



Σχολή Εφαρμοσμένων Τεχνών

Ακουστικός Σχεδιασμός και Ψηφιακός Ήχος

Διπλωματική Εργασία

«Μελέτη αντικειμενικών χαρακτηριστικών σε μουσικά σήματα για
ανίχνευση συναισθηματικής απόκρισης»

Χατζησταμάτη Στέλλα

Επιβλέπων καθηγητής: Χατζηαντωνίου Παναγιώτης

Β' Επιβλέπων καθηγητής: Καλλίρης Γεώργιος

Αθήνα, Ιούλιος 2024

Η παρούσα εργασία αποτελεί πνευματική ιδιοκτησία του φοιτητή («συγγραφέας/δημιουργός») που την εκπόνησε. Στο πλαίσιο της πολιτικής ανοικτής πρόσβασης ο συγγραφέας/δημιουργός εκχωρεί στο ΕΑΠ, μη αποκλειστική άδεια χρήσης του δικαιώματος αναπαραγωγής, προσαρμογής, δημόσιου δανεισμού, παρουσίασης στο κοινό και ψηφιακής διάχυσής τους διεθνώς, σε ηλεκτρονική μορφή και σε οποιοδήποτε μέσο, για διδακτικούς και ερευνητικούς σκοπούς, άνευ ανταλλάγματος και για όλο το χρόνο διάρκειας των δικαιωμάτων πνευματικής ιδιοκτησίας. Η ανοικτή πρόσβαση στο πλήρες κείμενο για μελέτη και ανάγνωση δεν σημαίνει καθ' οιονδήποτε τρόπο παραχώρηση δικαιωμάτων διανοητικής ιδιοκτησίας του συγγραφέα/δημιουργού ούτε επιτρέπει την αναπαραγωγή, αναδημοσίευση, αντιγραφή, αποθήκευση, πώληση, εμπορική χρήση, μετάδοση, διανομή, έκδοση, εκτέλεση, «μεταφόρτωση» (downloading), «ανάρτηση» (uploading), μετάφραση, τροποποίηση με οποιονδήποτε τρόπο, τμηματικά ή περιληπτικά της εργασίας, χωρίς τη ρητή προηγούμενη έγγραφη συναίνεση του συγγραφέα/δημιουργού. Ο συγγραφέας/δημιουργός διατηρεί το σύνολο των ηθικών και περιουσιακών του δικαιωμάτων.



«Μελέτη αντικειμενικών χαρακτηριστικών σε μουσικά σήματα για
ανίχνευση συναισθηματικής απόκρισης»

Χατζησταμάτη Στέλλα

Επιτροπή Επίβλεψης Διπλωματικής Εργασίας

Επιβλέπων καθηγητής:
Χατζηαντωνίου Παναγιώτης

Συν-Επιβλέπων καθηγητής:
Καλλίρης Γεώργιος

Αθήνα, Ιούλιος 2024

Ευχαριστίες

Κατά τη διαδικασία εκπόνησης της Διπλωματικής Εργασίας, παρακινήθηκα και βοηθήθηκα από τον επιβλέποντα καθηγητή μου, κ. Χατζηαντωνίου Παναγιώτη. Τον ευχαριστώ θερμά για την υποστήριξη και την κατανόηση σε αυτή μου την προσπάθεια. Επίσης, θα ήθελα να ευχαριστήσω του γονείς μου για όλη την στήριξη.

Η εργασία αυτή αφιερώνεται στους αγαπημένους μου Έκτορα και Άρη, που είναι πάντα δίπλα μου και με βοηθούν να βελτιώνω την ποιότητα της ζωής μου μέρα με τη μέρα.

Περίληψη

Η παρούσα εργασία βασίζεται στη σημασιολογική ανάλυση των ετικετών (annotations) μουσικών κομματιών, τα οποία συλλέχθηκαν από μια δημοφιλή εφαρμογή. Κάθε ετικέτα περιλαμβάνει έναν αριθμό κομματιών. Σκοπός της διπλωματικής εργασίας είναι να ταυτοποιήσουμε αν τα μουσικά κομμάτια έχουν τη σωστή ετικέτα συναισθηματικής κατηγοριοποίησης. Αυτή η ταυτοποίηση πραγματοποιείται εξάγοντας τα μουσικά χαρακτηριστικά των κομματιών. Τα χαρακτηριστικά αυτά υπολογίστηκαν με τη βοήθεια της βιβλιοθήκης του λογισμικού Python. Τα αποτελέσματα που εξήχθησαν αποτέλεσαν τη βάση για την τελική αξιολόγηση καθώς και για την παρουσίαση προτάσεων σε μελλοντικές προεκτάσεις και βελτιώσεις.

Abstract

This paper is based on the semantic analysis of music track annotations collected from a popular application. Each tag contains a number of tracks. The aim of the thesis is to identify whether the music tracks have the correct emotional categorization tag. This identification is performed by extracting the musical features of the tracks. These features were computed with the help of the Python software library. The extracted results formed the basis for the final evaluation as well as for presenting suggestions for future extensions and improvements.

Περιεχόμενα

Περίληψη	v
Abstract	vii
Περιεχόμενα	ix
Κατάλογος Εικόνων - Πινάκων	xi
Συντομογραφίες & Ακρωνύμια	xii
1. Εισαγωγή.....	11
1.1 Κίνητρο της Διπλωματικής Εργασίας	11
1.2 Σκοπός της Διπλωματικής Εργασίας	11
1.3 Διάρθρωση της ΔΕ	11
2. Θεωρητικό Υπόβαθρο	12
2.1 Τι είναι το συναίσθημα – Ορισμός	12
2.2 Κατηγοριοποίηση Συναισθημάτων	12
2.2.1 Κατηγορηματική προσέγγιση.....	12
2.2.2 Διαστατική προσέγγιση.....	14
2.3 Πως επηρεάζει η μουσική το συναίσθημα.....	17
2.3.1 Εισαγωγή.....	17
2.3.2 Προβλήματα στην αναγνώριση συναισθημάτων.....	18
2.3.2.1 Ασάφεια ή σύγχυση και διακρίσιμότητα.....	18
2.3.2.2 Απαιτούμενες γνώσεις.....	19
2.3.2.3 Υποκειμενικότητα συναισθηματικής αντίληψης.....	19
2.3.2.4 Σημαιολογικό χάσμα ανάμεσα στα χαμηλού επιπέδου ηχητικά σήματα και στην υψηλού επιπέδου ανθρώπινη αντίληψη.....	19
2.4 Μουσικά χαρακτηριστικά	20
2.4.1 Ενεργειακά χαρακτηριστικά.....	20
2.4.2 Χρονικά χαρακτηριστικά.....	23
2.4.3 Ρυθμικά χαρακτηριστικά.....	25
2.4.4 Αρμονικά χαρακτηριστικά.....	26
2.4.5 Φασματικά χαρακτηριστικά.....	27
3. Μεθοδολογία.....	33
3.1 Διαδικασία μελέτης.....	33
3.2 Συλλογή και επεξεργασία δεδομένων.....	33
3.4 Εξαγωγή μουσικών παραμέτρων.....	34
4. Αποτελέσματα.....	37
4.1 Εφαρμογή μεθόδου και Στατιστική ανάλυση	37
4.2 Συγκριτική ανάλυση αποτελεσμάτων.....	99
5. Συμπεράσματα και Μελλοντικές Προεκτάσεις.....	114
5.1 Αξιοποίηση Αποτελεσμάτων - Πιθανές Εφαρμογές	114
5.2 Μελλοντικές Προεκτάσεις - Βελτιώσεις	114
Βιβλιογραφία	115
Παράρτημα.....	121

Κατάλογος Εικόνων-Πινάκων

Εικόνα 2.1: Κυκλικό διάγραμμα με οκτώ κατηγορίες συναισθήματος.....	13
Εικόνα 2.2: Εννιά κατηγορίες συναισθήματος.....	13
Εικόνα 2.3: Πέντε κατηγορίες συναισθήματος της κοινότητας MIREX.....	14
Εικόνα 2.4: Κυκλικό μοντέλο με 28 συναισθήματα.....	15
Εικόνα 2.5: Μοντέλο Thayer.....	16
Εικόνα 2.6: Μοντέλο Yang βασισμένο στο μοντέλο του Thayer.....	16
Εικόνα 2.7: Κυκλικό μοντέλο με οκτώ κατηγορίες συναισθήματος.....	17
Εικόνα 2.8: Παράδειγμα απεικόνισης του Audio Power (AP).....	21
Εικόνα 2.9: Παράδειγμα απεικόνισης του Total Loudness (TL).....	22
Εικόνα 2.10: Παράδειγμα απεικόνισης των Specific Loudness Sensation Coefficients (SONE).....	23
Εικόνα 2.11: Παράδειγμα απεικόνισης του Zero Crossing Rate.....	24
Εικόνα 2.12: Παράδειγμα απεικόνισης του Spectral Centroid.....	27
Εικόνα 2.13: Παράδειγμα απεικόνισης του Spectral Rolloff.....	28
Εικόνα 2.14: Παράδειγμα απεικόνισης των MFCC.....	32
Εικόνα 3.1: Κατηγορίες συναισθήματος και μουσικά κομμάτια.....	34
Εικόνα 3.2: Μπλοκ διάγραμμα επεξεργασίας μουσικών χαρακτηριστικών.....	35
Εικόνα 4.1: Τυπική απόκλιση των χαρακτηριστικών για την κατηγορία angry.....	37
Εικόνα 4.2: Τυπική απόκλιση των χαρακτηριστικών για την κατηγορία angry.....	37
Εικόνα 4.3: Μέση τιμή των χαρακτηριστικών για την κατηγορία angry.....	38
Εικόνα 4.4: Μέση τιμή των χαρακτηριστικών για την κατηγορία angry.....	38
Εικόνα 4.5: Μέση τιμή των MFCC για την κατηγορία angry.....	39
Εικόνα 4.6: Διακύμανση των χαρακτηριστικών για την κατηγορία angry.....	39
Εικόνα 4.7: Διακύμανση των χαρακτηριστικών για την κατηγορία angry.....	40
Εικόνα 4.8: Κύρτωση των χαρακτηριστικών για την κατηγορία angry.....	40
Εικόνα 4.9: Κύρτωση των MFCC (1 ^{ος} έως 4 ^{ος}) για την κατηγορία angry.....	41
Εικόνα 4.10: Κύρτωση των MFCC (5 ^{ος} έως 8 ^{ος}) για την κατηγορία angry.....	41
Εικόνα 4.11: Κύρτωση του spectral contrast στις 6 ζώνες για την κατηγορία angry.....	42
Εικόνα 4.12: Λοξότητα των χαρακτηριστικών για την κατηγορία angry.....	42
Εικόνα 4.13: Λοξότητα του spectral flatness για την κατηγορία angry.....	43
Εικόνα 4.14: Λοξότητα των MFCC (οι πρώτοι 8) για την κατηγορία angry.....	43
Εικόνα 4.15: Λοξότητα των spectral contrast στις 6 ζώνες για την κατηγορία angry.....	44
Εικόνα 4.16: Τυπική απόκλιση των χαρακτηριστικών για την κατηγορία chill.....	44

Εικόνα 4.17: Τυπική απόκλιση των χαρακτηριστικών για την κατηγορία chill.....	45
Εικόνα 4.18: Μέση τιμή των χαρακτηριστικών για την κατηγορία chill.....	45
Εικόνα 4.19: Μέση τιμή των χαρακτηριστικών για την κατηγορία chill.....	46
Εικόνα 4.20: Μέση τιμή των MFCC για την κατηγορία chill.....	46
Εικόνα 4.21: Διακύμανση των χαρακτηριστικών για την κατηγορία chill.....	47
Εικόνα 4.22: Διακύμανση των χαρακτηριστικών για την κατηγορία chill.....	47
Εικόνα 4.23: Κύρτωση των χαρακτηριστικών για την κατηγορία chill.....	48
Εικόνα 4.24: Κύρτωση των MFCC (1 ^{ος} έως 4 ^{ος}) για την κατηγορία chill.....	48
Εικόνα 4.25: Κύρτωση των MFCC (5 ^{ος} έως 8 ^{ος}) για την κατηγορία chill.....	49
Εικόνα 4.26: Κύρτωση του spectral contrast στις 6 ζώνες για την κατηγορία chill.....	49
Εικόνα 4.27: Λοξότητα των χαρακτηριστικών για την κατηγορία chill.....	50
Εικόνα 4.28: Λοξότητα του spectral flatness για την κατηγορία chill.....	50
Εικόνα 4.29: Λοξότητα των MFCC (1 ^{ος} έως 4 ^{ος}) για την κατηγορία chill.....	51
Εικόνα 4.30: Λοξότητα των MFCC (5 ^{ος} έως 8 ^{ος}) για την κατηγορία chill.....	51
Εικόνα 4.31: Λοξότητα του spectral contrast στις 6 ζώνες για την κατηγορία chill.....	52
Εικόνα 4.32: Τυπική απόκλιση των χαρακτηριστικών για την κατηγορία dance.....	52
Εικόνα 4.33: Τυπική απόκλιση των χαρακτηριστικών για την κατηγορία dance.....	53
Εικόνα 4.34: Μέση τιμή των χαρακτηριστικών για την κατηγορία dance	53
Εικόνα 4.35: Μέση τιμή των χαρακτηριστικών για την κατηγορία dance.....	54
Εικόνα 4.36: Μέση τιμή των MFCC για την κατηγορία dance.....	54
Εικόνα 4.37: Διακύμανση των χαρακτηριστικών για την κατηγορία dance.....	55
Εικόνα 4.38: Διακύμανση των χαρακτηριστικών για την κατηγορία dance.....	55
Εικόνα 4.39: Κύρτωση των χαρακτηριστικών για την κατηγορία dance.....	56
Εικόνα 4.40: Κύρτωση των MFCC (1 ^{ος} έως 4 ^{ος}) για την κατηγορία dance.....	56
Εικόνα 4.41: Κύρτωση των MFCC (5 ^{ος} έως 8 ^{ος}) για την κατηγορία dance.....	57
Εικόνα 4.42: Κύρτωση του spectral contrast στις 6 ζώνες για την κατηγορία dance.....	57
Εικόνα 4.43: Λοξότητα των χαρακτηριστικών για την κατηγορία dance.....	58
Εικόνα 4.44: Λοξότητα του spectral flatness για την κατηγορία dance.....	58
Εικόνα 4.45: Λοξότητα των MFCC (1 ^{ος} έως 4 ^{ος}) για την κατηγορία dance.....	59
Εικόνα 4.46: Λοξότητα των MFCC (5 ^{ος} έως 8 ^{ος}) για την κατηγορία dance.....	59
Εικόνα 4.47: Λοξότητα του spectral contrast στις 6 ζώνες για την κατηγορία dance.....	60
Εικόνα 4.48: Τυπική απόκλιση των χαρακτηριστικών για την κατηγορία focus.....	60
Εικόνα 4.49: Τυπική απόκλιση των χαρακτηριστικών για την κατηγορία focus.....	61

Εικόνα 4.50: Μέση τιμή των χαρακτηριστικών για την κατηγορία focus.....	61
Εικόνα 4.51: Μέση τιμή των χαρακτηριστικών για την κατηγορία focus.....	62
Εικόνα 4.52: Μέση τιμή των MFCC για την κατηγορία focus.....	62
Εικόνα 4.53: Διακύμανση των χαρακτηριστικών για την κατηγορία focus.....	63
Εικόνα 4.54: Διακύμανση των χαρακτηριστικών για την κατηγορία focus.....	63
Εικόνα 4.55: Κύρτωση των χαρακτηριστικών για την κατηγορία focus.....	64
Εικόνα 4.56: Κύρτωση των MFCC (1 ^{ος} έως 4 ^{ος}) για την κατηγορία focus.....	64
Εικόνα 4.57: Κύρτωση των MFCC (5 ^{ος} έως 8 ^{ος}) για την κατηγορία focus.....	65
Εικόνα 4.58: Κύρτωση του spectral contrast στις 6 ζώνες για την κατηγορία focus.....	65
Εικόνα 4.59: Λοξότητα των χαρακτηριστικών για την κατηγορία focus.....	66
Εικόνα 4.60: Λοξότητα του spectral flatness για την κατηγορία focus.....	66
Εικόνα 4.61: Λοξότητα των MFCC (1 ^{ος} έως 4 ^{ος}) για την κατηγορία focus.....	67
Εικόνα 4.62: Λοξότητα των MFCC (5 ^{ος} έως 8 ^{ος}) για την κατηγορία focus.....	67
Εικόνα 4.63: Λοξότητα του spectral contrast στις 6 ζώνες για την κατηγορία focus.....	68
Εικόνα 4.64: Τυπική απόκλιση των χαρακτηριστικών για την κατηγορία happy.....	68
Εικόνα 4.65: Τυπική απόκλιση των χαρακτηριστικών για την κατηγορία happy.....	69
Εικόνα 4.66: Μέση τιμή των χαρακτηριστικών για την κατηγορία happy.....	69
Εικόνα 4.67: Μέση τιμή των χαρακτηριστικών για την κατηγορία happy.....	70
Εικόνα 4.68: Μέση τιμή των MFCC για την κατηγορία happy.....	70
Εικόνα 4.69: Διακύμανση των χαρακτηριστικών για την κατηγορία happy.....	71
Εικόνα 4.70: Διακύμανση των χαρακτηριστικών για την κατηγορία happy.....	71
Εικόνα 4.71: Κύρτωση των χαρακτηριστικών για την κατηγορία happy.....	72
Εικόνα 4.72: Κύρτωση των χαρακτηριστικών για την κατηγορία happy.....	72
Εικόνα 4.73: Κύρτωση των MFCC (1 ^{ος} έως 4 ^{ος}) για την κατηγορία happy.....	73
Εικόνα 4.74: Κύρτωση των MFCC (5 ^{ος} έως 8 ^{ος}) για την κατηγορία happy.....	73
Εικόνα 4.75: Κύρτωση του spectral contrast στις 6 ζώνες για την κατηγορία happy.....	74
Εικόνα 4.76: Λοξότητα των χαρακτηριστικών για την κατηγορία happy.....	74
Εικόνα 4.77: Λοξότητα των χαρακτηριστικών για την κατηγορία happy.....	75
Εικόνα 4.78: Λοξότητα του spectral flatness για την κατηγορία happy.....	75
Εικόνα 4.79: Κύρτωση των MFCC (1 ^{ος} έως 4 ^{ος}) για την κατηγορία happy.....	76
Εικόνα 4.80: Κύρτωση των MFCC (5 ^{ος} έως 8 ^{ος}) για την κατηγορία happy.....	76
Εικόνα 4.81: Λοξότητα του spectral contrast στις 6 ζώνες για την κατηγορία happy.....	77

Εικόνα 4.82: Τυπική απόκλιση των χαρακτηριστικών για την κατηγορία romance.....	78
Εικόνα 4.83: Τυπική απόκλιση των χαρακτηριστικών για την κατηγορία romance.....	78
Εικόνα 4.84: Μέση τιμή των χαρακτηριστικών για την κατηγορία romance.....	79
Εικόνα 4.85: Μέση τιμή των χαρακτηριστικών για την κατηγορία romance.....	79
Εικόνα 4.86: Μέση τιμή των MFCC για την κατηγορία romance.....	80
Εικόνα 4.87: Διακύμανση των χαρακτηριστικών για την κατηγορία romance.....	80
Εικόνα 4.88: Διακύμανση των χαρακτηριστικών για την κατηγορία romance.....	81
Εικόνα 4.89: Διακύμανση των χαρακτηριστικών για την κατηγορία romance.....	81
Εικόνα 4.90: Κύρτωση των χαρακτηριστικών για την κατηγορία romance.....	82
Εικόνα 4.91: Κύρτωση των χαρακτηριστικών για την κατηγορία romance.....	82
Εικόνα 4.92: Κύρτωση των MFCC (1 ^{ος} έως 4 ^{ος}) για την κατηγορία romance.....	83
Εικόνα 4.93: Κύρτωση των MFCC (5 ^{ος} έως 8 ^{ος}) για την κατηγορία romance.....	83
Εικόνα 4.94: Κύρτωση του spectral contrast (1 ^η έως 3 ^η ζώνη) για την κατηγορία romance.....	84
Εικόνα 4.95: Κύρτωση του spectral contrast (4 ^η έως 6 ^η ζώνη) για την κατηγορία romance.....	84
Εικόνα 4.96: Λοξότητα των χαρακτηριστικών για την κατηγορία romance.....	85
Εικόνα 4.97: Λοξότητα των χαρακτηριστικών για την κατηγορία romance.....	85
Εικόνα 4.98: Λοξότητα του spectral flatness για την κατηγορία romance.....	86
Εικόνα 4.99: Λοξότητα των MFCC (1 ^{ος} έως 4 ^{ος}) για την κατηγορία romance.....	86
Εικόνα 4.100: Λοξότητα των MFCC (5 ^{ος} έως 8 ^{ος}) για την κατηγορία romance.....	87
Εικόνα 4.101: Λοξότητα του spectral contrast (1 ^η έως 3 ^η ζώνη) για την κατηγορία romance.....	87
Εικόνα 4.102: Λοξότητα του spectral contrast (4 ^η έως 6 ^η ζώνη) για την κατηγορία romance.....	88
Εικόνα 4.103: Τυπική απόκλιση των χαρακτηριστικών για την κατηγορία sad.....	89
Εικόνα 4.104: Τυπική απόκλιση των χαρακτηριστικών για την κατηγορία sad.....	89
Εικόνα 4.105: Μέση Τιμή των χαρακτηριστικών για την κατηγορία sad.....	90
Εικόνα 4.106: Μέση Τιμή των χαρακτηριστικών για την κατηγορία sad.....	90
Εικόνα 4.107: Μέση Τιμή των MFCC για την κατηγορία sad.....	91
Εικόνα 4.108: Διακύμανση των χαρακτηριστικών για την κατηγορία sad.....	91
Εικόνα 4.109: Διακύμανση των χαρακτηριστικών για την κατηγορία sad.....	92
Εικόνα 4.110: Διακύμανση των χαρακτηριστικών για την κατηγορία sad.....	92
Εικόνα 4.111: Κύρτωση των χαρακτηριστικών για την κατηγορία sad.....	93
Εικόνα 4.112: Κύρτωση των χαρακτηριστικών για την κατηγορία sad.....	93
Εικόνα 4.113: Κύρτωση των MFCC (1 ^{ος} έως 4 ^{ος}) για την κατηγορία sad.....	94

Εικόνα 4.114: Κύρτωση των MFCC (5 ^{ος} έως 8 ^{ος}) για την κατηγορία sad.....	94
Εικόνα 4.115: Κύρτωση του spectral contrast (1 ^η έως 3 ^η ζώνη) για την κατηγορία sad.....	95
Εικόνα 4.116: Κύρτωση του spectral contrast (4 ^η έως 6 ^η ζώνη) για την κατηγορία sad.....	95
Εικόνα 4.117: Λοξότητα των χαρακτηριστικών για την κατηγορία sad.....	96
Εικόνα 4.118: Λοξότητα των χαρακτηριστικών για την κατηγορία sad.....	96
Εικόνα 4.119: Λοξότητα του spectral flatness για την κατηγορία sad.....	97
Εικόνα 4.120: Λοξότητα των MFCC (1 ^{ος} έως 4 ^{ος}) για την κατηγορία sad.....	97
Εικόνα 4.121: Λοξότητα των MFCC (5 ^{ος} έως 8 ^{ος}) για την κατηγορία sad.....	98
Εικόνα 4.122: Λοξότητα του spectral contrast (1 ^η έως 3 ^η ζώνη) για την κατηγορία sad.....	98
Εικόνα 4.123: Λοξότητα του spectral contrast (4 ^η έως 6 ^η ζώνη) για την κατηγορία sad.....	99
Εικόνα 4.124: Μέση τιμή όλων των συναισθημάτων για το zero crossing.....	100
Εικόνα 4.125: Μέση τιμή όλων των συναισθημάτων για το spectral contrast.....	101
Εικόνα 4.126: Μέση τιμή όλων των συναισθημάτων για το spectral centroid.....	102
Εικόνα 4.127: Μέση τιμή όλων των συναισθημάτων για το spectral bandwidth.....	103
Εικόνα 4.128: Μέση τιμή όλων των συναισθημάτων για το spectral flatness.....	104
Εικόνα 4.129: Μέση τιμή όλων των συναισθημάτων για το spectral rolloff.....	105
Εικόνα 4.130: Μέση τιμή όλων των συναισθημάτων για τους MFCC.....	106
Εικόνα 4.131: Μέση τιμή όλων των συναισθημάτων για τους RMS.....	107
Εικόνα 4.132: Μέση τιμή όλων των συναισθημάτων για τους Chroma STFT.....	108
Εικόνα 4.133: Μέσος όρος μέσης τιμής όλων των συναισθημάτων για το Zero Crossing.....	109
Εικόνα 4.134: Μέσος όρος μέσης τιμής όλων των συναισθημάτων για το Spectral Contrast.....	109
Εικόνα 4.135: Μέσος όρος μέσης τιμής όλων των συναισθημάτων για το Spectral Centroid.....	110
Εικόνα 4.136: Μέσος όρος μέσης τιμής όλων των συναισθημάτων για το Spectral Bandwidth.....	110
Εικόνα 4.137: Μέσος όρος μέσης τιμής όλων των συναισθημάτων για το Spectral Flatness.....	111
Εικόνα 4.138: Μέσος όρος μέσης τιμής όλων των συναισθημάτων για το Spectral Rolloff.....	111
Εικόνα 4.139: Μέσος όρος μέσης τιμής όλων των συναισθημάτων για τους MFCC.....	112
Εικόνα 4.140: Μέσος όρος μέσης τιμής όλων των συναισθημάτων για την RMS.....	112
Εικόνα 4.141: Μέσος όρος μέσης τιμής όλων των συναισθημάτων για το Chroma Stft.....	113

Συντομογραφίες & Ακρωνύμια

ΔΕ	Διπλωματική Εργασία
ΕΑΠ	Ελληνικό Ανοικτό Πανεπιστήμιο
AMC	Audio Mood Classification
AP	Audio Power
DCT	Discrete Cosine Transform
DWCH	Daubechies Wavelets Coefficient Histogram
FFT	Fast Fourier Transform
GMM	Gaussian Mixture Models
MA	Music Analysis Toolbox
MER	Music Emotion Recognition
MFCC	Mel-Frequency Cepstral Coefficient
MIR	Music Information Retrieval Toolbox
MIREX	Music Information Retrieval Evaluation eXchange
RMS	Root Mean Square
RP	Rhythm Pattern Extractor
SCF	Spectral Crest Factors
SDT	Sound Description Toolbox
SFM	Spectral Flatness Measures
SONE	Specific Loudness Sensation Coefficients
STFT	Short Time Fourier Transform
SWIPE	Sawtooth Waveform Inspired Pitch Estimator
TL	Total Loudness
VA	Valence-Arousal Diagram

1. Εισαγωγή

1.1 Κίνητρο της Διπλωματικής Εργασίας

Η συνεχής διάδοση της μουσικής μέσω του διαδικτύου καθώς και η δημιουργία εφαρμογών με δωρεάν πρόσβαση στο χρήστη οδήγησαν στην δημιουργία της συγκεκριμένης μελέτης. Διάφορες εφαρμογές όπως το Spotify, δίνουν τη δυνατότητα στο χρήστη να προσθέτει ετικέτα σε μουσικά κομμάτια χαρακτηρίζοντάς τα με κάποιο συναίσθημα. Αυτοί οι χαρακτηρισμοί βοηθούν άλλους χρήστες να επιλέξουν τα μουσικά κομμάτια που θέλουν να ακούσουν.

1.1.1 Αυτόματη αναγνώριση συναισθημάτων από μουσική

Αυτόματη αναγνώριση συναισθημάτων από μουσική είναι ένα πεδίο που απασχολεί τους ερευνητές εδώ και αρκετές δεκαετίες. Έχουν γίνει σημαντικά βήματα στην εξέλιξη αυτού του πεδίου με λογισμικά και γλώσσες προγραμματισμού.

1.2 Σκοπός της Διπλωματικής Εργασίας

Ο σκοπός της παρούσας ΔΕ, είναι η ταυτοποίηση συναισθημάτων από σχολιασμούς μουσικών κομματιών, εξάγοντας τα μουσικά χαρακτηριστικά τους.

1.3 Διάρθρωση της ΔΕ

Η ΔΕ δομείται σε κεφάλαια ως εξής:

Στο κεφάλαιο 2, δίδεται ένα θεωρητικό υπόβαθρο, κυρίως από την επιστήμη της Ψυχολογίας, όπου αναφέρεται η έννοια του συναισθήματος στον άνθρωπο και παρουσιάζονται οι σημαντικότερες θεωρίες κατηγοριοποίησης συναισθημάτων. Επιπλέον, αναλύονται τα προβλήματα στην αναγνώριση συναισθημάτων καθώς και τα μουσικά χαρακτηριστικά.

Στο κεφάλαιο 3, περιγράφεται η μεθοδολογία της μελέτης. Η μεθοδολογία περιλαμβάνει την συλλογή και επεξεργασία δεδομένων και την εξαγωγή των μουσικών παραμέτρων.

Στο κεφάλαιο 4, παρουσιάζεται η εφαρμογή της μεθόδου και η στατιστική ανάλυση καθώς επίσης και η συγκριτική ανάλυση αποτελεσμάτων.

Τέλος, στο κεφάλαιο 5, γίνεται σχολιασμός των πειραματικών αποτελεσμάτων και τρόποι αξιοποίησης της μεθόδου με αναφορά σε μελλοντικές προεκτάσεις για περαιτέρω βελτίωση.

2. Θεωρητικό Υπόβαθρο

‘Music is the shorthand of emotion’

Leo Tolstoy

2.1 Τι είναι το συναίσθημα – Ορισμός

Συναίσθημα είναι μία σύνθετη υποκειμενική συνειδητή εμπειρία. Συνδυάζει δηλαδή νοητικές και ψυχοσωματικές καταστάσεις και εκφράζεται ως κάτι εσωτερικό που επιδρά στο σώμα. Το συναίσθημα γίνεται αντιληπτό μέσω του προσώπου, της φωνής ή της στάσης του σώματος. Θα μπορούσαμε να αναφέρουμε ως βασικά συναισθήματα τα εξής: αγάπη, χαρά, έκπληξη, θυμός, λύπη και φόβος (Parrot, 2001· Wikipedia).

2.2 Κατηγοριοποίηση Συναισθημάτων

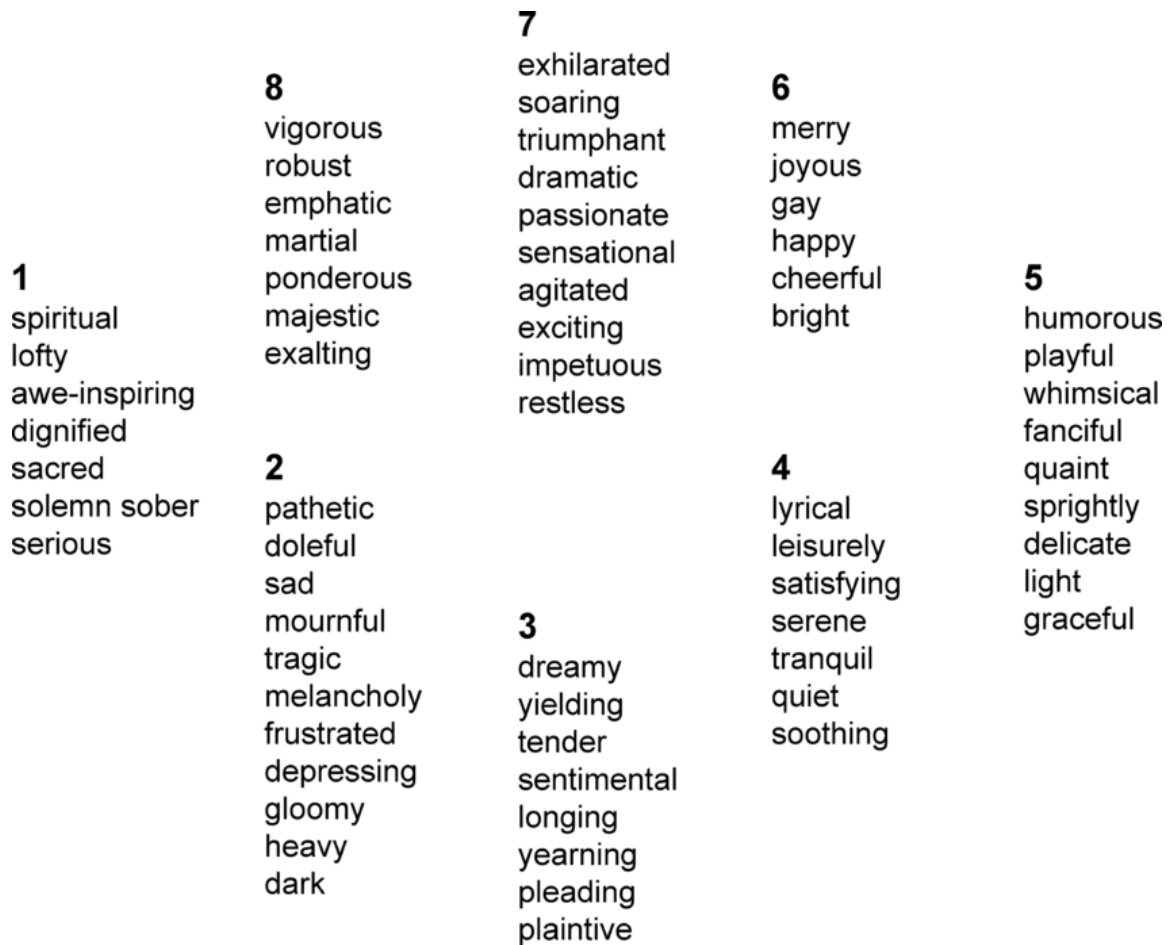
Για δεκαετίες, οι ψυχολόγοι μελετούν τη σχέση ανάμεσα στο συναίσθημα και τη μουσική. Τα θέματα που αντιμετώπισαν στην έρευνα τους αφορούσαν, το αν τα καθημερινά συναισθήματα είναι τα ίδια με εκείνα που λαμβάνονται από τη μουσική, με εκείνα τα συναισθήματα που λαμβάνονται από τον ακροατή ή με εκείνα που αισθάνεται ο ακροατής. Επίσης, κατά πόσο οι παράγοντες μουσικοί, προσωπικοί και καταστατικοί επηρεάζουν την συναισθηματική αντίληψη και πως πρέπει να αντιληφθούμε το μουσικό συναίσθημα (Juslin, Sloboda, 2001). Έτσι κατέληξαν σε δυο προσεγγίσεις για την μουσική αντίληψη, την κατηγορηματική και τη διαστατική.

2.2.1 Κατηγορηματική προσέγγιση

Σε αυτή την προσέγγιση ο ακροατής αντιλαμβάνεται τα συναισθήματα ως κατηγορίες διαφορετικές μεταξύ τους. Πρωταρχικό στοιχείο στην κατηγορηματική προσέγγιση είναι η έννοια των βασικών συναισθημάτων, δηλαδή η ιδέα του να υπάρχουν έμφυτες και παγκόσμιες κατηγορίες συναισθήματος όπως *anger*, *fear*, *sad*, *happiness*, από τις οποίες έχουν δημιουργηθεί όλες οι άλλες τάξεις συναισθήματος (Schuller, Hage, Schuller, Rigoll, 2010· Ekman, 1992). Ο ερευνητής Paul Eckman (Ekman, 1992), για παράδειγμα χρησιμοποίησε την εξής κατηγοριοποίηση, *anger*, *fear*, *sadness*, *happiness*, και *disgust*.

Μια τυπική κατηγορηματική προσέγγιση μπορεί να αποτελείται από επτά διακριτές τάξεις. Η αντίληψη των βασικών συναισθημάτων διαφοροποιείται. Διαφορετικοί ερευνητές εντόπισαν διαφορετικά σύνολα βασικών συναισθημάτων (Sloboda, Juslin, 2001). Ένα μεγάλο μειονέκτημά της προσέγγισης αυτής είναι ότι ο άνθρωπος αντιλαμβάνεται πολύ περισσότερα συναισθήματα από τον αριθμό των κύριων τάξεων συναισθημάτων που ορίζει αυτή. Επιπλέον, η χρήση μεγάλου αριθμού τάξεων συναισθημάτων μπερδεύει τους ακροατές και είναι μη πρακτική για ψυχολογικές μελέτες (Sloboda, Juslin, 2001).

Η πρώτη μελέτη, που βασίστηκε στην κατηγορηματική προσέγγιση έγινε από την Kate Hevner (1936), η οποία δημιούργησε ένα κυκλικό μοντέλο, χωρίζοντας τα συναισθήματα σε οκτώ κατηγορίες, *dignified*, *sad*, *dreamy*, *serene*, *graceful*, *happy*, *exciting*, *vigorous* (Fernandes, Paiva, 2010· Hevner, 1936). Το ίδιο μοντέλο χρησιμοποιήθηκε αργότερα και από τον Owen Craigie Meyers (2007).



Εικόνα 2.1: Κυκλικό διάγραμμα με οκτώ κατηγορίες συναισθήματος (Hevner, 1936)

Η Emery Schubert το 2003 χρησιμοποίησε επίσης, την κατηγορηματική προσέγγιση και δημιούργησε ένα νέο μοντέλο με εννιά κατηγορίες (Schubert, 2003).

<i>Cluster</i>	<i>Emotions in Each Cluster</i>
1	Bright, cheerful, happy, joyous
2	Humorous, light, lyrical, merry, playful
3	Calm, delicate, graceful, quiet, relaxed, serene, soothing, tender, tranquil
4	Dreamy, sentimental
5	Dark, depressing, gloomy, melancholy, mournful, sad, solemn
6	Heavy, majestic, sacred, serious, spiritual, vigorous
7	Tragic, yearning
8	Agitated, angry, restless, tense
9	Dramatic, exciting, exhilarated, passionate, sensational, soaring, triumphant

Εικόνα 2.2: Εννιά κατηγορίες συναισθήματος (Schubert, 2003· Yang, Chen, 2011)

Για πρώτη φορά το 2007, η κοινότητα MIREX (Music Information Retrieval Evaluation eXchange) παρουσίασε τη δική της ταξινόμηση συναισθημάτων, χωρίζοντας τα συναισθήματα σε πέντε κατηγορίες (Panda, Paiva, 2011· Cardoso, Panda, Paiva, 2011· Panda, Paiva, 2012· Panda, Paiva, 2012).

<i>Cluster</i>	<i>Description</i>
1	Passionate, rousing, confident, boisterous, rowdy
2	Rollicking, cheerful, fun, sweet, amiable/good-natured
3	Literate, poignant, wistful, bittersweet, autumnal, brooding
4	Humorous, silly, campy, quirky, whimsical, witty, wry
5	Aggressive, fiery, tense/anxious, intense, volatile, visceral

Εικόνα 2.3: Πέντε κατηγορίες συναισθήματος της κοινότητας MIREX (Hu, Downie, Laurier, Bay, Ehmann, 2008)

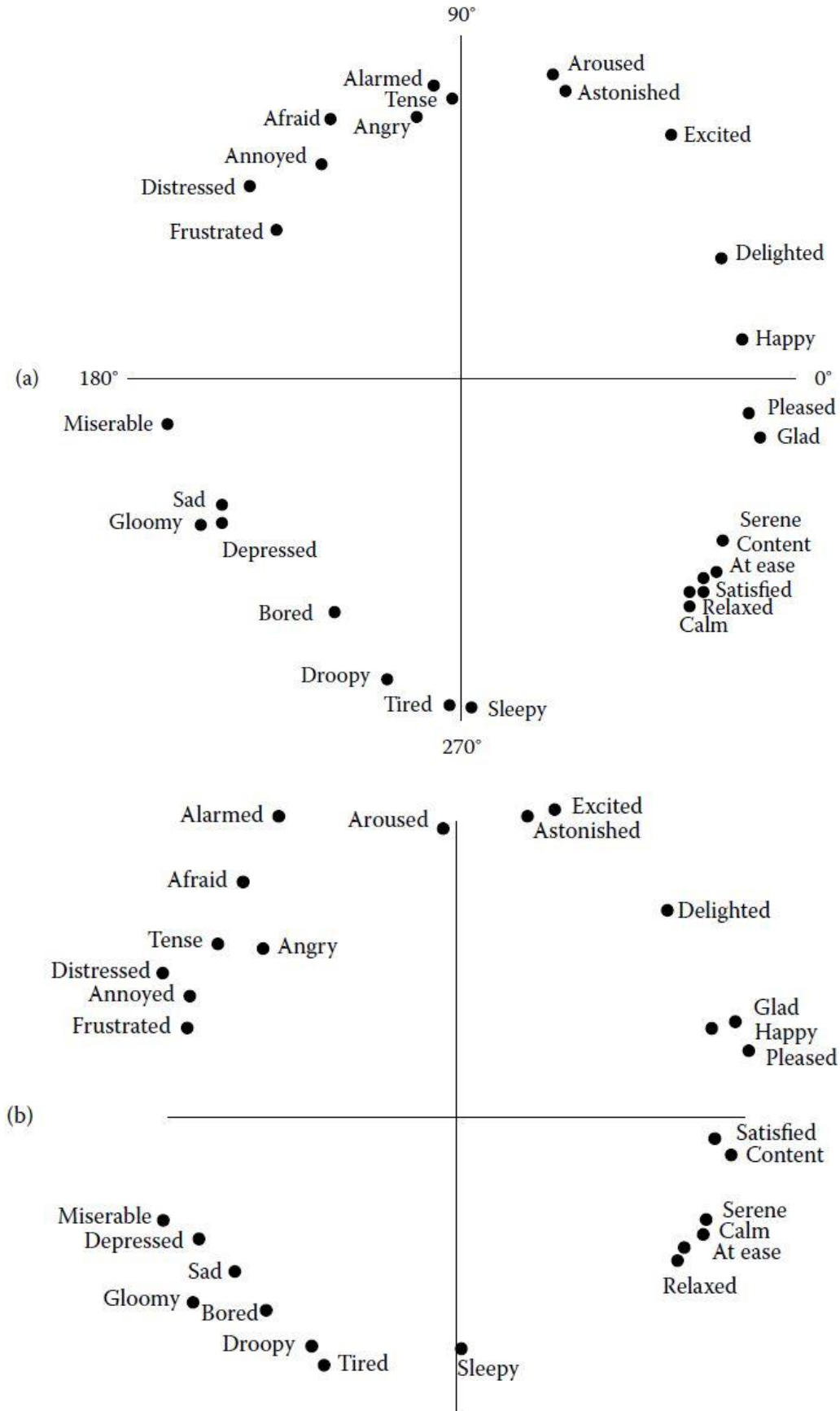
2.2.2 Διαστατική προσέγγιση

Η προσέγγιση αυτή εστιάζει σε ταυτοποιημένα συναισθήματα, βασισμένα στις θέσεις τους σε ένα επίπεδο συναισθημάτων με συντεταγμένες, οι οποίες αντιπροσωπεύουν τις εσωτερικές ανθρώπινες εκφράσεις του συναισθήματος. Αυτές οι εσωτερικές διαστάσεις συναισθήματος εντοπίζονται αναλύοντας τον συσχετισμό ανάμεσα σε συναισθηματικούς όρους. Αυτό πραγματοποιείται ζητώντας από ανθρώπους να σχολιάσουν συναισθηματικά ένα μουσικό κομμάτι με την χρήση μιας κλίμακας αξιολόγησης από ποικίλα συναισθήματα.

Στη συνέχεια, με τεχνικές ανάλυσης παραγόντων, συλλέγονται τα αποτελέσματα και αποκτάται ο συσχετισμός ανάμεσα στους συναισθηματικούς όρους. Στους δυο άξονες χρησιμοποιούνται οι μεταβλητές *valence* και *arousal*. Στους τρεις άξονες χρησιμοποιούνται οι μεταβλητές, *valence*, *arousal*, και *dominance*. Το *valence* ορίζει πόσο συναρπαστικό ή ήρεμο είναι ένα μουσικό κομμάτι, το *arousal* πόσο θετικό ή αρνητικό είναι το συναίσθημα που προκαλείται και το *dominance* δείχνει την δραστηριότητα ή τον έλεγχο του συναισθήματος (Scherer, 2004).

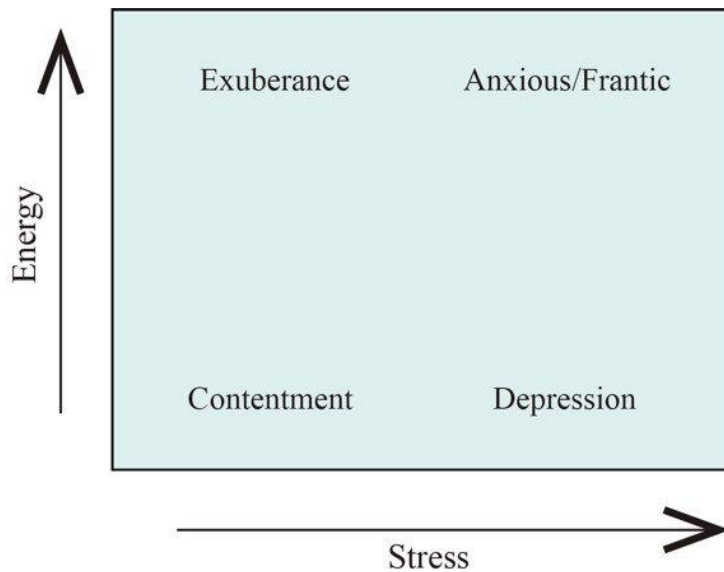
Ωστόσο, αυτή η προσέγγιση έχει επικριθεί. Υπάρχει μια διαφωνία από ερευνητές, διότι, μπερδεύει σημαντικές ψυχολογικές διαφορές και κατ' επέκταση καλύπτει σημαντικές πλευρές στη συναισθηματική διαδικασία (Ali, 2006). Για παράδειγμα, τα συναισθήματα *angry* και *fear* που ανήκουν στο δεύτερο τεταρτημόριο του διαγράμματος συντεταγμένων *valence-arousal* όπως δείχνει η παρακάτω εικόνα έχουν πολύ διαφορετικές επιπτώσεις στον οργανισμό. Το ίδιο ισχύει και για τα συναισθήματα *boredom* και *melancholy* (εικόνα παρακάτω).

Χρησιμοποιώντας το δισδιάστατο μοντέλο, δεν διακρίνονται τα μουσικά κομμάτια που παράγουν ένα συναίσθημα από ένα άλλο και δεν επιτρέπουν την θεωρητική εξέταση της προέλευσης και των μηχανισμών τέτοιων επιδραστικών επιπτώσεων (Sloboda, Juslin, 2001). Επιπλέον, υπάρχει και η διαφωνία ότι χρησιμοποιώντας λίγες διαστάσεις συναισθήματος δεν μπορούν να περιγραφούν όλα τα συναισθήματα χωρίς σημαντικά σφάλματα (Collier, 2007).



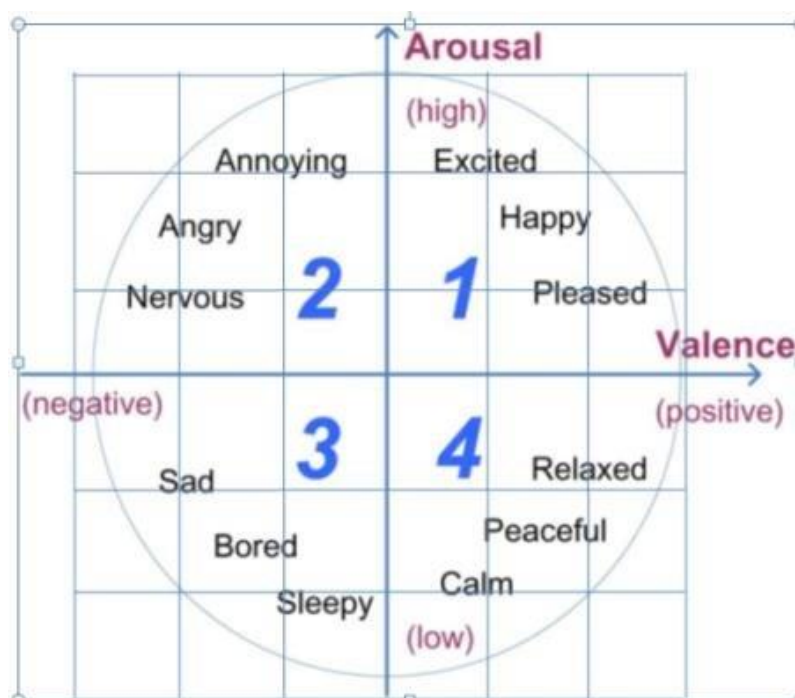
Εικόνα 2.4: Κυκλικό μοντέλο με 28 συναισθήματα (Russell, 1980· Yang, Chen, 2011)

Λίγα χρόνια αργότερα, ο ερευνητής Robert Thayer (1989) χρησιμοποίησε την ίδια προσέγγιση, δημιουργώντας ένα διάγραμμα συντεταγμένων με τις ίδιες μεταβλητές (Thayer, 1989· Fernandes, Paiva, 2010).



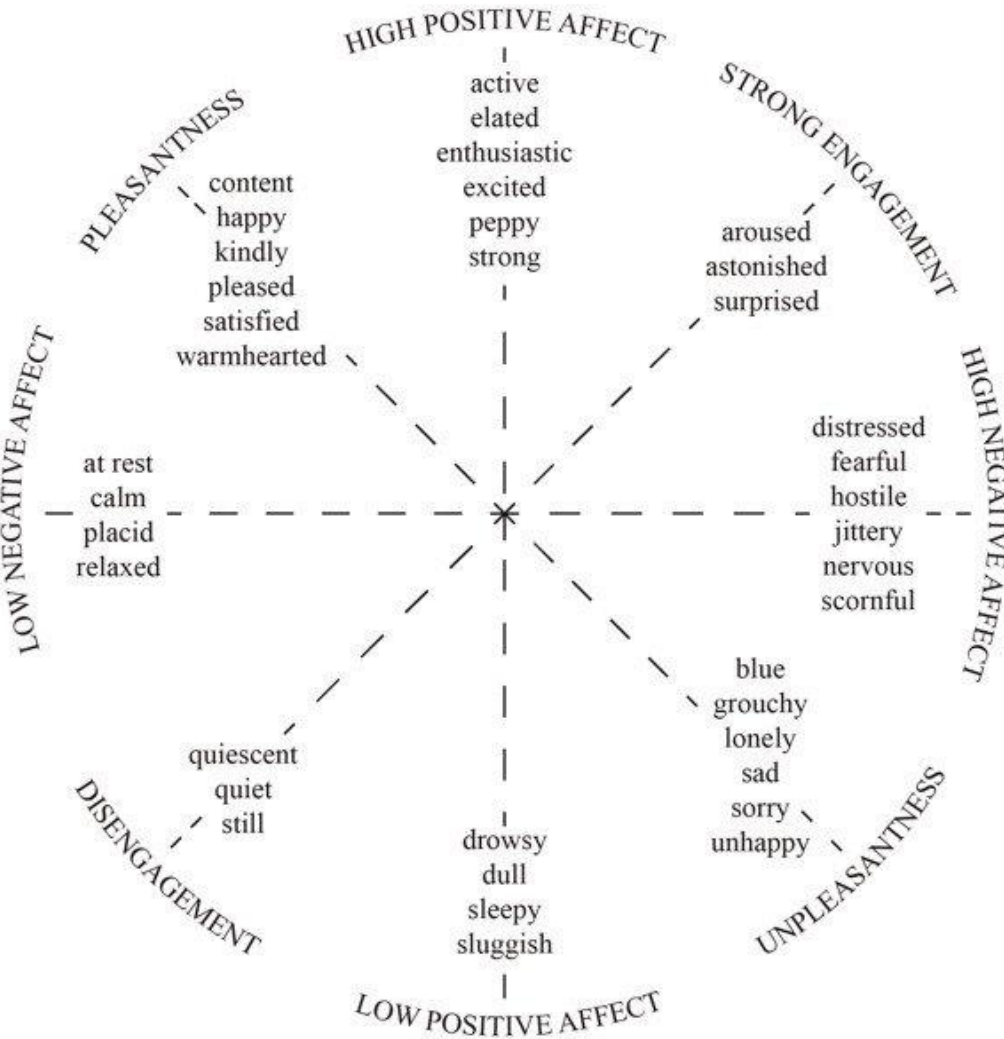
Εικόνα 2.5: Μοντέλο Thayer (Thayer, 1989· Fernandes, Paiva, 2010)

Ο Yang, για πρώτη φορά το 2007, χρησιμοποίησε τη διαστατική προσέγγιση βασισμένος στο μοντέλο του Thayer (Thayer, 1989· Yang, Liu, Chen, 2006· Yang, Lin, Chen, 2009· Yang, Su, Lin, Chen, 2007).



Εικόνα 2.6: Μοντέλο Yang βασισμένο στο μοντέλο του Thayer (2007) (Yang, Chen, 2012)

Οι ερευνητές Tellegen-Watson-Clark (1999), χρησιμοποίησαν την διαστατική προσέγγιση με έναν διαφορετικό τρόπο, δημιουργώντας ένα διάγραμμα με μεταβλητές *high* και *low positive affect* και *high* και *low negative affect* (Tellegen, Watson, Clark, 1999). Στο διάγραμμα αυτό βασίστηκε και ο Laar το 2006 (Tellegen, Watson, Clark, 1999· Laar, 2006· Fernandes, Paiva, 2010).



Εικόνα 2.7: Κυκλικό μοντέλο με οκτώ κατηγορίες συναισθήματος (Watson, Weiss, Vaidya, Tellegen, 1999)

2.3 Πως επηρεάζει η μουσική το συναίσθημα

2.3.1 Εισαγωγή

Η μουσική είναι ένα σημαντικό κομμάτι στη ζωή του ανθρώπου, ιδιαίτερα την σημερινή εποχή που η διάδοση της μουσικής είναι ευκολότερη και οι χρήστες έχουν άμεση και πιο γρήγορη πρόσβαση σε αυτήν. Η μελέτη συναισθηματικής απόκρισης σε μουσικά κομμάτια αποτελεί μία πρόκληση. Αυτό συμβαίνει, γιατί η συναισθηματική αντίληψη είναι υποκειμενική και οι άνθρωποι μπορούν να αντιληφθούν διαφορετικά συναισθήματα για το ίδιο κομμάτι. Το ζήτημα της υποκειμενικότητας είναι η βασική δυσκολία μιας τέτοιας μελέτης αφού δεν υπάρχει συμφωνία στο αποτέλεσμα της ταξινόμησης. Επιπλέον, είναι δύσκολη η περιγραφή ενός συναισθήματος γιατί τα επίθετα που το περιγράφουν είναι διφορούμενα και η χρήση των επιθέτων για το ίδιο συναίσθημα ποικίλει από άτομο σε άτομο. Τέλος, είναι ακόμα ανεξήγητο πώς η μουσική προκαλεί συναισθήματα. Είναι δύσκολο να κατανοηθεί, ποιο είναι αυτό το στοιχείο της μουσικής που δημιουργεί ένα ιδιαίτερο συναίσθημα στον ακροατή.

Η σχέση ανάμεσα σε μουσική και συναίσθημα αποτελεί ένα θέμα που έχει απασχολήσει πολλούς ερευνητές από διάφορους τομείς όπως ανθρωπολογία, φιλοσοφία, ψυχολογία, μουσικολογία, κοινωνιολογία, βιολογία (Sloboda, Juslin, 2001). Ένας μεγάλος αριθμός από μελέτες, έχει γίνει στην ψυχολογία. Σε αυτές τις μελέτες, το συναίσθημα κατηγοριοποιείται σε λίστες: εκφραζόμενο,

λαμβάνόμενο και το προκαλούμενο (Sloboda, Juslin, 2001· Gabrielsson, 2002· Hallam, Cross, Thaut, 2008· Huron, 2006). Το εκφραζόμενο αναφέρεται στο συναίσθημα με το οποίο ο ερμηνευτής προσπαθεί να επικοινωνήσει με τους ακροατές. Αντίθετα, το λαμβανόμενο και το προκαλούμενο αναφέρονται στην επιδραστική ανταπόκριση των ακροατών. Τα δυο τελευταία, ιδιαίτερα το προκαλούμενο εξαρτάται από την αλληλεπίδραση παραγόντων όπως μουσικοί, προσωπικοί και καταστατικοί. Το λαμβανόμενο συναίσθημα επηρεάζεται λιγότερο από καταστατικούς παράγοντες.

Μια τυπική προσέγγιση ενός συστήματος αναγνώρισης συναισθήματος από μουσική κατηγοριοποιεί τα συναισθήματα σε ομάδες-κατηγορίες (όπως *χαρούμενος*, *θυμωμένος*, *λυπημένος* και *χαλαρωμένος*) και εφαρμόζει τεχνικές αναγνώρισης προτύπων για να εκπαιδεύσει έναν ταξινομητή (Hu, Downie, Laurier, Bay, Ehmann, 2008· Katayose, Imai, Inokuchi, 1998· Laurier, Grivolla, Herrera, 2008· Li, Ogihara, 2003· Liu, Yang, Wu, Chen, 2006· Livingstone, Brown, 2005· Schuller, Hage, Schuller, Rigoll, 2010· Skowronek, McKinney, Van de Par, 2006· Trohidis, Tsoumakas, Kalliris, Vlahavas, 2008· Wang, Zhang, Zhu, 2004· Wu, Jeng, 2008). Άλλες πάλι, λόγω της ασάφειας των χαρακτηρισμών, εφαρμόζουν τεχνικές που ορίζουν τις συναισθηματικές τάξεις με ένα διάγραμμα valence-arousal. Το valence ορίζει πόσο συναρπαστικό ή ήρεμο είναι ένα τραγούδι ενώ το arousal πόσο θετικό ή αρνητικό είναι το συναίσθημα που προκαλεί. Ένα αντιπροσωπευτικό παράδειγμα είναι το διάγραμμα VA του Thayer που χωρίζει τις τάξεις των συναισθημάτων σε τεταρτημόρια (Chang, Lo, Wang, Chung, 2010), όπως φαίνεται στην ενότητα 2.2.2 (Εικόνα 2.6). Παρ' όλα αυτά υπάρχει μια ασάφεια στην ταξινόμηση των συναισθημάτων σε κατηγορίες. Για παράδειγμα στο πρώτο τεταρτημόριο, βρίσκονται τα συναισθήματα *exciting*, *happy* και *pleasure* τα οποία είναι διαφορετικά εκ φύσεως. Η ασάφεια αυτή, μπερδεύει τόσο τους ακροατές στα υποκειμενικά τεστ όσο και τους χρήστες όταν ανακτούν ένα μουσικό κομμάτι σύμφωνα με την συναισθηματική τους κατάσταση.

2.3.2 Προβλήματα στην αναγνώριση συναισθημάτων

Η αναγνώριση συναισθημάτων αποτελεί μία περίπλοκη διαδικασία. Προκύπτουν, τέσσερα πολύ σημαντικά ζητήματα που πρέπει να επιλυθούν. Η ασάφεια και διακριτότητα της περιγραφής του συναισθήματος, το βαρύ γνωστικό φορτίο του σχολιασμού συναισθήματος, η υποκειμενικότητα της συναισθηματικής αντίληψης και τέλος το σημασιολογικό χάσμα ανάμεσα στα χαμηλού επιπέδου ηχητικά σήματα και στην υψηλού επιπέδου ανθρώπινη αντίληψη (Yang, Chen, 2011).

2.3.2.1 Ασάφεια ή σύγχυση και διακριτότητα

Η ασάφεια ή σύγχυση είναι ένα χαρακτηριστικό της φυσικής γλώσσας των κατηγοριών (Hofmann, 1999). Τα συναισθήματα είναι πολύ συγκεχυμένες έννοιες. Για να αποφευχθεί αυτή η ασάφεια των επιδραστικών όρων και να μειωθεί η προσπάθεια του αναπτυγμένου συστήματος, πολλοί ερευνητές χρησιμοποίησαν τα βασικά συναισθήματα όπως *happy*, *sad*, *angry* και *relaxing* ή ομάδες συναισθημάτων, ταξινομώντας τα σε κατηγορίες (Lu, Liu, Zhang, 2006· Wang, Zhang, Zhu, 2004· Feng, Zhuang, Pan 2003· Hu, Downie, 2007· Yang, Lee, 2004). Η διακριτότητα αναφέρεται στον αριθμό των τάξεων των συναισθημάτων. Δηλαδή η ταξινόμηση του συναισθήματος είναι πολύ μικρότερη σε σύγκριση με την πληθώρα συναισθημάτων που αισθάνεται ο ακροατής. Αυτό είναι ανεπιθύμητο επειδή ένα σύστημα σύστασης με περιορισμένο λεξιλόγιο συναισθημάτων, ίσως δεν ικανοποιήσει τις απαιτήσεις του χρήστη στον πραγματικό κόσμο των εφαρμογών μουσικής ανάκτησης. Παρ' όλα αυτά, χρησιμοποιώντας μια λεπτομερή περιγραφή των συναισθημάτων, δεν είναι απαραίτητο να διευθετηθεί η διακριτότητα, επειδή μειώνεται η ασάφεια μεταξύ των συναισθηματικών όρων (Bartoszewski, Kwasnicka, Markowska-Kaczmar, Myszkowski, 2008) και

επειδή η ανάπτυξη ενός αυτόματου συστήματος που ταξινομεί με μικρό σφάλμα την μουσική σε ένα μεγάλο αριθμό τάξεων είναι πολύ δύσκολο (Yang, Chen, 2011).

2.3.2.2 Απαιτούμενες γνώσεις

Για να συλλέξουμε το ground truth (ακρίβεια του τρόπου εκπαίδευσης του συστήματος) εκπαιδεύοντας ένα αυτόματο μοντέλο, μια υποκειμενική δοκιμή τυπικά γίνεται προσκαλώντας εθελοντές να χαρακτηρίσουν συναισθηματικά ένα μουσικό κομμάτι. Για να μειωθεί ο χρόνος εξαγωγής των αποτελεσμάτων, κάθε μουσικό κομμάτι χαρακτηρίζεται το λιγότερο από τρεις εθελοντές (Li, Ogihara, 2003· Lu, Liu, Zhang, 2006· Skowronek, McKinney, Van de Par, 2006· Trohidis, Tsoumakas, Kalliris, Vlahavas, 2008). Αυτή η πρακτική είναι προβληματική γιατί τα καθημερινά σχόλια αυτών που έχουν γνώσεις πάνω στο αντικείμενο είναι πολύ διαφορετικά από αυτά των εθελοντών που δεν έχουν καμία γνώση και απαιτείται ξεχωριστή μεταχείριση (Sloboda, O'Neill, Ivaldi, 2001]. Το γνωστικό φορτίο συλλογής ετικετών για τα συναισθήματα ενός κατηγορηματικού συστήματος έχει λυθεί πρόσφατα με την αύξηση ιστοσελίδων που καταχωρούν οι χρήστες ετικέτες στα συναισθήματα όπως το All Music Guide (<http://www.allmusic.com/>), ο Last.fm (<http://www.last.fm/>) και το Sporify. Αντίθετα στο διαστατικό σύστημα απαιτούνται αξιολογήσεις συναισθήματος στο επίπεδο που ορίζεται από τους δύο άξονες valence και arousal όπως περιγράφηκε στην ενότητα 2.3.2. Αυτές οι αξιολογήσεις δεν αποκτούνται διαδικτυακά αλλά εντοπίζονται χρησιμοποιώντας είτε σταθερή κλίμακα αξιολόγησης είτε γραφική κλίμακα αξιολόγησης, δίνοντας στους εθελοντές τις απαραίτητες γνώσεις. Είναι δύσκολο να βεβαιώσουμε μια σταθερή κλίμακα αξιολόγησης ανάμεσα σε διαφορετικούς εθελοντές αλλά και στον ίδιο εθελοντή (Ovadia, 2004). Ως αποτέλεσμα, η ποιότητα των τιμών του ground truth, μπορεί να υποβαθμίσει την ακρίβεια του συστήματος (Yang, Chen, 2011).

2.3.2.3 Υποκειμενικότητα συναισθηματικής αντίληψης

Η μουσική αντίληψη είναι υποκειμενική και εξαρτάται από πολλούς παράγοντες όπως η ηλικία, το γένος, η προσωπικότητα και άλλα (Abeles, Chung, 1996). Οι επιδράσεις ανάμεσα στην μουσική και τον ακροατή εξαρτώνται από τις προτιμήσεις του πάνω στη μουσική (Holbrook, Schindler, 1989· Jargreaves, North, 1997) και από την οικειότητά του με αυτήν (Jargreaves, North, 1997). Εξαιτίας της υποκειμενικότητας είναι δύσκολο να υπάρξει ομοφωνία σχετικά με το ποιος συναισθηματικός όρος χαρακτηρίζει καλύτερα ένα μουσικό κομμάτι. Στην κατηγορηματική προσέγγιση που κάθε μουσικό κομμάτι χαρακτηρίζεται από μια τάξη συναισθήματος, το θέμα της υποκειμενικότητας δεν αντιμετωπίζεται σωστά. Αντίθετα, στην διαστατική προσέγγιση το πρόβλημα λύνεται επειδή κάθε χρήστης απαντά διαφορετικά στο ίδιο κομμάτι. Παρά το γεγονός ότι η υποκειμενική φύση της συναισθηματικής αντίληψης είναι αναγνωρισμένη, δεν έχουν γίνει αρκετές προσπάθειες για την επίλυση του θέματος (Yang, Chen, 2011).

2.3.2.4 Σημασιολογικό χάσμα ανάμεσα στα χαμηλού επιπέδου ηχητικά σήματα και στην υψηλού επιπέδου ανθρώπινη αντίληψη

Εξαιτίας αυτού του χάσματος είναι δύσκολο να υπολογιστούν με ακρίβεια οι τιμές του συναισθήματος, ιδιαίτερα του valence (Lu, Liu & Zhang, 2006· Schubert, 1999· Tolos, Tato, Kemp, 2005· Yang, Lin, Su, Chen 2008). Το στοιχείο αυτό της μουσικής που προκαλεί ένα συναίσθημα στον ακροατή, δεν έχει κατανοηθεί πλήρως. Κατά συνέπεια, η απόδοση των συμβατικών μεθόδων που αξιοποιούν μόνο τα χαρακτηριστικά ήχου χαμηλού επιπέδου φαίνεται να έχουν φτάσει σε ένα όριο. Για παράδειγμα, στην Audio Mood Classification (AMC), έχει πραγματοποιηθεί από το 2007 το MIREX (Music Information Retrieval Evaluation Exchange), με σκοπό την προώθηση της έρευνας

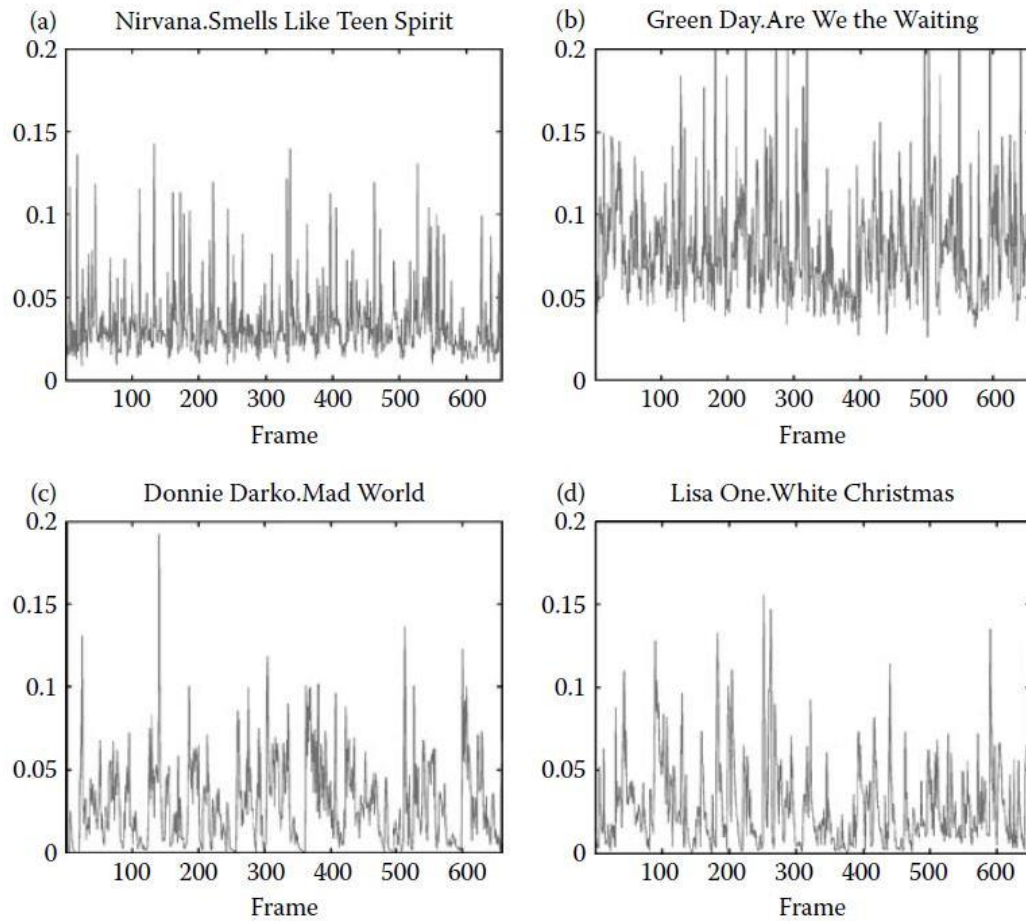
του συστήματος Music Emotion Recognition (MER) και την παροχή συγκρίσεων αναφοράς (Panda, Malheiro, Rocha, Oliveira, Paiva, 2008· Panda, Paiva, 2012· Panda, Paiva, 2011· Hu, Downie, Laurier, Bay, Ehmann, 2008). Χρησιμοποιούνται πέντε ομάδες συναισθήματος: *passionate, rollicking, literate, humorous και aggressive* (Hu, Downie, 2007). Τα ποσοστά ακρίβειας της ταξινόμησης από το 2007 έως το 2010 δεν ξεπέρασαν ποτέ το 70% (Yang, Chen, 2011).

2.4 Μουσικά χαρακτηριστικά

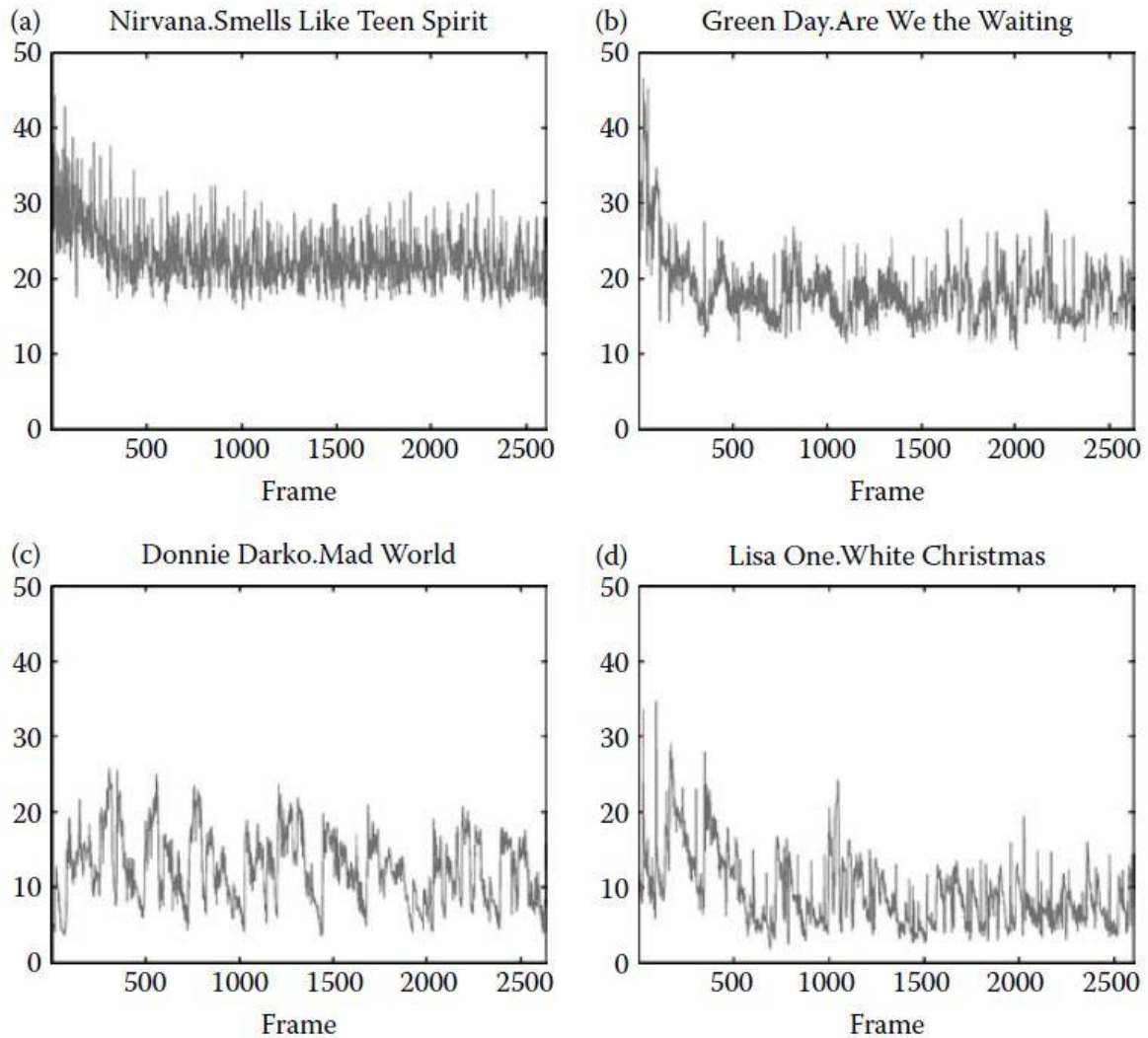
Η εμπειρία του να ακούς μουσική είναι πολυδιάστατη. Για παράδειγμα το arousal έχει σχέση με το ρυθμό, τον τόνο, την ένταση και το ηχόχρωμα, ενώ το valence έχει σχέση με την λειτουργία και την αρμονία (Gabrielsson, Lindstrom, 2001). Η συναισθηματική αντίληψη σχετίζεται με τον συνδυασμό πολλών παραγόντων της μουσικής (Hevner, 1935· Rigg, 1964). Δηλαδή, μουσικά κομμάτια με υψηλής έντασης συγχορδίες έχουν πιο θετικό valence σε αντίθεση με τις συγχορδίες χαμηλής έντασης. Τα μουσικά χαρακτηριστικά εξάγονται για να αναπαραστήσουν τον ακουστικό χαρακτήρα ενός μουσικού κομματιού. Μπορούν να χωριστούν στις εξής κατηγορίες: ενεργειακά, διαχρονικά, ρυθμικά, αρμονικά και φασματικά (Yang, Chen, 2011).

