



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ
ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΤΕΧΝΟΛΟΓΙΑΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ
ΥΠΟΛΟΓΙΣΤΩΝ

Προς το Ψηφιακό Μουσικό Αναλόγιο:

αλγόριθμος εντοπισμού γραμμών πενταγράμμου και διαστολών

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Μαριάννα Γ. Τζωρμπατζάκη

Επιβλέπων: Στέφανος Κόλλιας
Καθηγητής Ε.Μ.Π.

Αθήνα, Ιούλιος 2012



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ
ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΤΕΧΝΟΛΟΓΙΑΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ
ΥΠΟΛΟΓΙΣΤΩΝ

Προς το Ψηφιακό Μουσικό Αναλόγιο:

αλγόριθμος εντοπισμού γραμμών πενταγράμμου και διαστολών

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Μαριάννα Γ. Τζωρμπατζάκη

Επιβλέπων: Στέφανος Κόλλιας
Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 19^η Ιουλίου 2012.

.....
Στέφανος Κόλλιας
Καθηγητής Ε.Μ.Π.

.....
Ανδρέας- Γεώργιος
Σταφυλοπάτης
Καθηγητής Ε.Μ.Π.

.....
Τατιάνα Ταμπουρατζή
Αν. Καθηγήτρια
Πανεπιστημίου
Πειραιά

Αθήνα, Ιούλιος 2012

.....

Μαριάννα Γ. Τζωρμπατζάκη

Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

Copyright© Μαριάννα Γ. Τζωρμπατζάκη, 2012

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

Περίληψη

Στην παρούσα διπλωματική εργασία προτείνεται ένας νέος αλγόριθμος εντοπισμού των γραμμών πενταγράμμου και των διαστολών σε μουσικά κείμενα για πιάνο. Κύριος σκοπός του εντοπισμού είναι να αποτελέσει το προπαρασκευαστικό στάδιο της υλοποίησης ενός ψηφιακού μουσικού αναλογίου, που θα μπορέσει να εφαρμοστεί σε χειρόγραφες, σαρωμένες έντυπες ή ψηφιακά παραχθείσες παρτιτούρες.

Ο αλγόριθμος εστιάζεται στην ελάχιστη επεξεργασία. Στο πλαίσιο αυτό, χρησιμοποιεί απλές τεχνικές επεξεργασίας της εικόνας του μουσικού κειμένου, συνδυασμένες με τέτοιο τρόπο ώστε να επιτυγχάνεται το βέλτιστο αποτέλεσμα αναγνώρισης. Ακολουθώντας τη μέθοδο hill climbing, βρίσκει και συμπληρώνει σταδιακά όλες τις γραμμές του πενταγράμμου που αποτελούν τη βάση της παρτιτούρας, διορθώνοντας και εμπλουτίζοντας το αποτέλεσμα σε κάθε στάδιο της διαδικασίας. Μετά τον εντοπισμό των πενταγράμμων και την ομαδοποίησή τους σε συστήματα, ο αλγόριθμος προχωρά στον εντοπισμό των διαστολών. Η αναγνώριση θα αποτελέσει καθοριστικό παράγοντα στην παρουσίαση της παρτιτούρας κατά την υλοποίηση του ψηφιακού μουσικού αναλογίου, καθιστώντας δυνατή την παραμετροποιήσιμη απεικόνιση μουσικών παρτιτούρων στην οθόνη του χρήστη.

Ο αλγόριθμος λειτουργεί με μη παρεμβατικό τρόπο ως προς την εικόνα του μουσικού κειμένου ώστε να εξασφαλίζει κατάλληλη είσοδο για το ψηφιακό μουσικό αναλόγιο. Οι σελίδες του μουσικού κειμένου υφίστανται επεξεργασία μία προς μία, χωρίς την απαίτηση ευθυγράμμισης της κάθε σελίδας ή της αφαίρεσης των γραμμών του πενταγράμμου, προτείνοντας με αυτό τον τρόπο μια εναλλακτική μέθοδο Οπτικής Αναγνώρισης Μουσικής.

Λέξεις Κλειδιά

Οπτική Αναγνώριση Μουσικής, επεξεργασία εικόνας, μουσικό κείμενο, παρτιτούρα πιάνου, μέθοδος hill climbing, εντοπισμός γραμμών πενταγράμμου, μη απομάκρυνση γραμμών πενταγράμμου, οριζόντια προβολή

Abstract

The current thesis puts forward a novel methodology for staff and bar line detection of piano music manuscripts. The main aim of the methodology is the construction of an efficient as well as accurate preparatory stage of the digital music stand that can be uniformly applied to handwritten, scanned and born-digital music manuscripts.

The methodology is focused upon minimum processing. In this context, simple image processing techniques are combined in such a way as to obtain the best possible result concerning bar segmentation and staff line identification. By constructing a repetitive hill climbing approach, the methodology gradually detects all the staff lines and uncovers all the staves, correcting and enhancing the result at every repetition. After staff line and staff identification is completed, the staves are grouped into systems. This is followed by bar line detection, whereby the bars of the music manuscript are identified. This constitutes the key input to the presentation side of the digital music rest, which can then proceed with on-line custom-made on-screen music manuscript presentation.

The algorithm operates page-by-page, without any pre-processing regarding the music manuscript. By not requiring manuscript alignment or staff line removal, the digital music stand puts forward an alternative Optical Music Recognition methodology.

KeyWords

Optical Music Recognition, image processing, music sheet, piano sheet, hill climbing method, staff line tracing, non-removal of staff lines, horizontal projection

Ευχαριστίες

Θα ήθελα να ευχαριστήσω τον επιβλέποντα καθηγητή κ. Στέφανο Κόλλια, για την εμπιστοσύνη που μου έδειξε κατά την ανάθεση αυτής της εργασίας, την ευκαιρία που μου έδωσε να ασχοληθώ με ένα τόσο ενδιαφέρον θέμα, καθώς και για τα εποικοδομητικά ερωτήματά του που με βοήθησαν να συνθέσω μια πιο ολοκληρωμένη περιγραφή του αλγορίθμου αναζήτησης γραμμών.

Ιδιαίτερα θα ήθελα να ευχαριστήσω την καθηγήτρια κ. Τατιάνα Ταμπουρατζή, για τη συνολική της καθοδήγηση πάνω στο θέμα της διπλωματικής, η οποία βασίστηκε σε δική της ιδέα, καθώς και στις εργασίες της πάνω στο αντικείμενο αλλά και στον κώδικά της. Επίσης, για τη βοήθεια και την ενθάρρυνση που μου προσέφερε καθ' όλη τη διάρκεια της υλοποίησης, αλλά και για τα αναλυτικά σχόλια και διορθώσεις κατά τη διάρκεια της συγγραφής. Η επιστημονική προοπτική που μου άνοιξε και το ενδιαφέρον που μου έδειξε είναι για μένα πολύ σημαντικά.

Στο ρουλάκι και στο τιτάκι

Περιεχόμενα

Κεφάλαιο 1 ^ο	19
Εισαγωγή.....	21
1.1 Το τονικό μουσικό σύστημα και η γραφική αναπαράσταση των ήχων	21
1.2 Γραφική αναπαράσταση της μουσικής.....	23
1.2.1. Το πεντάγραμμο	23
1.2.2 Τα κλειδιά.....	24
1.2.3 Αλλοιώσεις και οπλισμός.....	25
1.2.4 Τα μέτρα και ο χρόνος.....	27
1.3 Η παρτιτούρα του πιάνου.....	29
1.4 Πλαίσιο εργασίας.....	32
Κεφάλαιο 2 ^ο	37
Βιβλιογραφική ανασκόπηση.....	37
2.1 Γενικό πλαίσιο.....	37
2.2 Σύνοψη υπάρχουσών προσεγγίσεων αναγνώρισης γραμμών του πενταγράμμου- Σχετική βιβλιογραφία.....	38
Κεφάλαιο 3 ^ο	43
Περιγραφή του αλγορίθμου	43
3.1 Συνοπτική περιγραφή	43
3.2 Τα στάδια του αλγορίθμου	48
3.2.1 Στάδιο 1 ^ο : Προ-επεξεργασία εικόνας εισόδου	48
3.2.2 Στάδιο 2 ^ο : Αρχική εκτίμηση των γραμμών του πενταγράμμου	52
3.2.3: Στάδιο 3 ^ο : Συμπλήρωση γραμμών εντοπισμένων πενταγράμμων.....	60
3.2.4: Στάδιο 4 ^ο : Εντοπισμός ολόκληρων πενταγράμμων.....	72
3.2.5 Στάδιο 5 ^ο : Αναθεώρηση κατωφλίου.....	94
3.2.6 Στάδιο 6 ^ο : Αναδιάταξη εντοπισμένων γραμμών	99
3.2.7 Στάδιο 7 ^ο : Απομάκρυνση μη πραγματικών πενταγράμμων.....	106
3.2.8: Στάδιο 8 ^ο : Εντοπισμός διαστολών- Χωρισμός μέτρων	113
Κεφάλαιο 4 ^ο	118
Αποτελέσματα, Συμπεράσματα και Μελλοντική εξέλιξη	118
4.1 Αποτελέσματα και παρατηρήσεις.....	118
4.2 Συμπεράσματα	124
4.3 Μελλοντικές εξελίξεις	124
Βιβλιογραφία	125

Κατάλογος Εικόνων

Εικόνα 1. 1: Η αντιστοίχιση των νοτών στα λευκά πλήκτρα του πιάνου.	21
Εικόνα 1. 2: Σύνολο πλήκτρων πιάνου (κλαβιέ).	22
Εικόνα 1. 3: Δύο γειτονικά πρότυπα του πάνω στο πιάνο.	22
Εικόνα 1. 4: Οι αξίες των νοτών.	23
Εικόνα 1. 5: Παρεστιγμένες αξίες.	23
Εικόνα 1. 6: Το πεντάγραμμο.	24
Εικόνα 1. 7: Το κλειδί του σολ.	24
Εικόνα 1. 8: Απεικόνιση των διαδοχικών φθόγγων και των θέσεων τους στο κλειδί του σολ.	24
Εικόνα 1. 9: Το κλειδί του φα.	25
Εικόνα 1. 10: Απεικόνιση διαδοχικών φθόγγων και των θέσεων τους στο κλειδί του φα.	25
Εικόνα 1. 11: Από αριστερά προς τα δεξιά, πριν από κάθε νότα: δίεση, ύφεση, αναίρεση.	26
Εικόνα 1. 12: Σειρά διέσεων σε κλειδί σολ (επάνω) και κλειδί φα (κάτω).	26
Εικόνα 1. 13: Σειρά υφέσεων σε κλειδί σολ (επάνω) και κλειδί φα (κάτω).	26
Εικόνα 1. 14: Παράδειγμα χρήσης αλλοιώσεων στον οπλισμό.	27
Εικόνα 1.15: Δύο διαστολές μέσα στο πεντάγραμμο σχηματίζουν ένα μέτρο.	28
Εικόνα 1. 16: Χρόνος κομματιού.	28
Εικόνα 1. 17: Δείγμα παρτιτούρας πιάνου.	29
Εικόνα 1. 18: Απόσταση μεταξύ δύο διαδοχικών γραμμών του πενταγράμμου.	31
Εικόνα 1. 19: scanned παρτιτούρα-born digital παρτιτούρα.	32
Εικόνα 1. 20: Παράδειγμα scanned παρτιτούρας καλής ποιότητας.	33
Εικόνα 1. 21: Στιγμιότυπο λειτουργίας του Sibelius 7 demo version.	34
Εικόνα 1. 22: Επίδειξη λειτουργία Photoscore Lite demo version.	35
Εικόνα 2. 1: Απόσπασμα παρτιτούρας πριν και μετά από την αφαίρεση γραμμών.	38
Εικόνα 3. 1: Συνοπτικό διάγραμμα ροής του προτεινόμενου αλγορίθμου εντοπισμού γραμμών.	47
Εικόνα 3. 2: Διάγραμμα ροής σταδίου προ-επεξεργασίας εικόνας.	48
Εικόνα 3. 3: Επίδειξη περικοπής της ασπρόμαυρης εικόνας του μουσικού κειμένου.	50
Εικόνα 3. 4: Παράδειγμα περικοπής εικόνας.	50
Εικόνα 3. 5: Επίδειξη προβληματικής λειτουργίας υψηλού κατωφλίου σε παρτιτούρα με μη ενιαίο υπόβαθρο.	51
Εικόνα 3. 6: Αποτέλεσμα κατωφλίωσης με χρήση χαμηλότερου κατωφλίου για τη μετατροπή της εικόνας εισόδου σε ασπρόμαυρη.	52
Εικόνα 3. 7: Διάγραμμα ροής σταδίου αρχικής εκτίμησης γραμμών.	52
Εικόνα 3. 8: Αρχική εκτίμηση των γραμμών.	54
Εικόνα 3. 9: Δείγματα αρχικής εκτίμησης γραμμών.	55

