, με τους παράγοντες Evaluation και Activity να βρίσκονται κάτω από το -0,5 ενώ ο παράγοντας Potency βρίσκεται -0,13.



Γράφημα 6.8: Σύγκριση των factor scores για διαφορετικά bitrates στο είδος κλασσική (με ορχήστρα)

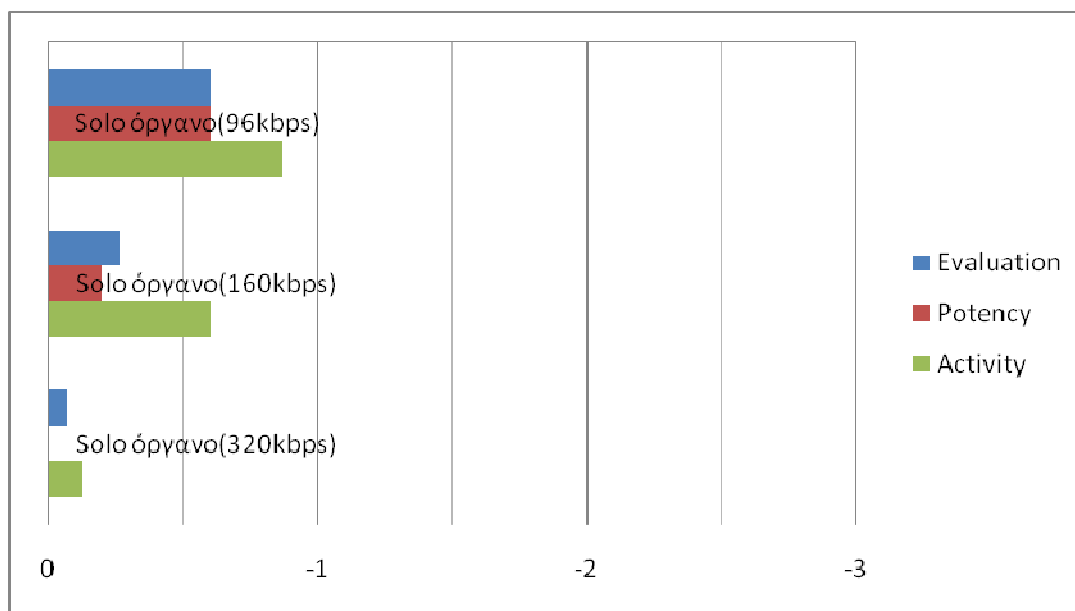
6.6.5 Solo όργανο

Φτάνοντας στο τελευταίο είδος, αυτό του solo οργάνου, η απόδοση του MP3 σε επίπεδο κλιμάκων φαίνεται να μην έχει τόσο μεγάλα προβλήματα όσο τα προηγούμενα είδη μουσικής, όπως φαίνεται και στο Γράφημα 6.9. Στα 96kbps, οι βαθμολογίες κυμαίνονται από το -1,2 έως το -0,4, με πιο προβληματική την κλίμακα των υψηλών συχνοτήτων (Λιγότερο Λαμπρό – Περισσότερο Λαμπρό). Στο bitrate των 160kbps, η ποιότητα βελτιώνεται με τη μικρότερη βαθμολογία να βρίσκεται στο -0,6 και τη μεγαλύτερη στο 0. Στα 320kbps, οι διαφορές ανάμεσα στο συμπιεσμένο και το ασυμπίεστο γίνονται ιδιαίτερα δυσδιάκριτες, με τις βαθμολογίες των κλιμάκων να κυμαίνονται ανάμεσα στο -0,4 και το 0,2.



Γράφημα 6.9: Σύγκριση των διαφορετικών bitrates για το μουσικό είδος solo όργανο

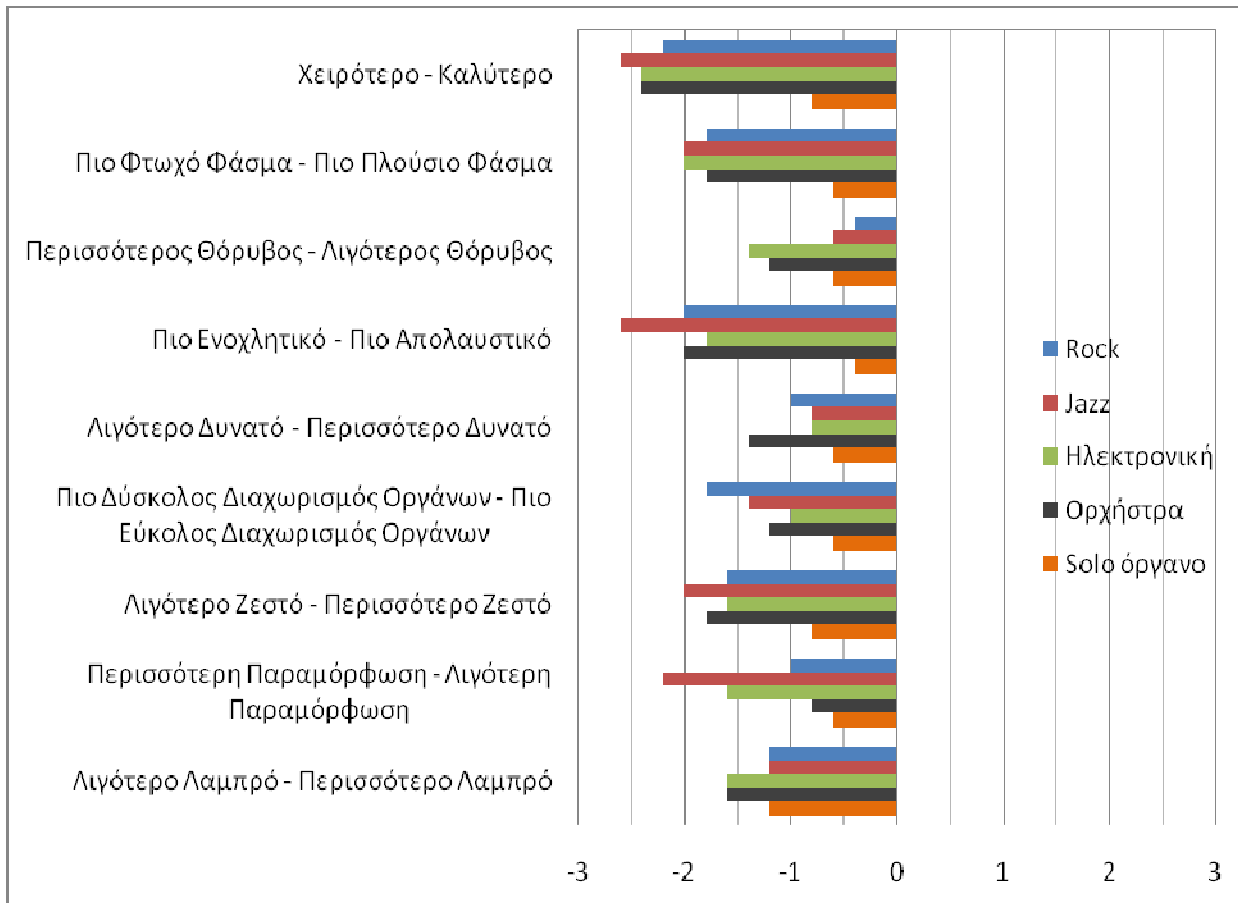
Σε επίπεδο factor scores, οι βαθμολογίες που συγκεντρώνει το συγκεκριμένο είδος είναι σχετικά μεγάλες, όπως φαίνεται και στο Γράφημα 6.10. Στα 96kbps, το πρόβλημα εστιάζεται κυρίως στον παράγοντα Activity, με score στο -0,87, ενώ ακολουθούν οι παράγοντες Evaluation και Potency, με βαθμολογίες -0,6. Περνώντας στο bitrate των 160kbps, το χειρότερο score εντοπίζεται και πάλι στον παράγοντα Activity, στο -0,6, ενώ στους δύο άλλους παράγοντες οι βαθμολογίες δεν πέφτουν χαμηλότερα από το -0,27. Τέλος, στα 320kbps, φαίνεται η βελτίωση σε όλους τους παράγοντες, καθιστώντας έτσι το πρόβλημα σχεδόν ανύπαρκτο, με scores που δεν πέφτουν κάτω από το -0,13.