2.4.1 Ενεργειακά χαρακτηριστικά

Η ενέργεια σχετίζεται σε μεγάλο βαθμό με το arousal (Gabrielsson, Lindstrom, 2001). Η μέτρηση της λαμβανόμενης έντασης γίνεται με την χρήση του μοντέλου δυναμικής έντασης (*Dynamic Loudness*) των Chalupper και Fastl (Chalupper, Fastl, 2002) και εφαρμόζεται στο PsySound (Cabrera, 1999· Ricard, 2004). Το PsySound είναι ένα πρόγραμμα υπολογιστή που οι παράμετροι των μοντέλων της ακουστικής αίσθησης βασίζονται σε ψυχο-ακουστικά μοντέλα. Τέτοια μοντέλα είναι, η Bark critical band (Zwicker, 1961) για μοντελοποίηση ακουστικών φίλτρων στα αυτιά μας, το ακουστικό μοντέλο χρονικής ολοκλήρωσης, και το μοντέλο για την μοντελοποίηση της ευκρίνειας των Zwicker και Fastl (υποκειμενική μέτρηση του ήχου σε μια επεκτατική κλίμακα από το πολύ βαρετό στο οξύ) (Zwicker, Fastl, 1999). Η Sound Description Toolbox (SDT) εξάγει έναν αριθμό MPEG-7 Standard Descriptors και άλλους από ηχητικά σήματα (Benetos, Kotti, Kotropoulos, 2007). Τα βασικά ενεργειακά χαρακτηριστικά είναι το *Audio Power (AP)*, *Total Loudness (TL)*, και *Specific Loudness Sensation Coefficients (SONE)*. Το AP είναι η δύναμη του ηχητικού σήματος. Η εξαγωγή των Total Loudness (TL) και Specific Loudness Sensation Coefficients (SONE) βασίζεται στα αντιληπτικά μοντέλα που εφαρμόζονται στην εργαλειοθήκη μουσικής ανάλυσης (Music Analysis toolbox - MA toolbox) (Pampalk, 2004), συμπεριλαμβάνοντας το Outer Ear Model, την Bark Critical-Band Rate Scale και την φασματική συγκάλυψη (spectral masking). Το φάσμα ισχύος που προκύπτει, αντικατοπτρίζει την ένταση της ανθρώπινης αίσθησης και ονομάζεται Sonogram (Pampalk, Rauber, Merkl, 2002). Οι Specific Loudness Sensation Coefficients (SONE) είναι οι συντελεστές που υπολογίζονται από το sonogram, το οποίο αποτελείται από έως και 24 Bark Critical Bands (ο ακριβής αριθμός εξαρτάται από την συχνότητα δειγματοληψίας του ηχητικού σήματος). Το Total Loudness (TL) υπολογίζεται ως μια συνάθροιση βασισμένη στη μέθοδο του Stevens (Hartmann, 1998), η οποία παίρνει το σύνολο του μεγαλύτερου SONE συντελεστή και μια αναλογία 0,15 του συνόλου των υπολειπόμενων συντελεστών. Στις παρακάτω εικόνες δίνονται παραδείγματα των ενεργειακών χαρακτηριστικών για τέσσερα τυχαία κομμάτια. Το επίπεδο έντασης των τεσσάρων κομματιών έχει κανονικοποιηθεί με τον καθορισμό της υψηλότερης έντασης της κυματομορφής ήχου. Στις εικόνες 2.8 και 2.9 παρουσιάζονται η ηχητική ισχύς και η συνολική ένταση του ήχου για κάθε καρέ των τεσσάρων κομματιών, αντίστοιχα.

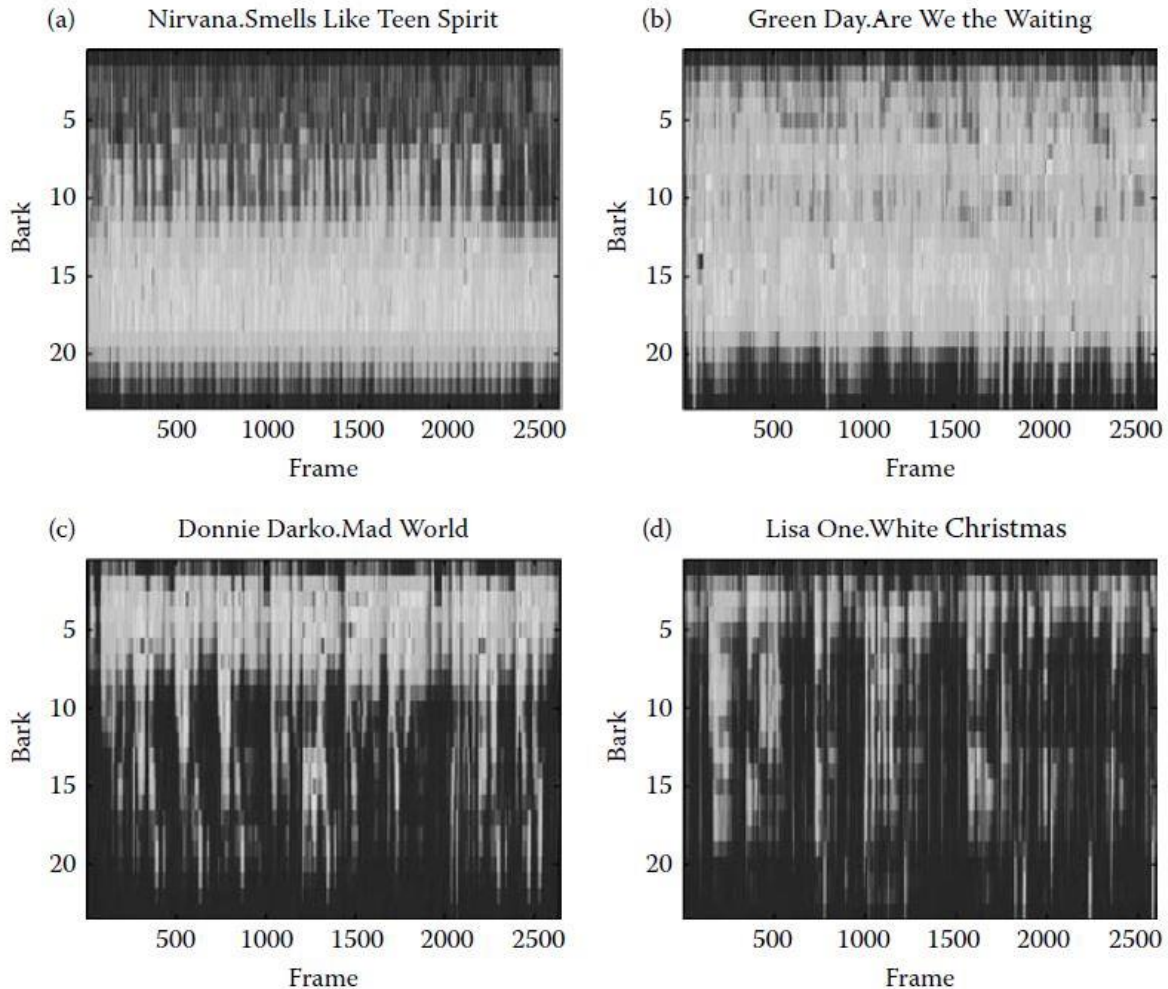


Εικόνα 2.8: Παράδειγμα απεικόνισης του Audio Power (AP) σε τέσσερα μουσικά κομμάτια (Benetos, Kotti, Kotropoulos, 2007· Pampalk, 2004· Yang, Chen, 2011)



Εικόνα 2.9: Παράδειγμα απεικόνισης του Total Loudness (TL) σε τέσσερα μουσικά κομμάτια (Benetos, Kotti, Kotropoulos, 2007· Pampalk, 2004· Yang, Chen, 2011)

Από τις Εικόνες 2.8 και 2.9 παρατηρείται ότι αυτά τα δύο χαρακτηριστικά σχετίζονται αρκετά με την αντίληψη της διέγερσης των μουσικών κομματιών. Τα κομμάτια (α) και (β) έχουν υψηλότερο Audio Power και Total Loudness σε σχέση με τα κομμάτια (γ) και (δ). Παράλληλα, το κομμάτι (β) έχει υψηλότερο Audio Power από το τραγούδι (α). Στην εικόνα 2.10 παρουσιάζονται οι συντελεστές SONE των τεσσάρων κομματιών σε 23 Bark Critical Bands. Ο μεγαλύτερος δείκτης της κλίμακας Bark αντιστοιχεί στην υψηλότερη ζώνη συχνοτήτων.



Εικόνα 2.10: Παράδειγμα απεικόνισης των Specific Loudness Sensation Coefficients σε τέσσερα μουσικά κομμάτια (SONE) (Benetos, Kotti, Kotropoulos, 2007· Pampalk, 2004· Yang, Chen, 2011)

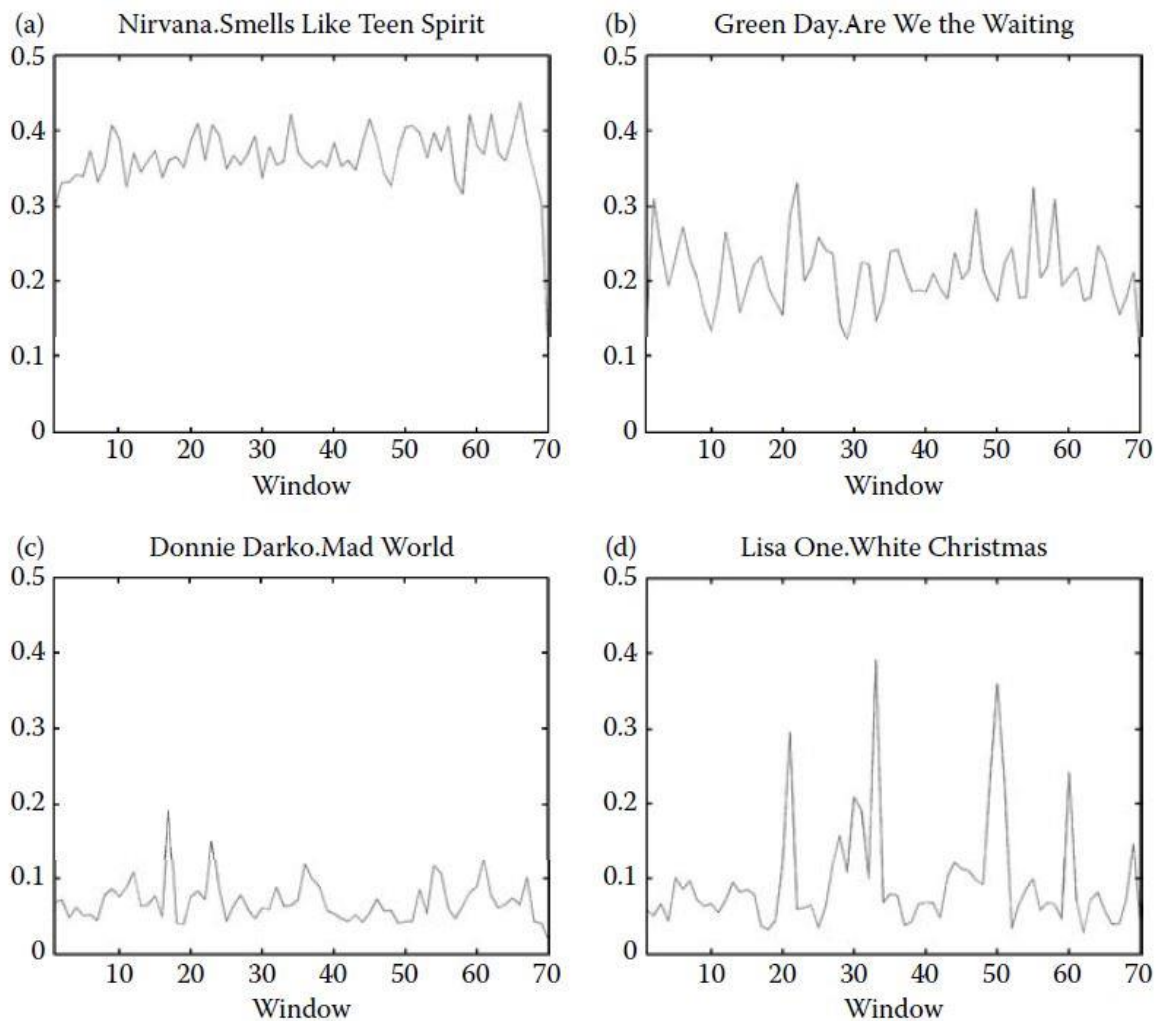
Από την εικόνα 2.10 παρατηρείται η ενέργεια του κομματιού (α) είναι ιδιαίτερα υψηλή στις υψηλές συχνότητες, ενώ στο κομμάτι (γ) είναι ιδιαίτερα υψηλή στη χαμηλή συχνότητα. Οι κατανομές της ενέργειας των τραγουδιών (β) και (δ) είναι σχετικά ομοιόμορφες σε όλες τις ζώνες συχνοτήτων. Αυτό το αποτέλεσμα δείχνει ότι οι συντελεστές SONE των κομματιών διαφορετικών συναισθημάτων μπορεί να έχουν διαφορετικά χαρακτηριστικά.

2.4.2 Χρονικά χαρακτηριστικά

Στα χρονικά χαρακτηριστικά περιλαμβάνονται τα *Zero Crossing Rate*, *Temporal Centroid*, *Log Attack Time* και *Root Mean Square (RMS)* για να βρεθεί η χρονική ποιότητα της μουσικής και εξάγονται από την εργαλειοθήκη Sound Description (Sound Description toolbox – SDT) (Benetos, Kotti, Kotropoulos, 2007). Το Zero Crossing Rate ή αλλιώς αριθμός μηδενικών διαβάσεων, μετρά το θόρυβο του σήματος που υπολογίζεται παίρνοντας την μέση και σταθερή απόκλιση του αριθμού των τιμών του σήματος που περνούν από το άξονα μηδέν σε κάθε χρόνο. Το Zero Crossing Rate υπολογίζεται από την παρακάτω σχέση:

$$\text{zero-crossing rate} = \frac{1}{T} \sum_{t=m-T+1}^m \frac{|\text{sgn}(s_t) - \text{sgn}(s_{t-1})|}{2} w(m-t),$$

Όπου το T είναι το μήκος του παράθυρου του χρόνου είναι το πλάτος των t δειγμάτων χρόνου και $w(\cdot)$ είναι ένα ορθογώνιο παράθυρο. Το αρχικό μήκος του παραθύρου του χρόνου είναι 1,6% του συνολικού μήκους του εισαγόμενου σήματος με 10% επικάλυψη (Magalhaes, Barros, Loureiro, 2019). Αξίζει να σημειωθεί ότι, η σταθερή διακύμανση του Zero Crossing Rate μπορεί να χρησιμεύσει στην πρόβλεψη του valence (Yang, Chen, 2011). Η παρακάτω εικόνα απεικονίζει ένα παράδειγμα του Zero Crossing Rate για τέσσερα μουσικά κομμάτια.



Εικόνα 2.11: Παράδειγμα απεικόνισης του Zero Crossing Rate σε τέσσερα μουσικά κομμάτια (Benetos, Kotti, Kotropoulos, 2007· Yang, Chen, 2011)

Από την Εικόνα 2.11 διαπιστώνεται ότι το κομμάτι (α) έχει το υψηλότερο Zero Crossing Rate και το αμέσως επόμενο παρατηρείται στο κομμάτι (β). Από αυτό το αποτέλεσμα φαίνεται ότι το Zero Crossing Rate είναι χρήσιμο. Επιπλέον, διαπιστώνουμε ότι το Zero Crossing Rate για το κομμάτι (γ) είναι πολύ μικρότερο από του κομματιού (δ). Το υψηλότερο Zero Crossing Rate υποδεικνύει μεγαλύτερο ποσοστό θορύβου στο σήμα. Σύμφωνα με την εικόνα 2.11, το κομμάτι (α) έχει μεγαλύτερο ποσοστό θορύβου.

Τα Temporal Centroid και Log Attack Time είναι δυο MPEG-7 harmonic instrument timbre περιγραφείς που περιγράφουν την ενέργεια του σήματος (Allamanche, Herre, Helmuth, Froba, Kasten, Cremer, 2001). Και τα δυο χρησιμοποιούνται στην ταξινόμηση μουσικού οργάνου. Η Root Mean Square ή αλλιώς ενεργός τιμή ορίζεται ως η τετραγωνική ρίζα της μέσης τιμής των τετραγώνων του πλάτους του σήματος. Περιγράφει την ενέργεια του σήματος (Magalhaes, Barros, Loureiro, 2019). Δίνεται από την παρακάτω σχέση:

$$x_{rms} = \sqrt{\frac{1}{n} \cdot \sum_{i=1}^n x_i^2}$$

όπου n ο αριθμός των δειγμάτων που χρησιμοποιήθηκαν.

Τα Zero Crossing Rate και Root Mean Square μπορούν να εξαχθούν και με τις βιβλιοθήκες LibROSA και Iracema. Η βιβλιοθήκη LibROSA είναι ένα πακέτο της Python που χρησιμοποιείται στην επεξεργασία ήχου και μουσικών σημάτων (McFee, Raffel, Liang, Ellis, McVicar, Battenberg, Nieto 2015). Η βιβλιοθήκη Iracema είναι επίσης ένα πακέτο της Python για την ανάλυση μουσικού περιεχομένου που στοχεύει στην εμπειρική έρευνα για τη μουσική εκτέλεση (Magalhaes, Barros, Loureiro, 2019).

2.4.3 Ρυθμικά χαρακτηριστικά

Ρυθμός είναι το μοτίβο των παλμών/νότων ποικίλης δύναμης. Συνήθως περιγράφεται με τους όρους τέμπο, μέτρο ή διατύπωση. Ένα κομμάτι με γρήγορο ρυθμό έχει υψηλό arousal, με συνεχή ρυθμό έχει θετικό valence και με σταθερό ρυθμό έχει αρνητικό valence (Gabrielsson, Lindstrom, 2001). Για την εξαγωγή των ρυθμικών χαρακτηριστικών χρησιμοποιούνται πολλές εργαλειοθήκες όπως είναι το Marsyas (Tzanetakis, Cook, 2002). Το Marsyas είναι ένα δωρεάν λογισμικό το οποίο χρησιμοποιείται για να αναπτύξει γρήγορα και να αξιολογήσει εφαρμογές ήχου στον υπολογιστή. Το Marsyas χρησιμοποιείται για να υπολογίσει το *Beat Histogram* της μουσικής και να εξάγει 6 χαρακτηριστικά από αυτό. Τα βασικότερα από αυτά τα χαρακτηριστικά είναι: Beat Strength, Amplitude, Period της πρώτης και της δεύτερης κορυφής του Beat Histogram, και Ratio της δύναμης των δυο κορυφών σε χτύπους ανά λεπτό (bpm). Το Beat Histogram κατασκευάζεται υπολογίζοντας την αυτοσυσχέτιση του σήματος σε κάθε οκτάβα ζώνης συχνότητας. Οι κυρίαρχες κορυφές της συνάρτησης αντιστοιχούν σε διάφορες περιοδικότητες του σήματος. Η εργαλειοθήκη μουσικής ανάλυσης (MA Toolbox) (Pampalk, 2004) χρησιμοποιείται για να εξάγει το ρυθμικό μοτίβο ή αλλιώς *Rhythm Pattern* που περιέχει πληροφορία για το πόσο δυνατοί και γρήγοροι είναι οι χτύποι που παίζονται μέσα στις αντίστοιχες ζώνες συχνότητας (Pampalk, Rauber, Merkl, 2002). Αυτή η εργαλειοθήκη χρησιμοποιεί τον μετασχηματισμό STFT (Short-Time Fourier Transform) για να αποκτηθεί η διαμόρφωση πλάτους του SONE για κάθε τμήμα 6 δευτερολέπτων του μουσικού κομματιού. Το Rhythm Pattern ενσωματώνεται σε ένα ιστογράμμο ρυθμού (*Rhythm histogram*) των 60-bin αθροίζοντας τους συντελεστές διαμόρφωσης πλάτους κατά μήκος των ζωνών. Η μέση τιμή του ιστογράμματος ρυθμού είναι και ο μέσος όρος του τέμπο. Αυτά τα χαρακτηριστικά εξάγονται από

τον Rhythm Pattern (RP) Extractor (Lidy, Rauber, 2005). Παράλληλα, με την χρήση της εργαλειοθήκης ανάκτησης μουσικής πληροφορίας (MIR toolbox) (Lartillot, Toivianinen, 2007), όπως προτείνεται σε έρευνες, εξάγονται πέντε ρυθμικά χαρακτηριστικά: *Rhythm Strength*, *Rhythm Regularity*, *Rhythm Clarity*, *Average Onset Frequency*, και *Average Tempo* (Lu, Liu, Zhang, 2006). Το Rhythm Strength υπολογίζεται από τον μέσο όρο έναρξης της δύναμης της καμπύλης εντοπισμού έναρξης, με βάση τον αλγόριθμο του Klapuri (Klapuri, 1999). Ο όρος έναρξη αναφέρεται στον χρόνο έναρξης του τραγουδιού. Τα Rhythm Regularity και Rhythm Clarity υπολογίζονται κάνοντας αυτοσυσχέτιση στην καμπύλη εντοπισμού έναρξης. Το Average Onset Frequency υπολογίζεται ως ο αριθμός των ενάρξεων ανά δευτερόλεπτο, ενώ το Average Tempo υπολογίζεται εντοπίζοντας την περιοδικότητα από την καμπύλη εντοπισμού έναρξης.

2.4.4 Αρμονικά χαρακτηριστικά

Τα αρμονικά χαρακτηριστικά είναι χαρακτηριστικά που υπολογίζονται από την ημιτονοειδή αρμονική μοντελοποίηση του σήματος (Peeters, 2004). Για αυτά, χρησιμοποιείται η εργαλειοθήκη μουσικής ανάκτησης πληροφορίας ή αλλιώς Music Information Retrieval (MIR toolbox) (Lartillot, Toivianinen, 2007) και εξάγονται δυο χαρακτηριστικά τόνου, *Salient Pitch*, *Chromagram Centroid* και τρία χαρακτηριστικά τονικότητας, *Key Clarity*, *Musical Mode*, *Harmonic Change*. Η MIRtoolbox εκτιμά το τονικό ύψος (Pitch), ή την αντιλαμβανόμενη θεμελιώδη συχνότητα (Perceived Fundamental Frequency) του κάθε βραχυπρόθεσμου καρέ (50 ms, 1/2 επικάλυψη) βασισμένη στον αλγόριθμο multi-pitch detection των Tolonen και Karjalainen (Tolonen, Karjalainen, 2000). Ο αλγόριθμος αποσυνθέτει μια ακουστική κυματομορφή σε δύο ζώνες συχνότητας (πάνω και κάτω του 1kHz), υπολογίζει τη συνάρτηση αυτοσυσχέτισης σε κάθε υποζώνη και τελικά παράγει εκτιμήσεις τονικού ύψους επιλέγοντας τις κορυφές από τις δύο λειτουργίες αυτοσυσχέτισης. Έτσι το εκτιμώμενο τονικό ύψος αντιστοιχεί στην υψηλότερη κορυφή και ονομάζεται εξέχον τονικό ύψος (Salient Pitch). Επιπλέον, υπολογίζει το Wrapped Chromagram, ή το Pitch Class Profile, για κάθε καρέ (100 ms και 1/8 επικάλυψης) και χρησιμοποιεί το κέντρο βάρους του χρωμογράμματος (Chromagram Centroid) ως άλλη μια εκτίμηση της θεμελιώδους συχνότητας. Το Wrapped Chromagram σχεδιάζει το φάσμα συχνότητας σε 12 bins εκπροσωπώντας το καθένα 12 διαφορετικά ημιτόνια (χρώματα) της μουσικής οκτάβας. Κάθε bin αντιστοιχεί σε μια από τις 12 τάξεις ημιτόνου σε μια κλίμακα δυτικού δωδεκάτονου ίσης ιδιοσυγκρασίας. Επόμενο βήμα, είναι η σύγκριση ενός χρωμογράμματος σε 24 μικρότερα και μέγιστα προφίλ κλειδιών (Gomez, 2006) για να εκτελεστεί ο εντοπισμός κλειδιού και να εκτιμηθεί η δύναμη του καρέ σε συνδυασμό με κάθε κλειδί. Η δύναμη που σχετίζεται με το καλύτερο κλειδί, που είναι αυτό με την υψηλότερη δύναμη, επιστρέφεται ως Key Clarity. Η διαφορά ανάμεσα στο μέγιστο και στο μικρότερο κλειδί είναι η δύναμη του κλειδιού που επιστρέφεται ως μια εκτίμηση της μουσικής λειτουργίας (Musical Mode), η οποία περιγράφει μια καθορισμένη διάταξη των διατονικών τόνων μιας οκτάβας (Oliveira & Cardoso, 2008). Η μουσική λειτουργία σχετίζεται με την αντίληψη της μουσικής (Gabrielsson, Lindstrom, 2001· Oliveira, Cardoso, 2009). Ωστόσο, δεν είναι σίγουρο το αν οι τιμές της μουσικής λειτουργίας συσχετίζονται με το valence. Η εργαλειοθήκη MIR χρησιμοποιεί έναν αλγόριθμο (Harte, Sandler, Gasser, 2006) για να υπολογίσει ένα διάνυσμα χαρακτηριστικού έξι διαστάσεων που ονομάζεται Tonal Centroid από το χρωμόγραμμα και εντοπίζει αρμονικές αλλαγές (Harmonic Changes) σε μουσικό ήχο. Η υψηλή αρμονική αλλαγή δείχνει μεγάλη διαφορά στο αρμονικό περιεχόμενο ανάμεσα σε συνεχόμενα καρέ. Τα βραχυπρόθεσμα χαρακτηριστικά συγκεντρώνονται λαμβάνοντας την μέση και τυπική απόκλιση (Meng, Ahrendt, Larsen, Hansen, 2007). Η εργαλειοθήκη Marsyas υπολογίζει το *Pitch Histogram* και εξάγει τα εξής χαρακτηριστικά: Tonic, Main Pitch Class, Octave Range του Dominant Pitch, Main Tonal Interval Relation, και Overall Pitch Strength (Tzanetakis, Cook, 2002). Τέλος, το PsySound συγκρίνει 16 χαρακτηριστικά τόνου, όπως Mean, Standard Deviation, Skewness, Kurtosis του τόνου και Pitch

Strength Time Series που εκτιμούνται από το *SWIPE* (*Sawtooth Waveform Inspired Pitch Estimator*) και το *SWIPE'*. Το τελευταίο περιορίζει την ανάλυση των πρώτων και κύριων αρμονικών χαρακτηριστικών (Cabrera, 1999· Camacho, 2007).

2.4.5 Φασματικά χαρακτηριστικά

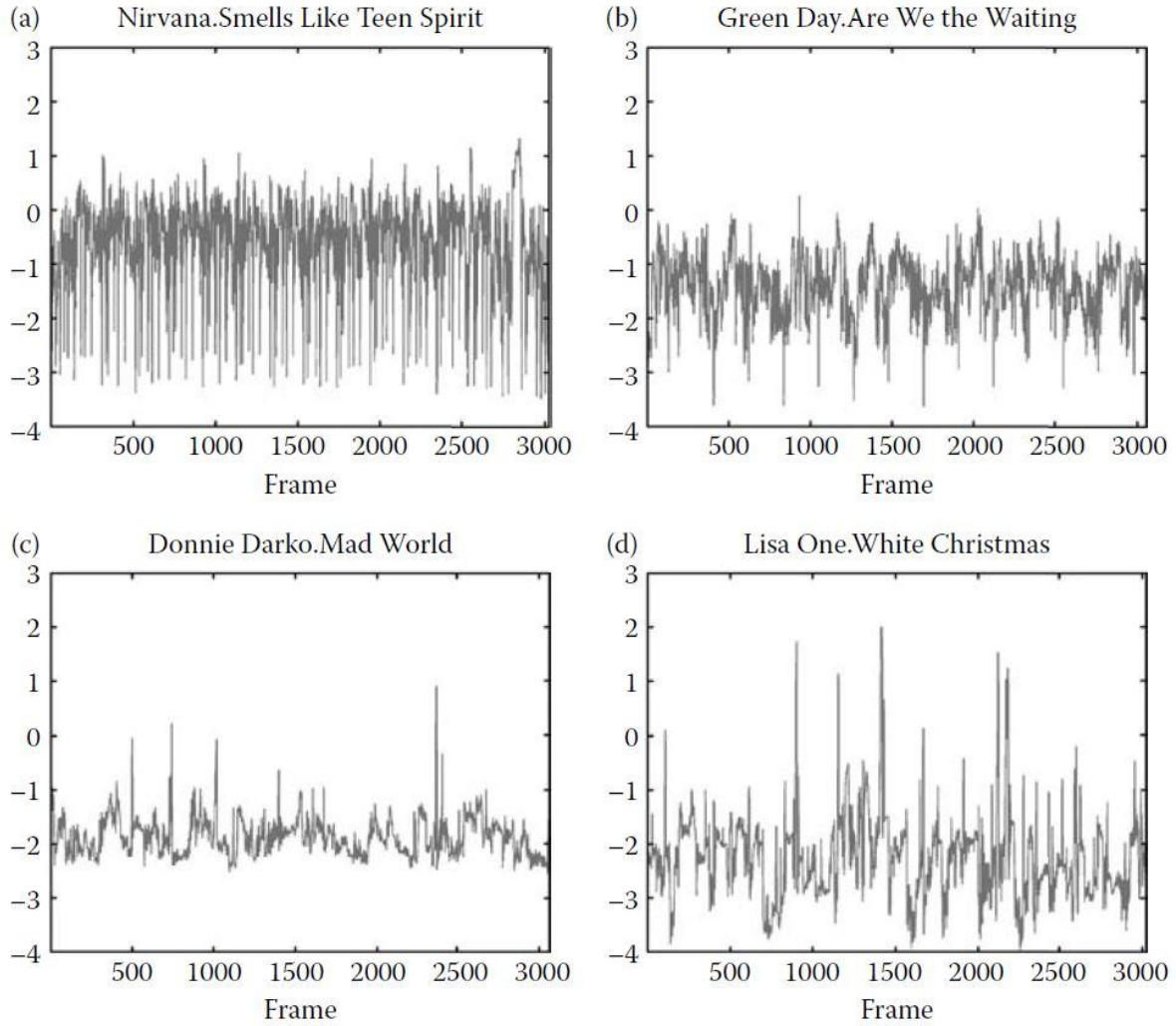
Τα φασματικά χαρακτηριστικά υπολογίζονται από το μετασχηματισμό STFT του ηχητικού σήματος (Peeters, 2004). Σε αυτή την κατηγορία περιλαμβάνονται τα *Spectral Centroid*, *Spectral Rolloff*, *Spectral Flux*, *Spectral Flatness Measures (SFM)*, *Spectral Crest Factors (SCF)*, *Spectral Entropy*, *Spectral Spread*, *Spectral Bandwidth*, *Spectral contrast*, *Chroma STFT* και *Mel Frequency Cepstral Coefficient (MFCC)*.

Spectral centroid

Το *Spectral Centroid* είναι το κέντρο βάρους του φασματικού πλάτους του μετασχηματισμού STFT. Το κέντρο βάρους είναι η μέτρηση της φασματικής μορφής και υπολογίζεται από την παρακάτω σχέση:

$$\text{spectral centroid} = \frac{\sum_{n=1}^N n A_t^n}{\sum_{n=1}^N A_t^n},$$

όπου A_t^n είναι πλάτος του φάσματος των t καρέ και n η συχνότητα των bin, και N ο συνολικός αριθμός των bins (Yang, Chen, 2011·). Η παρακάτω εικόνα απεικονίζει ένα παράδειγμα του Spectral Centroid για τέσσερα μουσικά κομμάτια.



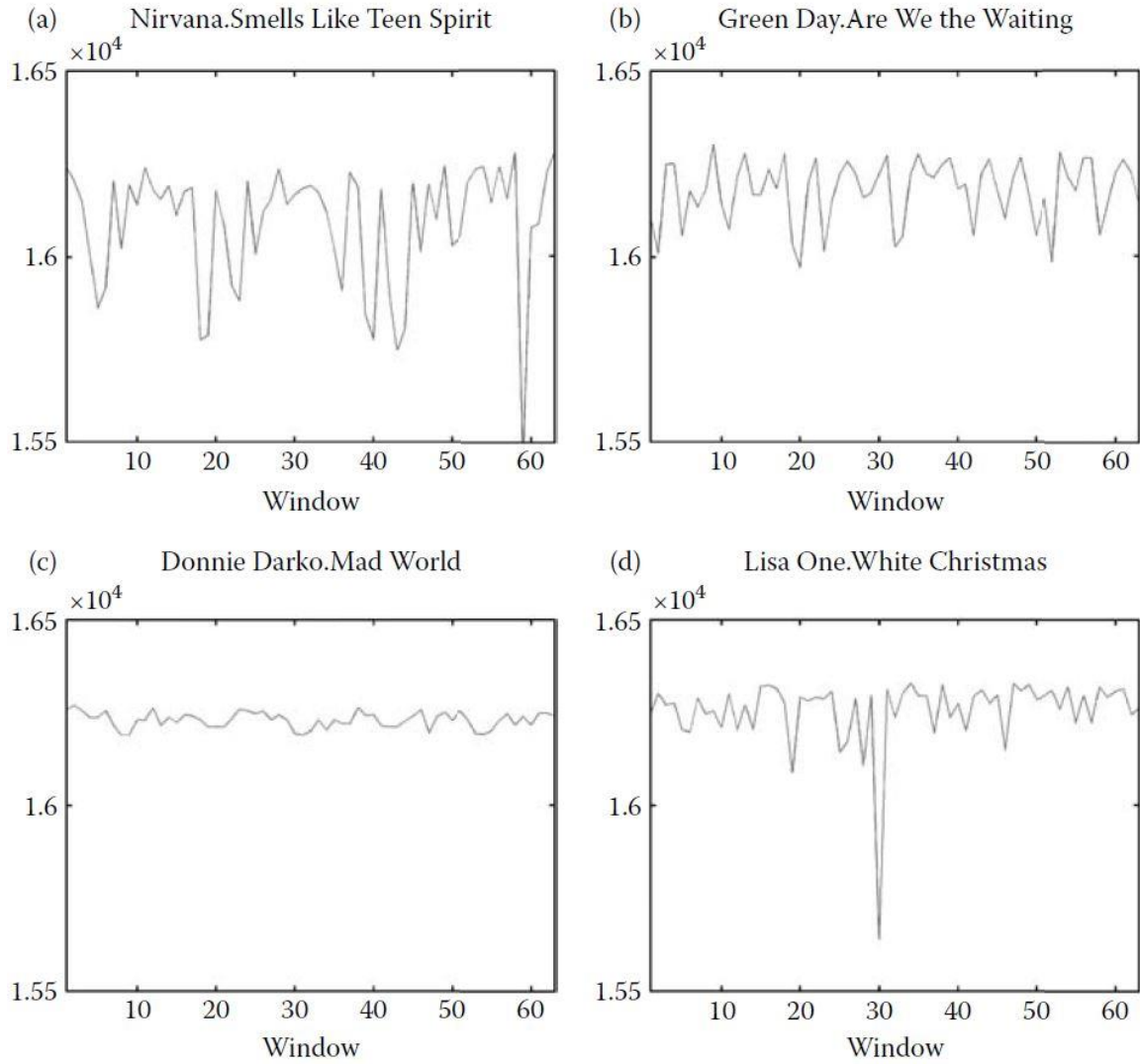
Εικόνα 2.12: Παράδειγμα απεικόνισης του Spectral Centroid σε τέσσερα μουσικά κομμάτια (Benetos, 2007· Yang, Chen, 2011)

Spectral rolloff

Το *spectral rolloff* είναι η ποσότητα της υψηλής συχνότητας στο σήμα και είναι και αυτό μέτρηση της φασματικής μορφής. Το Spectral Rolloff ορίζεται ως η συχνότητα k_t στην παρακάτω σχέση η οποία περιέχει ένα ορισμένο κλάσμα της συνολικής ενέργειας. Η αναλογία είναι 0,85.

$$\sum_{n=1}^{k_t} A_t^n = 0.85 * \sum_{n=1}^N A_t^n.$$

όπου A_t^n είναι πλάτος του φάσματος των t καρέ και n η συχνότητα των bin, και N ο συνολικός αριθμός των bins (Yang, Chen, 2011· Magalhaes, Barros, Loureiro, 2019· Βαρδάνης, 2019). Η παρακάτω εικόνα απεικονίζει ένα παράδειγμα του Spectral Rolloff για τέσσερα μουσικά κομμάτια.



Εικόνα 2.13: Παράδειγμα απεικόνισης του Spectral Rolloff σε τέσσερα μουσικά κομμάτια (Benetos, 2007· Yang, Chen, 2011)

Spectral flux

Το *Spectral Flux* εκτιμάει την ποσότητα της γενικής φασματικής αλλαγής. Το Spectral Flux ορίζεται ως ο κύκλος της διαφοράς ανάμεσα στα ισορροπημένα πλάτη των επιτυχημένων καρέ και δίνεται από την παρακάτω σχέση:

$$\text{spectral flux} = \sum_{n=1}^N (a_t^n - a_{t-1}^n)^2,$$

Όπου το a_t^n δηλώνει το ισορροπημένο πλάτος του φάσματος (ισορροπημένο για κάθε καρέ) των t καρέ και n η συχνότητα των bin, και N ο συνολικός αριθμός των bins.

Spectral Flatness Measures και Spectral Crest Factors

Τα *spectral flatness measures (SFM)* και οι *spectral crest factors (SCF)* σχετίζονται με την τονικότητα του ηχητικού σήματος (Allamanche, Herre, Helmuth, Froba, Kasten, Cremer, 2001). Η

τονικότητα έχει να κάνει με το valence. Για παράδειγμα χαρούμενες και γαλήνιες μελωδίες είναι τονικές και οι μελωδίες που δείχνουν θυμό είναι άτονες (Thompson, Robitaille, 1992). Οι Spectral Flatness Measures (SFM) είναι η αναλογία ανάμεσα στο γεωμετρικό μέσο της φασματικής δύναμης και στο αριθμητικό μέσο. Τα Spectral Flatness Measures υπολογίζονται από την σχέση:

$$\text{spectral flatness measure} = \frac{\left(\prod_{n \in B^k} A_t^n\right)^{1/N_k}}{\frac{1}{N_k} \sum_{n \in B^k} A_t^n},$$

όπου A_t^n είναι πλάτος του φάσματος των t καρέ, το B^k δηλώνει τις k υποζώνες συχνότητας, και N_k είναι ο αριθμός των bins στο B^k . Αντίθετα, οι Spectral Crest Factors (SCF) είναι η αναλογία ανάμεσα στην κορυφή πλάτους και την ρίζα του μέσου τετραγωνικού πλάτους.

$$\text{spectral crest factor} = \frac{\max_{n \in B^k} A_t^n}{\frac{1}{N_k} \sum_{n=1}^N A_t^n},$$

όπου το B^k δηλώνει τις k υποζώνες συχνότητας, και N_k είναι ο αριθμός των bins στο B^k .

Spectral entropy

Το *Spectral Entropy* συνιστά, ουσιαστικά, έναν δείκτη που μετράει τις αλλαγές στην κατανομή της φασματικής ενέργειας του σήματος. Σε περίπτωση που το φάσμα εμπεριέχει μία μόνο κυρίαρχη κορυφή, η εντροπία, είναι ελάχιστη (Magalhaes, Barros, Loureiro, 2019· Βαρδάνης, 2019). Δίνεται από την παρακάτω σχέση.

$$H(X) = - \sum_{i=1}^{N/2} p_i \cdot \log_2(p_i)$$

όπου p_i η συνάρτηση πυκνότητας πιθανότητας του φάσματος και N ο αριθμός των σημείων του FFT.

Spectral spread

Το *Spectral Spread* δίνει μια εκτίμηση της εξάπλωσης του φασματικής ενέργειας γύρω από το Spectral Centroid. Εκφράζει την τυπική απόκλιση γύρω από τη μέση τιμή της φασματικής κατανομής, και υποδηλώνει πόσο θορυβώδης είναι ένας ήχος. Όσο μεγαλύτερη είναι η φασματική διασπορά τόσο πιο θορυβώδης είναι ο ήχος (Magalhaes, Barros, Loureiro, 2019· Βαρδάνης, 2019). Δίνεται από την παρακάτω σχέση.

$$V_{SS} = \sqrt{\frac{\sum_{k=1}^{N/2} (f(k) - V_{SC})^2 \cdot |X(k)|}{\sum_{k=1}^{N/2} f|X(k)|}}$$

όπου N ο αριθμός των σημείων του FFT και $f(k)$ και $X(k)$ η συχνότητα και το πλάτος στο k-οστό σημείο του FFT (Magalhaes, Barros, Loureiro, 2019).

Spectral bandwidth

Το *Spectral Bandwidth* υπολογίζει το φασματικό εύρος ζώνης της τάξης-p και δίνεται από την παρακάτω σχέση.

$$\left(\sum_k S(k)(f(k) - f_c)^p \right)^{\frac{1}{p}}$$

Όπου $S(k)$ είναι το φασματικό μέγεθος στη συχνότητα bin k, $f(k)$ είναι η συχνότητα bin k και f_c είναι το Spectral Centroid. Όταν $p=2$, πρόκειται για μια σταθμισμένη τυπική απόκλιση. (https://colab.research.google.com/github/iranroman/musicinformationretrieval.com/blob/gh-pages/spectral_features.ipynb#scrollTo=hYpfj_Meenm).

Spectral contrast

Το *Spectral Contrast* εξετάζει τη φασματική κορυφή, τη φασματική κοιλάδα και τη διαφορά τους σε κάθε υποζώνη συχνότητας. Συγκεκριμένα, υπολογίζει τη φασματική αντίθεση για έξι υποζώνες για κάθε καρέ.

Τα Spectral Centroid, Spectral Rolloff, Spectral Flatness Measures (SFM), μπορούν να εξαχθούν τόσο από τα Music Analysis Toolbox (MA), Marsyas, Sound Description Toolbox (SDT), όσο και από τις βιβλιοθήκες της Python, LibROSA και Ircacema. Το Spectral Flux μπορεί να εξαχθεί τόσο από τα Marsyas, Sound Description Toolbox (SDT) όσο και από την βιβλιοθήκη Ircacema της Python. Το Spectral Entropy μπορεί να εξαχθεί από το MATLAB και την βιβλιοθήκη Ircacema της Python. Το Spectral Contrast εξάγεται από την βιβλιοθήκη LibROSA της Python αλλά και από το MATLAB. Το Spectral Spread εξάγεται από τη βιβλιοθήκη Ircacema της Python ενώ το Spectral Bandwidth από την βιβλιοθήκη LibROSA της Python (Tzanetakis, Cook, 2002· Magalhaes, Barros, Loureiro, 2019· McFee, Raffel, Liang, Ellis, McVicar, Battenberg, Nieto, 2015 · Yang, Chen, 2011).

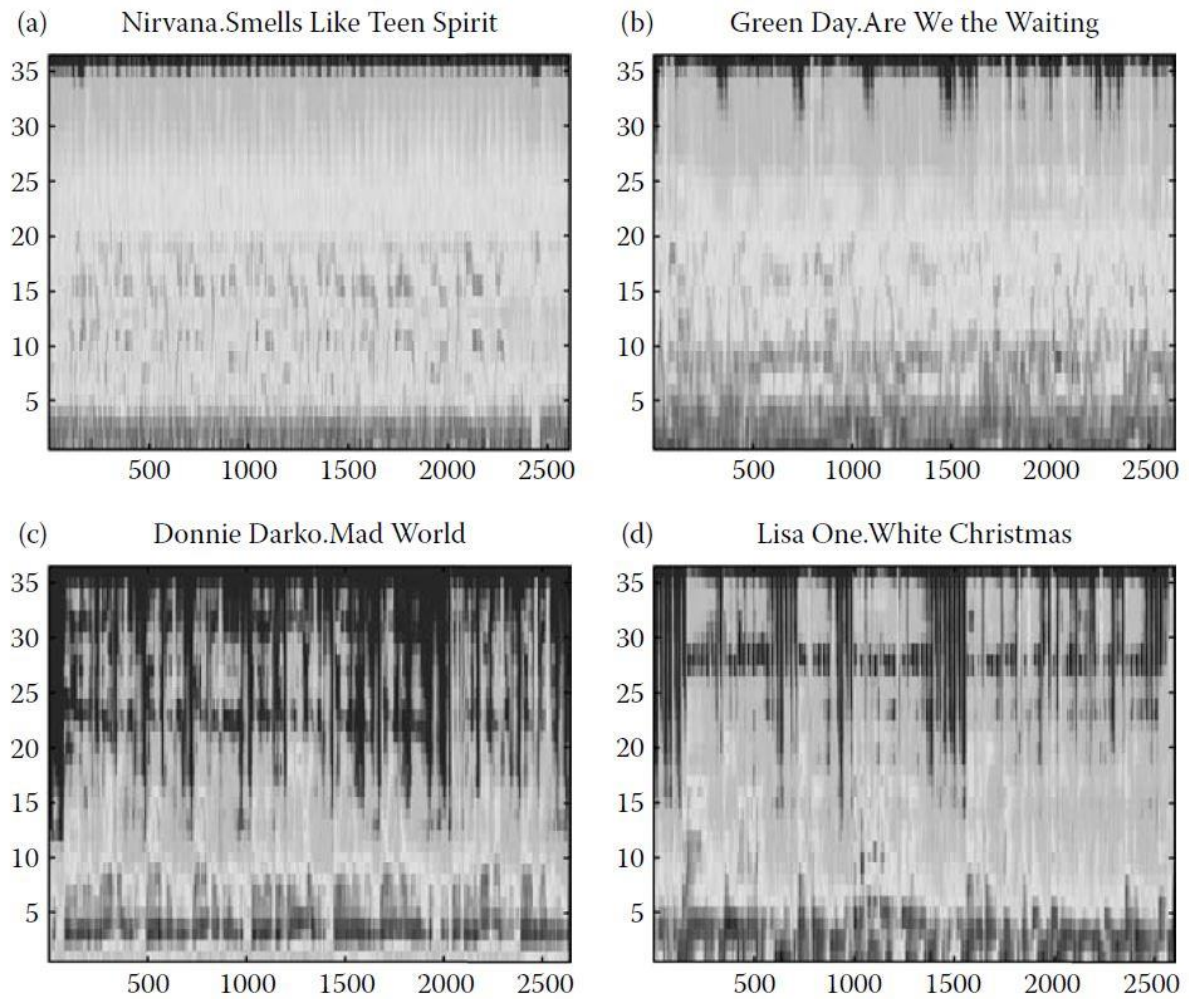
MFCC

Στο λογισμικό Marsyas χρησιμοποιούνται 24 υποζώνες. Μια από αυτές χρησιμοποιεί την εργαλειοθήκη μουσικής ανάλυσης (MA toolbox) (Pampalk, 2002) για να εξάγει τους *Mel-frequency cepstral coefficients* (MFCCs). Τους MFCC μπορεί να εξάγει και η βιβλιοθήκη LibROSA της Python. Οι MFCC είναι οι συντελεστές του μετασχηματισμού διακριτού συνημίτονου ή αλλιώς Discrete Cosine Transform (DCT) κάθε βραχυπρόθεσμου λογάριθμου του φάσματος ισχύος που εκφράζεται σε μια κλίμακα Mel-frequency μη γραμμικής αντίληψης για να αναπαραστήσει μορφώττες κορυφές

του φάσματος (Davis, Mermelstein, 1980). Η κανονική συχνότητα f hertz μπορεί να μετατραπεί στο εύρος mel και δίνεται από τον παρακάτω τύπο:

$$m = 1127.01048 \log (1 + f / 700)$$

Λαμβάνεται έτσι, η μέση και τυπική απόκλιση των πρώτων 13 ή 20 MFCC του κάθε καρέ. Ο MFCC είναι ευρέως διαδεδομένος στην επεξεργασία σήματος ομιλίας και στην έρευνα Music Information Retrieval (Casey, Veltkamp, Goto, Leman, Rhodes, Slaney, 2008). Παρόλα αυτά ο συντελεστής MFCC χάνει την σχετική φασματική πληροφορία (Lu, Liu, Zhang, 2006). Για αυτό το λόγο προτείνεται να βρεθεί η σχετική διανομή ενέργειας στα αρμονικά συστατικά του φάσματος. Έτσι η φασματική κορυφή (Spectral Peak), η φασματική κοιλάδα (Spectral Valley) και οι δυναμικές τους σε κάθε υποζώνη αντανakλούν την σχετική κατανομή των αρμονικών και μη αρμονικών συστατικών στο φάσμα. Αυτό δείχνει την υπεροχή των φασματικών χαρακτηριστικών αντίθεσης (Spectral Contrast Features) (Jiang, Lu, Zhang, Tao, Cai, 2002) έναντι του MFCC για την μουσική ταξινόμηση. Άλλο ένα φασματικό χαρακτηριστικό που μπορεί να χρησιμοποιηθεί στο παρόν σύστημα είναι το *Daubechies Wavelets Coefficient Histogram* (DWCH) (Li, Ogihara, 2003· Li, Ogihara, 2004· Li, Ogihara, 2006). Αυτό, υπολογίζεται σε διαφορετικές υποζώνες συχνότητας με διαφορετικές αναλύσεις. Τόσο τα χαρακτηριστικά αντίθεσης όσο και το DWCH εφαρμόζονται στο MATLAB (<http://www.mathworks.com/products/wavelet/>). Μπορεί επιπλέον να χρησιμοποιηθεί η εργαλειοθήκη μουσικής ανάκτησης πληροφορίας (MIRtoolbox) (Lartillot, Toivainen, 2007) για να παραχθούν τρία αισθητήρια χαρακτηριστικά, *Roughness*, *Irregularity*, *Inharmonicity*. Το πρώτο μετράει τον θόρυβο του φάσματος (Sethares, 1998). Το δεύτερο μετράει το βαθμό παραλλαγής των διαδοχικών κορυφών [Fujinaga, McMillan, 2000· Jensen, 1999]. Το τρίτο αναπαριστά την απόκλιση των φασματικών συστατικών του σήματος από ένα καθαρά αρμονικό σήμα (Osgood, Suci, Tannenbaum, 1957). Οι ψυχολογικές μελέτες δείχνουν ότι το valence σχετίζεται με το αρμονικό περιεχόμενο των μουσικών σημάτων. Τέλος, άλλα τρία χαρακτηριστικά που χρησιμοποιούνται είναι τα *Tristimulus*, *Even-Harm*, και *Odd-Harm* (Zhu, Shi, Kim, Eom, 2006· Wieczorkowska, Synak, Lewis, Ras, 2005· Wieczorkowska, Synak, Ras, 2006). Το πρώτο είναι μια παράμετρος ενώ τα άλλα δυο αναπαριστούν τις άρτιες και περιττές αρμονικές του φάσματος. Η παρακάτω εικόνα απεικονίζει ένα παράδειγμα των MFCC για τέσσερα μουσικά κομμάτια.



Εικόνα 2.14: Παράδειγμα απεικόνισης των MFCC σε τέσσερα μουσικά κομμάτια (Pampalk, 2004· Yang, Chen, 2011)

3 Μεθοδολογία

Κατά καιρούς, στις μελέτες που έχουν πραγματοποιηθεί, χρησιμοποιήθηκαν αλγόριθμοι ταξινόμησης. Οι αλγόριθμοι ταξινόμησης χρησιμοποιούνται για να βρουν τη σχέση ανάμεσα σε μουσικά χαρακτηριστικά και ετικέτες συναισθήματος όπως είναι οι support vector machines (Hu, Downie, Laurier, Bay, Ehmann, 2008· Li, Ogihara, 2003· Wang, Zhang, Zhu, 2004· Bischoff, Firan, Paiu, Nejd, Laurier, Sordo, 2009), τα Gaussian Mixture Models (GMM) (Fernandes, Paiva, 2010· Liu, Zhang, Lu, 2003), τα neural networks (Feng, Zhuang, Pan, 2003) και τα k-nearest neighbor (Yang, Liu, Chen, 2006· Wiczorkowska, 2004).

3.1 Διαδικασία μελέτης

Στην παρούσα μελέτη δε χρησιμοποιήσαμε κάποιον αλγόριθμό ταξινόμησης αλλά την γλώσσα προγραμματισμού Python. Η Python είναι διερμηνευόμενη, γενικού σκοπού και υψηλού επιπέδου, γλώσσα προγραμματισμού. Ανήκει στις γλώσσες προστακτικού προγραμματισμού. Επίσης, υποστηρίζει τόσο το διαδικαστικό όσο και το αντικειμενοστρεφές προγραμματιστικό υπόδειγμα.

Το λογισμικό που εγκαταστήσαμε για την χρήση αυτής της γλώσσας προγραμματισμού είναι το PyCharm. Το PyCharm είναι ένα ολοκληρωμένο περιβάλλον ανάπτυξης (IDE) που χρησιμοποιείται για τον προγραμματισμό σε Python. Παρέχει ανάλυση κώδικα, ένα γραφικό πρόγραμμα εντοπισμού σφαλμάτων, έναν ολοκληρωμένο ελεγκτή μονάδων, ενσωμάτωση με συστήματα ελέγχου εκδόσεων και υποστηρίζει την ανάπτυξη ιστοσελίδων με το Django.

Στο λογισμικό PyCharm, χρειάστηκε να εισάγουμε τη βιβλιοθήκη LibROSA. Η βιβλιοθήκη LibROSA όπως αναφέραμε και στην υποενότητα 2.4.2 είναι ένα πακέτο της Python που χρησιμοποιείται στην επεξεργασία ήχου και μουσικών σημάτων.

Απαραίτητη ήταν επίσης η εγκατάσταση της εφαρμογής Spotify. Το Spotify είναι σουηδική υπηρεσία αναπαραγωγής μουσικής. Η υπηρεσία είναι διαθέσιμη σε χώρες της Αμερικής, της Ευρώπης και της Ωκεανίας. Επιτρέπει την πλοήγηση και την αναζήτηση μουσικής ανά καλλιτέχνη, άλμπουμ, είδος, λίστα αναπαραγωγής ή δισκογραφικής εταιρείας.

Τέλος, εγκαταστήσαμε το λογισμικό Audacity. Το Audacity είναι πρόγραμμα ψηφιακής επεξεργασίας ήχου και ηχογράφησης, που κυκλοφορεί σαν ελεύθερο λογισμικό και είναι ανεξάρτητο πλατφόρμας.

3.2 Συλλογή και επεξεργασία δεδομένων

Αρχικά, συλλέξαμε τα κομμάτια που θα χρησιμοποιηθούν στη μελέτη, από την εφαρμογή spotify. Η εφαρμογή διαθέτει λίστες αναπαραγωγής με ετικέτα συναισθήματος. Πραγματοποιούμε μία επιλογή στις ετικέτες συναισθήματος και στη συνέχεια επιλέγουμε από κάθε λίστα μουσικά κομμάτια διαφορετικού είδους. Τα μουσικά κομμάτια που επιλέχθηκαν είναι 28, στο σύνολο. Τα συναισθήματα που έχουν επιλεγεί είναι *angry*, *chill*, *dance*, *focus*, *happy*, *romance*, και *sad*. Στον παρακάτω πίνακα παρουσιάζονται οι κατηγορίες συναισθημάτων με τα μουσικά κομμάτια.

ANGRY
Metal - Arctic Monkeys-R U Mine
Rnb - Justin Timberlake-Cry Me A River
Rock - Imagine Dragons-Believer
CHILL
Hip hop - The Weeknd-The Hills

Pop - Niall Horan-This Town
Rock - Radiohead-No Surprises
DANCE
Disco - Boney M-Rasputin
House - David Guetta Anne-Marie Coi Leray-Baby Dont Hurt Me
Pop - Madonna-Like A Prayer
FOCUS
Jazz - Figgy mold-All My Tomorrows
Rnb - Drake-Not You Too (Audio) ft. Chris Brown
Κλασσική - Beethoven-Tempo di Menuetto
HAPPY
Disco - Elton John Britney Spears-Hold Me Closer
House - David Guetta Bebe Rexha-Im Good (Blue)
Pop - Harry Styles-As It Was
Pop ελληνικό - Μέλισσες-30-40(Oh Yes i do)
ROMANCE
Jazz - Frank Sinatra-Fly Me To The Moon
Metal - Aerosmith-Dream On
Rock - Bon Jovi-Always
Rock ελληνικό - Παύλος Σιδηρόπουλος-Να Μ Αγαπάς
Έντεχνο - Ελεωνόρα Ζουγανέλη-Η Επιμονή σου
Κλασσική - W.A. Mozart-Serenade
Λαϊκό - Δημήτρης Μητροπάνος-Όταν έχω εσένα
SAD
Pop - Adele-Easy On Me
Pop ελληνικό - Στέλιος Ρόκκος-Είμαι δικός σου
Rock ελληνικό - Φίλιππος Πλιάτσικας-Ποιος έχει λόγο στην αγάπη
Έντεχνο - Alex Sid - Μάτια Κλειστά
Λαϊκό - Κωνσταντίνος Αργυρός-Δεν Έχω Πολλά

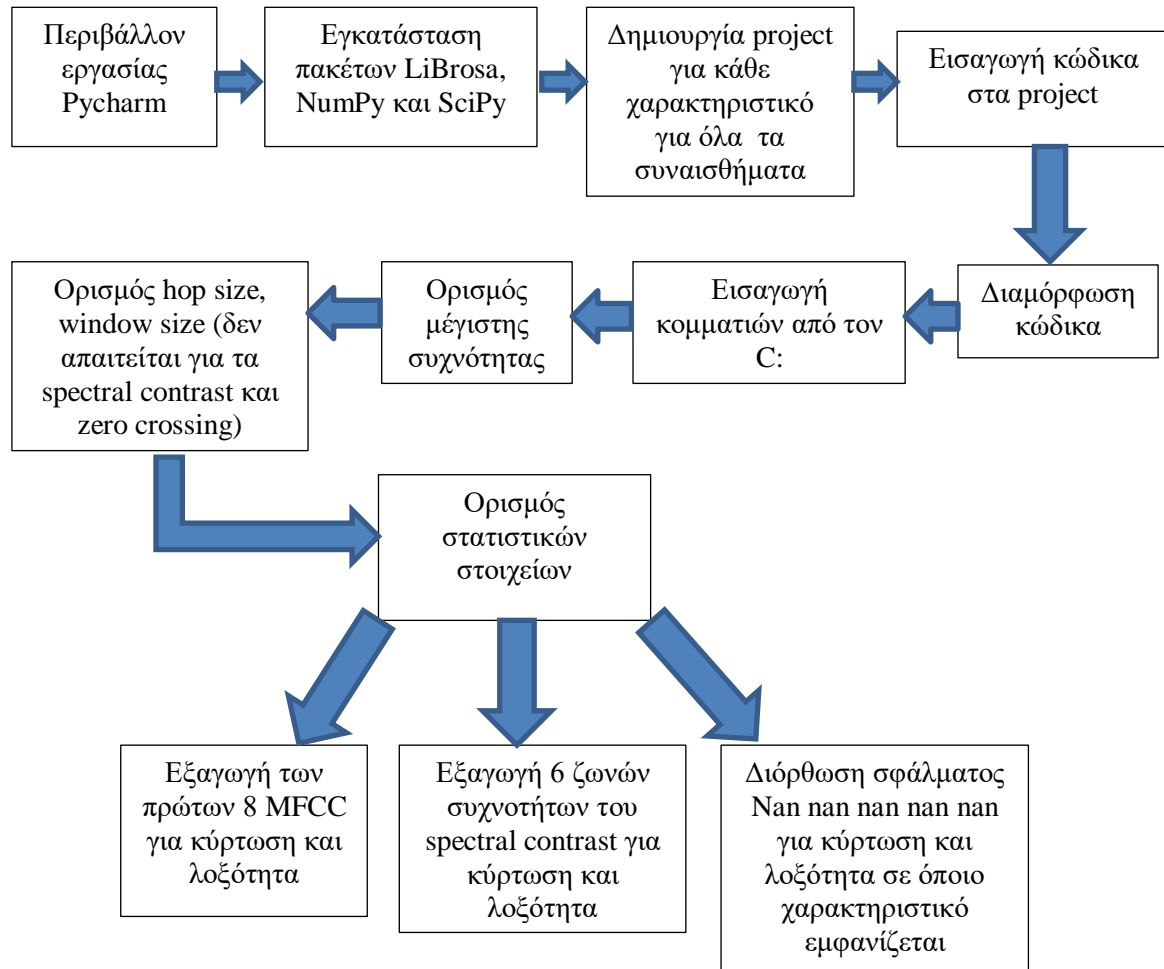
Εικόνα 3.1: Κατηγορίες συναισθήματος και μουσικά κομμάτια

Τα μουσικά κομμάτια έχουν μορφή .wav. Μετά τη συλλογή των κομματιών, είναι απαραίτητο να ελεγχθούν για ψαλιδισμό. Χρησιμοποιήσαμε το λογισμικό Audacity. Τα κομμάτια που παρουσίαζαν ψαλιδισμό κανονικοποιήθηκαν στο -1 και αποθηκεύτηκαν εκ νέου. Για να εισαχθούν τα μουσικά κομμάτια στο λογισμικό PyCharm, ήταν απαραίτητο να αποθηκευτούν σε φάκελο στον Τοπικό Δίσκο (C:).

3.3 Εξαγωγή μουσικών παραμέτρων

Τα μουσικά χαρακτηριστικά που εξάγονται είναι τα φασματικά και τα χρονικά. Τα φασματικά χαρακτηριστικά που εξάγονται είναι τα spectral contrast, spectral centroid, spectral bandwidth, spectral flatness, spectral rolloff, MFCC και chroma stft. Τα χρονικά χαρακτηριστικά που εξάγονται είναι τα zero crossing rate και rms. Δίνεται έμφαση στα φασματικά χαρακτηριστικά γιατί

είναι ιδιαίτερα σημαντικά στην ταξινόμηση ηχητικών σημάτων και δίνουν πληροφορίες για τη συχνότητα, την ισχύ και άλλα χαρακτηριστικά του σήματος (Banitalebi-Dehkordi, Banitalebi-Dehkordi, 2018). Τα παραπάνω χαρακτηριστικά έχουν αναλυθεί στο κεφάλαιο 2, υποενότητες 2.4.2, 2.4.4, 2.4.5. Εφόσον έχουμε ολοκληρώσει τα βήματα της προηγούμενης ενότητας, χρειάζεται να γίνουν κάποιες απαραίτητες ενέργειες για να μας δώσει ο κώδικας τα επιθυμητά αποτελέσματα. Οι ενέργειες αυτές παρουσιάζονται στο παρακάτω μπλοκ διάγραμμα.



Εικόνα 3.2: Μπλοκ διάγραμμα επεξεργασίας μουσικών χαρακτηριστικών

Τα πακέτα που είναι σημαντικό να εγκατασταθούν είναι τα NumPy και SciPy. Το NumPy είναι μια βιβλιοθήκη ανοιχτού κώδικα για εργασίες προγραμματισμού Python που πραγματοποιεί μαθηματικούς και επιστημονικούς υπολογισμούς. Το SciPy είναι μία βιβλιοθήκη ανοιχτού κώδικα για εργασίες προγραμματισμού Python που χρησιμοποιείται για την επίλυση επιστημονικών και μαθηματικών προβλημάτων.

Η μέγιστη συχνότητα των κομματιών είναι 48000Hz.

Το Window Size είναι το μέγεθος σε δείγματα της ανάλυσης παραθύρου. Επιλέγεται μέγεθος παραθύρου 1024.

Το Hop size είναι το μέγεθος σε δείγματα του μεγέθους ανάλυσης μετάβασης παραθύρου. Επιλέγεται μέγεθος 768.

Τα αποτελέσματα των παραπάνω χαρακτηριστικών εξάγονται σε μορφή πίνακα. Έτσι για να είναι τα αποτελέσματά μας πιο σαφή εξάγουμε πέντε στατιστικά στοιχεία: μέση τιμή, τυπική απόκλιση, διακύμανση, κύρτωση και λοξότητα.

Η μέση τιμή είναι η μαθηματική πράξη ανεύρεσης της «μέσης απόστασης» ανάμεσα σε δύο ή περισσότερους αριθμούς.

Η τυπική απόκλιση είναι ένα μέτρο που χρησιμοποιείται για να υπολογιστεί το ποσό της μεταβολής ή της διασποράς ενός συνόλου τιμών δεδομένων.

Η διακύμανση είναι η αναμενόμενη τιμή της τετραγωνικής απόκλισης της τυχαίας μεταβλητής από τη μέση τιμή, και άτυπα μετρά πόσο μακριά ένα σύνολο (τυχαίων) αριθμών απλώνεται από τη μέση τιμή του.

Η κύρτωση εκφράζει το βαθμό συγκέντρωσης των τιμών μιας κατανομής γύρω από το μέσον της. Ο προσδιορισμός της κύρτωσης γίνεται όπως και στην ασυμμετρία σε σχέση με την κανονική κατανομή.

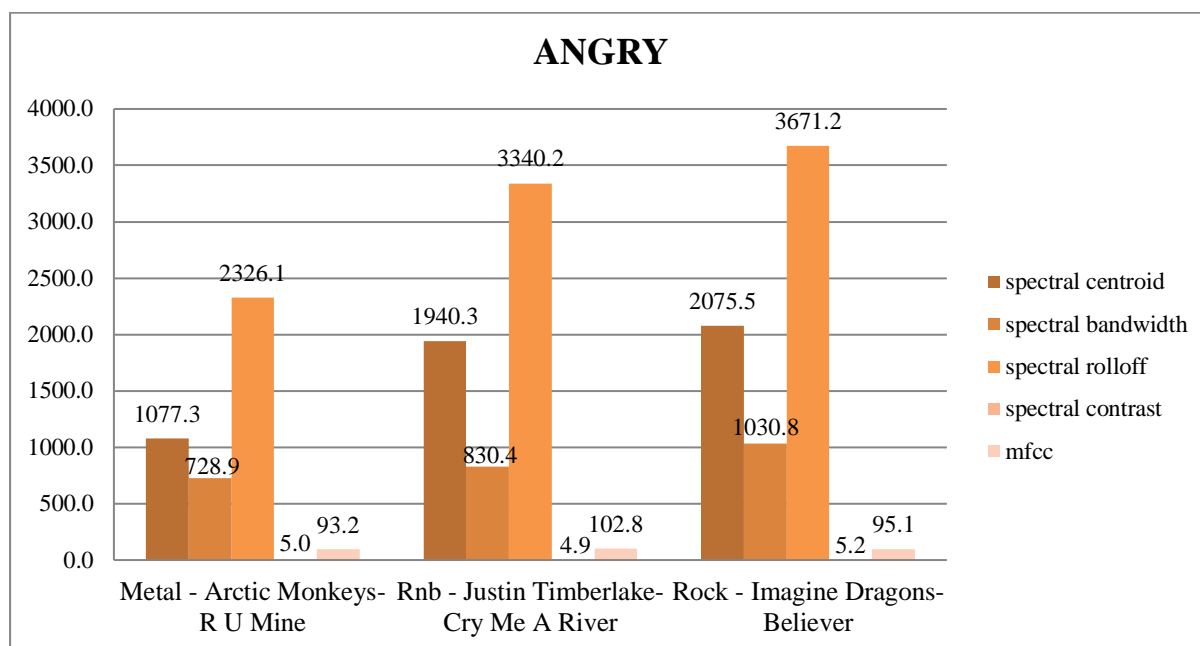
Η λοξότητα εκφράζει την εκτροπή μιας κατανομής από την κανονικότητα. Η εκτροπή αυτή μπορεί να εμφανίζεται είτε από δεξιά είτε από αριστερά της κατανομής, οπότε η λοξότητα χαρακτηρίζεται αντίστοιχα θετική ή αρνητική.

Για την αποφυγή σφαλμάτων στα αποτελέσματα του κώδικα ήταν απαραίτητο να διαμορφώσουμε τον κώδικα. Το σφάλμα που εμφανίστηκε 6 από τα 9 χαρακτηριστικά (εκτός από τα spectral contrast, zero crossing, MFCC) είναι το μήνυμα που εμφανίζεται στα αποτελέσματα του κώδικα: Nan nan nan nan nan. Το συγκεκριμένο σφάλμα είναι μια ειδική τιμή που αντιπροσωπεύει ελλιπή ή απροσδιόριστα δεδομένα της Python και παρουσιάστηκε στα αποτελέσματα της κύρτωσης και της λοξότητας.

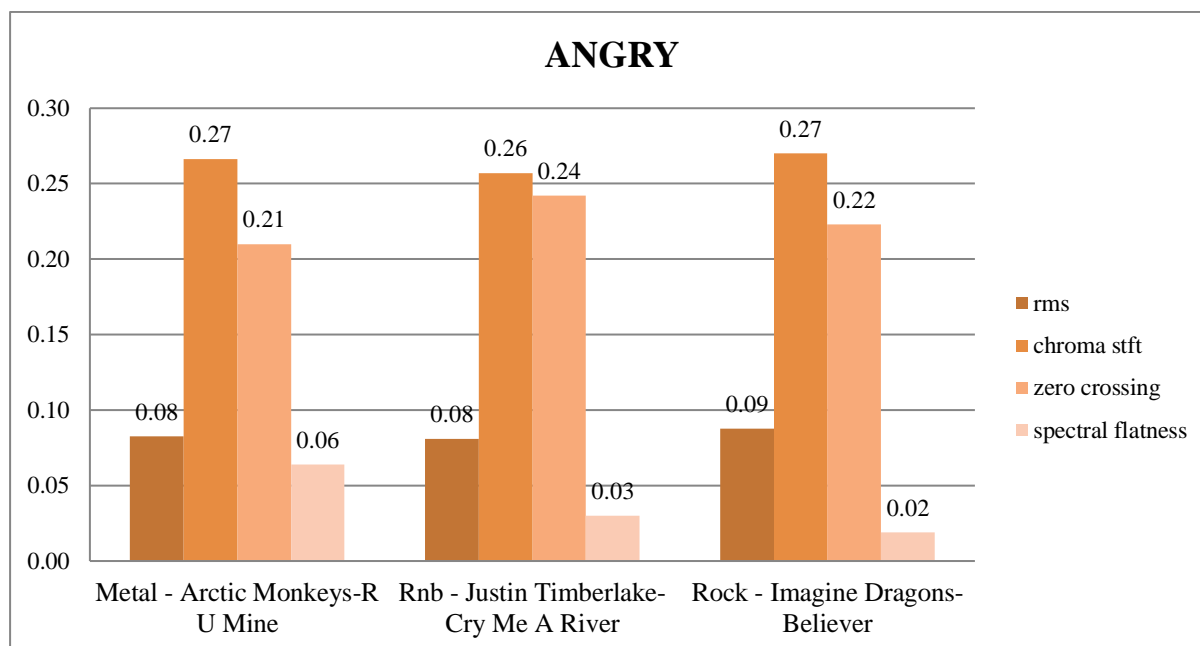
4. Αποτελέσματα

4.1 Εφαρμογή μεθόδου και Στατιστική ανάλυση

Όλα τα αποτελέσματα συγκεντρώθηκαν στο Microsoft Excel για να εξαχθούν τα γραφήματα. Στις παρακάτω εικόνες παρουσιάζονται τα γραφήματα για το συναίσθημα angry στα πέντε στατιστικά στοιχεία. Στα κομμάτια που παρατηρούνται ίδιες τιμές έχει γίνει στρογγυλοποίηση από το Microsoft Excel.