Εικόνα 3. 10: Οριζόντια προβολή των εικόνων που εμφανίζονται στην Εικόνα 3.9.	56
Εικόνα 3. 11: Αποτέλεσμα περιορισμού πολλαπλών γραμμών.	57
Εικόνα 3. 12: Εντοπισμός μη πραγματικής γραμμής πενταγράμμου.	58
Εικόνα 3. 13: Η αρχική εκτίμηση γραμμών είναι αρκετή για τον υπολογισμό της staff-line-distance, παρά την παράλειψη πολλών πενταγράμμων.	59
Εικόνα 3. 14: Διάγραμμα ροής σταδίου συμπλήρωσης γραμμών.	63
Εικόνα 3. 15: Παραδείγματα γραμμών που λείπουν ενδιάμεσα ήδη εντοπισμένων γραμμών.	64
Εικόνα 3. 16: Δύο διαφορετικά παραδείγματα μοναδικής εντοπισμένης γραμμής σε πεντάγραμμο.	65
Εικόνα 3. 17: Έλλειψη μιας και τριών γραμμών αντίστοιχα από το άνω άκρο του πενταγράμμου.	65
Εικόνα 3. 18: Έλλειψη τριών γραμμών.	65
Εικόνα 3. 19: Παράδειγμα συμπλήρωσης μίας ενδιάμεσης γραμμής.	67
Εικόνα 3. 20: Συμπλήρωση μίας ενδιάμεσης γραμμής και συμπλήρωση τεσσάρων γραμμών από μοναδική γραμμή του πενταγράμμου.	67
Εικόνα 3. 21: Συμπλήρωση τεσσάρων γραμμών από μοναδική γραμμή του πενταγράμμου.	68
Εικόνα 3. 22: Συμπλήρωση δύο συνεχόμενων ενδιάμεσων γραμμών πενταγράμμου.	68
Εικόνα 3. 23: Συμπλήρωση μοναδικής γραμμής από το άνω άκρο του πενταγράμμου.	69
Εικόνα 3. 24: Συμπλήρωση τριών γραμμών στα άκρα του πενταγράμμου, δύο πάνω και ένα κάτω και συμπλήρωση δύο γραμμών στο κάτω άκρο του πενταγράμμου.	69
Εικόνα 3. 25: Συμπλήρωση πρώτης και πέμπτης γραμμής του πενταγράμμου.	69
Εικόνα 3. 26: Συμπλήρωση τριών διαδοχικών γραμμών στο πάνω μέρος του πενταγράμμου.	70
Εικόνα 3. 27: Διάγραμμα ροής σταδίου εντοπισμού πενταγράμμων.	73
Εικόνα 3. 28: Έλλειψη ενός πενταγράμμου από το κάτω μέρος της εικόνας - Έλλειψη του δεύτερου πενταγράμμου του πρώτου συστήματος.	74
Εικόνα 3. 29: Έλλειψη μεγάλου αριθμού πενταγράμμων.	75
Εικόνα 3. 30: Διαδοχικά στιγμιότυπα τελευταίου συστήματος παρτιτούρας.	81
Εικόνα 3. 31: Εμφάνιση διαδοχικών στιγμιότυπων του ίδιου αποσπάσματος της παρτιτούρας. Τοποθέτηση επιπλέον πενταγράμμου.	81
Εικόνα 3. 32: Διαδοχικά στιγμιότυπα εντοπισμού πενταγράμμων.	82
Εικόνα 3. 33: Δύο στιγμιότυπα έλλειψης και εντοπισμού.	82
Εικόνα 3. 34: Διαδοχικά στιγμιότυπα εντοπισμού γραμμών πενταγράμμου.	83
Εικόνα 3. 35: Εντοπισμός πενταγράμμων.	83
Εικόνα 3. 36: Εντοπισμός πενταγράμμων.	84
Εικόνα 3. 37: Διαδοχικά στιγμιότυπα εντοπισμού πενταγράμμων.	85
Εικόνα 3. 38: Εντοπισμός μη πραγματικού πενταγράμμου και στη συνέχεια διόρθωση.	86
Εικόνα 3. 39: Περιττός αριθμός πενταγράμμων.	87

Εικόνα 3. 40: Διάγραμμα ροής διαδικασίας περιορισμού υπεράριθμων γραμμών.....	89
Εικόνα 3. 41: Τρία διαδοχικά στιγμιότυπα του ίδιου πενταγράμμου, καθώς προχωράει η εκτέλεση του αλγορίθμου.....	92
Εικόνα 3. 42: Τέσσερα διαδοχικά στιγμιότυπα των βημάτων του αλγορίθμου..	93
Εικόνα 3. 43: Τρία διαδοχικά στιγμιότυπα του ίδιου πενταγράμμου:.....	93
Εικόνα 3. 44: Δύο διαδοχικά στιγμιότυπα του ίδιου πενταγράμμου.	94
Εικόνα 3. 45: Επίδειξη σταδίου αναθεώρησης κατωφλίου, για διαφορετικά στιγμιότυπα του ίδιου μουσικού κειμένου.	96
Εικόνα 3. 46: Συνοπτική επίδειξη αναθεώρησης κατωφλίου.	96
Εικόνα 3. 47: Το αποτέλεσμα της τρίτης αναθεώρησης του κατωφλίου δίνει τελικά μια σωστή εκτίμηση γραμμών.....	97
Εικόνα 3. 48: Επίδειξη επίδρασης κλίσης στην τοποθέτηση των γραμμών.....	98
Εικόνα 3. 49: Επίδειξη επίδρασης κλίσης	98
Εικόνα 3. 50: Διάγραμμα ροής σταδίου αναδιάταξης γραμμών.	99
Εικόνα 3. 51: Δύο διακριτές ομάδες γραμμών εντός του πενταγράμμου.....	100
Εικόνα 3. 52: Το αποτέλεσμα της επιδιόρθωσης των θέσεων των γραμμών του πενταγράμμου.	101
Εικόνα 3. 53: Επίδειξη επιδιόρθωσης θέσης γραμμών πενταγράμμου.	101
Εικόνα 3. 54: Αναδιάταξη γραμμών πενταγράμμου.	102
Εικόνα 3. 55: Το πρόβλημα της μη κατάταξης των γραμμών του ίδιου πενταγράμμου στο ίδιο πεντάγραμμο, παρά το αυστηρότερο κατώφλι.....	103
Εικόνα 3. 56: Η επίλυση του προβλήματος μέσω του ελέγχου αν διαδοχικά ελλιπή «πεντάγραμμο» παρουσιάζουν άθροισμα γραμμών ίσο με πέντε.	103
Εικόνα 3. 57: Δημιουργία μη πραγματικού πενταγράμμου.	104
Εικόνα 3. 58: Ο περιορισμός των υπεράριθμων γραμμών του πενταγράμμου απομακρύνει το μη πραγματικό πεντάγραμμο.	104
Εικόνα 3. 59: Διάγραμμα ροής σταδίου απομάκρυνσης μη πραγματικών πενταγράμμων.	107
Εικόνα 3. 60: Μη πραγματικό πεντάγραμμο, μετά την ολοκλήρωση του σταδίου αναδιάταξης των γραμμών.....	108
Εικόνα 3. 61: Δύο μη πραγματικά πεντάγραμμο, εξαιτίας του τίτλου και των σχολίων.....	108
Εικόνα 3. 62: Μη πραγματικό πεντάγραμμο ενδιαμέσως αναγνωρισμένων πενταγράμμων.	108
Εικόνα 3. 63: Δύο μη πραγματικά πεντάγραμμο.....	109
Εικόνα 3. 64: Η ακολουθία των αποστάσεων mid-staff-distance και mid-system-distance.	110
Εικόνα 3. 65: Παράδειγμα περιορισμού τριών μη πραγματικών πενταγράμμων.	112
Εικόνα 3. 66: Παράδειγμα περιορισμού δύο πενταγράμμων από την άνω πλευρά της εικόνας.....	112
Εικόνα 3. 67: Παράδειγμα περιορισμού δύο πενταγράμμων σε τυχαίες θέσεις.	113
Εικόνα 3. 68: Διάγραμμα ροής σταδίου εντοπισμού διαστολών.....	114

Εικόνα 3. 69: Δύο παραδείγματα γραμμών από νότες με ίδια εμφάνιση με τις διαστολές.....	115
Εικόνα 3. 70: Τελικό αποτέλεσμα αλγορίθμου, απεικονίζονται και θέσεις των διαστολών.....	117
Εικόνα 3. 71: Τελικό αποτέλεσμα αλγορίθμου, με εμφάνιση των θέσεων των διαστολών αλλά και των εντοπισμένων γραμμών.	117

Κατάλογος Πινάκων

Πίνακας 4. 1: Παρουσίαση χαρακτηριστικών συνόλου παρτιτούρων.....	119
Πίνακας 4. 2: Στατιστικά συνόλου παρτιτούρων σχετικά με πλήθος γραμμών και διαστολών ανά σελίδα.....	120
Πίνακας 4. 3: Αποτελέσματα εντοπισμού γραμμών και πενταγράμμων ανά μουσικό κομμάτι.....	121

Κεφάλαιο 1°

Εισαγωγή

Για την κατανόηση του περιβάλλοντος εργασίας για την υλοποίηση του προτεινόμενου αλγορίθμου, περιβάλλον το οποίο περιλαμβάνει κυρίως το μουσικό κείμενο (παρτιτούρα), κρίνεται απαραίτητη μια σύντομη εισαγωγή στη θεωρία και τη σημειογραφία της μουσικής, καθώς και στο υπάρχον πλαίσιο μέσα στο οποίο κινείται η προτεινόμενη υλοποίηση. Με αυτό τον τρόπο, ο αναγνώστης, ανεξάρτητα από τις μουσικές του γνώσεις, θα είναι σε θέση να κατανοήσει το πλαίσιο λειτουργίας καθώς και τα προβλήματα που ζητούν επίλυση κατά τη διαδικασία της προτεινόμενης υλοποίησης.