Γράφημα 6.10: Σύγκριση των factor scores για διαφορετικά bitrates στο είδος solo όργανο

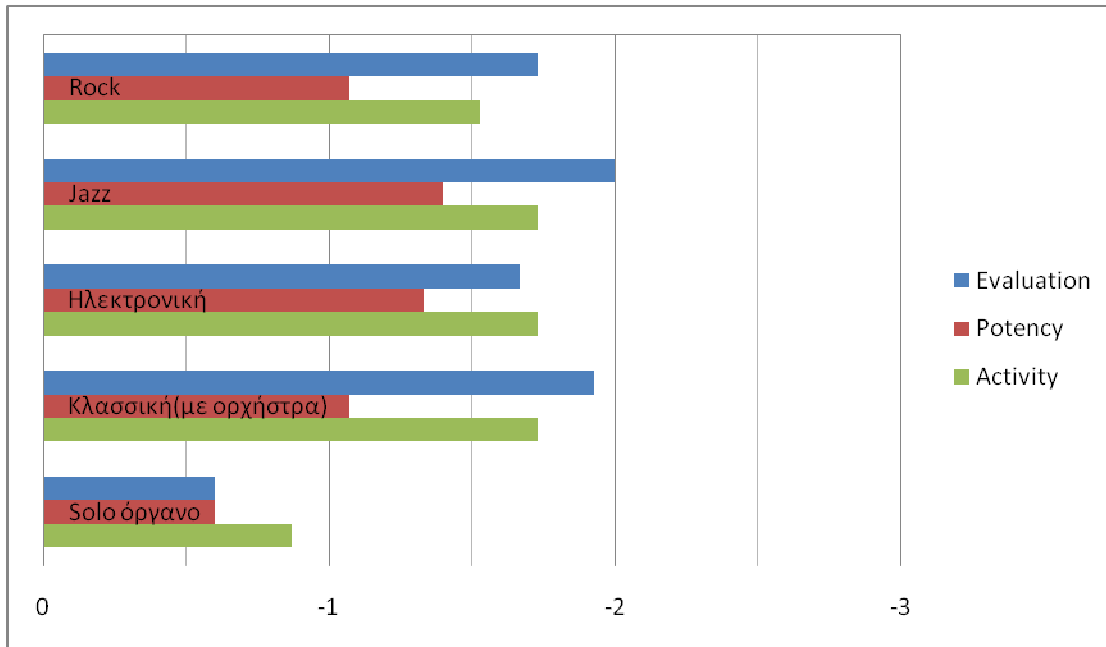
6.6.6 Bitrate στα 96kbps

Σε αυτό το σημείο ξεκινά η σύγκριση των αποτελεσμάτων ανά bitrate, συγκρίνοντας μεταξύ τους τα διαφορετικά μουσικά είδη. Στο Γράφημα 6.11 φαίνεται η σύγκριση των ειδών, σε επίπεδο κλιμάκων, για bitrate στα 96kbps. Από το γράφημα, προκύπτει ότι στα 96kbps, το solo όργανο είναι αυτό με την καλύτερη βαθμολογία ενώ jazz έχει τη χειρότερη βαθμολογία. Τα υπόλοιπα είδη μουσικής βρίσκονται κάπου στη μέση με τη rock να έχει καλύτερη βαθμολογία σε σχέση με την ηλεκτρονική και την κλασική μουσική με ορχήστρα.



Γράφημα 6.11: Σύγκριση των κλιμάκων για διαφορετικά είδη μουσικής με bitrate 96kbps

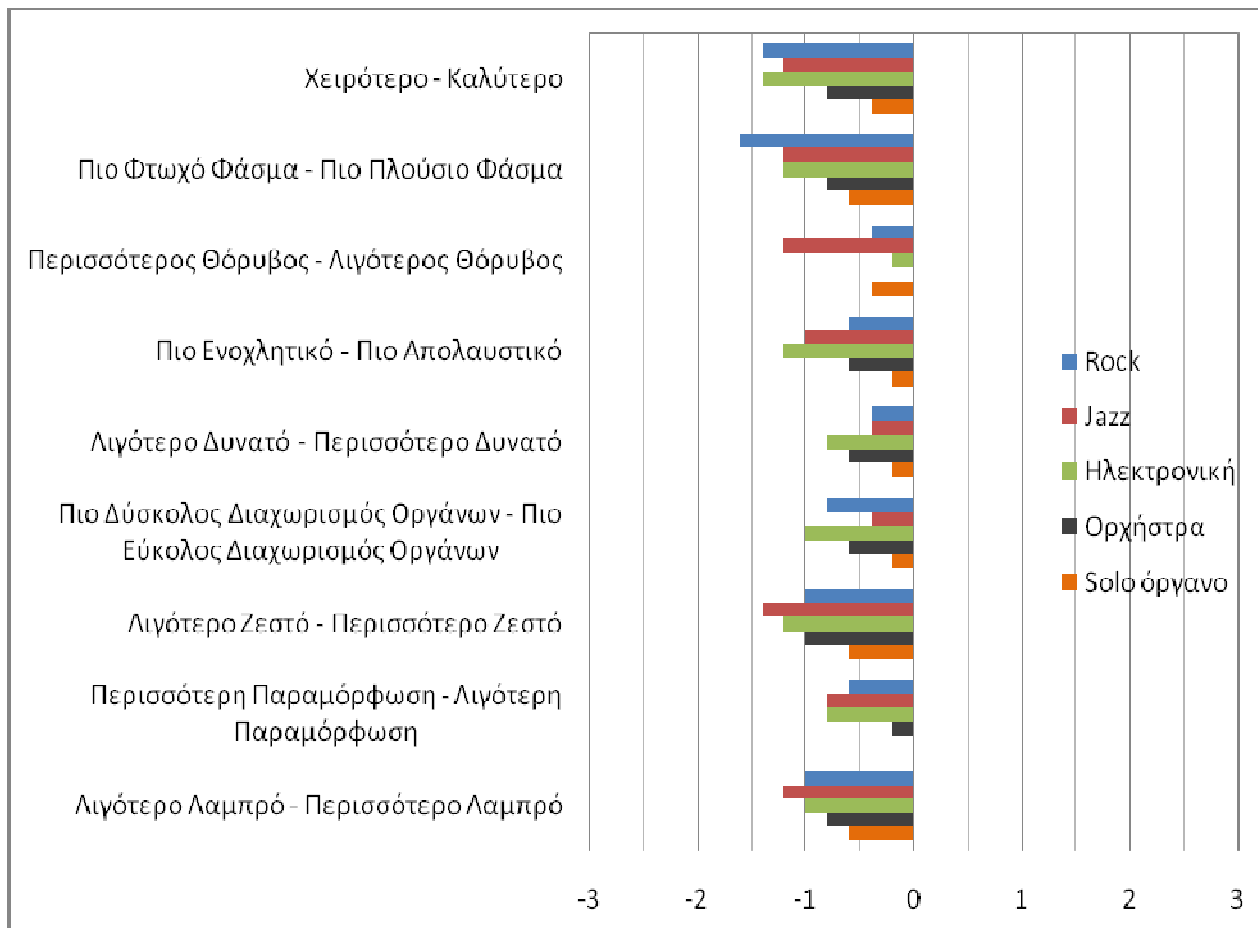
Σε επίπεδο factor scores, τα είδη συγκρίνονται στο Γράφημα 6.12. Εδώ πλέον γίνεται ξεκάθαρη η κατάσταση του κάθε είδους. Η jazz μουσική είναι αυτή που έχει τις χαμηλότερες βαθμολογίες ενώ το solo όργανο έχει τις υψηλότερες. Οι διαφορές είναι πολύ μεγάλες, ενώ ανάμεσα στα δύο αυτά είδη, από την καλύτερη στη χειρότερη βαθμολογία, βρίσκονται η rock, η ηλεκτρονική και η κλασική. Στο σημείο αυτό, παρατηρείται ότι ο παράγοντας Evaluation είναι αυτός που συγκεντρώνει τις χαμηλότερες βαθμολογίες στα περισσότερα μουσικά είδη, ενώ ο παράγοντας Potency συγκεντρώνει τις υψηλότερες.



Γράφημα 6.12: Σύγκριση των factor scores για διαφορετικά είδη μουσικής με bitrate 96kbps

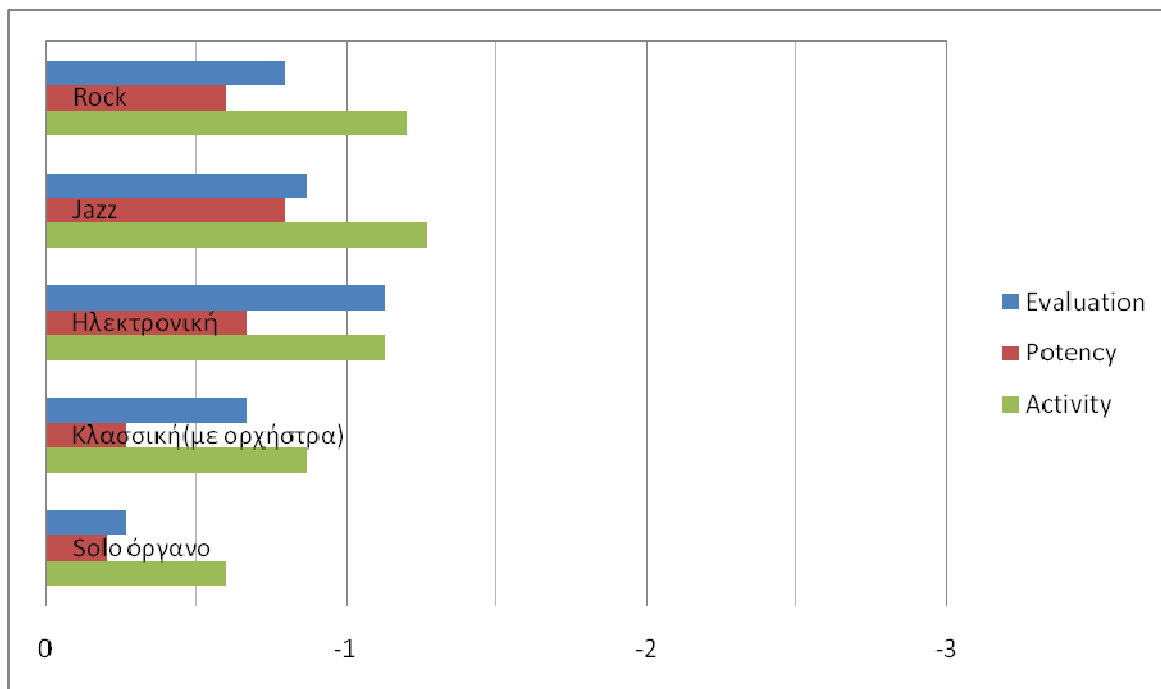
6.6.7 Bitrate στα 160kbps

Στα 160kbps, οι διαφορές σε επίπεδο κλιμάκων ανάμεσα στα μουσικά είδη φαίνονται και στο Γράφημα 6.13. Το solo όργανο και πάλι ξεχωρίζει συγκεντρώνοντας τις υψηλότερες βαθμολογίες, ενώ η jazz και η ηλεκτρονική μουσική συγκεντρώνουν τις χαμηλότερες, με τη rock και την κλασσική μουσική να ακολουθούν.