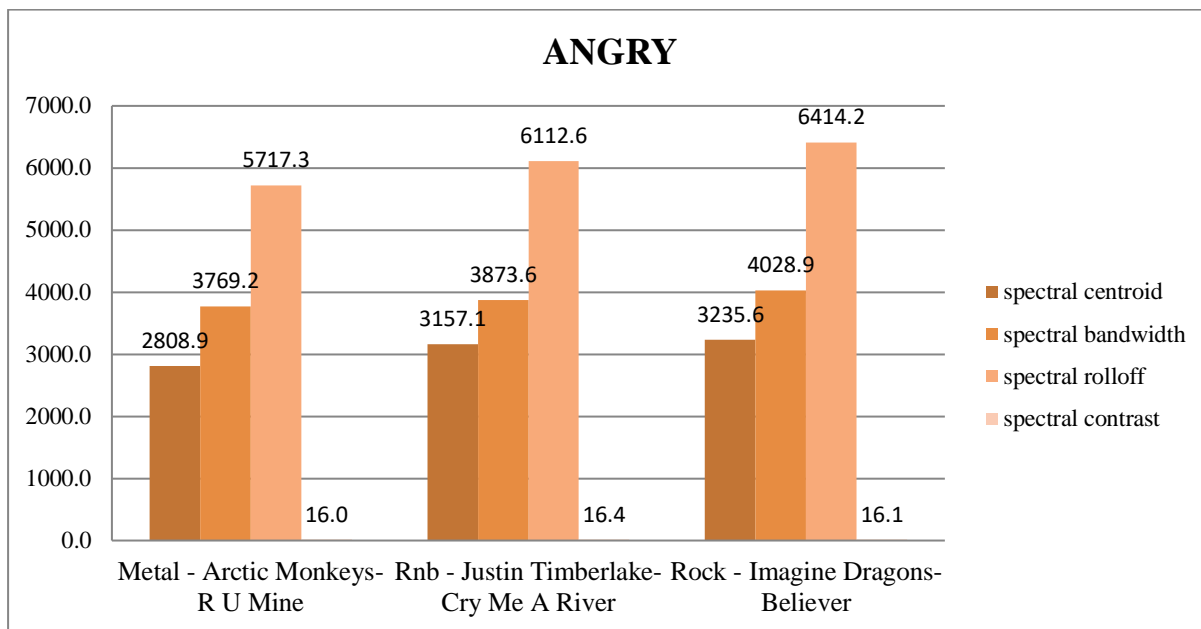


Εικόνα 4.1: Τυπική απόκλιση των χαρακτηριστικών για την κατηγορία angry

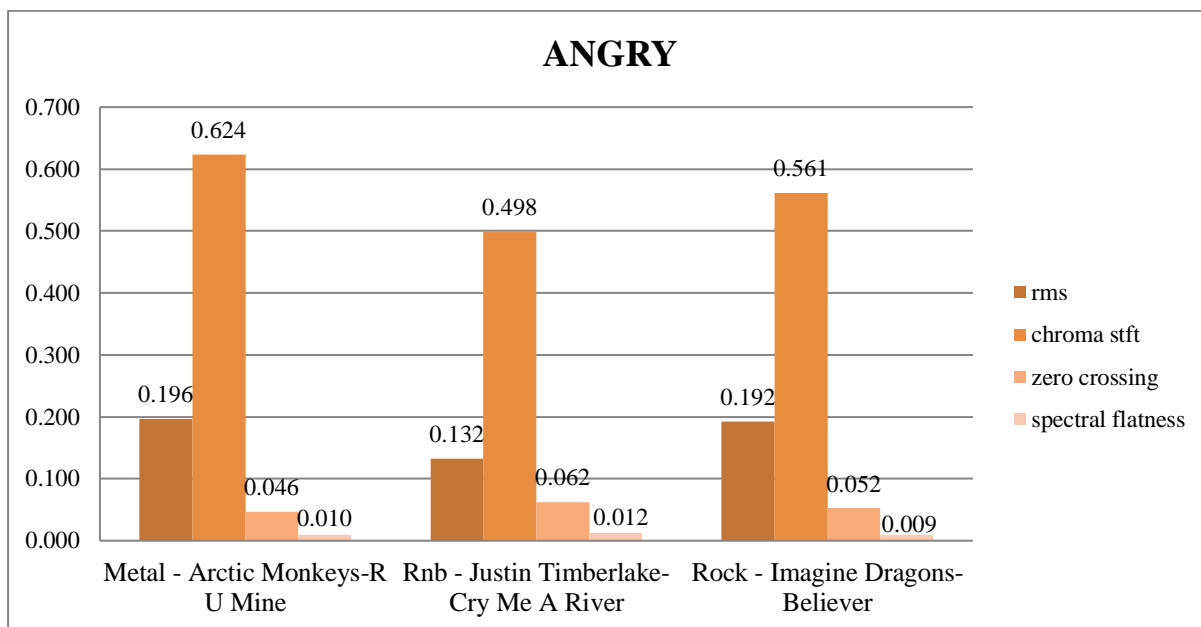


Εικόνα 4.2: Τυπική απόκλιση των χαρακτηριστικών για την κατηγορία angry

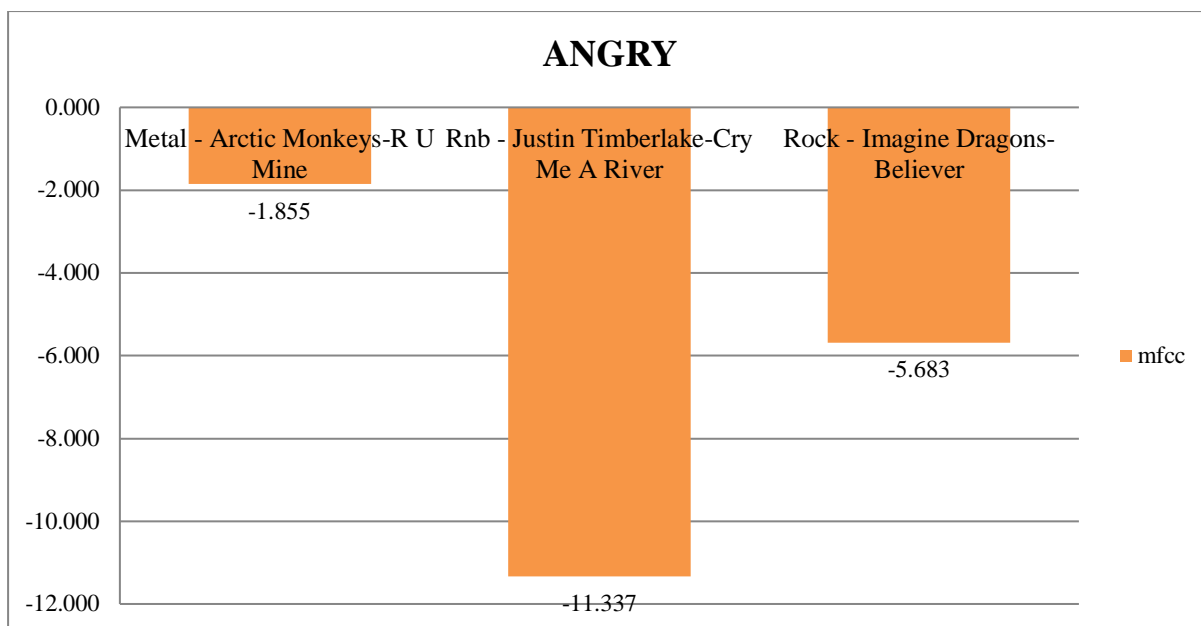
Σύμφωνα με τα διαγράμματα, οι τιμές των χαρακτηριστικών για τα τρία κομμάτια στην τυπική απόκλιση είναι σχετικά κοντά. Η υψηλότερη τιμή στο spectral centroid παρατηρείται στο Rock - Imagine Dragons-Believer καθώς και στα spectral bandwidth, spectral rolloff, spectral contrast, rms και chroma stft. Για τους mfcc και zero crossing, η υψηλότερη τιμή παρατηρείται στο Rnb - Justin Timberlake-Cry Me A River, ενώ στο spectral flatness η υψηλότερη τιμή εμφανίζεται στο Metal - Arctic Monkeys-R U Mine.



Εικόνα 4.3: Μέση τιμή των χαρακτηριστικών για την κατηγορία angry

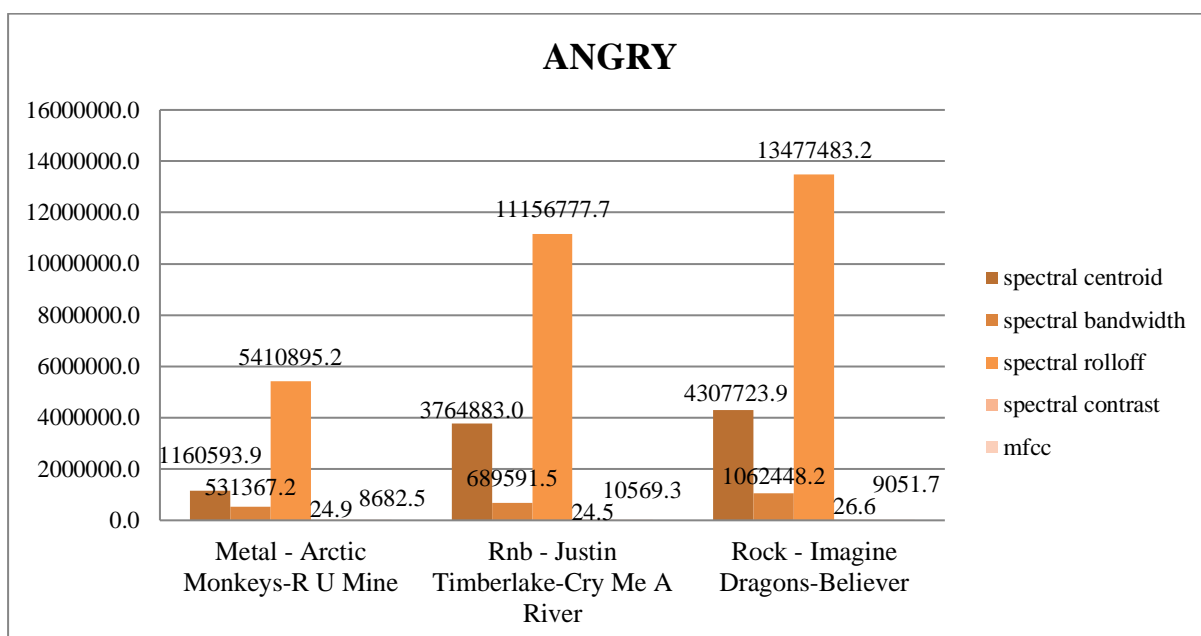


Εικόνα 4.4: Μέση τιμή των χαρακτηριστικών για την κατηγορία angry

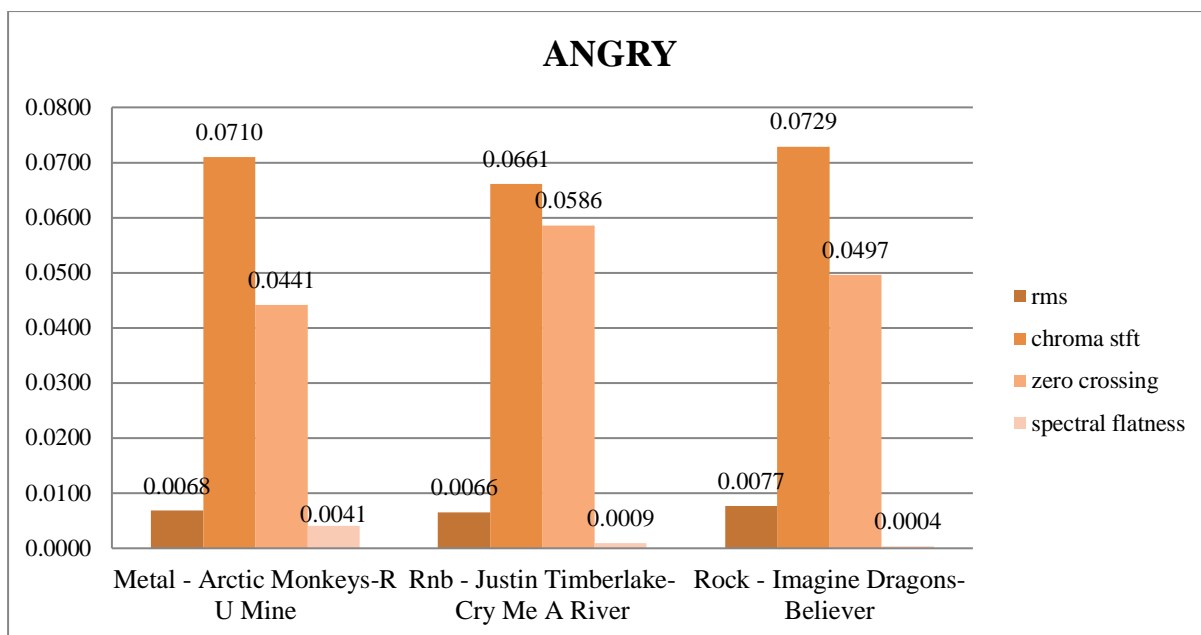


Εικόνα 4.5: Μέση τιμή των MFCC για την κατηγορία angry

Από τα διαγράμματα παρατηρείται ότι οι τιμές των τριών κομματιών για την μέση τιμή στα πρώτα τέσσερα χαρακτηριστικά είναι πολύ κοντά. Η υψηλότερη τιμή στο spectral centroid παρατηρείται στο Rock - Imagine Dragons-Believer καθώς και στα spectral bandwidth, spectral rolloff, ενώ στα spectral contrast, zero crossing και spectral flatness η υψηλότερη τιμή παρατηρείται στο Rnb - Justin Timberlake-Cry Me A River. Στην rms η υψηλότερη τιμή παρατηρείται στο Metal - Arctic Monkeys-R U Mine καθώς και στο chroma stft. Στους mfcc παρατηρούνται αρνητικές τιμές, με την υψηλότερη να παρατηρείται στο Rnb - Justin Timberlake-Cry Me A River.

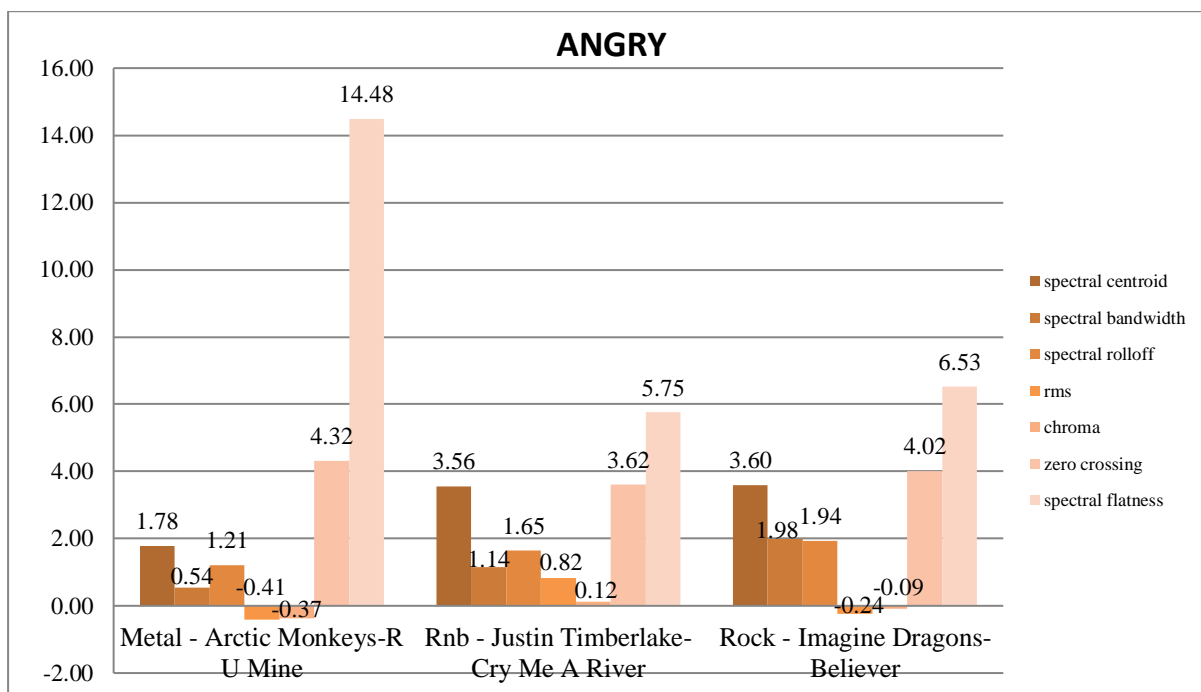


Εικόνα 4.6: Διακύμανση των χαρακτηριστικών για την κατηγορία angry



Εικόνα 4.7: Διακύμανση των χαρακτηριστικών για την κατηγορία angry

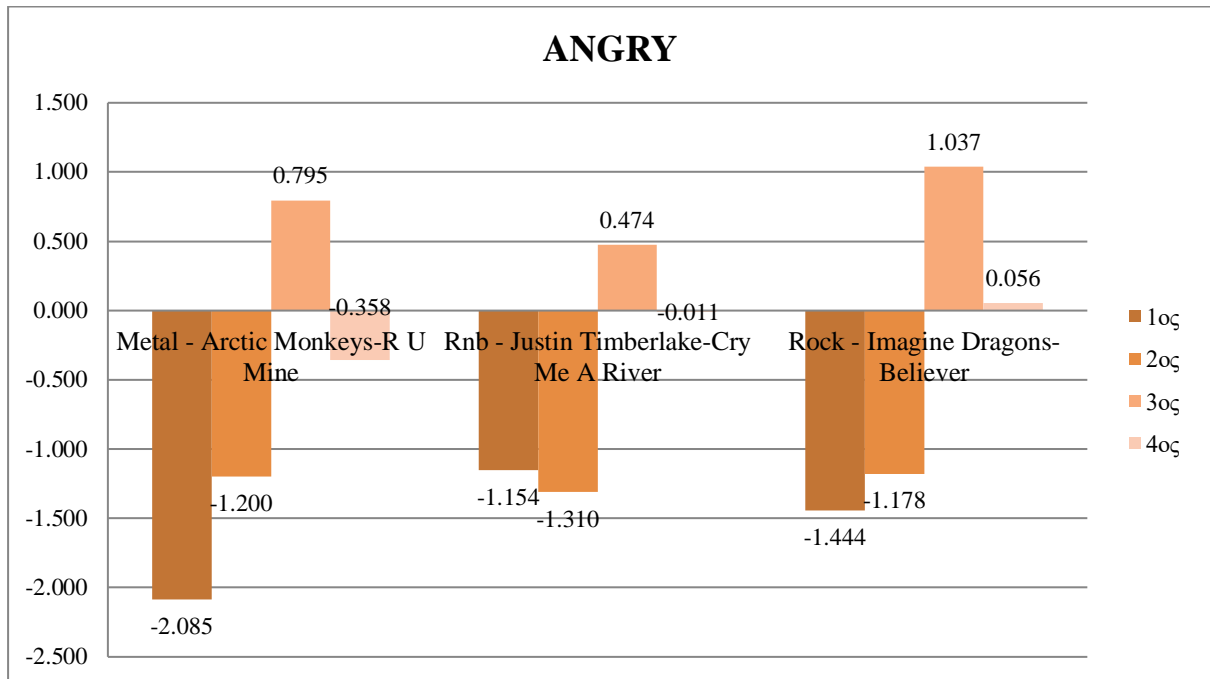
Από τα διαγράμματα παρατηρούμε ότι οι τιμές των τριών κομματιών για τη διακύμανση σε όλα τα χαρακτηριστικά, είναι πολύ κοντά. Η υψηλότερη τιμή στο spectral centroid παρατηρείται στο Metal - Arctic Monkeys-R U Mine, ενώ στα spectral bandwidth, spectral rolloff, spectral contrast, rms και chroma stft παρατηρείται στο Rock - Imagine Dragons-Believer. Η υψηλότερη τιμή στους mfcc παρατηρείται στο Rnb - Justin Timberlake-Cry Me A River καθώς και στα zero crossing και spectral flatness.



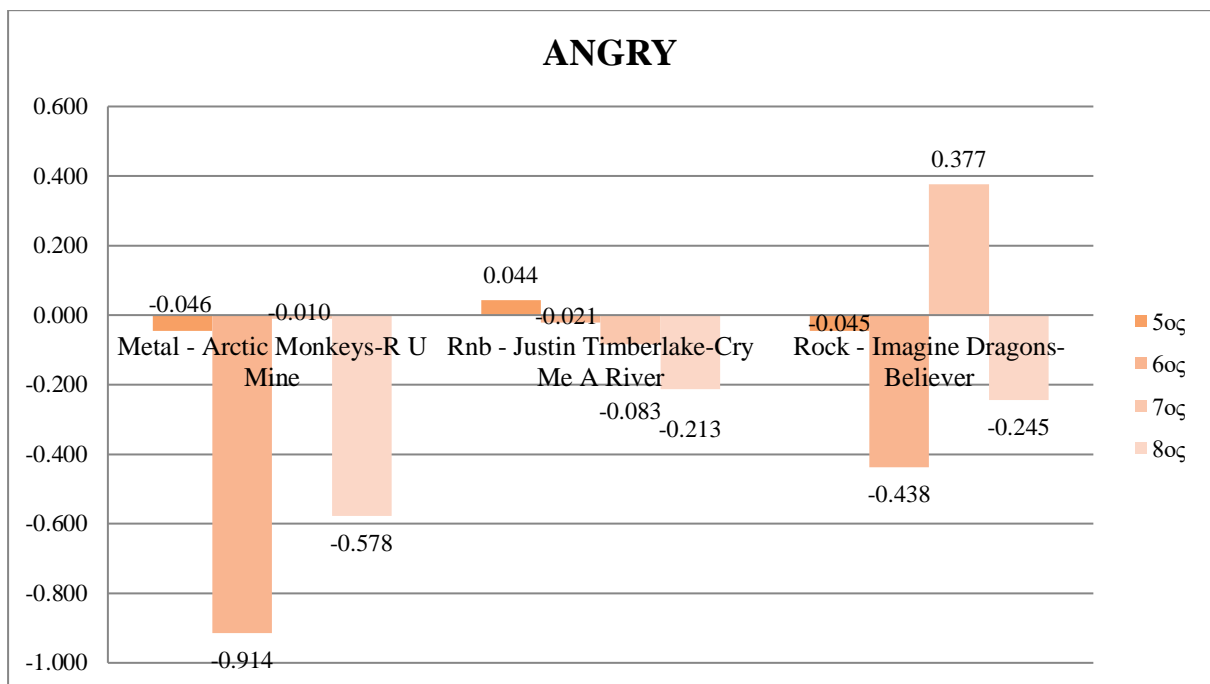
Εικόνα 4.8: Κύρτωση των χαρακτηριστικών για την κατηγορία angry

Σύμφωνα με τα διαγράμματα παρατηρούμε ότι οι τιμές των τριών κομματιών για την κύρτωση στα πρώτα επτά χαρακτηριστικά είναι σχετικά κοντά. Η υψηλότερη τιμή στο spectral centroid παρατηρείται στο Rock - Imagine Dragons-Believer καθώς και στα spectral bandwidth και spectral rolloff. Στην rms η υψηλότερη τιμή παρατηρείται στο Rnb - Justin Timberlake-Cry Me A River

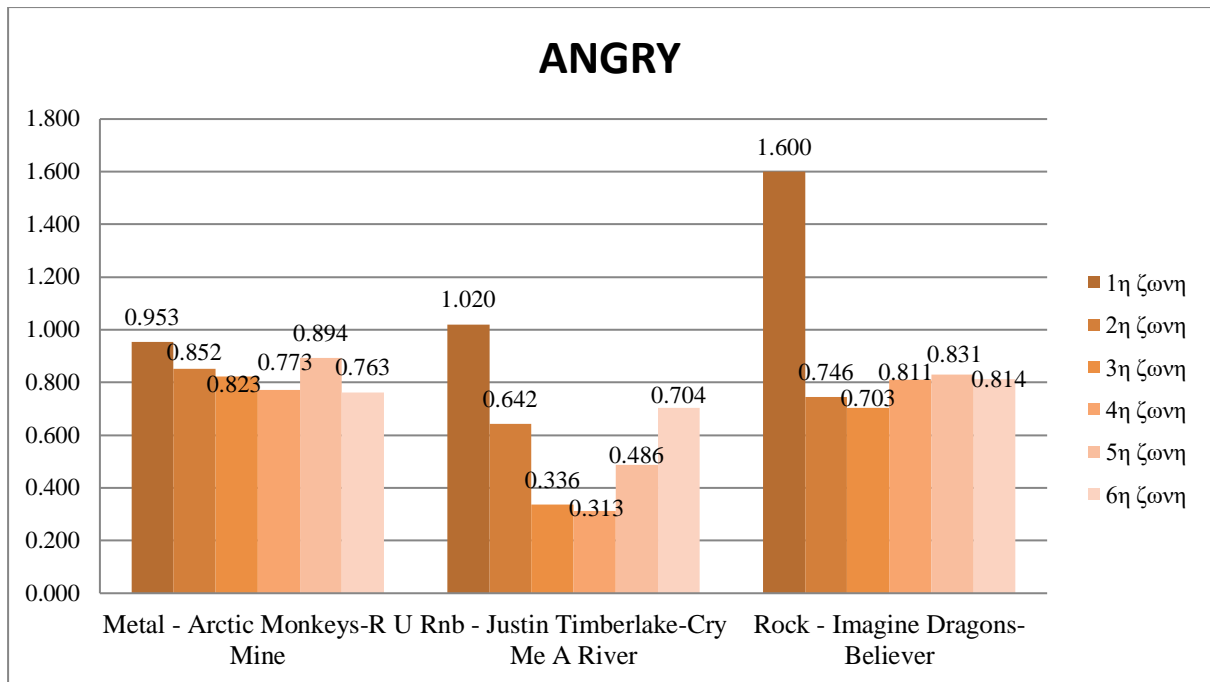
καθώς και στο chroma stft. Στο zero crossing, η υψηλότερη τιμή παρατηρείται στο Metal - Arctic Monkeys-R U Mine καθώς και στο spectral flatness.



Εικόνα 4.9: Κύρτωση των MFCC (1^{ος} έως 4^{ος}) για την κατηγορία angry

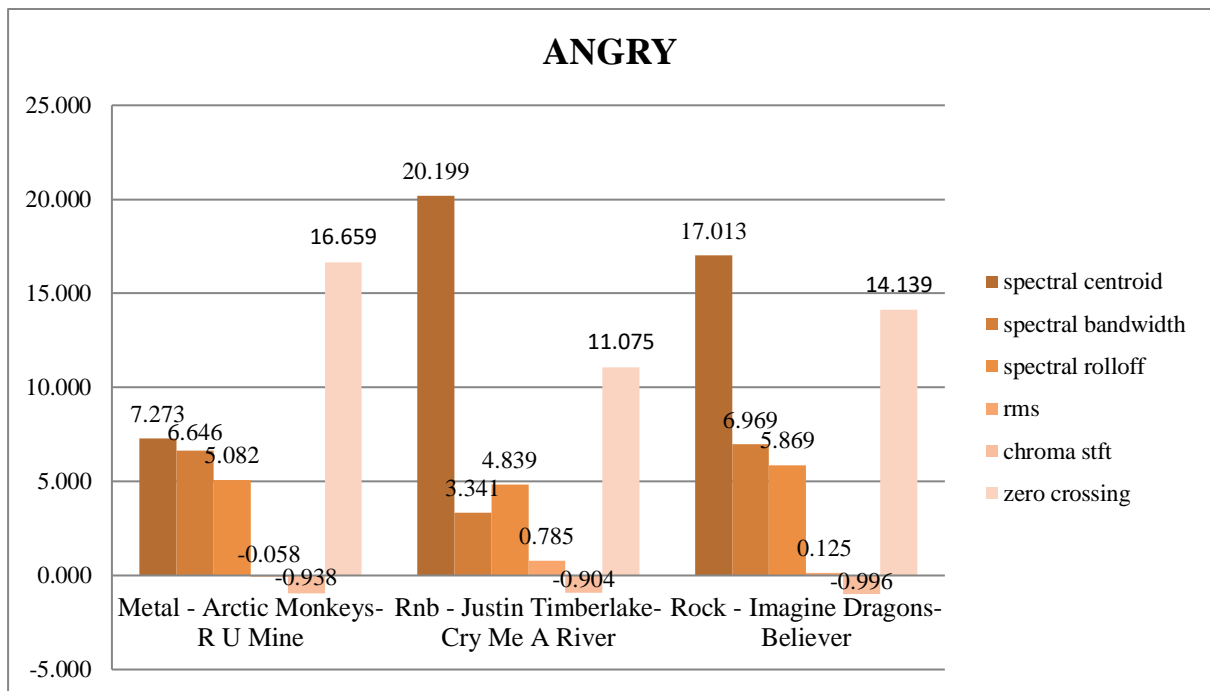


Εικόνα 4.10: Κύρτωση των MFCC (5^{ος} έως 8^{ος}) για την κατηγορία angry

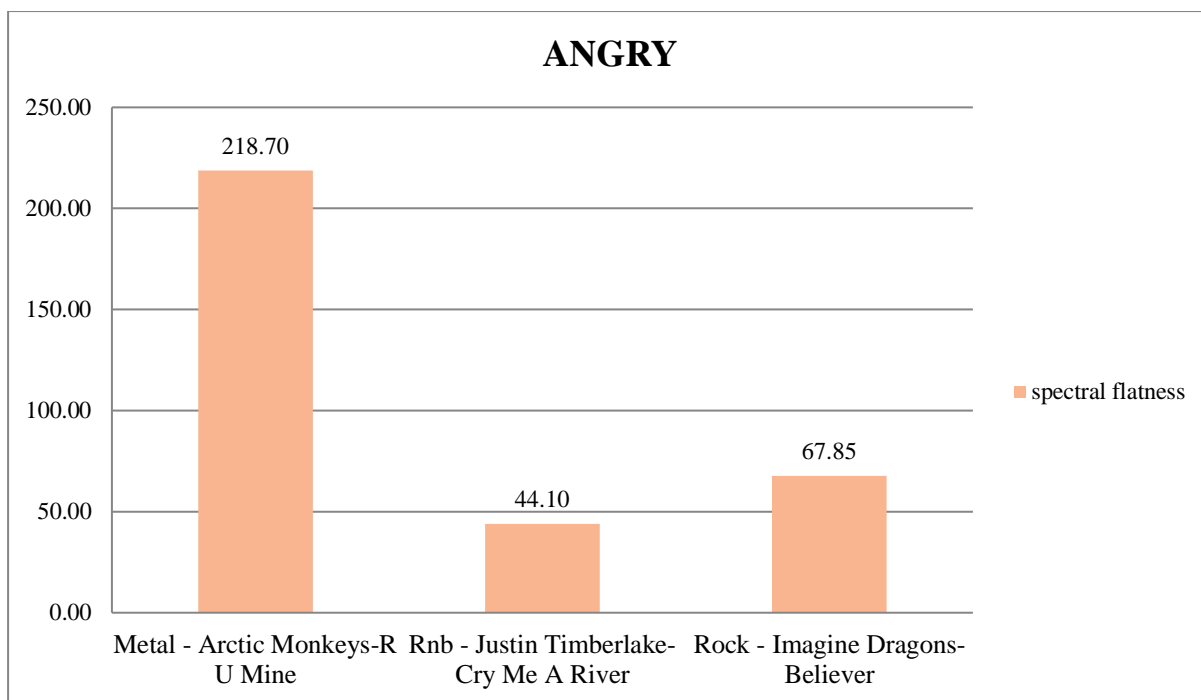


Εικόνα 4.11: Κύρτωση του spectral contrast στις 6 ζώνες για την κατηγορία angry

Όσον αφορά τους mfcc παρατηρούνται αρνητικές τιμές. Ο 1^{ος} και ο 2^{ος} mfcc έχουν υψηλότερες τιμές στο Metal - Arctic Monkeys-R U Mine, ενώ ο 3^{ος} και 4^{ος} στο Rock - Imagine Dragons-Believer. Ο 5^{ος} mfcc έχει υψηλότερη τιμή στο Rnb - Justin Timberlake-Cry Me A River. Στον 6^ο και 8^ο mfcc παρατηρείται υψηλότερη τιμή στο Metal - Arctic Monkeys-R U Mine, ενώ στον 7^ο στο Rock - Imagine Dragons-Believer. Στην περίπτωση του spectral contrast οι τιμές στις έξι ζώνες είναι πολύ κοντά και στα τρία κομμάτια. Στην 1^η, 4^η και 6^η ζώνη η υψηλότερη τιμή παρατηρείται στο Rock - Imagine Dragons-Believer ενώ στην 2^η, 3^η και 5^η ζώνη στο Metal - Arctic Monkeys-R U Mine.

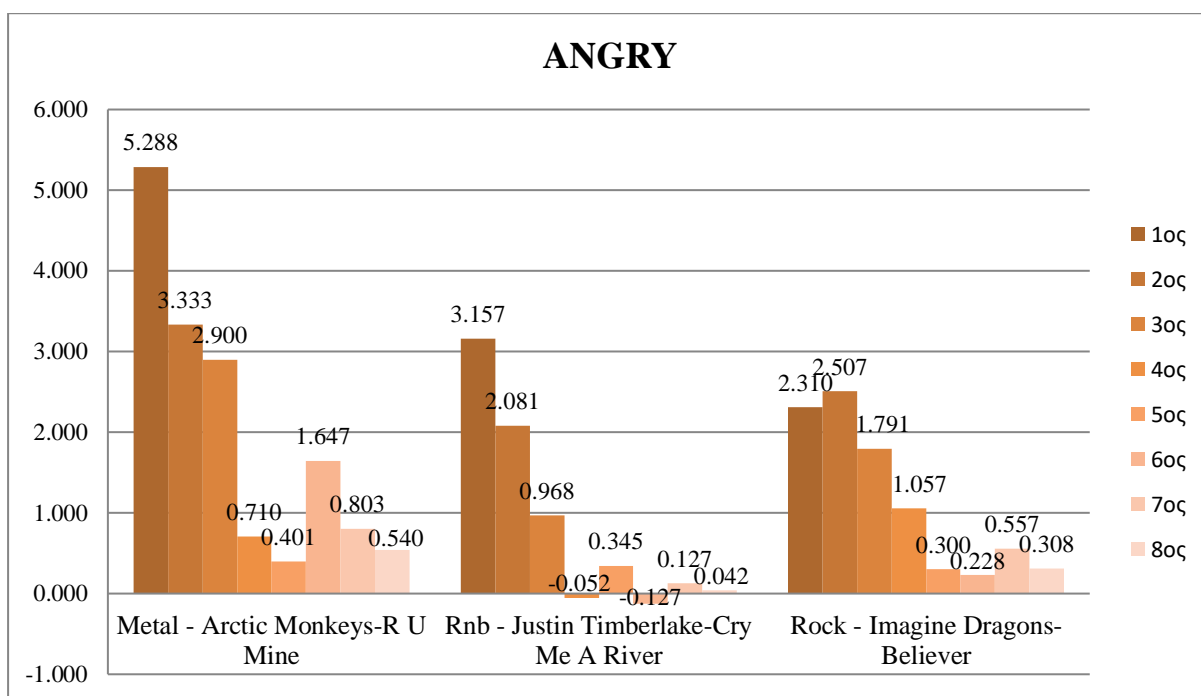


Εικόνα 4.12: Λοξότητα των χαρακτηριστικών για την κατηγορία angry

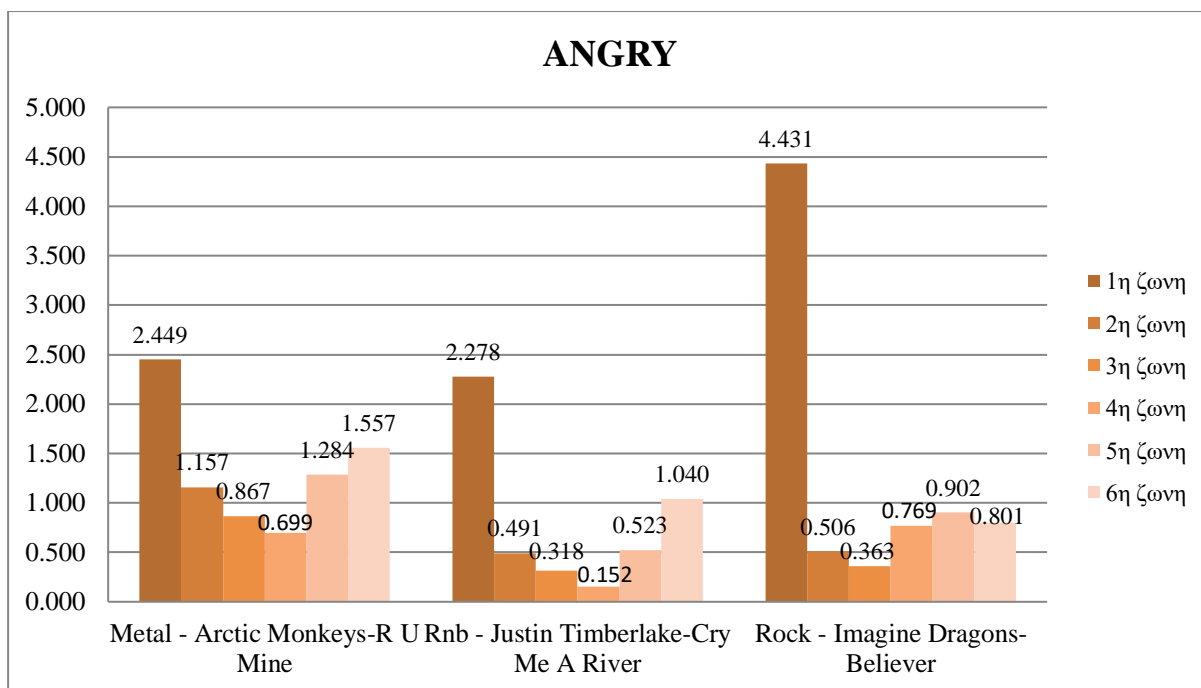


Εικόνα 4.13: Λοξότητα του spectral flatness για την κατηγορία angry

Από τα διαγράμματα παρατηρούμε ότι οι τιμές των τριών κομματιών για την λοξότητα στα πρώτα επτά χαρακτηριστικά παρουσιάζουν μία μικρή απόκλιση. Η υψηλότερη τιμή στο spectral centroid παρατηρείται στο Rnb - Justin Timberlake-Cry Me A River καθώς και στην rms. Στα spectral bandwidth, spectral rolloff, zero crossing και chroma stft η υψηλότερη τιμή παρατηρείται στο Rock - Imagine Dragons-Believer. Στο spectral flatness η υψηλότερη τιμή παρατηρείται στο Metal - Arctic Monkeys-R U Mine.



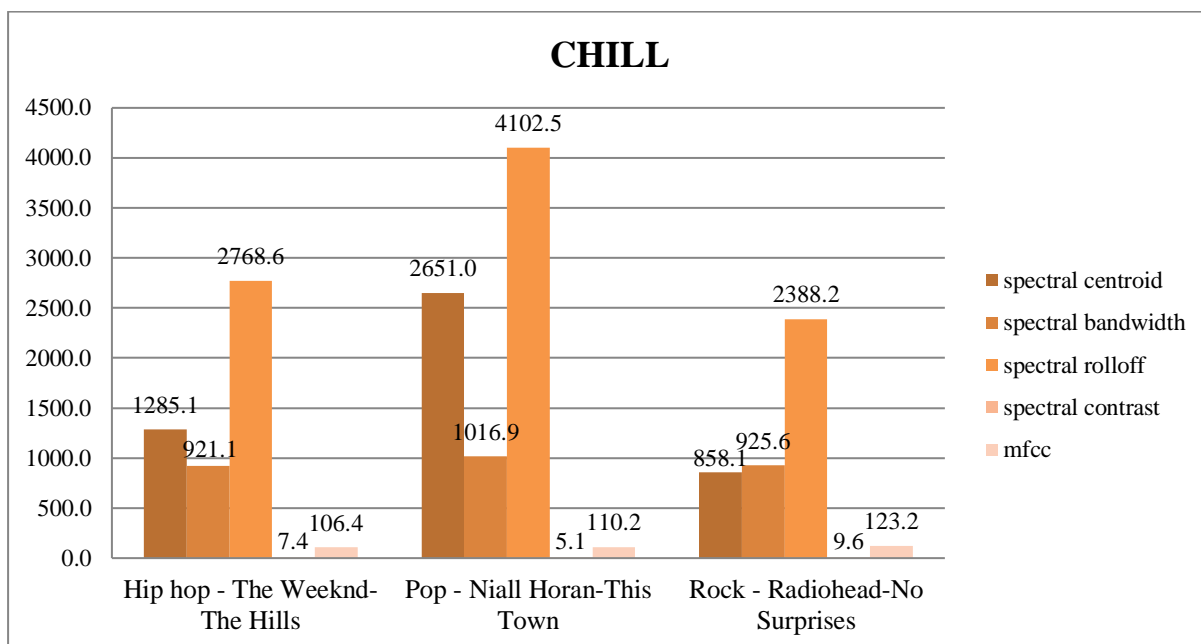
Εικόνα 4.14: Λοξότητα των MFCC (οι πρώτοι 8) για την κατηγορία angry



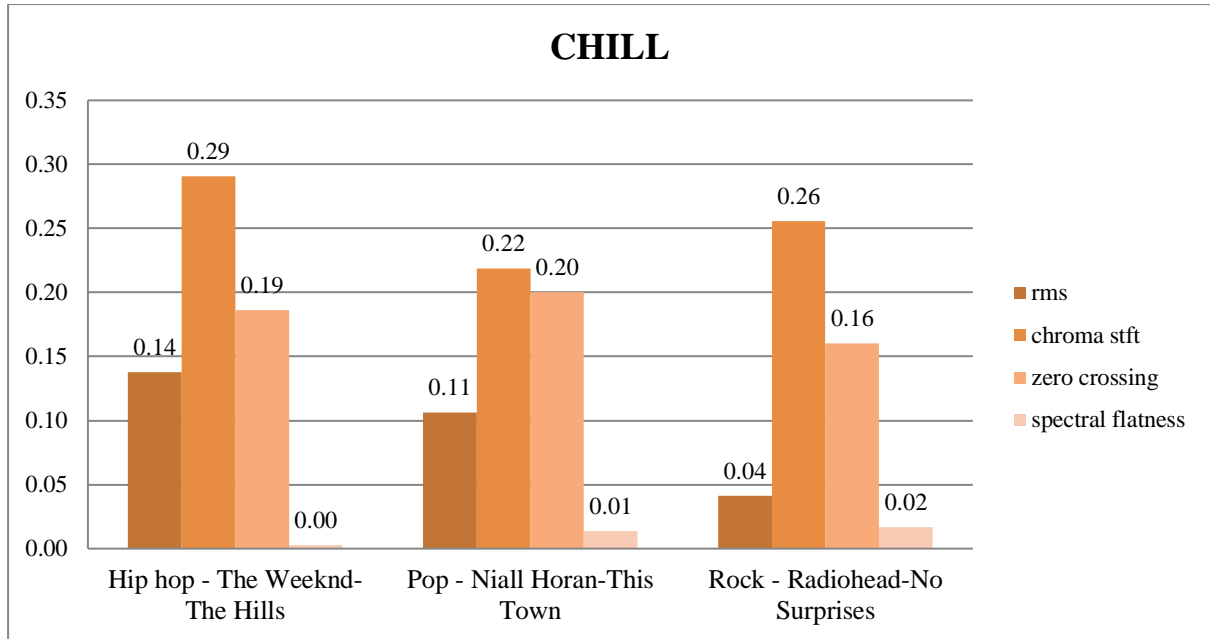
Εικόνα 4.15: Λοξότητα των spectral contrast στις 6 ζώνες για την κατηγορία angry

Όσον αφορά τους mfcc οι τιμές είναι αρκετά κοντά. Ο 1^{ος}, ο 2^{ος}, ο 3^{ος}, ο 5^{ος}, 6^{ος}, 7^{ος} και 8^{ος} mfcc έχουν υψηλότερες τιμές στο Metal - Arctic Monkeys-R U Mine, ενώ ο 4^{ος} στο Rock - Imagine Dragons-Believer. Στην περίπτωση του spectral contrast οι τιμές στις έξι ζώνες είναι πολύ κοντά και στα τρία κομμάτια. Στην 1^η και 4^η ζώνη η υψηλότερη τιμή παρατηρείται στο Rock - Imagine Dragons-Believer ενώ στην 2^η, 3^η, 5^η και 6^η στο Metal - Arctic Monkeys-R U Mine.

Στις παρακάτω εικόνες παρουσιάζονται τα γραφήματα για το συναίσθημα chill στα πέντε στατιστικά στοιχεία.

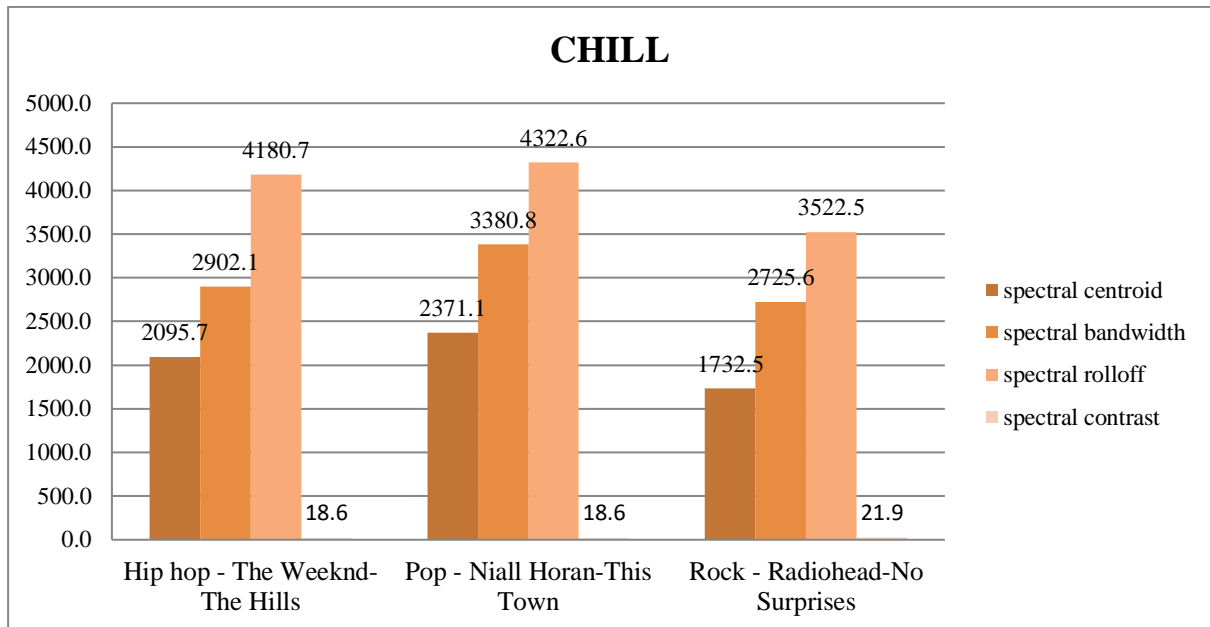


Εικόνα 4.16: Τυπική απόκλιση των χαρακτηριστικών για την κατηγορία chill

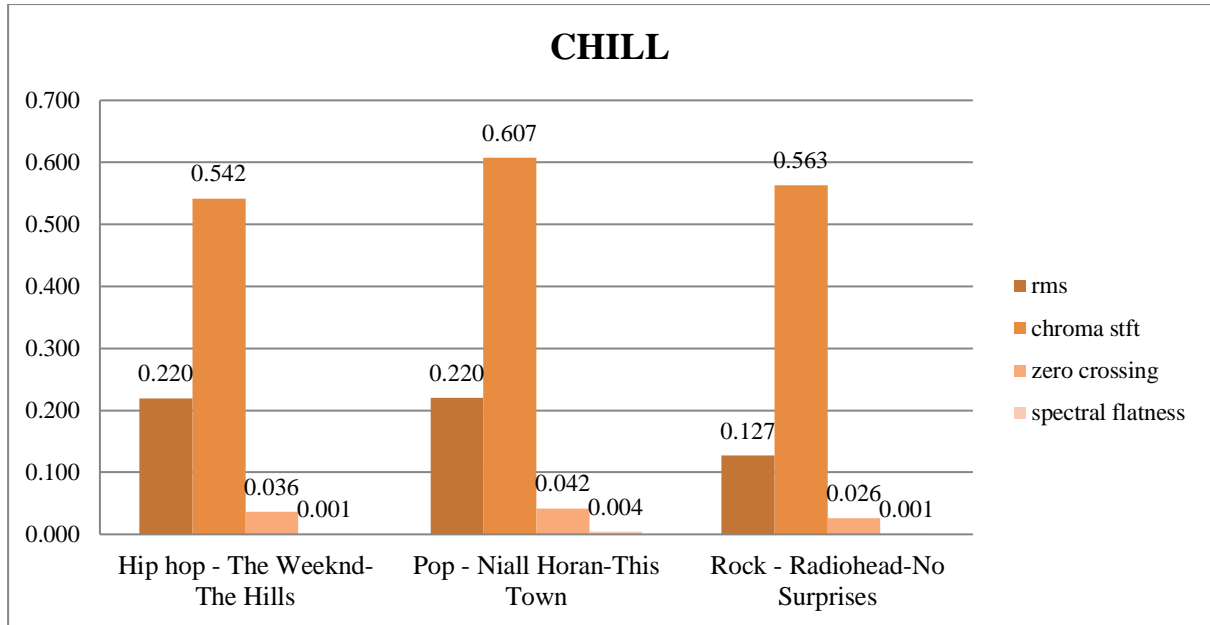


Εικόνα 4.17: Τυπική απόκλιση των χαρακτηριστικών για την κατηγορία chill

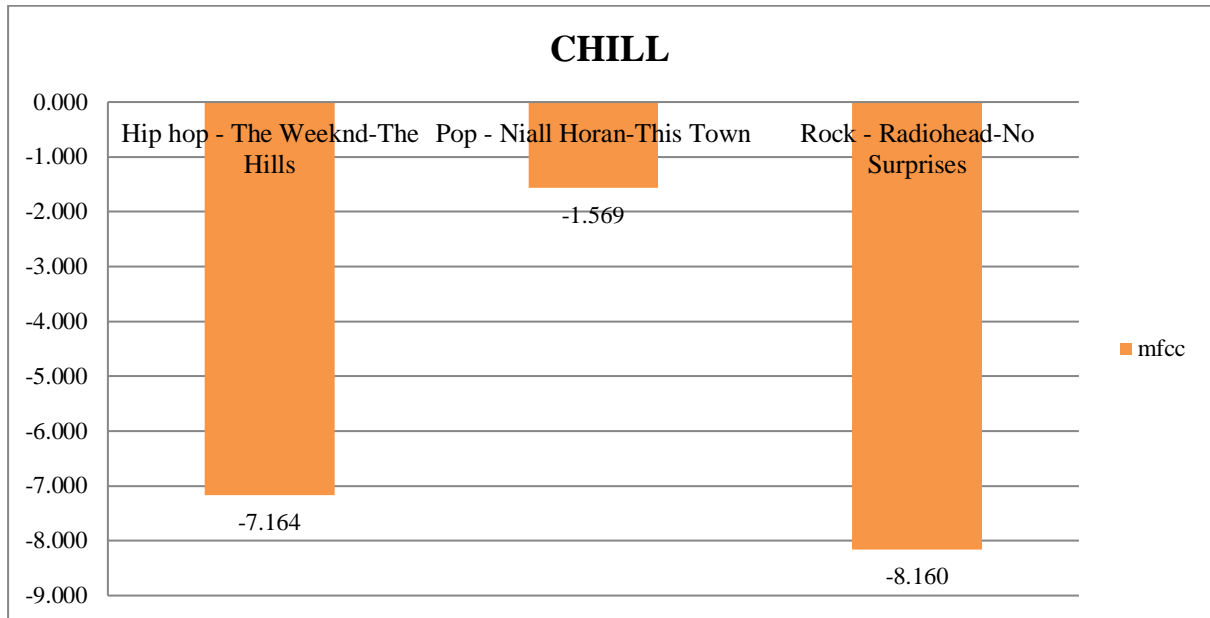
Όπως βλέπουμε, στις παραπάνω εικόνες οι τιμές των τριών κομματιών για την τυπική απόκλιση σε όλα τα χαρακτηριστικά είναι σχετικά κοντά. Η υψηλότερη τιμή στο spectral centroid παρατηρείται στο Pop - Niall Horan-This Town καθώς και στα spectral bandwidth, spectral rolloff και zero crossing. Στην rms η υψηλότερη τιμή παρατηρείται στο Hip hop - The Weeknd-The Hills καθώς και στο chroma stft. Στα spectral contrast, mfcc και spectral flatness η υψηλότερη τιμή παρατηρείται στο Rock - Radiohead-No Surprises.



Εικόνα 4.18: Μέση τιμή των χαρακτηριστικών για την κατηγορία chill

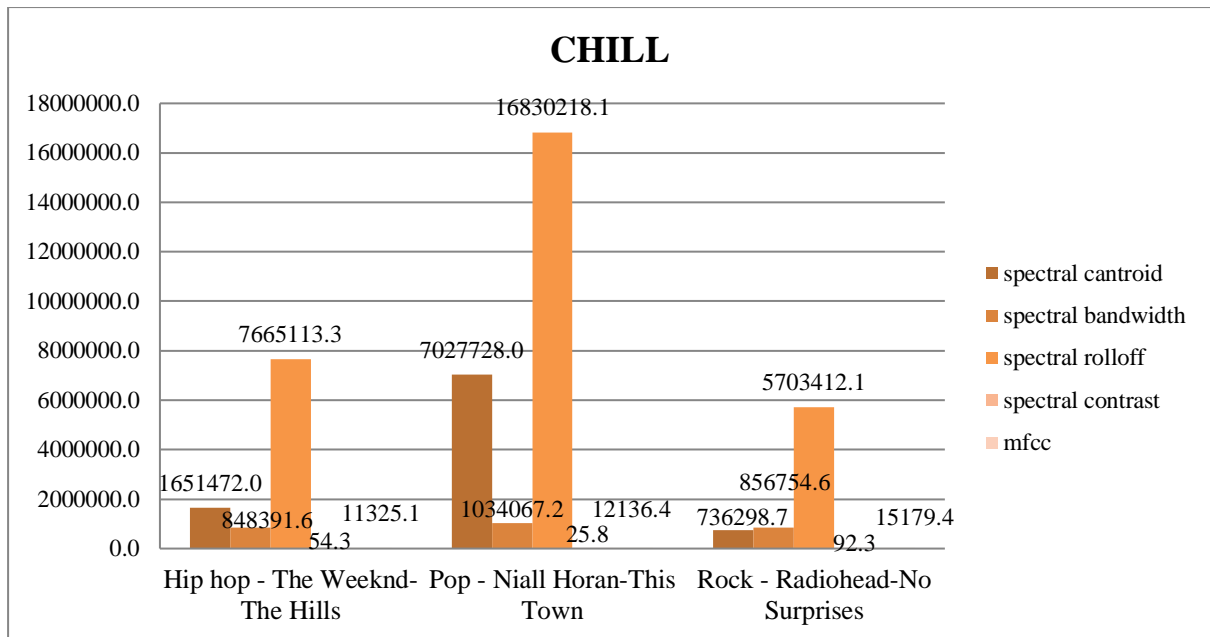


Εικόνα 4.19: Μέση τιμή των χαρακτηριστικών για την κατηγορία chill

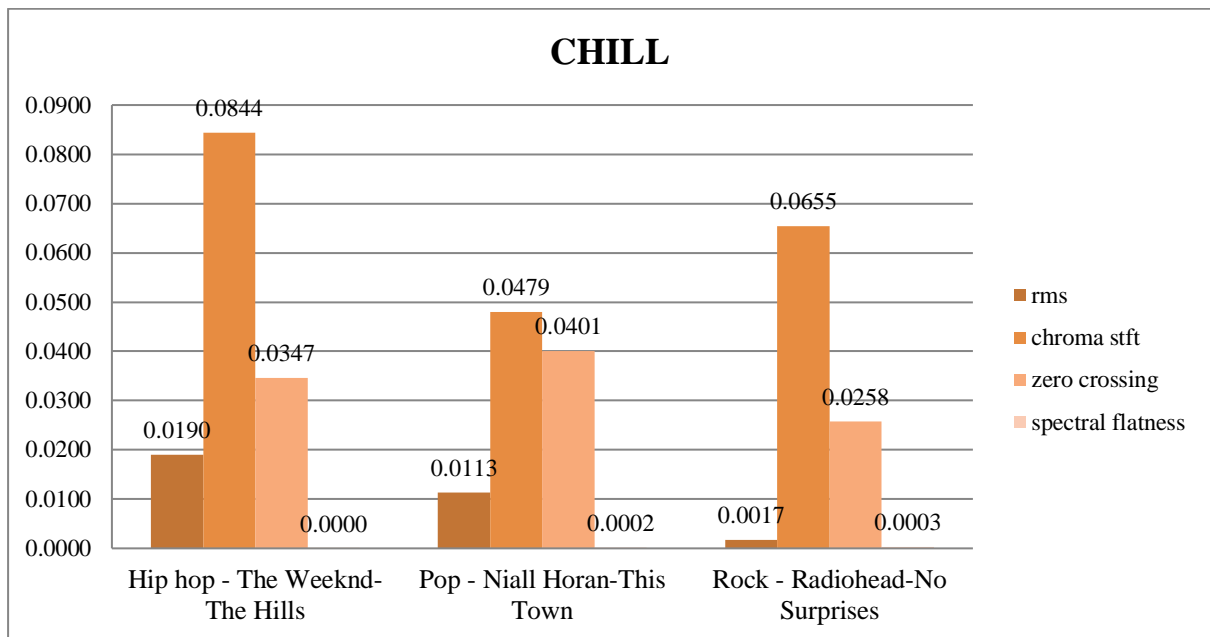


Εικόνα 4.20: Μέση τιμή των MFCC για την κατηγορία chill

Από τα διαγράμματα παρατηρούμε ότι οι τιμές των τριών κομματιών για την μέση τιμή στα πρώτα τέσσερα χαρακτηριστικά είναι πολύ κοντά. Η υψηλότερη τιμή στο spectral centroid παρατηρείται στο Pop - Niall Horan-This Town καθώς και στα spectral bandwidth, spectral rolloff, rms, chroma stft, zero crossing και spectral flatness. Στα spectral contrast και mfcc η υψηλότερη τιμή παρατηρείται στο Rock - Radiohead-No Surprises.

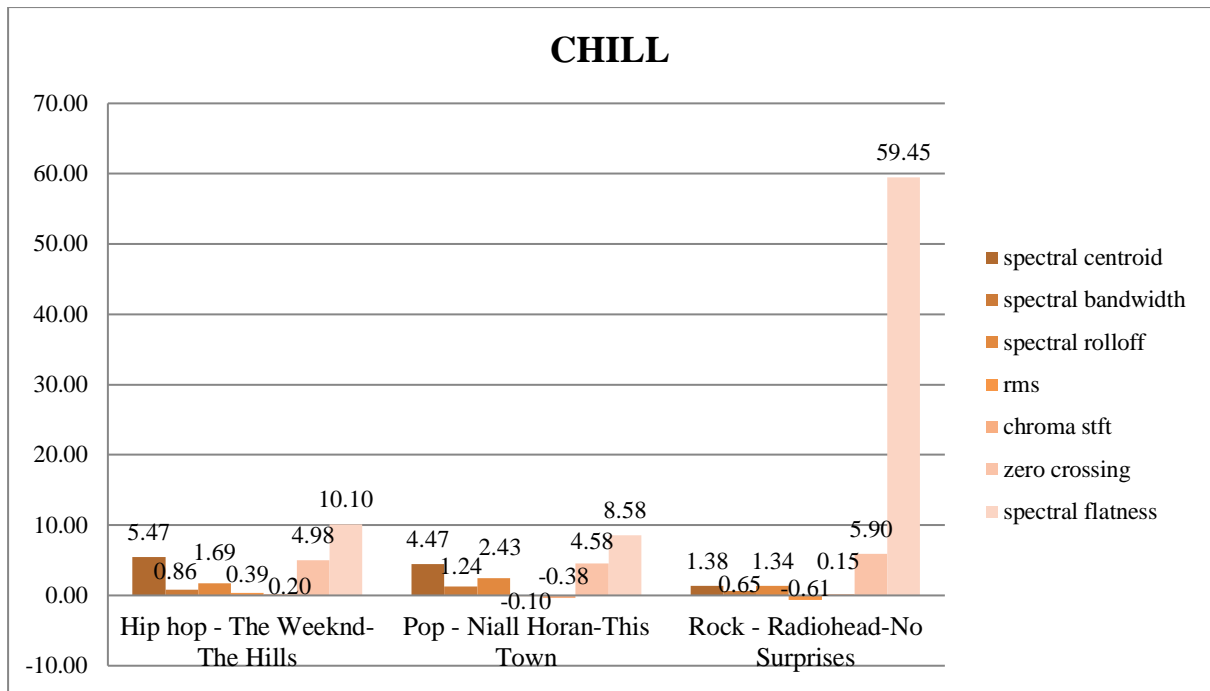


Εικόνα 4.21: Διακύμανση των χαρακτηριστικών για την κατηγορία chill



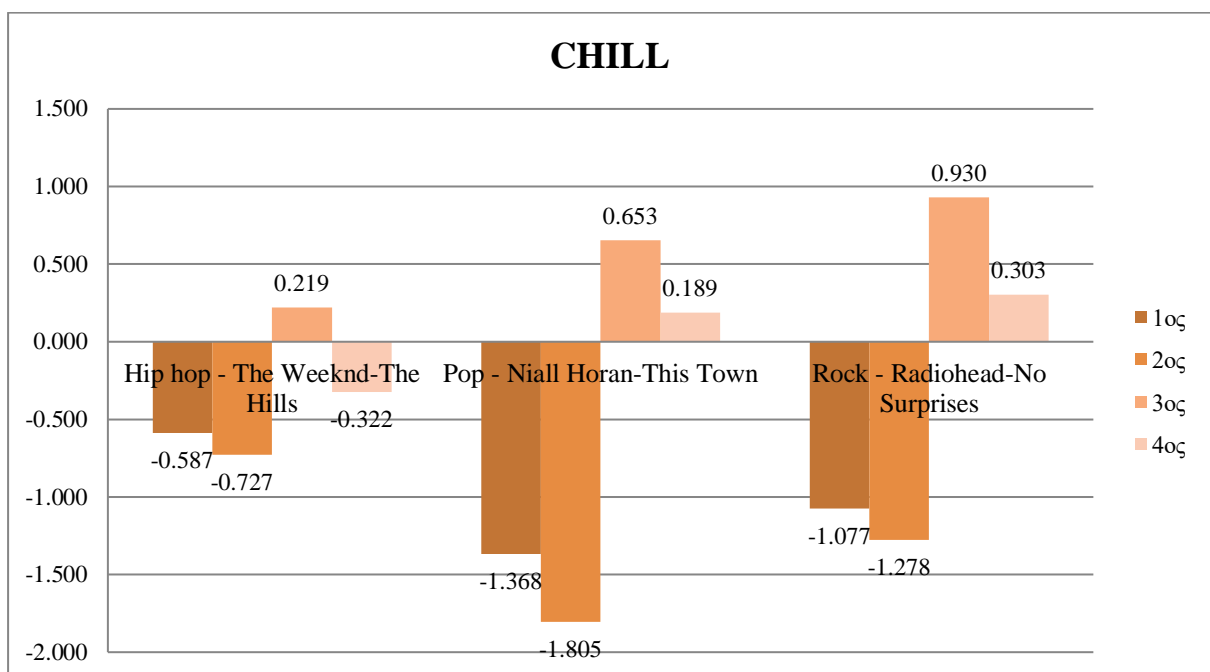
Εικόνα 4.22: Διακύμανση των χαρακτηριστικών για την κατηγορία chill

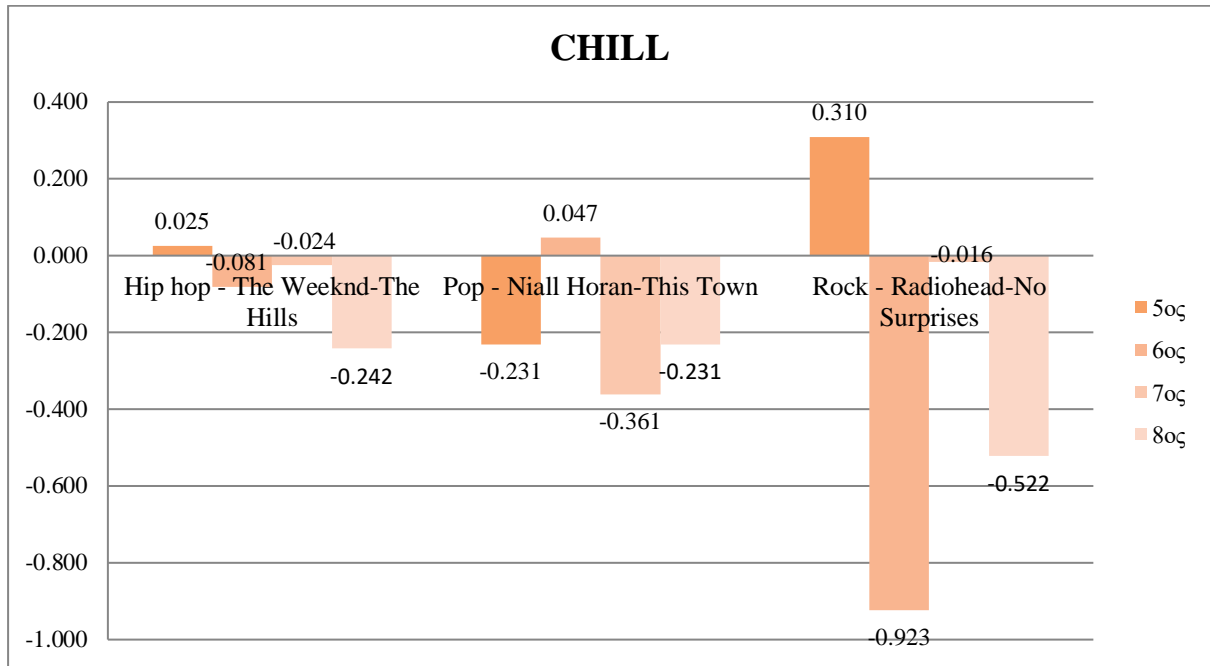
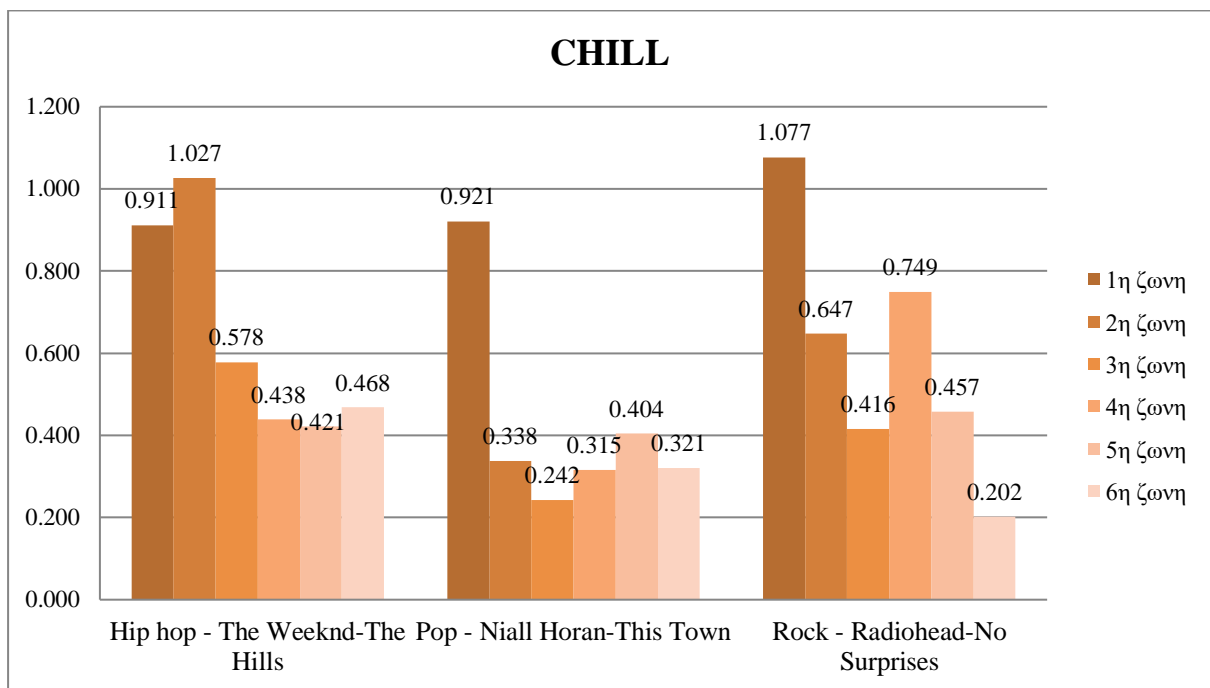
Σύμφωνα με τα διαγράμματα παρατηρούμε ότι οι τιμές των τριών κομματιών για την διακύμανση στα πρώτα τέσσερα χαρακτηριστικά παρουσιάζουν αρκετή απόκλιση σε αντίθεση με τα επόμενα τέσσερα που είναι πολύ κοντά. Η υψηλότερη τιμή στο spectral centroid παρατηρείται στο Pop - Niall Horan-This Town καθώς και στα spectral bandwidth, spectral rolloff και zero crossing. Στα spectral contrast, mfcc και spectral flatness η υψηλότερη τιμή παρατηρείται στο Rock - Radiohead-No Surprises. Στην rms η υψηλότερη τιμή παρατηρείται στο Hip hop - The Weeknd-The Hills καθώς και στο chroma stft.



Εικόνα 4.23: Κύρτωση των χαρακτηριστικών για την κατηγορία chill

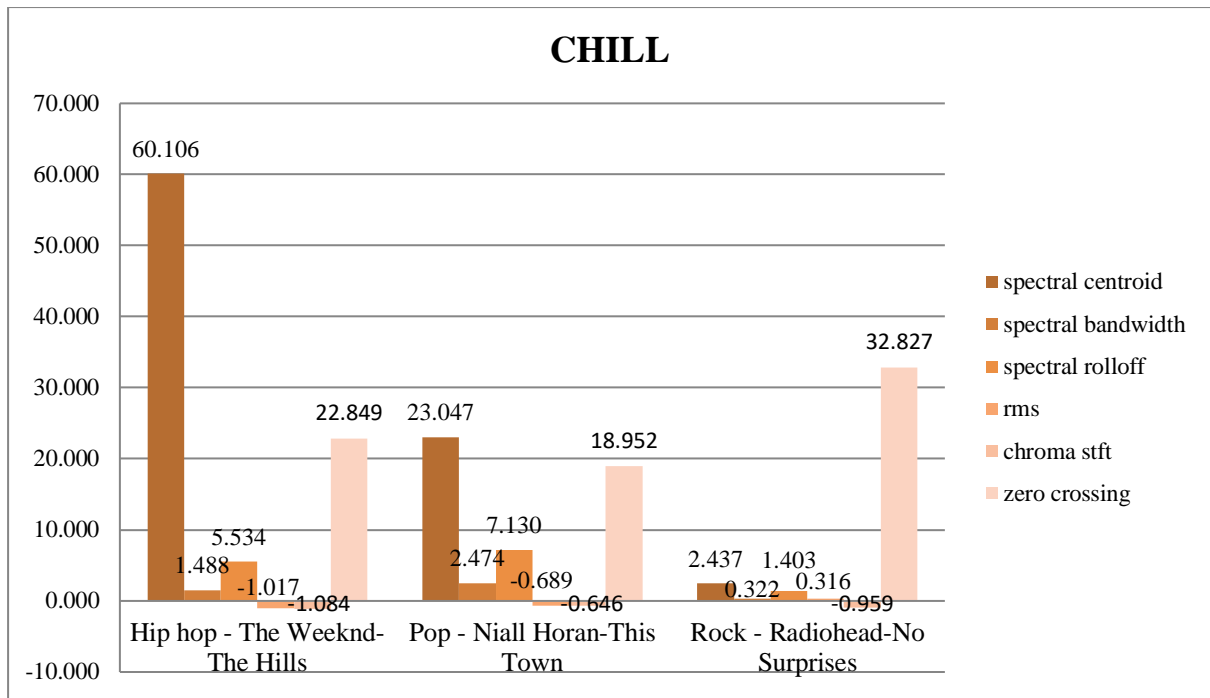
Από τα διαγράμματα παρατηρούμε ότι οι τιμές των τριών κομματιών για την κύρτωση στα πρώτα επτά χαρακτηριστικά είναι σχετικά κοντά. Η υψηλότερη τιμή στο spectral centroid, rms και chroma stft παρατηρείται στο Hip hop - The Weeknd-The Hills. Στα spectral bandwidth και spectral rolloff η υψηλότερη τιμή παρατηρείται στο Pop - Niall Horan-This Town. Στα zero crossing και spectral flatness η υψηλότερη τιμή παρατηρείται στο Rock - Radiohead-No Surprises.

Εικόνα 4.24: Κύρτωση των MFCC (1^{ος} έως 4^{ος}) για την κατηγορία chill

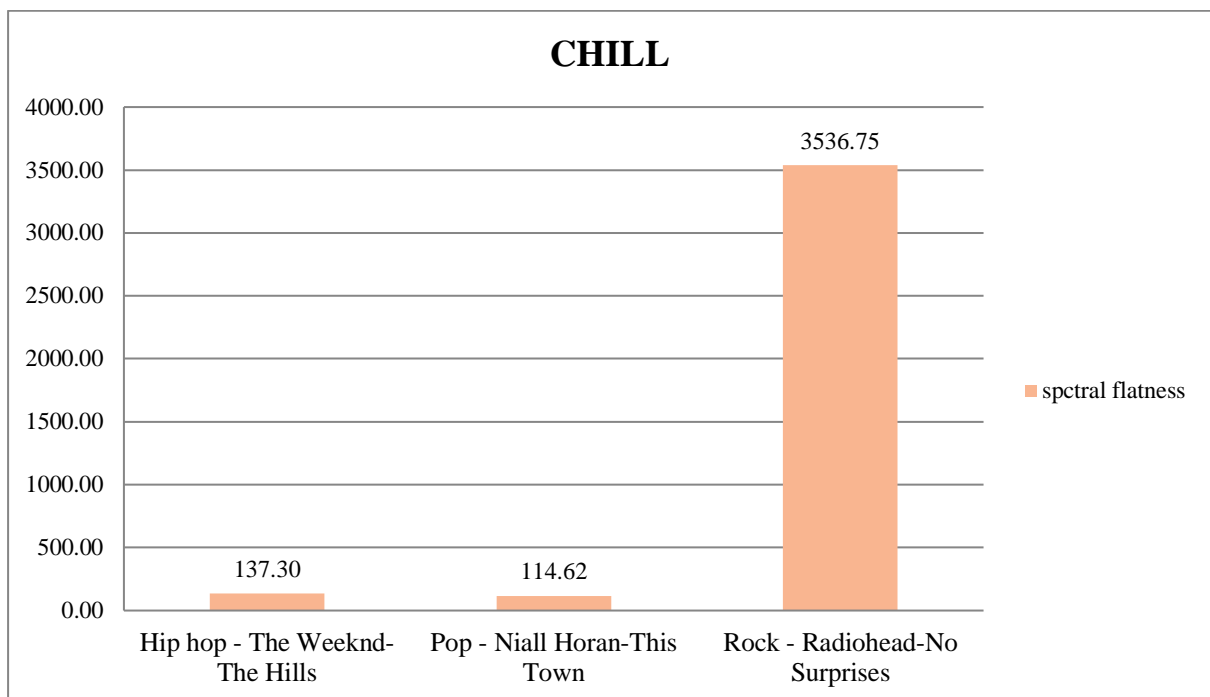
Εικόνα 4.25: Κύρτωση των MFCC (5^{ος} έως 8^{ος}) για την κατηγορία chill

Εικόνα 4.26: Κύρτωση του spectral contrast στις 6 ζώνες για την κατηγορία chill

Όσον αφορά τους mfcc παρατηρούνται και αρνητικές και θετικές τιμές. Στον 1^ο, 2^ο και 7^ο mfcc η υψηλότερη τιμή εμφανίζεται στο Pop - Niall Horan-This Town, ενώ στον 3^ο, 4^ο, 5^ο, 6^ο και 8^ο εμφανίζεται στο Rock - Radiohead-No Surprises. Στην περίπτωση του spectral contrast, στην 1^η, 4^η και 5^η ζώνη η υψηλότερη τιμή παρουσιάζεται στο Rock - Radiohead-No Surprises ενώ στη 2^η, 3^η και 6^η ζώνη εμφανίζεται στο Hip hop - The Weeknd-The Hills.