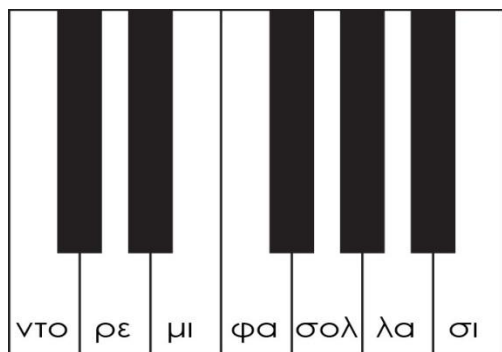
1.1 Το τονικό μουσικό σύστημα και η γραφική αναπαράσταση των ήχων

Ο όρος τονικό ή Ευρωπαϊκό μουσικό σύστημα χρησιμοποιείται για να περιγράψει το μουσικό σύστημα που άρχισε να καθιερώνεται από το 1700 και χρησιμοποιείται ευρέως μέχρι σήμερα. [1]

Οι μουσικοί φθόγγοι (ή νότες) είναι για τη μουσική ό,τι είναι η αλφάβητος για τη γλώσσα, δηλαδή τα βασικά δομικά στοιχεία της. Οι νότες είναι επτά και έχουν τα παρακάτω ονόματα:

ντο ρε μι φα σολ λα σι

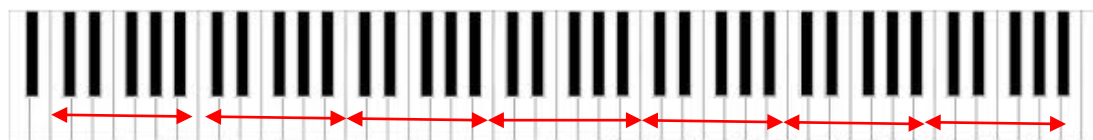
Η αντιστοίχιση τους στα πλήκτρα του πιάνου, φαίνεται ακολούθως:



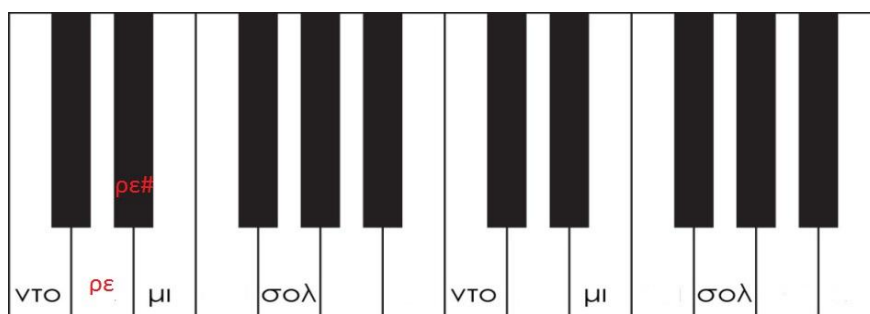
Εικόνα 1. 1: Η αντιστοίχιση των νοτών στα λευκά πλήκτρα του πιάνου.

Το πρότυπο που απεικονίζεται στην Εικόνα 1.1 παρουσιάζει επτά λευκά πλήκτρα του πιάνου (ένα για κάθε νότα) τα οποία είναι τοποθετημένα το ένα δίπλα στο άλλο, με τα μαύρα πλήκτρα να τοποθετούνται ανάμεσα τους ανά δύο και ανά τρία, αντίστοιχα. Αυτό το πρότυπο επαναλαμβάνεται επτά φορές στο σύνολο των πλήκτρων του πιάνου, όπως φαίνεται στην παρακάτω εικόνα (Εικόνα 1.2). Μαζί με τρία πλήκτρα στη χαμηλή έκταση (λα, λα δίεση, σι) και ένα επιπλέον στην υψηλή έκταση (ντο), δημιουργείται ένα σύνολο ογδόντα οκτώ (88) πλήκτρων. Τα πλήκτρα που έχουν την ίδια θέση σε δύο πρότυπα (βλ.

Εικόνα 1.3) στο πιάνο αντιπροσωπεύουν την ίδια νότα, π.χ. το λευκό πλήκτρο που είναι τοποθετημένο πριν την ομάδα των δύο μαύρων πλήκτρων είναι πάντα η νότα ντο.



Εικόνα 1. 2: Σύνολο πλήκτρων πιάνου (κλαβιέ). Παρατηρείται η επανάληψη του προτύπου που παρουσιάστηκε παραπάνω. [2]



Εικόνα 1. 3: Δύο γειτονικά πρότυπα του πάνω στο πιάνο. Παρατηρείται ότι τα πλήκτρα που βρίσκονται στην ίδια θέση σε δύο διαφορετικά πρότυπα αντιπροσωπεύουν την ίδια νότα (πχ. η ντο βρίσκεται πάντα πριν από την ομάδα των δύο μαύρων πλήκτρων, η μι μετά από αυτή την ομάδα, η σολ είναι το πέμπτο λευκό πλήκτρο (αν θεωρήσουμε ότι η σειρά ξεκινάει από το ντο) κλπ. Επιπλέον φαίνεται η νότα ρε, ενδιάμεσα της ντο και της μι καθώς και η ρε δίεση, το μαύρο πλήκτρο που ακολουθεί τη νότα ρε.

Οι νότες που αντιστοιχούν στα λευκά πλήκτρα του πιάνου ονομάζονται φυσικές νότες, ενώ αυτές που αντιστοιχούν στα μαύρα πλήκτρα ονομάζονται αλλοιωμένες. Δύο διαδοχικά πλήκτρα (λευκό- μαύρο, μαύρο- λευκό, ή λευκό- λευκό στις δύο περιπτώσεις μι-φα και σι-ντο όπου δεν παρεμβάλλεται μαύρο πλήκτρο) αντιπροσωπεύουν το διάστημα ημιτονίου, που είναι το μικρότερο διάστημα που μπορούν να απέχουν μεταξύ τους δύο νότες ως προς το ύψος συχνότητας [1]. Η συχνότητα των νοτών που αντιστοιχούν σε διαδοχικά πλήκτρα του πιάνου αυξάνεται κατά ένα ημιτόνιο από τα αριστερά προς τα δεξιά, συμπεριλαμβάνοντας τα μαύρα πλήκτρα στην ακολουθία των πλήκτρων. Προκειμένου να γίνει κατανοητή η σχέση συχνότητας, ορολογίας/σήμανσης και νοτών θα χρησιμοποιηθεί ως παράδειγμα η νότα ρε: Η νότα που βρίσκεται ένα ημιτόνιο ψηλότερα από τη νότα ρε (δηλαδή στο μαύρο πλήκτρο αμέσως δεξιά από το ρε, βλ. Εικόνα 1.3), αναφέρεται ως ρε δίεση (ρε #) ενώ όταν γίνεται αναφορά στη νότα που βρίσκεται ένα ημιτόνιο χαμηλότερα από τη νότα ρε (δηλαδή στο μαύρο πλήκτρο αμέσως αριστερά από το ρε), η νότα αναφέρεται ως ρε ύφεση (ρε \flat). Γίνεται κατανοητό ότι, με τη χρήση αυτού του μηχανισμού, κάθε νότα μπορεί να έχει δύο ονόματα, π.χ. η νότα ντο δίεση ταυτίζεται με το ρε ύφεση. Στον κανόνα αυτό υπάρχουν δύο εξαιρέσεις, και αυτές αφορούν στις νότες μι-φα και σι-ντο. Οι νότες αυτές απέχουν μεταξύ τους διάστημα ημιτονίου (και γι' αυτό δεν παρεμβάλλεται μεταξύ τους μαύρο πλήκτρο), επομένως δε νοείται η έννοια πχ. μι δίεση (ή αντίστοιχα φα ύφεση), καθώς το μι δίεση είναι

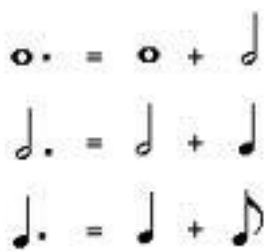
στην πραγματικότητα η νότα φα (και το φα ύφεση είναι στην πραγματικότητα η νότα μι). Ακριβώς το ίδιο ισχύει και για τις νότες σι-ντο, που επίσης απέχουν μεταξύ τους διάστημα ημιτονίου.

Η διάρκεια των νοτών αποδίδεται με διαφορετικά σύμβολα, που ονομάζονται αξίες. Για κάθε αξία νότας (μουσικού φθόγγου) υπάρχει και μια αντίστοιχη αξία παύσης, ενός συμβόλου δηλαδή που υποδεικνύει ότι κατά τη διάρκειά του δεν παίζεται καμία νότα. Συνοπτικά, οι αξίες (τόσο για τις νότες όσο και για τις παύσεις) είναι: ολόκληρο (1), μισό (1/2), όγδοο (1/8), δέκατο έκτο (1/16), τριακοστό δεύτερο (1/32) και εξηκοστό τέταρτο (1/64). Οι αξίες απεικονίζονται στην παρακάτω εικόνα (Εικόνα 1.4):



Εικόνα 1. 4: Οι αξίες των νοτών. Από αριστερά: ολόκληρο, μισό, τέταρτο, όγδοο, δέκατο έκτο, τριακοστό δεύτερο, εξηκοστό τέταρτο [3]

Μια τελεία τοποθετημένη αμέσως μετά από κάποια αξία νότας ή κάποια παύση, τη μεγαλώνει κατά το μισό της. Οι αξίες των νοτών ή των παύσεων που έχουν μεγαλώσει με αυτό τον τρόπο λέγονται παρεστιγμένες [1]:



Εικόνα 1. 5: Παρεστιγμένες αξίες [4]

1.2 Γραφική αναπαράσταση της μουσικής

1.2.1. Το πεντάγραμμο

Στη μουσική σημειογραφία, οι νότες, τοποθετούνται μέσα στο πεντάγραμμο, αλλά και επάνω (εκτός ορίων του) καθώς και κάτω (εκτός των ορίων του) από αυτό με την βοήθεια επιπλέον βοηθητικών γραμμών. Το πεντάγραμμο, όπως δηλώνει και το όνομά του, αποτελείται από πέντε διαδοχικές γραμμές τοποθετημένες σε ίσες μεταξύ τους αποστάσεις, καθώς και από τα διαστήματα που βρίσκονται ανάμεσα τους (βλ. Εικόνα 1.6). Οι νότες μπορούν να τοποθετηθούν είτε επάνω στις γραμμές είτε στα διαστήματα ανάμεσα σε δύο διαδοχικές γραμμές.



Εικόνα 1. 6: Το πεντάγραμμο [3]

Η αναφορά στις γραμμές του πενταγράμμου γίνεται από κάτω προς τα πάνω, δηλαδή ως πρώτη γραμμή χαρακτηρίζεται η γραμμή που βρίσκεται στο κάτω μέρος του πενταγράμμου, ως δεύτερη η ακριβώς από πάνω, κοκ.

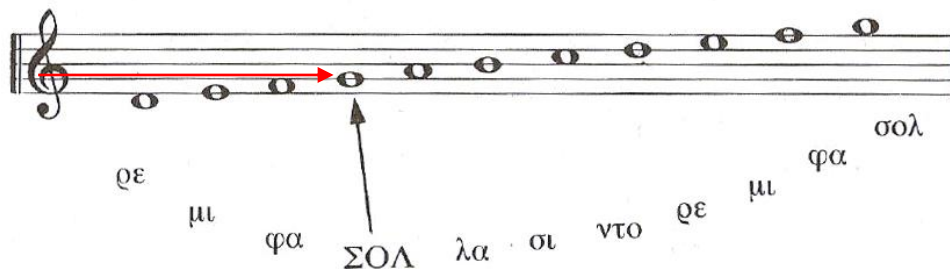
1.2.2 Τα κλειδιά

Η συχνότητα των νοτών που απεικονίζονται πάνω στο πεντάγραμμο (δηλαδή σε ποια θέση πρέπει να παιχτεί η κάθε νότα πάνω στο πιάνο) καθορίζεται από το κλειδί το οποίο εμφανίζεται στο αριστερό άκρο του πενταγράμμου. Δύο είδη κλειδιών απαντώνται στις παρτιτούρες πιάνου, το κλειδί του σολ και το κλειδί του φα.



Εικόνα 1. 7: Το κλειδί του σολ [3]

Το κλειδί του σολ χρησιμοποιείται για μελωδίες οι οποίες πρέπει να ακουστούν στην υψηλή έκταση (ή υψηλότερες συχνότητες) του πιάνου. Η θέση του κλειδιού πάνω στο πεντάγραμμο είναι προκαθορισμένη: το άκρο της καμπύλης (που σχηματίζει το κλειδί του σολ) που βρίσκεται στο εσωτερικό του κλειδιού πρέπει να ακουμπά στη δεύτερη γραμμή του πενταγράμμου υποδηλώνοντας ότι οποιαδήποτε αξίας νότα τοποθετηθεί πάνω σε αυτή τη γραμμή είναι η νότα σολ.