Γράφημα 6.13: Σύγκριση των κλιμάκων για διαφορετικά είδη μουσικής με bitrate 160kbps

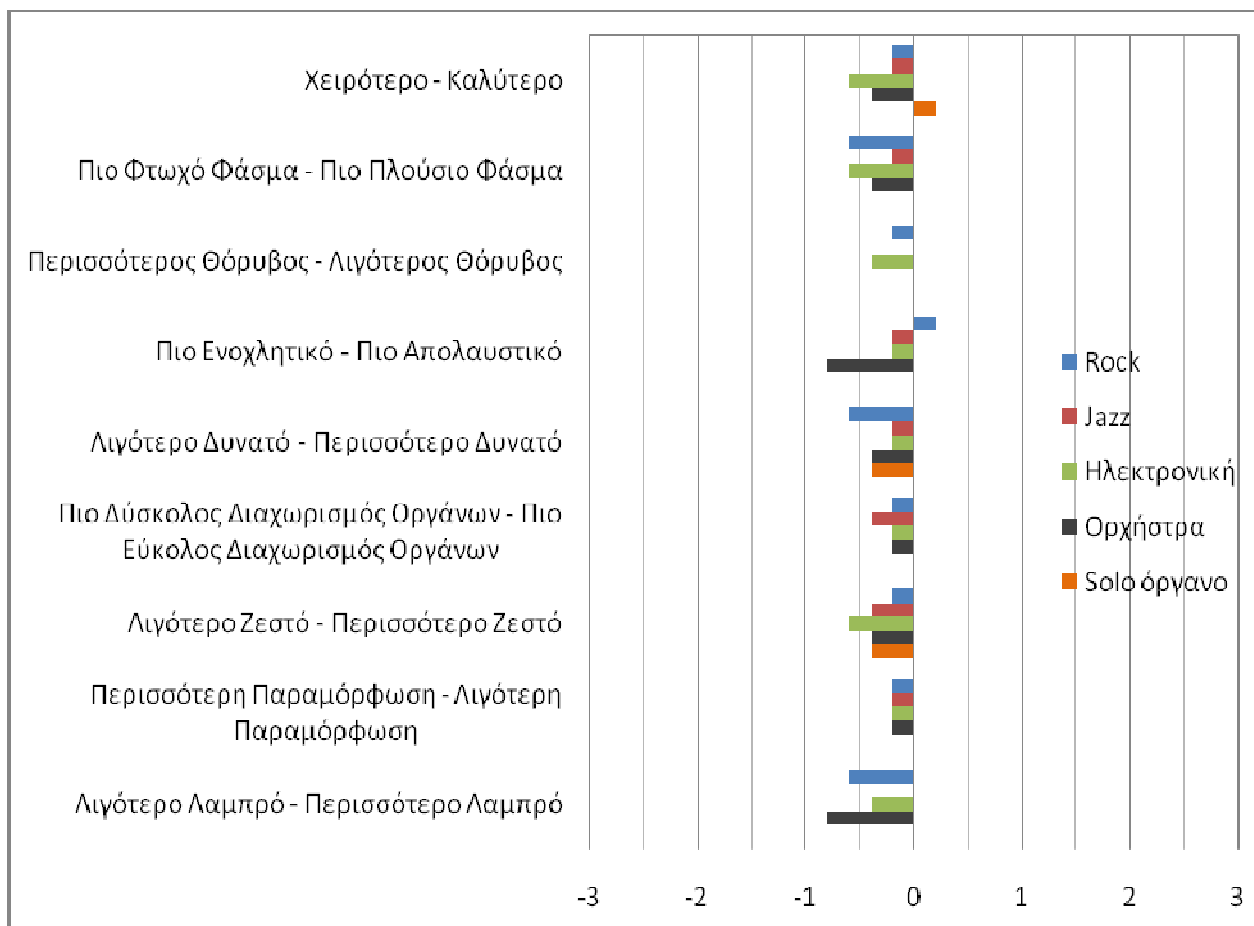
Στα factor scores, τα οποία φαίνονται στο Γράφημα 6.14, η σειρά γίνεται ακόμα πιο ξεκάθαρη με τον παράγοντα Activity να συγκεντρώνει τις χαμηλότερες βαθμολογίες σε όλα τα μουσικά είδη και τον παράγοντα Potency τις υψηλότερες. Παρ' όλες, τις βελτιώσεις που έχουν επέλθει σε σχέση με το προηγούμενο bitrate, τα προβλήματα εξακολουθούν να υπάρχουν με μεγαλύτερα αυτά στο είδος της jazz και της ηλεκτρονικής μουσικής.



Γράφημα 6.14: Σύγκριση των factor scores για διαφορετικά είδη μουσικής με bitrate 160kbps

6.6.8 Bitrate στα 320kbps

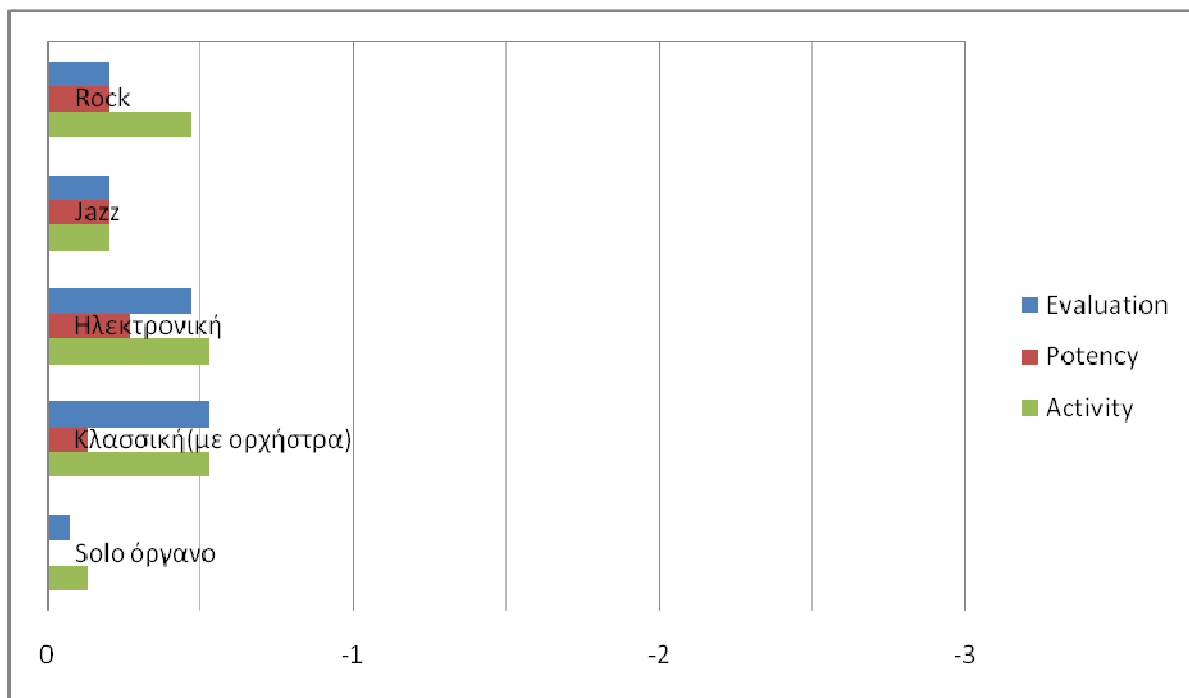
Στο τελευταίο και μεγαλύτερο bitrate, οι συγκρίσεις των κλιμάκων φαίνονται στο Γράφημα 6.15. Εδώ, η κατάταξη αλλάζει, με την κλασσική μουσική να συγκεντρώνει τις χαμηλότερες βαθμολογίες ακολουθούμενη από την ηλεκτρονική και τη rock. Το solo όργανο έχει και πάλι τις μεγαλύτερες βαθμολογίες, με τη jazz να έρχεται δεύτερη.