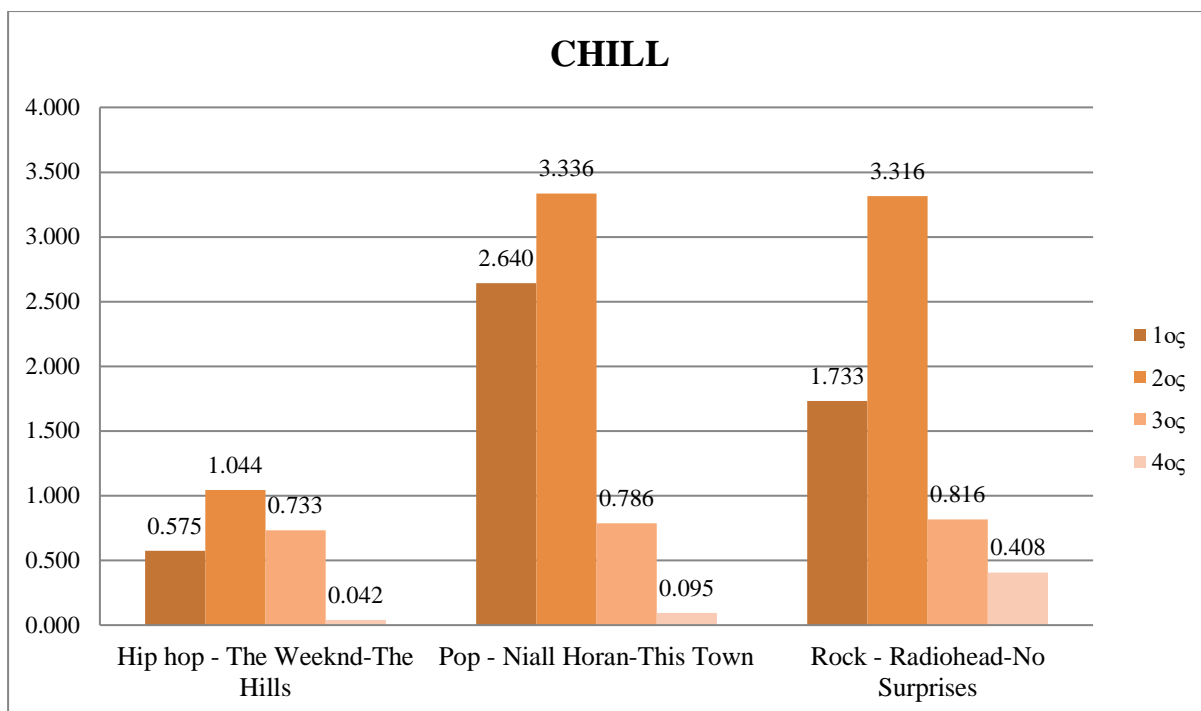
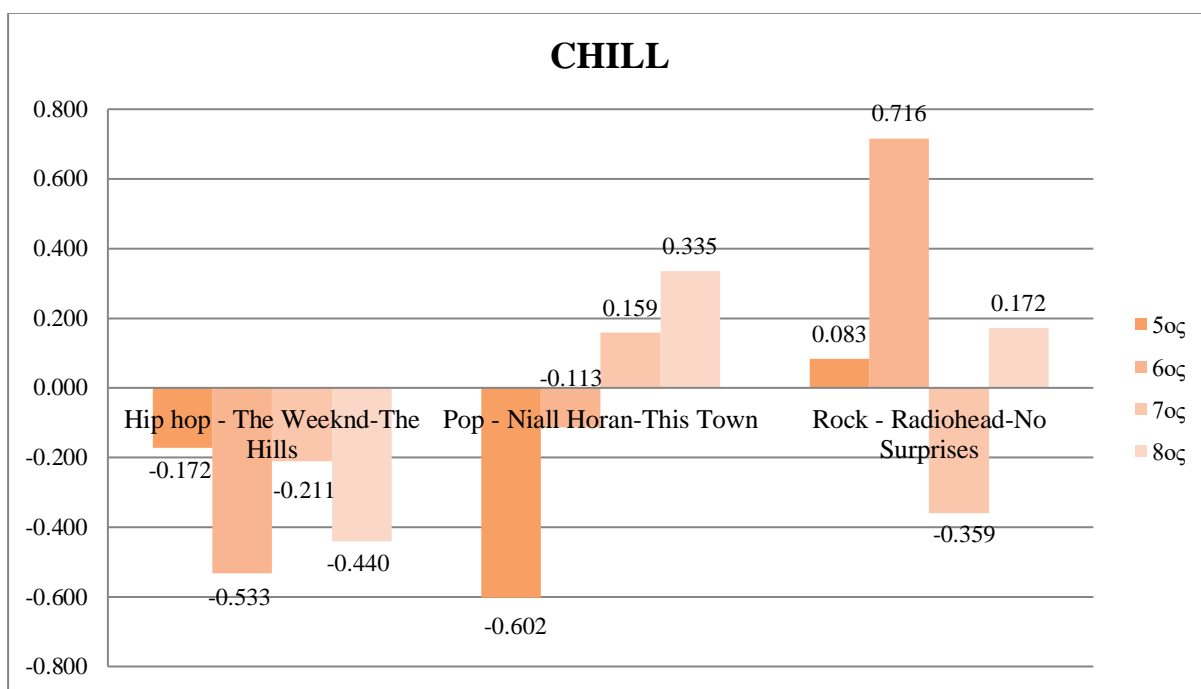


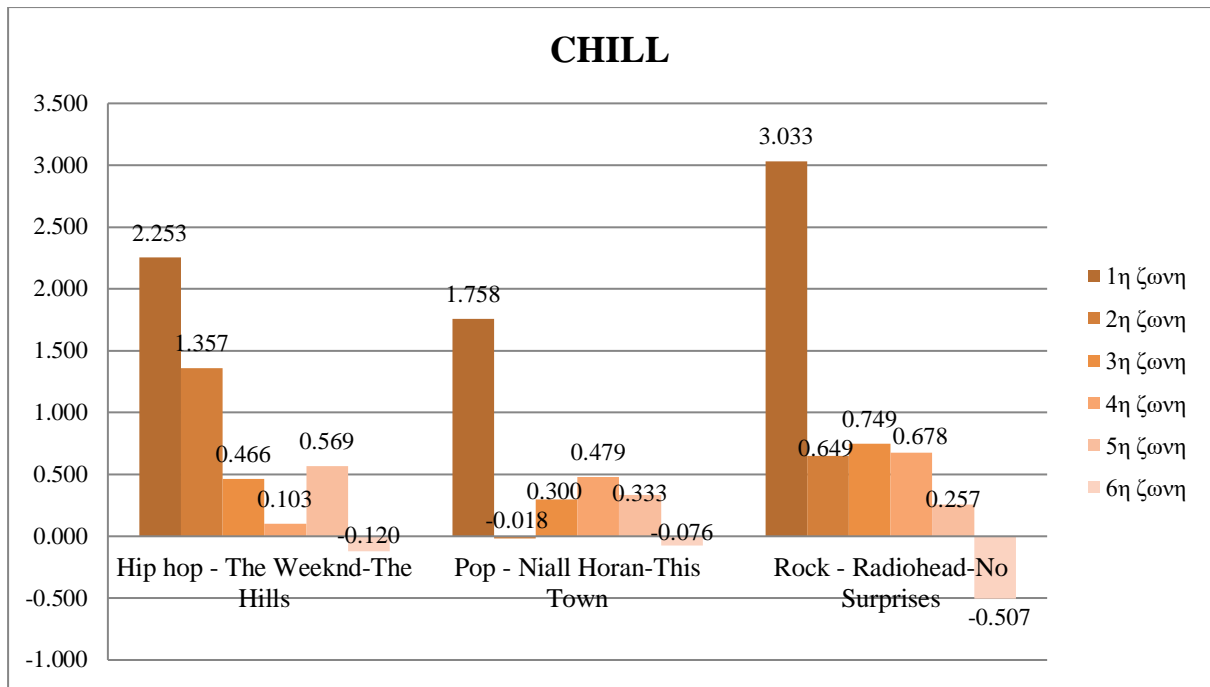
Εικόνα 4.27: Λοξότητα των χαρακτηριστικών για την κατηγορία chill



Εικόνα 4.28: Λοξότητα του spectral flatness για την κατηγορία chill

Από τα διαγράμματα παρατηρούμε ότι οι τιμές των τριών κομματιών για την τυπική απόκλιση στα πρώτα επτά χαρακτηριστικά, παρατηρούνται αποκλίσεις μεταξύ των τιμών κάθε τραγουδιού. Η υψηλότερη τιμή στα spectral centroid και chroma stft παρατηρείται στο Hip hop - The Weeknd-The Hills, ενώ στα spectral bandwidth και spectral rolloff, η υψηλότερη τιμή εμφανίζεται στο Pop - Niall Horan-This Town. Στα rms, zero crossing και spectral flatness η υψηλότερη τιμή παρουσιάζεται στο Rock - Radiohead-No Surprises.

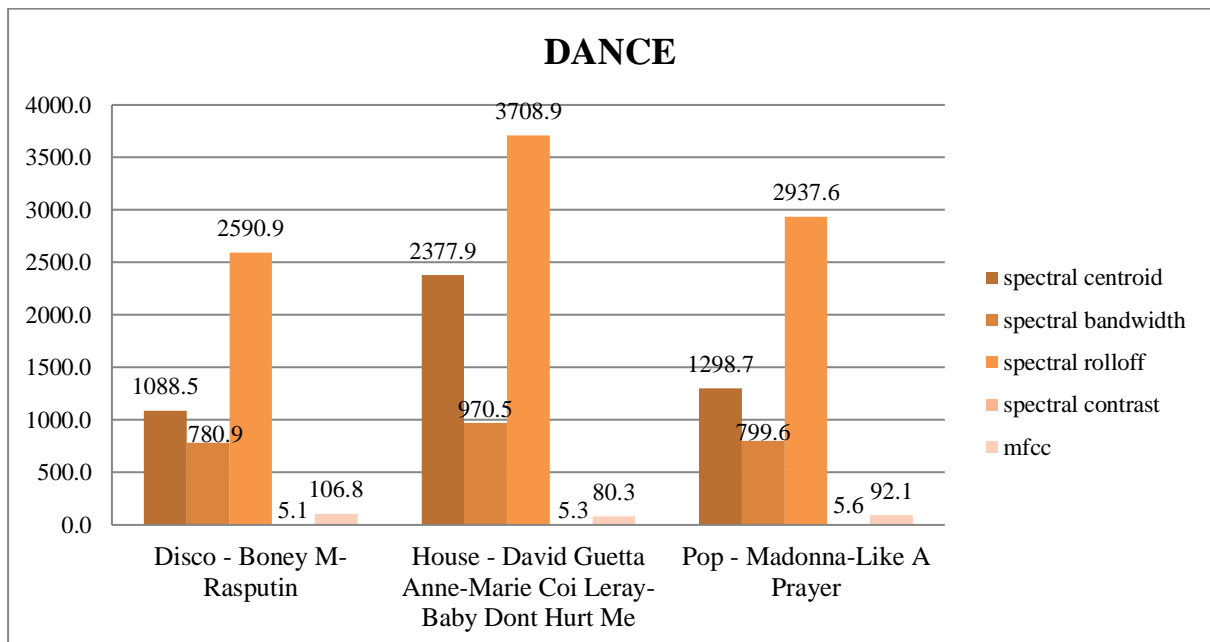
Εικόνα 4.29: Λοξότητα των MFCC (1^{ος} έως 4^{ος}) για την κατηγορία chillΕικόνα 4.30: Λοξότητα των MFCC (5^{ος} έως 8^{ος}) για την κατηγορία chill



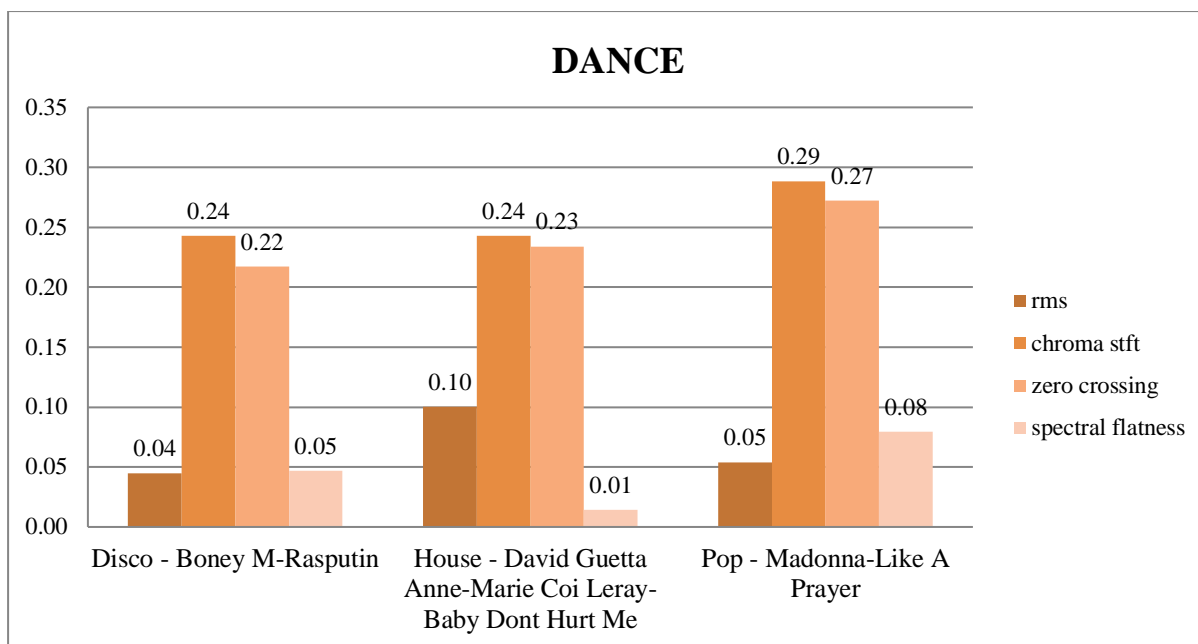
Εικόνα 4.31: Λοξότητα του spectral contrast στις 6 ζώνες για την κατηγορία chill

Όσον αφορά τους mfcc παρατηρούνται και αρνητικές και θετικές τιμές. Στον 1^ο, 2^ο και 7^ο mfcc η υψηλότερη τιμή εμφανίζεται στο Pop - Niall Horan-This Town, ενώ στον 3^ο, 4^ο, 5^ο, 6^ο και 8^ο εμφανίζεται στο Rock - Radiohead-No Surprises. Στην περίπτωση του spectral contrast, στην 1^η, 4^η και 5^η ζώνη η υψηλότερη τιμή παρουσιάζεται στο Rock - Radiohead-No Surprises ενώ στη 2^η, 3^η και 6^η ζώνη εμφανίζεται στο Hip hop - The Weeknd-The Hills.

Στις παρακάτω εικόνες παρουσιάζονται τα γραφήματα για το συναίσθημα dance στα πέντε στατιστικά στοιχεία.

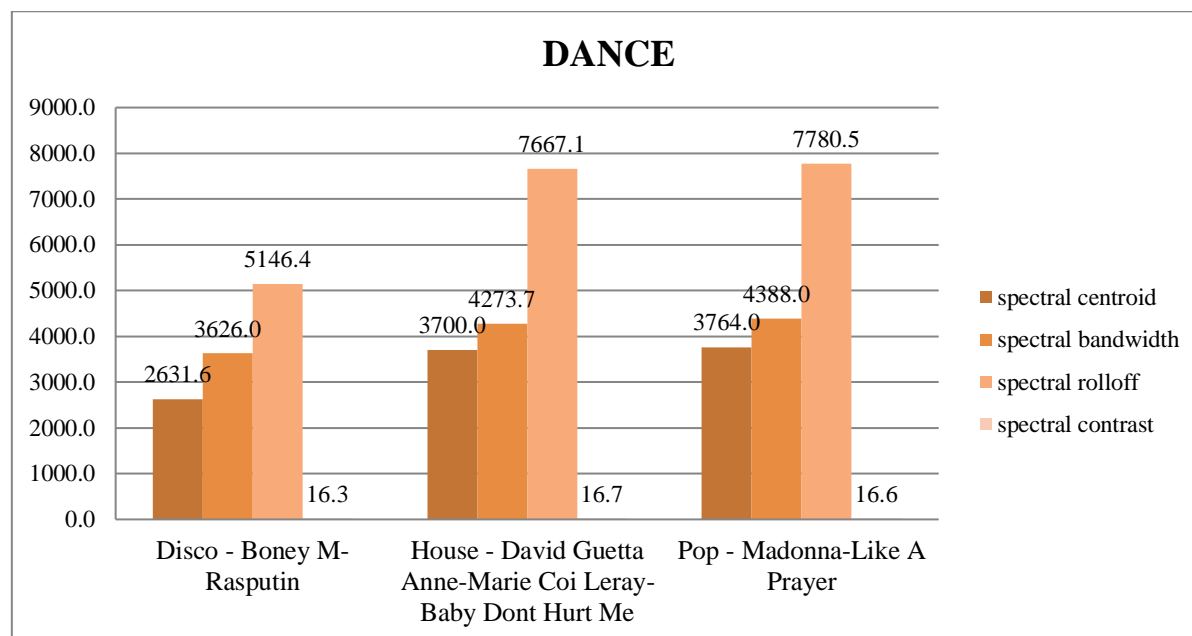


Εικόνα 4.32: Τυπική απόκλιση των χαρακτηριστικών για την κατηγορία dance

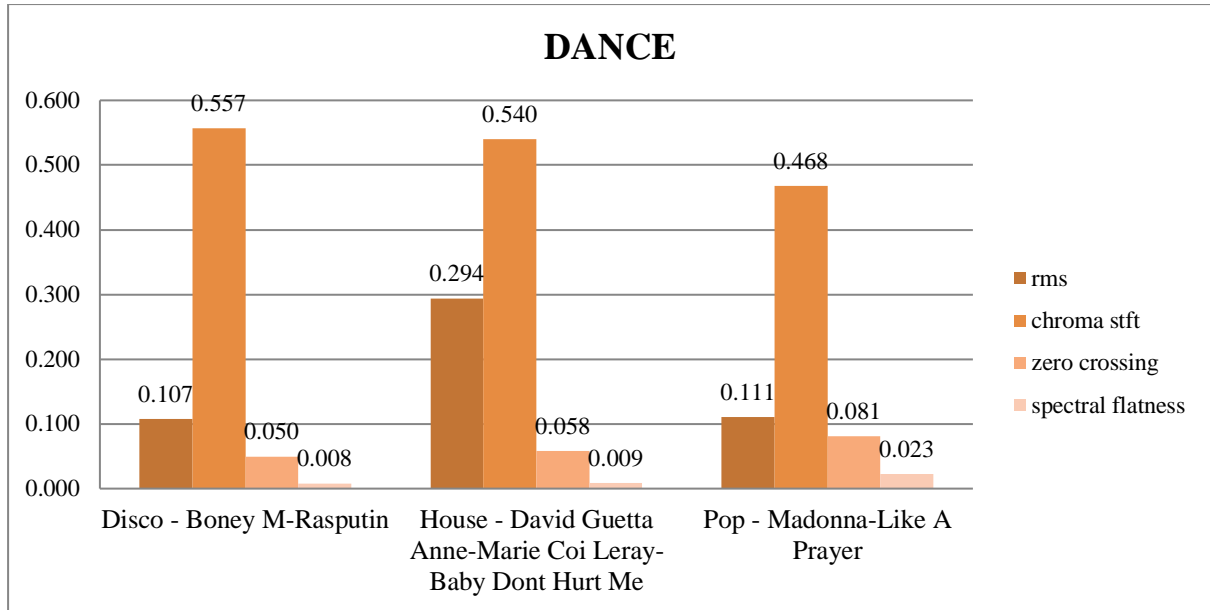


Εικόνα 4.33: Τυπική απόκλιση των χαρακτηριστικών για την κατηγορία dance

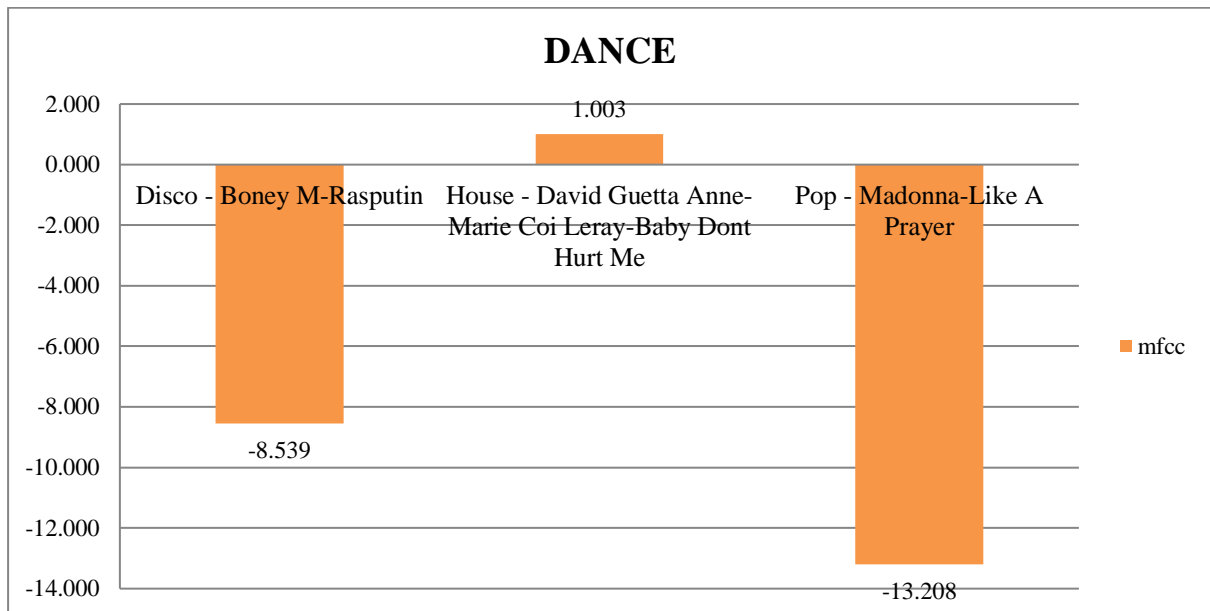
Από τα διαγράμματα παρατηρούμε ότι οι τιμές των τριών κομματιών για την τυπική απόκλιση στα πρώτα πέντε χαρακτηριστικά είναι αρκετά κοντά. Η υψηλότερη τιμή στο spectral centroid παρατηρείται στο House - David Guetta Anne-Marie Coi Leray-Baby Dont Hurt Me καθώς και στα rms, spectral bandwidth και spectral rolloff. Στους mfcc, η υψηλότερη τιμή παρατηρείται στο Disco - Boney M-Rasputin. Στα spectral contrast, chroma stft, zero crossing και spectral flatness η υψηλότερη τιμή εμφανίζεται στο Pop - Madonna-Like A Prayer.



Εικόνα 4.34: Μέση τιμή των χαρακτηριστικών για την κατηγορία dance

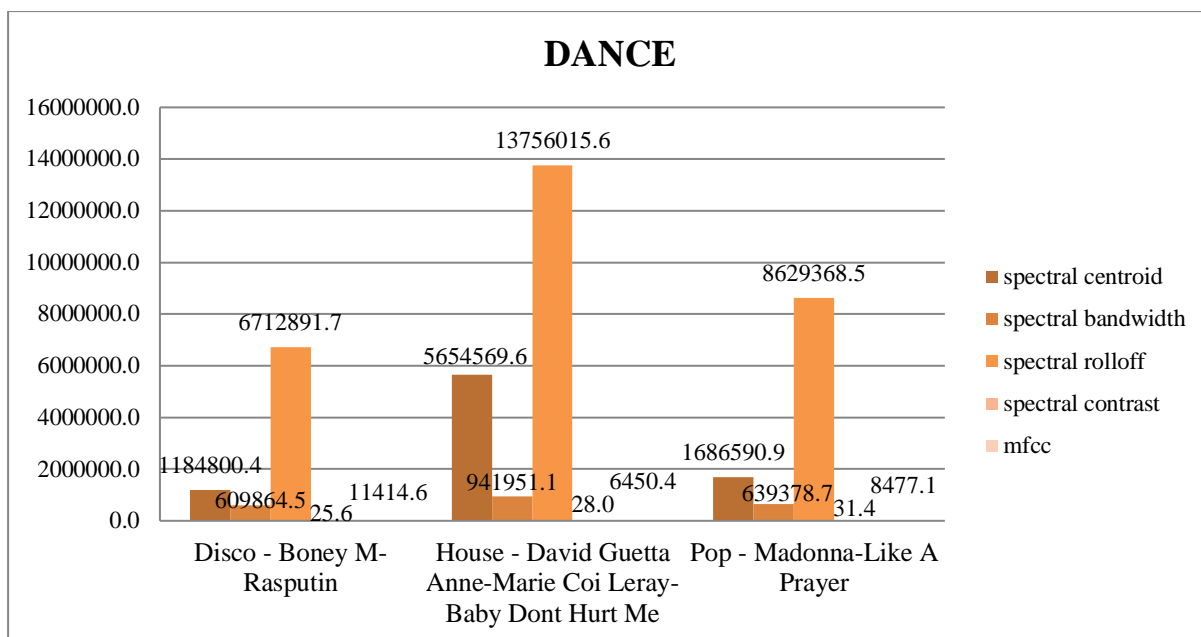


Εικόνα 4.35: Μέση τιμή των χαρακτηριστικών για την κατηγορία dance

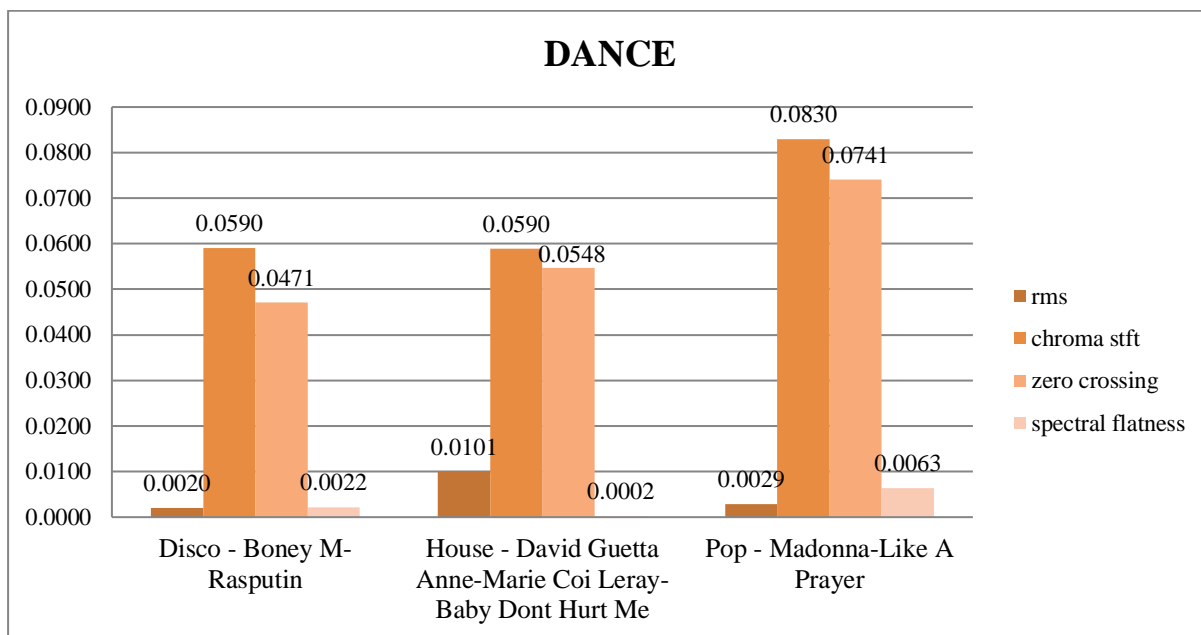


Εικόνα 4.36: Μέση τιμή των MFCC για την κατηγορία dance

Από τα διαγράμματα παρατηρούμε ότι οι τιμές των τριών κομματιών για την μέση τιμή σε όλα χαρακτηριστικά είναι πολύ κοντά. Η υψηλότερη τιμή στο spectral centroid παρατηρείται στο Pop - Madonna-Like A Prayer καθώς και στα spectral bandwidth, spectral rolloff, zero crossing, spectral flatness και mfcc. Στα spectral contrast και rms η υψηλότερη τιμή παρατηρείται στο House - David Guetta Anne-Marie Coi Leray-Baby Dont Hurt Me. Στο chroma stft η υψηλότερη τιμή εμφανίζεται στο Disco - Boney M-Rasputin.

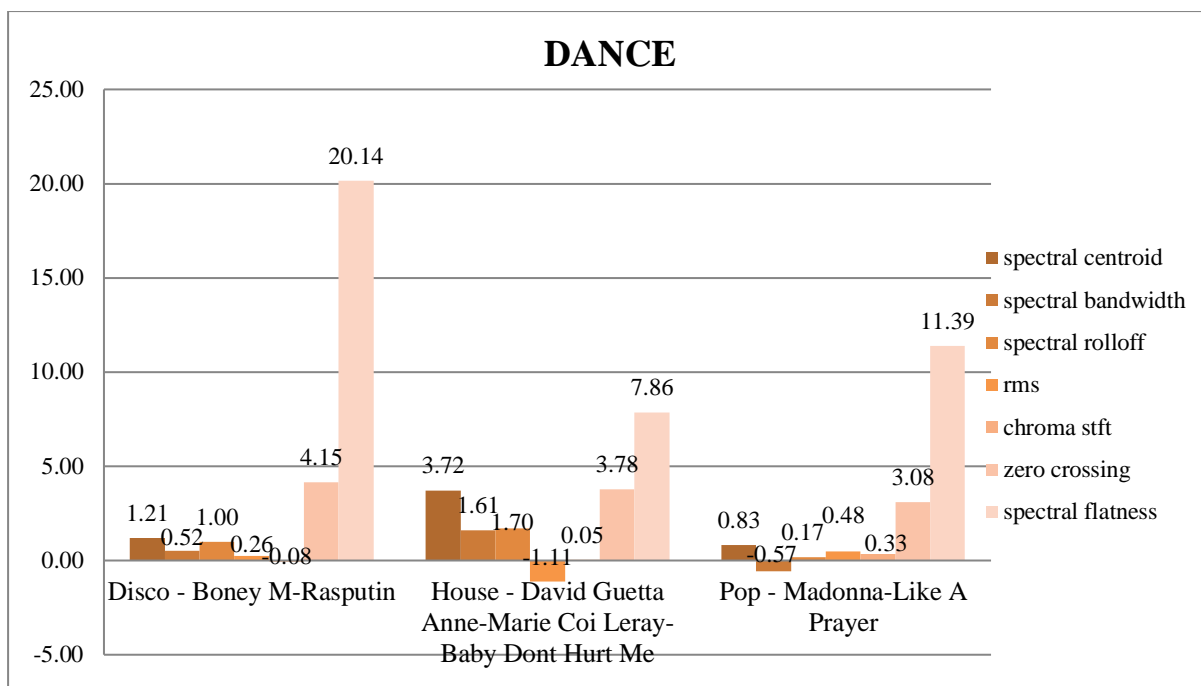


Εικόνα 4.37: Διακύμανση των χαρακτηριστικών για την κατηγορία dance



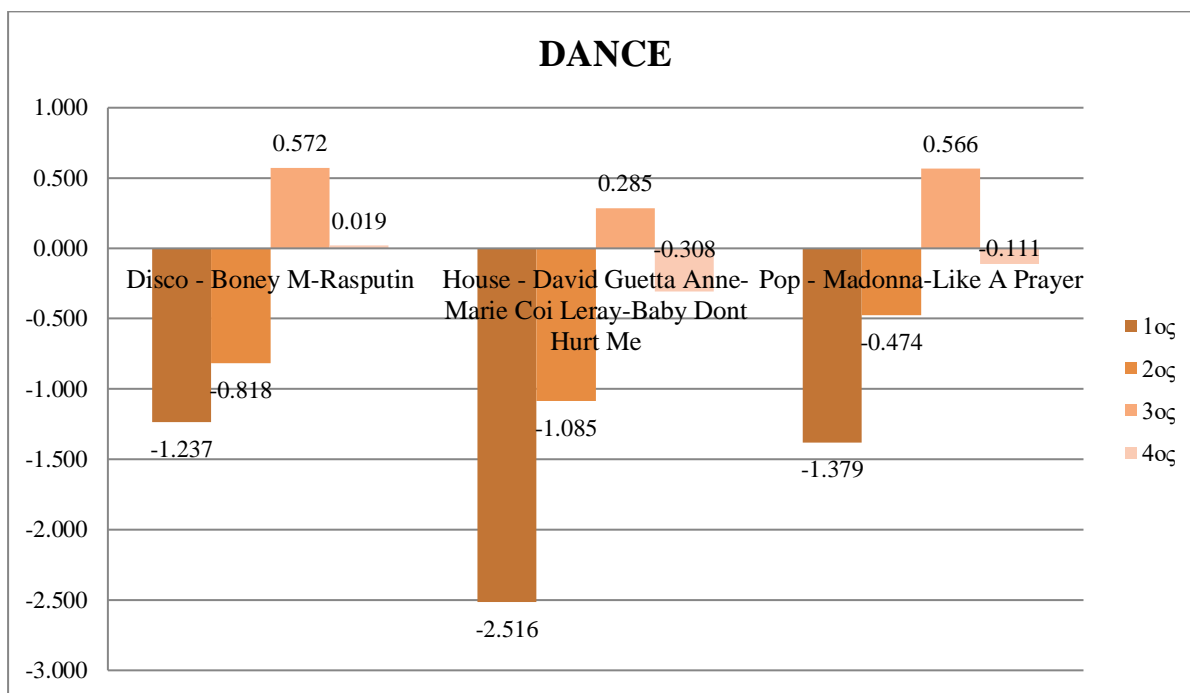
Εικόνα 4.38: Διακύμανση των χαρακτηριστικών για την κατηγορία dance

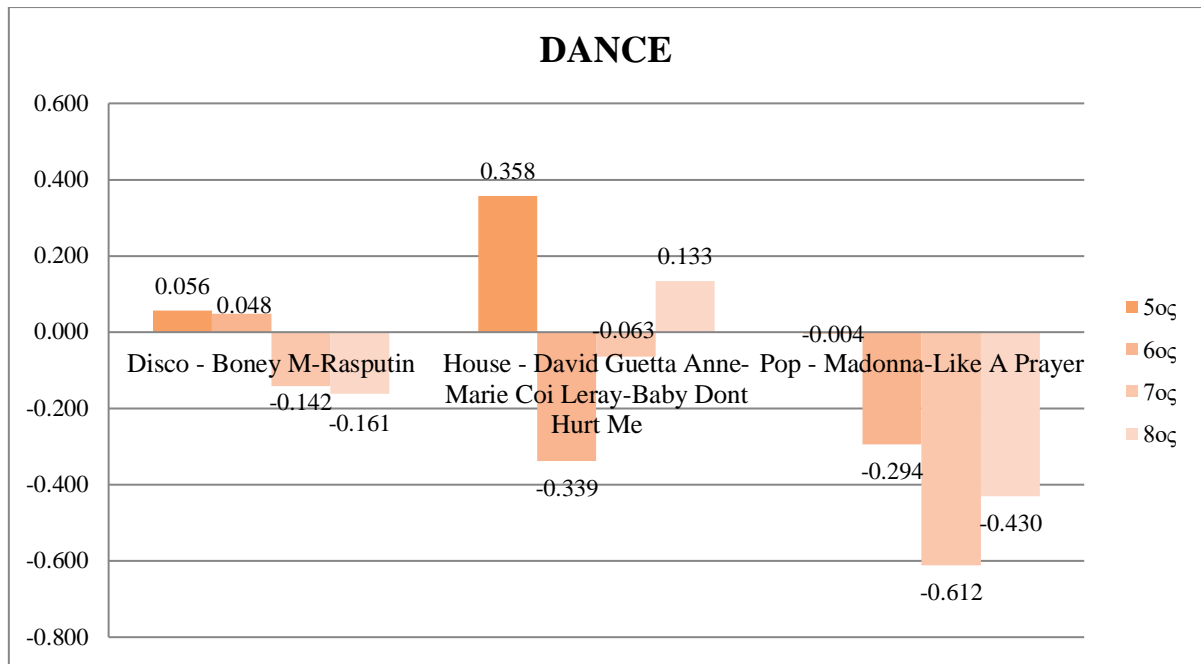
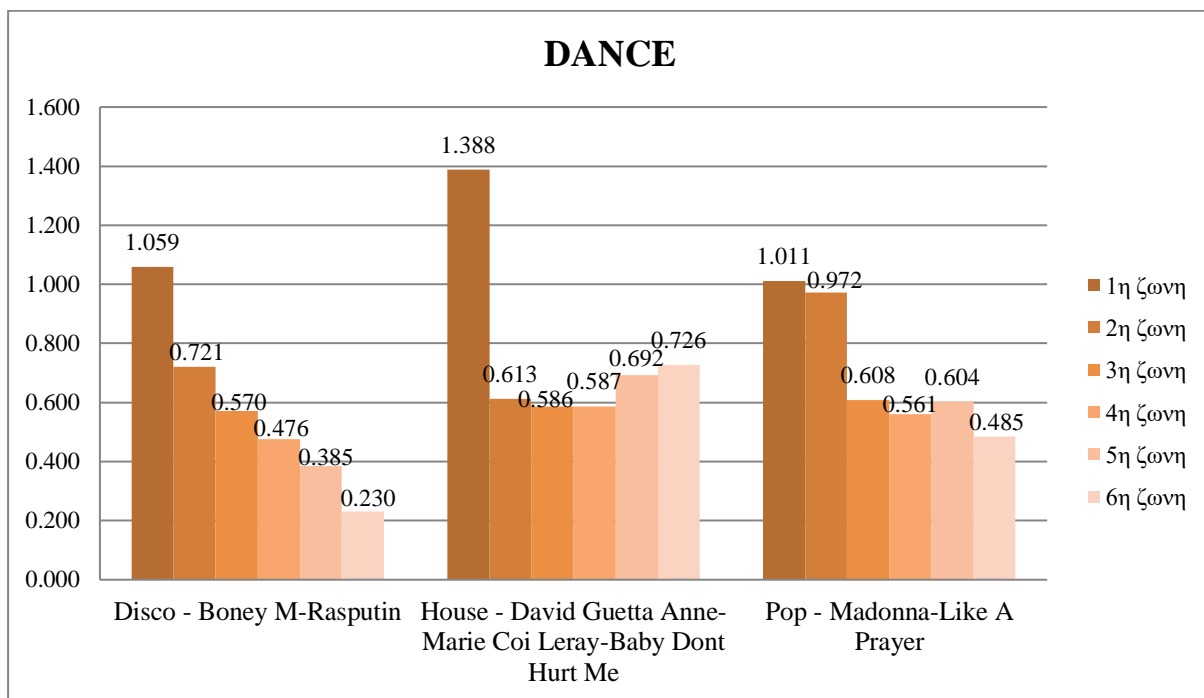
Από τα διαγράμματα παρατηρούμε ότι οι τιμές των τριών κομματιών για την διακύμανση σε όλα τα χαρακτηριστικά παρουσιάζουν μία μικρή απόκλιση. Η υψηλότερη τιμή στο spectral centroid παρατηρείται στο House - David Guetta Anne-Marie Coi Leray-Baby Dont Hurt Me καθώς και στα spectral rolloff, spectral bandwidth και rms. Στα spectral contrast, chroma stft, zero crossing και spectral flatness η υψηλότερη τιμή παρατηρείται στο Pop - Madonna-Like A Prayer, ενώ στους mfcc η υψηλότερη τιμή εμφανίζεται στο Disco - Boney M-Rasputin.



Εικόνα 4.39: Κύρτωση των χαρακτηριστικών για την κατηγορία dance

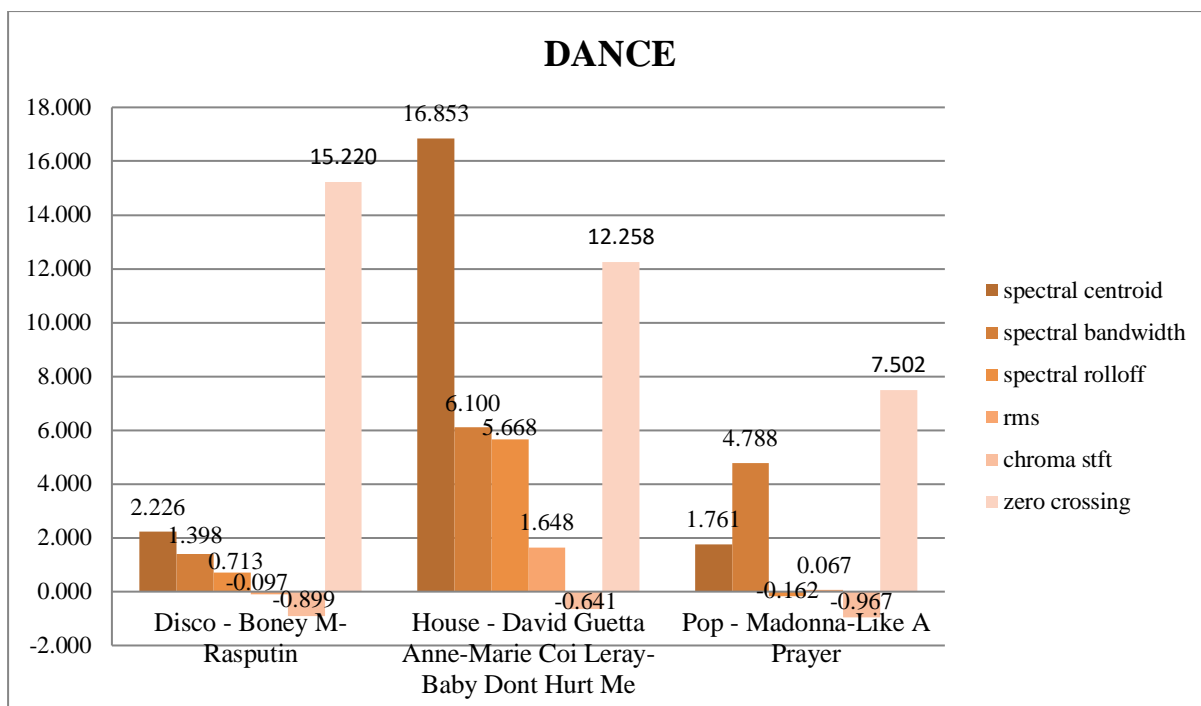
Από τα διαγράμματα παρατηρούμε ότι οι τιμές των τριών κομματιών για την κύρτωση στα πρώτα επτά χαρακτηριστικά είναι αρκετά κοντά. Η υψηλότερη τιμή στο spectral centroid παρατηρείται στο House - David Guetta Anne-Marie Coi Leray-Baby Dont Hurt Me καθώς και στα spectral bandwidth και spectral rolloff. Στην rms η υψηλότερη τιμή παρατηρείται στο Pop - Madonna-Like A Prayer καθώς και στο chroma stft. Στο zero crossing η υψηλότερη τιμή παρατηρείται στο Disco - Boney M-Rasputin καθώς και στο spectral flatness.

Εικόνα 4.40: Κύρτωση των MFCC (1^{ος} έως 4^{ος}) για την κατηγορία dance

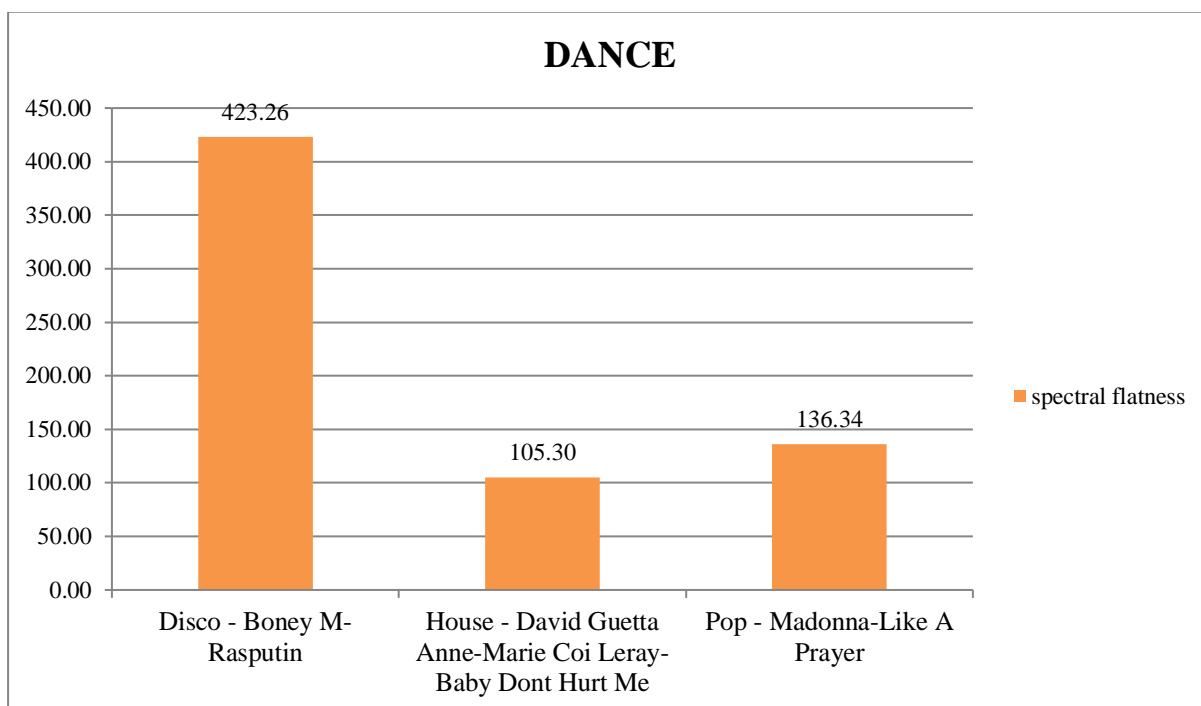
Εικόνα 4.41: Κύρτωση των MFCC (5^{ος} έως 8^{ος}) για την κατηγορία dance

Εικόνα 4.42: Κύρτωση του spectral contrast στις 6 ζώνες για την κατηγορία dance

Όσον αφορά τους mfcc παρατηρούνται περισσότερες αρνητικές τιμές. Στον 1^ο, 2^ο και 5^ο mfcc η υψηλότερη τιμή εμφανίζεται στο House - David Guetta Anne-Marie Coi Leray-Baby Dont Hurt Me, ενώ στον 3^ο, 4^ο, 6^ο και 8^ο εμφανίζεται στο Disco - Boney M-Rasputin. Στον 7^ο mfcc η υψηλότερη τιμή εμφανίζεται στο Pop - Madonna-Like A Prayer. Στην περίπτωση του spectral contrast, στην 1^η, 4^η, 5^η και 6^η ζώνη η υψηλότερη τιμή παρουσιάζεται στο House - David Guetta Anne-Marie Coi Leray-Baby Dont Hurt Me ενώ στη 2^η 3^η ζώνη εμφανίζεται στο Pop - Madonna-Like A Prayer.

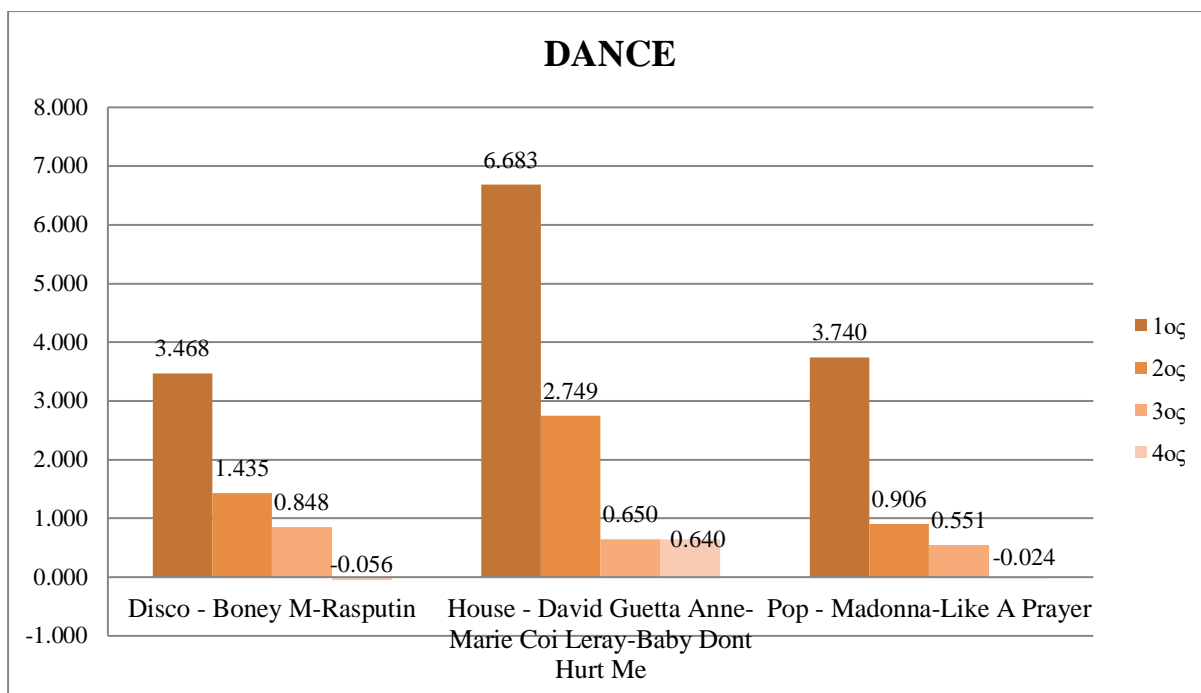
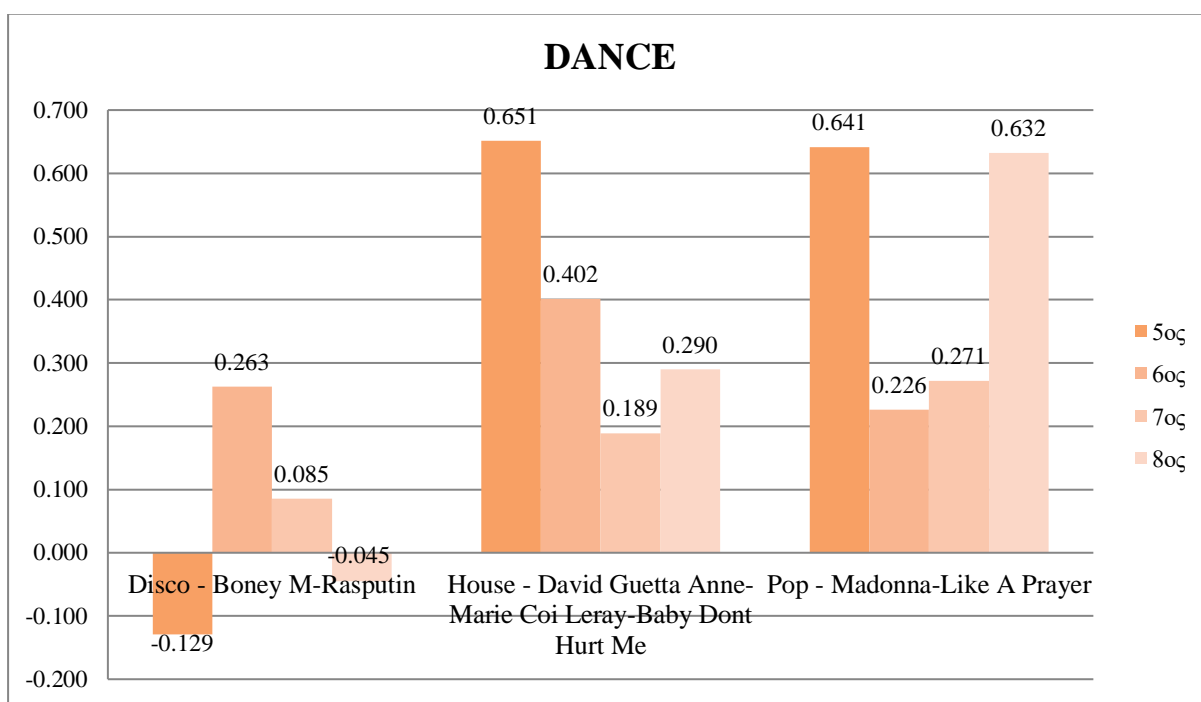


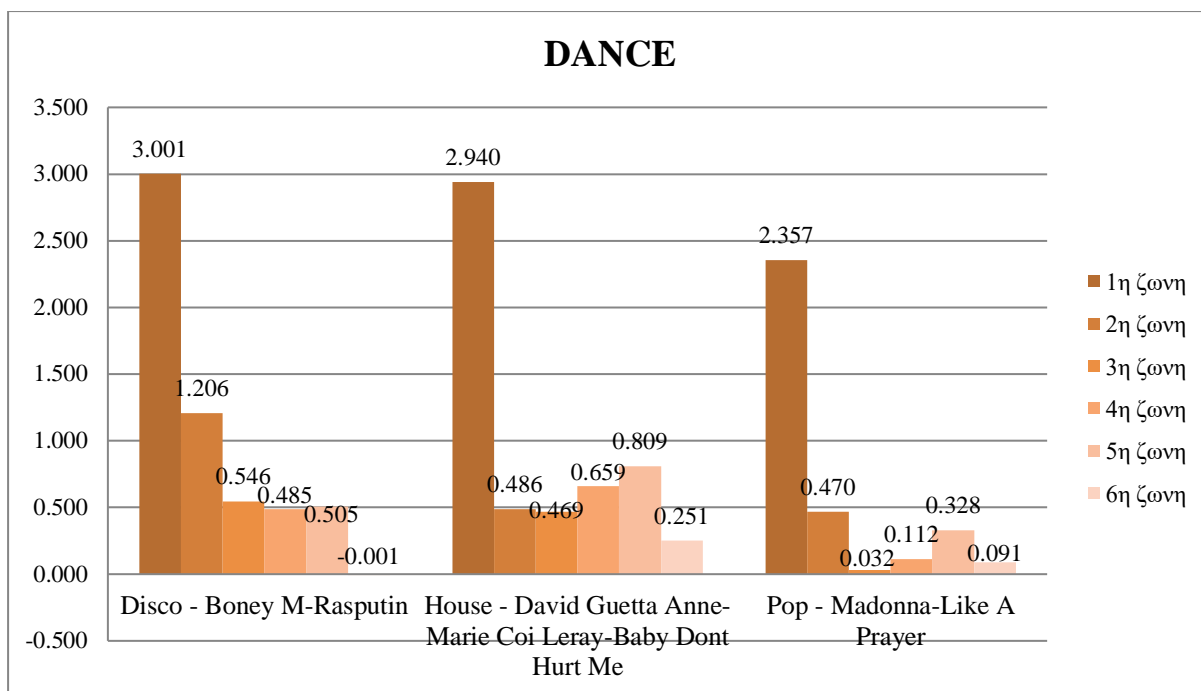
Εικόνα 4.43: Λοξότητα των χαρακτηριστικών για την κατηγορία dance



Εικόνα 4.44: Λοξότητα του spectral flatness για την κατηγορία dance

Από τα διαγράμματα παρατηρούμε ότι οι τιμές των τριών κομματιών για την λοξότητα στα πρώτα επτά χαρακτηριστικά είναι αρκετά κοντά. Η υψηλότερη τιμή στο spectral centroid παρατηρείται στο House - David Guetta Anne-Marie Coi Leray-Baby Dont Hurt Me καθώς και στα spectral bandwidth, spectral rolloff και rms. Στο chroma stft η υψηλότερη τιμή παρατηρείται στο Pop - Madonna-Like A Prayer. Στο zero crossing η υψηλότερη τιμή παρατηρείται στο Disco - Boney M-Rasputin καθώς και στο spectral flatness.

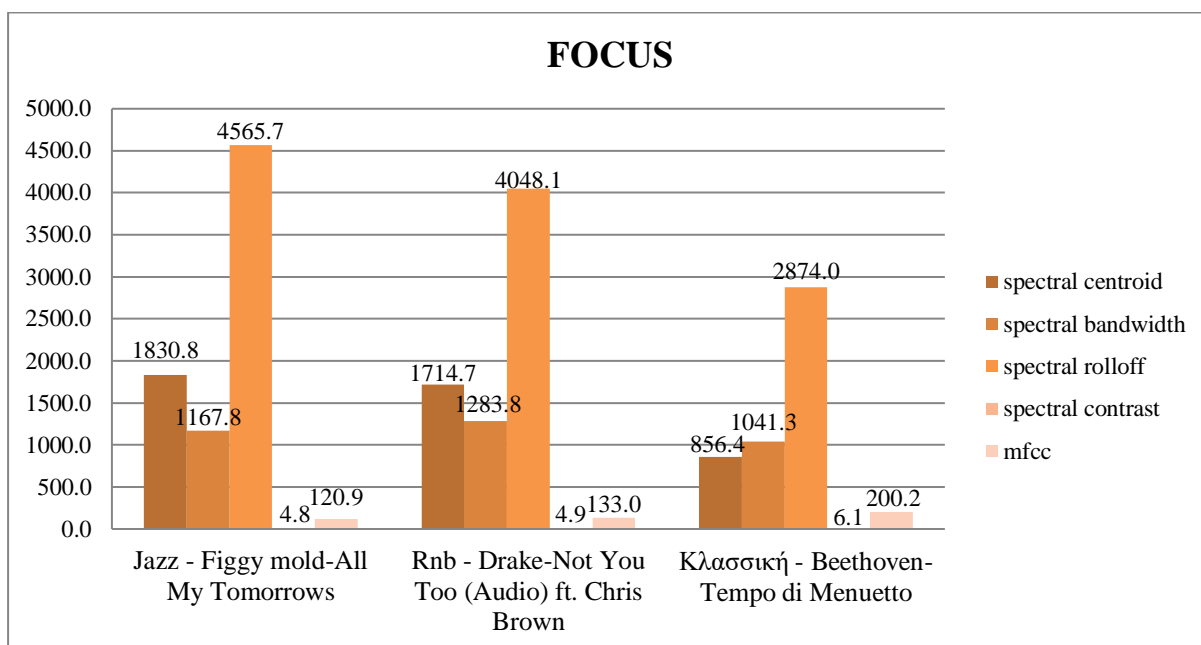
Εικόνα 4.45: Λοξότητα των MFCC (1^{ος} έως 4^{ος}) για την κατηγορία danceΕικόνα 4.46: Λοξότητα των MFCC (5^{ος} έως 8^{ος}) για την κατηγορία dance



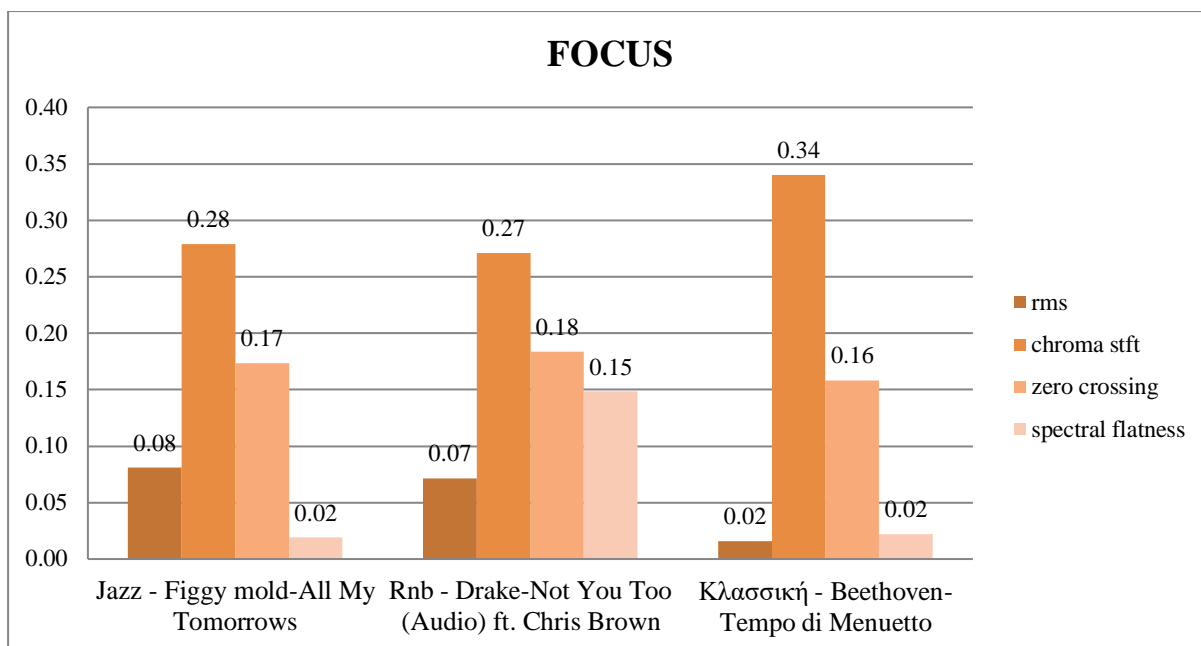
Εικόνα 4.47: Λοξότητα του spectral contrast στις 6 ζώνες για την κατηγορία dance

Όσον αφορά τους mfcc οι τιμές είναι αρκετά κοντά. Στον 1°, 2°, 4°, 5° και 6° mfcc η υψηλότερη τιμή εμφανίζεται στο House - David Guetta Anne-Marie Coi Leray-Baby Dont Hurt Me, ενώ στον 3° εμφανίζεται στο Disco - Boney M-Rasputin. Στον 7° και 8° mfcc η υψηλότερη τιμή εμφανίζεται στο Pop - Madonna-Like A Prayer. Στην περίπτωση του spectral contrast, στην 1η, 2η και 3η ζώνη η υψηλότερη τιμή παρουσιάζεται στο Disco - Boney M-Rasputin ενώ στη 4η, 5η και 6η ζώνη εμφανίζεται στο House - David Guetta Anne-Marie Coi Leray-Baby Dont Hurt Me.

Στις παρακάτω εικόνες παρουσιάζονται τα γραφήματα για το συναίσθημα focus στα πέντε στατιστικά στοιχεία.

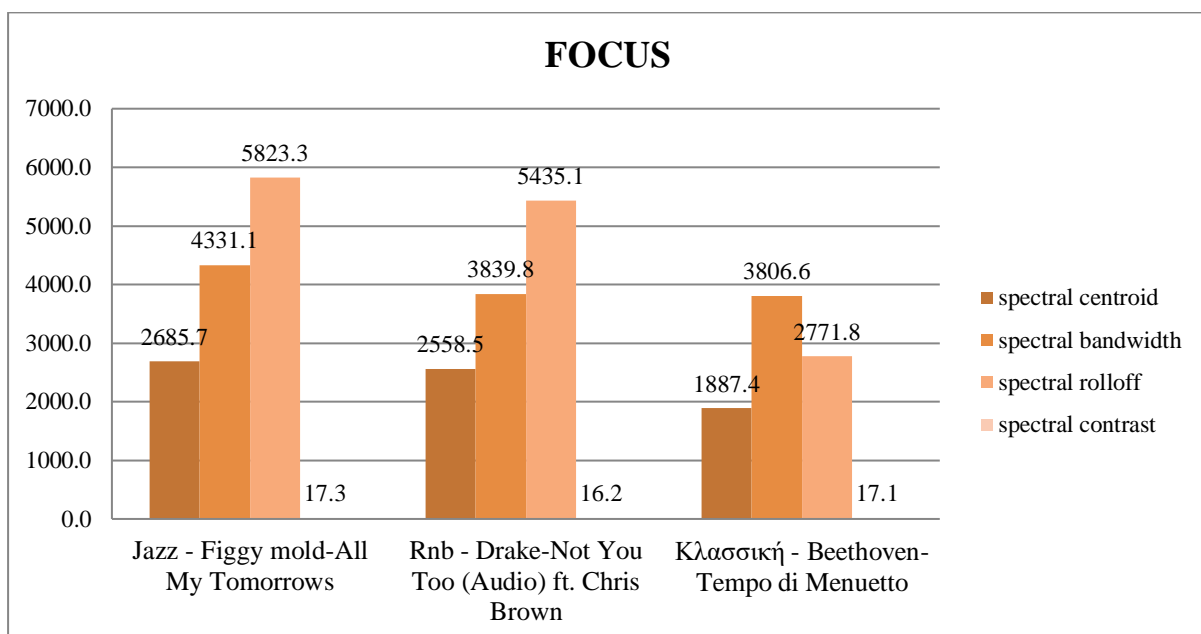


Εικόνα 4.48: Τυπική απόκλιση των χαρακτηριστικών για την κατηγορία focus

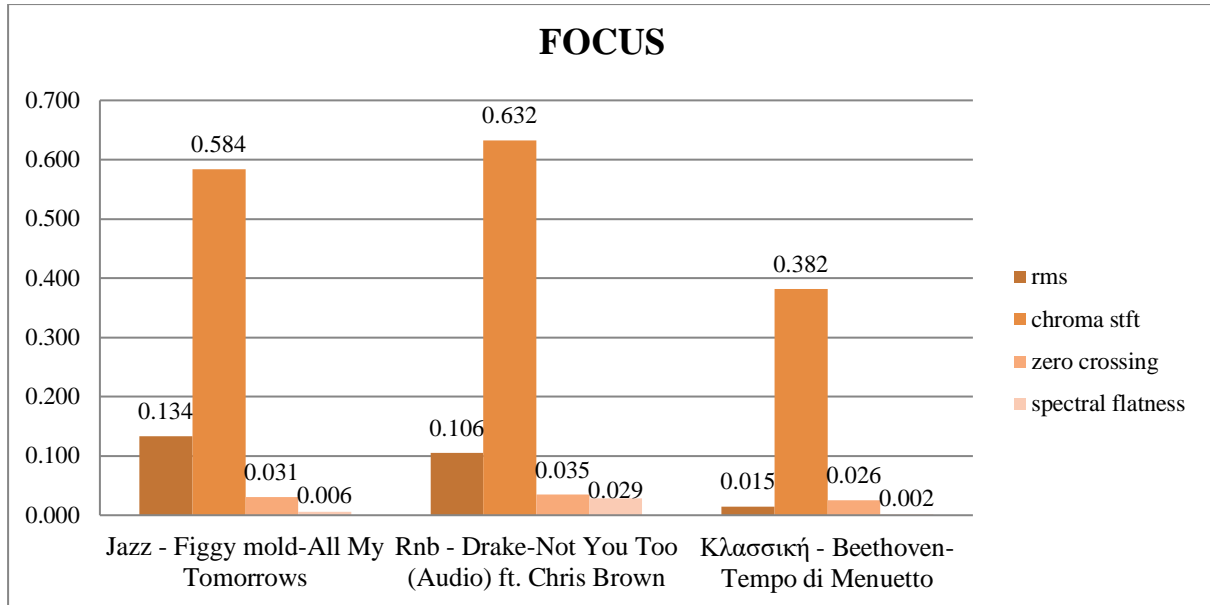


Εικόνα 4.49: Τυπική απόκλιση των χαρακτηριστικών για την κατηγορία focus

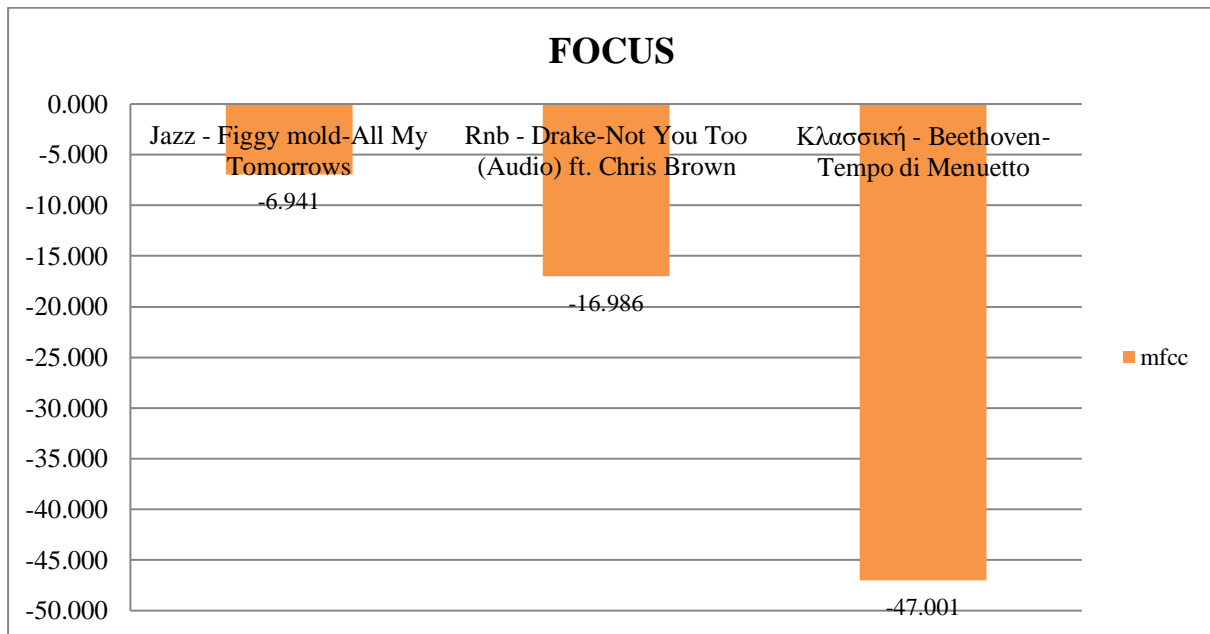
Από τα διαγράμματα παρατηρούμε ότι οι τιμές των τριών κομματιών για την τυπική απόκλιση σε όλα τα χαρακτηριστικά είναι πολύ κοντά. Η υψηλότερη τιμή στο spectral centroid παρατηρείται στο Jazz - Figgy mold-All My Tomorrows καθώς και στα spectral rolloff και rms. Στα spectral bandwidth, zero crossing και spectral flatness η υψηλότερη τιμή εμφανίζεται στο Rnb - Drake-Not You Too (Audio) ft. Chris Brown. Στα spectral contrast, mfcc και chroma stft η υψηλότερη τιμή παρατηρείται στο Κλασσική - Beethoven-Tempo di Menuetto.



Εικόνα 4.50: Μέση τιμή των χαρακτηριστικών για την κατηγορία focus

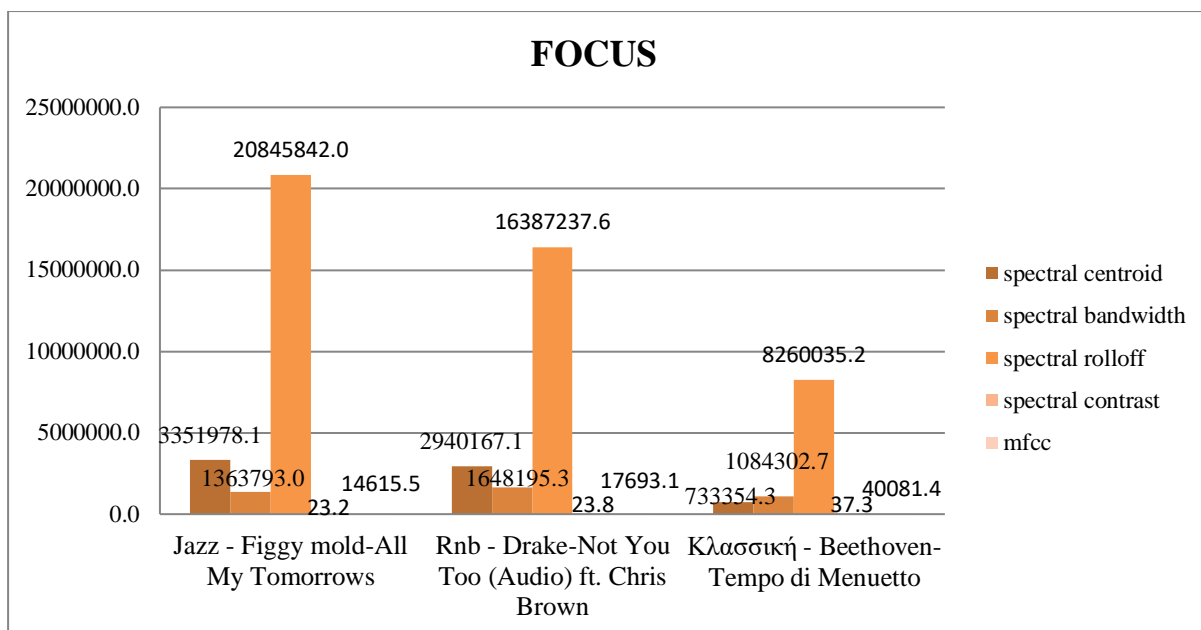


Εικόνα 4.51: Μέση τιμή των χαρακτηριστικών για την κατηγορία focus

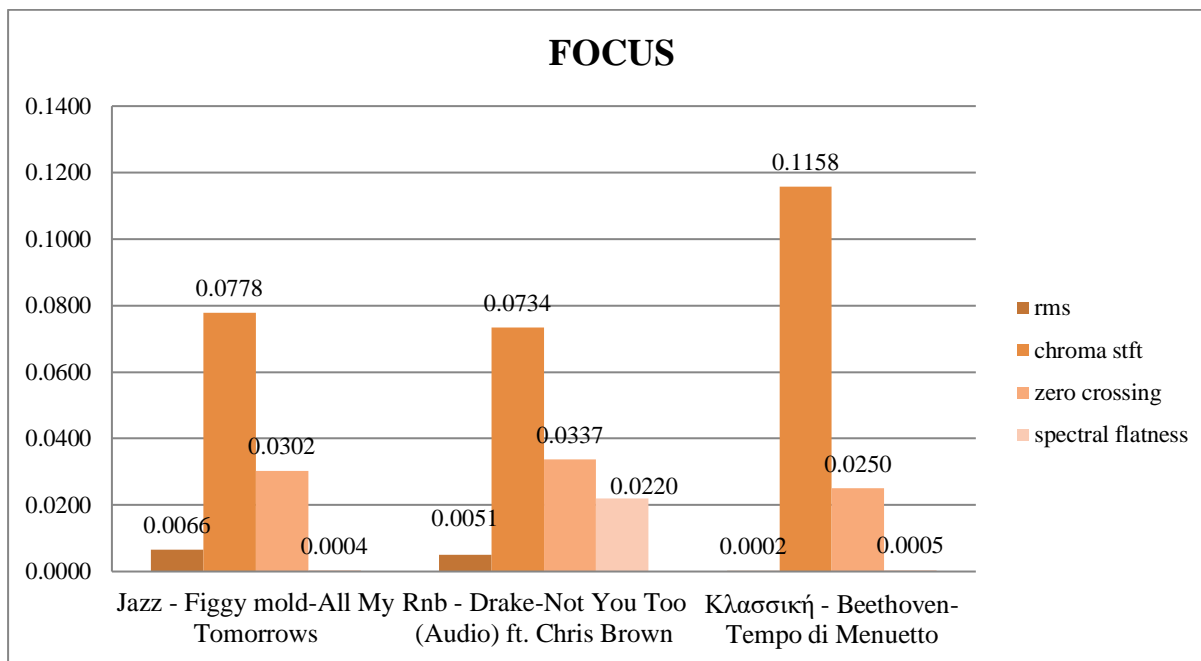


Εικόνα 4.52: Μέση τιμή των MFCC για την κατηγορία focus

Από τα διαγράμματα παρατηρούμε ότι οι τιμές των τριών κομματιών για την μέση τιμή σε όλα τα χαρακτηριστικά είναι πολύ κοντά. Η υψηλότερη τιμή στο spectral centroid παρατηρείται στο Jazz - Figgy mold-All My Tomorrows καθώς και στα spectral rolloff, spectral bandwidth, spectral contrast και rms. Στα chroma stft, zero crossing και spectral flatness η υψηλότερη τιμή παρατηρείται στο Rnb - Drake-Not You Too (Audio) ft. Chris Brown. Στους mfcc η υψηλότερη τιμή παρατηρείται στο Κλασσική - Beethoven-Tempo di Menuetto.