Εικόνα 1. 8: Απεικόνιση των διαδοχικών φθόγγων και των θέσεων τους στο κλειδί του σολ [1]

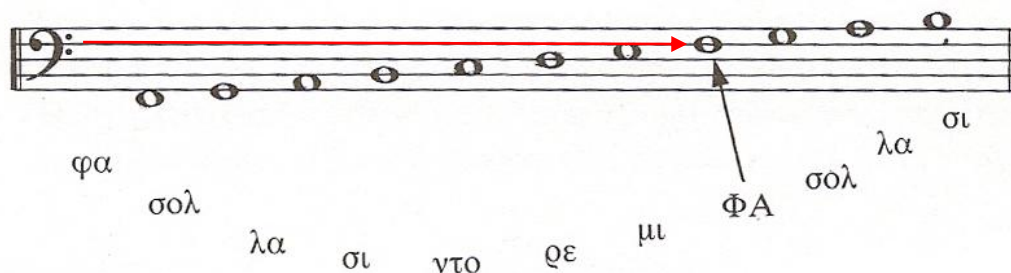
Όπως φαίνεται στην παραπάνω εικόνα (Εικόνα 1.8), οι νότες που τοποθετούνται εκατέρωθεν της νότας σολ πάνω στο πεντάγραμμο ακολουθούν τη σειρά των

νοτών που προαναφέρθηκε. Έτσι, πάνω από τη νότα σολ τοποθετείται η νότα λα, ενώ κάτω η νότα φα, κοκ.



Εικόνα 1. 9: Το κλειδί του φα [3]

Το κλειδί του φα χρησιμοποιείται για μελωδίες οι οποίες πρέπει να ακουστούν στην χαμηλή έκταση (ή χαμηλότερες συχνότητες) του πιάνου. Η θέση στην οποία τοποθετείται πάνω στο πεντάγραμμο είναι και πάλι προκαθορισμένη, δηλαδή το άνω άκρο της καμπύλης που σχηματίζει το κλειδί πρέπει να ακουμπά στην τέταρτη γραμμή του πενταγράμμου υποδηλώνοντας ότι νότα οποιασδήποτε αξίας τοποθετηθεί πάνω στην ίδια γραμμή είναι η νότα φα. Εναλλακτικά, οι δύο τελείες που μοιάζουν με άνω και κάτω τελεία και συνοδεύουν πάντοτε το κλειδί του φα πρέπει να βρίσκονται εκατέρωθεν της τέταρτης γραμμής του πενταγράμμου.



Εικόνα 1. 10: Απεικόνιση διαδοχικών φθόγγων και των θέσεων τους στο κλειδί του φα [1]

Γίνεται κατανοητό με αυτό τον τρόπο και με άμεση σύγκριση των εικόνων 1.8 και 1.10 ότι, για διαφορετικά κλειδιά, οι νότες στις οποίες αντιστοιχούν οι ίδιες θέσεις πάνω στο πεντάγραμμο είναι διαφορετικές. Για παράδειγμα, η νότα που γράφεται στην πρώτη γραμμή σε κλειδί σολ είναι η μι και στο πρώτο διάστημα είναι η φα, ενώ σε κλειδί φα η νότα που γράφεται στην πρώτη γραμμή είναι η σολ και στο πρώτο διάστημα η λα. Το κλειδί δηλαδή είναι αυτό που καθορίζει σε τι συχνότητα αντιστοιχεί κάθε νότα πάνω στο πεντάγραμμο. [5]

1.2.3 Αλλοιώσεις και οπλισμός

Η συχνότητα κάποιων από τις νότες που τοποθετούνται μέσα στο μουσικό κείμενο μπορεί να μεταβληθεί κατά ένα ημιτόνιο προς τα πάνω ή προς τα κάτω με τη χρήση των αλλοιώσεων. Υπάρχουν τρία είδη αλλοιώσεων, η δίεση, η ύφεση και η αναίρεση (βλ. Εικόνα 1.11), τα οποία μπορεί να έχουν είτε καθολική (μόνιμη) είτε τοπική δράση.

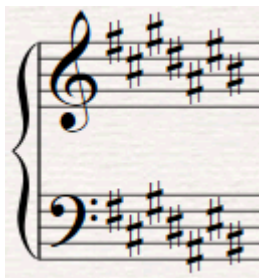


Εικόνα 1. 11: Από αριστερά προς τα δεξιά, πριν από κάθε νότα: διέση, ύφεση, αναίρεση [3]

Όταν μία ή περισσότερες αλλοιώσεις τοποθετούνται στην αρχή του πενταγράμμου αμέσως μετά το κλειδί, ονομάζονται οπλισμός και έχουν καθολική/μόνιμη δράση . Ο οπλισμός:

(α) Συμβολίζεται μέσω αλλοιώσεων (είτε διέσεων είτε υφέσεων) οι οποίες τοποθετούνται στις θέσεις που αντιστοιχούν στις νότες στις οποίες εφαρμόζονται (βλέπε Εικόνα 1.12). Στη γενική περίπτωση, οι αλλοιώσεις μπορούν να φτάσουν τις επτά διέσεις ή τις επτά υφέσεις. Ο συνδυασμός διέσεων και υφέσεων στον οπλισμό δεν επιτρέπεται.

(β) Εμφανίζεται στην παρτιτούρα μέσω προκαθορισμένης σειράς των αλλοιώσεων (διέσεων ή υφέσεων), συγκεκριμένα φα ντο σολ ρε λα μι σι για τις διέσεις:



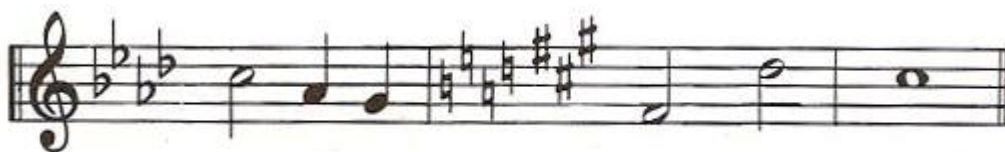
Εικόνα 1. 12: Σειρά διέσεων σε κλειδί σολ (επάνω) και κλειδί φα (κάτω) [6]

και σι μι λα ρε σολ ντο φα για τις υφέσεις (σειρά αντίστροφη αυτής των διέσεων):



Εικόνα 1. 13: Σειρά υφέσεων σε κλειδί σολ (επάνω) και κλειδί φα (κάτω) [6]

(γ) Αλλοιώνει κατά ένα ημιτόνιο την συχνότητα όλων των νοτών μέσα στο μουσικό κομμάτι που αντιστοιχούν στην θέση της κάθε αλλοίωσης, ανεξάρτητα από το ύψος στο οποίο βρίσκονται οι αντίστοιχες νότες. Αναιρέσεις μπορούν να χρησιμοποιηθούν στον οπλισμό μόνο στην περίπτωση αλλαγής του οπλισμού κατά τη διάρκεια του κομματιού, όπως φαίνεται στην παρακάτω εικόνα (Εικόνα 1.14). Παρατηρείται ότι, προκειμένου να τοποθετηθεί ο νέος οπλισμός των τριών διέσεων, απαιτείται να αναιρεθούν οι υφέσεις του αρχικού οπλισμού. Οι αναιρέσεις τοποθετούνται στις αντίστοιχες θέσεις των υφέσεων που ακυρώνονται.



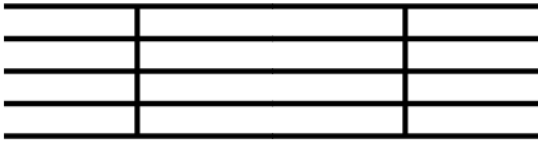
Εικόνα 1. 14: Παράδειγμα χρήσης αλλοιώσεων στον οπλισμό. Ο αρχικός οπλισμός τεσσάρων υφέσεων σι μι λα ρε, υποδηλώνει ότι οποιοδήποτε σι, μι, λα ή ρε εμφανιστεί στο μουσικό κείμενο θα είναι αλλοιωμένο κατά ένα ημιτόνιο προς τα κάτω, πχ. η δεύτερη νότα του πρώτου μέτρου (λα) ενώ εμφανίζεται ως φυσική, θα πρέπει να παιχτεί ένα ημιτόνιο κάτω (ως λα ύφεση), εξαιτίας του οπλισμού. Αυτό παύει να ισχύει μετά την αλλαγή του οπλισμού στο δεύτερο μέτρο, μετά από το οποίο όλα τα φα, ντο ή σολ που εμφανίζονται στο μουσικό κείμενο πρέπει να παιχτούν αλλοιωμένα κατά ένα ημιτόνιο προς τα πάνω. Για παράδειγμα, το φα του δεύτερου μέτρου, αν και δε συνοδεύεται από δίεση, θα πρέπει να παιχτεί ένα ημιτόνιο προς τα πάνω, αφού η δίεση που θα έπρεπε να υπάρχει είναι ήδη τοποθετημένη στον οπλισμό. [1]

Όταν μία αλλοίωση τοποθετείται μπροστά από μία νότα (βλ. Εικόνα 1. 11), έχει τοπική δράση. Έτσι, η δίεση υψώνει την συχνότητα της νότας μπροστά από την οποία τοποθετείται κατά ένα ημιτόνιο, η ύφεση την χαμηλώνει κατά ένα ημιτόνιο, και η αναίρεση ακυρώνει οποιοδήποτε σύμβολο αλλοίωσης της νότας μπροστά από την οποία τοποθετείται. [1]

1.2.4 Τα μέτρα και ο χρόνος¹

Μέσω των παραπάνω μουσικών συμβόλων (κλειδιά, νότες, παύσεις, αξίες και αλλοιώσεις), ένα μουσικό κομμάτι μπορεί να εκτελεστεί με ακρίβεια. Με την προσθήκη ερμηνευτικών σχολίων και τονισμών (οι τονισμοί αφορούν σε νότες που πρέπει να παιχτούν περισσότερο τονισμένες σε σχέση με τις γειτονικές τους), ο ερμηνευτής έχει μια πλήρη εικόνα του ύφους που πρέπει να ακολουθήσει για τη σωστή εκτέλεση του κομματιού. Επιπλέον, για την καλύτερη οργάνωση και την ευκολότερη ανάγνωση της παρτιτούρας, έχει καθιερωθεί η υποδιαίρεση του κομματιού σε ισόποσα χρονικά διαστήματα, τα μέτρα. Κάθε μέτρο (α) διαχωρίζεται από τα προηγούμενα και τα επόμενα μέτρα στο πεντάγραμμο μέσω κάθετων γραμμών οι οποίες διανύουν το πεντάγραμμο και ονομάζονται διαστολές (Εικόνα 1.15) και (β) έχει ίση χρονική διάρκεια με κάθε άλλο μέτρο μέσα στο ίδιο μουσικό κείμενο.

¹ Η ελληνική ορολογία χρησιμοποιεί επίσης και τον όρο «μέτρο», η χρήση του οποίου αποφεύγεται



Εικόνα 1.15: Δύο διαστολές μέσα στο πεντάγραμμο σχηματίζουν ένα μέτρο (κέντρο) [3]

Για να είναι δυνατό το μέτρημα μέσα σε κάθε μέτρο καθώς και σε όλο το μουσικό κομμάτι, πρέπει να υπάρχει μια απόλυτα ισόποση διαίρεση του χρόνου, σε αντιστοιχία με τη διαίρεση του χρόνου σε ώρες, λεπτά, δευτερόλεπτα κ.λπ. Στη μουσική, η διαίρεση αυτή γίνεται με μια υποθετική ροή ισόχρονων χτύπων, που ονομάζονται παλμοί. Οι παλμοί μπορεί να είναι γρήγοροι ή αργοί, πρέπει οπωσδήποτε όμως να είναι ισόχρονοι, έτσι ώστε να επιτρέπουν τη διαίρεση του χρόνου σε ισόποσα διαστήματα. Υπ' αυτή την έννοια, τα μέτρα του κομματιού αποτελούν και μέσο ομαδοποίησης των παλμών σε ισόποσες ομάδες, με τέτοιο τρόπο ώστε η πραγματική διάρκεια του κομματιού να μπορεί να μετρηθεί με ακρίβεια. [1]

Ως βάση για τη γραφική αναπαράσταση των παλμών του μέτρου χρησιμοποιείται κάποια από τις αξίες νοτών που προαναφέρθηκαν (ολόκληρο, μισό, τέταρτο κλπ.). Αφενός ο αριθμός των παλμών που περιέχεται μέσα σε κάθε μέτρο και αφετέρου ο αριθμός που αντιπροσωπεύει το είδος της αξίας που επιλέχθηκε για να αναπαραστήσει τον κάθε παλμό, τοποθετούνται μαζί σε μορφή κλάσματος στην αρχή του κομματιού. Το κλάσμα αυτό ονομάζεται χρόνος του κομματιού, και έχει σαν αριθμητή των αριθμό των παλμών ανά μέτρο και σαν παρονομαστή την αξία των παλμών (Εικόνα 1. 16).