Γράφημα 6.15: Σύγκριση των κλιμάκων για διαφορετικά είδη μουσικής με bitrate 320kbps

Τα πράγματα γίνονται πολύ πιο ξεκάθαρα στο Γράφημα 6.16, όπου και αναπαρίστανται τα factor scores των ειδών. Η χειρότερη βαθμολογία είναι αυτή της κλασικής μουσικής με ορχήστρα, μιας και σε δύο από τους τρεις παράγοντες (Evaluation και Activity) βαθμολογείται με score μικρότερο από το -0,5. Στα υπόλοιπα είδη μουσικής, το solo όργανο είναι και πάλι στην καλύτερη θέση, με τον παράγοντα Potency να βρίσκεται στο μηδέν, ενώ στη δεύτερη θέση ακολουθεί η jazz, με το score και των τριών παραγόντων να βρίσκεται στο -0,2. Η ηλεκτρονική και η rock μουσική, βρίσκονται κοντά στις βαθμολογίες, με τα μεγαλύτερα προβλήματα να εμφανίζονται στους παράγοντες Evaluation και Activity για την πρώτη και στον παράγοντα Activity για τη δεύτερη. Πάντως, σε γενικές γραμμές, οι βαθμολογίες που συγκεντρώνουν όλα τα μουσικά είδη

είναι πολύ καλές δείχνοντας πως οι διαφορές που εντοπίζονται σε σχέση με τα πρωτότυπα αρχεία WAV είναι μικρές.



Γράφημα 6.16: Σύγκριση των factor scores για διαφορετικά είδη μουσικής με bitrate 320kbps

ΚΕΦΑΛΑΙΟ 7^ο: ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΠΕΡΙΟΡΙΣΜΟΙ ΤΗΣ ΕΡΕΥΝΑΣ

Εισαγωγή 7^ο Κεφαλαίου

Στο κεφάλαιο αυτό, απαντώνται τα ερευνητικά ερωτήματα που τέθηκαν, ερμηνεύοντας τα αποτελέσματα που παρουσιάστηκαν στο προηγούμενο κεφάλαιο και εξάγονται τα συμπεράσματα σχετικά με την επίδοση του αλγορίθμου του MP3 ανάλογα με το bitrate και το είδος μουσικής που συμπιέζεται.

Πρέπει να σημειωθεί εδώ ότι τα συμπεράσματα της συγκεκριμένης έρευνας είναι δύσκολο να γενικευτούν για ολόκληρο τον πληθυσμό, μιας και υπόκεινται σε μια σειρά περιορισμών, όπως αυτοί θα αναλυθούν παρακάτω.

7.1 Συμπεράσματα

Σύμφωνα με τα αποτελέσματα που ελήφθησαν από την έρευνα, εξάγονται τα παρακάτω συμπεράσματα σχετικά με την αξιολόγηση του αλγορίθμου MPEG-1 Layer III για τα μουσικά είδη rock, jazz, ηλεκτρονική, κλασική (με ορχήστρα) και solo όργανο:

- Το πρώτο και πιο φανερό συμπέρασμα που εξάγεται είναι πως ο αλγόριθμος του MP3, δεν αποδίδει εξίσου καλά, ούτε σε όλα τα είδη μουσικής, ούτε σε όλα τα bitrates. Συνεπώς, προκειμένου να επιτευχθεί η μέγιστη απόδοση και να γίνει σωστή επιλογή bitrate, πρέπει κάθε φορά να λαμβάνεται υπ' όψιν το προς συμπίεση μουσικό είδος.
- Όσον αφορά ξεχωριστά το κάθε ένα από τα πέντε είδη μουσικής³⁷:
 - Για το είδος της rock, τα bitrates των 96kbps και 160kbps δεν έχουν καλή απόδοση. Συνεπώς, το συγκεκριμένο είδος χρειάζεται σίγουρα μεγαλύτερα bitrates προκειμένου να αποφευχθούν αντιληπτές αλλοιώσεις του ηχητικού

³⁷ Στα συμπεράσματα δεν θα γίνει αναφορά στην επίδοση του αλγορίθμου με bitrate στα 320kbps, μιας και σε όλα τα είδη μουσικής, το MP3 έχει ιδιαίτερα υψηλές επιδόσεις στο συγκεκριμένο bitrate.

σήματος λόγω της συμπίεσης. Πιο συγκεκριμένα, το είδος αυτό, στα δύο συγκεκριμένα bitrates, επηρεάζεται κατά πολύ στους τομείς των συχνοτήτων και της έντασης και λιγότερο στον τομέα του θορύβου και της παραμόρφωσης. Όμως, μιας και η ένταση και το συχνοτικό φάσμα είναι βασικοί παράγοντες για την γενική αισθητική ενός κομματιού, αναπόφευκτα επηρεάζεται και αυτή.

- Στο είδος της jazz, παρατηρήθηκε η κακή επίδοση του αλγορίθμου στα bitrates των 96kbps και 160kbps, με πιο μελανά σημεία αυτά του συχνοτικού φάσματος και της γενικής αισθητικής του κομματιού. Αξίζει να σημειωθεί, ότι για το bitrate των 96kbps, παρατηρήθηκε μεγάλο ποσοστό παραμόρφωσης.
- Για την ηλεκτρονική μουσική, τα αποτελέσματα της έρευνας έδειξαν πως η συμπίεση με bitrate στα 96kbps και 160kbps είναι μια πολύ κακή επιλογή, μιας και τόσο ηχοχρωματικά όσο και σε θέμα έντασης αλλά και γενικής εντύπωσης, ο αλγόριθμος αλλοιώνει αρκετά το ηχητικό σήμα. Βέβαια, στο bitrate των 160kbps, τα αποτελέσματα παρουσιάζονται αρκετά βελτιωμένα σε σχέση με αυτά των 96kbps, χωρίς όμως η βελτίωση αυτή να είναι αρκετή έτσι ώστε να αποτελεί ικανοποιητική επιλογή.
- Περνώντας στο είδος της κλασσικής μουσικής με ορχήστρα, τα αποτελέσματα ήταν μάλλον αναμενόμενα μιας και στα 96kbps και τα 160kbps η συμπίεση γίνεται αντιληπτή κυρίως στον τομέα του ηχοχρώματος και της γενικής εντύπωσης. Φυσικά, και εδώ, το bitrate των 160kbps δίνει σαφώς καλύτερα αποτελέσματα από αυτό των 96kbps, χωρίς όμως να φτάνει στο επιθυμητό αποτέλεσμα. Συνεπώς, για το συγκεκριμένο είδος, είναι απαραίτητη η χρήση ενός bitrate μεγαλύτερου από αυτό των 160kbps κατά τη συμπίεση του ηχητικού σήματος, προκειμένου οι διαφορές ανάμεσα στο πρωτότυπο και το συμπιεσμένο αρχείο να μη γίνονται αντιληπτές από τους ακροατές. Αξιοσημείωτο είναι το γεγονός πως ακόμη και στο bitrate 320kbps, οι διαφορές ανάμεσα στο συμπιεσμένο και το πρωτότυπο ηχητικό σήμα είναι μεγαλύτερες απ' ό,τι στα υπόλοιπα μουσικά είδη.

- Στο τελευταίο μουσικό είδος, αυτό του solo οργάνου, παρατηρήθηκε η κακή επίδοση του αλγορίθμου στο bitrate των 96kbps, με τις περισσότερες αλλοιώσεις να εντοπίζονται στον τομέα του ηχοχρώματος και της έντασης του ηχητικού σήματος. Στο bitrate των 160kbps, η επίδοση του αλγορίθμου βελτιώθηκε κατά πολύ και οι ακουστές αλλοιώσεις σχεδόν εξαφανίστηκαν, καθιστώντας το έτσι πολύ ικανοποιητική επιλογή για τη συμπίεση του συγκεκριμένου είδους.

Στον Πίνακα 7.1 παρατίθενται συγκεντρωτικά τα προτεινόμενα bitrates για κάθε ένα από τα πέντε μουσικά είδη, όπως αυτά προέκυψαν από τα αποτελέσματα της έρευνας.

Μουσικό Είδος	Προτεινόμενο Bitrate
Rock	Μεγαλύτερο από τα 160kbps
Jazz	Μεγαλύτερο από τα 160kbps
Ηλεκτρονική	Μεγαλύτερο από τα 160kbps
Κλασσική (με ορχήστρα)	Μεγαλύτερο από τα 160kbps
Solo όργανο	160kbps

Πίνακας 7.1: Προτεινόμενα bitrates για κάθε μουσικό είδος

Συνοψίζοντας, είναι φανερό ότι:

- Για συμπίεση με λίγες αντιληπτές αλλοιώσεις, είναι καλό τα bitrates που επιλέγονται να μην είναι μικρότερα από αυτό των 160kbps. Σίγουρα, λοιπόν, το bitrate των 96kbps, σε κάθε περίπτωση, κρίνεται απαγορευτικό.
- Ηχητικά σήματα με μικρό δυναμικό εύρος (π.χ. solo όργανο) μπορούν να κωδικοποιηθούν σε σχετικά μικρά bitrates, χωρίς οι αλλοιώσεις στο σήμα να είναι αντιληπτές. Αντίθετα, σήματα με μεγάλο δυναμικό εύρος (π.χ. κλασσική μουσική με ορχήστρα) χρειάζονται μεγαλύτερο bitrate προκειμένου οι αλλοιώσεις τις συμπίεσης να μην είναι ιδιαίτερα διακριτές.
- Αν και η μείωση των bits (χαμηλό bitrate), για την κωδικοποίηση του ηχητικού σήματος είναι η βασικότερη διεργασία που επιτελούν οι αλγόριθμοι συμπίεσης, ο θόρυβος κβαντισμού γίνεται λιγότερο αντιληπτός σε σχέση με τις συχνοτικές και δυναμικές αλλοιώσεις. Όπως φαίνεται και από τα αποτελέσματα της έρευνας, τα μεγαλύτερα προβλήματα προέκυψαν στις κλίμακες που αφορούσαν το φάσμα των μουσικών κομματιών και όχι τόσο της εισαγωγή θορύβου.