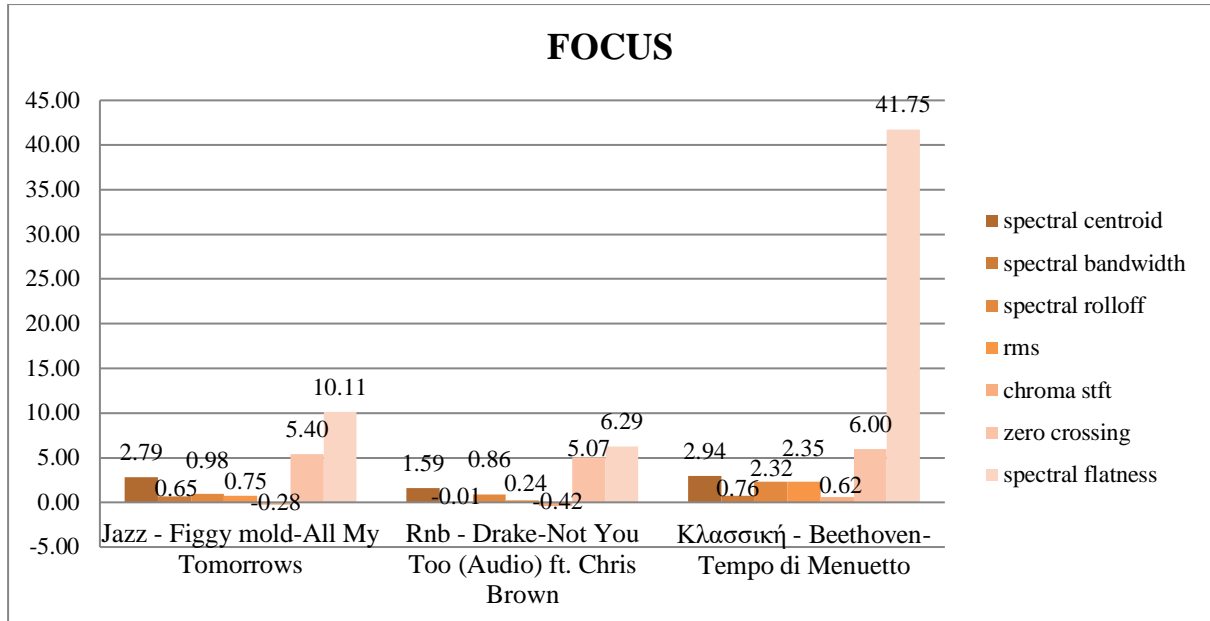


Εικόνα 4.53: Διακύμανση των χαρακτηριστικών για την κατηγορία focus



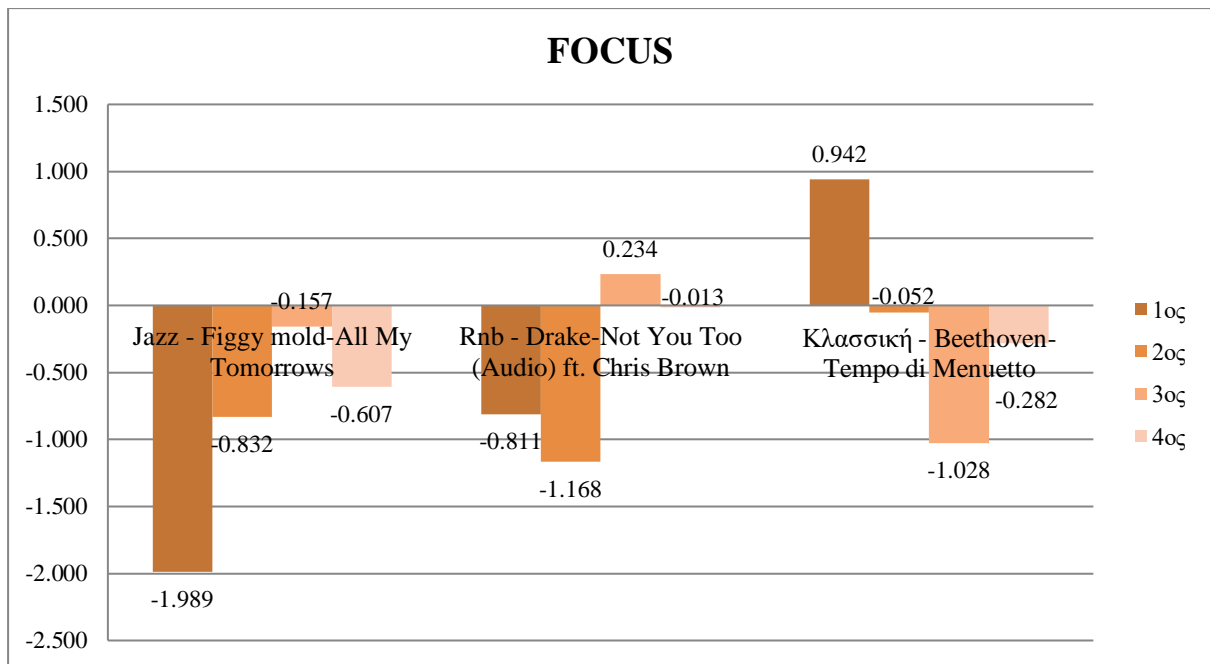
Εικόνα 4.54: Διακύμανση των χαρακτηριστικών για την κατηγορία focus

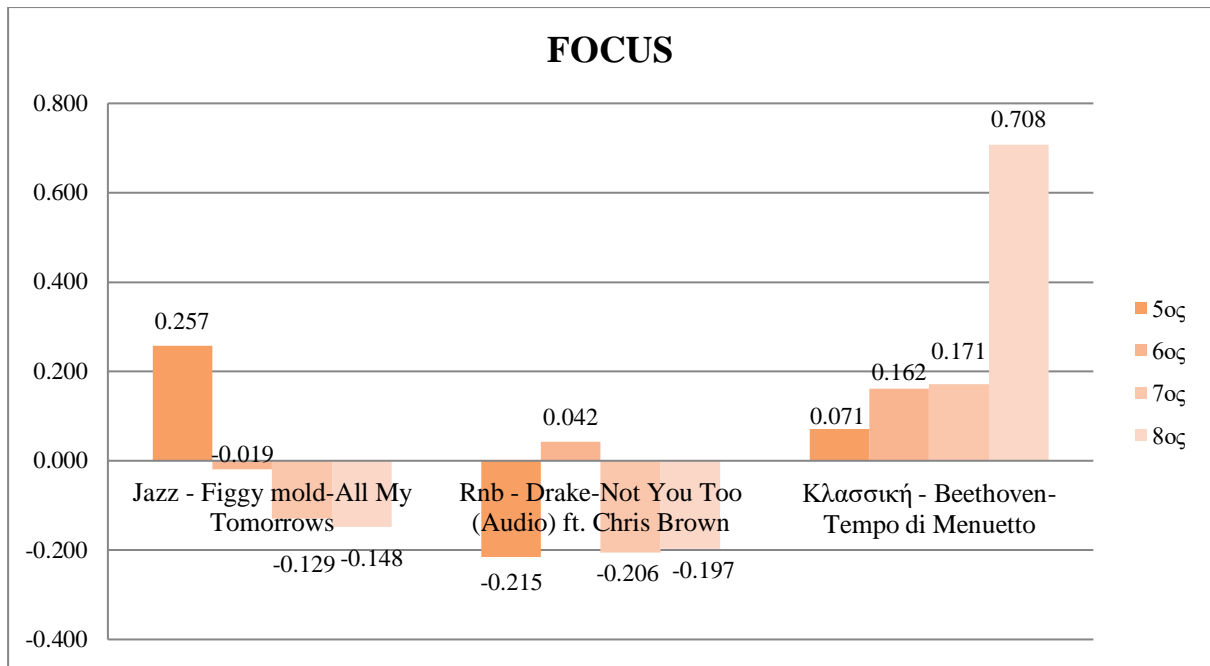
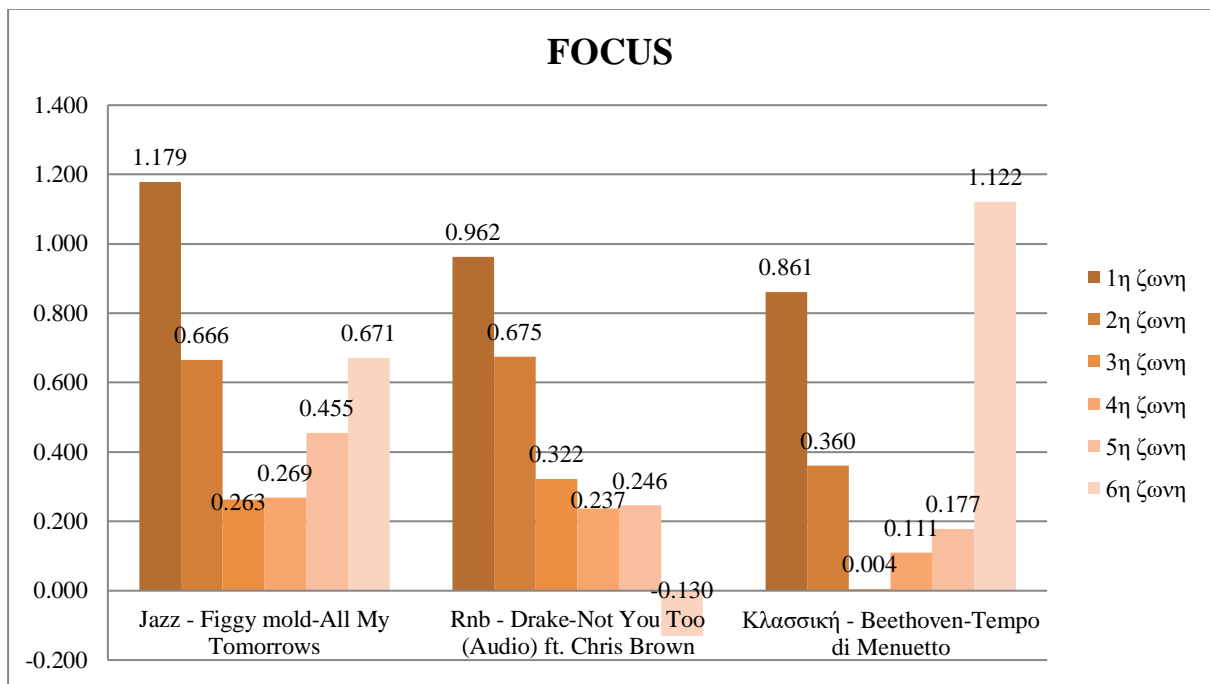
Από τα διαγράμματα παρατηρούμε ότι οι τιμές των τριών κομματιών για την διακύμανση σε όλα τα χαρακτηριστικά είναι πολύ κοντά. Η υψηλότερη τιμή στο spectral centroid παρατηρείται στο Jazz - Figgy mold-All My Tomorrows καθώς και στα spectral rolloff και rms. Στο spectral bandwidth η υψηλότερη τιμή παρατηρείται στο Rnb - Drake-Not You Too (Audio) ft. Chris Brown καθώς και στα zero crossing και spectral flatness. Στα spectral contrast και mfcc παρατηρείται στο Κλασσική - Beethoven-Tempo di Menuetto καθώς και στο chroma stft.



Εικόνα 4.55: Κύρτωση των χαρακτηριστικών για την κατηγορία focus

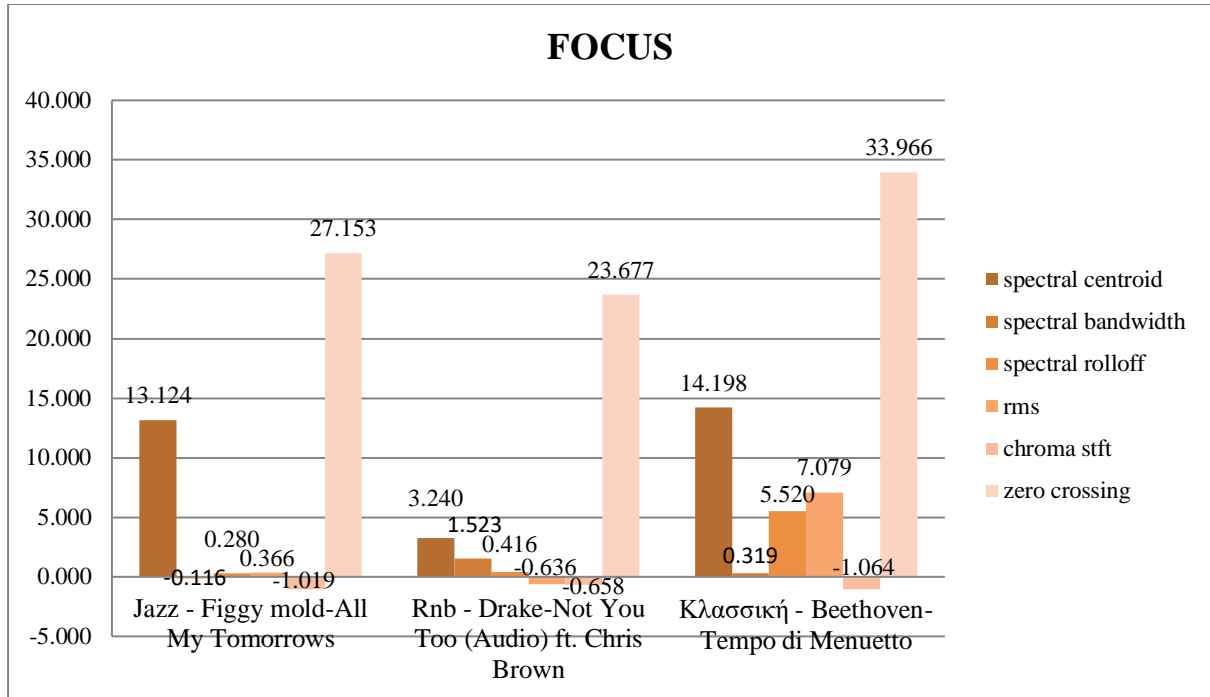
Από τα διαγράμματα παρατηρούμε ότι οι τιμές των τριών κομματιών για την κύρτωση στα πρώτα επτά χαρακτηριστικά είναι αρκετά κοντά. Η υψηλότερη τιμή στα πρώτα επτά χαρακτηριστικά παρατηρείται στο Κλασσική - Beethoven-Tempo di Menuetto.

Εικόνα 4.56: Κύρτωση των MFCC (1^{ος} έως 4^{ος}) για την κατηγορία focus

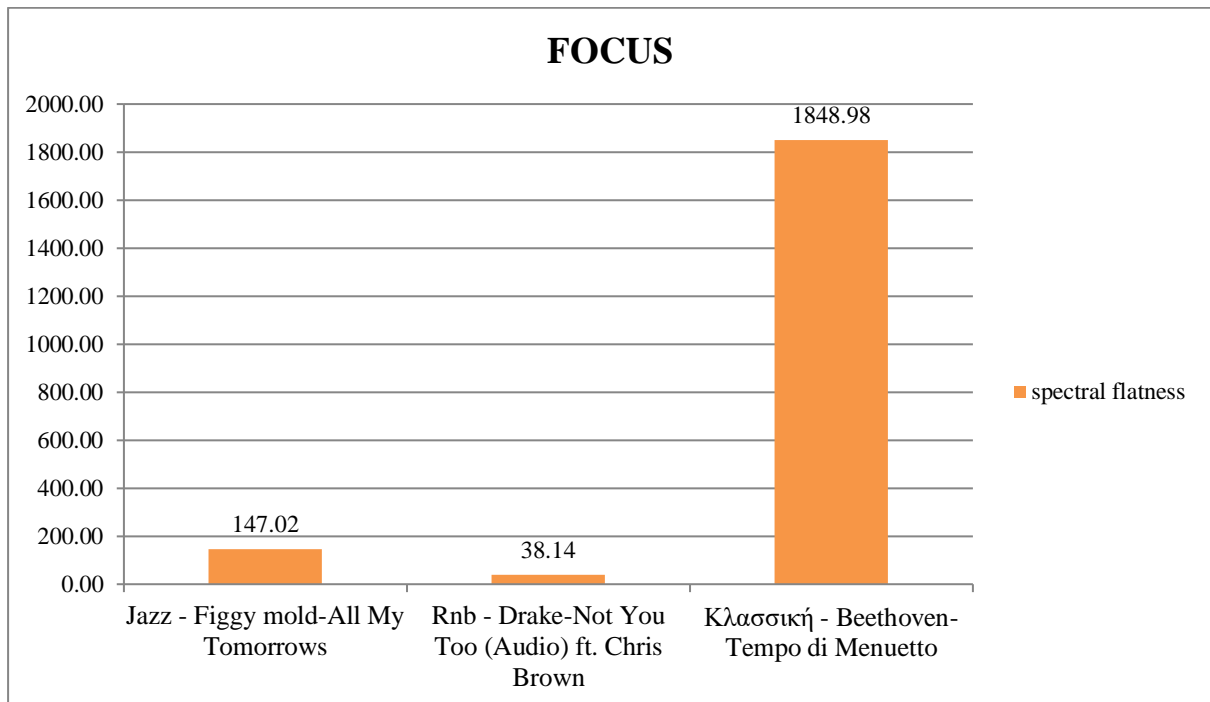
Εικόνα 4.57: Κύρτωση των MFCC (5^{ος} έως 8^{ος}) για την κατηγορία focus

Εικόνα 4.58: Κύρτωση του spectral contrast στις 6 ζώνες για την κατηγορία focus

Όσον αφορά τους mfcc παρατηρούνται περισσότερες αρνητικές τιμές. Στον 1^ο, 6^ο, 7^ο και 8^ο mfcc η υψηλότερη τιμή εμφανίζεται στο Κλασσική - Beethoven-Tempo di Menuetto, ενώ στον 2^ο 3^ο εμφανίζεται στο Rnb - Drake-Not You Too (Audio) ft. Chris Brown. Στον 4^ο και 5^ο mfcc η υψηλότερη τιμή εμφανίζεται στο Jazz - Figgy mold-All My Tomorrows. Στην περίπτωση του spectral contrast, στην 1^η, 2^η, 3^η, 4^η και 5^η ζώνη η υψηλότερη τιμή παρουσιάζεται στο Jazz - Figgy mold-All My Tomorrows ενώ στην 6η ζώνη εμφανίζεται στο Κλασσική - Beethoven-Tempo di Menuetto.

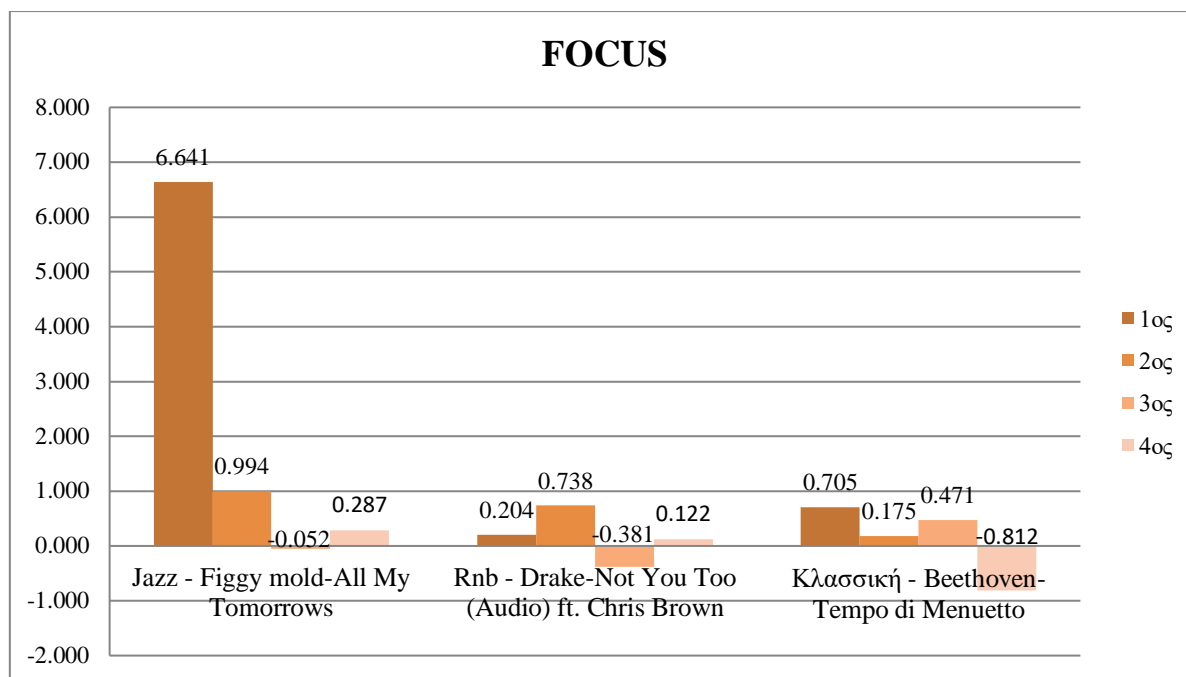
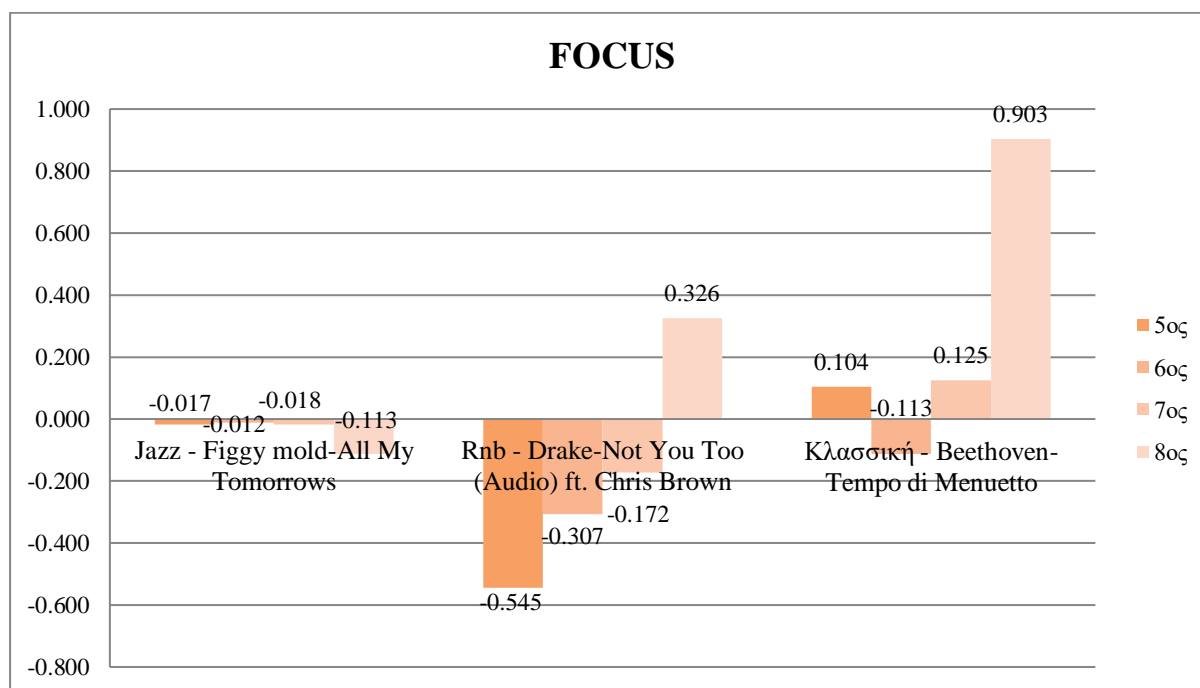


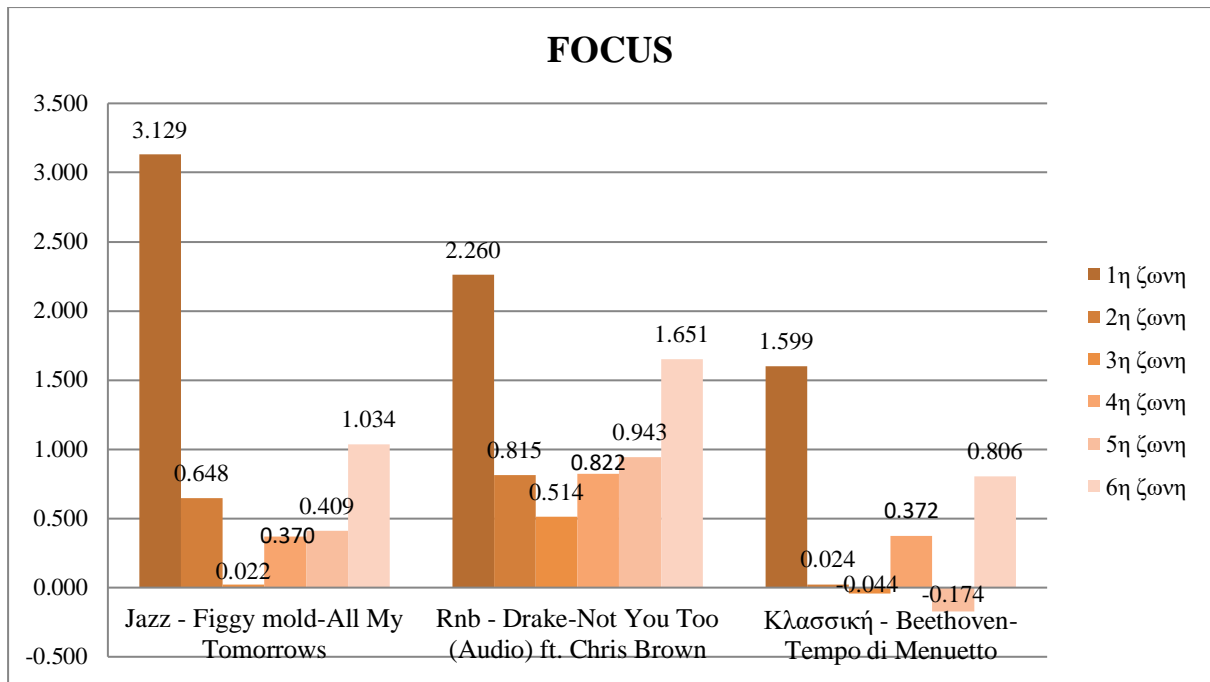
Εικόνα 4.59: Λοξότητα των χαρακτηριστικών για την κατηγορία focus



Εικόνα 4.60: Λοξότητα του spectral flatness για την κατηγορία focus

Από τα διαγράμματα παρατηρούμε ότι οι τιμές των τριών κομματιών για την κύρτωση στα πρώτα επτά χαρακτηριστικά είναι αρκετά κοντά. Η υψηλότερη τιμή στα πρώτα επτά χαρακτηριστικά παρατηρείται στο Κλασσική - Beethoven-Tempo di Menuetto.

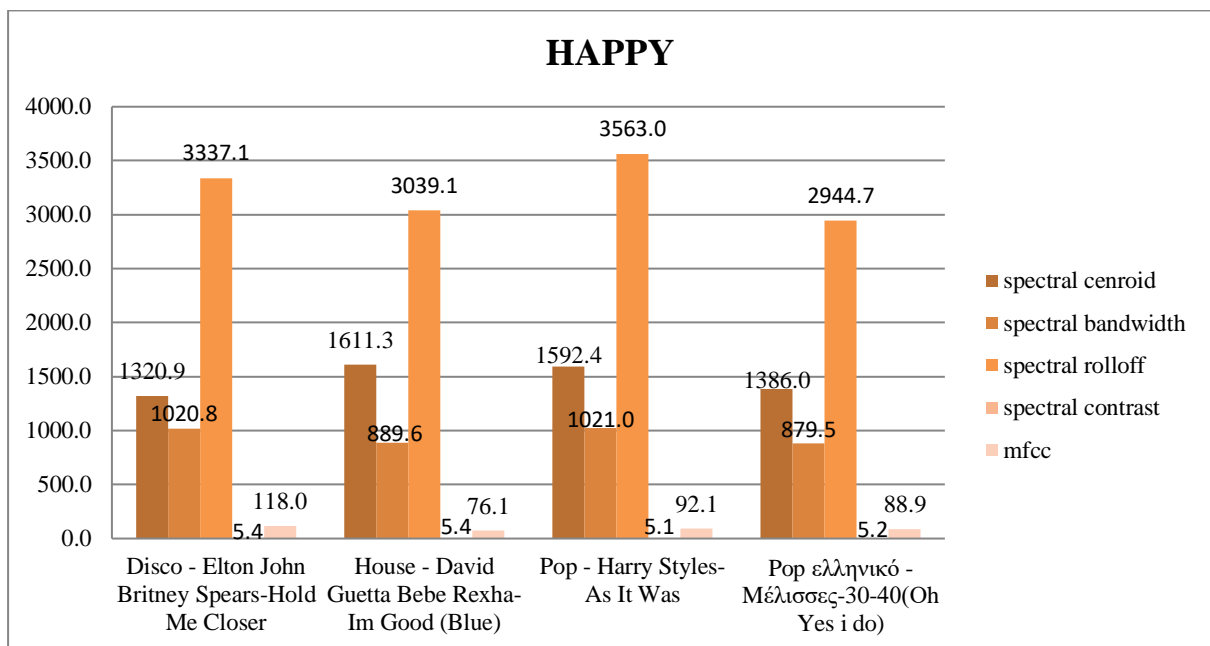
Εικόνα 4.61: Λοξότητα των MFCC (1^{ος} έως 4^{ος}) για την κατηγορία focusΕικόνα 4.62: Λοξότητα των MFCC (5^{ος} έως 8^{ος}) για την κατηγορία focus



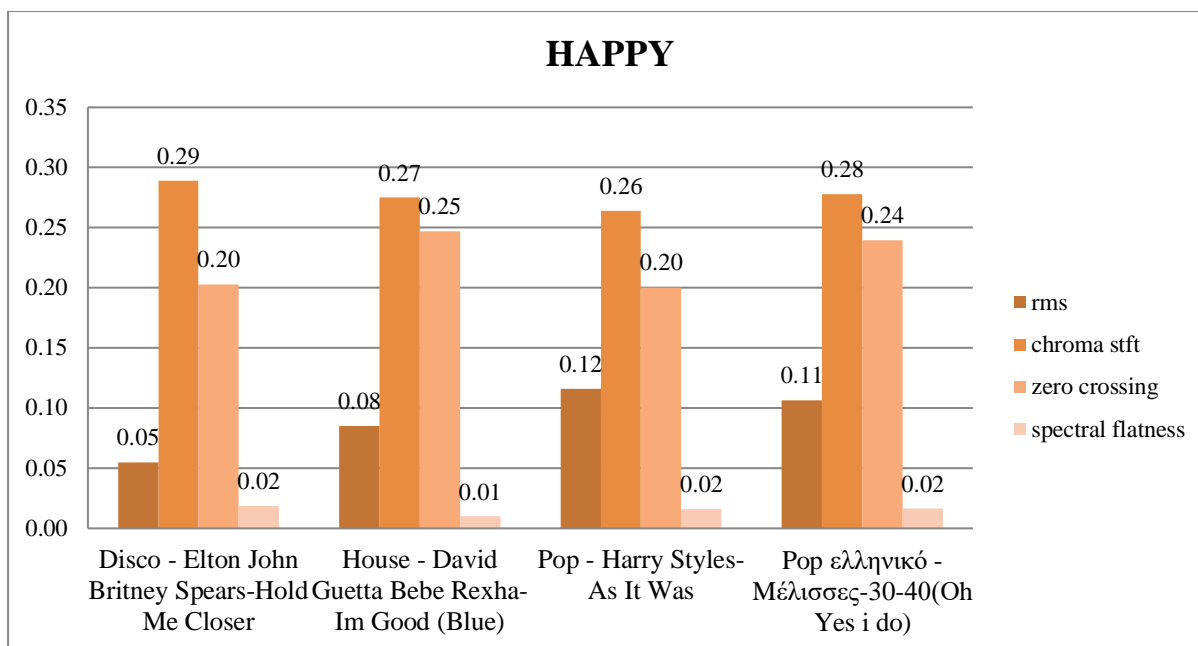
Εικόνα 4.63: Λοξότητα του spectral contrast στις 6 ζώνες για την κατηγορία focus

Όσον αφορά τους mfcc παρατηρούνται περισσότερες αρνητικές τιμές. Στον 1°, 6°, 7° και 8° mfcc η υψηλότερη τιμή εμφανίζεται στο Κλασσική - Beethoven-Tempo di Menuetto, ενώ στον 2° 3° εμφανίζεται στο Rnb - Drake-Not You Too (Audio) ft. Chris Brown. Στον 4° και 5° mfcc η υψηλότερη τιμή εμφανίζεται στο Jazz - Figgy mold-All My Tomorrows. Στην περίπτωση του spectral contrast, στην 1η ζώνη η υψηλότερη τιμή παρουσιάζεται στο Jazz - Figgy mold-All My Tomorrows, ενώ στην 2η 3η 4η 5η 6η ζώνη εμφανίζεται στο Rnb - Drake-Not You Too (Audio) ft. Chris Brown.

Στις παρακάτω εικόνες παρουσιάζονται τα γραφήματα για το συναίσθημα happy στα πέντε στατιστικά στοιχεία.

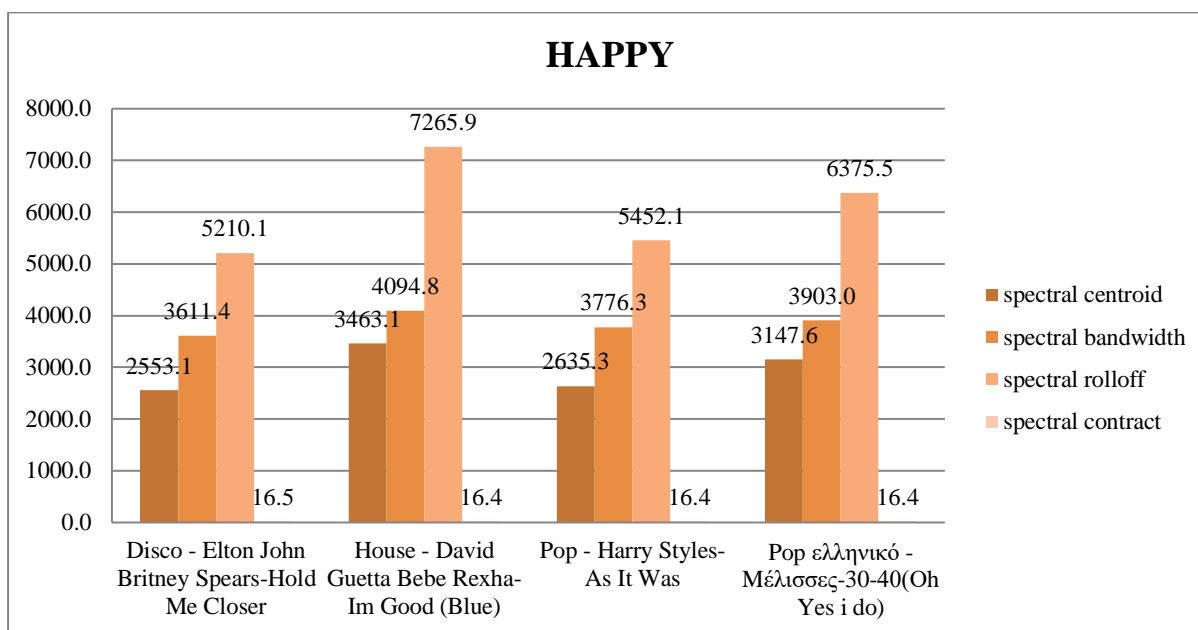


Εικόνα 4.64: Τυπική απόκλιση των χαρακτηριστικών για την κατηγορία happy

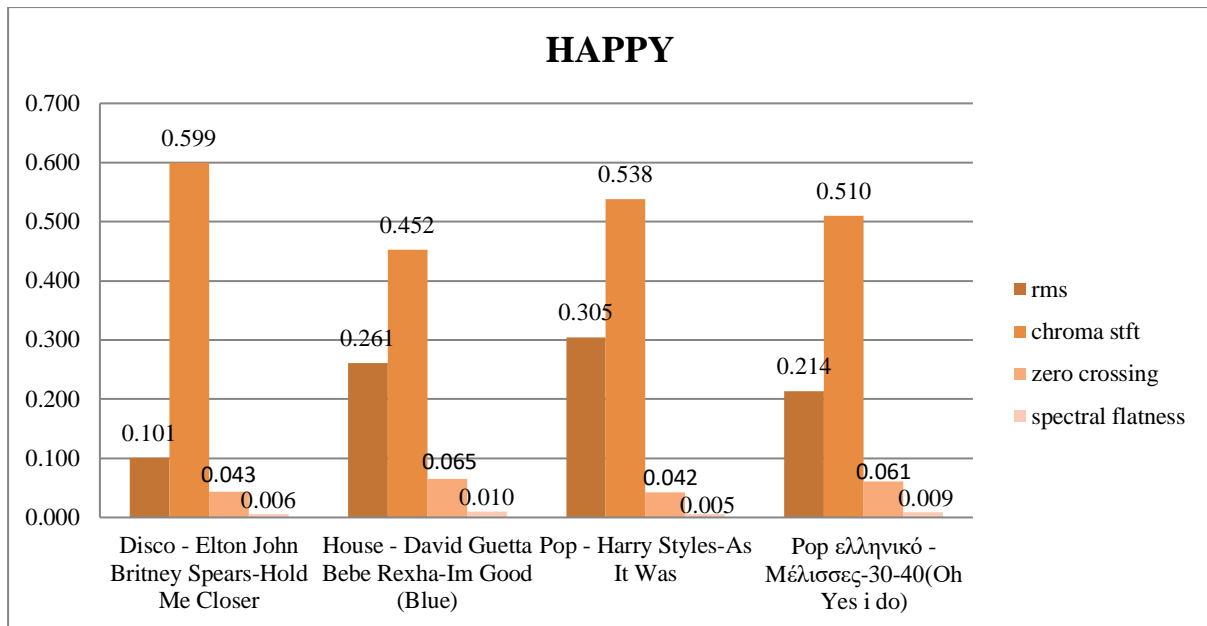


Εικόνα 4.65: Τυπική απόκλιση των χαρακτηριστικών για την κατηγορία happy

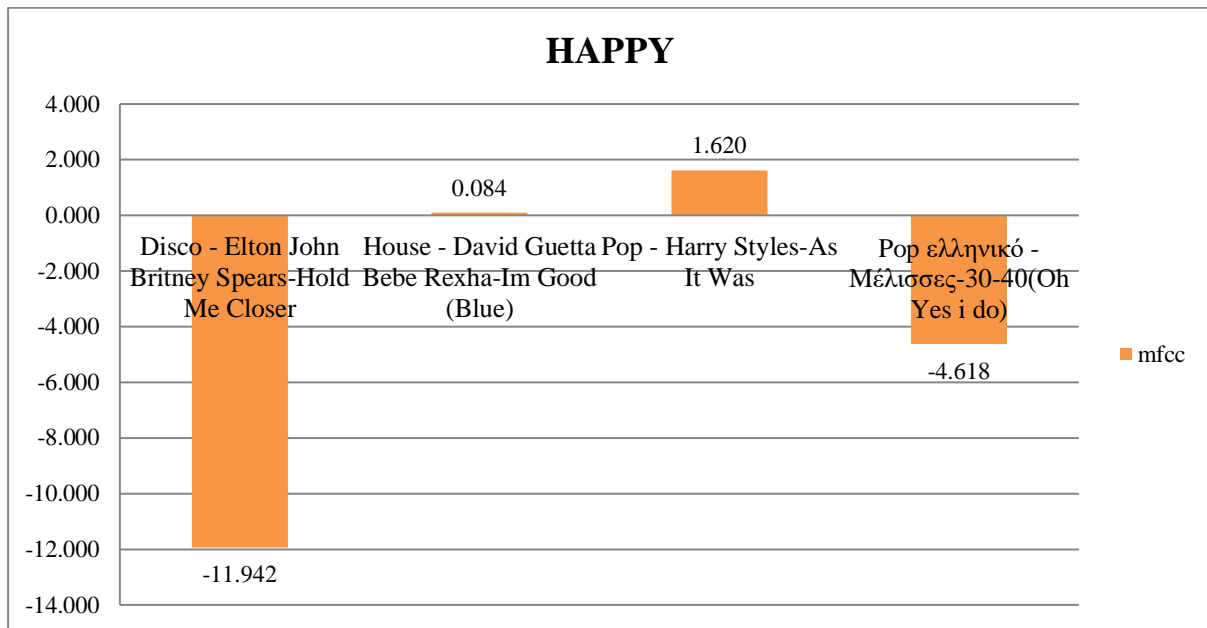
Από τα διαγράμματα παρατηρούμε ότι οι τιμές των τεσσάρων κομματιών για την τυπική απόκλιση στα πρώτα πέντε χαρακτηριστικά είναι αρκετά κοντά. Η υψηλότερη τιμή στα spectral centroid, spectral contrast και zero crossing παρατηρείται στο House - David Guetta Bebe Rexha-Im Good (Blue). Στα spectral bandwidth, spectral rolloff και rms η υψηλότερη τιμή εμφανίζεται στο Pop - Harry Styles-As It Was. Στους mfcc η υψηλότερη τιμή παρατηρείται στο Disco - Elton John Britney Spears-Hold Me Closer καθώς και στα chroma stft και spectral flatness.



Εικόνα 4.66: Μέση τιμή των χαρακτηριστικών για την κατηγορία happy

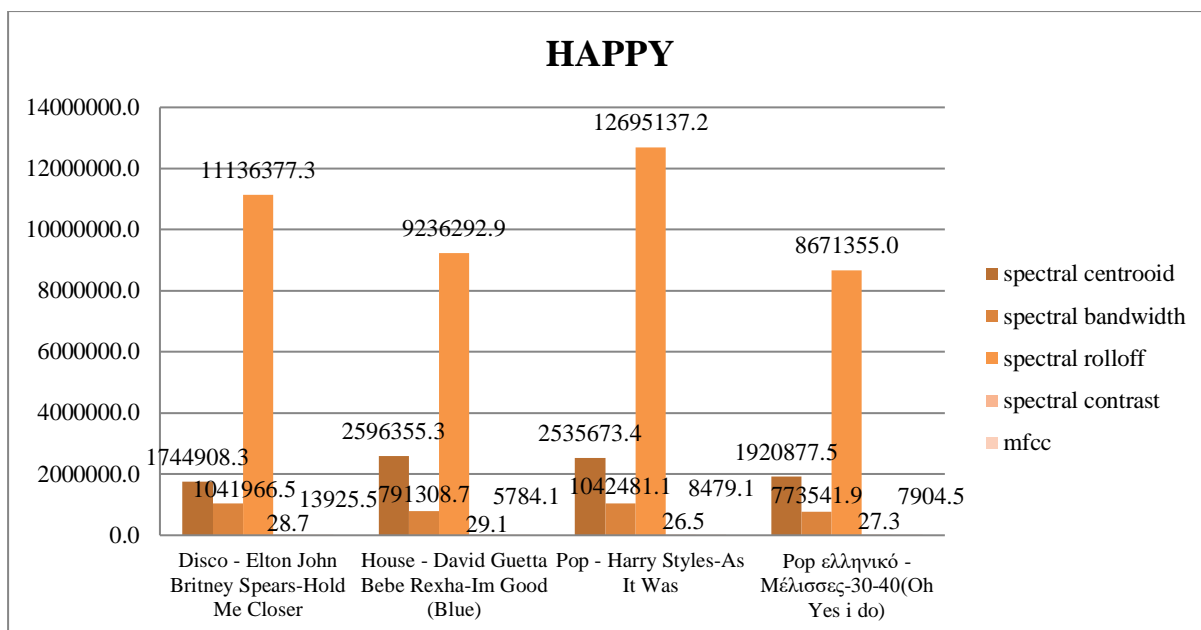


Εικόνα 4.67: Μέση τιμή των χαρακτηριστικών για την κατηγορία happy

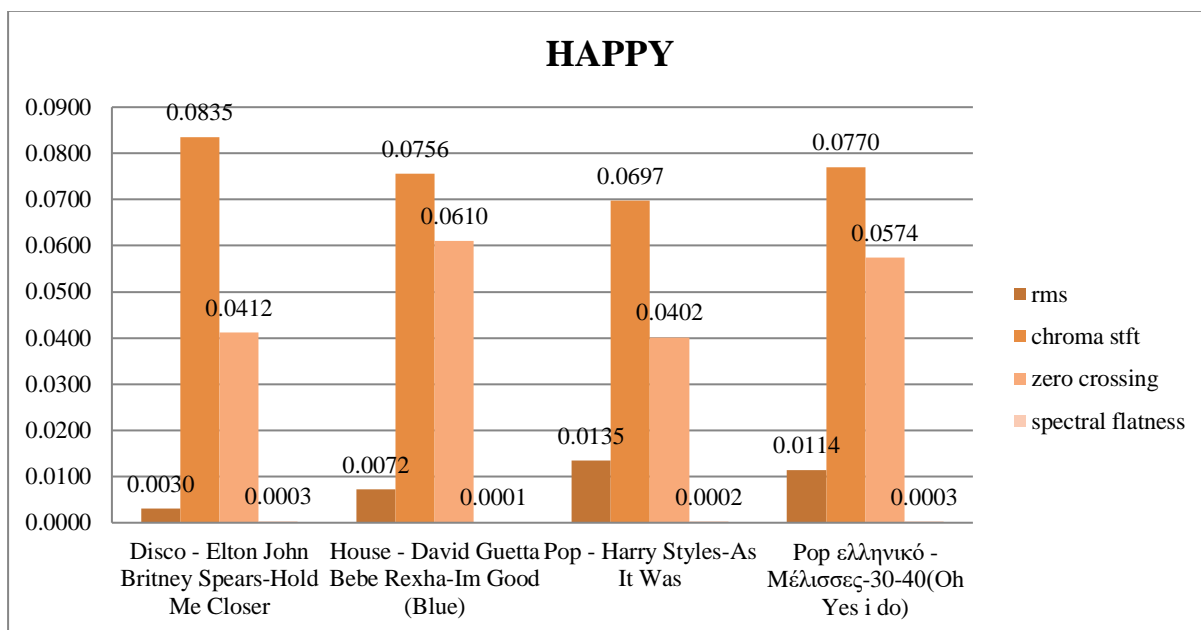


Εικόνα 4.68: Μέση τιμή των MFCC για την κατηγορία happy

Από τα διαγράμματα παρατηρούμε ότι οι τιμές των τεσσάρων κομματιών για την μέση τιμή σε όλα τα χαρακτηριστικά είναι αρκετά κοντά. Η υψηλότερη τιμή στα spectral centroid, spectral bandwidth, spectral rolloff, zero crossing και spectral flatness παρατηρείται στο House - David Guetta Bebe Rexha-Im Good (Blue). Στα spectral contrast, chroma stft και mfcc παρατηρείται στο Disco - Elton John Britney Spears-Hold Me Closer. Στην rms η υψηλότερη τιμή παρατηρείται στο Pop - Harry Styles-As It Was.

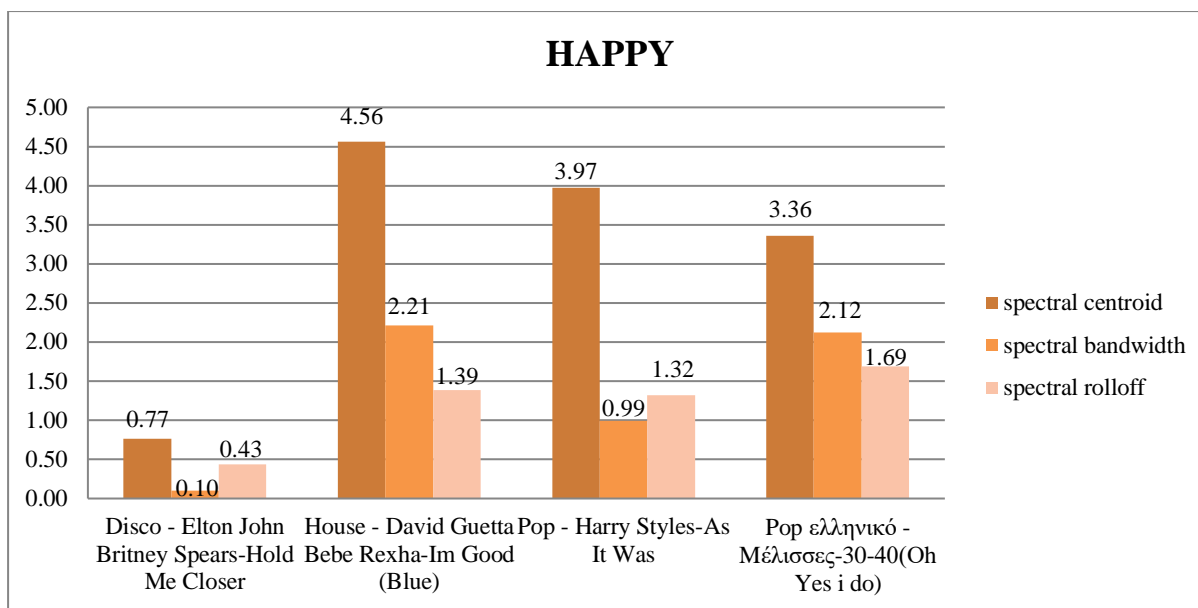


Εικόνα 4.69: Διακύμανση των χαρακτηριστικών για την κατηγορία happy

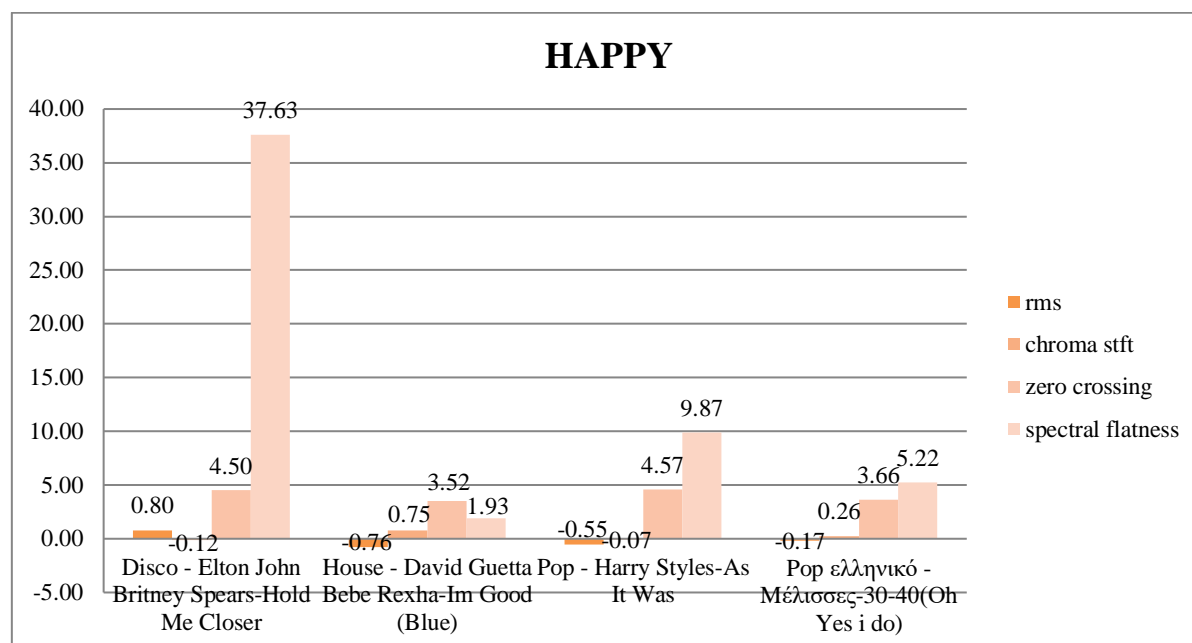


Εικόνα 4.70: Διακύμανση των χαρακτηριστικών για την κατηγορία happy

Από τα διαγράμματα παρατηρούμε ότι οι τιμές των τεσσάρων κομματιών για την διακύμανση στα πρώτα τέσσερα χαρακτηριστικά παρουσιάζουν αρκετή απόκλιση. Η υψηλότερη τιμή στο spectral centroid παρατηρείται στο House - David Guetta Bebe Rexha-Im Good (Blue) καθώς και στα spectral contrast και zero crossing. Στα spectral bandwidth, mfcc, chroma stft και spectral flatness η υψηλότερη τιμή παρατηρείται στο Disco - Elton John Britney Spears-Hold Me Closer. Στα spectral rolloff και rms η υψηλότερη τιμή παρατηρείται στο Pop - Harry Styles-As It Was.

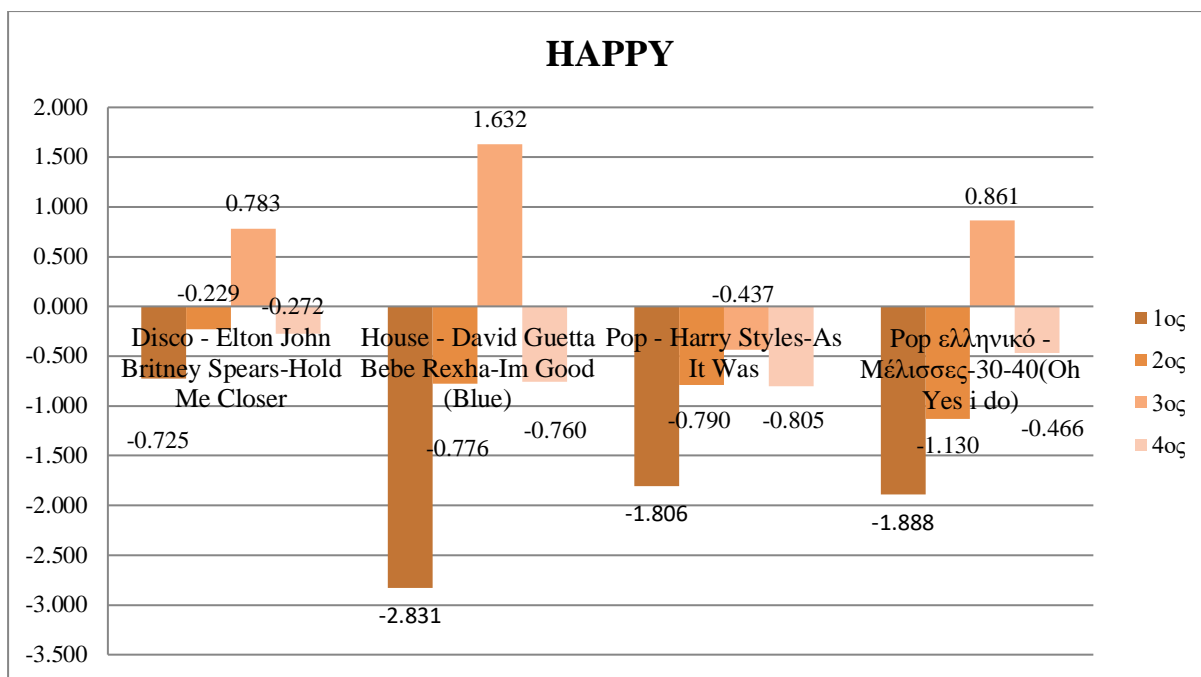
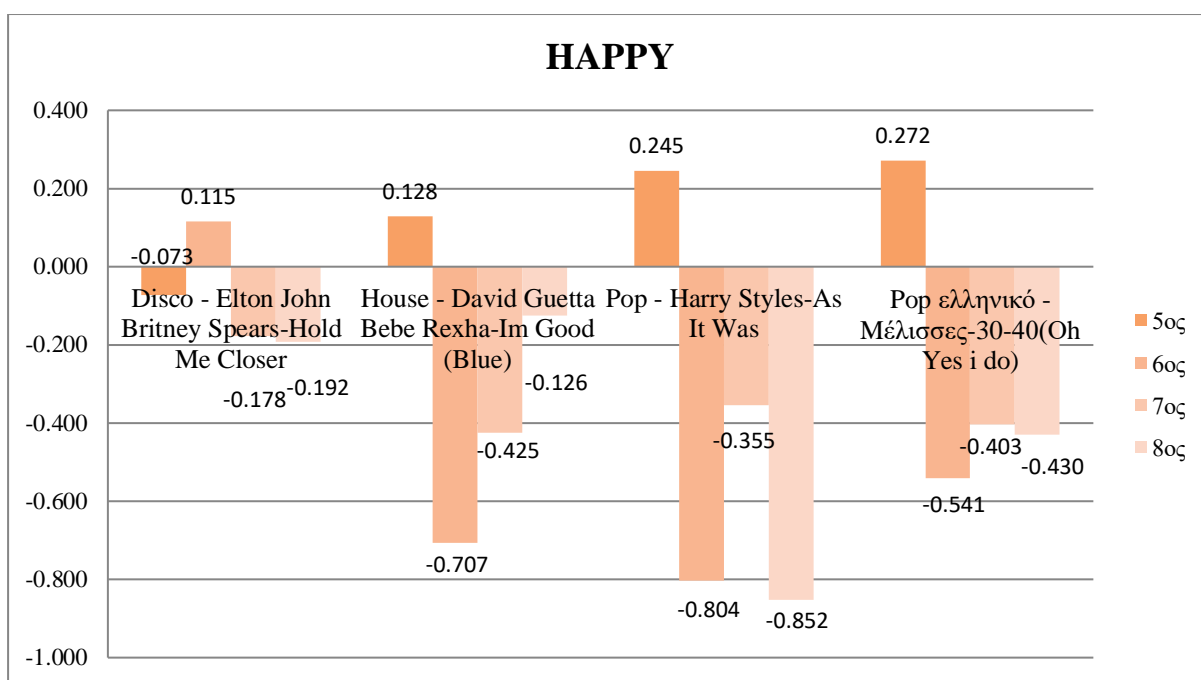


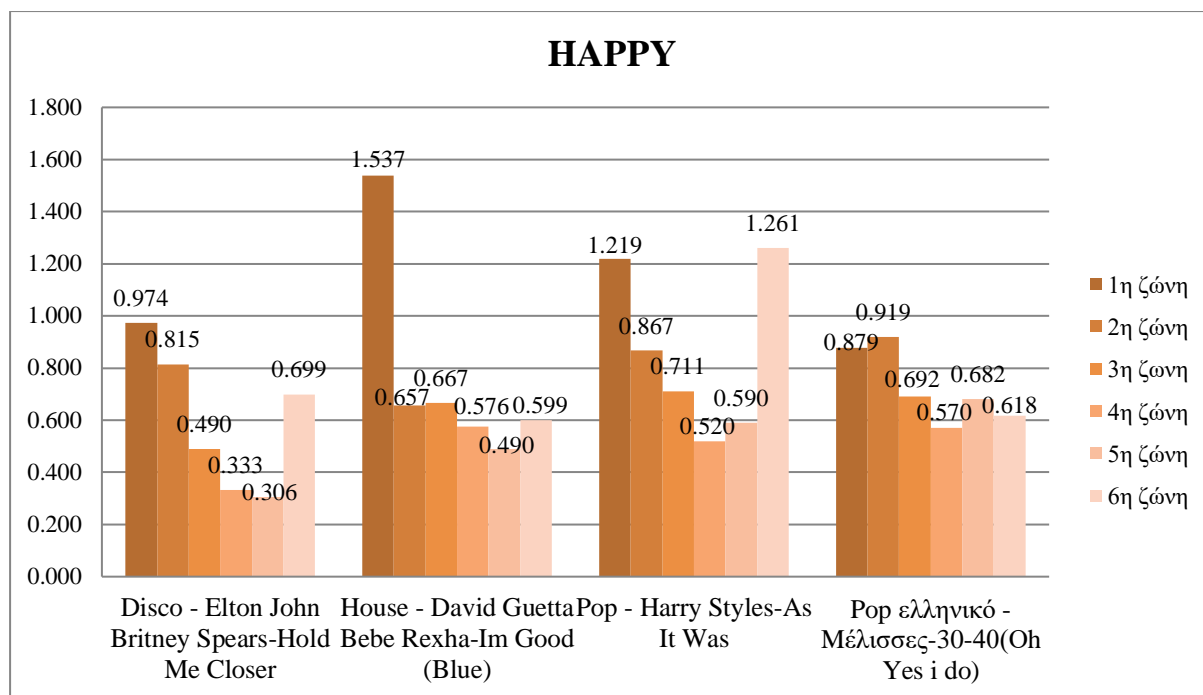
Εικόνα 4.71: Κύρτωση των χαρακτηριστικών για την κατηγορία happy



Εικόνα 4.72: Κύρτωση των χαρακτηριστικών για την κατηγορία happy

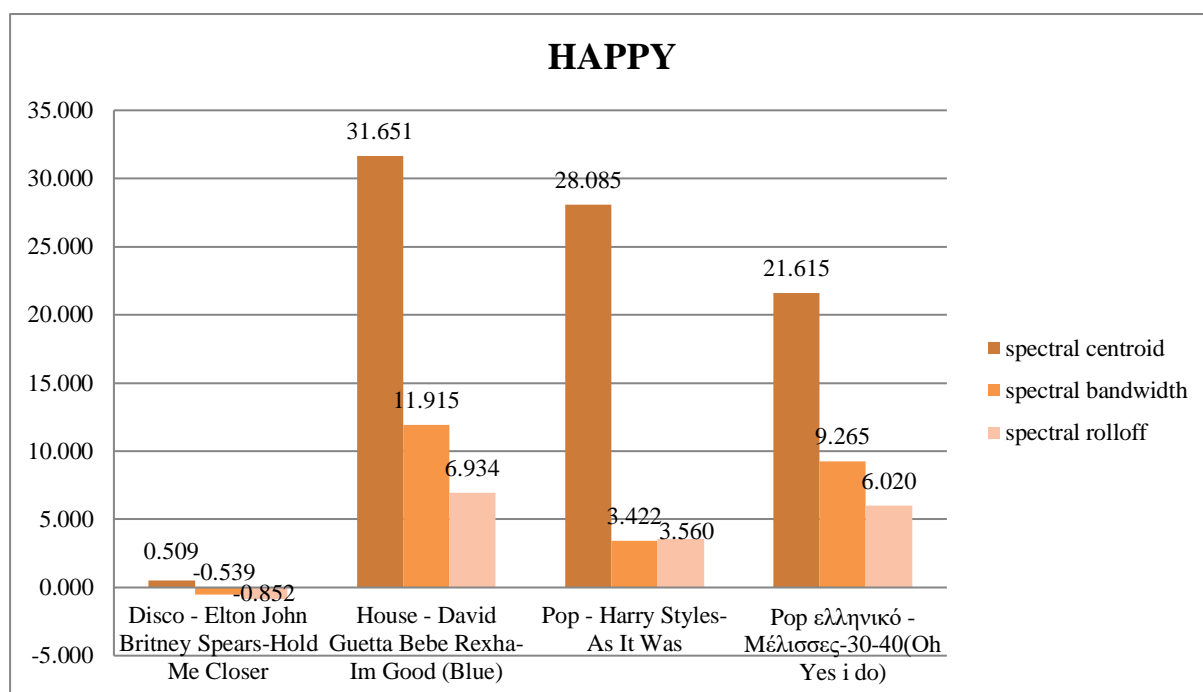
Από τα διαγράμματα παρατηρούμε ότι οι τιμές των τεσσάρων κομματιών για την κύρτωση σε όλα τα χαρακτηριστικά είναι πολύ κοντά. Η υψηλότερη τιμή στο spectral centroid παρατηρείται στο House - David Guetta Bebe Rexha-Im Good (Blue) καθώς και στα spectral bandwidth και chroma stft. Στο spectral rolloff η υψηλότερη τιμή εμφανίζεται στο Pop ελληνικό - Μέλισσες-30-40(Oh Yes i do) ενώ στα rms και spectral flatness παρατηρείται στο Disco - Elton John Britney Spears-Hold Me Closer. Στο zero crossing η υψηλότερη τιμή παρουσιάζεται στο Pop - Harry Styles-As It Was.

Εικόνα 4.73: Κύρτωση των MFCC (1^{ος} έως 4^{ος}) για την κατηγορία happyΕικόνα 4.74: Κύρτωση των MFCC (5^{ος} έως 8^{ος}) για την κατηγορία happy

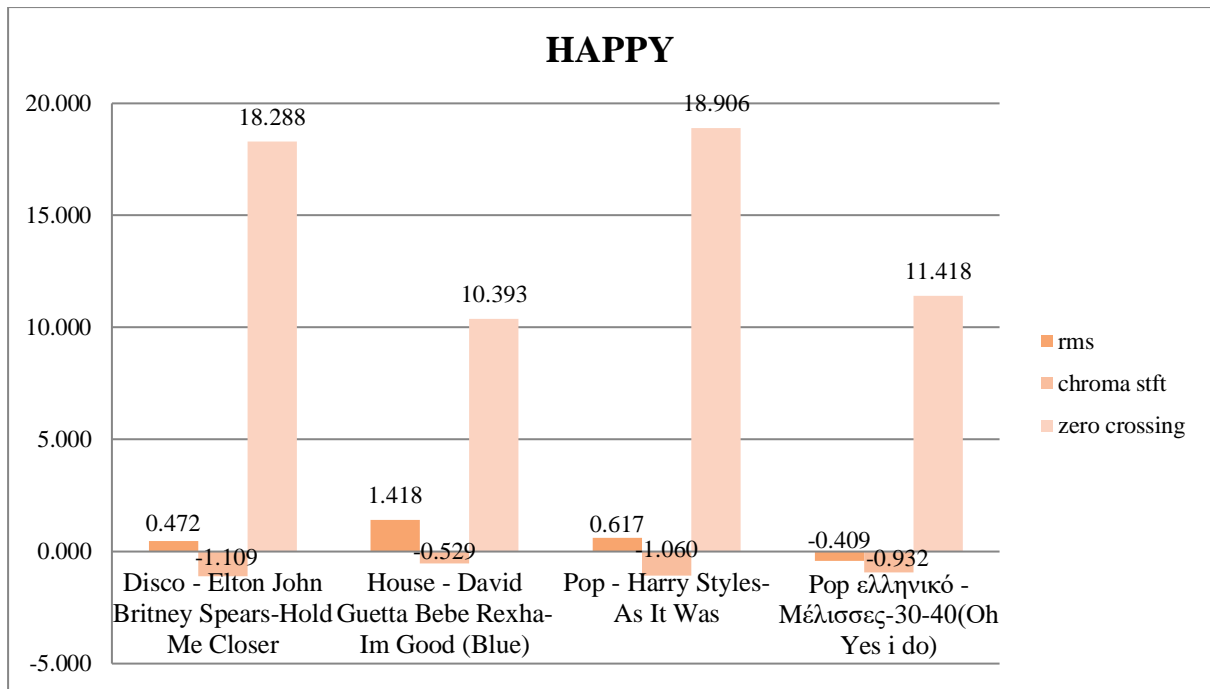


Εικόνα 4.75: Κύρτωση του spectral contrast στις 6 ζώνες για την κατηγορία happy

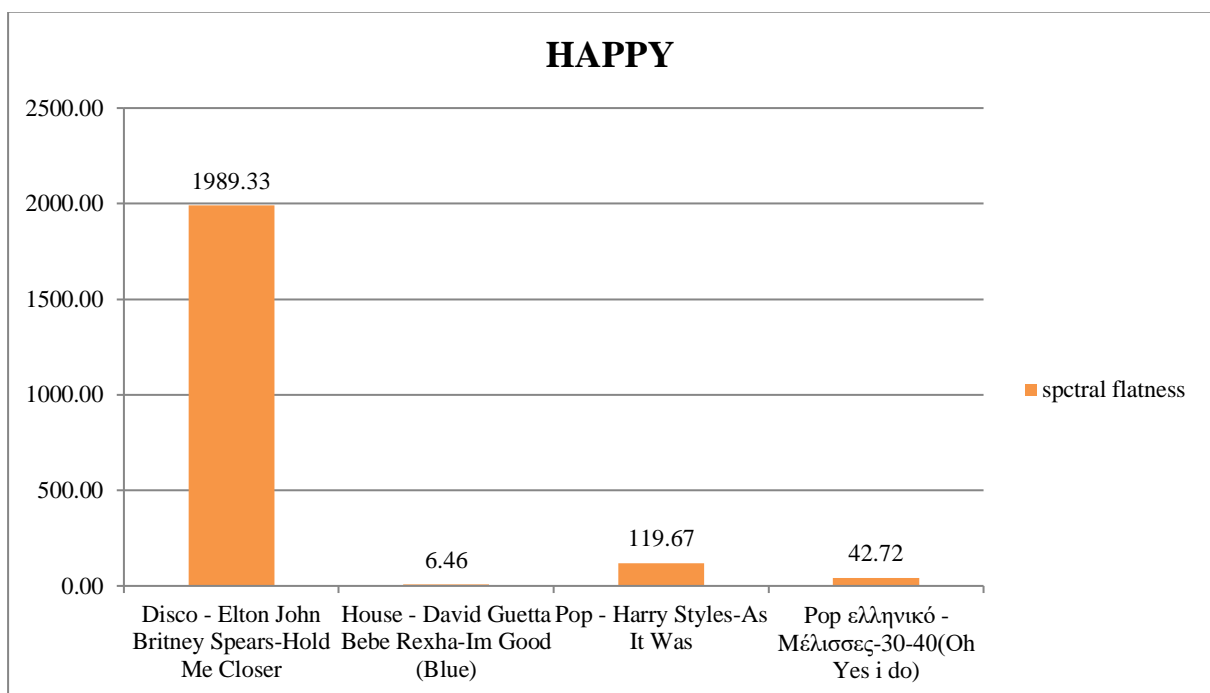
Όσον αφορά τους mfcc παρατηρούνται περισσότερες αρνητικές τιμές. Στον 1°, 3° και 7° mfcc η υψηλότερη τιμή εμφανίζεται στο House - David Guetta Bebe Rexha-Im Good (Blue), ενώ στον 2°, 5° και 8° εμφανίζεται στο Pop ελληνικό - Μέλισσες-30-40(Oh Yes i do). Στον 4° mfcc η υψηλότερη τιμή εμφανίζεται στο Pop - Harry Styles-As It Was. Στον 6° mfcc η υψηλότερη τιμή εμφανίζεται στο Disco - Elton John Britney Spears-Hold Me Closer. Στην περίπτωση του spectral contrast, στην 1^η και 4^η ζώνη η υψηλότερη τιμή παρουσιάζεται στο House - David Guetta Bebe Rexha-Im Good (Blue), ενώ στην 2^η και 5^η ζώνη εμφανίζεται στο Pop ελληνικό - Μέλισσες-30-40(Oh Yes i do). Στην 3^η και 6^η ζώνη η υψηλότερη τιμή παρουσιάζεται στο Pop - Harry Styles-As It Was.