Εικόνα 1. 16: Χρόνος κομματιού. Υποδηλώνει ότι η αξία των παλμών του κομματιού είναι το τέταρτο, και σε κάθε μέτρο βρίσκονται τρεις παλμοί. Με άλλα λόγια σε κάθε μέτρο εμφανίζονται νότες συνολικής αξίας τριών τετάρτων. [3]

Υπάρχουν δύο περιπτώσεις χρόνων, των οποίων ο συμβολισμός δεν ακολουθεί

τον παραπάνω κανόνα. Αυτά είναι: το σύμβολο \mathbf{C} [7] που αντιπροσωπεύει τον

χρόνο 4/4 σε μέτρημα τεσσάρων παλμών και το σύμβολο \mathbf{C} [8] που αντιπροσωπεύει και πάλι το χρόνο 4/4 αλλά σε μέτρημα δύο παλμών, δηλαδή το χρόνο 2/2.

1.3 Η παρτιτούρα του πιάνου

Συμπληρωματικά με τις προαναφερθείσες γενικές αρχές που ισχύουν για κάθε παρτιτούρα, η παρτιτούρα του πιάνου εμφανίζει κάποια επιπλέον χαρακτηριστικά ως προς τη δομή και το περιεχόμενο. Τα στοιχεία αυτά (μαζί με τα βασικά σύμβολα) εμφανίζονται εποπτικά στην παρακάτω παρτιτούρα (βλ. Εικόνα 1.17): Το προτεινόμενο σύστημα έχει υλοποιηθεί για την επεξεργασία αποκλειστικά παρτιτούρων για πιάνο, καθώς αυτές εμφανίζουν μια πιο περίπλοκη μορφή σε σχέση με άλλα μουσικά όργανα (πχ. βιολί, κιθάρα). Επίσης γίνεται παρουσίαση και των μεταβλητών που θα χρησιμοποιηθούν κατά την υλοποίηση του αλγορίθμου εντοπισμού γραμμών:

The image shows a musical score for 'Arlequin' by Schumann, specifically the piano accompaniment. The score is written for two staves (treble and bass clef) and is titled 'Arlequin' with the tempo 'Vivo M.M. $\text{♩} = 96$ '. The score is annotated with red circles and numbers 1 through 10, highlighting specific musical features. Annotations include: 1. A red circle around the first measure of the bass staff. 2. A red circle around the second measure of the bass staff. 3. A red circle around the third measure of the bass staff. 4. A red circle around the tempo marking 'Vivo M.M. $\text{♩} = 96$ '. 5. A red circle around the fifth measure of the bass staff. 6. A red circle around the sixth measure of the bass staff. 7. A red circle around the seventh measure of the bass staff. 8. A red circle around the eighth measure of the bass staff. 9. A red circle around the ninth measure of the bass staff. 10. A red circle around the tenth measure of the bass staff. There are also red arrows pointing to specific notes and dynamics like 'p', 'sf', and 'β'. The score is divided into three systems, with the first system containing measures 1-7, the second system containing measures 8-9, and the third system containing measure 10.

Εικόνα 1. 17: Δείγμα παρτιτούρας πιάνου, απόσπασμα από σελίδα 1, Schumann Carnival [9]

α) Σύστημα: ζεύγος δύο πενταγράμμων που συνδέονται μέσω μιας αγκύλης. Αποτελεί τη μορφή εμφάνισης των πενταγράμμων στις παρτιτούρες πιάνου, όπου το επάνω πεντάγραμμο αναφέρεται στο δεξί χέρι του ερμηνευτή ενώ το κάτω στο αριστερό. Ο ερμηνευτής παίζει ταυτόχρονα με τα δύο χέρια διαβάζοντας από το αντίστοιχο πεντάγραμμο για κάθε χέρι.

β) Απόσταση μεταξύ δύο πενταγράμμων του ίδιου συστήματος (mid-staff-distance). Αυτή η απόσταση εμφανίζει μικρές διακυμάνσεις μέσα στο μουσικό κομμάτι που αφορούν στη δομή της παρτιτούρας (μέγεθος των

μουσικών συμβόλων και των παρατηρήσεων τα οποία βρίσκονται ανάμεσα στα δύο πεντάγραμματα του συστήματος), αλλά δεν σχετίζεται με κανέναν τρόπο με το μουσικό κείμενο ή με την εκτέλεση του κομματιού. Είναι χρήσιμη για την εύρεση των γραμμών πενταγράμμου που είναι απαραίτητη για την υλοποίηση του προτεινόμενου αλγορίθμου, ως ένα μέτρο της υπάρχουσας απόστασης μεταξύ των πενταγράμμων.

γ) Απόσταση μεταξύ δύο διαδοχικών συστημάτων (mid-system-distance). Ομοίως με τη mid-staff-distance, η απόσταση αυτή δε σχετίζεται με το μουσικό κείμενο. Είναι χρήσιμη για την υλοποίηση της αναζήτησης των γραμμών πενταγράμμου, ως ένα μέτρο της υπάρχουσας απόστασης μεταξύ των πενταγράμμων. Η υπέρβαση του μέτρου αυτού αποτελεί ένδειξη για πιθανά πεντάγραμματα που έχουν παραμείνει μη εντοπισμένα από τον αλγόριθμο.

Τέλος, τα επόμενα χαρακτηριστικά που ακολουθούν, με αρίθμηση από το 1 έως το 10 αποτελούν στοιχεία της μουσικής σημειογραφίας που σχετίζονται άμεσα με τον τρόπο εκτέλεσης του μουσικού κομματιού:

1. Κλειδιά σολ και φα: ο ερμηνευτής διαβάζει ταυτόχρονα σε κλειδί σολ από το πάνω πεντάγραμμο και σε κλειδί φα από το κάτω πεντάγραμμο.
2. Χρόνος του κομματιού: ο βασικός παλμός που χρησιμοποιείται είναι το τέταρτο και υπάρχουν τρεις παλμοί σε κάθε μέτρο. Παράλληλα το άθροισμα των αξιών των νοτών σε κάθε μέτρο είναι ίσο με τρία τέταρτα.
3. Οπλισμός του κομματιού: δύο υφέσεις, υποχρεωτικά οι δύο πρώτες στην καθορισμένη σειρά των υφέσεων σι και μι. Όλες οι νότες σι και μι που υπάρχουν στο κομμάτι σε όλες τις συχνότητες (δηλαδή ανεξάρτητα από τη θέση του προτύπου στο οποίο ανήκουν πάνω στο πιάνο) πρέπει να παιχτούν ένα ημιτόνιο χαμηλότερα, εκτός κι αν υπάρχει τοπική αναίρεση.
4. Ύψος του κομματιού: η ένδειξη *Vivo* (συνήθως όλες οι ενδείξεις παρόμοιου τύπου είναι στα ιταλικά και κάποιες φορές στα γαλλικά) σημαίνει ότι το κομμάτι πρέπει να παιχτεί σε ζωηρό ύψος. Συμπληρωματικά, το 96 που φαίνεται ακριβώς δίπλα και ακολουθεί μια νότα συνολικής αξίας τριών τετάρτων (μισού παρεσιγμένου, βλ. Εικόνα 1.5) ορίζει την ταχύτητα με την οποία πρέπει να παιχτεί το κομμάτι σε παλμούς ανά λεπτό, όπου ο κάθε παλμός διαρκεί τρία τέταρτα. Δηλαδή σε κάθε επόμενο χτύπο του παλμού (με 96 παλμούς ανά λεπτό) πρέπει να παίζεται ένα ολόκληρο μέτρο.

5,6. Διαστολές που ορίζουν ένα μέτρο

7. Ένδειξη για την ένταση με την οποία πρέπει να παιχτεί το συγκεκριμένο σημείο του μουσικού κειμένου (συντομογραφία από τα ιταλικά). Τέτοιες ενδείξεις εμφανίζονται με μεγάλη συχνότητα στις παρτιτούρες προκειμένου να καθοδηγήσουν τον ερμηνευτή στο κατάλληλο ύψος ερμηνείας του κομματιού (πχ. *piano*, *forte*, *crescendo*, *diminuendo* κλπ).

8. Αλλαγή κλειδιού. Στο σημείο αυτό, και ενώ ο εκτελεστής διάβαζε το αριστερό χέρι (κάτω πεντάγραμμο) σε κλειδί φα, εμφανίζεται ένα κλειδί σολ που σημαίνει ότι όλες οι νότες που βρίσκονται δεξιά του νέου κλειδιού θα πρέπει να διαβαστούν στις θέσεις του κλειδιού σολ. Στην πράξη αυτό σημαίνει ότι οι νότες που εμφανίζονται στο κάτω πεντάγραμμο, εξακολουθούν μεν να αφορούν στο αριστερό χέρι, αλλά βρίσκονται σε ψηλότερη συχνότητα, δηλαδή πιο δεξιά πάνω στο κλαβιέ πιάνου (υψηλή έκταση του πιάνου, ή υψηλότερες συχνότητες). Το νέο κλειδί έχει ισχύ μέχρις ότου αντικατασταθεί από κάποιο κλειδί φα (εάν κάτι τέτοιο συμβεί) μέχρι το τέλος του κομματιού. Μέσα στο μουσικό κείμενο μπορεί να υπάρχουν πολλαπλές αλλαγές κλειδιών.

9. Εκ νέου αλλαγή κλειδιού. Το προηγούμενο σε ισχύ κλειδί (σολ) στο κάτω πεντάγραμμο αντικαθίσταται από κλειδί φα, επομένως όλες οι νότες που ακολουθούν το νέο κλειδί διαβάζονται στις θέσεις του κλειδιού φα. Όπως παρατηρείται, το νέο κλειδί φα συνεχίζει να ισχύει και στα επόμενα συστήματα της παρτιτούρας (μέχρι νέας αλλαγής κλειδιού).

10. Σημείο έναρξης επανάληψης. Η διπλή γραμμή με τις άνω και κάτω τελείες στα δεξιά υποδηλώνει ότι το κομμάτι του μουσικού κειμένου που ξεκινά από αυτό το σημείο και ολοκληρώνεται στο σημείο όπου εμφανίζονται αντίστοιχες άνω και κάτω τελείες και διπλή γραμμή πρέπει να επαναληφθεί κατά την εκτέλεση. Με άλλα λόγια, τα σημεία επανάληψης περικλείουν το υπό επανάληψη μουσικό κείμενο, το οποίο πρέπει να παιχτεί δύο φορές συνεχόμενα. Στην εικόνα δεν εμφανίζεται το μέτρο στο οποίο τελειώνει η επανάληψη.

Το μοναδικό μορφολογικό χαρακτηριστικό ιδιαίτερης σημασίας για την υλοποίηση του ψηφιακού μουσικού αναλογίου στο οποίο δεν έχει γίνει ακόμα αναφορά είναι η απόσταση μεταξύ δύο διαδοχικών γραμμών εντός ενός πενταγράμμου (*staff-line-distance*). Η απόσταση αυτή διατηρείται σταθερή κατά κανόνα σε όλη την έκταση του κομματιού, δεδομένου ότι το μέγεθος των νοτών δε μεταβάλλεται μέσα στο ίδιο μουσικό κείμενο (βλ. Εικόνα 1.18).



Εικόνα 1. 18: Απόσταση μεταξύ δύο διαδοχικών γραμμών του πενταγράμμου, Schumann *Carnaval*, σελίδα 1 [9]. Η *staff line distance* είναι η απόσταση μεταξύ των δύο τελευταίων (πάνω) γραμμών που κάτω πενταγράμμου και επισημαίνεται με το βέλος.