- Στις υψηλές συχνότητες εντοπίζεται το μεγαλύτερο πρόβλημα του αλγορίθμου, μιας και είναι αρκετά έντονες οι αλλοιώσεις σε όλα τα εξεταζόμενα μουσικά είδη. Φυσικά, αυτό είναι αναμενόμενο μιας και ο αλγόριθμος είναι σχεδιασμένος έτσι ώστε να χρησιμοποιεί όσο λιγότερα bits γίνεται. Οι υψηλές συχνότητες είναι αυτές που χρειάζονται πολλά bits για την κωδικοποίησή τους. Συνεπώς, είναι και αυτές που χάνονται πρώτες κατά τη συμπίεση.
- Τα ηχητικά σήματα τα οποία δεν έχουν ιδιαίτερες αυξομειώσεις στη δυναμική τους και περιέχουν φυσικά μουσικά όργανα (solo όργανο (πίانو)) μπορούν να κωδικοποιηθούν σε μικρότερα bitrates σε σχέση με τα αντίστοιχα σήματα τα οποία περιέχουν ηλεκτρονικά μουσικά όργανα (rock).

Τα αποτελέσματα που βρέθηκαν σε αυτή την εργασία συμβαδίζουν σε μεγάλο βαθμό με αυτά παρόμοιων ερευνών όπως αυτή του Evan P. Ruzanski³⁸ και αυτή των Pras et al³⁹.

Πιο συγκεκριμένα:

- Στην έρευνα των Pras et al, αναφέρεται ότι ακόμα και επαγγελματίες στο χώρο του ήχου με δεκαετή πείρα σε studio, δεν ήταν σε θέση να διακρίνουν ιδιαίτερες διαφορές ανάμεσα στα πρωτότυπα αρχεία WAV και τα mp3 με bitrate στα 320kbps, για μουσικά είδη όπως η rock και η κλασική μουσική με ορχήστρα, ένα συμπέρασμα το οποίο εξήχθη και στην παρούσα εργασία. Το συγκεκριμένο αποτέλεσμα βρέθηκε και στην έρευνα του Ruzanski, στην οποία εξετάστηκαν παρόμοια μουσικά είδη (heavy metal, soft rock, κλασική μουσική, solo πιάνο).
- Στην έρευνα των Pras et al, οι φασματικές αλλοιώσεις ήταν οι πιο εμφανείς και ακουστές σε σχέση με το θόρυβο που εισάγεται λόγω του κβαντισμού, ένα αποτέλεσμα το οποίο βρέθηκε και στην παρούσα εργασία. Επίσης, βρέθηκαν σημαντικές αλλοιώσεις στον τομέα των δυναμικών, λόγω της μείωσης των bits.

³⁸ Ruzanski, E. P. (2006). Effects of MP3 encoding on the sounds of music. IEEE Potentials, 43-45, 25.

³⁹ Pras, A., Zimmerman, R., Levitin, D., Guastavino, C. (2009). Subjective evaluation of mp3 compression for different musical genres. AES 127th Convention, New York, NY, USA.

- Στην έρευνα του Ruzanski, βρέθηκε ότι το είδος του solo οργάνου είναι αυτό που απαιτεί το μικρότερο bitrate προκειμένου να συμπιεσθεί χωρίς ιδιαίτερα ακουστές αλλοιώσεις, γεγονός που παρουσιάστηκε και στην παρούσα εργασία, με το solo όργανο να είναι το μόνο είδος το οποίο δεν απαιτεί bitrate μεγαλύτερο των 160kbps.
- Στην ίδια έρευνα, βρέθηκε επίσης, ότι τα μουσικά είδη που χρειάζονται συμπίεση σε υψηλότερα bitrates είναι αυτά τα οποία έχουν μεγάλο δυναμικό εύρος, όπως η κλασική μουσική με ορχήστρα, γεγονός που επιβεβαιώνεται και στην παρούσα εργασία. Επίσης, αναφέρεται πως είδη τα οποία περιέχουν μεγάλα ποσά ηλεκτρονικών ήχων και μουσικών οργάνων, χρειάζονται μεγαλύτερα bitrates κατά τη συμπίεση σε σχέση με τα είδη που περιέχουν φυσικά μουσικά όργανα. Φυσικά, μεγάλο ρόλο σε αυτό το διαχωρισμό παίζει το δυναμικό εύρος, όπως αναφέρθηκε προηγουμένως, συμπέρασμα το οποίο εξήχθη και στην παρούσα εργασία.

7.2 Περιορισμοί της Έρευνας

Στη συνέχεια, παρατίθενται ορισμένοι περιορισμοί που αφορούν στον τρόπο διεξαγωγής της πειραματικής διαδικασίας.

- Όπως αναφέρθηκε και σε προηγούμενο κεφάλαιο, το δείγμα που επιλέχθηκε για τη διεξαγωγή των listening tests, δεν ήταν τυχαίο, αλλά μη πιθανοτικό. Αυτό σημαίνει, πως το δείγμα δεν επιλέγεται τυχαία, αλλά με βάση ορισμένα κριτήρια που θέτει ο ερευνητής (τεχνικές γνώσεις, ηλικία κλπ) και συνεπώς τα συμπεράσματα της έρευνας δεν μπορούν να γενικευτούν σε ολόκληρο τον πληθυσμό.
- Το μέγεθος του δείγματος (30 άτομα), είναι μεν επαρκές για έρευνα με τη συγκεκριμένη μεθοδολογία, άλλα είναι και οριακό (σύμφωνα με τη μεθοδολογία το δείγμα πρέπει να είναι $N \geq 30$). Πιθανώς, να υπάρχουν σφάλματα στις μετρήσεις λόγω οριακά μικρού δείγματος.

- Ο χώρος στον οποίο διεξήχθη η πειραματική διαδικασία διαθέτει ηχομόνωση και χαμηλή στάθμη θορύβου, όχι όμως σε επίπεδο που να μπορεί να εξασφαλιστεί η απόλυτη σιγή. Πιθανώς, λοιπόν, τα υποκείμενα να έδωσαν εσφαλμένες απαντήσεις οι οποίες να προήλθαν από κάποιον εξωγενή θόρυβο.
- Η διάρκεια της συμπλήρωσης των ερωτηματολογίων, όπως αναφέρθηκε και σε προηγούμενο κεφάλαιο, ήταν κατά μέσο όρο, 50 λεπτά. Πιθανώς, ορισμένα υποκείμενα να κουράστηκαν από μια τόσο μακροσκελή ακρόαση και να μην έδωσαν προσεκτικές απαντήσεις.

Βιβλιογραφία

Ελληνική

Παπαδάκης, Ν. (2009). Πανεπιστημιακές σημειώσεις για το μάθημα «Εισαγωγή στην Ψυχοακουστική». ΑΤΕΙ Κρήτης, Παράρτημα Ρεθύμνου, Τμήμα Μουσικής Τεχνολογίας και Ακουστικής.

Ξενόγλωσση

Al-Hindawe, J. (1996). Considerations when constructing a semantic differential scale. *La Trobe working papers in linguistics*, Vol.9.

Arehart, K. H., Hansen, J. H. L., Gallant, S., Kalstein, L. (2003). Evaluation of an auditory masked threshold noise suppression algorithm in normal-hearing and hearing-impaired listeners. *Speech communication*. 40, 575-592

Avelar, D., Morissette, B., Forbes, D., Albert, G. (2008). Audio codecs: Evaluation and comparison of popular formats. *ECE 560*.

Bech, S. & Zacharov, N.. (2006). *Perceptual audio evaluation: Theory, Method and Application*. John Wiley & Sons Ltd, Great Britain

Bergman, P., Skold, A., Vastfjall, D., Fransson, N. (2009). Perceptual and emotional categorization of sound. *Journal of acoustic society of America*. 126, No.6, 3156-3167

Bopp, Joan. (1955). A quantitative semantic analysis of word association in schizophrenia. *Unpublished doctor's dissertation*, University of Illinois, 131, 170, 223-26, 227

Bosi, M & Goldberg R. E. (2003). *Introduction to digital audio coding and standards*. Kluwer Academic Publishers, Dordrecht, The Netherlands