Εικόνα 4.76: Λοξότητα των χαρακτηριστικών για την κατηγορία happy

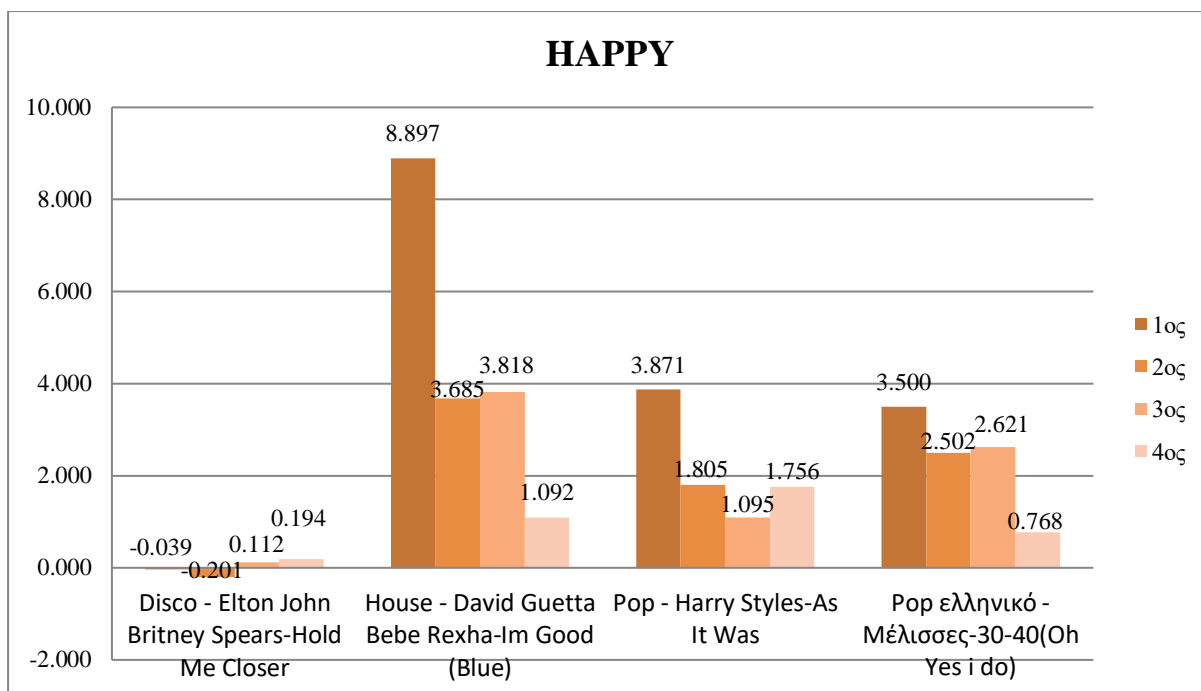
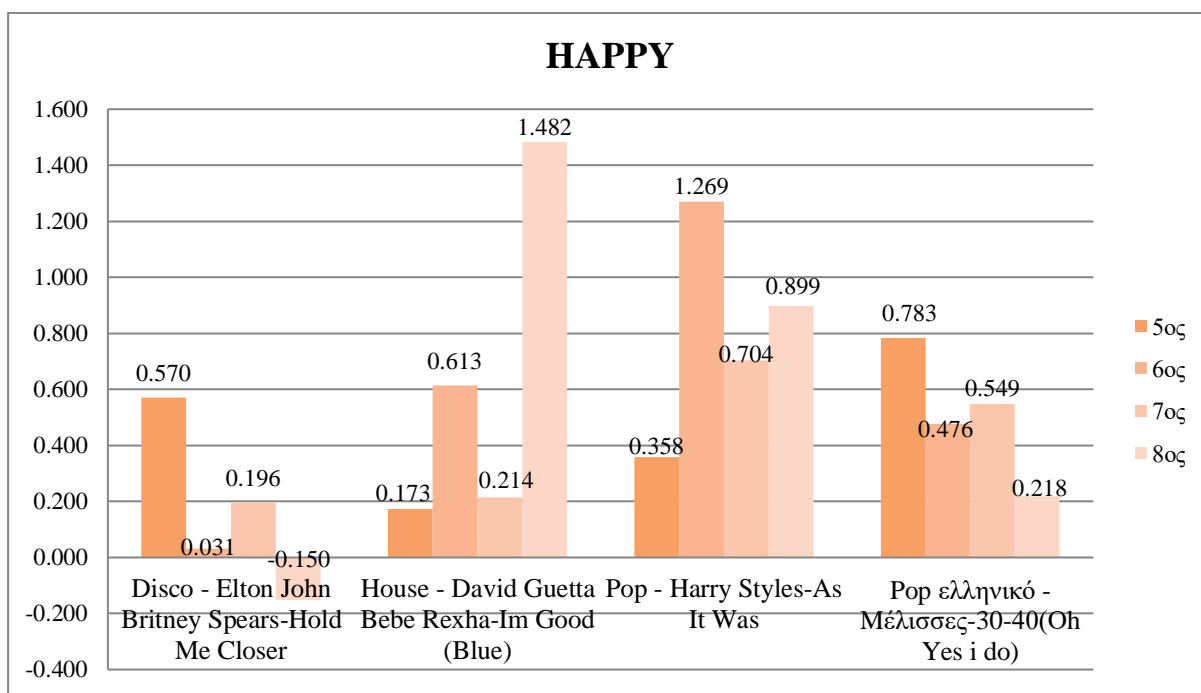


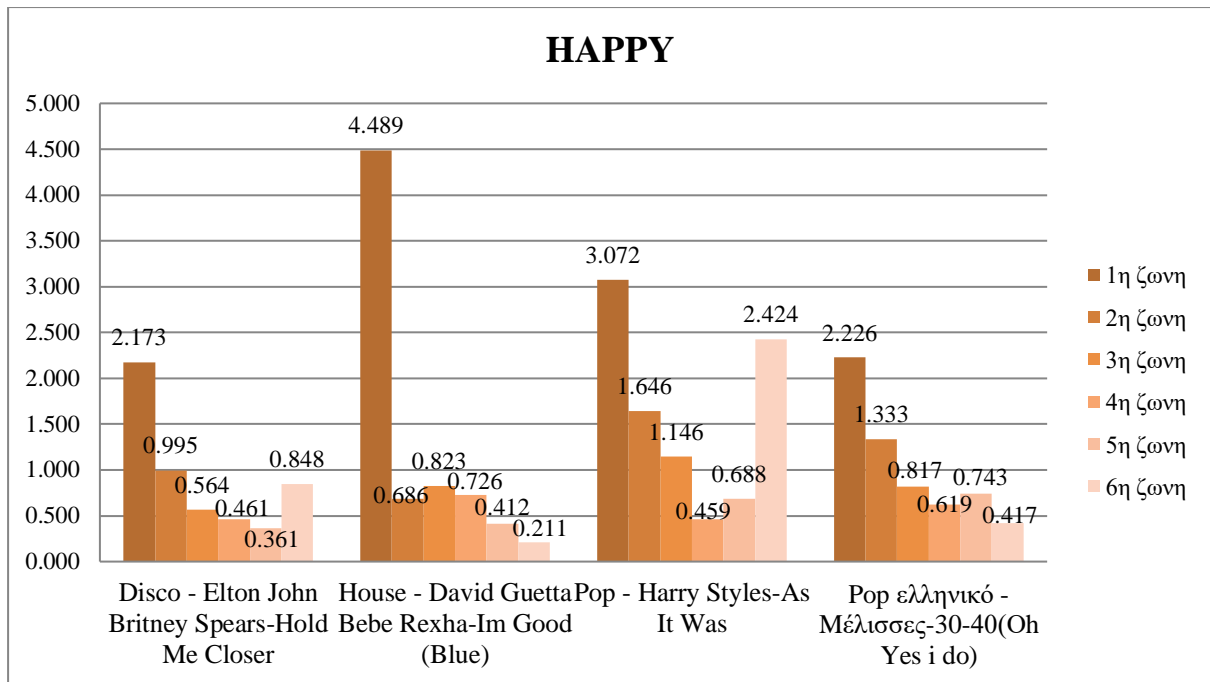
Εικόνα 4.77: Λοξότητα των χαρακτηριστικών για την κατηγορία happy



Εικόνα 4.78: Λοξότητα του spectral flatness για την κατηγορία happy

Από τα διαγράμματα παρατηρούμε ότι οι τιμές των τεσσάρων κομματιών για την λοξότητα παρουσιάζουν μία μικρή απόκλιση. Η υψηλότερη τιμή στο spectral centroid παρατηρείται στο House - David Guetta Bebe Rexha-Im Good (Blue) καθώς και στα spectral bandwidth spectral rolloff και rms. Στο chroma stft η υψηλότερη τιμή εμφανίζεται στο Disco - Elton John Britney Spears-Hold Me Closer καθώς στα zero crossing και spectral flatness.

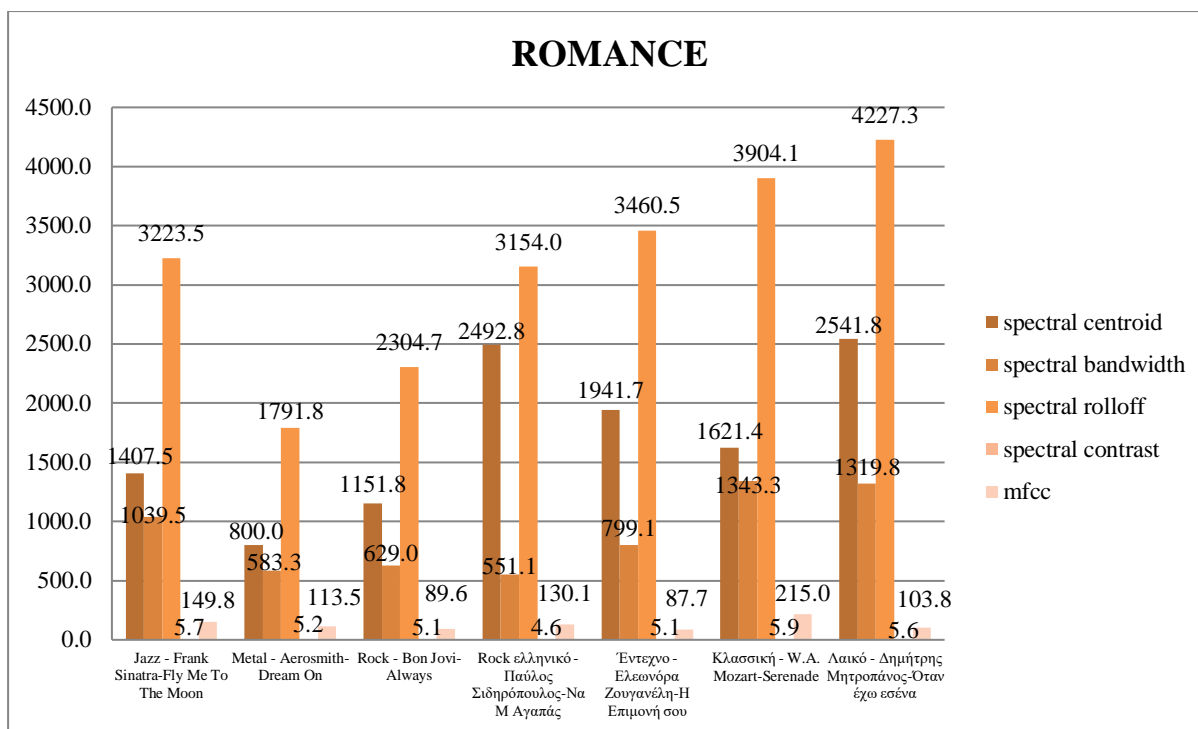
Εικόνα 4.79: Κύρτωση των MFCC (1^{ος} έως 4^{ος}) για την κατηγορία happyΕικόνα 4.80: Κύρτωση των MFCC (5^{ος} έως 8^{ος}) για την κατηγορία happy



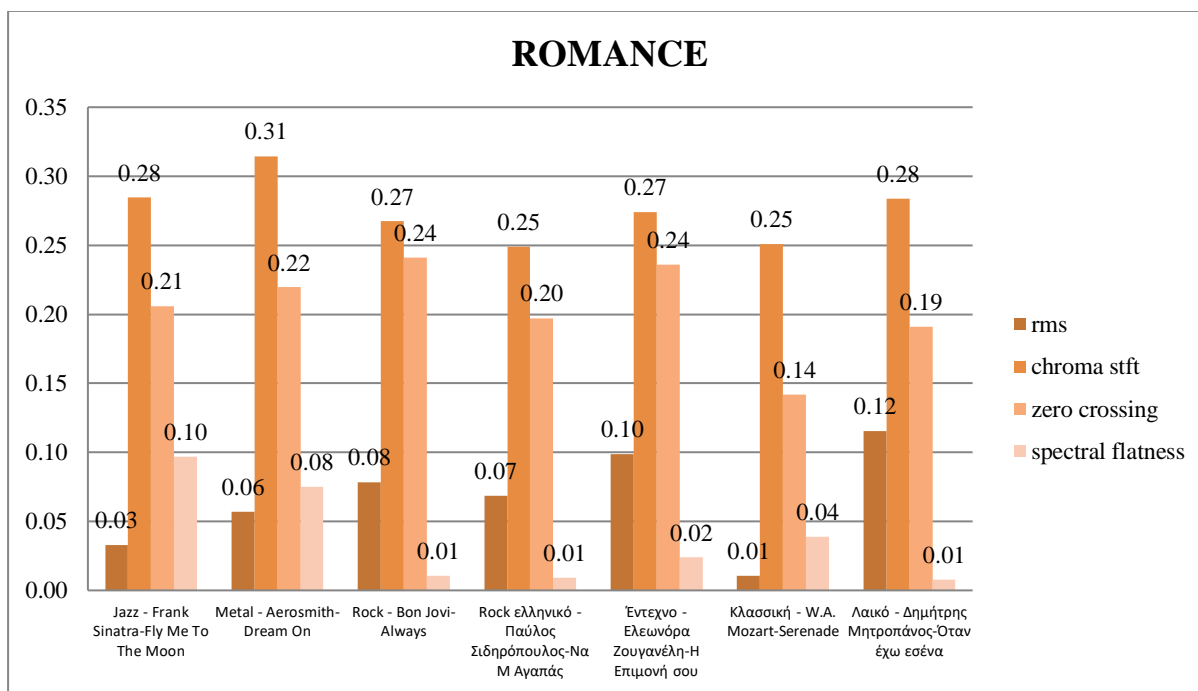
Εικόνα 4.81: Λοξότητα του spectral contrast στις 6 ζώνες για την κατηγορία happy

Όσον αφορά τους mfcc παρατηρούνται περισσότερες αρνητικές τιμές. Στον 1^ο, 2^ο, 3^ο και 8^ο mfcc η υψηλότερη τιμή εμφανίζεται στο House - David Guetta Bebe Rexha-Im Good (Blue), ενώ στον 4^ο, 6^ο και 7^ο εμφανίζεται στο Pop - Harry Styles-As It Was. Στον 5^ο mfcc η υψηλότερη τιμή εμφανίζεται στο Pop ελληνικό - Μέλισσες-30-40(Oh Yes i do). Στην περίπτωση του spectral contrast, στην 1^η και 4^η ζώνη η υψηλότερη τιμή παρουσιάζεται στο House - David Guetta Bebe Rexha-Im Good (Blue), ενώ στην 2^η, 3^η και 6^η εμφανίζεται στο Pop - Harry Styles-As It Was. Στην 5^η ζώνη η υψηλότερη τιμή παρουσιάζεται στο Pop ελληνικό - Μέλισσες-30-40(Oh Yes i do).

Στις παρακάτω εικόνες παρουσιάζονται τα γραφήματα για το συναίσθημα romance στα πέντε στατιστικά στοιχεία.

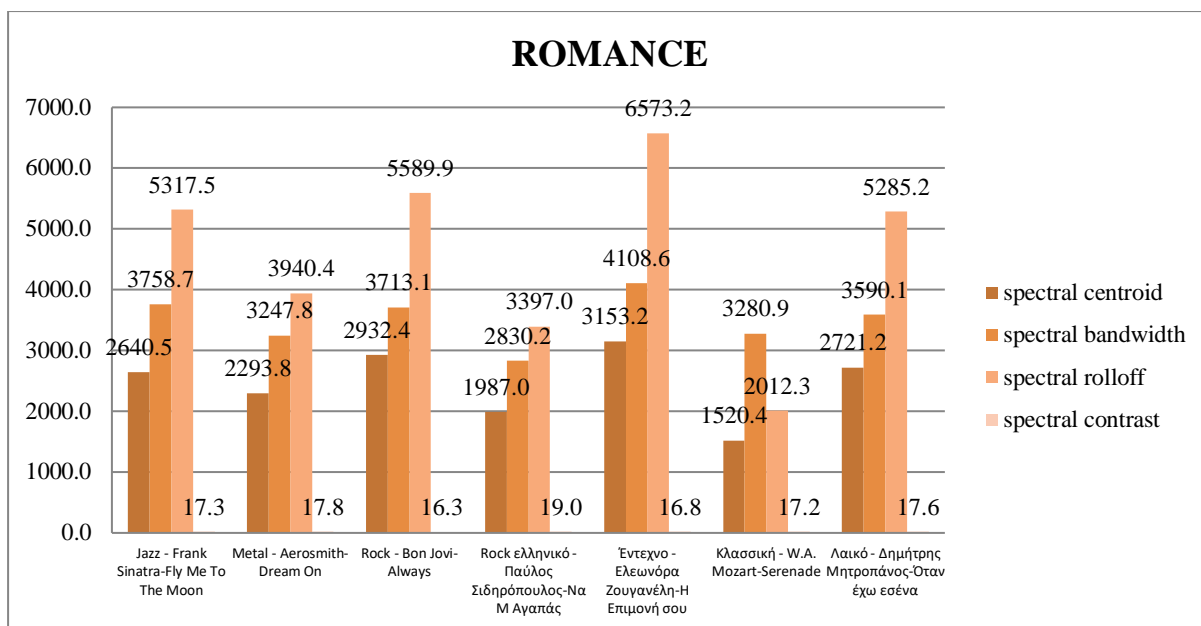


Εικόνα 4.82: Τυπική απόκλιση των χαρακτηριστικών για την κατηγορία romance

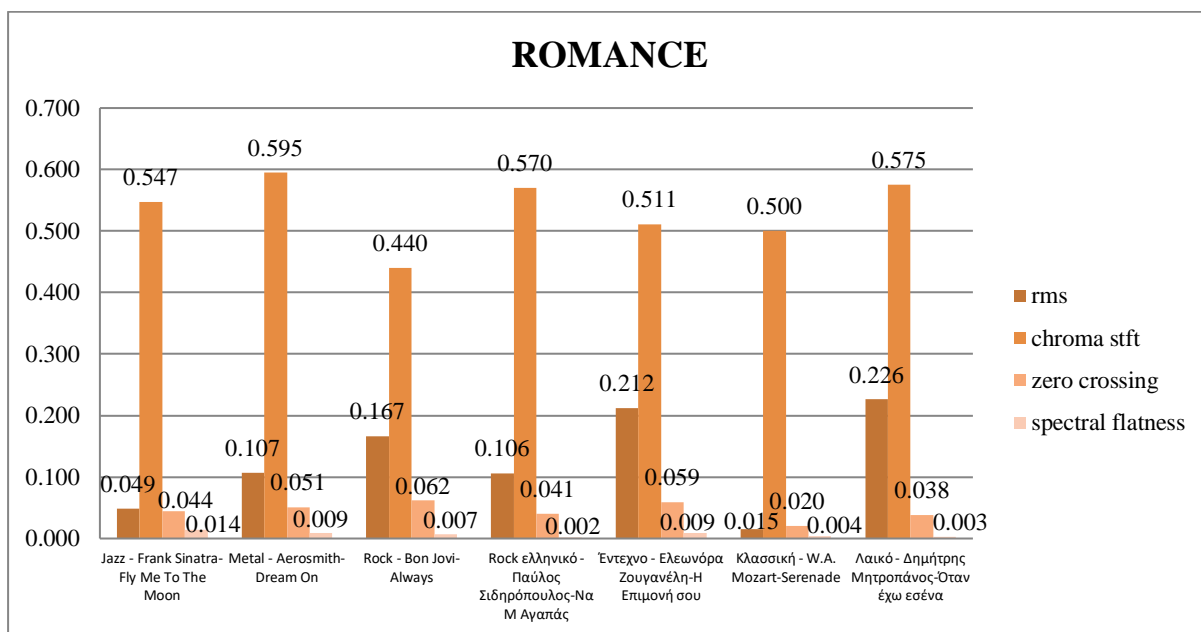


Εικόνα 4.83: Τυπική απόκλιση των χαρακτηριστικών για την κατηγορία romance

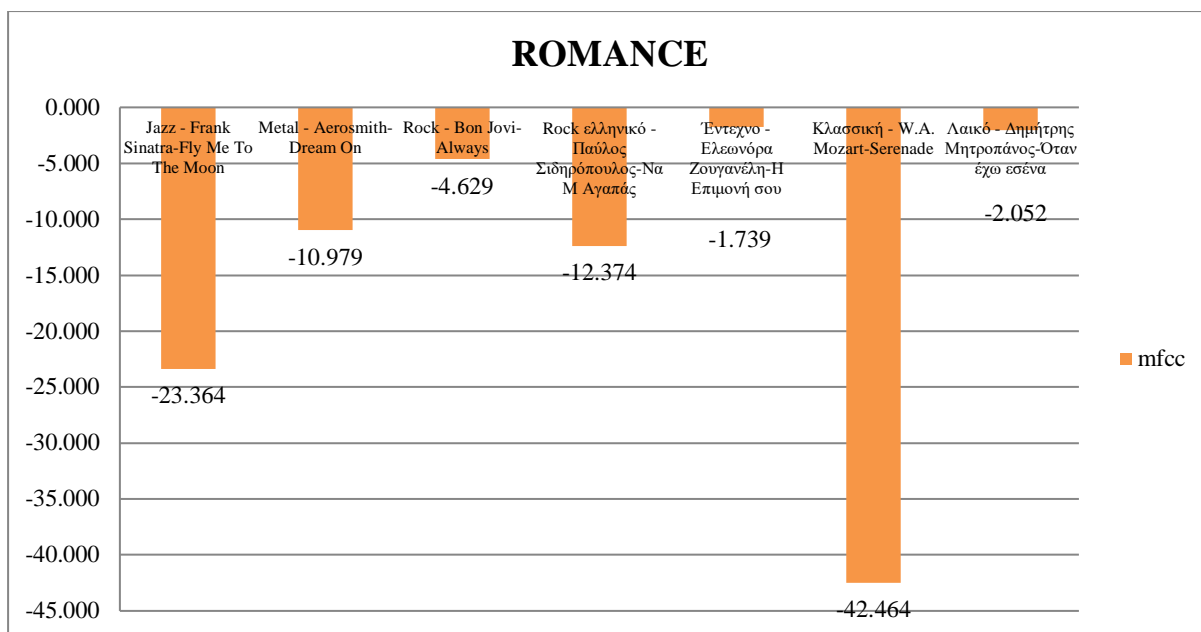
Από τα διαγράμματα παρατηρούμε ότι οι τιμές των επτά κομματιών για την τυπική απόκλιση σε όλα τα χαρακτηριστικά είναι σχετικά κοντά. Η υψηλότερη τιμή στα spectral centroid, spectral rolloff, spectral contrast και mfcc rms παρατηρείται στο Λαϊκό - Δημήτρης Μητροπάνος-Όταν έχω εσένα ενώ στο spectral bandwidth παρατηρείται στο Κλασσική - W.A. Mozart-Serenade. Στο chroma stft η υψηλότερη τιμή παρατηρείται στο Metal - Aerosmith-Dream On. Στο zero crossing η υψηλότερη τιμή παρατηρείται στο Rock - Bon Jovi-Always ενώ στο spectral flatness η υψηλότερη τιμή παρατηρείται στο Jazz - Frank Sinatra-Fly Me To The Moon.



Εικόνα 4.84: Μέση τιμή των χαρακτηριστικών για την κατηγορία romance

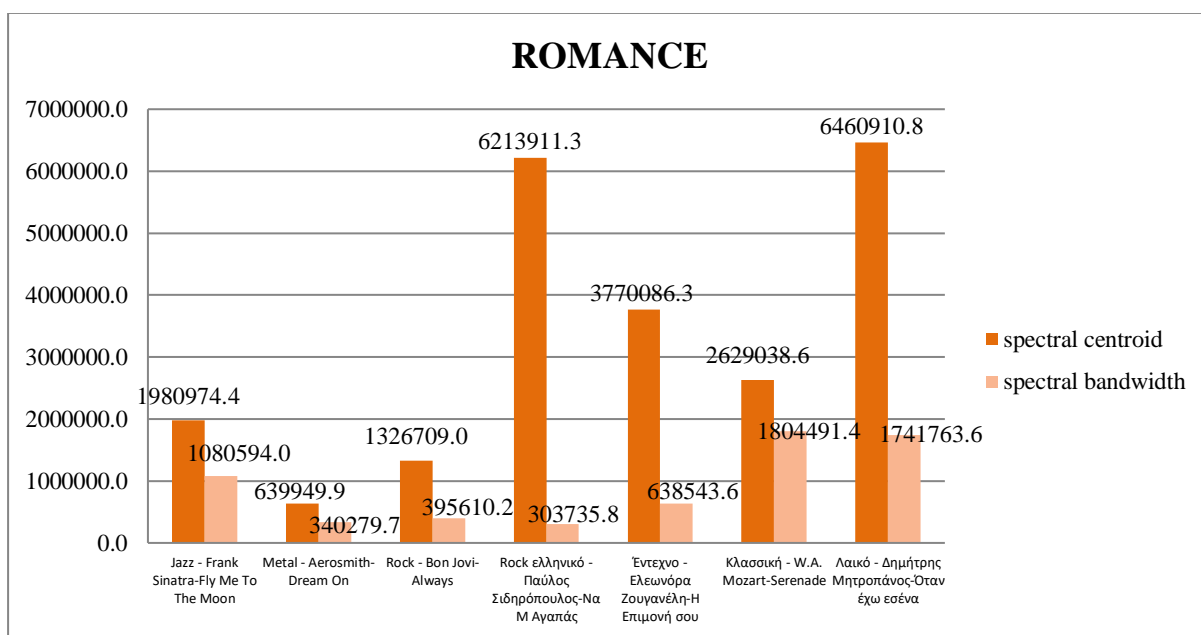


Εικόνα 4.85: Μέση τιμή των χαρακτηριστικών για την κατηγορία romance

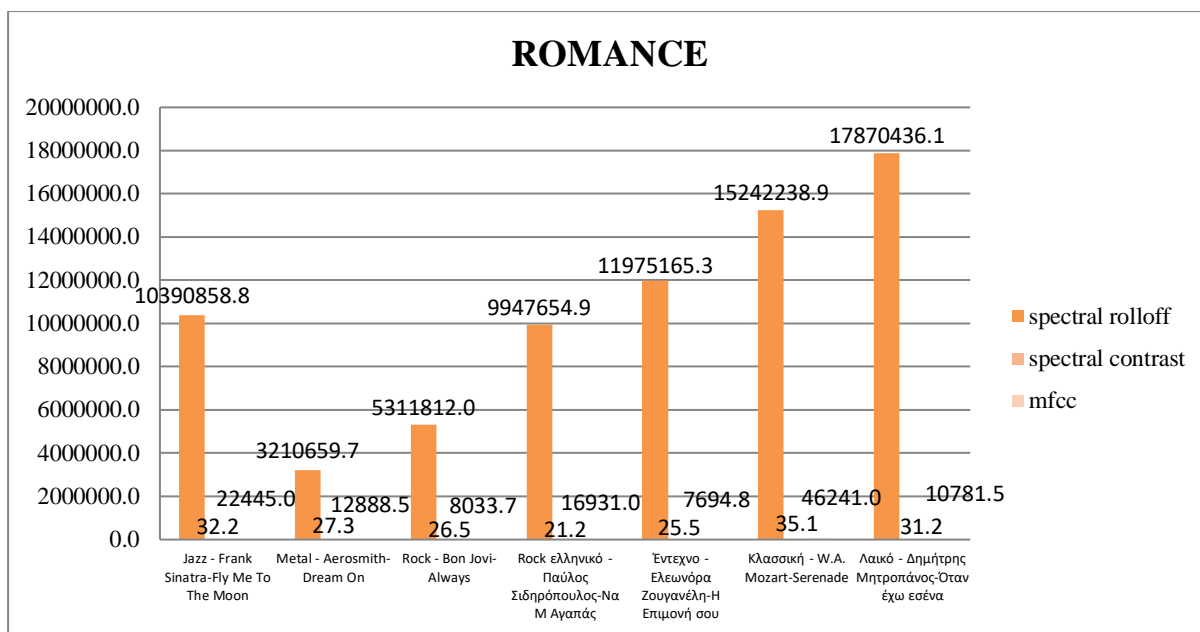


Εικόνα 4.86: Μέση τιμή των MFCC για την κατηγορία romance

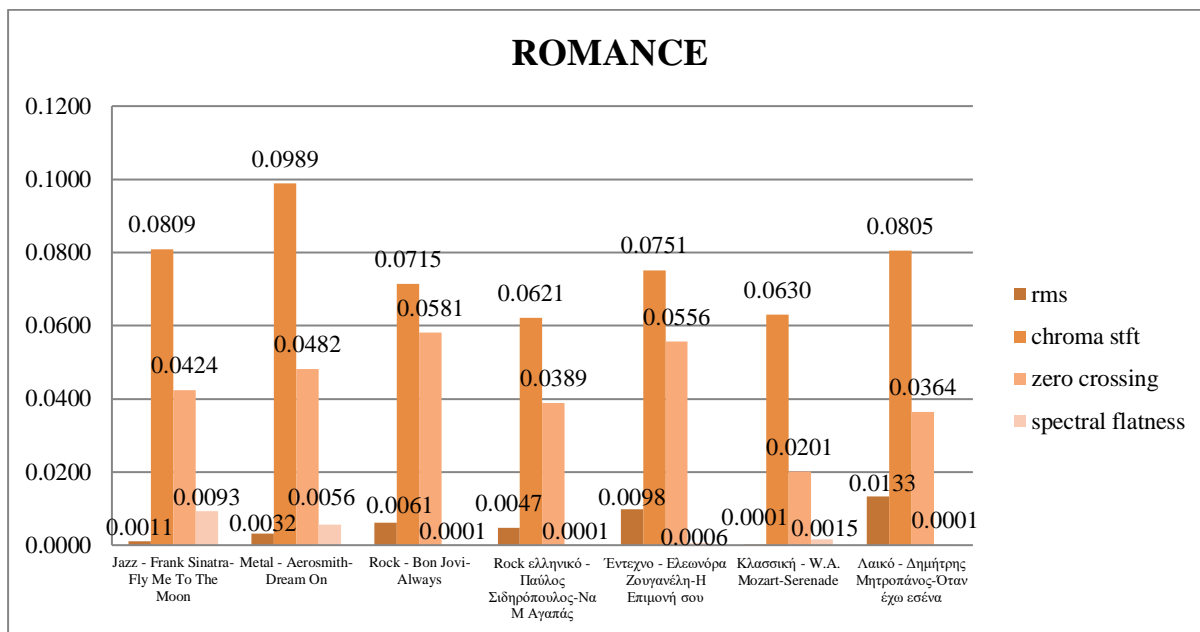
Από τα διαγράμματα παρατηρούμε ότι οι τιμές των επτά κομματιών για την μέση τιμή σε όλα τα χαρακτηριστικά είναι αρκετά κοντά. Η υψηλότερη τιμή στο spectral centroid παρατηρείται στο Έντεχνο - Ελεωνόρα Ζουγανέλη-Η Επιμονή σου καθώς και στα spectral bandwidth και spectral rolloff. Στα spectral contrast η υψηλότερη τιμή παρατηρείται στο Rock ελληνικό - Παύλος Σιδηρόπουλος-Να Μ' Αγαπάς. Στην rms η υψηλότερη τιμή παρατηρείται στο Λαϊκό - Δημήτρης Μητροπάνος-Όταν έχω εσένα ενώ στο chroma stft εμφανίζεται στο Metal - Aerosmith-Dream On. Στο zero crossing η υψηλότερη τιμή παρατηρείται στο Rock - Bon Jovi-Always ενώ στο spectral flatness η υψηλότερη τιμή εμφανίζεται στο Jazz - Frank Sinatra-Fly Me To The Moon. Στους mfcc η υψηλότερη τιμή παρατηρείται στο Κλασσική - W.A. Mozart-Serenade.



Εικόνα 4.87: Διακύμανση των χαρακτηριστικών για την κατηγορία romance

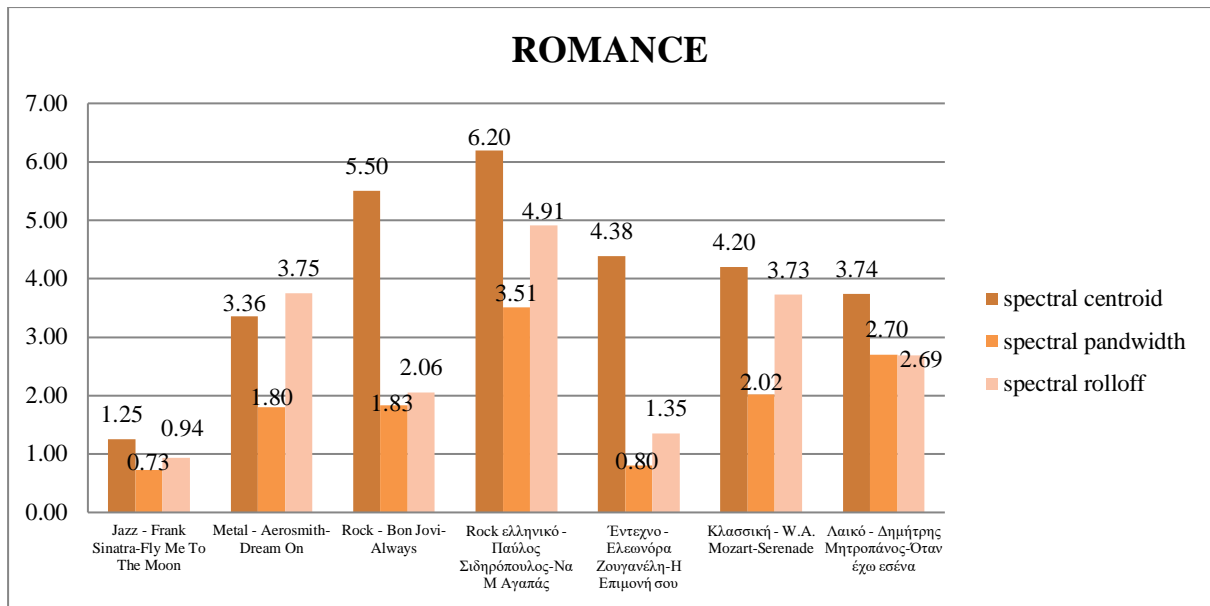


Εικόνα 4.88: Διακύμανση των χαρακτηριστικών για την κατηγορία romance

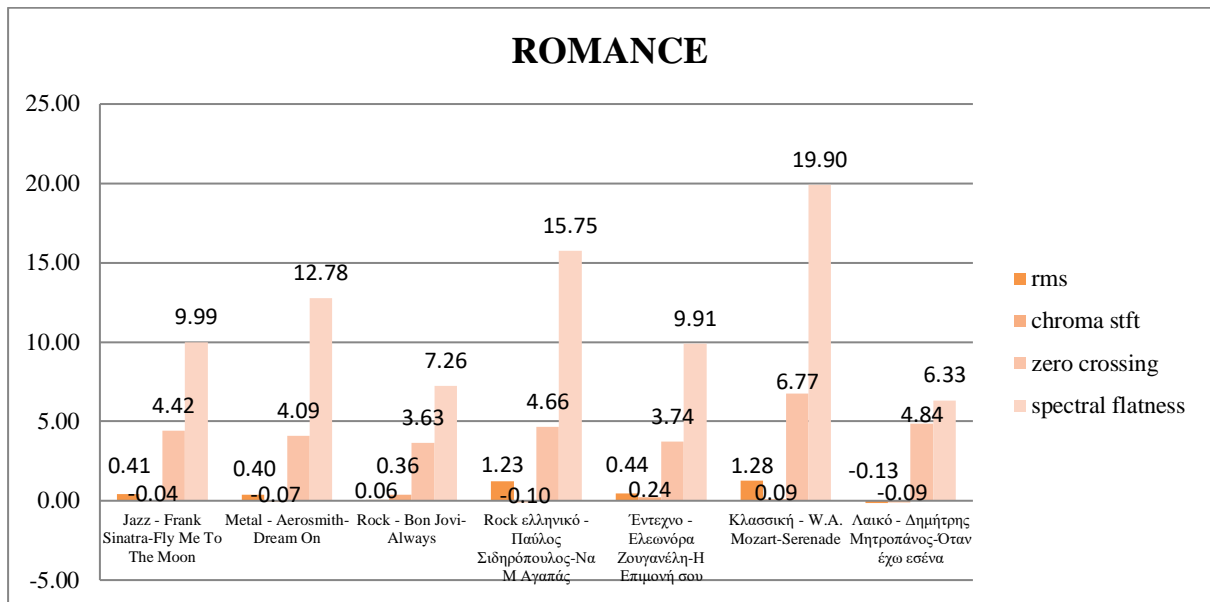


Εικόνα 4.89: Διακύμανση των χαρακτηριστικών για την κατηγορία romance

Από τα διαγράμματα παρατηρούμε ότι οι τιμές των επτά κομματιών για την διακύμανση σε όλα τα χαρακτηριστικά είναι αρκετά κοντά. Η υψηλότερη τιμή στο spectral centroid παρατηρείται στο Λαϊκό - Δημήτρης Μητροπάνος-Όταν έχω εσένα καθώς και στα spectral rolloff και rms. Στο spectral bandwidth η υψηλότερη τιμή παρατηρείται στο Κλασσική - W.A. Mozart-Serenade καθώς και στο spectral contrast. Στους mfcc και spectral flatness η υψηλότερη τιμή παρατηρείται στο Jazz - Frank Sinatra-Fly Me To The Moon. Στο chroma stft η υψηλότερη τιμή παρατηρείται στο Metal - Aerosmith-Dream On. Στο zero crossing η υψηλότερη τιμή παρατηρείται στο Rock - Bon Jovi-Always.

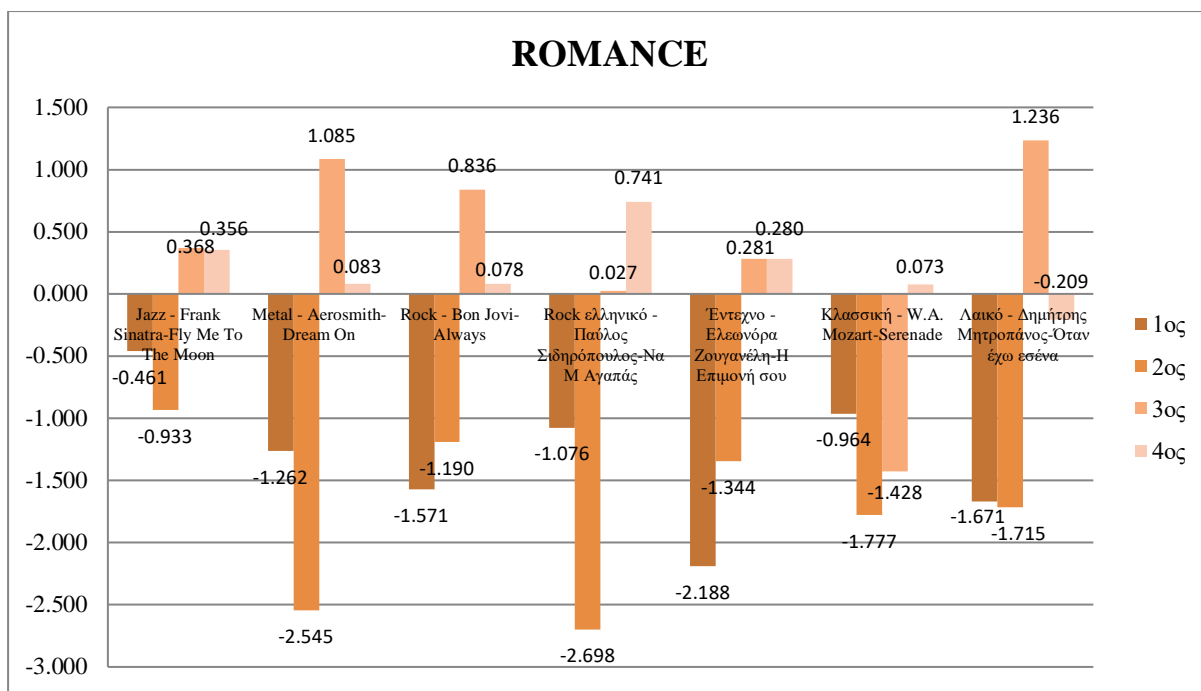
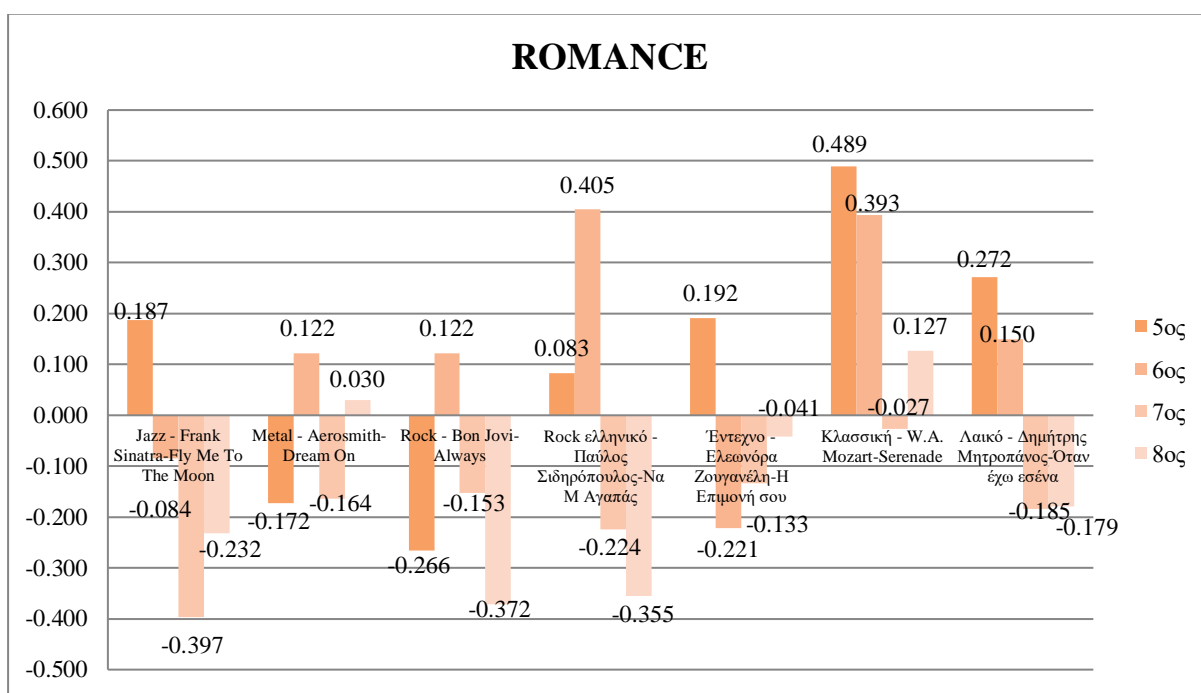


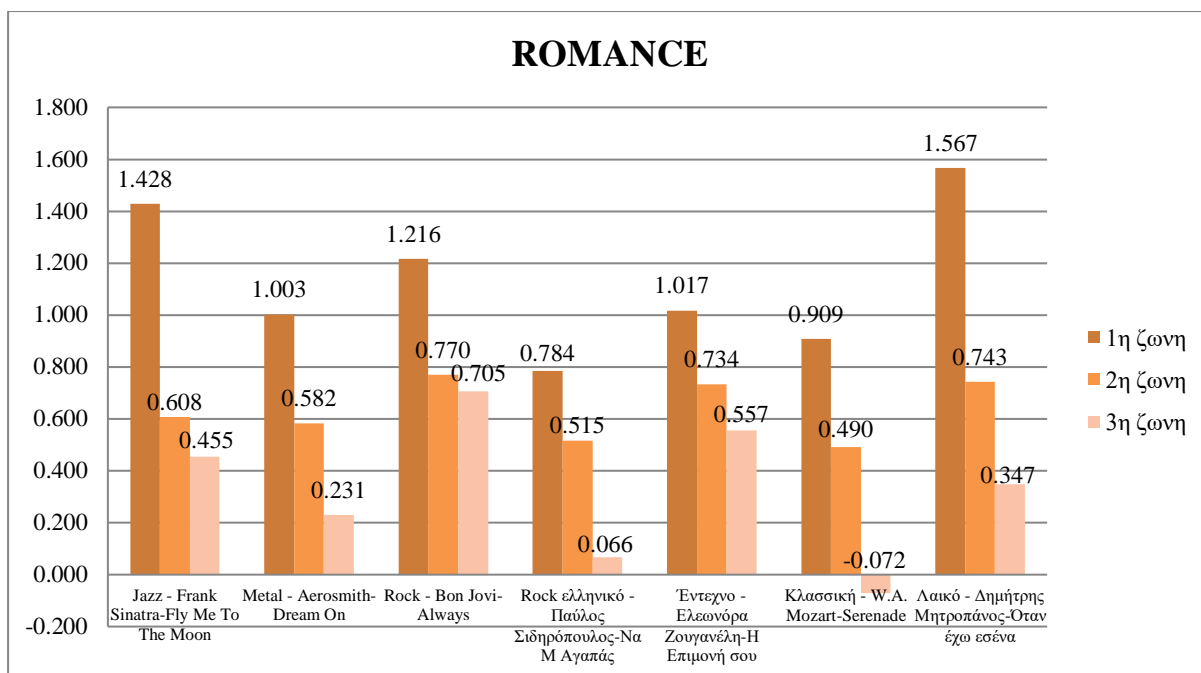
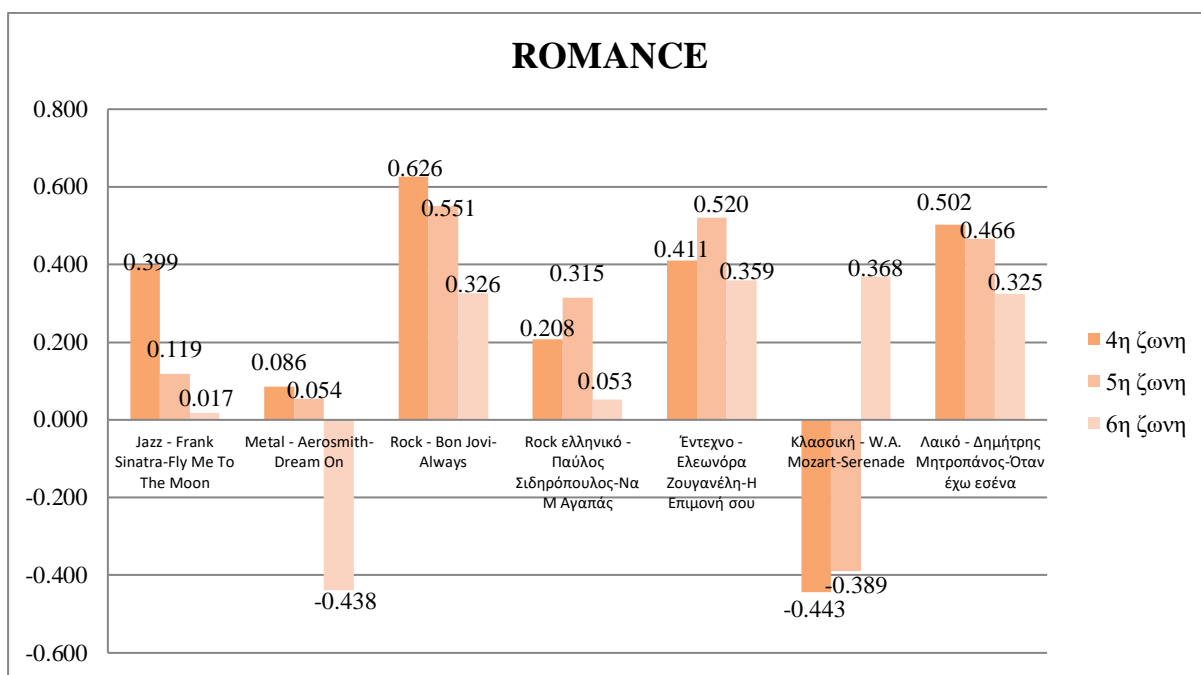
Εικόνα 4.90: Κύρτωση των χαρακτηριστικών για την κατηγορία romance



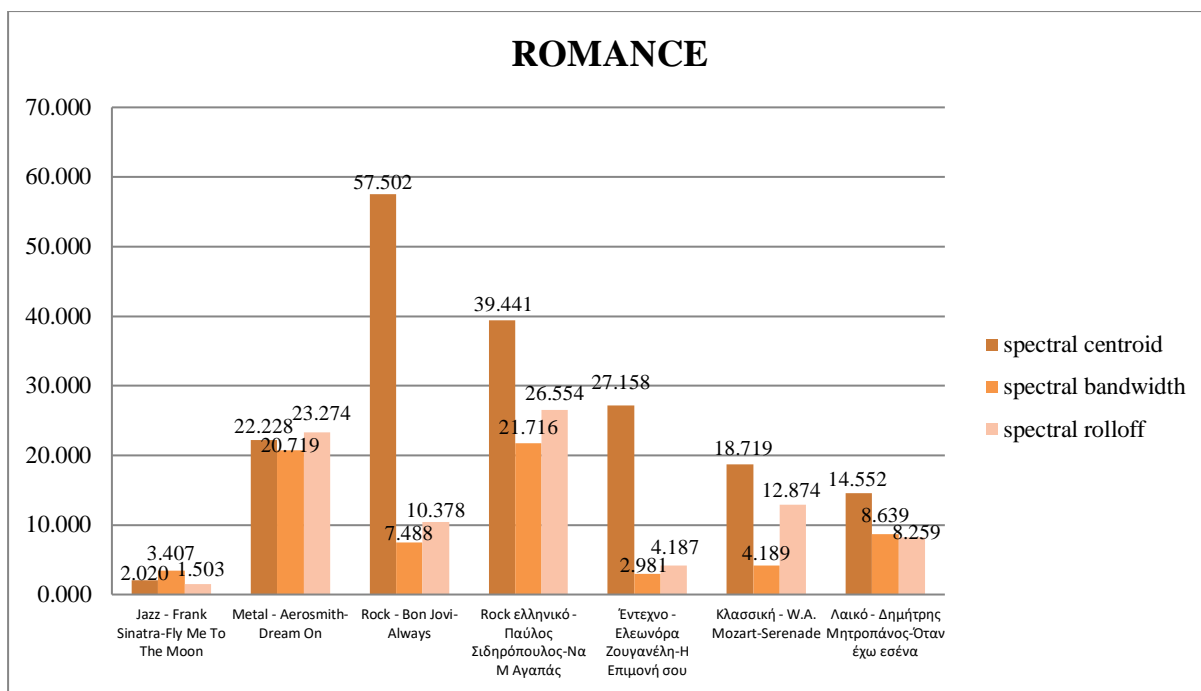
Εικόνα 4.91: Κύρτωση των χαρακτηριστικών για την κατηγορία romance

Από τα διαγράμματα παρατηρούμε ότι οι τιμές των επτά κομματιών για την κύρτωση σε όλα τα χαρακτηριστικά είναι σχετικά κοντά. Η υψηλότερη τιμή στο spectral centroid παρατηρείται στο Rock ελληνικό - Παύλος Σιδηρόπουλος-Να Μ'Αγαπάς καθώς και στα spectral bandwidth και spectral rolloff. Στα rms, zero crossing και spectral flatness η υψηλότερη τιμή παρατηρείται στο Κλασσική - W.A. Mozart-Serenade ενώ στο chroma stft παρατηρείται στο Rock - Bon Jovi-Always.

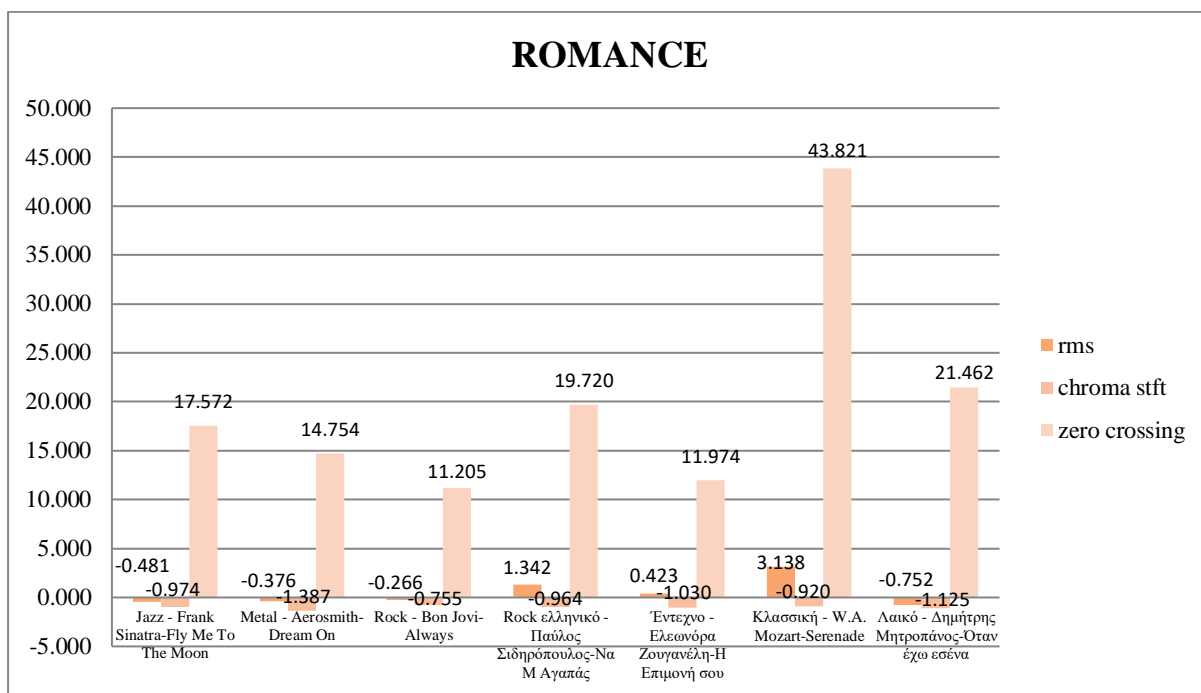
Εικόνα 4.92: Κύρτωση των MFCC (1^{ος} έως 4^{ος}) για την κατηγορία romanceΕικόνα 4.93: Κύρτωση των MFCC (5^{ος} έως 8^{ος}) για την κατηγορία romance

Εικόνα 4.94: Κύρτωση του spectral contrast (1^η έως 3^η ζώνη) για την κατηγορία romanceΕικόνα 4.95: Κύρτωση του spectral contrast (4^η έως 6^η ζώνη) για την κατηγορία romance

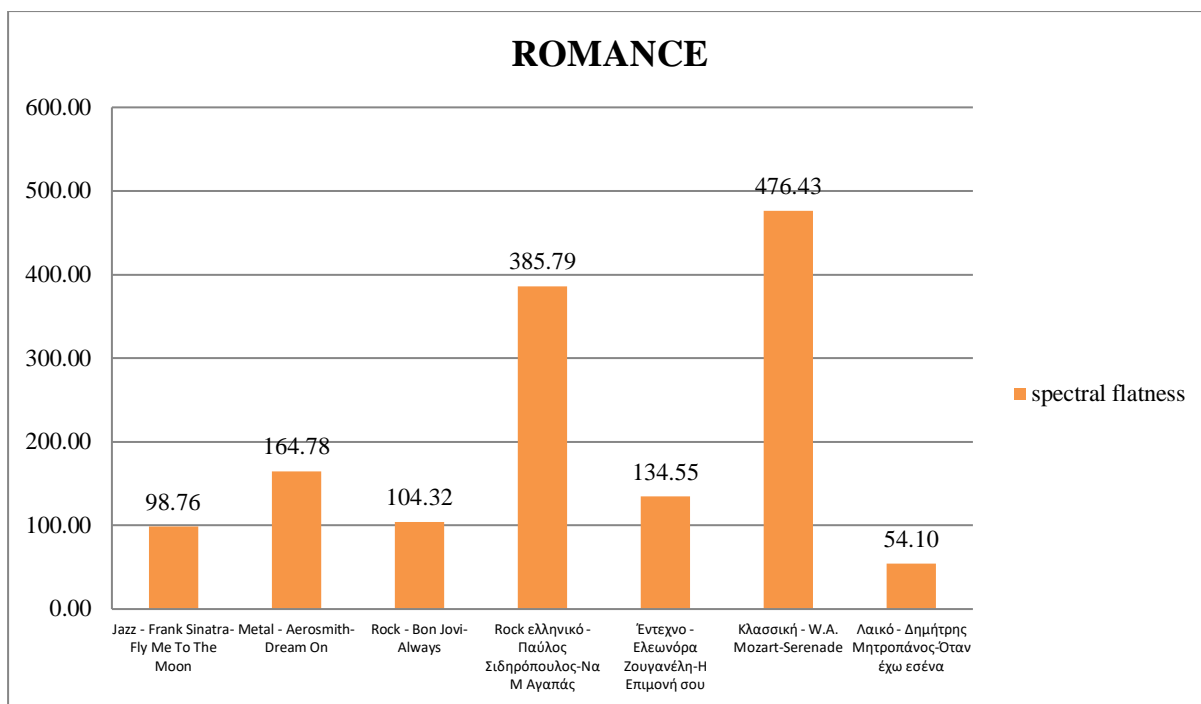
Όσον αφορά τους mfcc παρατηρούνται περισσότερες αρνητικές τιμές. Στον 1^ο mfcc η υψηλότερη τιμή εμφανίζεται στο Έντεχνο - Ελεωνόρα Ζουγανέλη-Η Επιμονή σου, ενώ στον 2^ο 4ο 6ο εμφανίζεται στο Rock ελληνικό - Παύλος Σιδηρόπουλος-Να Μ Αγαπάς. Στον 3^ο mfcc η υψηλότερη τιμή εμφανίζεται στο Λαϊκό - Δημήτρης Μητροπάνος-Όταν έχω εσένα. Στον 5^ο mfcc η υψηλότερη τιμή εμφανίζεται στο Κλασσική - W.A. Mozart-Serenade. Στον 7^ο και 8^ο mfcc η υψηλότερη τιμή εμφανίζεται στο Jazz - Frank Sinatra-Fly Me To The Moon. Στην περίπτωση του spectral contrast, στην 1^η ζώνη η υψηλότερη τιμή παρουσιάζεται στο Λαϊκό - Δημήτρης Μητροπάνος-Όταν έχω εσένα, ενώ στην 2^η, 3^η, 4^η και 5^η εμφανίζεται στο Rock - Bon Jovi-Always. Στην 6^η ζώνη η υψηλότερη τιμή παρουσιάζεται στο Έντεχνο - Ελεωνόρα Ζουγανέλη-Η Επιμονή σου.



Εικόνα 4.96: Λοξότητα των χαρακτηριστικών για την κατηγορία romance

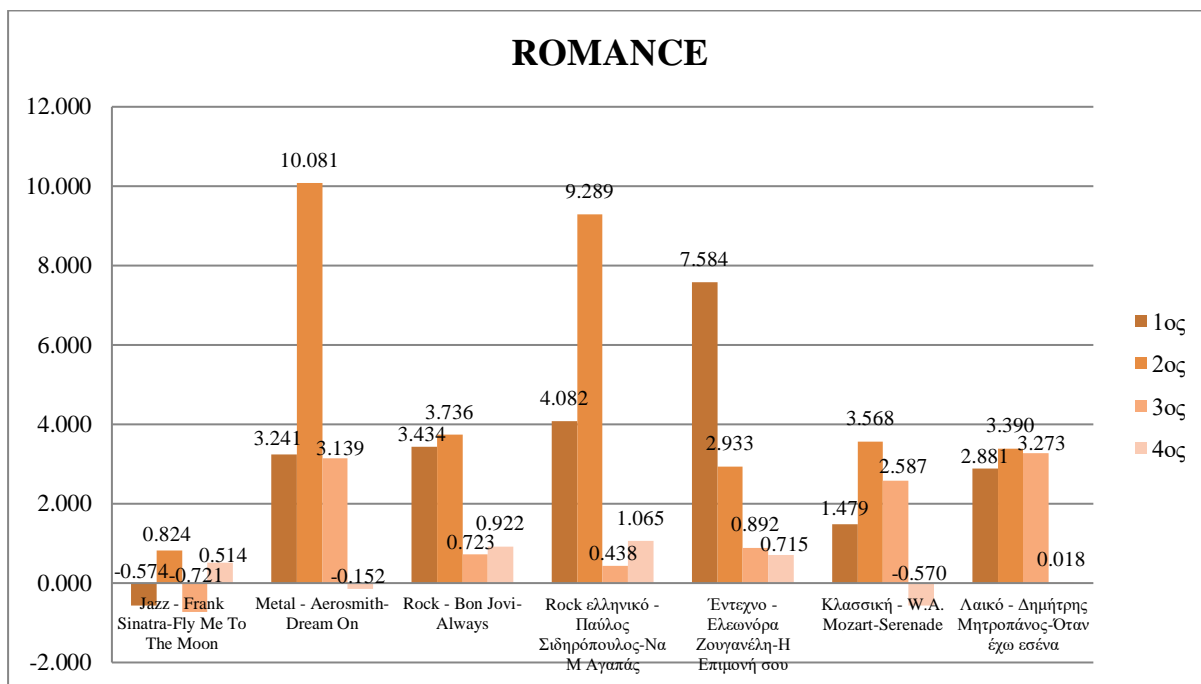


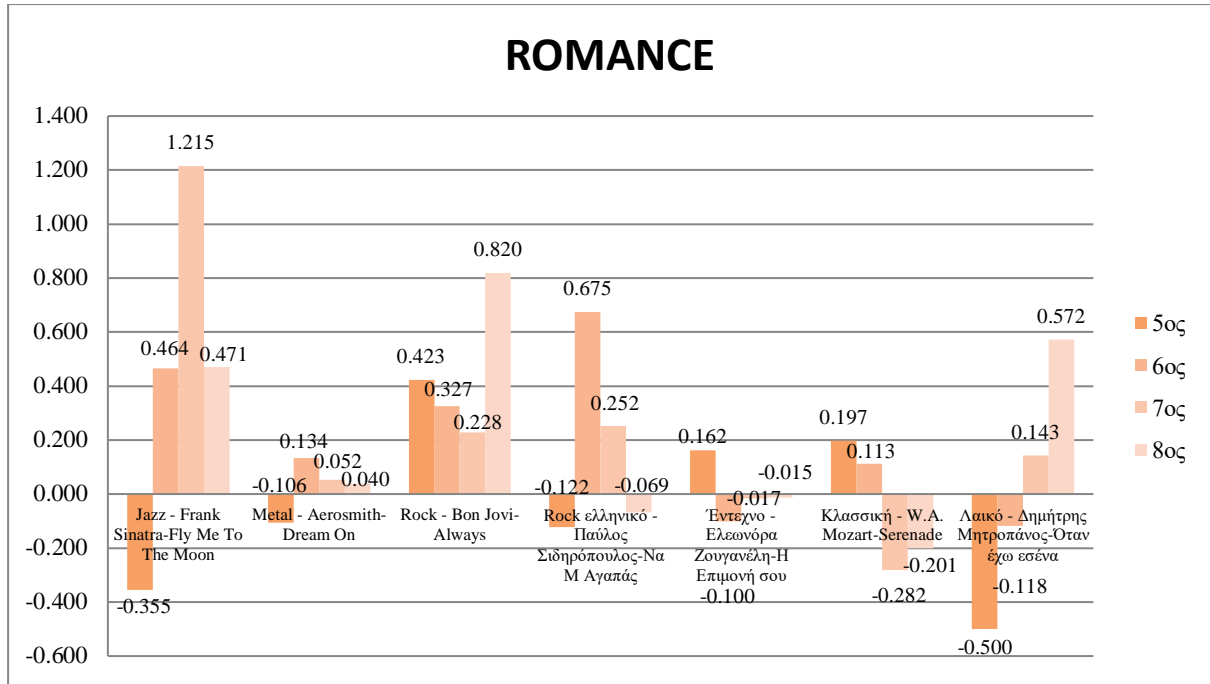
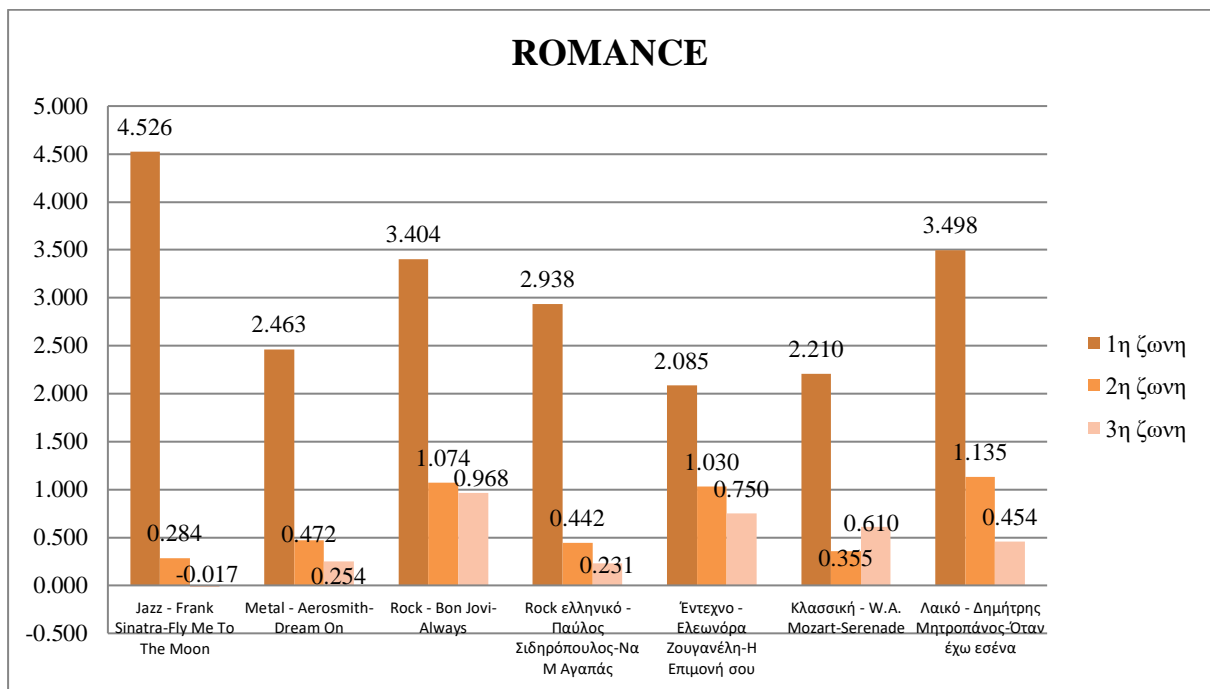
Εικόνα 4.97: Λοξότητα των χαρακτηριστικών για την κατηγορία romance

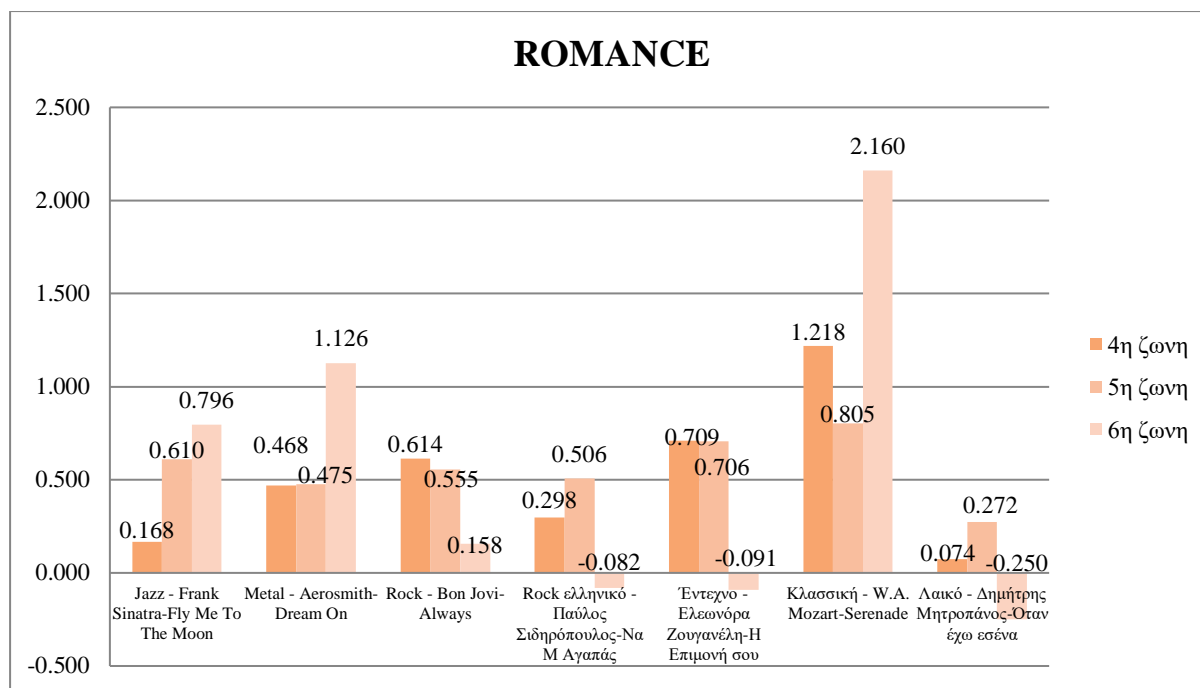


Εικόνα 4.98: Λοξότητα του spectral flatness για την κατηγορία romance

Από τα διαγράμματα παρατηρούμε ότι οι τιμές των επτά κομματιών για την λοξότητα σε όλα τα χαρακτηριστικά είναι σχετικά κοντά. Η υψηλότερη τιμή στο spectral centroid παρατηρείται στο Rock - Bon Jovi - Always ενώ στα spectral bandwidth και spectral rolloff παρατηρείται στο Rock ελληνικό - Παύλος Σιδηρόπουλος - Να Μ' Αγαπάς. Στα rms, zero crossing και spectral flatness η υψηλότερη τιμή παρατηρείται στο Κλασσική - W.A. Mozart - Serenade ενώ στο chroma stft παρατηρείται στο Metal - Aerosmith - Dream On.

Εικόνα 4.99: Λοξότητα των MFCC (1^{ος} έως 4^{ος}) για την κατηγορία romance

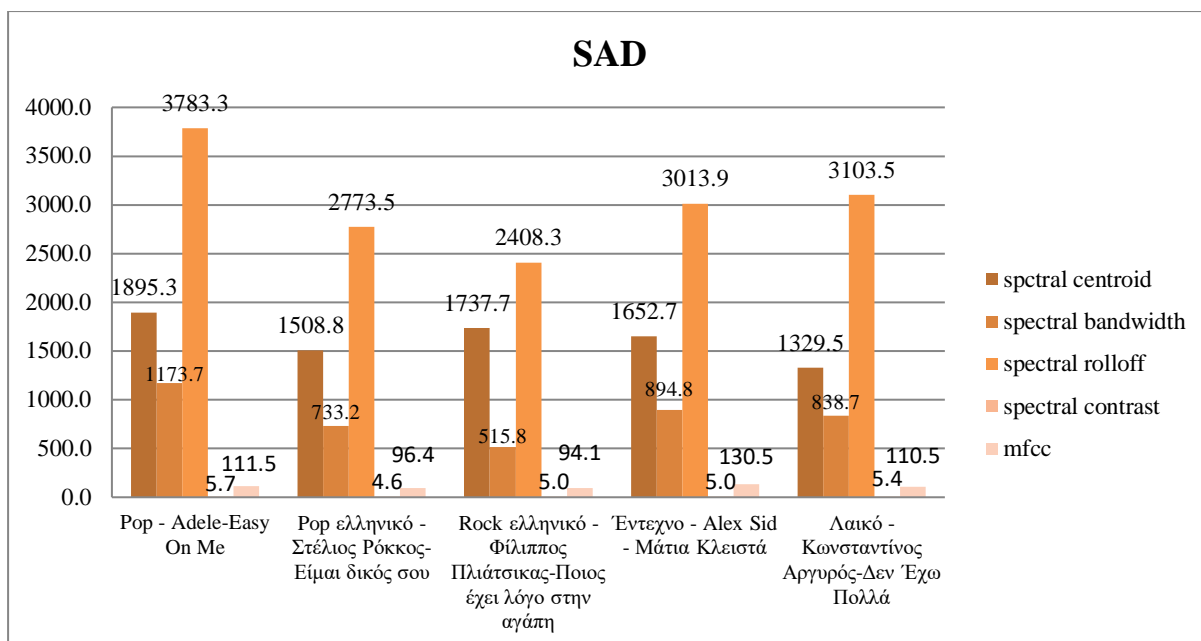
Εικόνα 4.100: Λοξότητα των MFCC (5^{ος} έως 8^{ος}) για την κατηγορία romanceΕικόνα 4.101: Λοξότητα του spectral contrast (1^η έως 3^η ζώνη) για την κατηγορία romance



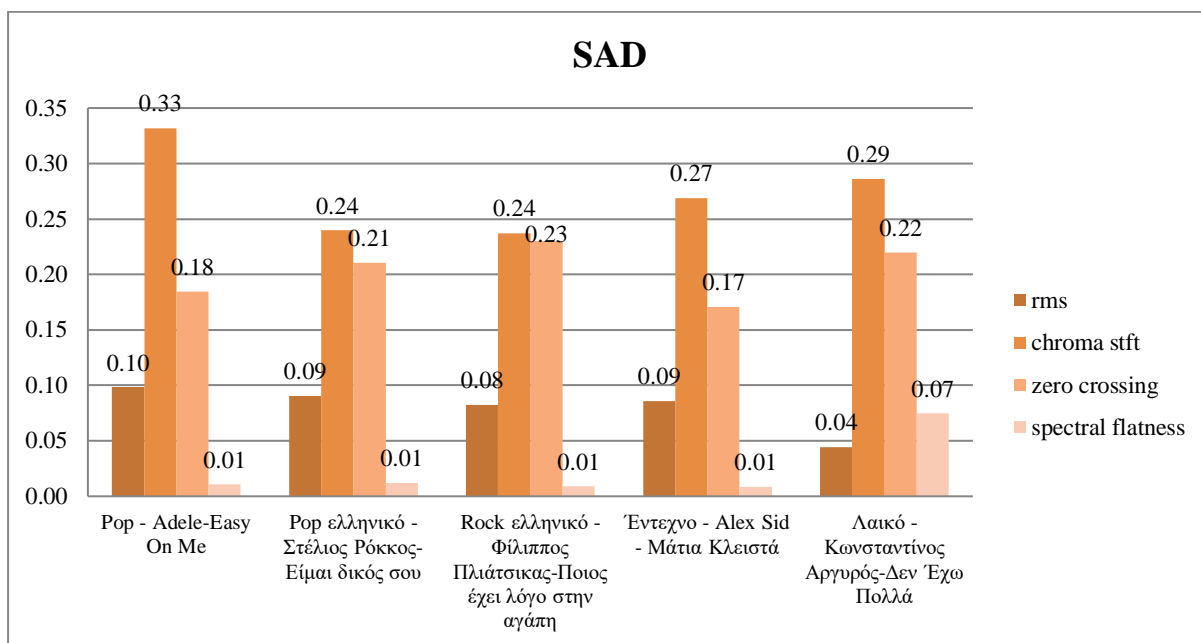
Εικόνα 4.102: Λοξότητα του spectral contrast (4^η έως 6^η ζώνη) για την κατηγορία romance

Όσον αφορά τους mfcc παρατηρούνται περισσότερες αρνητικές τιμές. Στον 1^ο mfcc η υψηλότερη τιμή εμφανίζεται στο Έντεχνο - Ελεωνόρα Ζουγανέλη-Η Επιμονή σου, ενώ στον 2^ο εμφανίζεται στο Metal - Aerosmith-Dream On. Στον 3^ο mfcc η υψηλότερη τιμή εμφανίζεται στο Λαϊκό - Δημήτρης Μητροπάνος-Όταν έχω εσένα. Στον 4^ο και 6^ο mfcc η υψηλότερη τιμή εμφανίζεται στο Rock ελληνικό - Παύλος Σιδηρόπουλος-Να Μ Αγαπάς. Στον 5^ο και 8^ο mfcc η υψηλότερη τιμή εμφανίζεται στο Rock - Bon Jovi-Always. Στον 7^ο mfcc η υψηλότερη τιμή εμφανίζεται στο Jazz - Frank Sinatra-Fly Me To The Moon. Στην περίπτωση του spectral contrast, στην 1^η ζώνη η υψηλότερη τιμή παρουσιάζεται στο Jazz - Frank Sinatra-Fly Me To The Moon, ενώ στην 2^η εμφανίζεται στο Λαϊκό - Δημήτρης Μητροπάνος-Όταν έχω εσένα. Στην 3^η ζώνη η υψηλότερη τιμή παρουσιάζεται στο Rock - Bon Jovi-Always. Στην 4^η, 5^η και 6^η ζώνη η υψηλότερη τιμή παρουσιάζεται στο Κλασσική - W.A. Mozart-Serenade.

Στις παρακάτω εικόνες παρουσιάζονται τα γραφήματα για το συναίσθημα sad στα πέντε στατιστικά στοιχεία.