Γίνεται κατανοητό ότι η staff-line-distance, επαναλαμβάνεται τέσσερις φορές στο ύψος ενός πενταγράμμου, αφού μεταξύ των πέντε γραμμών υπάρχουν τέσσερις ίσες μεταξύ τους κενές αποστάσεις.

1.4 Πλαίσιο εργασίας

Το επιστημονικό πλαίσιο μέσα στο οποίο τοποθετείται η προτεινόμενη υλοποίηση, θα αναλυθεί ξεχωριστά στο επόμενο κεφάλαιο. Στην ενότητα αυτή θα παρουσιαστούν τα σχετικά (με την υλοποίηση) προγράμματα που κυκλοφορούν στην αγορά και χρήσιμοι όροι για την κατανόηση του περιβάλλοντος και του αντικειμένου του αλγορίθμου αναζήτησης γραμμών, καθώς και της θέσης του μέσα στο υπάρχον πλαίσιο.

Πριν την παρουσίαση του σκοπού των προγραμμάτων, είναι χρήσιμο να επισημανθεί μια σημαντική διάκριση μεταξύ του είδους των παρτιτούρων: Όλα τα υπάρχοντα συστήματα, του προτεινόμενου μη εξαιρετέου, δέχονται σαν είσοδο μια παρτιτούρα σε μορφή εικόνας ή αρχείου, δηλαδή ένα αρχείο ψηφιακή μορφή. Ο τρόπος με τον οποίο έχει παραχθεί αυτή η ψηφιακή μορφή επιβάλλει τη διάκριση των παρτιτούρων σε δύο κατηγορίες:

- αυτές που έχουν προκύψει από σάρωση (scanned) ενός χάρτινου πρωτότυπου, και
- αυτές που δημιουργήθηκαν με ψηφιακό τρόπο από κάποιο πρόγραμμα (born digital)

Μια σύγκριση μεταξύ scanned και born digital παρτιτούρας για πιάνο, φαίνεται στο παρακάτω παράδειγμα (Εικόνα 1.19):

The image displays two side-by-side examples of musical notation for piano. On the left is a scanned page of Schumann's 'Novelletes', page 2, showing five systems of music with a yellowish, aged paper texture and some ink bleed-through. On the right is a born digital page of Beethoven's 'Elise', page 1, showing six systems of music with a clean, white background and sharp, clear notation. The contrast highlights the differences in paper quality and digital clarity between the two methods.

Εικόνα 1. 19: Αριστερά: scanned παρτιτούρα- είναι εμφανής ο χρωματισμός των σελίδων λόγω ποιότητας χαρτιού και παλαιότητας καθώς και μια μικρή κλίση των πενταγράμμων, Schumann Novelletes, σελίδα 2, [10]. Δεξιά: born digital παρτιτούρα, εξαιρετικά καθαρή και ευκρινής, Beethoven Elise, σελίδα 1, [11].

Φυσικά δεν είναι η ποιότητα όλων των scanned παρτιτούρων ίδια με αυτήν που απεικονίζεται στην παραπάνω εικόνα (Εικόνα 1.19), αλλά υπάρχουν scanned παρτιτούρες καλύτερης ποιότητας (βλ. Εικόνα 1.20):

Two Cadenzas for Beethoven's Piano Concerto in G Major, Op. 58 free-scores.com

Two Cadenzas
for Beethoven's
Piano Concerto in G Major, Op. 58

First Movement

1

Εικόνα 1. 20: Παράδειγμα scanned παρτιτούρας καλής ποιότητας, Brahms Cadenzas, σελίδα 1, [12]

Καθώς οι scanned παρτιτούρες είναι απλώς το αποτέλεσμα της σάρωσης της τυπωμένης έκδοσης του μουσικού κομματιού, εμφανίζουν (εκτός από τα βασικά χαρακτηριστικά τους) και κάποια επιπρόσθετα που έχουν προκύψει από τον τρόπο συντήρησης του πρωτότυπου ή από τον τρόπο σάρωσης. Για παράδειγμα, η σάρωση μιας πρωτότυπης έκδοσης (πχ. του 1800) είναι πιθανό να εμφανίζει σημάδια που προέκυψαν λόγω παλαιότητας, να έχει φθαρμένο χαρτί (βλ. Εικόνα 1.19), να έχει ελαττωμένη ένταση γραμμών και νοτών μέσα στο μουσικό κείμενο, ή ακόμα να έχει αποκτήσει κλίση λόγω απροσεξίας στη σάρωση. Αυτό δημιουργεί μια σειρά προβλημάτων στην εικόνα του μουσικού κειμένου, όπως η ύπαρξη κατακερματισμένων γραμμών ή γραμμών μη ενιαίου πάχους.

Από την άλλη πλευρά, οι born digital παρτιτούρες, καθώς έχουν προκύψει με ψηφιακό τρόπο, θα μπορούσαν να χαρακτηριστούν ιδανικές από την άποψη της ποιότητας, καθώς τα χρώματα τους είναι άρτια, δεν υπάρχει θόρυβος στην εικόνα, ούτε κάποια φθορά στην ποιότητα της. Ο αριθμός τους όμως είναι περιορισμένος (σε σχέση με το πλήθος των διαθέσιμων scanned) και αφορά κυρίως σε σύγχρονες συνθέσεις ή σε απλά έργα μικρού μεγέθους. Τα περισσότερα μουσικά έργα κλασσικής μουσικής, είναι διαθέσιμα σε μορφή scanned και η επεξήγηση είναι απλή: Η μέθοδος της εκτύπωσης είναι πολύ παλαιότερη της μεθόδου ψηφιακής δημιουργίας των παρτιτούρων και επιπλέον τα έργα της κλασσικής μουσικής είναι αρκετά σύνθετα και περίπλοκα: οι παλαιές εκδόσεις είναι εξαιρετικά αξιόπιστες και επιπλέον το έργο της ψηφιακής «αντιγραφής» τους είναι εξαιρετικά χρονοβόρο, ασύμφορο, ενώ

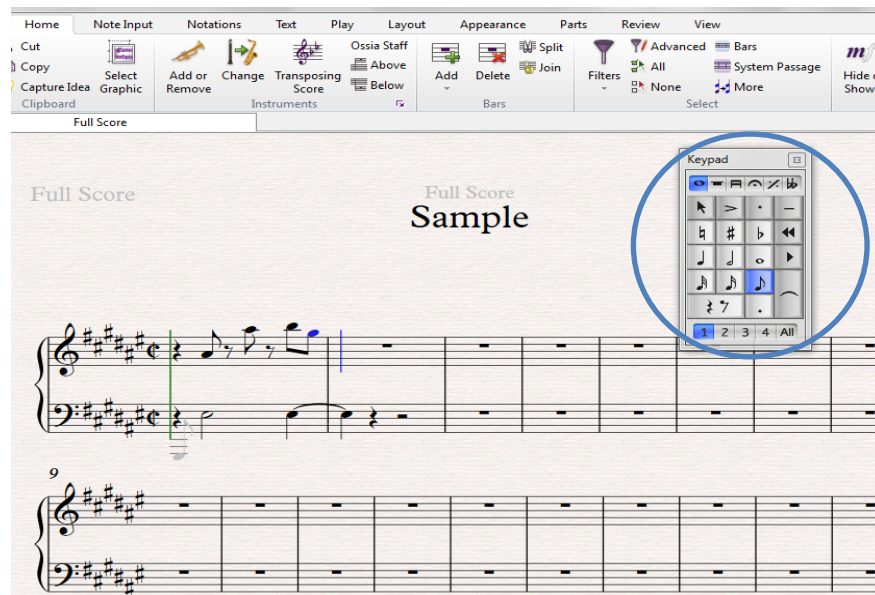
προκύπτει και το θέμα της αξιοπιστίας. Η αυτοματοποίηση του έργου της ψηφιοποίησης ανήκει κατεξοχήν στους σκοπούς της επιστημονικής έρευνας.

Διαθέσιμα προγράμματα

Υπάρχουν τρία είδη εμπορικών προγραμμάτων που σχετίζονται με τη χρήση παρτιτούρων:

- Προγράμματα σύνθεσης μουσικών κειμένων σε ψηφιακή (digital) μορφή
- Προγράμματα απλής απεικόνισης παρτιτούρων
- Προγράμματα μετατροπής scanned παρτιτούρων σε born digital μορφή

Τα προγράμματα σύνθεσης μουσικών κειμένων σε ψηφιακή μορφή παρέχουν στο χρήστη ένα περιβάλλον στο οποίο συνθέτει την παρτιτούρα είτε σε γραφικό περιβάλλον (πχ.Sibelius [13], Finale [14]), είτε σε περισσότερο προγραμματιστικό περιβάλλον (πχ.Lilypond [15]). Τα προγράμματα αυτά, παρουσιάζουν στην πλειοψηφία τους, ένα σύνολο διευρυμένων δυνατοτήτων, εκτός από την απλή συγγραφή του μουσικού κειμένου, όπως μετατροπή σε άλλη τονικότητα, ηχητική εκτέλεση του υπό συγγραφή κομματιού, μετατροπή σε άλλη τονικότητα, αλλαγή χρώματος και μεγέθους γραμματοσειράς, δυνατότητα συγγραφής για όλα τα μουσικά όργανα και τις επιμέρους ιδιαίτερες μορφές παρτιτούρας που χρειάζεται το καθένα και πλήθος άλλων δυνατοτήτων. Τα προγράμματα αυτά παράγουν απευθείας ψηφιακά έγγραφα μουσικού κειμένου (born digital παρτιτούρες). Καθώς σκοπός δεν είναι η παρουσίαση όλων των λειτουργιών για όλα τα προγράμματα, απεικονίζεται ένα στιγμιότυπο λειτουργίας ενός από αυτά (Sibelius 7, demo version, [6]):



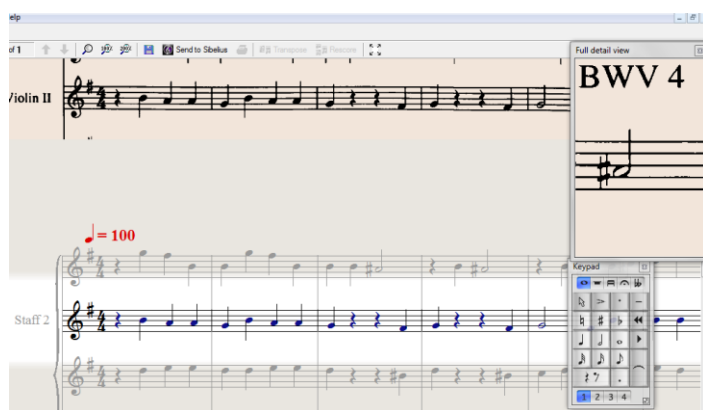
Εικόνα 1. 21: Στιγμιότυπο λειτουργίας του Sibelius 7 demo version, [6]: ο χρήστης επιλέγει την επιθυμητή αξία νότα από ένα πληκτρολόγιο (εμφανίζεται κυκλωμένο) και με drag and drop τρόπο την τοποθετεί στην επιθυμητή θέση. Μπορεί να διαλέξει τον οπλισμό, τις τυχαίες αλλοιώσεις, τις παύσεις και τα κλειδιά, καθώς και όλα τα στοιχεία που απαρτίζουν ένα μουσικό κείμενο.

Τα προγράμματα απλής απεικόνισης παρτιτούρων έχουν σκοπό να συνδράμουν τον εκτελεστή ενός μουσικού οργάνου στην ανάγνωση της μουσικής. Καθώς ο εκτελεστής διαβάζει την παρτιτούρα, παίζοντας ταυτόχρονα

τις νότες που απεικονίζονται στο μουσικό κείμενο, προκύπτει η ανάγκη της αλλαγής σελίδας: η αλλαγή σελίδας προϋποθέτει ο εκτελεστής να αφήσει το ένα του χέρι από το μουσικό όργανο προκειμένου να αλλάξει σελίδα και να μπορέσει να συνεχίσει την εκτέλεση του κομματιού. Η αλλαγή πρέπει να γίνει τόσο γρήγορα ώστε να μη διακοπεί η κανονική ροή του ακούσματος του μουσικού κομματιού-όπως γίνεται αντιληπτό αυτό είναι αρκετά δύσκολο, και σε πολλές περιπτώσεις υπάρχει κάποιος που συνοδεύει τον ερμηνευτή ώστε να του γυρίζει τις σελίδες.