- Brandenburg, K. (1999). MP3 and AAC explained. *AES 17th International Conference on High Quality Audio Coding*.
- Brandenburg, K. & Popp, H. (2000). An introduction to MPEG Layer 3. *EBU Technical Review*.
- Buss, S., Schulte-Fortkamp, B., Muckel, P. (2000). Combining methods to evaluate sound quality. Το έγγραφο αυτό είναι διαθέσιμο στην ιστοσελίδα www.physik.uni-oldenburg.de/docs/aku/literatur/Buss/buss_in2000.pdf
- Carter, R. F., Ruggels, W. L., Chaffee, S. H. The semantic differential in opinion measurement. Το έγγραφο αυτό είναι διαθέσιμο στην ιστοσελίδα www.poq.oxfordjournals.org
- Chasin, M. & Russo, F. A. (2004). Hearing aids and music. *Trends in amplification*. 8, No. 2, 35-47
- Cliff, N. (1959). Adverbs as multipliers. *Psychological Review*, 66, 27-44.
- Damiani, A., JERIC, D., Imberty, M., Olivetti Belardinelli, M. (2004). Evaluation of the musical style using the semantic differential technique and a fuzzy graphic rating scale. *ICMPC8*, Evanston, IL, USA
- DiVesta, F. J., & Dick, W. (1966). The test-retest reliability of children's ratings on the semantic differential. *Educational and Psychological Measurement*, 26, 605-616.
- Fastl, H. (2005). Chapter 6: Psycho-acoustics and sound quality. *Communication Acoustics*. Springer-Verlag Berlin Heidelberg, The Netherlands. 139-162.
- Gulliksen, H. (1954). A least squares solution for successive intervals assuming unequal standard deviations. *Psychometrika*, 19, 117-139. 146.
- Green, R.F., & Goldfried, M. R. (1965). On the bipolarity of semantic space. *Psychological Monographs* 79 (6, Whole No.601).
- Hansen, H., Weber, R. (2009). Semantic evaluations of noise with tonal components in Japan, France and Germany-A cross-cultural comparison. *Journal of acoustic society of America*. 125, 2, 850-862

- Heise, D. R. (1959). Semantic differential profiles for 1,000 most frequent English words. *Psychological Monographs*, 79 (8, Whole No. 601).
- Heise, D. R. (1969). Some methodological issues in semantic differential. *Psychological Bulletin*, 72 6. 406-422.
- Heise, D. R. (1970). Chapter 14: The semantic differential and attitude research. *Attitude measurement*. Chicago. 235-253
- Howard, D.M., Angus, J. (2006). *Acoustics and psychoacoustics*. Focal Press, Great Britain
- Howe, E. S. (1962). Probabilistic adverbial qualification of adjectives. *Journal of Verbal Learning and Verbal Behavior*, 225-242.
- Howe, E. S. (1966). Associative structure of quantifiers. *Journal of Verbal Learning and Verbal Behavior*, 5, 156-162.
- Ishimitsu, S., Sakamoto, K., Arai, T., Yoshimi, T., Fujimoto, Y., Kawasaki, K. (2008). A study of evaluating the button sounds. *Acoustics '08, Paris*. 3153-3158.
- Iwanaga, M., Moroki, Y. (1999). Subjective and physiological responses to music stimuli controlled over activity and preference. *Journal of Music Therapy*. XXXVI, 26-38.
- Jakobovits, L. A. (1966). Comparative psycholinguistics in the study of cultures. *International Journal of Psychology*, 1, 15-37.
- Jobert S. Jacaba. (2001). *Audio compression using modified discrete cosine transform: The MP3 coding standard*. Undergraduate research paper. Department of Mathematics, College of Science, The University of Philippines, Diliman, Quezon City.
- Kubo, N., Mellert, V., Weber, R., Mesche, J. (2004). Categorisation of engine sounds. *33rd International Congress and Exposition on Noise Control Engineering*. Prague, Czech Republic

- Kumata, H., & W. Schramm. (1956). A pilot study of cross-cultural meaning. *Publ. Opin. Quart.*, 20, 229-238. 171-176, 222.
- Langerstrom, K. (2001). *Design and implementation of an MPEG-1 Layer III audio decoder*. Master's Thesis. Chalmers University of Technology, Gothenburg, Sweeden.
- Liu, C. M., Hsu, H. W., Lee, W. C. (2008). Compression artifacts in perceptual audio coding. *IEEE Transaction on Audio, Speech and Language Processing*. 16 2, 681-695
- Messick, S. J. (1957). Metric properties of the semantic differential. *Educational and Psychological Measurement*, 17, 200-206.
- Namba, S., Kuwano, S., Kinoshita, K., Kurakata, K. (1992). Loudness and timbre of broad-band noise mixed with frequency-modulated sounds. *J. Acoust. Soc. Jpn. (E)*, 13, 49-58.
- Norman, W. (1959). Stability characteristics of the semantic differential. *American Journal of Psychology*, 72, 581-584.
- Osgood, C. E. (1964). Semantic differential technique in the comparative study of cultures. *American Anthropologist*, 66(3, Pt. 2), 171-200.
- Osgood, C. E. & Suci, G. J. & Tannenbaum, P. H. (1957). *The measurement of meaning*. University of Illinois.
- Pohlmann, K. C. (2005). *Principles of digital audio 5th edition*. McGraw-Hill, New York
- Praveen Spirada (2006). *MP3 decoder in theory and practice*. Master's Thesis Report. Department of Signal Processing and Telecommunications. Bleckinge Institute of Technology, Ronneby, Sweeden
- Pras, A., Zimmerman, R., Levitin, D., Guastavino, C. (2009). Subjective evaluation of mp3 compression for different musical genres. *AES 127th Convention*, New York, NY, USA.

- Rassol Raissi (2002). *The theory behind MP3*. Το έγγραφο αυτό είναι διαθέσιμο στην ιστοσελίδα www.mp3-tech.org.
- Rowan, T. C. (1954). Some developments in multidimensional scaling applied to semantic relationships. *Unpublished doctor's dissertation*, University of Illinois. 143-46.
- Ruzanski, E. P. (2006). Effects of MP3 encoding on the sounds of music. *IEEE Potentials*, 43-45, 25.
- Saffir, M. (1937). A comparative study of scales constructed by three psychophysical methods. *Psychometrika*, 2, 179-198.
- Sayood, K. (2006). *Introduction to data compression 3rd edition*. Morgan Kaufmann Publishers, USA
- Schutte, M., Muller, U., Sandrock, S., Griefahn, B., Lavandier, C., Barbot, B. (2009). Perceived quality features of aircraft sounds-An analysis of the measurement characteristics of a newly created semantic differential. *Applied acoustics*. 70, 903-914
- Shlien, S., Soulodre, G. (1996). Measuring the characteristics of “expert” listeners. *AES 101st Convention*. Los Angeles.
- Solomon, L. N. (1954). A factorial study of complex auditory stimuli (passive sonar sounds). *Unpublished doctor's dissertation*, University of Illinois, 66-68, 128.
- Spanias, A. & Painter, T. & Atti, V. (2007). *Audio Signal Processing and Coding*. John Wiley & Sons, Inc.
- Takeda, S. et al (1992). Age variation in the upper limit of hearing. *European journal of applied physiology*. 65, 403-408
- Valois, P., Godin, G. (1991). The importance of selecting appropriate adjective pairs for measuring attitude based on the semantic differential method. *Quality & quantity*. 25, 57-68.
- Watkinson, J. (2001). *The art of digital audio 3rd edition*. Focal Press, Great Britain

Yen Pan, D. (1993). Digital audio compression. *Digital Technical Journal*, 5, No.2.

You, J., Reiter, U., Hannuksela, M. M., Gabbouj, M., Perkis, A. (2010) Perceptual-based quality assessment for audio-visual services: A survey. *Signal processing: Image communication*. 25, 482-501

ΠΑΡΑΡΤΗΜΑ Α

Σκοπός της μελέτης είναι η αξιολόγηση του προτύπου MPEG-1 Layer III (MP3) για διάφορα είδη μουσικής. Η αξιολόγηση θα γίνει από ένα πλήθος ανθρώπων μέσω ορισμένων κλιμάκων. Το ερωτηματολόγιο είναι κατασκευασμένο έτσι ώστε να αποτυπώσει τη δική τις, υποκειμενική κρίση. Παρακαλείστε, να διαβάσετε προσεκτικά τις οδηγίες για τη συμπλήρωση, προκειμένου να κατανοήσετε τη διαδικασία η οποία θα ακολουθήσει.

ΟΔΗΓΙΕΣ ΣΥΜΠΛΗΡΩΣΗΣ ΕΡΩΤΗΜΑΤΟΛΟΓΙΟΥ

Το ερωτηματολόγιο αυτό αποτελείται από 9 κλίμακες για κάθε ζεύγος μουσικών κομματιών. Θα ακούσετε τα ζεύγη αυτά και θα τοποθετήσετε ένα X σε κάθε κλίμακα, ανάλογα με την κρίση τις. Η συμπλήρωσή γίνεται με τον εξής τρόπο:

- Εάν θεωρείτε ότι το άκουσμα είναι **ΠΟΛΥ** κοντά σε κάποιο από τα άκρα τις κλίμακας, τοποθετείστε το X τις φαίνεται παρακάτω:

	Πολύ	Αρκετά	Λίγο	Το Ίδιο	Λίγο	Αρκετά	Πολύ	
Χειρότερο	X							Καλύτερο

ή

	Πολύ	Αρκετά	Λίγο	Το Ίδιο	Λίγο	Αρκετά	Πολύ	
Χειρότερο							X	Καλύτερο

- Εάν θεωρείτε ότι το άκουσμα είναι **ΑΡΚΕΤΑ** κοντά σε κάποιο από τα άκρα τις κλίμακας, τοποθετείστε το X τις φαίνεται παρακάτω:

	Πολύ	Αρκετά	Λίγο	Το Ίδιο	Λίγο	Αρκετά	Πολύ	
Χειρότερο		X						Καλύτερο

ή

	Πολύ	Αρκετά	Λίγο	Το Ίδιο	Λίγο	Αρκετά	Πολύ	
Χειρότερο						X		Καλύτερο

- Εάν θεωρείτε ότι το άκουσμα είναι **ΛΙΓΟ** κοντά σε κάποιο από τα άκρα τις κλίμακας, τοποθετείστε το X τις φαίνεται παρακάτω:

	Πολύ	Αρκετά	Λίγο	Το Ίδιο	Λίγο	Αρκετά	Πολύ	
Χειρότερο			X					Καλύτερο

ή

	Πολύ	Αρκετά	Λίγο	Το Ίδιο	Λίγο	Αρκετά	Πολύ	
Χειρότερο					X			Καλύτερο

- Εάν θεωρείτε ότι το άκουσμα απέχει **ΕΞΙΣΟΥ** από τα δύο άκρα τις κλίμακας, τοποθετείστε το X τις φαίνεται παρακάτω:

	Πολύ	Αρκετά	Λίγο	Το Ίδιο	Λίγο	Αρκετά	Πολύ	
Χειρότερο				X				Καλύτερο

ΠΡΟΣΟΧΗ

1. Σιγουρευτείτε ότι δώσατε απάντηση σε όλες τις κλίμακες.
2. Μη βάλετε παραπάνω από ένα X σε κάθε κλίμακα.