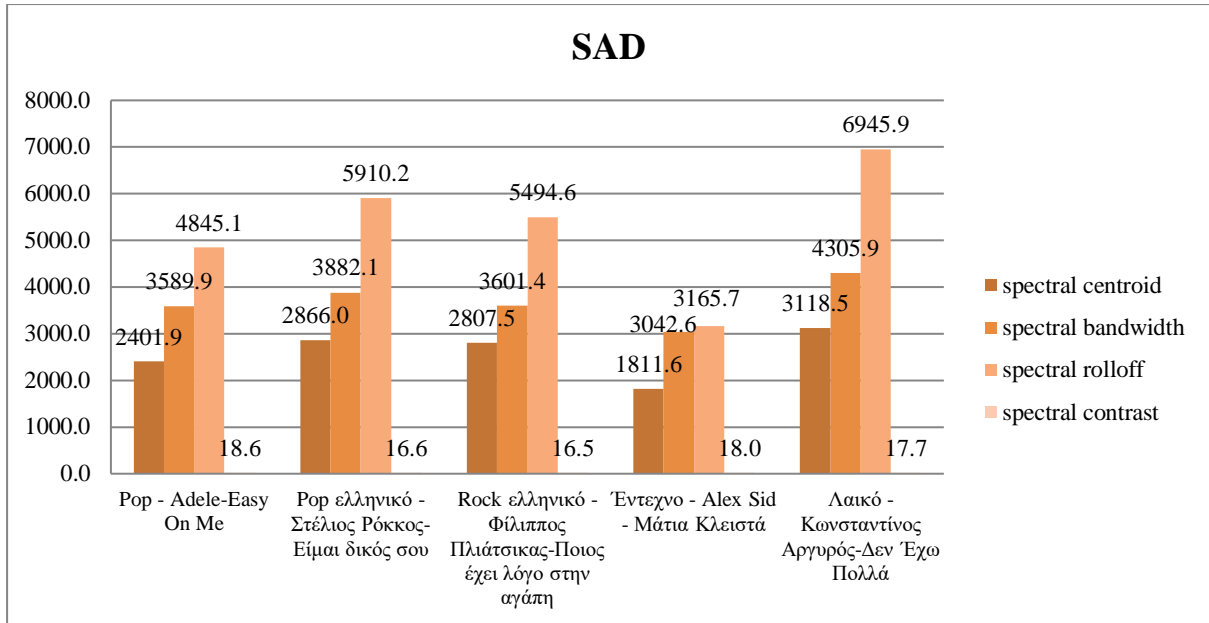


Εικόνα 4.103: Τυπική απόκλιση των χαρακτηριστικών για την κατηγορία sad

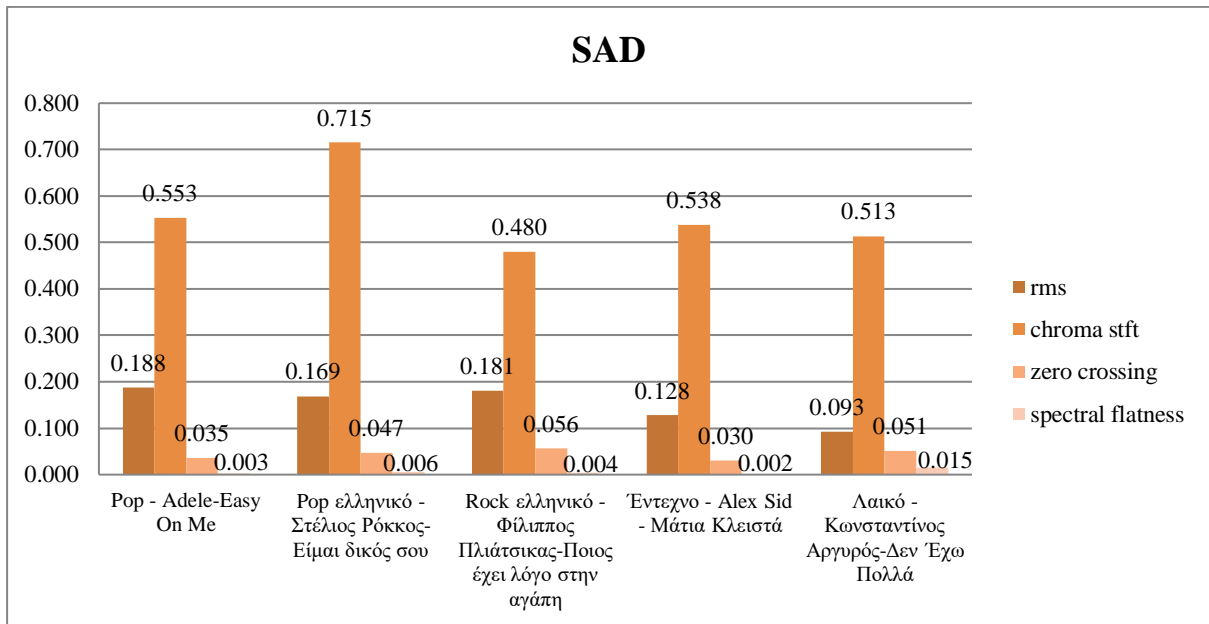


Εικόνα 4.104: Τυπική απόκλιση των χαρακτηριστικών για την κατηγορία sad

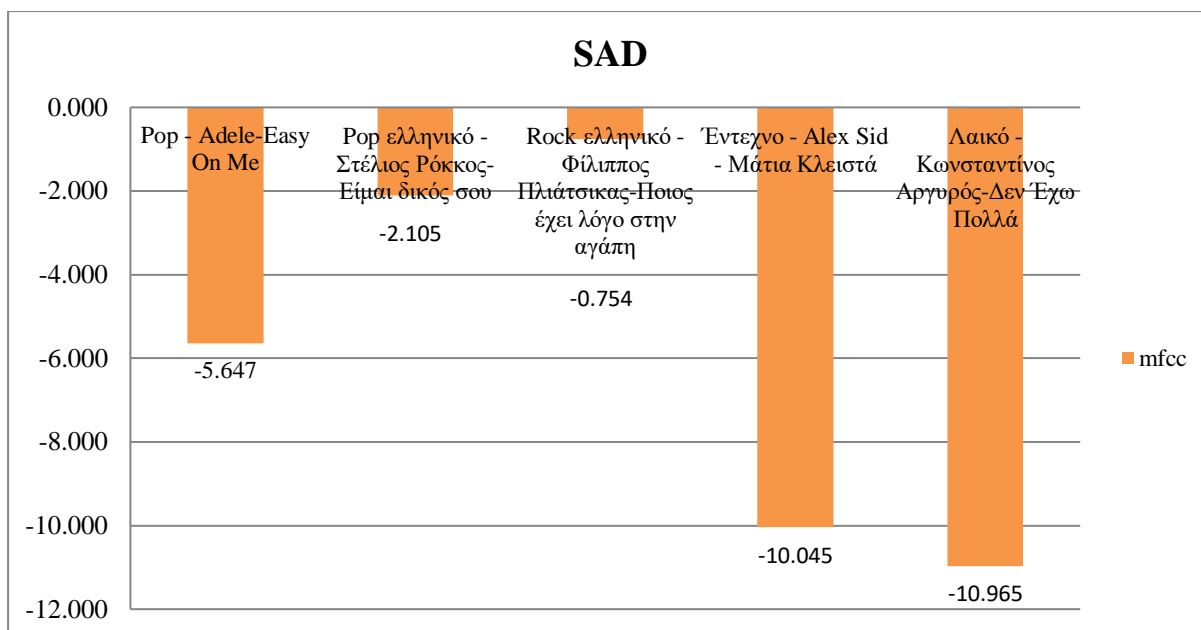
Από τα διαγράμματα παρατηρούμε ότι οι τιμές των πέντε κομματιών για την τυπική απόκλιση σε όλα τα χαρακτηριστικά είναι πολύ κοντά. Η υψηλότερη τιμή στο spectral centroid παρατηρείται στο Pop - Adele-Easy On Me καθώς και στα spectral bandwidth, spectral rolloff, spectral contrast, rms και chroma stft. Στους mfcc η υψηλότερη τιμή παρατηρείται στο Έντεχνο - Alex Sid - Μάτια Κλειστά. Στο zero crossing η υψηλότερη τιμή παρατηρείται στο Rock ελληνικό - Φίλιππος Πλιάτσικας-Ποιος έχει λόγο στην αγάπη ενώ στο spectral flatness η υψηλότερη τιμή παρατηρείται στο Λαϊκό - Κωνσταντίνος Αργυρός-Δεν Έχω Πολλά.



Εικόνα 4.105: Μέση Τιμή των χαρακτηριστικών για την κατηγορία sad

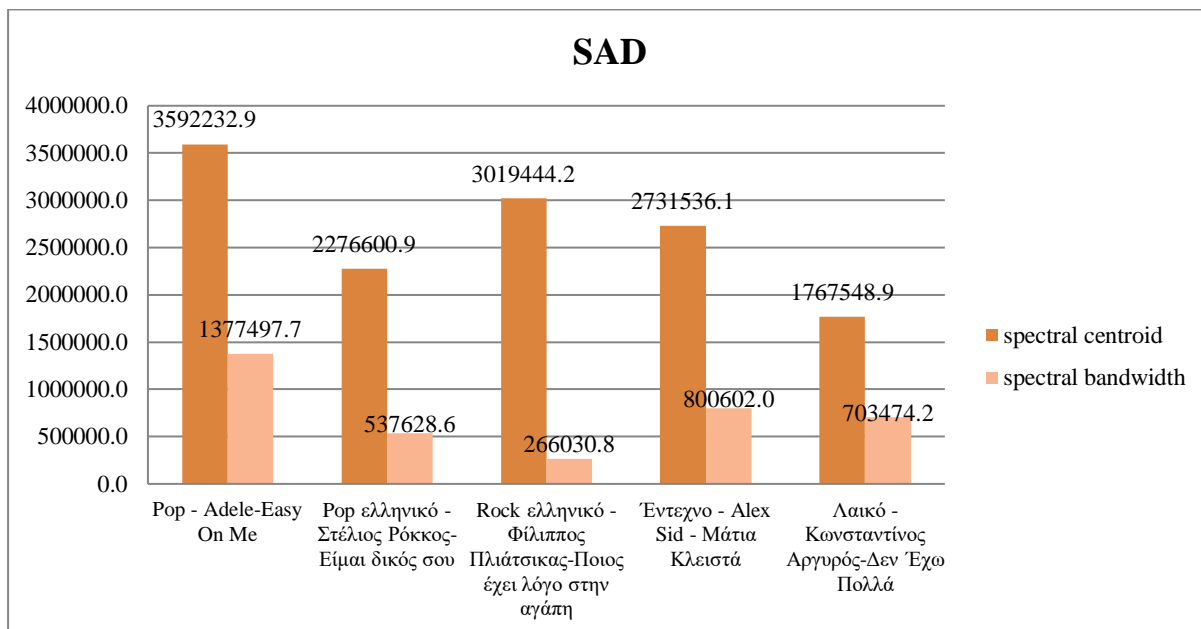


Εικόνα 4.106: Μέση Τιμή των χαρακτηριστικών για την κατηγορία sad

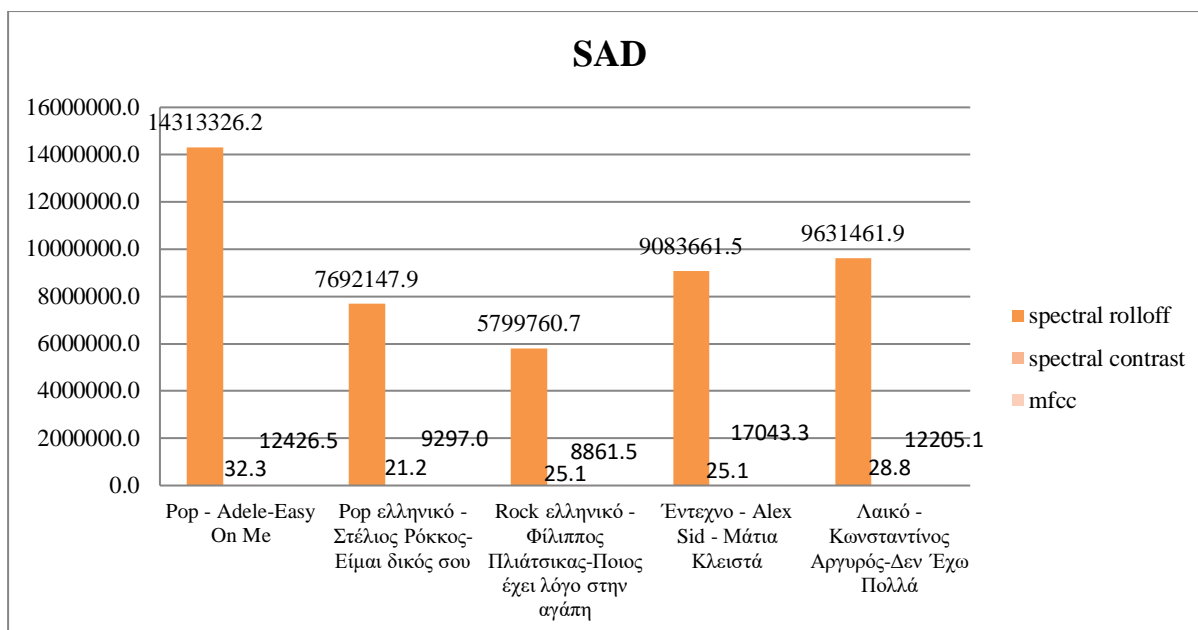


Εικόνα 4.107: Μέση Τιμή των MFCC για την κατηγορία sad

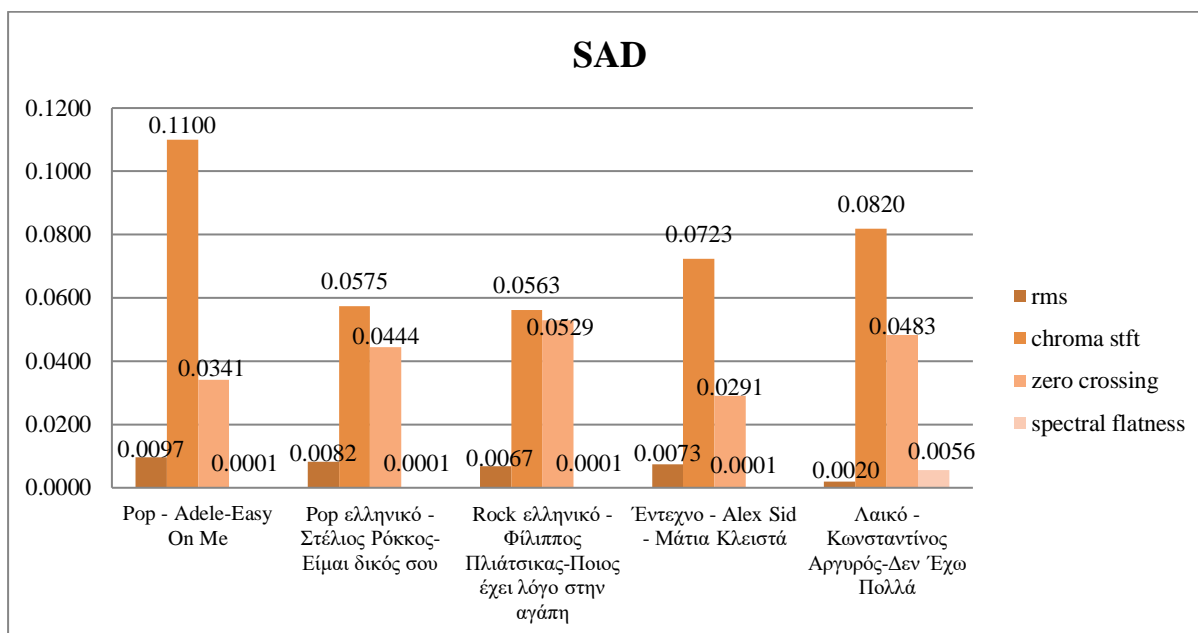
Από τα διαγράμματα παρατηρούμε ότι οι τιμές των πέντε κομματιών για την μέση τιμή σε όλα τα χαρακτηριστικά είναι πολύ κοντά. Η υψηλότερη τιμή στο spectral centroid παρατηρείται στο Λαϊκό - Κωνσταντίνος Αργυρός-Δεν Έχω Πολλά καθώς και στα spectral bandwidth και spectral rolloff, mfcc και spectral flatness. Στα spectral contrast, rms και chroma stft η υψηλότερη τιμή παρατηρείται στο Pop - Adele-Easy On Me. Στο zero crossing η υψηλότερη τιμή παρατηρείται στο Rock ελληνικό - Φίλιππος Πλιάτσικας-Ποιος έχει λόγο στην αγάπη.



Εικόνα 4.108: Διακύμανση των χαρακτηριστικών για την κατηγορία sad

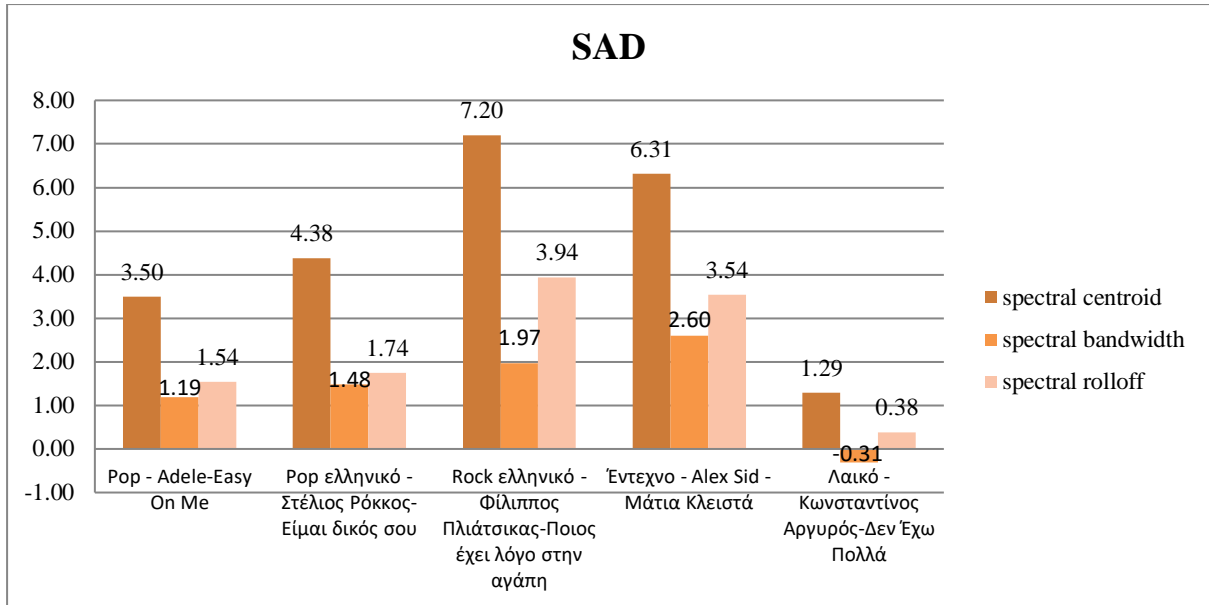


Εικόνα 4.109: Διακύμανση των χαρακτηριστικών για την κατηγορία sad

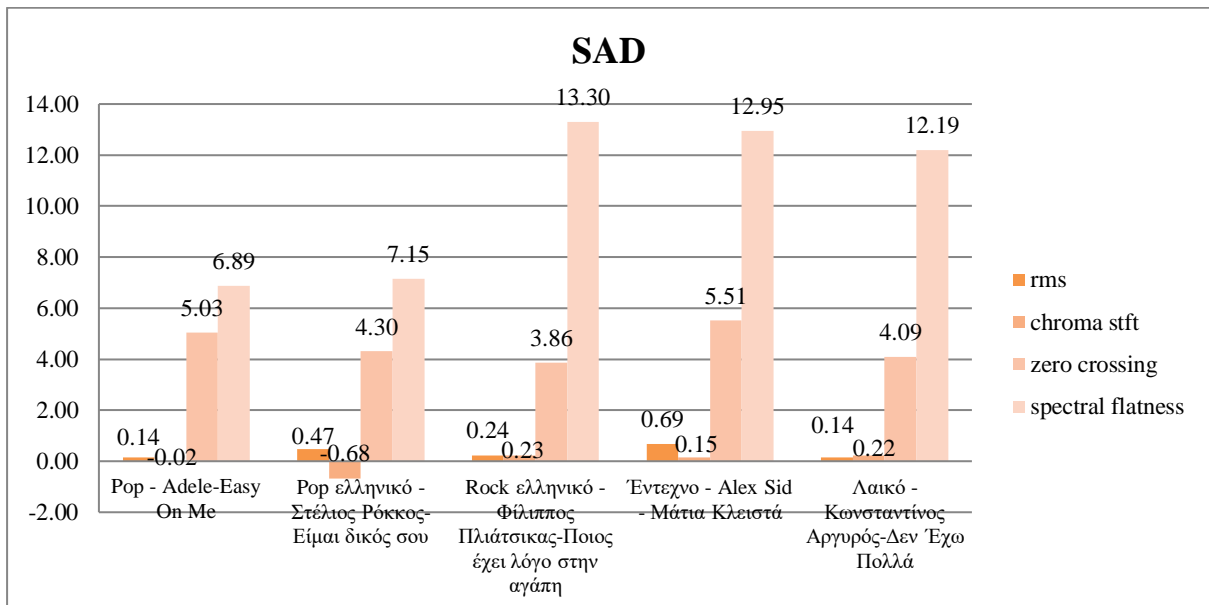


Εικόνα 4.110: Διακύμανση των χαρακτηριστικών για την κατηγορία sad

Από τα διαγράμματα παρατηρούμε ότι οι τιμές των πέντε κομματιών για την μέση τιμή σε όλα τα χαρακτηριστικά είναι πολύ κοντά. Η υψηλότερη τιμή στο spectral centroid παρατηρείται στο Pop - Adele-Easy On Me καθώς και στα spectral bandwidth, spectral rolloff, spectral contrast, rms και chroma stft. Για τους mfcc, η υψηλότερη τιμή παρατηρείται στο Έντεχνο - Alex Sid - Μάτια Κλειστά. Στο zero crossing, η υψηλότερη τιμή παρατηρείται στο Rock ελληνικό - Φίλιππος Πλιάτσικας-Ποιος έχει λόγο στην αγάπη ενώ στο spectral flatness η υψηλότερη τιμή εμφανίζεται στο Λαϊκό - Κωνσταντίνος Αργυρός-Δεν Έχω Πολλά.

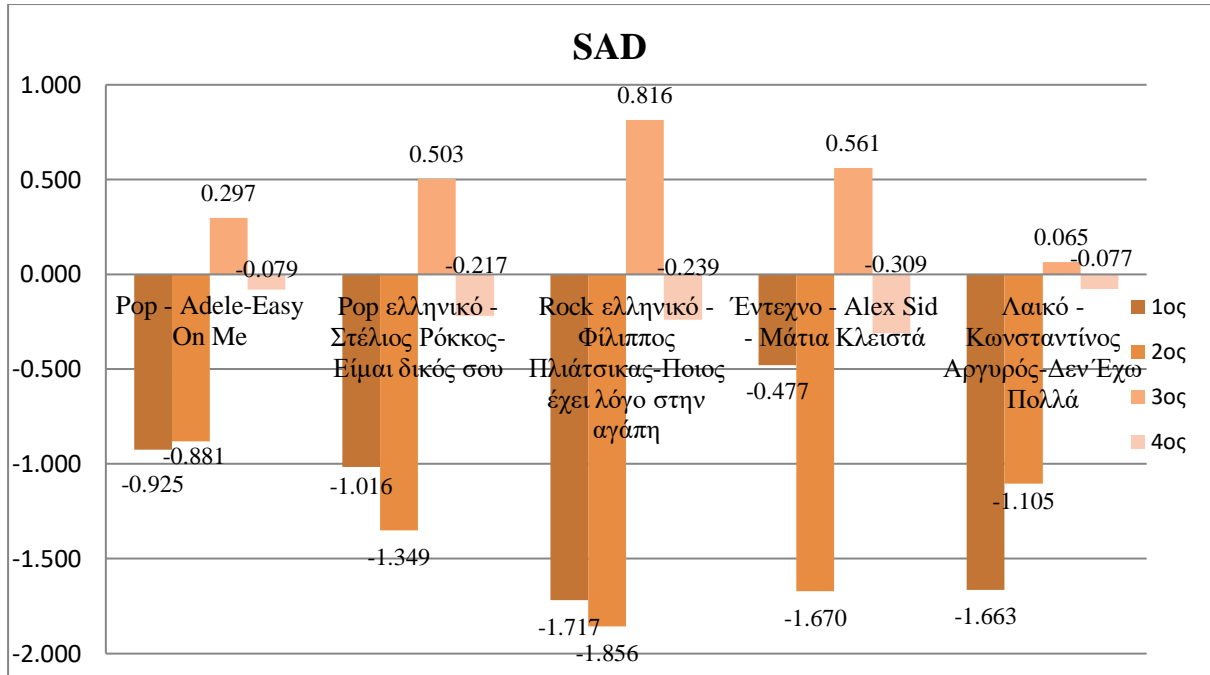
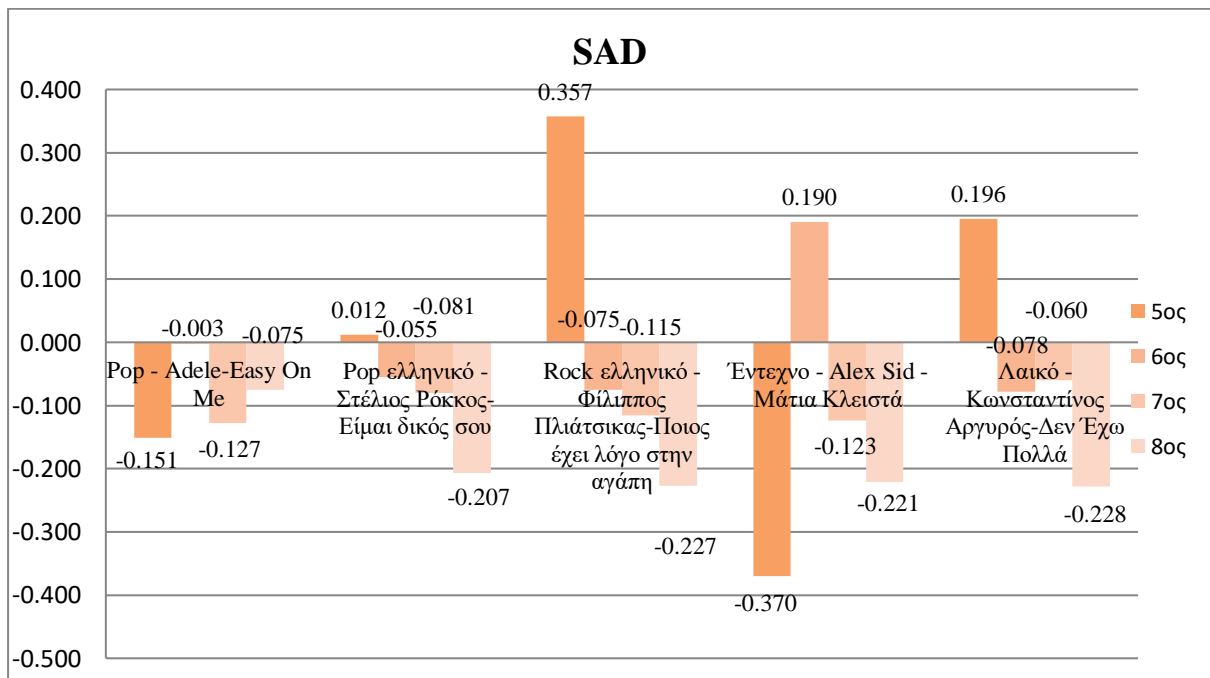


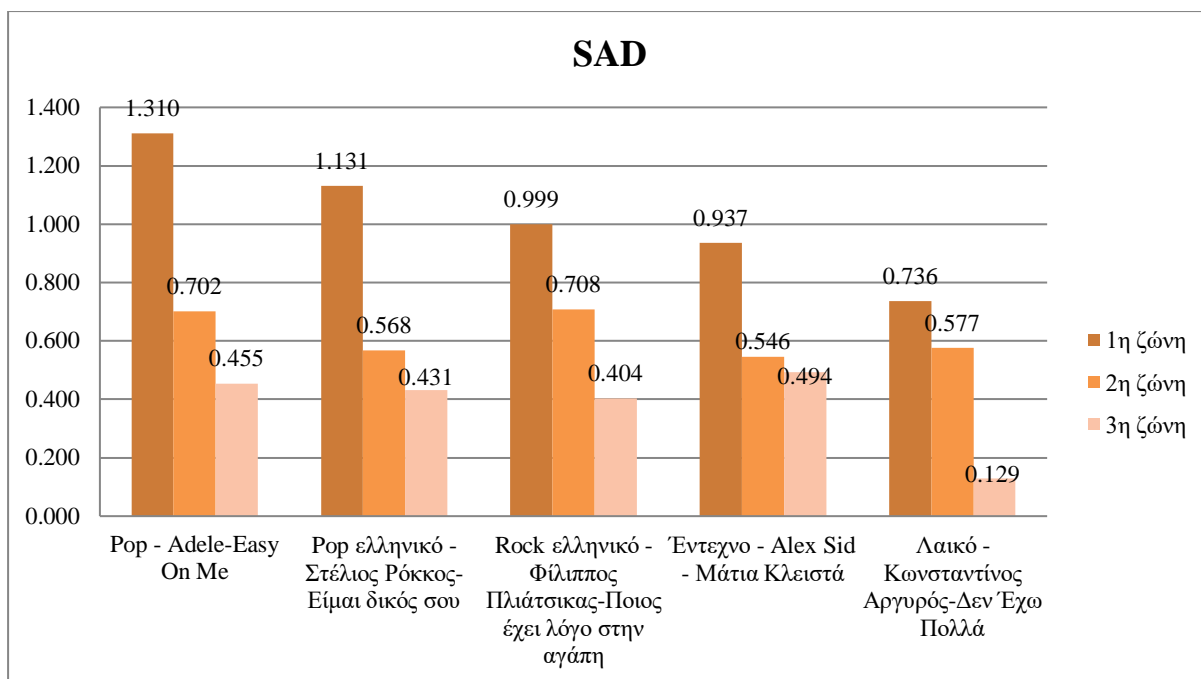
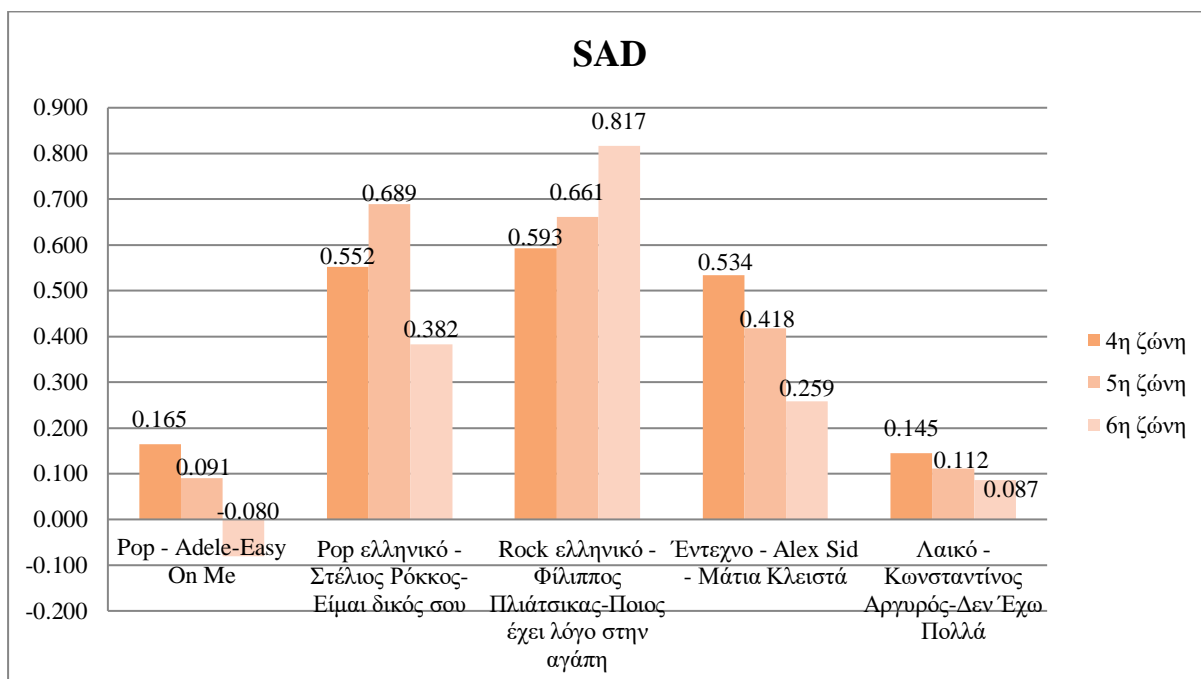
Εικόνα 4.111: Κύρτωση των χαρακτηριστικών για την κατηγορία sad



Εικόνα 4.112: Κύρτωση των χαρακτηριστικών για την κατηγορία sad

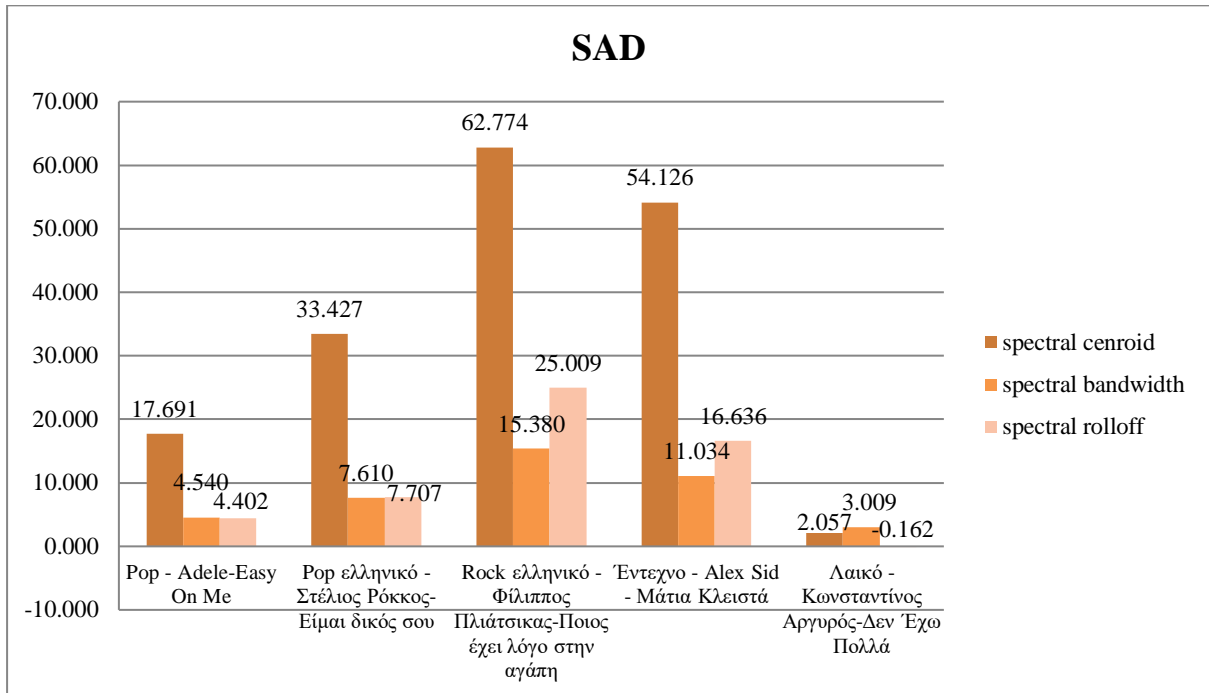
Από τα διαγράμματα παρατηρούμε ότι οι τιμές των πέντε κομματιών για την κύρτωση σε όλα τα χαρακτηριστικά είναι σχετικά κοντά. Η υψηλότερη τιμή στα spectral centroid, spectral rolloff, chroma stft και spectral flatness παρατηρείται στο Rock ελληνικό - Φίλιππος Πλιάτσικας-Ποιος έχει λόγο στην αγάπη ενώ στα spectral bandwidth, rms και zero crossing παρατηρείται στο Έντεχνο - Alex Sid - Μάτια Κλειστά.

Εικόνα 4.113: Κύρτωση των MFCC (1^{ος} έως 4^{ος}) για την κατηγορία sadΕικόνα 4.114: Κύρτωση των MFCC (5^{ος} έως 8^{ος}) για την κατηγορία sad

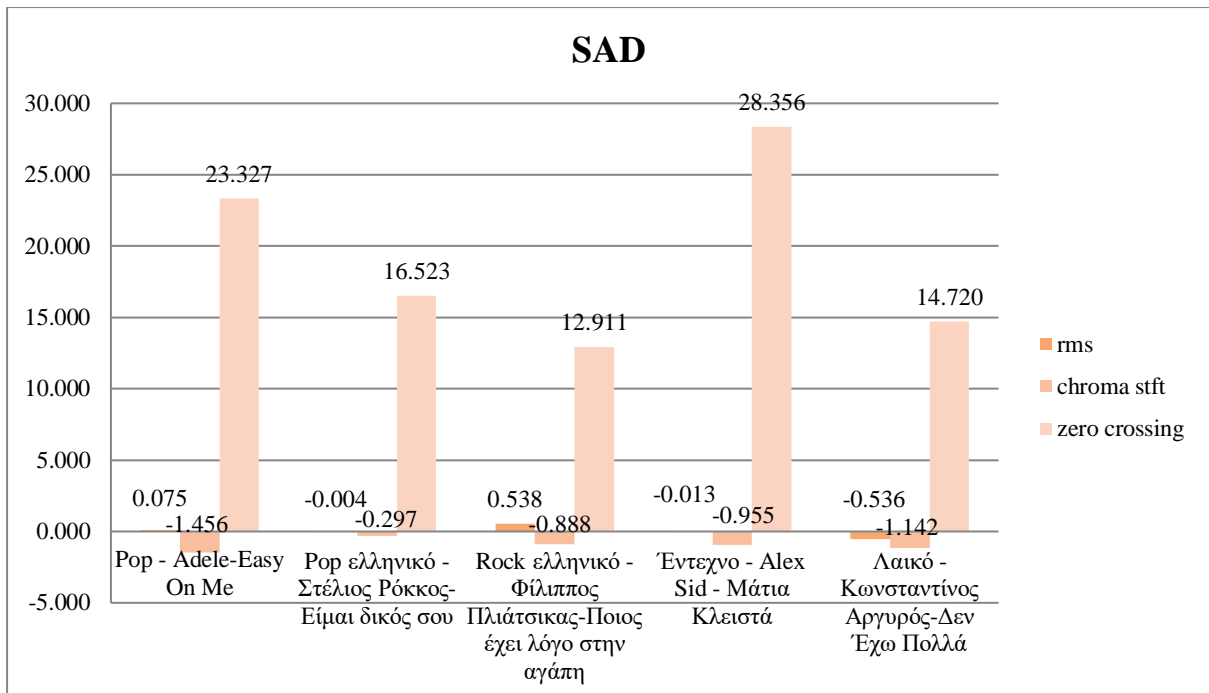
Εικόνα 4.115: Κύρτωση του spectral contrast (1^η έως 3^η ζώνη) για την κατηγορία sadΕικόνα 4.116: Κύρτωση του spectral contrast (4^η έως 6^η ζώνη) για την κατηγορία sad

Όσον αφορά τους mfcc παρατηρούνται περισσότερες αρνητικές τιμές. Στον 1^ο 2^ο 3^ο και 5^ο mfcc η υψηλότερη τιμή εμφανίζεται στο Rock ελληνικό - Φίλιππος Πλιάτσικας-Ποιος έχει λόγο στην αγάπη. Στον 4^ο και 6^ο mfcc η υψηλότερη τιμή εμφανίζεται στο Έντεχνο - Alex Sid - Μάτια Κλειστά. Στον 7^ο mfcc η υψηλότερη τιμή εμφανίζεται στο Pop - Adele-Easy On Me. Στον 8^ο mfcc η υψηλότερη τιμή εμφανίζεται στο Λαϊκό - Κωνσταντίνος Αργυρός-Δεν Έχω Πολλά. Στην περίπτωση του spectral contrast, στην 1^η ζώνη η υψηλότερη τιμή παρουσιάζεται στο Pop - Adele-Easy On Me, ενώ στην 2^η, 4^η και 6^η εμφανίζεται στο Rock ελληνικό - Φίλιππος Πλιάτσικας-Ποιος έχει λόγο στην αγάπη. Στην 3^η

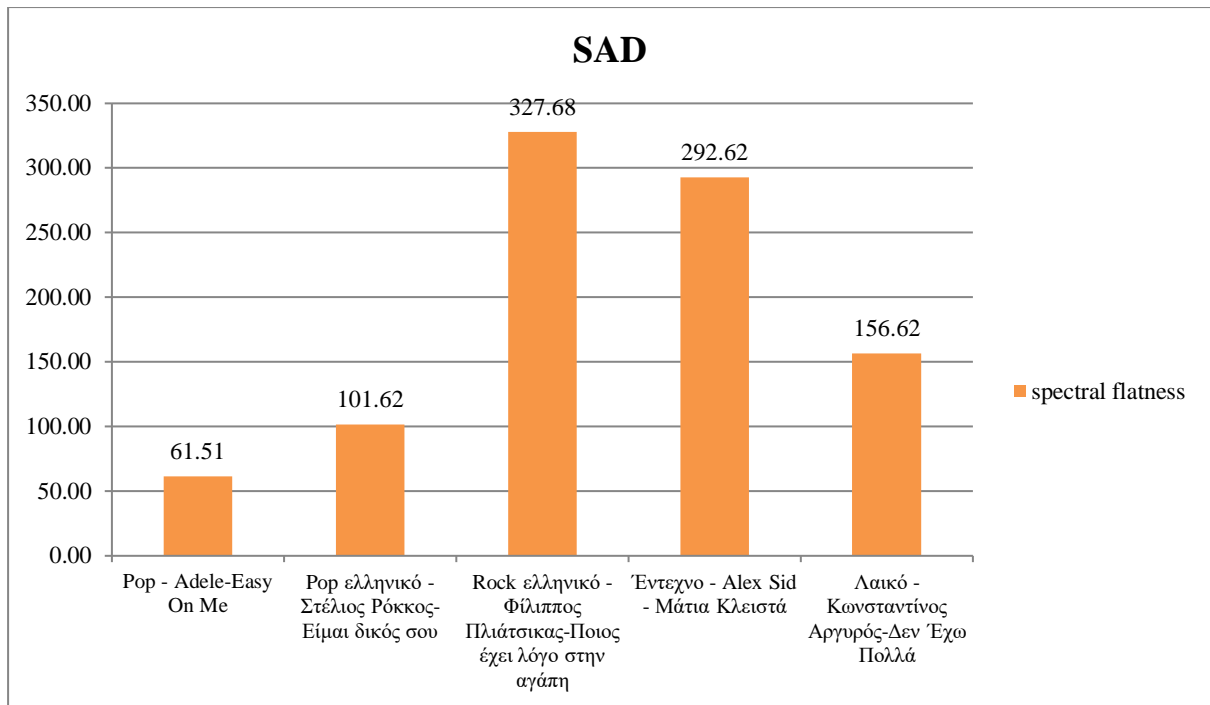
ζώνη η υψηλότερη τιμή παρουσιάζεται στο Έντεχνο - Alex Sid - Μάτια Κλειστά. Στην 5^η ζώνη η υψηλότερη τιμή παρουσιάζεται στο Pop ελληνικό - Στέλιος Ρόκκος-Είμαι δικός σου.



Εικόνα 4.117: Λοξότητα των χαρακτηριστικών για την κατηγορία sad

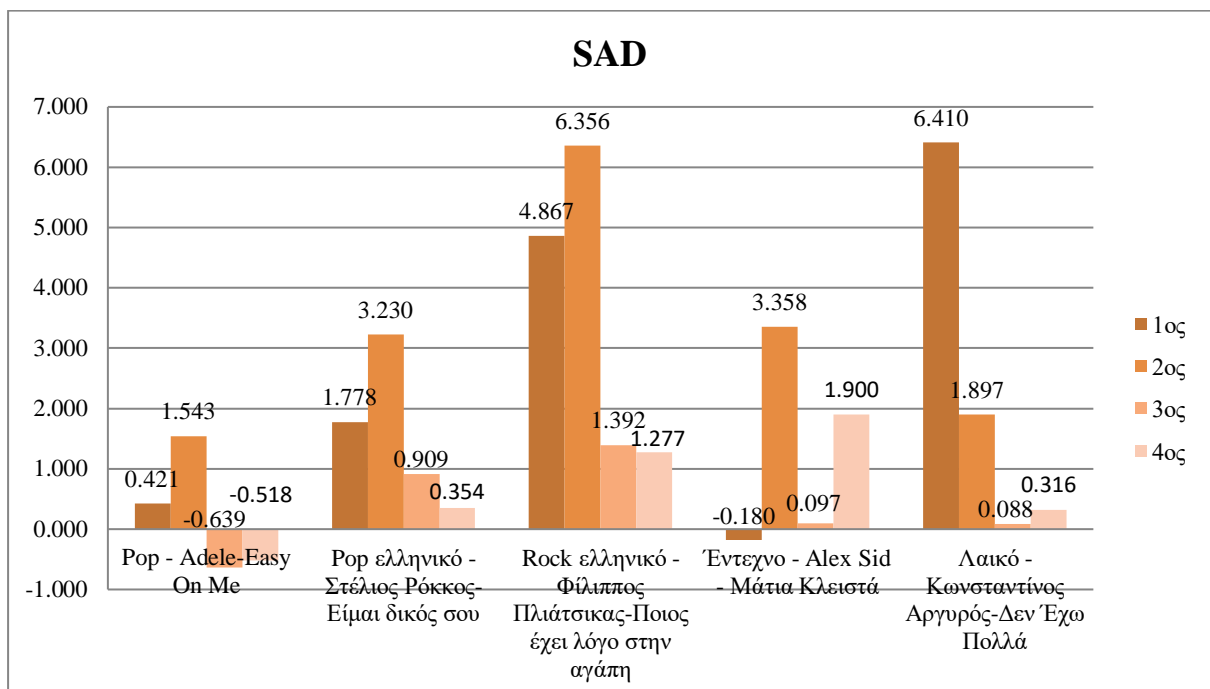


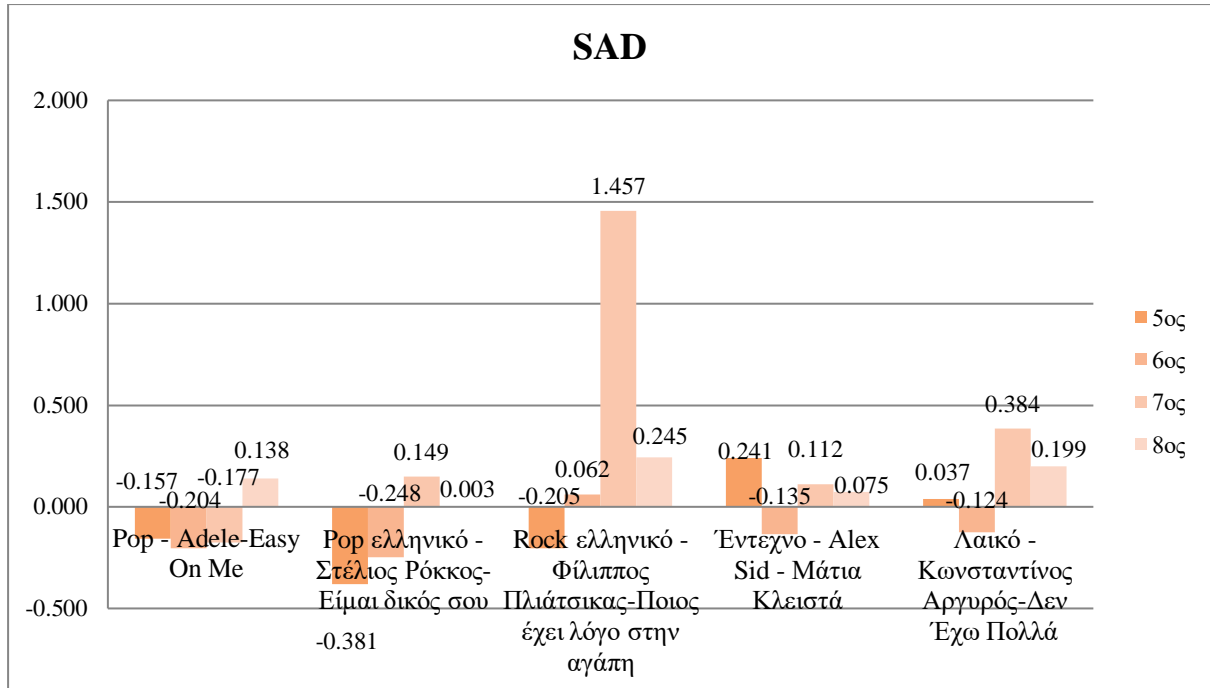
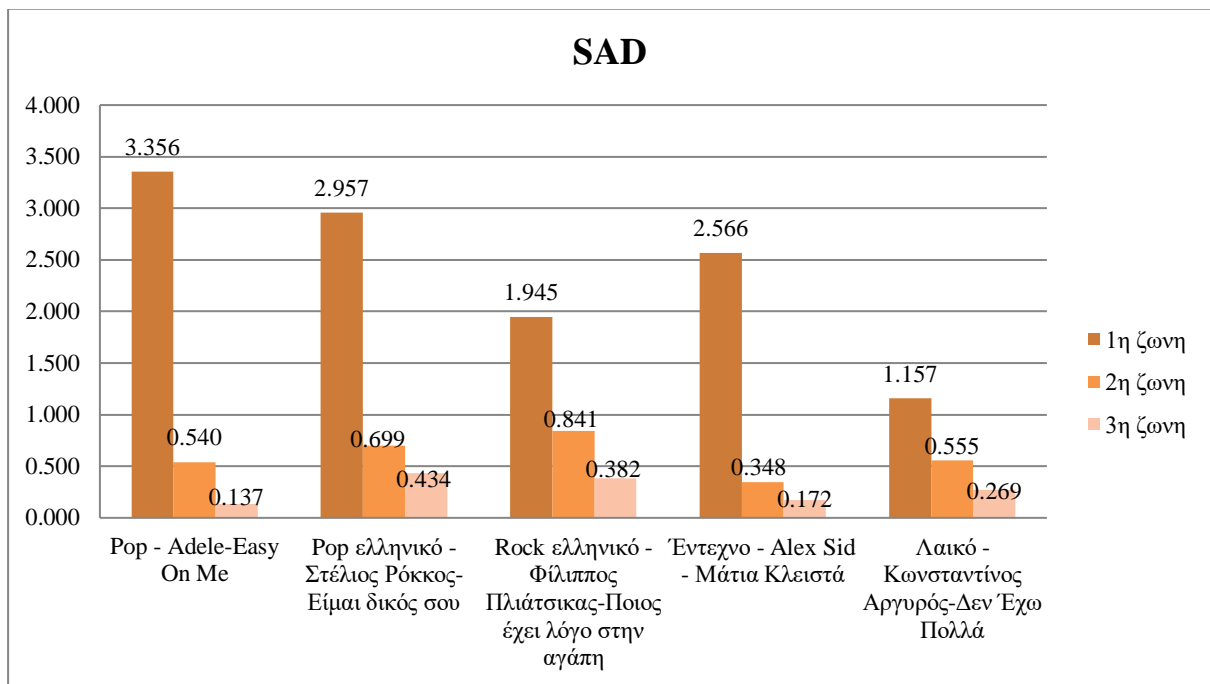
Εικόνα 4.118: Λοξότητα των χαρακτηριστικών για την κατηγορία sad

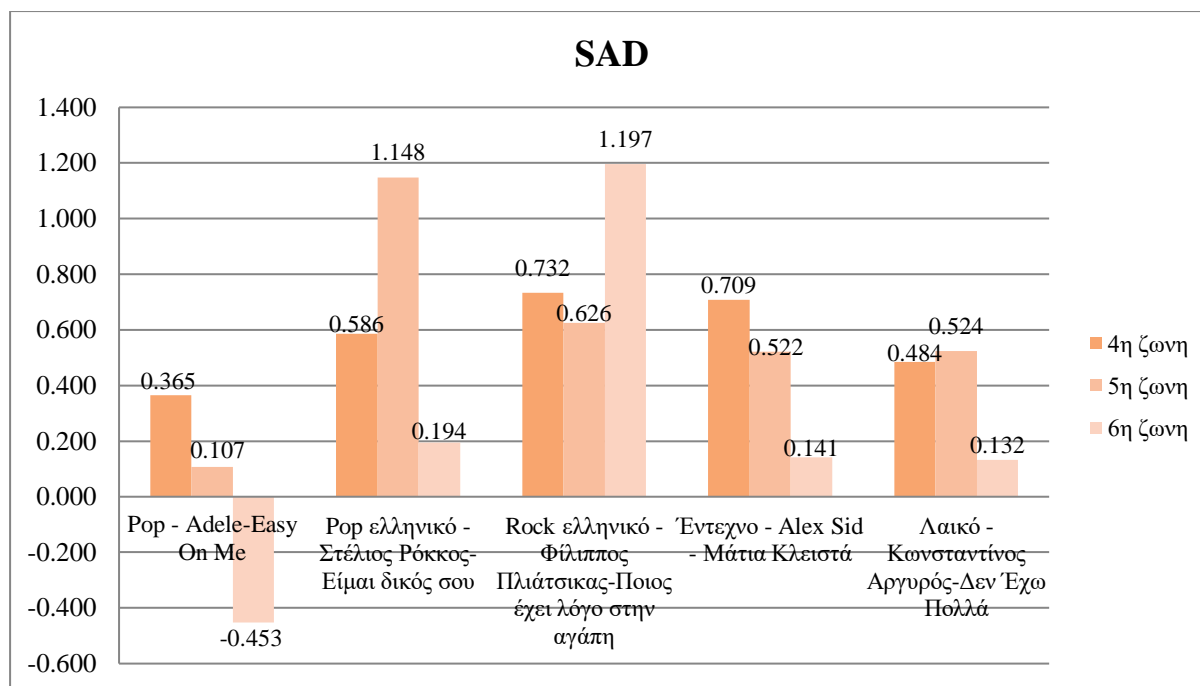


Εικόνα 4.119: Λοξότητα του spectral flatness για την κατηγορία sad

Από τα διαγράμματα παρατηρούμε ότι οι τιμές των πέντε κομματιών για την λοξότητα σε όλα τα χαρακτηριστικά είναι σχετικά κοντά. Η υψηλότερη τιμή στα spectral centroid, spectral bandwidth, spectral rolloff, rms και spectral flatness, παρατηρείται στο Rock ελληνικό - Φίλιππος Πλιάτσικας-Ποιος έχει λόγο στην αγάπη ενώ στα chroma stft και zero crossing παρατηρείται στο Έντεχνο - Alex Sid - Μάτια Κλειστά.

Εικόνα 4.120: Λοξότητα των MFCC (1^{ος} έως 4^{ος}) για την κατηγορία sad

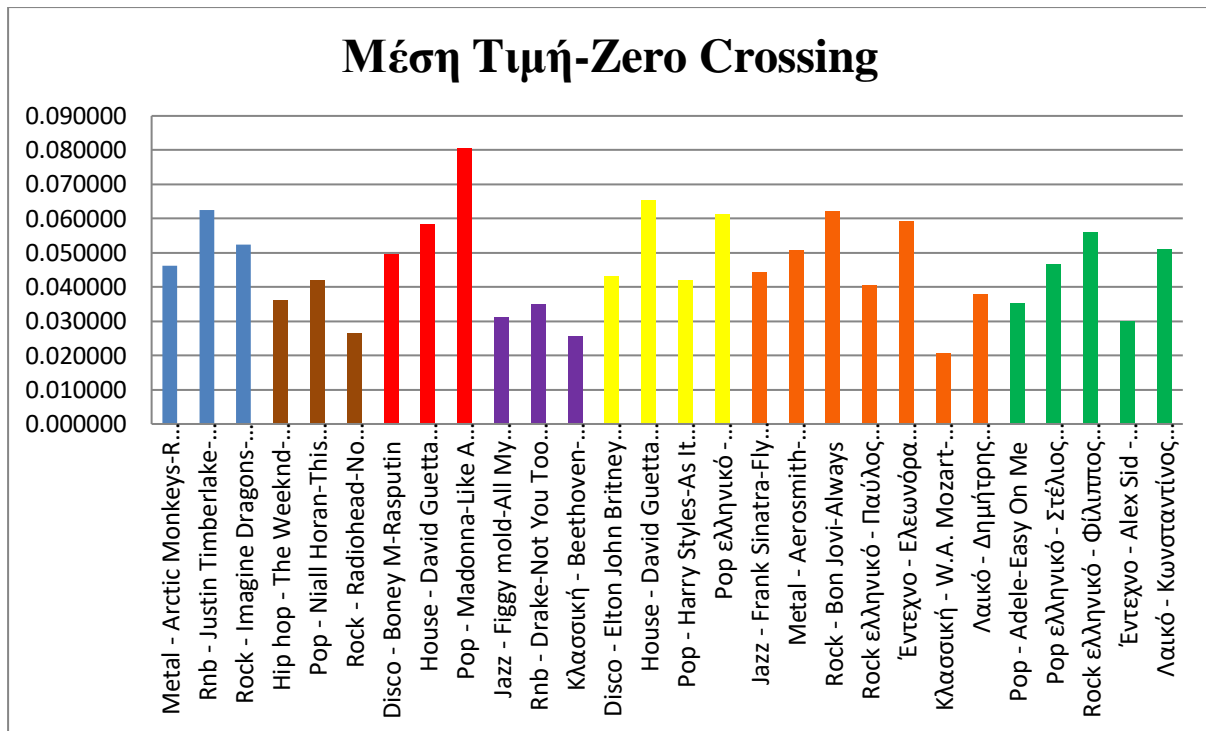
Εικόνα 4.121: Λοξότητα των MFCC (5^{ος} έως 8^{ος}) για την κατηγορία sadΕικόνα 4.122: Λοξότητα του spectral contrast (1^η έως 3^η ζώνη) για την κατηγορία sad

Εικόνα 4.123: Λοξότητα του spectral contrast (4^η έως 6^η ζώνη) για την κατηγορία sad

Όσον αφορά τους mfcc παρατηρούνται περισσότερες αρνητικές τιμές. Στον 1^ο mfcc η υψηλότερη τιμή εμφανίζεται στο Λαϊκό - Κωνσταντίνος Αργυρός-Δεν Έχω Πολλά. Στον 2^ο 4^ο 6^ο 7^ο και 8^ο mfcc η υψηλότερη τιμή εμφανίζεται στο Rock ελληνικό - Φίλιππος Πλιάτσικας-Ποιος έχει λόγο στην αγάπη. Στον 3^ο και 5^ο mfcc η υψηλότερη τιμή εμφανίζεται στο Έντεχνο - Alex Sid - Μάτια Κλειστά. Στην περίπτωση του spectral contrast, στην 1^η ζώνη η υψηλότερη τιμή παρουσιάζεται στο Pop - Adele-Easy On Me, ενώ στην 2^η, 4^η και 6^η εμφανίζεται στο Rock ελληνικό - Φίλιππος Πλιάτσικας-Ποιος έχει λόγο στην αγάπη. Στην 3^η και 5^η ζώνη η υψηλότερη τιμή παρουσιάζεται στο Pop ελληνικό - Στέλιος Ρόκκος-Είμαι δικός σου.

4.2 . Συγκριτική ανάλυση αποτελεσμάτων

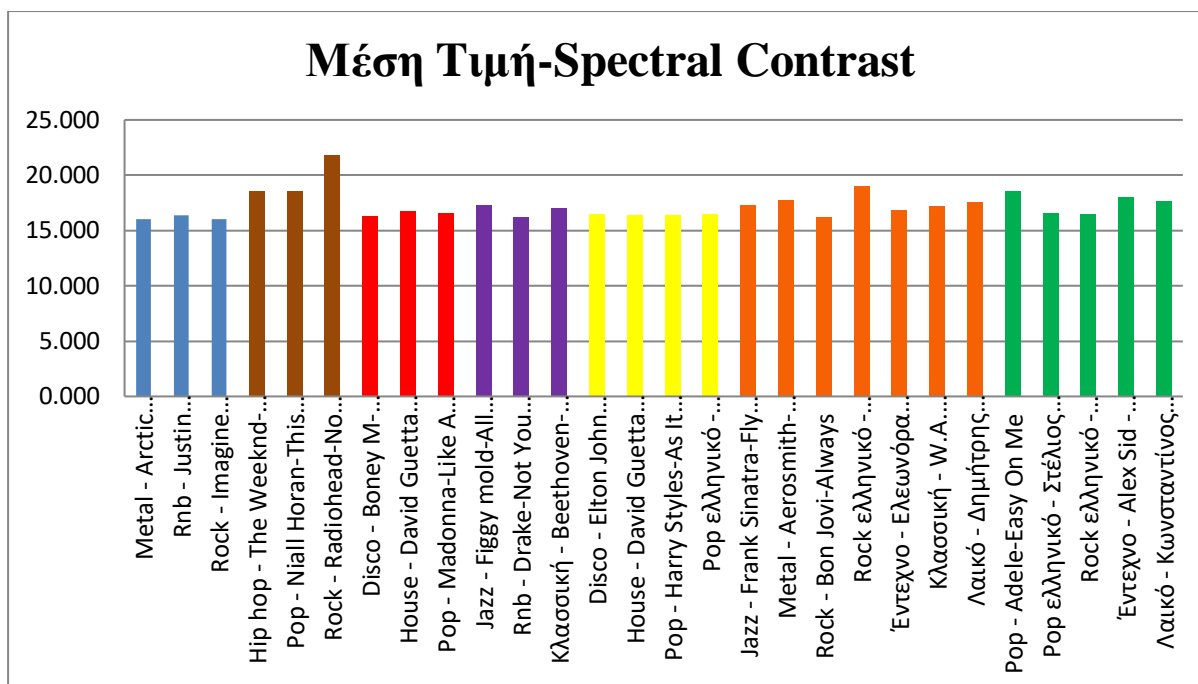
Για να μπορέσουμε να συγκρίνουμε με καλύτερο τρόπο τα αποτελέσματα, δημιουργήσαμε γραφήματα με τη μέση τιμή όλων των χαρακτηριστικών, για όλα τα συναισθήματα. Κάθε συναίσθημα παρουσιάζεται με διαφορετικό χρώμα. Το συναίσθημα angry παρουσιάζεται με χρώμα μπλε ενώ το συναίσθημα chill με χρώμα καφέ. Το συναίσθημα dance απεικονίζεται με χρώμα κόκκινο ενώ το συναίσθημα focus με χρώμα μωβ. Το συναίσθημα happy παρουσιάζεται με χρώμα κίτρινο ενώ το συναίσθημα romance παρουσιάζεται με χρώμα πορτοκαλί. Τέλος, το συναίσθημα sad απεικονίζεται με χρώμα πράσινο. Στην παρακάτω εικόνα παρουσιάζεται το μουσικό χαρακτηριστικό zero crossing.



Εικόνα 4.124: Μέση τιμή όλων των συναισθημάτων για το zero crossing, angry (μπλέ), chill (καφέ), dance (κόκκινο), focus (μωβ), happy (κίτρινο), romance (πορτοκαλί), sad (πράσινο)

Αρχικά, για το συναίσθημα angry, η υψηλότερη μέση τιμή παρατηρείται στο Rnb - Justin Timberlake-Cry Me A River. Στο συναίσθημα chill, η υψηλότερη μέση τιμή εμφανίζεται στο Pop - Niall Horan-This Town ενώ στο συναίσθημα dance παρουσιάζεται στο Pop - Madonna-Like A Prayer. Παράλληλα, στο συναίσθημα focus η υψηλότερη μέση τιμή εμφανίζεται στο Rnb - Drake-Not You Too (Audio) ft. Chris Brown ενώ στο συναίσθημα happy η υψηλότερη μέση τιμή παρουσιάζεται στο House - David Guetta Bebe Rexha-Im Good (Blue). Τέλος, στο συναίσθημα romance η υψηλότερη μέση τιμή εμφανίζεται στο Rock - Bon Jovi-Always ενώ στο συναίσθημα sad η υψηλότερη μέση τιμή παρουσιάζεται στο Rock ελληνικό - Φίλιππος Πλιάτσικας-Ποιος έχει λόγο στην αγάπη. Συγκρίνοντας όλα τα συναισθήματα μαζί η υψηλότερη μέση τιμή του zero crossing εμφανίζεται στο Pop - Madonna-Like A Prayer.

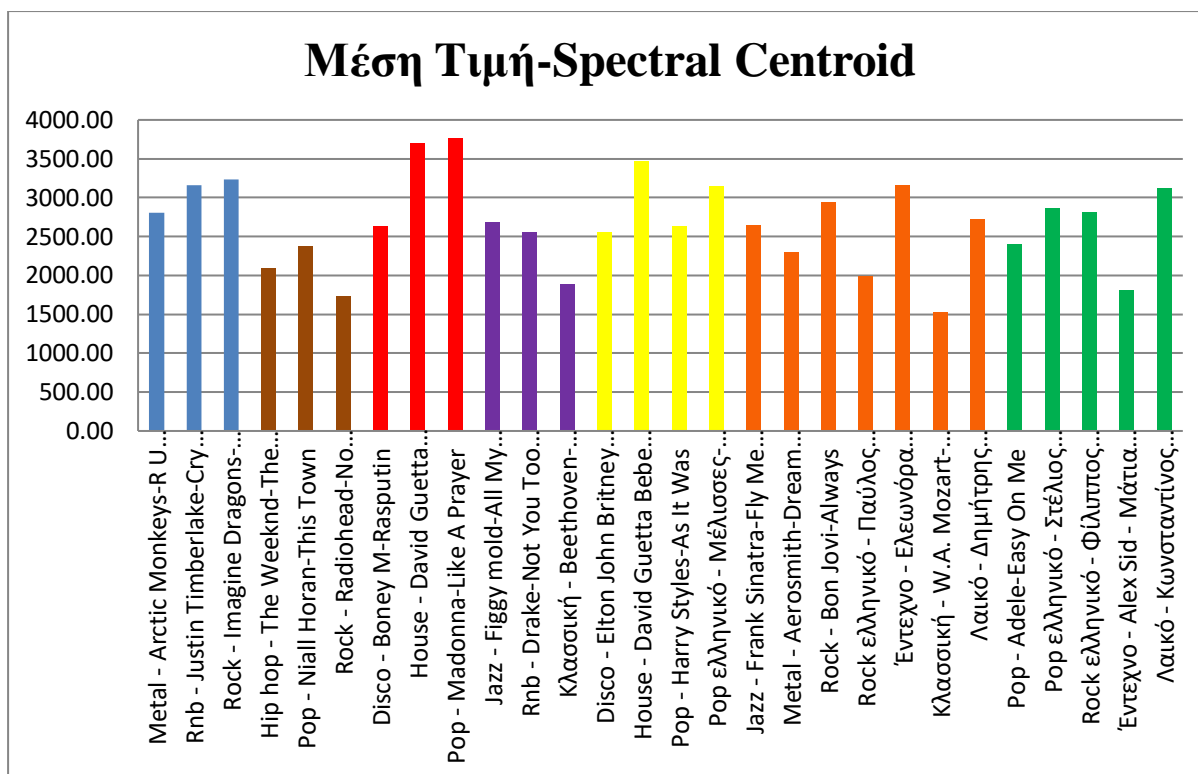
Στην παρακάτω εικόνα παρουσιάζεται το μουσικό χαρακτηριστικό spectral contrast.



Εικόνα 4.125: Μέση τιμή όλων των συναισθημάτων για το spectral contrast, angry (μπλέ), chill (καφέ), dance (κόκκινο), focus (μωβ), happy (κίτρινο), romance (πορτοκαλί), sad (πράσινο)

Αρχικά, για το συναίσθημα angry, η υψηλότερη μέση τιμή παρατηρείται στο Rnb - Justin Timberlake-Cry Me A River. Στο συναίσθημα chill, η υψηλότερη μέση τιμή εμφανίζεται στο Rock - Radiohead-No Surprises ενώ στο συναίσθημα dance παρουσιάζεται στο House - David Guetta Anne-Marie Coi Leray-Baby Dont Hurt Me. Παράλληλα, στο συναίσθημα focus η υψηλότερη μέση τιμή εμφανίζεται στο Jazz - Figgy mold-All My Tomorrows ενώ στο συναίσθημα happy η υψηλότερη μέση τιμή παρουσιάζεται στο Disco - Elton John Britney Spears-Hold Me Closer. Τέλος, στο συναίσθημα romance η υψηλότερη μέση τιμή εμφανίζεται στο Rock ελληνικό - Παύλος Σιδηρόπουλος-Να Μ Αγαπάς ενώ στο συναίσθημα sad η υψηλότερη μέση τιμή παρουσιάζεται στο Pop - Adele-Easy On Me. Συγκρίνοντας όλα τα συναισθήματα μαζί η υψηλότερη μέση τιμή του spectral contrast εμφανίζεται στο Rock - Radiohead-No Surprises.

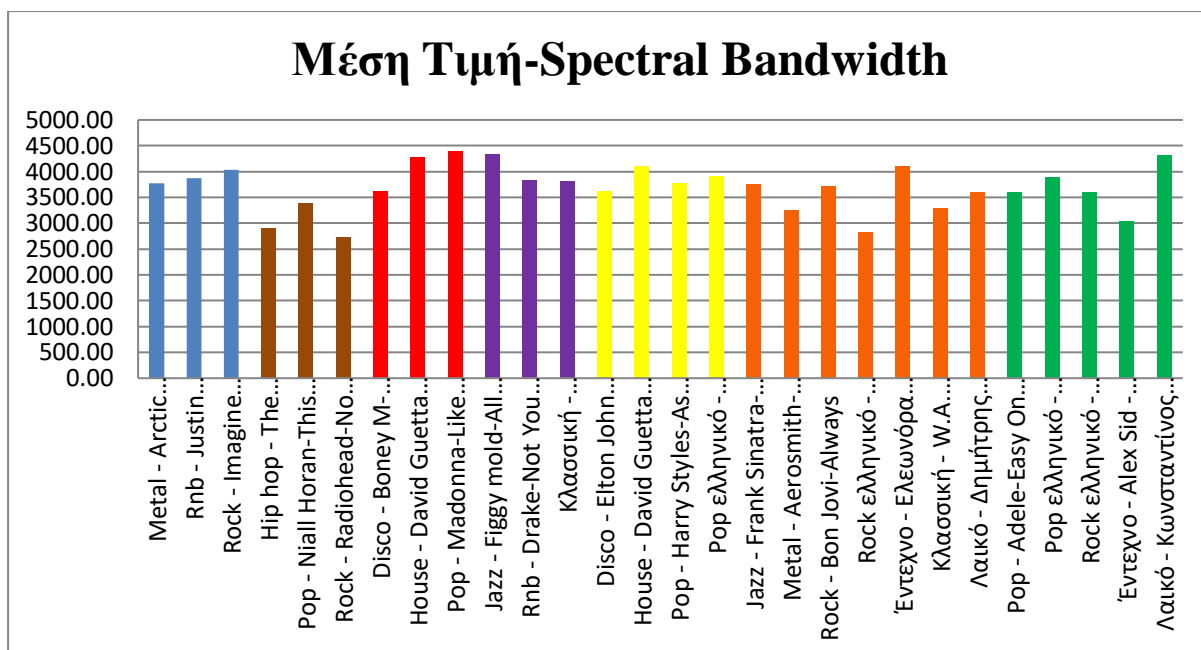
Στην παρακάτω εικόνα παρουσιάζεται το μουσικό χαρακτηριστικό spectral centroid.



Εικόνα 4.126: Μέση τιμή όλων των συναισθημάτων για το spectral centroid, angry (μπλέ), chill (καφέ), dance (κόκκινο), focus (μωβ), happy (κίτρινο), romance (πορτοκαλί), sad (πράσινο)

Αρχικά, για το συναίσθημα angry, η υψηλότερη μέση τιμή παρατηρείται στο Rock - Imagine Dragons-Believer. Στο συναίσθημα chill, η υψηλότερη μέση τιμή εμφανίζεται στο Pop - Niall Horan-This Town ενώ στο συναίσθημα dance παρουσιάζεται στο Pop - Madonna-Like A Prayer. Παράλληλα, στο συναίσθημα focus η υψηλότερη μέση τιμή εμφανίζεται στο Jazz - Figgy mold-All My Tomorrows ενώ στο συναίσθημα happy η υψηλότερη μέση τιμή παρουσιάζεται στο House - David Guetta Bebe Rexha-Im Good (Blue). Τέλος, στο συναίσθημα romance η υψηλότερη μέση τιμή εμφανίζεται στο Έντεχνο - Ελεωνόρα Ζουγανέλη-Η Επιμονή σου ενώ στο συναίσθημα sad η υψηλότερη μέση τιμή παρουσιάζεται στο Λαϊκό - Κωνσταντίνος Αργυρός-Δεν Έχω Πολλά. Συγκρίνοντας όλα τα συναισθήματα μαζί η υψηλότερη μέση τιμή του spectral centroid εμφανίζεται στο Pop - Madonna-Like A Prayer.