Τα προγράμματα αυτά λοιπόν, σκοπό έχουν να αλλάζουν αυτόματα τη σελίδα που απεικονίζεται, βάσει της προκαθορισμένης ταχύτητας με την οποία πρέπει να παιχτεί το κομμάτι. Συχνά, το καθένα από τα προγράμματα αυτά, συνεργάζεται με κάποιο πρόγραμμα συγγραφής ψηφιακών παρτιτούρων και δέχεται σαν είσοδο μόνο παρτιτούρες που έχουν παραχθεί από το πρόγραμμα αυτό (πχ. Scorch [16], με Sibelius- το Scorch ενσωματώνεται στο φυλλομετρητή- web browser- ιστοσελίδων για απεικόνιση των παρτιτούρων στο Internet). Από την άλλη υπάρχουν και αντίστοιχα προγράμματα που δέχονται σαν είσοδο αρχεία pdf και έχουν διευρυμένες δυνατότητες απεικόνισης και παρέμβασης στο αρχικό αρχείο- χωρίς όμως την εγγύηση της λειτουργίας σε όλα τα είδη παρτιτούρων και της μέγιστης ακριβείας ή αυτόματης αλλαγής σελίδας χωρίς τροφοδότηση (feedback) από το χρήστη (πχ.forScore [17]).

Τα προγράμματα μετατροπής scanned παρτιτούρων σε born digital μορφή, επίσης συνεργάζονται με κάποιο πρόγραμμα συγγραφής μουσικού κειμένου (πχ.PhotoScore [18]). Σκοπός τους είναι να μετατρέπουν μια εικόνα (ή ένα pdf αρχείο) scanned ή χειρόγραφης παρτιτούρας σε μορφή born digital, δηλαδή σαν να έχει παραχθεί από το ψηφιακό πρόγραμμα συγγραφής μουσικού κειμένου. Αν και αυτό μοιάζει να επιλύει το πρόβλημα της ψηφιοποίησης των scanned παρτιτούρων, η πραγματικότητα είναι αρκετά διαφορετική, καθώς η ακρίβεια που προσφέρεται δε φτάνει σε καμία περίπτωση το 100%, καθώς προκύπτουν λάθη στην αναγνώριση γραμμών, μουσικών συμβόλων, αξιών, κλπ., η διόρθωση των οποίων είναι εξαιρετικά χρονοβόρα.



Εικόνα 1. 22: Επίδειξη λειτουργία Photoscore Lite demo version, [18], επάνω φαίνεται η αρχική παρτιτούρα-το ίδιο και σε μεγέθυνση στο παράθυρο δεξιά, ενώ κάτω το αποτέλεσμα της ψηφιοποίησης, με φωτεινό το αντίστοιχο πεντάγραμμα που παρουσιάζεται και επάνω, για να διευκολύνονται οι διορθώσεις.

Η θέση του ψηφιακού μουσικού αναλογίου

Ο προτεινόμενος αλγόριθμος προσβλέπει στην υλοποίηση του ψηφιακού μουσικού αναλογίου, το οποίο τοποθετείται σε μια ενδιάμεση θέση σε σχέση με τα υπάρχοντα προγράμματα. Σκοπός του δεν είναι η συγγραφή ψηφιακών μουσικών κειμένων ούτε η μετατροπή scanned παρτιτούρων σε born digital μορφή: για την ακρίβεια η επεξεργασία του περιεχομένου του μουσικού κειμένου δεν ανήκει στους σκοπούς του. Πλησιάζει περισσότερο στο σκοπό των προγραμμάτων απεικόνισης: Προσβλέπει στη διευκόλυνση του χρήστη στην παρουσίαση της παρτιτούρας κατά τη μελέτη ή την εκτέλεση ενός μουσικού κομματιού, με χρήση όμως της ελάχιστης δυνατής επεξεργασίας και την προσδοκία επίτευξης άρτιων αποτελεσμάτων.

Όπως αναφέρθηκε, η επεξεργασία του περιεχομένου του μουσικού κειμένου είναι έξω από τους σκοπούς της προτεινόμενης υλοποίησης. Στην πραγματικότητα, η αναγνώριση επιμέρους στοιχείων του μουσικού κειμένου περιορίζεται στο ελάχιστο δυνατό προκειμένου να εξασφαλίζεται στο χρήστη η δυνατότητα πολλαπλής παρέμβασης στον τρόπο απεικόνισης της παρτιτούρας. Για το λόγο αυτό, γίνεται αναγνώριση των γραμμών πενταγράμμου και των διαστολών, ενώ για τη μελλοντική εξέλιξη πρέπει να υπάρξει ένας τρόπος αναγνώρισης των κλειδιών, του οπλισμού και του χρόνου του κομματιού.

Η ιδέα και η υλοποίηση του αλγορίθμου βασίστηκαν στις εργασίες της Tambouratzis [19], [20], [21] και στον σχετικό κώδικα. Στις εργασίες αυτές αναλύεται πληρέστερα το πλαίσιο και οι επιμέρους λειτουργίες που αναμένεται να έχει το ψηφιακό μουσικό αναλόγιο και προτείνεται μια μεθοδολογία με χρήση μετασχηματισμού Fourier και προβολών για τον εντοπισμό των συστημάτων, των γραμμών και των διαστολών του μουσικού κειμένου καθώς και των μουσικών στοιχείων. Η παρούσα υλοποίηση εφαρμόζει επιμέρους τεχνικές που χρησιμοποιήθηκαν στον σχετικό με τις εργασίες κώδικα, όπως η περικοπή των λευκών περιθωρίων και η τεχνική της μακρύτερης ενιαίας γραμμής πενταγράμμου (longest unbroken black line).

Για την εξαγωγή των αποτελεσμάτων ακρίβειας της λειτουργίας του προτεινόμενου αλγορίθμου χρησιμοποιήθηκε ένα σύνολο 175 διαφορετικών εικόνων παρτιτούρων, εκ των οποίων οι 153 born digital (καθότι παρουσιάζουν το μεγαλύτερο ενδιαφέρον) και 22 born digital, ενώ δόθηκε μεγάλη προσοχή στην εξασφάλιση μιας ποικιλίας παρτιτούρων τόσο από άποψη ποιότητας και μεγέθους όσο και από άποψη περιεχομένου (αριθμός συστημάτων ανά σελίδα, αριθμός μέτρων κλπ.) Περισσότερες λεπτομέρειες θα αναφερθούν στο κεφάλαιο παρουσίασης των αποτελεσμάτων (Κεφάλαιο 4).

Κεφάλαιο 2°

Βιβλιογραφική ανασκόπηση

2.1 Γενικό πλαίσιο

Το επιστημονικό πεδίο της Οπτικής Αναγνώρισης Μουσικής (OAM, Optical Music Recognition, OMR) αφορά (α) στην κατά το δυνατόν ακριβέστερη αναγνώριση των μουσικών συμβόλων που απαρτίζουν το κάθε μουσικό κείμενο και (β) στην μετέπειτα κωδικοποίηση των αναγνωρισμένων μουσικών συμβόλων με τέτοιο τρόπο ώστε να είναι δυνατή η αυτόματη μετατροπή της παρτιτούρας σε μορφή επεξεργάσιμη από τον υπολογιστή. [22]

Η αναγνώριση μουσικού κειμένου έχει πολλαπλές εφαρμογές, για παράδειγμα:

(α) συντήρηση, συμπλήρωση, διόρθωση και ψηφιοποίηση παλιών καθώς και νέων παρτιτούρων,

(β) μεταγραφή της παρτιτούρας σε άλλο όργανο, άλλη τονικότητα ή άλλη μουσική σήμανση

(γ) εμφάνιση της παρτιτούρας στην οθόνη του Η/Υ του ερμηνευτή,

(δ) ταυτοποίηση του συνθέτη, αναγνώριση εικόνων παρτιτούρων κατά την αναζήτηση σε πλήθος εικόνων διάφορων τύπων κ.ά. (βλ. [23], [24], [25])

Στις περισσότερες υλοποιήσεις OAM, και πριν από την καθεαυτή αναγνώριση των μουσικών συμβόλων (η οποία δεν αποτελεί προ-απαιτήση της παρούσας εργασίας), υπάρχει ένα (ή και περισσότερα) στάδια προ-επεξεργασίας της εικόνας που συνήθως εφαρμόζεται ανεξάρτητα σε κάθε μία σελίδα της παρτιτούρας και αφορά είτε στην ταυτοποίηση των γραμμών του πενταγράμμου είτε στον εντοπισμό των πενταγράμμων σε μορφή αντικειμένων (blocks- χωρίς δηλαδή να ενδιαφέρει η ακριβής θέση των γραμμών του πενταγράμμου). Κατόπιν, οι αναγνωρισμένες γραμμές είτε απομακρύνονται (βλ. Εικόνα 2.1), προκειμένου να διευκολυνθεί η διαδικασία κατηγοριοποίησης και αναγνώρισης των μουσικών συμβόλων, είτε διατηρούνται προκειμένου να μην τεμαχιστούν τα μουσικά σύμβολα τα οποία τέμνονται από αυτές.



Εικόνα 2. 1: Αριστερά φαίνεται το απόσπασμα της αρχικής παρτιτούρας, δεξιά το αποτέλεσμα μετά την αφαίρεση των γραμμών. Τα μουσικά στοιχεία που έχουν αναγνωριστεί ως διαφορετικά, περικλείονται από διαφορετικό κουτί. Στη συνέχεια τα κουτιά που θεωρείται ότι αφορούν στο ίδιο στοιχείο συνενώνονται (δεν απεικονίζεται) [26]

Το στάδιο της αναγνώρισης των γραμμών είναι θεμελιώδες για την σωστή επεξεργασία της παρτιτούρας [11] [27], καθώς οι γραμμές των πενταγράμμων δημιουργούν ένα δισδιάστατο σύστημα/σκελετό του μουσικού κειμένου [28], με τον ίδιο τρόπο που οι γραμμές σε ένα τετράδιο αποτελούν τον οδηγό σε ένα χειρόγραφο κείμενο. Η αναγνώριση των γραμμών αποτελεί βασικό κομμάτι της παρούσας εργασίας, όπου προτείνεται μια επαναληπτική σταδιακή προσέγγιση του προβλήματος της αναγνώρισης, με τελικό σκοπό την επιτυχή αναγνώριση όλων των γραμμών αλλά και την εκμετάλλευση αυτής της πληροφορίας για την υλοποίηση του ψηφιακού μουσικού αναλογίου. Για το λόγο αυτό, αλλά και με σκοπό να αποκτήσει ο αναγνώστης μια σφαιρική αντίληψη της πολυπλοκότητας του προβλήματος αλλά και των διαφορετικών τρόπων που προτείνονται για την επίλυσή του, ακολουθεί περιγραφή των βασικών (ως προς το πλήθος των μεθόδων) προσεγγίσεων που έχουν δημοσιευθεί αναφορικά με τον εντοπισμό των γραμμών των πενταγράμμων.