1. Ακούγοντας και συγκρίνοντας τα κομμάτια 1 και 2, θα λέγατε ότι το **κομμάτι 1** είναι/έχει (σε σχέση με το 2):

	Πολύ	Αρκετά	Λίγο	Το Ίδιο	Λίγο	Αρκετά	Πολύ	
Χειρότερο								Καλύτερο
Πιο Φτωχό Φάσμα								Πιο Πλούσιο Φάσμα
Περισσότερο Θόρυβο								Λιγότερο Θόρυβο
Πιο Ενοχλητικό								Πιο Απολαυστικό
Λιγότερο Δυνατό								Περισσότερο Δυνατό
Πιο Δύσκολο Διαχωρισμό Οργάνων								Πιο Εύκολο Διαχωρισμό Οργάνων
Λιγότερο Ζεστό								Περισσότερο Ζεστό
Περισσότερη Παραμόρφωση								Λιγότερη Παραμόρφωση
Λιγότερο Λαμπρό								Περισσότερο Λαμπρό

2. Ακούγοντας και συγκρίνοντας τα κομμάτια 3 και 4, θα λέγατε ότι το **κομμάτι 3** είναι/έχει (σε σχέση με το 4):

	Πολύ	Αρκετά	Λίγο	Το Ίδιο	Λίγο	Αρκετά	Πολύ	
Χειρότερο								Καλύτερο
Πιο Φτωχό Φάσμα								Πιο Πλούσιο Φάσμα
Περισσότερο Θόρυβο								Λιγότερο Θόρυβο
Πιο Ενοχλητικό								Πιο Απολαυστικό
Λιγότερο Δυνατό								Περισσότερο Δυνατό
Πιο Δύσκολο Διαχωρισμό Οργάνων								Πιο Εύκολο Διαχωρισμό Οργάνων
Λιγότερο Ζεστό								Περισσότερο Ζεστό
Περισσότερη Παραμόρφωση								Λιγότερη Παραμόρφωση
Λιγότερο Λαμπρό								Περισσότερο Λαμπρό

3. Ακούγοντας και συγκρίνοντας τα κομμάτια 5 και 6, θα λέγατε ότι το **κομμάτι 5** είναι/έχει (σε σχέση με το 6):

	Πολύ	Αρκετά	Λίγο	Το Ίδιο	Λίγο	Αρκετά	Πολύ	
Χειρότερο								Καλύτερο
Πιο Φτωχό Φάσμα								Πιο Πλούσιο Φάσμα
Περισσότερο Θόρυβο								Λιγότερο Θόρυβο
Πιο Ενοχλητικό								Πιο Απολαυστικό
Λιγότερο Δυνατό								Περισσότερο Δυνατό
Πιο Δύσκολο Διαχωρισμό Οργάνων								Πιο Εύκολο Διαχωρισμό Οργάνων
Λιγότερο Ζεστό								Περισσότερο Ζεστό
Περισσότερη Παραμόρφωση								Λιγότερη Παραμόρφωση
Λιγότερο Λαμπρό								Περισσότερο Λαμπρό

4. Ακούγοντας και συγκρίνοντας τα κομμάτια 7 και 8, θα λέγατε ότι το **κομμάτι 7** είναι/έχει (σε σχέση με το 8):

	Πολύ	Αρκετά	Λίγο	Το Ίδιο	Λίγο	Αρκετά	Πολύ	
Χειρότερο								Καλύτερο
Πιο Φτωχό Φάσμα								Πιο Πλούσιο Φάσμα
Περισσότερο Θόρυβο								Λιγότερο Θόρυβο
Πιο Ενοχλητικό								Πιο Απολαυστικό
Λιγότερο Δυνατό								Περισσότερο Δυνατό
Πιο Δύσκολο Διαχωρισμό Οργάνων								Πιο Εύκολο Διαχωρισμό Οργάνων
Λιγότερο Ζεστό								Περισσότερο Ζεστό
Περισσότερη Παραμόρφωση								Λιγότερη Παραμόρφωση
Λιγότερο Λαμπρό								Περισσότερο Λαμπρό

5. Ακούγοντας και συγκρίνοντας τα κομμάτια 9 και 10, θα λέγατε ότι το **κομμάτι 9** είναι/έχει (σε σχέση με το 10):

	Πολύ	Αρκετά	Λίγο	Το Ίδιο	Λίγο	Αρκετά	Πολύ	
Χειρότερο								Καλύτερο
Πιο Φτωχό Φάσμα								Πιο Πλούσιο Φάσμα
Περισσότερο Θόρυβο								Λιγότερο Θόρυβο
Πιο Ενοχλητικό								Πιο Απολαυστικό
Λιγότερο Δυνατό								Περισσότερο Δυνατό
Πιο Δύσκολο Διαχωρισμό Οργάνων								Πιο Εύκολο Διαχωρισμό Οργάνων
Λιγότερο Ζεστό								Περισσότερο Ζεστό
Περισσότερη Παραμόρφωση								Λιγότερη Παραμόρφωση
Λιγότερο Λαμπρό								Περισσότερο Λαμπρό

6. Ακούγοντας και συγκρίνοντας τα κομμάτια 11 και 12, θα λέγατε ότι το **κομμάτι 11** είναι/έχει (σε σχέση με το 12):

	Πολύ	Αρκετά	Λίγο	Το Ίδιο	Λίγο	Αρκετά	Πολύ	
Χειρότερο								Καλύτερο
Πιο Φτωχό Φάσμα								Πιο Πλούσιο Φάσμα
Περισσότερο Θόρυβο								Λιγότερο Θόρυβο
Πιο Ενοχλητικό								Πιο Απολαυστικό
Λιγότερο Δυνατό								Περισσότερο Δυνατό
Πιο Δύσκολο Διαχωρισμό Οργάνων								Πιο Εύκολο Διαχωρισμό Οργάνων
Λιγότερο Ζεστό								Περισσότερο Ζεστό
Περισσότερη Παραμόρφωση								Λιγότερη Παραμόρφωση
Λιγότερο Λαμπρό								Περισσότερο Λαμπρό

7. Ακούγοντας και συγκρίνοντας τα κομμάτια 13 και 14, θα λέγατε ότι το **κομμάτι 13** είναι/έχει (σε σχέση με το 14):

	Πολύ	Αρκετά	Λίγο	Το Ίδιο	Λίγο	Αρκετά	Πολύ	
Χειρότερο								Καλύτερο
Πιο Φτωχό Φάσμα								Πιο Πλούσιο Φάσμα
Περισσότερο Θόρυβο								Λιγότερο Θόρυβο
Πιο Ενοχλητικό								Πιο Απολαυστικό
Λιγότερο Δυνατό								Περισσότερο Δυνατό
Πιο Δύσκολο Διαχωρισμό Οργάνων								Πιο Εύκολο Διαχωρισμό Οργάνων
Λιγότερο Ζεστό								Περισσότερο Ζεστό
Περισσότερη Παραμόρφωση								Λιγότερη Παραμόρφωση
Λιγότερο Λαμπρό								Περισσότερο Λαμπρό

8. Ακούγοντας και συγκρίνοντας τα κομμάτια 15 και 16, θα λέγατε ότι το **κομμάτι 15** είναι/έχει (σε σχέση με το 16):

	Πολύ	Αρκετά	Λίγο	Το Ίδιο	Λίγο	Αρκετά	Πολύ	
Χειρότερο								Καλύτερο
Πιο Φτωχό Φάσμα								Πιο Πλούσιο Φάσμα
Περισσότερο Θόρυβο								Λιγότερο Θόρυβο
Πιο Ενοχλητικό								Πιο Απολαυστικό
Λιγότερο Δυνατό								Περισσότερο Δυνατό
Πιο Δύσκολο Διαχωρισμό Οργάνων								Πιο Εύκολο Διαχωρισμό Οργάνων
Λιγότερο Ζεστό								Περισσότερο Ζεστό
Περισσότερη Παραμόρφωση								Λιγότερη Παραμόρφωση
Λιγότερο Λαμπρό								Περισσότερο Λαμπρό