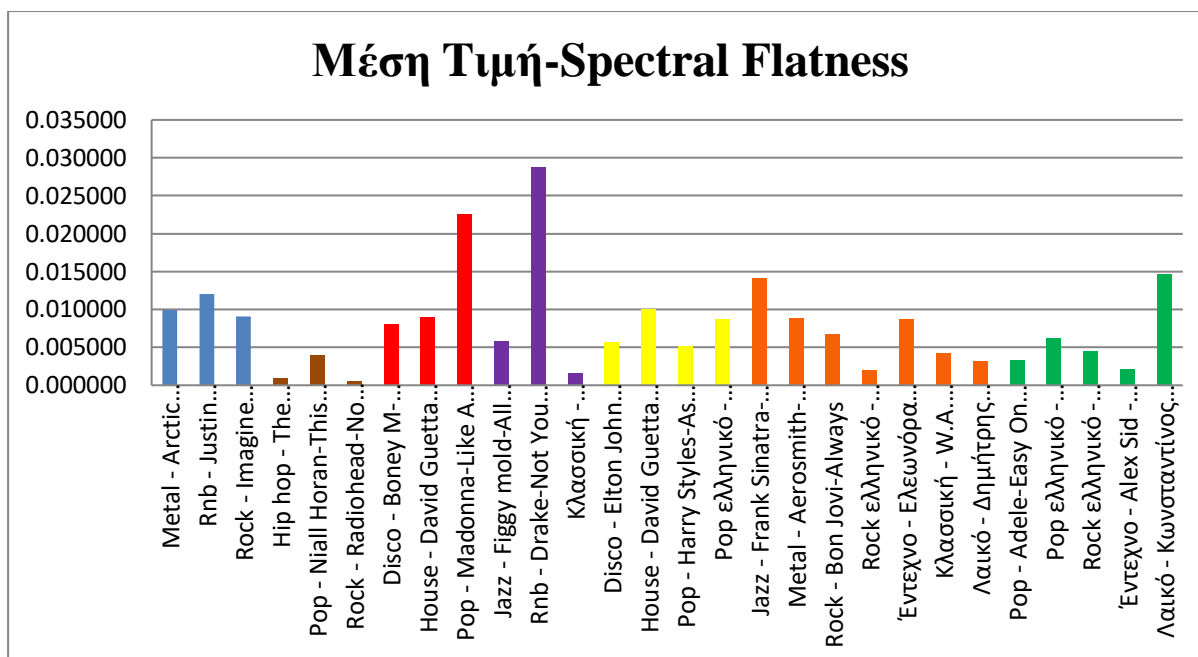
Στην παρακάτω εικόνα παρουσιάζεται το μουσικό χαρακτηριστικό spectral bandwidth.



Εικόνα 4.127: Μέση τιμή όλων των συναισθημάτων για το spectral bandwidth, angry (μπλέ), chill (καφέ), dance (κόκκινο), focus (μωβ), happy (κίτρινο), romance (πορτοκαλί), sad (πράσινο)

Αρχικά, για το συναίσθημα angry, η υψηλότερη μέση τιμή παρατηρείται στο Rock - Imagine Dragons-Believer. Στο συναίσθημα chill, η υψηλότερη μέση τιμή εμφανίζεται στο Pop - Niall Horan-This Town ενώ στο συναίσθημα dance παρουσιάζεται στο Pop - Madonna-Like A Prayer. Παράλληλα, στο συναίσθημα focus η υψηλότερη μέση τιμή εμφανίζεται στο Jazz - Figgy mold-All My Tomorrows ενώ στο συναίσθημα happy η υψηλότερη μέση τιμή παρουσιάζεται στο House - David Guetta Bebe Rexha-Im Good (Blue). Τέλος, στο συναίσθημα romance η υψηλότερη μέση τιμή εμφανίζεται στο Έντεχνο - Ελεωνόρα Ζουγανέλη-Η Επιμονή σου ενώ στο συναίσθημα sad η υψηλότερη μέση τιμή παρουσιάζεται στο Λαϊκό - Κωνσταντίνος Αργυρός-Δεν Έχω Πολλά. Συγκρίνοντας όλα τα συναισθήματα μαζί η υψηλότερη μέση τιμή του spectral bandwidth εμφανίζεται στο Pop - Madonna-Like A Prayer.

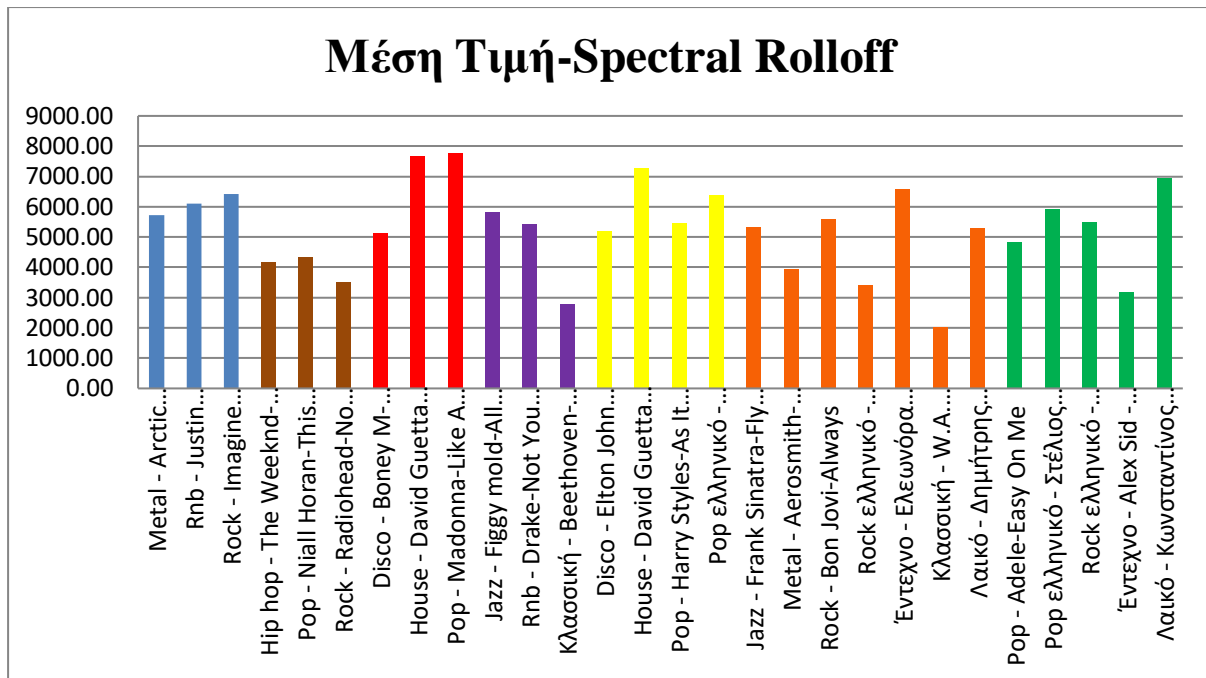
Στην παρακάτω εικόνα παρουσιάζεται το μουσικό χαρακτηριστικό spectral flatness.



Εικόνα 4.128: Μέση τιμή όλων των συναισθημάτων για το spectral flatness, angry (μπλέ), chill (καφέ), dance (κόκκινο), focus (μωβ), happy (κίτρινο), romance (πορτοκαλί), sad (πράσινο)

Αρχικά, για το συναίσθημα angry, η υψηλότερη μέση τιμή παρατηρείται στο Rnb - Justin Timberlake-Cry Me A River. Στο συναίσθημα chill, η υψηλότερη μέση τιμή εμφανίζεται στο Pop - Niall Horan-This Town ενώ στο συναίσθημα dance παρουσιάζεται στο Pop - Madonna-Like A Prayer. Παράλληλα, στο συναίσθημα focus η υψηλότερη μέση τιμή εμφανίζεται στο Rnb - Drake-Not You Too (Audio) ft. Chris Brown ενώ στο συναίσθημα happy η υψηλότερη μέση τιμή παρουσιάζεται στο House - David Guetta Bebe Rexha-Im Good (Blue). Τέλος, στο συναίσθημα romance η υψηλότερη μέση τιμή εμφανίζεται στο Jazz - Frank Sinatra-Fly Me To The Moon ενώ στο συναίσθημα sad η υψηλότερη μέση τιμή παρουσιάζεται στο Λαϊκό - Κωνσταντίνος Αργυρός-Δεν Έχω Πολλά. Συγκρίνοντας όλα τα συναισθήματα μαζί η υψηλότερη μέση τιμή του spectral flatness εμφανίζεται στο Rnb - Drake-Not You Too (Audio) ft. Chris Brown.

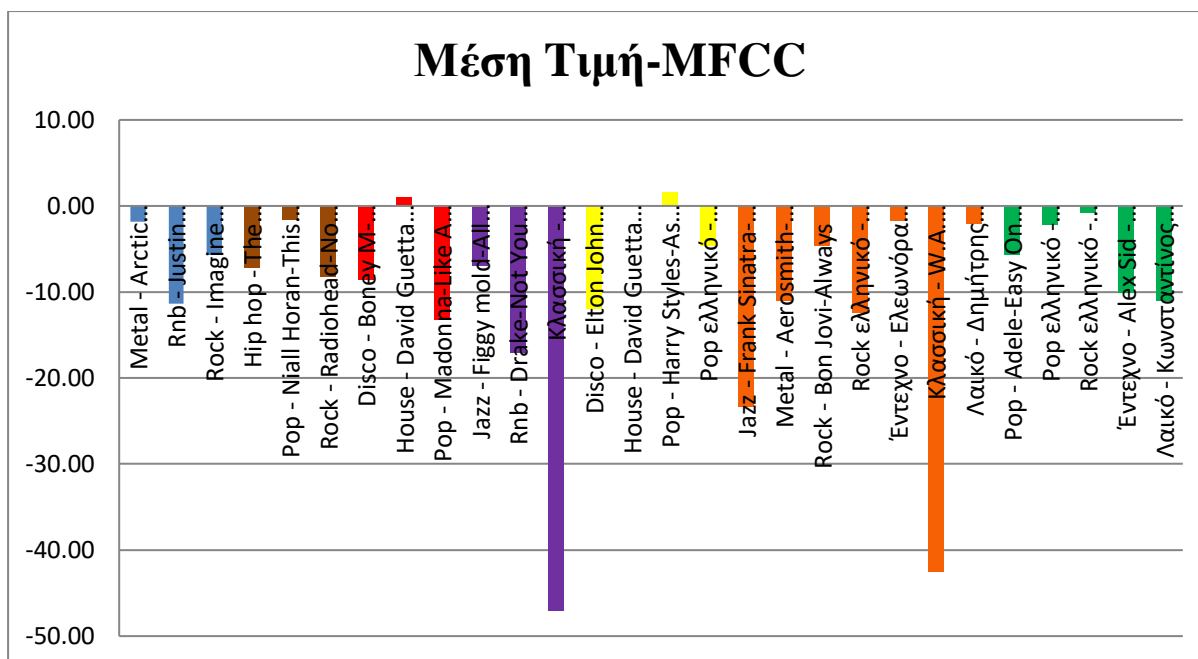
Στην παρακάτω εικόνα παρουσιάζεται το μουσικό χαρακτηριστικό spectral rolloff.



Εικόνα 4.129: Μέση τιμή όλων των συναισθημάτων για το spectral rolloff, angry (μπλέ), chill (καφέ), dance (κόκκινο), focus (μωβ), happy (κίτρινο), romance (πορτοκαλί), sad (πράσινο)

Αρχικά, για το συναίσθημα angry, η υψηλότερη μέση τιμή παρατηρείται στο Rock - Imagine Dragons-Believer. Στο συναίσθημα chill, η υψηλότερη μέση τιμή εμφανίζεται στο Pop - Niall Horan-This Town ενώ στο συναίσθημα dance παρουσιάζεται στο Pop - Madonna-Like A Prayer. Παράλληλα, στο συναίσθημα focus η υψηλότερη μέση τιμή εμφανίζεται στο Jazz - Figgy mold-All My Tomorrows ενώ στο συναίσθημα happy η υψηλότερη μέση τιμή παρουσιάζεται στο House - David Guetta Bebe Rexha-Im Good (Blue). Τέλος, στο συναίσθημα romance η υψηλότερη μέση τιμή εμφανίζεται στο Έντεχνο - Ελεωνόρα Ζουγανέλη-Η Επιμονή σου ενώ στο συναίσθημα sad η υψηλότερη μέση τιμή παρουσιάζεται στο Λαϊκό - Κωνσταντίνος Αργυρός-Δεν Έχω Πολλά. Συγκρίνοντας όλα τα συναισθήματα μαζί η υψηλότερη μέση τιμή του spectral rolloff εμφανίζεται στο Pop - Madonna-Like A Prayer.

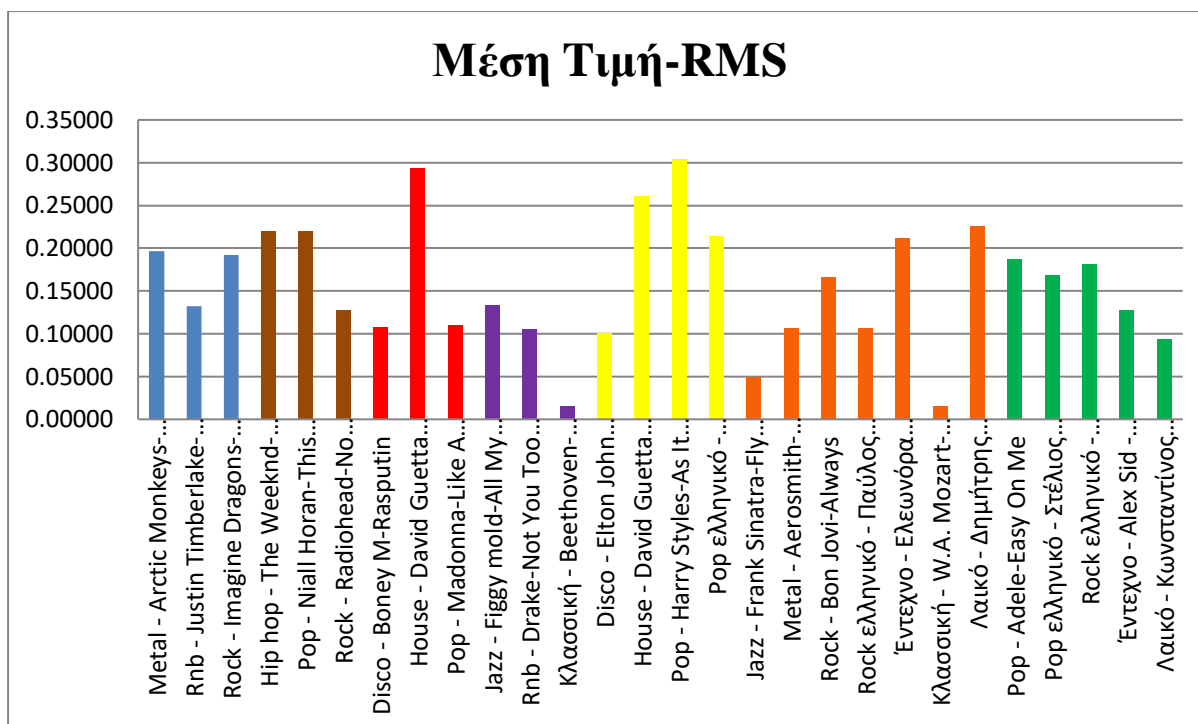
Στην παρακάτω εικόνα παρουσιάζεται το μουσικό χαρακτηριστικό MFCC.



Εικόνα 4.130: Μέση τιμή όλων των συναισθημάτων για τους MFCC, angry (μπλέ), chill (καφέ), dance (κόκκινο), focus (μωβ), happy (κίτρινο), romance (πορτοκαλί), sad (πράσινο)

Αρχικά, για το συναίσθημα angry, η υψηλότερη μέση τιμή παρατηρείται στο Rnb - Justin Timberlake-Cry Me A River. Στο συναίσθημα chill, η υψηλότερη μέση τιμή εμφανίζεται στο Rock - Radiohead-No Surprises ενώ στο συναίσθημα dance παρουσιάζεται στο House - David Guetta Anne-Marie Coi Leray-Baby Dont Hurt Me. Παράλληλα, στο συναίσθημα focus η υψηλότερη μέση τιμή εμφανίζεται στο Κλασσική - Beethoven-Tempo di Menuetto ενώ στο συναίσθημα happy η υψηλότερη μέση τιμή παρουσιάζεται στο Pop - Harry Styles-As It Was. Τέλος, στο συναίσθημα romance η υψηλότερη μέση τιμή εμφανίζεται στο Κλασσική - W.A. Mozart-Serenade ενώ στο συναίσθημα sad η υψηλότερη μέση τιμή παρουσιάζεται στο Λαϊκό - Κωνσταντίνος Αργυρός-Δεν Έχω Πολλά. Συγκρίνοντας όλα τα συναισθήματα μαζί η υψηλότερη μέση τιμή των MFCC εμφανίζεται στο Κλασσική - Beethoven-Tempo di Menuetto.

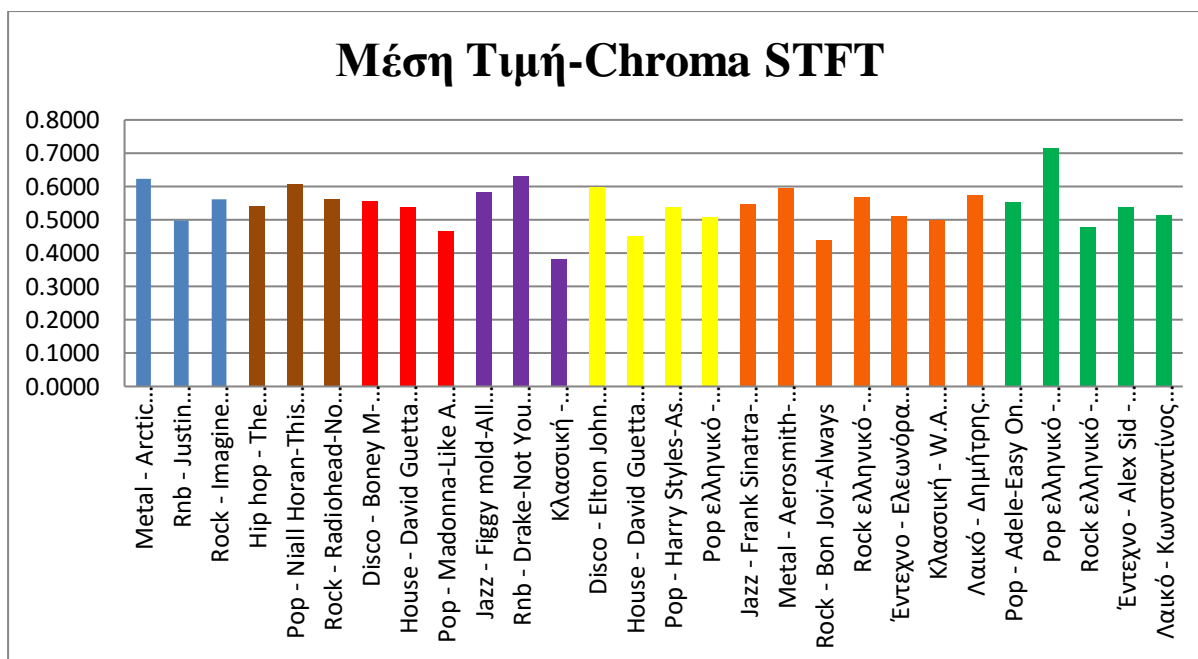
Στην παρακάτω εικόνα παρουσιάζεται το μουσικό χαρακτηριστικό RMS.



Εικόνα 4.131: Μέση τιμή όλων των συναισθημάτων για τους RMS, angry (μπλέ), chill (καφέ), dance (κόκκινο), focus (μωβ), happy (κίτρινο), romance (πορτοκαλί), sad (πράσινο)

Αρχικά, για το συναίσθημα angry, η υψηλότερη μέση τιμή παρατηρείται στο Metal - Arctic Monkeys-R U Mine. Στο συναίσθημα chill, η υψηλότερη μέση τιμή εμφανίζεται στο Pop - Niall Horan-This Town ενώ στο συναίσθημα dance παρουσιάζεται στο House - David Guetta Anne-Marie Coi Leray-Baby Dont Hurt Me. Παράλληλα, στο συναίσθημα focus η υψηλότερη μέση τιμή εμφανίζεται στο Jazz - Figgy mold-All My Tomorrows ενώ στο συναίσθημα happy η υψηλότερη μέση τιμή παρουσιάζεται στο Pop - Harry Styles-As It Was. Τέλος, στο συναίσθημα romance η υψηλότερη μέση τιμή εμφανίζεται στο Λαϊκό - Δημήτρης Μητροπάνος-Όταν έχω εσένα ενώ στο συναίσθημα sad η υψηλότερη μέση τιμή παρουσιάζεται στο Pop - Adele-Easy On Me. Συγκρίνοντας όλα τα συναισθήματα μαζί η υψηλότερη μέση τιμή των RMS εμφανίζεται στο Pop - Harry Styles-As It Was.

Στην παρακάτω εικόνα παρουσιάζεται το μουσικό χαρακτηριστικό Chroma STFT.

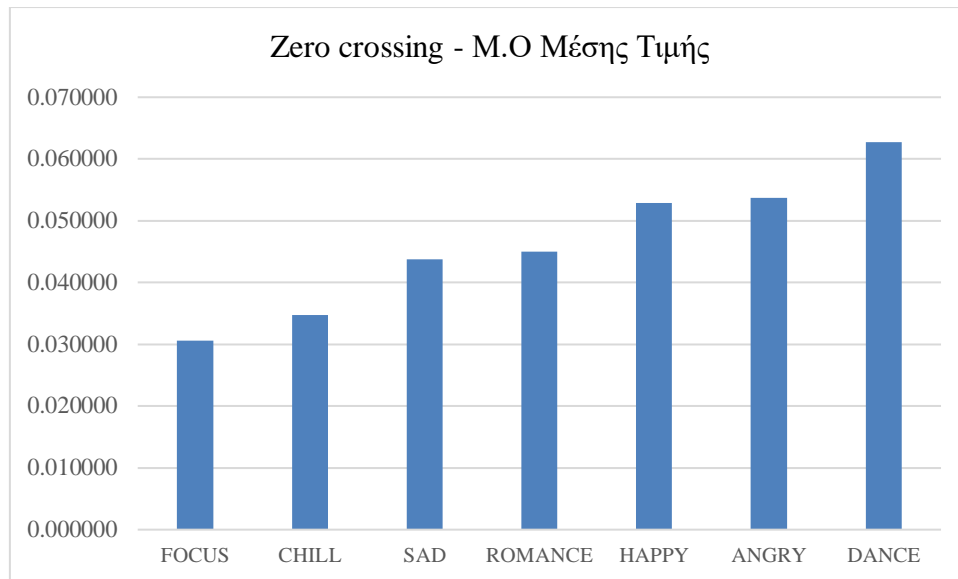


Εικόνα 4.132: Μέση τιμή όλων των συναισθημάτων για τους Chroma STFT, angry (μπλέ), chill (καφέ), dance (κόκκινο), focus (μωβ), happy (κίτρινο), romance (πορτοκαλί), sad (πράσινο)

Αρχικά, για το συναίσθημα angry, η υψηλότερη μέση τιμή παρατηρείται στο Metal - Arctic Monkeys-R U Mine. Στο συναίσθημα chill, η υψηλότερη μέση τιμή εμφανίζεται στο Pop - Niall Horan-This Town ενώ στο συναίσθημα dance παρουσιάζεται στο Disco - Boney M-Rasputin. Παράλληλα, στο συναίσθημα focus η υψηλότερη μέση τιμή εμφανίζεται στο Rnb - Drake-Not You Too (Audio) ft. Chris Brown ενώ στο συναίσθημα happy η υψηλότερη μέση τιμή παρουσιάζεται στο Disco - Elton John Britney Spears-Hold Me Closer. Τέλος, στο συναίσθημα romance η υψηλότερη μέση τιμή εμφανίζεται στο Metal - Aerosmith-Dream On ενώ στο συναίσθημα sad η υψηλότερη μέση τιμή παρουσιάζεται στο Pop ελληνικό - Στέλιος Ρόκκος-Είμαι δικός σου. Συγκρίνοντας όλα τα συναισθήματα μαζί η υψηλότερη μέση τιμή των Chroma STFT εμφανίζεται στο Pop ελληνικό - Στέλιος Ρόκκος-Είμαι δικός σου.

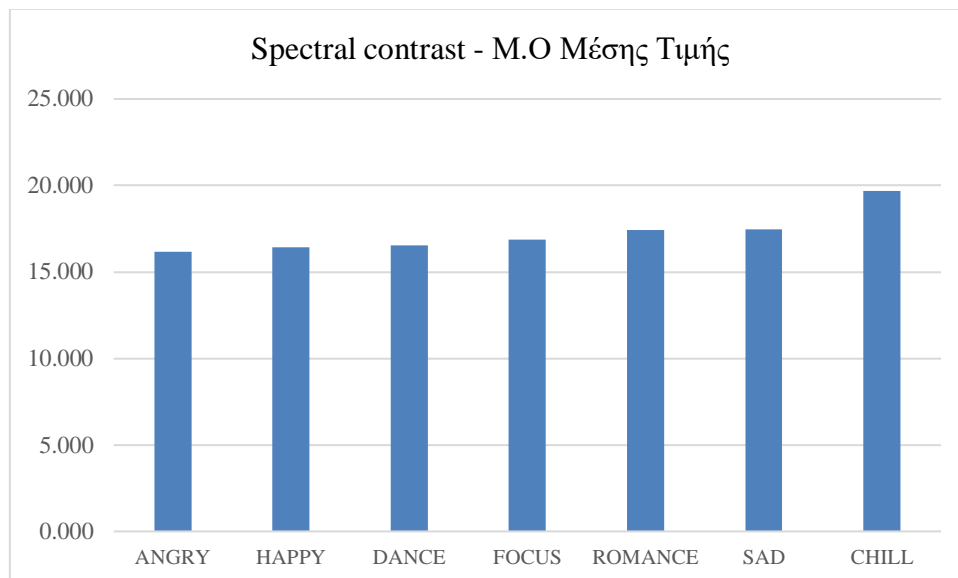
Σύμφωνα με τα παραπάνω αποτελέσματα στο συναίσθημα angry, η υψηλότερη μέση τιμή στα περισσότερα χαρακτηριστικά (4/9) εμφανίζεται στο Rnb - Justin Timberlake-Cry Me A River. Στο συναίσθημα chill η υψηλότερη μέση τιμή στα περισσότερα χαρακτηριστικά (7/9) εμφανίζεται στο Pop - Niall Horan-This Town. Στο συναίσθημα dance η υψηλότερη μέση τιμή στα περισσότερα χαρακτηριστικά (5/9) εμφανίζεται στο Pop - Madonna-Like A Prayer. Στο συναίσθημα focus η υψηλότερη μέση τιμή στα περισσότερα χαρακτηριστικά (5/9) εμφανίζεται στο Jazz - Figgy mold-All My Tomorrows. Στο συναίσθημα happy η υψηλότερη μέση τιμή στα περισσότερα χαρακτηριστικά (5/9) εμφανίζεται στο House - David Guetta Bebe Rexha-Im Good (Blue). Στο συναίσθημα romance η υψηλότερη μέση τιμή στα περισσότερα χαρακτηριστικά (3/9) εμφανίζεται στο Έντεχνο - Έλεωνόρα Ζουγανέλη-Η Επιμονή σου. Στο συναίσθημα sad η υψηλότερη μέση τιμή στα περισσότερα χαρακτηριστικά (3/9) εμφανίζεται στο Λαϊκό - Κωνσταντίνος Αργυρός-Δεν Έχω Πολλά.

Για να μπορέσουμε να συγκρίνουμε τις κατηγορίες συναισθήματος μεταξύ τους, υπολογίσαμε τον μέσο όρο της μέσης τιμής για όλα τα μουσικά χαρακτηριστικά. Τα διαγράμματα με τα αποτελέσματα παρατίθενται στις παρακάτω εικόνες.



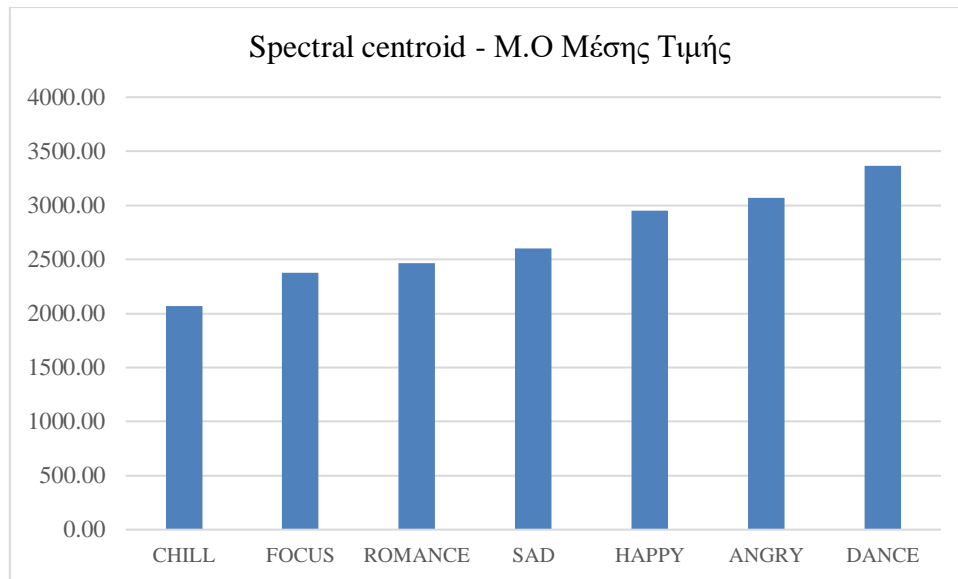
Εικόνα 4.133: Μέσος όρος μέσης τιμής όλων των συναισθημάτων για το Zero Crossing

Για το Zero Crossing δεν παρατηρούνται ιδιαίτερες διαφορές μεταξύ των συναισθημάτων ώστε να εξάγουμε κάποιο συμπέρασμα. Ωστόσο, το συγκεκριμένο χαρακτηριστικό μετρά το θόρυβο. Η υψηλότερη τιμή παρατηρείται στα happy, angry, dance, το οποίο πιθανόν να οφείλεται στο γεγονός ότι αυτές οι κατηγορίες περιλαμβάνουν κομμάτια house και metal.



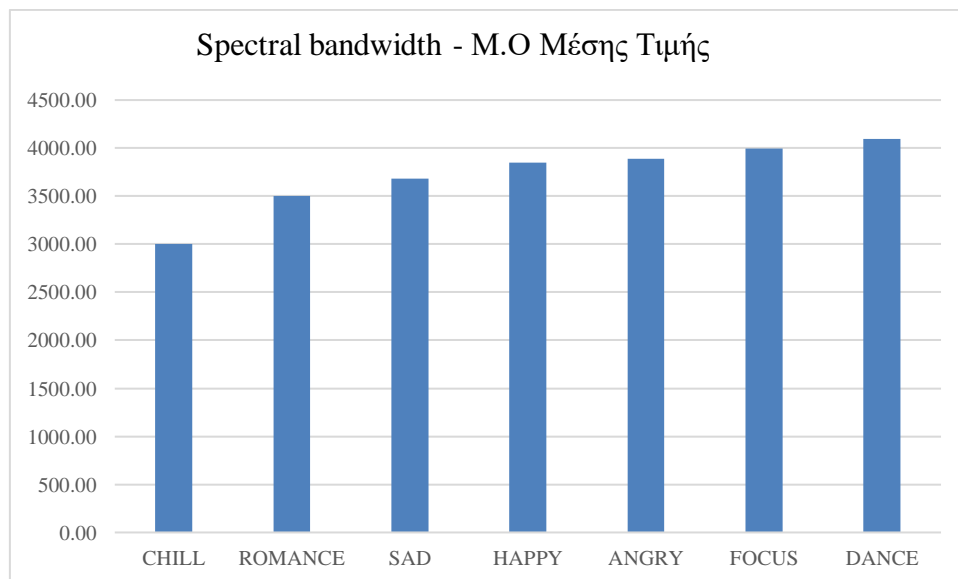
Εικόνα 4.134: Μέσος όρος μέσης τιμής όλων των συναισθημάτων για το Spectral Contrast

Για το Spectral Contrast δεν παρατηρούνται ιδιαίτερες διαφορές μεταξύ των συναισθημάτων ώστε να εξάγουμε κάποιο συμπέρασμα.



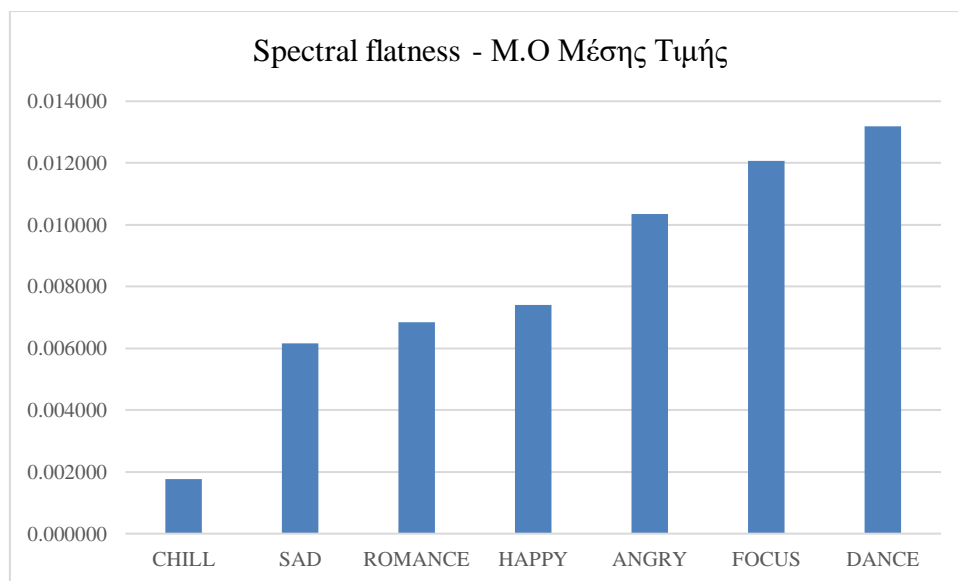
Εικόνα 4.135: Μέσος όρος μέσης τιμής όλων των συναισθημάτων για το Spectral Centroid

Για το Spectral Centroid δεν παρατηρούνται ιδιαίτερες διαφορές μεταξύ των συναισθημάτων ώστε να εξάγουμε κάποιο συμπέρασμα. Ωστόσο, το συγκεκριμένο χαρακτηριστικό υπολογίζει το φασματικό κέντρο. Η υψηλότερη τιμή παρατηρείται στα happy, angry, dance, το οποίο πιθανόν να οφείλεται στα διαφορετικά είδη μουσικής.



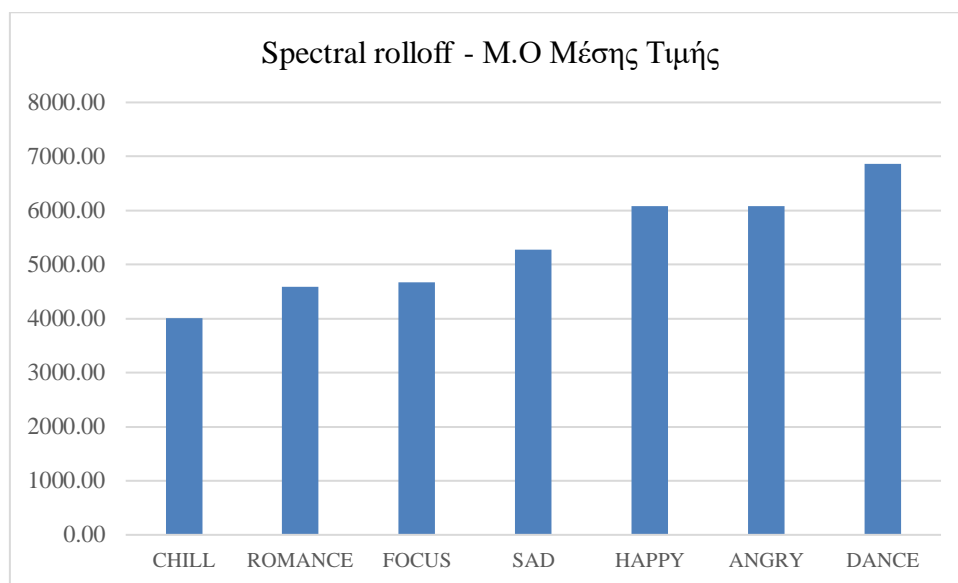
Εικόνα 4.136: Μέσος όρος μέσης τιμής όλων των συναισθημάτων για το Spectral Bandwidth

Για το Spectral Bandwidth δεν παρατηρούνται ιδιαίτερες διαφορές μεταξύ των συναισθημάτων ώστε να εξάγουμε κάποιο συμπέρασμα.



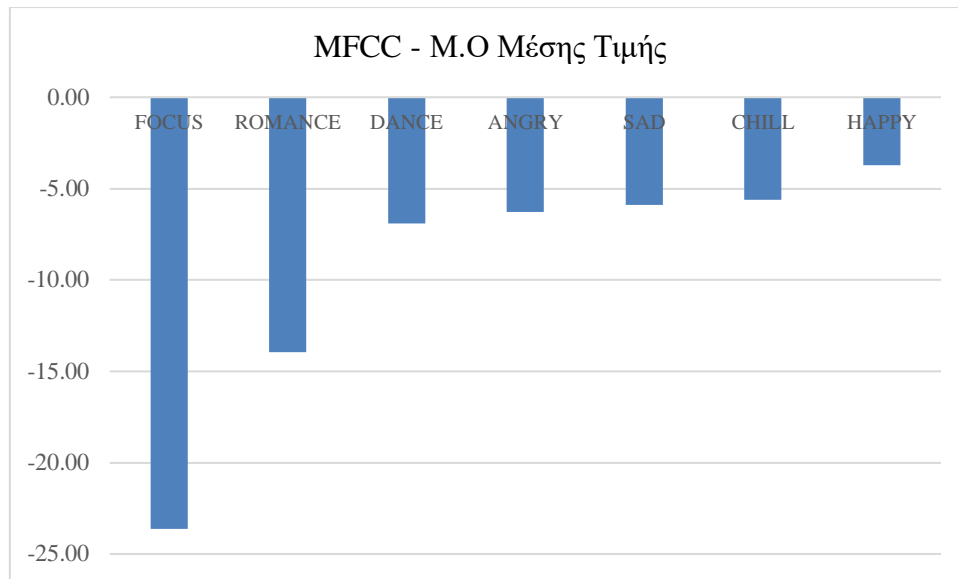
Εικόνα 4.137: Μέσος όρος μέσης τιμής όλων των συναισθημάτων για το Spectral Flatness

Για το Spectral Flatness δεν παρατηρούνται ιδιαίτερες διαφορές μεταξύ των συναισθημάτων ώστε να εξάγουμε κάποιο συμπέρασμα.

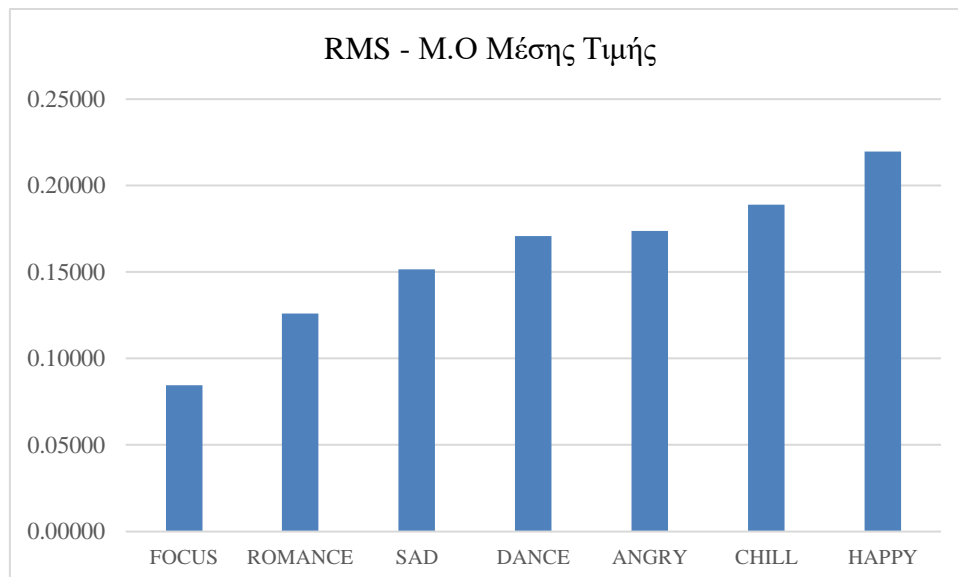


Εικόνα 4.138: Μέσος όρος μέσης τιμής όλων των συναισθημάτων για το Spectral Rolloff

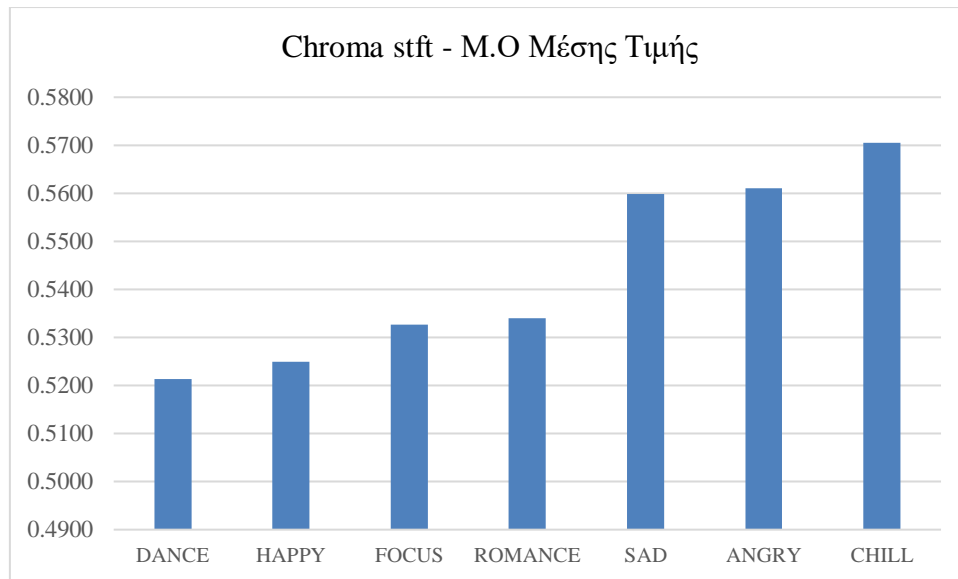
Για το Spectral Rolloff δεν παρατηρούνται ιδιαίτερες διαφορές μεταξύ των συναισθημάτων ώστε να εξάγουμε κάποιο συμπέρασμα.



Εικόνα 4.139: Μέσος όρος μέσης τιμής όλων των συναισθημάτων για τους MFCC



Εικόνα 4.140: Μέσος όρος μέσης τιμής όλων των συναισθημάτων για την RMS



Εικόνα 4.141: Μέσος όρος μέσης τιμής όλων των συναισθημάτων για το Chroma Stft

Αντίστοιχα, για τα MFCC, RMS, και Chroma Stft, δεν παρατηρούνται ιδιαίτερες διαφορές μεταξύ των συναισθημάτων ώστε να εξάγουμε κάποιο συμπέρασμα.

Συγκρίνοντας όλα τα αποτελέσματα παρουσιάζονται αποκλίσεις στις τιμές, κάτι που αποδεικνύει ότι δεν μπορεί να υπάρξει ταύτιση των κομματιών, καθώς αναφερόμαστε σε διαφορετικά είδη μουσικής. Αξίζει να σημειωθεί ότι, στα συναισθήματα romance και sad παρουσιάζονται οι μεγαλύτερες αποκλίσεις. Αυτό οφείλεται στο γεγονός ότι περιλαμβάνουν μεγαλύτερο αριθμό κομματιών σε σχέση με τα υπόλοιπα συναισθήματα αλλά και περισσότερα είδη μουσικής. Επομένως, είναι φυσικό να μην υπάρχει ταύτιση μεταξύ τους. Τέλος, συγκρίνοντας τις κατηγορίες μεταξύ τους παρατηρούνται πολύ μικρές διαφορές στις τιμές κάτι το οποίο δεν μας βοηθάει να εξάγουμε κάποιο συμπέρασμα.

5. Συμπεράσματα και Μελλοντικές Προεκτάσεις

Μελετώντας τα αποτελέσματα από τα διαγράμματα του κεφαλαίου 4, μπορούμε να πούμε πως τα κομμάτια που έχουν σχολιαστεί με τα αντίστοιχα συναισθήματα, δεν ταυτίζονται απόλυτα. Υπάρχουν αποκλίσεις στα μουσικά χαρακτηριστικά και ενδεχομένως αυτό οφείλεται στο γεγονός ότι τα είδη μουσικής είναι διαφορετικά. Σε κάθε είδος μουσικής, το φασματικό κέντρο (spectral centroid) είναι πολύ λογικό να διαφέρει καθώς επίσης και το spectral flatness που σχετίζεται με την τονικότητα. Εύκολα μπορούμε να πούμε και για το spectral rolloff ότι υπάρχουν διαφορές επίσης, μιας και σχετίζεται με την ποσότητα υψηλής συχνότητας στο σήμα. Συγκρίνοντας τη μέση τιμή όλων των χαρακτηριστικών σε όλα τα συναισθήματα παρατηρούμε ότι, στις κατηγορίες angry και chill παρουσιάζονται οι μικρότερες αποκλίσεις. Από αυτό το αποτέλεσμα συμπεραίνουμε ότι τα κομμάτια πιθανόν να ταίριαζαν στην αντίστοιχη κατηγορία.

5.1 Αξιοποίηση Αποτελεσμάτων - Πιθανές Εφαρμογές

Τα αποτελέσματα που έχουν εξαχθεί θα μπορούσαν να χρησιμοποιηθούν και σε μελλοντικές μελέτες, στις οποίες δεν είναι εφικτός ο έλεγχος των χρηστών που σχολιάζουν μουσικά κομμάτια σε εφαρμογές. Αρχικά, θα μπορούσε να αποτελέσει μια πολύ μικρή βάση δεδομένων με τα κομμάτια και τα μουσικά χαρακτηριστικά που εξήχθησαν και να ελεγχθεί η πιθανότητα ένταξης κάθε κομματιού σε διαφορετικές κατηγορίες συναισθήματος. Επιπλέον, θα μπορούσαν να χρησιμοποιηθούν σε μελέτες αξιολόγησης της εφαρμογής, για το κατά πόσο ταυτοποιείται η κατηγοριοποίηση των κομματιών στα αντίστοιχα συναισθήματα. Τέλος, θα μπορούσε να χρησιμοποιηθεί και ως πρόταση για βελτίωση της εφαρμογής.

5.2 Μελλοντικές Προεκτάσεις – Βελτιώσεις

Στη συγκεκριμένη μελέτη έγινε επιλογή συναισθημάτων από μία ποικιλία που διέθετε το Spotify. Παρόλα αυτά, υπάρχουν κατηγορίες που θα μπορούσαν να συμπτυχθούν καθώς τα κομμάτια που περιλαμβάνουν μπορούν να ανήκουν σε παραπάνω από μία κατηγορία. Για παράδειγμα θα μπορούσε η κατηγορία happy να ενωθεί με την κατηγορία dance. Επίσης, η κατηγορία sad θα μπορούσε να συμπτυχθεί με την κατηγορία romance. Τέλος, θα μπορούσε να γίνει προσπάθεια εύρεσης κοινών κατηγοριών συναισθημάτων και κομματιών σε άλλες εφαρμογές και να γίνει σύγκριση μεταξύ τους.

Βιβλιογραφία

Abeles, H., F., Chung, J., W. (1996). *Responses to Music*. IMR Press. San Antonio.

Ali., O. (2006). *Songs and emotions: are lyrics and melodies equal partners*. Psychology of Music. doi:[10.1177/0305735606067168](https://doi.org/10.1177/0305735606067168).

Allamanche, E., Herre, J., Helmuth, O., Froba, B., Kasten, T., Cremer, T. (2001). *Content-based identification of audio material using MPEG-7 low level description*. In Proc. Int. Conf. Music Information Retrieval. Erlangen, Germany.

Banitalebi-Dehkordi, M., Banitalebi-Dehkordi, A. (2018). *Music Genre Classification Using Spectral Analysis and Sparse Representation of the Signals*. Journal of Signal Processing Systems. <https://doi.org/10.48550/arXiv.1803.04652>.

Βαρδάνης, Π. (2019). Στατιστική ανάλυση της επίδρασης του mastering σε ηχητικά σήματα. Πανεπιστήμιο Πατρών, Τμήμα Ηλεκτρολόγων Μηχανικών και Τεχνολογίας Υπολογιστών, Τομέας: Τηλεπικοινωνιών & Τεχνολογίας Πληροφορίας, Εργαστήριο: Ακουστικής & Τεχνολογίας ήχου, Πάτρα.

Bartoszewski, H., M., Kwasnicka, H., Markowska-Kaczmar, U., Myszkowski, P., B. (2008). *Extraction of emotional content from music data*. In Proc. Computer Information Systems and Industrial Management Applications. doi:[10.1109/CISIM.2008.46](https://doi.org/10.1109/CISIM.2008.46).

Benetos, E., Kotti, M., Kotropoulos, C. (2007). *Large scale musical instrument identification*. In Proc. Int. Conf. Music Information Retrieval. Lefkada, Greece.

Bischoff, K., Firan, C., S., Paiu, R., Nejd, W., Laurier, C., Sordo., M. (2009). *Music mood and theme classification-a hybrid approach*. In Proc. Int. Conf. Music Information Retrieval.

Cabrera, D. (1999). *PsySound: A computer program for psycho-acoustical analysis*. In Proc. Australian Acoustic Society Conf. Melbourne, Australia.

Camacho, A. (2007). *SWIPE: A Sawtooth Waveform Inspired Pitch Estimator for Speech and Music*. Univ. Florida. doi:[10.1121/1.2951592](https://doi.org/10.1121/1.2951592).

Cardoso, L., Panda, R., Paiva, P. (2011). *MOODetector: A Prototype Software Tool for Mood-based Playlist Generation*. University of Coimbra. INForum 2011.

Casey, M., A., Veltkamp, R., Goto, M., Leman, M., Rhodes, C., Slaney, M. (2008). 'Content-based music information retrieval: Current directions and future challenges'. Proceedings of the IEEE. doi:[10.1109/jproc.2008.916370](https://doi.org/10.1109/jproc.2008.916370).

Chalupper, J., Fastl, H. (2002). *Dynamic loudness model for normal and hearing-impaired listeners*. Institute for Human-Machine Communication, Technical University of Munich. Munich, Germany.

Chang, C., Y., Lo, C., Y., Wang, C., J., Chung, P., C. (2010). *A music recommendation system with consideration of personal emotion*. IEEE. Department of Nursing, Institute of Computer and Communication Engineering. Taiwan.

- Collier, G. (2007). *Beyond valence and activity in the emotional connotations of music*. Psychology of Music. <https://doi.org/10.1177/0305735607068890>.
- Davis, S., Mermelstein, P. (1980). *Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences*. IEEE Trans. Acoustics, Speech & Signal Processing. University of Illinois.
- Ekman, P. (1992). *An argument for basic emotions*. Cognition and Emotion. American Psychological Association. <https://doi.org/10.1080/02699939208411068>.
- Feng, Y., Zhuang, Y., Pan, Y. (2003). *Popular music retrieval by detecting mood*. In Proc. ACM Int. Conf. Information Retrieval. Toronto, Canada. doi:[10.1145/860435.860508](https://doi.org/10.1145/860435.860508).
- Fernandes, J., Paiva, P. (2010). Automatic Playlist Generation via Music Mood Analysis. University of Coimbra, Faculty of Sciences and Technology Department of Informatics Engineering.
- Fujinaga, I., McMillan, K. (2000). *Real-time recognition of orchestral instruments*. In Proc. Int. Computer Music Conf.
- Gabrielsson, A. (2002). *Emotion perceived and emotion felt: Same or different?* Musicae Scientiae. doi:[10.1177/10298649020050s105](https://doi.org/10.1177/10298649020050s105).
- Gabrielsson, A., Lindstrom, E. (2001). *The influence of musical structure on emotional expression*. In P. N. Juslin and J. A. Sloboda, (Editors), Music and Emotion: Theory and Research. Oxford University Press, New York.
- Hallam, S., Cross, I., Thaut, M. (2008). The Oxford Handbook of Music Psychology, Oxford University Press, New York.
- Hargreaves, J., D., North, C., A. (1997). The Social Psychology of Music. Oxford University Press, Oxford, UK.
- Harte, C., Sandler, M., Gasser, M. (2006). *Detecting harmonic change in musical audio*. In Proc. ACM Workshop on Audio and Music Computing Multimedia. Santa Barbara, USA. doi:[10.1145/1178723.1178727](https://doi.org/10.1145/1178723.1178727).
- Hartmann, W., M. (1998). Signals, Sound, and Sensation. Springer, New York.
- Hevner., K. (1936). *Experimental studies of the elements of expression in music*. American Journal of Psychology. University of Illinois Press. <https://doi.org/10.2307/1415746>.
- Holbrook, M., B., Schindler R., M. (1989). *Some exploratory findings on the development of musical tastes*. J. Consumer Research. <https://doi.org/10.1086/209200>.
- Hu, X., Downie, J., S. (2007). *Exploring mood metadata: Relationships with genre, artist and usage metadata*. In Proc. Int. Conf. Music Information Retrieval. School of Information Sciences, National Center for Supercomputing Applications. Vienna, Austria.
- Hu, X., Downie, J., S., Laurier, C., Bay, M., Ehmann, A., F. (2008). *The 2007 MIREX audio mood classification task: lessons learned*. International Music Information Retrieval System Evaluation Laboratory University of Illinois at Urbana-Champaign.

- Huron., D. (2006). *Sweet Anticipation: Music and the Psychology of Expectation*. MIT Press, Cambridge, Massachusetts. doi:[10.7551/mitpress/6575.001.0001](https://doi.org/10.7551/mitpress/6575.001.0001).
- Jensen, K. (1999). Timbre models of musical sounds. Department of Computer Science, University of Copenhagen.
- Jiang, D., N., Lu, L., Zhang, H., J., Tao, J., H., Cai, L., H. (2002). *Music type classification by spectral contrast features*. In Proc. IEEE Int. Conf. Multimedia Expo. Department of Computer Science and Technology, Tsinghua University, China.
- Katayose, H., Imai, M., Inokuchi, S. (1998). *Sentiment extraction in music*. In Proc. Int. Conf. Pattern Recognition. Osaka University, Japan. doi: [10.1109/ICPR.1988.28447](https://doi.org/10.1109/ICPR.1988.28447).
- Klapuri, A. (1999). *Sound onset detection by applying psychoacoustic knowledge*. In Proc. Int. Conf. Acoustics, Speech, and Signal Processing. <https://doi.org/10.1109/ICASSP.1999.757494>.
- Laar, B. (2006). *Emotion detection in music, a survey*. 4th Twente Student Conference on IT. University of Twente, Faculty of Electrical Engineering, Mathematics and Computer Science.
- Lartillot. O., Toivainen, P. (2007). *MIR in MATLAB (II): A toolbox for musical feature extraction from audio*. In Proc. Int. Conf. Music Information Retrieval. University of Jyväskylä, Finland.
- Laurier, C., Grivolla, J., Herrera, P. (2008). *Multimodal music mood classification using audio and lyrics*. In Proc. Int. Conf. Machine Learning and Applications.
- Li, T., Ogihara, M., Li, Q. (2003). *A comparative study on content-based music genre classification*. In Proc. SIGIR, 2003. Toronto, Canada. doi:[10.1145/860435.860487](https://doi.org/10.1145/860435.860487).
- Li, T., Ogihara, M. (2004). *Content-based music similarity search and emotion detection*. In Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing.
- Li, T., Ogihara, M. (2006). *Toward intelligent music information retrieval*. IEEE Trans. Multi-media. doi:[10.1109/TMM.2006.870730](https://doi.org/10.1109/TMM.2006.870730).
- Lidy, T., Rauber, A. (2005). *Evaluation of feature extractors and psycho-acoustic transformations for music genre classification*. In Proc. Int. Conf. Music Information Retrieval. London, UK.
- Livingstone, S., R., Brown, A., R. (2005). *Dynamic response: A real-time adaptation for music emotion*. In Proc. Australian Conf. Interactive Entertainment. Sydney. doi:[10.13140/RG.2.1.3608.7841/1](https://doi.org/10.13140/RG.2.1.3608.7841/1).
- Liu, D., Lu, L., Zhang H., J. (2003). *Automatic music mood detection from acoustic music data*. In Proc. Int. Conf. Music Information Retrieval. Baltimore, Maryland.
- Liu, C., C., Yang, Y., H., Wu, P., H., Chen, H., H. (2006). *Detecting and classifying emotion in popular music*. In Proc. Joint Int. Conf. Information Sciences. Graduate Institute of Communication Engineering, National Taiwan University.
- Lu, L., Liu, D., Zhang, H. (2006). *Automatic mood detection and tracking of music audio signals*. IEEE Trans. Audio, Speech & Language Processing. China. doi: [10.1109/TSA.2005.860344](https://doi.org/10.1109/TSA.2005.860344).

- Magalhaes, T., N., Barros, F., B., Loureiro, M., A. (2019). *Iracema: a Python library for audio content analysis*. 17th Brazilian Symposium on Computer Music.
- McFee, B., Raffel, C., Liang, D., Ellis, D., McVicar, M., Battenberg, E., Nieto, O. (2015). *librosa: Audio and Music Signal Analysis in Python*. Proc. Of the 14th Python in Science Conf.
- Meng, A., Ahrendt, P., Larsen, J., Hansen, L., K. (2007). *Temporal feature integration for music genre classification*. IEEE Trans. Audio, Speech & Language Processing. Vaerloese, Denmark. doi:[10.1109/TASL.2007.899293](https://doi.org/10.1109/TASL.2007.899293).
- Oliveira, A., P., Cardoso, A. (2009). *Automatic manipulation of music to express desired emotions*. In Proc. Sound and Music Computing Conf. University of Coimbra, Portugal.
- Oliveira, A., P., Cardoso, A. (2008). *Controlling music affective content: A symbolic approach*. In Proc. Conf. Interdisciplinary Musicology.
- Osgood, C., E., Suci, G., J., Tannenbaum, P., H. (1957). *The measurement of meaning*. University of Illinois Press.
- Ovadia, S. (2004). *Ratings and rankings: Reconsidering the structure of values and their measurement*. International Journal of Social Research Methodology. University of Michigan. doi:[10.1080/1364557032000081654](https://doi.org/10.1080/1364557032000081654).
- Pampalk, E. (2004). *A MATLAB toolbox to compute music similarity from audio*. In Proc. Int. Conf. Music Information Retrieval. Barcelona, Spain.
- Pampalk, E., Rauber, A., Merkl, D. (2002). *Content-based organization and visualization of music archives*. In Proc. ACM Int. Conf. Multimedia. Juan les Pins, France. <https://doi.org/10.1145/641007.641121>.
- Panda, R., Malheiro, R., Rocha, B., Oliveira, A., Paiva, P. (2008). *Multi-Modal Music Emotion Recognition: A New Dataset, Methodology and Comparative Analysis*. University of Coimbra, Portugal. CISUC 2008.
- Panda, R., Paiva, P. (2011). *Automatic creation of mood playlists in the Thayer plane: a methodology and a comparative study*. University of Coimbra. CISUC 2011.
- Panda, R., Paiva, P. (2012). *Mirex 2012: mood classification tasks submission*. CISUC 2012.
- Panda, R., Paiva, P. (2012). *Music emotion classification: dataset acquisition and comparative analysis*, Proc. of the 15th Int. Conference on Digital Audio Effects (DAFx-12), York, UK.
- Parrott, W., G. (2001) *Emotions in Social Psychology: Essential Readings*. Psychology Press, USA.
- Peeters, G. (2004). *A large set of audio features for sound description (similarity and classification) in the CUIDADO project*. Technical report, IRCAM. Paris, France.
- Ricard, J. (2004). *Towards Computational Morphological Description of Sound*. Univ. Pompeu Fabra, Barcelona.
- Russell, J. (1980). *A circumplex model of affect*. Journal of Personality and Social Psychology. <https://doi.org/10.1037/h0077714>.

- Scherer., K. (2004). *Which emotions can be induced by music? what are the underlying mechanisms? and how can we measure them.* Journal New Music Research. doi:[10.1080/0929821042000317822](https://doi.org/10.1080/0929821042000317822).
- Schubert., E. (1999) Measurement and Time Series Analysis of Emotion in Music. School of Music Education University, New South Wales, Sydney, Australia. <https://doi.org/10.26190/unsworks/6566>.
- Schuller, B., Hage, C., Schuller, D., Rigoll., D. (2010). *Mister D.J., Cheer Me Up!’: Musical and textual features for automatic mood classification.* Journal of New Music Research. <https://doi.org/10.1080/09298210903430475>.
- Skowronek, J., McKinney, M., F., Van de Par, S. (2006). *Ground truth for automatic music mood classification.* In Proc. Int. Conf. Music Information Retrieval. University of Victoria.
- Sloboda, J., A., Juslin, P., N. (2001). *Music and Emotion: Theory and Research.* Oxford University Press, New York.
- Sloboda, J., A., O’Neill, A., S., Ivaldi, A. (2001). *Functions of music in everyday life: An exploratory study using the experience sampling methodology.* Musicae Scientiae. Department of Psychology, Keele University.
- Thayer, R. (1989). *The Biopsychology of Mood and Arousal.* Oxford University Press, New York.
- Thompson, W., F., Robitaille, B. (1992). *Can composers express emotions through music? Empirical Studies of the Arts.* Amityville, New York. <https://doi.org/10.2190/NBNY-AKDK-GW58-MTEL>.
- Tellegen, A., Watson, D., Clark., D. (1999). *On the dimensional and hierarchical structure of affect.* Psychological Science. <https://doi.org/10.1111/1467-9280.00157>.
- Tolonen, T., Karjalainen, M. (2000). *A computationally efficient multipitch analysis model.* IEEE Trans. Speech Audio Processing. Espoo, Finland.
- Tolos, M., Tato, R., Kemp, T. (2005). *Mood-based navigation through large collections of musical data.* In Proc. IEEE Consumer Communications & Network Conf.
- Trohidis, K., Tsoumakas, G., Kalliris, G., Vlahavas., I. (2008). *Multi-label classification of music into emotions.* In Proc. Int. Conf. Music Information Retrieval. Drexel University, Philadelphia, PA, USA. ISMIR 2008. doi:[10.1186/1687-4722-2011-426793](https://doi.org/10.1186/1687-4722-2011-426793).
- Tzanetakis, G., Cook, P. (2002). *Musical genre classification of audio signals.* IEEE Trans. Speech & Audio Processing. doi:[10.1109/TSA.2002.800560](https://doi.org/10.1109/TSA.2002.800560).
- Wang, M., Zhang, N., Zhu, H. (2004). *User-adaptive music emotion recognition.* In Proc. IEEE Int. Conf. Signal Processing. Beijing, China. doi: [10.1109/ICOSP.2004.1441576](https://doi.org/10.1109/ICOSP.2004.1441576).
- Watson, D., David Wiese, D., Vaidya, J., Tellegen, A. (1999). *The Two General Activation Systems of Affect: Structural Findings, Evolutionary Considerations, and Psychobiological Evidence.* Journal of Personality and Social Psychology.
- Wieczorkowska, A. (2006). *Towards extracting emotions from music.* In Proc. Int. Workshop on Intelligent Media Technology for Communicative Intelligence. Warsaw, Poland. doi:[10.1007/11558637_23](https://doi.org/10.1007/11558637_23).

Wieczorkowska, A., Synak, P., Lewis, R., A., Ras, Z., W. (2005). *Extracting emotions from music data*. In Proc. Int. Symp. Intelligent Systems. Saratoga Springs, New York, USA.

Wieczorkowska, A., Synak, P., Ras, Z., W. (2006). *Multi-label classification of emotions in music*. In Proc. Intelligent Information Processing and Web Mining. doi:[10.1007/3-540-33521-8_30](https://doi.org/10.1007/3-540-33521-8_30).

Wu, L., T., Jeng, K., S. (2008). *Probabilistic estimation of a novel music emotion model*. In Proc. Int. Multimedia Modeling Conf. Kyoto, Japan. doi:[10.1007/978-3-540-77409-9_46](https://doi.org/10.1007/978-3-540-77409-9_46).

Yang, Y., H., Chen, H., H. (2012). *Machine Recognition of Music Emotion: A Review*. ACM Trans. Intell. Syst. Technol. 3, 3, Article 40. National Taiwan University.
<https://doi.org/10.1145/2168752.2168754>.

Yang, D., Lee W., S. (2004). *Disambiguating music emotion using software agents*. In Proc. Int. Conf. Music Information Retrieval. Barcelona, Spain.

Yang, Y. H., Chen, H., H. (2011). *Music Emotion Recognition*. CRC Press, New York.

Yang, Y., H., Lin, Y., C., Chen, H., H. (2009). *Personalized music emotion recognition*. In Proc. ACM Int. Conf. Information Retrieval.

Yang, Y., H., Liu, C., C., Chen, H., H. (2006). *Music emotion classification: A fuzzy approach*. In Proc. ACM Int. Conf. Multimedia. <https://doi.org/10.1145/1180639.1180665>.

Yang, Y., H., Su, Y., F., Lin, Y., C., Chen, H., H. (2008). *A regression approach to music emotion recognition*. IEEE Trans. Audio, Speech & Language Processing. doi:[10.1109/TASL.2007.911513](https://doi.org/10.1109/TASL.2007.911513).

Yang, Y., H., Su, Y., F., Lin, Y., C., Chen, H., H. (2007). *Music emotion recognition: The role of individuality*. In Proc. ACM Int. Workshop on Human-Centered Multimedia. doi:[10.1145/1290128.1290132](https://doi.org/10.1145/1290128.1290132).

Zhu, X., Shi, Y., Y., Kim, H., G., Eom, K., W. (2006). *An integrated music recommendation system*. IEEE Transactions on Consumer Electronics. <https://doi.org/10.1109/TCE.2006.1706489>.

Zwicker, E. (1961). *Subdivision of the audible frequency range into critical bands*. J. Acoustical Society of America. <https://doi.org/10.1121/1.1908630>.

Zwicker, E., Fastl, H. (1999). *Psychoacoustics: Facts and Models*. Springer, New York.

Παράρτημα

Παρακάτω παρατίθενται τα Script της Python για την εξαγωγή των μουσικών παραμέτρων.

1. Chroma Stft

```

1 import librosa
2 import numpy as np
3 from scipy.stats import skew
4 from scipy.stats import kurtosis
5 y, sr = librosa.load('../../audio/romance/Αστικό - Δημήτρης Μπιτροπάνος-Όταν έχω εσένα.wav', sr=48000, mono=True, offset=0.0, duration=None)
6 chroma = librosa.feature.chroma_stft(y=y, sr=sr, n_fft = 1024, hop_length= 768, n_chroma=12)
7 chroma = chroma[0]
8 print(np.std(chroma))
9 print(np.mean(chroma))
10 print(np.var(chroma))
11 print(skew(chroma, nan_policy='omit'))
12 print(kurtosis(chroma, nan_policy='omit'))
13

```

2. MFCC

```

1 import librosa
2 import numpy as np
3 from scipy.stats import skew
4 from scipy.stats import kurtosis
5 y, sr = librosa.load('../../audio/angry/Metal - Arctic Monkeys-R U Mine.wav', sr=48000, mono=True, offset=0.0, duration=None)
6 mfcc = librosa.feature.mfcc(y=y, sr=sr, n_mfcc=8, hop_length=768, n_fft=1024)
7 print(np.std(mfcc))
8 print(np.mean(mfcc))
9 print(np.var(mfcc))
10 mfcc_0=mfcc[0,:]
11 mfcc_1=mfcc[1,:]
12 mfcc_2=mfcc[2,:]
13 mfcc_3=mfcc[3,:]
14 mfcc_4=mfcc[4,:]
15 mfcc_5=mfcc[5,:]
16 mfcc_6=mfcc[6,:]
17 mfcc_7=mfcc[7,:]
18 print('=====')
19 print(skew(mfcc_0, axis=0, bias=True))
20 print(skew(mfcc_1, axis=0, bias=True))
21 print(skew(mfcc_2, axis=0, bias=True))
22 print(skew(mfcc_3, axis=0, bias=True))
23 print(skew(mfcc_4, axis=0, bias=True))
24 print(skew(mfcc_5, axis=0, bias=True))
25 print(skew(mfcc_6, axis=0, bias=True))
26 print(skew(mfcc_7, axis=0, bias=True))
27 print('=====')
28 print(kurtosis(mfcc_0, axis=0, fisher=True, bias=True))
29 print(kurtosis(mfcc_1, axis=0, fisher=True, bias=True))
30 print(kurtosis(mfcc_2, axis=0, fisher=True, bias=True))
31 print(kurtosis(mfcc_3, axis=0, fisher=True, bias=True))
32 print(kurtosis(mfcc_4, axis=0, fisher=True, bias=True))
33 print(kurtosis(mfcc_5, axis=0, fisher=True, bias=True))
34 print(kurtosis(mfcc_6, axis=0, fisher=True, bias=True))
35 print(kurtosis(mfcc_7, axis=0, fisher=True, bias=True))

```

3. RMS

```

1 import librosa
2 import numpy as np
3 from scipy.stats import skew
4 from scipy.stats import kurtosis
5 y, sr = librosa.load('../../audio/chill/Rock - Radiohead-No Surprises.wav', sr=48000, mono=True, offset=0.0, duration=None)
6 rms = librosa.feature.rms(y=y, hop_length=768, frame_length=1024)
7 rms = rms[0]
8 print(np.std(rms))
9 print(np.mean(rms))
10 print(np.var(rms))
11 print(skew(rms, nan_policy='omit'))
12 print(kurtosis(rms, nan_policy='omit'))

```

4. Spectral contrast

```

1 import librosa
2 import numpy as np
3 from scipy.stats import skew
4 from scipy.stats import kurtosis
5 y, sr = librosa.load('../../audio/chill/Rock - Radiohead-No Surprises.wav', sr=48000, mono=True, offset=0.0, duration=None)
6 S = np.abs(librosa.stft(y))
7 contrast = librosa.feature.spectral_contrast(S=S, sr=sr, n_bands=6)
8 print(np.std(contrast))
9 print(np.mean(contrast))
10 print(np.var(contrast))
11 contrast_0=contrast[0,:]
12 contrast_1=contrast[1,:]
13 contrast_2=contrast[2,:]
14 contrast_3=contrast[3,:]
15 contrast_4=contrast[4,:]
16 contrast_5=contrast[5,:]
17 print('=====')
18 print(skew(contrast_0, axis=0, bias=True))
19 print(skew(contrast_1, axis=0, bias=True))
20 print(skew(contrast_2, axis=0, bias=True))
21 print(skew(contrast_3, axis=0, bias=True))
22 print(skew(contrast_4, axis=0, bias=True))
23 print(skew(contrast_5, axis=0, bias=True))
24 print('=====')
25 print(kurtosis(contrast_0, axis=0, fisher=True, bias=True))
26 print(kurtosis(contrast_1, axis=0, fisher=True, bias=True))
27 print(kurtosis(contrast_2, axis=0, fisher=True, bias=True))
28 print(kurtosis(contrast_3, axis=0, fisher=True, bias=True))
29 print(kurtosis(contrast_4, axis=0, fisher=True, bias=True))
30 print(kurtosis(contrast_5, axis=0, fisher=True, bias=True))

```

5. Spectral Centroid

```

1 import librosa
2 import numpy as np
3 from scipy.stats import skew
4 from scipy.stats import kurtosis
5 y, sr = librosa.load('../../audio/dance/Pop - Madonna-Like A Prayer.wav', sr=48000, mono=True, offset=0.0, duration=None)
6 cent = librosa.feature.spectral_centroid(y=y, sr=sr, hop_length=768, win_length=1024)
7 cent = cent[0]
8 print(np.std(cent))
9 print(np.mean(cent))
10 print(np.var(cent))
11 print(skew(cent, nan_policy='omit'))
12 print(kurtosis(cent, nan_policy='omit'))

```

6. Spectral Flatness

```

1 import librosa
2 import numpy as np
3 from scipy.stats import skew
4 from scipy.stats import kurtosis
5 y, sr = librosa.load('../../audio/focus/Κλασσική - Beethoven-Tempo di Menuetto.wav', sr=48000, mono=True, offset=0.0, duration=None)
6 flatness = librosa.feature.spectral_flatness(y=y, hop_length=768, win_length=1024)
7 flatness = flatness[0]
8 print(np.std(flatness))
9 print(np.mean(flatness))
10 print(np.var(flatness))
11 print(skew(flatness, nan_policy='omit'))
12 print(kurtosis(flatness, nan_policy='omit'))

```

7. Zero Crossing

```

1 import librosa
2 import numpy as np
3 from scipy.stats import skew
4 from scipy.stats import kurtosis
5 y, sr = librosa.load('../../audio/sad/Pop - Adele-Easy On Me.wav', sr=48000, mono=True, offset=0.0, duration=None)
6 librosa.feature.zero_crossing_rate(y)
7 z = librosa.zero_crossings(y)
8 print(np.std(z))
9 print(np.mean(z))
10 print(np.var(z))
11 print(skew(z, axis=0, bias=True))
12 print(kurtosis(z, axis=0, bias=True))
13

```

8. Spectral Rolloff

```

1 import librosa
2 import numpy as np
3 from scipy.stats import skew
4 from scipy.stats import kurtosis
5 y, sr = librosa.load('../../audio/happy/Disco - Elton John Britney Spears-Hold Me Closer.wav', sr=48000, mono=True, offset=0.0, duration=None)
6 librosa.feature.spectral_rolloff(y=y, sr=sr)
7 rolloff = librosa.feature.spectral_rolloff(y=y, sr=sr, hop_length=768, win_length=1024)
8 rolloff = rolloff[0]
9 print(np.std(rolloff))
10 print(np.mean(rolloff))
11 print(np.var(rolloff))
12 print(skew(rolloff, nan_policy='omit'))
13 print(kurtosis(rolloff, nan_policy='omit'))
14

```

9. Spectral Bandwidth

```
1 import librosa
2 import numpy as np
3 from scipy.stats import skew
4 from scipy.stats import kurtosis
5 y, sr = librosa.load('../audio/angry/Rnb - Justin Timberlake-Cry Me A River.wav', sr=48000, mono=True, offset=0.0, duration=None)
6 spec_bw = librosa.feature.spectral_bandwidth(y=y, sr=sr, hop_length=768, win_length=1024)
7 spec_bw = spec_bw[0]
8 print(np.std(spec_bw))
9 print(np.mean(spec_bw))
10 print(np.var(spec_bw))
11 print(skew(spec_bw, nan_policy='omit'))
12 print(kurtosis(spec_bw, nan_policy='omit'))
13
14
```

Υπεύθυνη Δήλωση Συγγραφέα:

Δηλώνω ρητά ότι, σύμφωνα με το άρθρο 8 του Ν.1599/1986, η παρούσα εργασία αποτελεί αποκλειστικά προϊόν προσωπικής μου εργασίας, δεν προσβάλλει κάθε μορφής δικαιώματα διανοητικής ιδιοκτησίας, προσωπικότητας και προσωπικών δεδομένων τρίτων, δεν περιέχει έργα/εισφορές τρίτων για τα οποία απαιτείται άδεια των δημιουργών/δικαιούχων και δεν είναι προϊόν μερικής ή ολικής αντιγραφής, οι πηγές δε που χρησιμοποιήθηκαν περιορίζονται στις βιβλιογραφικές αναφορές και μόνον και πληρούν τους κανόνες της επιστημονικής παράθεσης.