9. Ακούγοντας και συγκρίνοντας τα κομμάτια 17 και 18, θα λέγατε ότι το **κομμάτι 17** είναι/έχει (σε σχέση με το 18):

	Πολύ	Αρκετά	Λίγο	Το Ίδιο	Λίγο	Αρκετά	Πολύ	
Χειρότερο								Καλύτερο
Πιο Φτωχό Φάσμα								Πιο Πλούσιο Φάσμα
Περισσότερο Θόρυβο								Λιγότερο Θόρυβο
Πιο Ενοχλητικό								Πιο Απολαυστικό
Λιγότερο Δυνατό								Περισσότερο Δυνατό
Πιο Δύσκολο Διαχωρισμό Οργάνων								Πιο Εύκολο Διαχωρισμό Οργάνων
Λιγότερο Ζεστό								Περισσότερο Ζεστό
Περισσότερη Παραμόρφωση								Λιγότερη Παραμόρφωση
Λιγότερο Λαμπρό								Περισσότερο Λαμπρό

10. Ακούγοντας και συγκρίνοντας τα κομμάτια 19 και 20, θα λέγατε ότι το **κομμάτι 19** είναι/έχει (σε σχέση με το 20):

	Πολύ	Αρκετά	Λίγο	Το Ίδιο	Λίγο	Αρκετά	Πολύ	
Χειρότερο								Καλύτερο
Πιο Φτωχό Φάσμα								Πιο Πλούσιο Φάσμα
Περισσότερο Θόρυβο								Λιγότερο Θόρυβο
Πιο Ενοχλητικό								Πιο Απολαυστικό
Λιγότερο Δυνατό								Περισσότερο Δυνατό
Πιο Δύσκολο Διαχωρισμό Οργάνων								Πιο Εύκολο Διαχωρισμό Οργάνων
Λιγότερο Ζεστό								Περισσότερο Ζεστό
Περισσότερη Παραμόρφωση								Λιγότερη Παραμόρφωση
Λιγότερο Λαμπρό								Περισσότερο Λαμπρό

11. Ακούγοντας και συγκρίνοντας τα κομμάτια 21 και 22, θα λέγατε ότι το **κομμάτι 21** είναι/έχει (σε σχέση με το 22):

	Πολύ	Αρκετά	Λίγο	Το Ίδιο	Λίγο	Αρκετά	Πολύ	
Χειρότερο								Καλύτερο
Πιο Φτωχό Φάσμα								Πιο Πλούσιο Φάσμα
Περισσότερο Θόρυβο								Λιγότερο Θόρυβο
Πιο Ενοχλητικό								Πιο Απολαυστικό
Λιγότερο Δυνατό								Περισσότερο Δυνατό
Πιο Δύσκολο Διαχωρισμό Οργάνων								Πιο Εύκολο Διαχωρισμό Οργάνων
Λιγότερο Ζεστό								Περισσότερο Ζεστό
Περισσότερη Παραμόρφωση								Λιγότερη Παραμόρφωση
Λιγότερο Λαμπρό								Περισσότερο Λαμπρό

12. Ακούγοντας και συγκρίνοντας τα κομμάτια 23 και 24, θα λέγατε ότι το **κομμάτι 23** είναι/έχει (σε σχέση με το 24):

	Πολύ	Αρκετά	Λίγο	Το Ίδιο	Λίγο	Αρκετά	Πολύ	
Χειρότερο								Καλύτερο
Πιο Φτωχό Φάσμα								Πιο Πλούσιο Φάσμα
Περισσότερο Θόρυβο								Λιγότερο Θόρυβο
Πιο Ενοχλητικό								Πιο Απολαυστικό
Λιγότερο Δυνατό								Περισσότερο Δυνατό
Πιο Δύσκολο Διαχωρισμό Οργάνων								Πιο Εύκολο Διαχωρισμό Οργάνων
Λιγότερο Ζεστό								Περισσότερο Ζεστό
Περισσότερη Παραμόρφωση								Λιγότερη Παραμόρφωση
Λιγότερο Λαμπρό								Περισσότερο Λαμπρό

13. Ακούγοντας και συγκρίνοντας τα κομμάτια 25 και 26, θα λέγατε ότι το **κομμάτι 25** είναι/έχει (σε σχέση με το 26):

	Πολύ	Αρκετά	Λίγο	Το Ίδιο	Λίγο	Αρκετά	Πολύ	
Χειρότερο								Καλύτερο
Πιο Φτωχό Φάσμα								Πιο Πλούσιο Φάσμα
Περισσότερο Θόρυβο								Λιγότερο Θόρυβο
Πιο Ενοχλητικό								Πιο Απολαυστικό
Λιγότερο Δυνατό								Περισσότερο Δυνατό
Πιο Δύσκολο Διαχωρισμό Οργάνων								Πιο Εύκολο Διαχωρισμό Οργάνων
Λιγότερο Ζεστό								Περισσότερο Ζεστό
Περισσότερη Παραμόρφωση								Λιγότερη Παραμόρφωση
Λιγότερο Λαμπρό								Περισσότερο Λαμπρό

14. Ακούγοντας και συγκρίνοντας τα κομμάτια 27 και 28, θα λέγατε ότι το **κομμάτι 27** είναι/έχει (σε σχέση με το 28):

	Πολύ	Αρκετά	Λίγο	Το Ίδιο	Λίγο	Αρκετά	Πολύ	
Χειρότερο								Καλύτερο
Πιο Φτωχό Φάσμα								Πιο Πλούσιο Φάσμα
Περισσότερο Θόρυβο								Λιγότερο Θόρυβο
Πιο Ενοχλητικό								Πιο Απολαυστικό
Λιγότερο Δυνατό								Περισσότερο Δυνατό
Πιο Δύσκολο Διαχωρισμό Οργάνων								Πιο Εύκολο Διαχωρισμό Οργάνων
Λιγότερο Ζεστό								Περισσότερο Ζεστό
Περισσότερη Παραμόρφωση								Λιγότερη Παραμόρφωση
Λιγότερο Λαμπρό								Περισσότερο Λαμπρό

15. Ακούγοντας και συγκρίνοντας τα κομμάτια 29 και 30, θα λέγατε ότι το **κομμάτι 29** είναι/έχει (σε σχέση με το 30):

	Πολύ	Αρκετά	Λίγο	Το Ίδιο	Λίγο	Αρκετά	Πολύ	
Χειρότερο								Καλύτερο
Πιο Φτωχό Φάσμα								Πιο Πλούσιο Φάσμα
Περισσότερο Θόρυβο								Λιγότερο Θόρυβο
Πιο Ενοχλητικό								Πιο Απολαυστικό
Λιγότερο Δυνατό								Περισσότερο Δυνατό
Πιο Δύσκολο Διαχωρισμό Οργάνων								Πιο Εύκολο Διαχωρισμό Οργάνων
Λιγότερο Ζεστό								Περισσότερο Ζεστό
Περισσότερη Παραμόρφωση								Λιγότερη Παραμόρφωση
Λιγότερο Λαμπρό								Περισσότερο Λαμπρό

ΠΑΡΑΡΤΗΜΑ Β

Τίτλος	Μουσικό Είδος	Μορφή Αρχείου
Κομμάτι 1	Rock	WAV
Κομμάτι 2	Rock	MP3 με bitrate 96kbps
Κομμάτι 3	Jazz	MP3 με bitrate 160kbps
Κομμάτι 4	Jazz	WAV
Κομμάτι 5	Solo όργανο	MP3 με bitrate 160kbps
Κομμάτι 6	Solo όργανο	WAV
Κομμάτι 7	Ηλεκτρονική	MP3 με bitrate 96kbps
Κομμάτι 8	Ηλεκτρονική	WAV
Κομμάτι 9	Κλασσική (με ορχήστρα)	MP3 με bitrate 320kbps
Κομμάτι 10	Κλασσική (με ορχήστρα)	WAV
Κομμάτι 11	Jazz	WAV
Κομμάτι 12	Jazz	MP3 με bitrate 96kbps
Κομμάτι 13	Rock	MP3 με bitrate 320kbps
Κομμάτι 14	Rock	WAV
Κομμάτι 15	Κλασσική (με ορχήστρα)	MP3 με bitrate 96kbps
Κομμάτι 16	Κλασσική (με ορχήστρα)	WAV
Κομμάτι 17	Ηλεκτρονική	WAV
Κομμάτι 18	Ηλεκτρονική	MP3 με bitrate 320kbps
Κομμάτι 19	Solo όργανο	WAV
Κομμάτι 20	Solo όργανο	MP3 με bitrate 96kbps
Κομμάτι 21	Jazz	WAV
Κομμάτι 22	Jazz	MP3 με bitrate 320kbps
Κομμάτι 23	Κλασσική (με ορχήστρα)	MP3 με bitrate 160kbps
Κομμάτι 24	Κλασσική (με ορχήστρα)	WAV
Κομμάτι 25	Ηλεκτρονική	MP3 με bitrate 160kbps
Κομμάτι 26	Ηλεκτρονική	WAV
Κομμάτι 27	Rock	MP3 με bitrate 160kbps
Κομμάτι 28	Rock	WAV
Κομμάτι 29	Solo όργανο	MP3 με bitrate 320kbps
Κομμάτι 30	Solo όργανο	WAV