2.2 Σύνοψη υπάρχουσών προσεγγίσεων αναγνώρισης γραμμών του πενταγράμμου- Σχετική βιβλιογραφία

Λόγω της σημαίνουσας θέσης της αναγνώρισης/εντοπισμού των γραμμών πενταγράμμου του στην οπτική αναγνώριση μουσικής, η σχετική βιβλιογραφία είναι εξαιρετικά εκτεταμένη και εξακολουθεί να αναπτύσσεται, καθώς δεν υπάρχει ακόμα κάποιο σύστημα ικανοποιητικής ακρίβειας και ενιαίας εφαρμογής σε όλες τις μορφές και μεγέθη παρτιτούρας το οποίο να χαιρεί καθολικής αποδοχής. [22], [29]

Η έρευνα της βιβλιογραφίας δεν περιορίζεται σε συγκεκριμένου τύπου μουσικά κείμενα, είτε από την άποψη της εκτύπωσης είτε από την άποψη του περιεχομένου, καθώς πρέπει – ιδανικά – η υλοποίηση να δύναται να επεξεργάζεται με τον ίδιο (ή ανάλογο) τρόπο χειρόγραφες, τυπωμένες καθώς και – σπανιότερα - ψηφιακές παρτιτούρες είτε τελικό σκοπό αποτελεί η ταυτοποίηση του συνθέτη είτε η εξαντλητική αναγνώριση και αποθήκευση όλων των μουσικών σύμβολων για διάσωση παλαιών παρτιτούρων ή μετατροπή του μουσικού κειμένου σε άλλη τονικότητα ή σε άλλο μουσικό όργανο.

Οι σχετικές εργασίες παρουσιάζονται κατά το δυνατό ομαδοποιημένες βάσει της προσέγγισης που χρησιμοποιούν για την αναγνώριση των γραμμών του πενταγράμμου και της σύνδεσης των μεθόδων που χρησιμοποιούν.

Η συνηθέστερη προσέγγιση για τον εντοπισμό των γραμμών του πενταγράμμου είναι η χρήση οριζόντιων ή/και κάθετων προβολών της εικόνας του μουσικού κειμένου (οι μέθοδοι αναλύονται λεπτομερέστερα κατά την παρουσίαση των σχετικών εργασιών, βλ. Εικόνα 3.5 για εποπτική απεικόνιση της οριζόντιας προβολής), με αρκετές παραλλαγές και επεκτάσεις:

Παρουσιάζοντας τις προκλήσεις της οπτικής αναγνώρισης μουσικής, οι Bainbridge και Bell [30] κατηγοριοποιούν τις μεθόδους εντοπισμού γραμμών πενταγράμμου σε δύο ομάδες, συγκεκριμένα αυτήν που χρησιμοποιεί οριζόντιες προβολές και αυτήν που χρησιμοποιεί κάθετες γραμμές σάρωσης της εικόνας. Στην οριζόντια προβολή, η δισδιάστατη εικόνα του μουσικού κειμένου αντιστοιχίζεται σε ένα μονοδιάστατο ιστόγραμμα, κάθε γραμμή του οποίου περιλαμβάνει το άθροισμα των μαύρων pixels στην αντίστοιχη γραμμή της εικόνας του κειμένου. Οι γραμμές πενταγράμμου είναι οι κορυφές του ιστογράμματος, γεγονός που καθιστά σχετικά εύκολο τον εντοπισμό τους. Στην κάθετη ανάλυση της εικόνας, οι αλγόριθμοι συλλέγουν δεδομένα από επικαλυπτόμενες περιοχές της εικόνας του μουσικού κειμένου, με επακόλουθη αναζήτηση για κάποιο κυρίαρχο χαρακτηριστικό το οποίο είναι είτε η staff-line-distance είτε το πάχος της γραμμής πενταγράμμου.

Σε μεταγενέστερη δημοσίευση, οι Bainbridge και Bell [25] παρουσιάζουν μια σειρά μεθόδων εντοπισμού των γραμμών πενταγράμμου στα πλαίσια της προσπάθειας αναγνώρισης εικόνων μουσικών κειμένων από μια συλλογή εικόνων ποικίλης θεματολογίας. Γίνεται χρήση (α) οριζόντιων προβολών και του μέτρου της πυκνότητας των pixels σε μια οθόνη, (β) μετασχηματισμού Hough [31] ως μεθόδου εντοπισμού της πιθανότητας ύπαρξης ενός παραμετροποιήσιμου χαρακτηριστικού μέσα σε μια εικόνα, (γ) αναλογιών μήκους διαδρομής (run-length) καθώς και (δ) διάφορων συνδυασμών των παραπάνω για βελτιωμένα αποτελέσματα.

Με σκοπό την ψηφιοποίηση μιας βιβλιοθήκης μουσικών κειμένων, οι Choudhury et al. [32] εφαρμόζουν πρακτικές ανάλυσης προβολών και κάθετης κωδικοποίησης μήκους διαδρομής (run-length coding [33]) για τον προσδιορισμό και την μετέπειτα αφαίρεση των γραμμών πενταγράμμου

Κάποιες φορές ο σκοπός δεν είναι η αναγνώριση όλων των γραμμών πενταγράμμου, αλλά η αναγνώριση των θέσεων των πενταγράμμων και του χώρου (αντικείμενα) που καταλαμβάνουν μέσα στην εικόνα του μουσικού κειμένου:

Οι Marinai και Nesi [34] επίσης χρησιμοποιούν οριζόντιες προβολές για τον εντοπισμό των πενταγράμμων. Σκοπό τους δεν αποτελεί ο εντοπισμός των γραμμών πενταγράμμου ξεχωριστά, αλλά η εύρεση του κάθε πενταγράμμου και των νοτών που περιλαμβάνει (ακόμα κι αν αυτές ξεφεύγουν από τα όρια του πενταγράμμου) ως ενιαίο αντικείμενο μέσα στο μουσικό κείμενο. Προτείνεται η

χρήση του μηχανισμού cut tolerance που διαστέλλει την περιοχή που έχει αναγνωριστεί ως πεντάγραμμο προκειμένου να περιλάβει και τυχόν νότες που έχουν ξεφύγει από την αρχική κατάτμηση. Επιπλέον, προτείνεται η προσέγγιση merge: ένας μηχανισμός συγκώνευσης δύο γειτονικών περιοχών πενταγράμμων που αλληλο-επικαλύπτονται περισσότερο από ένα προκαθορισμένο σταθερό όριο (merging value).

Την ίδια προσέγγιση του πενταγράμμου ως αντικείμενο ακολουθεί και ο Goecke [23], το σύστημα του οποίου αποσκοπεί στην ταυτοποίηση του συγγραφέα χειρόγραφων μουσικών κειμένων. Προτείνεται μία διαδικασία δύο σταδίων για τον εντοπισμό των γραμμών πενταγράμμου: αρχικά γίνεται αναζήτηση των πενταγράμμων σε επίπεδο σχήματος/αντικειμένου, χωρίς δηλαδή να ενδιαφέρει η θέση της κάθε γραμμής μέσα στο πεντάγραμμο, χρησιμοποιώντας έναν συνδυασμό ανίχνευσης ακμών (Sobel operator [35]) και οριζόντιων προβολών. Μετά τον εντοπισμένων των πενταγράμμων ως αντικειμένων, η εικόνα ακμών του μουσικού κειμένου σαρώνεται κάθετα από είκοσι «ακτίνες έρευνας» που βρίσκονται σε ίσες αποστάσεις μεταξύ τους και, κάθε φορά που μια ακτίνα συναντά μια ακμή, καταγράφεται η κάθετη θέση της. Στη συνέχεια εφαρμόζεται μια μέθοδος παρόμοια με αυτή που χρησιμοποιείται στο [36]: δημιουργείται ένα πρότυπο πενταγράμμου πλάτους εικοσιπέντε (25) pixels το οποίο κινείται εντός των αναγνωρισμένων πενταγράμμων, υπολογίζοντας την ετεροσυσχέτιση του προτύπου με το πραγματικό πεντάγραμμο. Ως γραμμές πενταγράμμου επιλέγονται αυτές που τοπικά δίνουν την καλύτερη ταύτιση του προτύπου με το πραγματικό πεντάγραμμο. Τέλος, οι γραμμές - οι οποίες είναι διακεκομμένες λόγω του τοπικού εντοπισμού τους ενώνονται μεταξύ τους, ενώ γίνεται και προσπάθεια συμπλήρωσης και κάποιων γραμμών που δεν έχουν αναγνωριστεί.

Λόγω μη ικανοποιητικής ποιότητας σάρωσης, παλαιότητας παρτιτούρας κλπ., προτείνονται κάποιες μέθοδοι που εστιάζουν στον εντοπισμό γραμμών πενταγράμμου οι οποίες είναι σημαντικά κατακερματισμένες, παρουσιάζουν μεγάλη κλίση ή είναι καμπυλωτές :

Οι Miyao και Okamoto [27], προτείνουν την εφαρμογή τεσσάρων σταδίων για την εύρεση των θέσεων των γραμμών του πενταγράμμου : α) καθορισμός των υποψήφια σημείων που ανήκουν σε γραμμές πενταγράμμου με κάθετη σάρωση της σελίδας της παρτιτούρας ανά διαστήματα ίσου ύψους, β) συνένωση των σημείων αυτών σε ενιαία γραμμή (dp matching), γ) χωρισμό γραμμών σε πεντάγραμμο, και δ) εντοπισμό των σημείων όπου τελειώνουν οι γραμμές των πενταγράμμων και επακόλουθη διόρθωση των θέσεων τους. Τα αποτελέσματα αυτής της μεθόδου καταγράφονται ως 99,52% ακριβή όσον αφορά στην εύρεση των θέσεων των πενταγράμμων σε σύνολο 104 τυπωμένων παρτιτούρων.

Ο Szwoch [37] προτείνει μια μέθοδο εντοπισμού παραμορφωμένων πενταγράμμων μέσω της χρήσης οριζόντιων προβολών σε τοπικά παράθυρα της εικόνας της παρτιτούρας, περαιτέρω επεξεργασία των ιστογραμμάτων που προκύπτουν, και εύρεση των μεταξύ τους συνδέσεων. Ο εφαρμοζόμενος αλγόριθμος μπορεί να δεχτεί ως είσοδο εικόνες που έχουν καταγραφεί από

διαφορετικά οπτικά μέσα σε μη πλήρως ελεγχόμενες συνθήκες, πχ. φωτογραφική μηχανή.

Εκτός από τη μέθοδο των προβολών υπάρχει πληθώρα διαφορετικών προσεγγίσεων σε σχέση με τον εντοπισμό των γραμμών μέσα στο μουσικό κείμενο:

Ο Rugin [38] χρησιμοποιεί κρυμμένα μοντέλα Markov (hidden Markov models [39]) για την οπτική αναγνώριση μουσικής σε πρώιμες εκδόσεις και εκτυπώσεις. Η γραμματοσειρά του μουσικού κειμένου αντιστοιχίζεται σε ένα σύνολο κρυμμένων μοντέλων Markov που χρησιμοποιούνται στη συνέχεια για την αναγνώριση των συμβόλων και των γραμμών. Ειδικά για τον εντοπισμό των γραμμών με καμπυλότητα, χρησιμοποιείται ένα κινούμενο παράθυρο ύψους 90 pixels, όπου μέσω της χρήσης τοπικών οριζόντιων προβολών - εξάγεται το ύψος του πενταγράμμου μέσα στο οποίο βρίσκονται τα υποψήφια προς αναγνώριση σύμβολα. Στη συνέχεια, βάσει του εξαχθέντος ύψους πενταγράμμου, δημιουργείται μια μάσκα (πρότυπο) πενταγράμμου. Διατρέχοντας εκ νέου την εικόνα της παρτιτούρας σε παράθυρα, αυτή τη φορά οριζόντια, εντοπίζονται οι καμπύλες γραμμές του πενταγράμμου ως το αποτέλεσμα της συνέλιξης της μάσκας με το παράθυρο της οριζόντιας προβολής.

Οι Cardoso et al. [40] εφαρμόζουν σταθερά μονοπάτια (stable paths) για την εύρεση των γραμμών του πενταγράμμου, όπου είναι δυνατός ο εντοπισμός γραμμών με κλίση, καμπυλότητα και ασυνέχειες. Η προτεινόμενη μέθοδος αντιμετωπίζει την εικόνα της παρτιτούρας σαν γράφο, όπου οι γραμμές των πενταγράμμων θεωρούνται συνδεδεμένα μονοπάτια μεταξύ των δύο πλευρικών περιθωρίων της εικόνας. Η προσέγγιση αυτή θεωρείται η μοναδική η οποία δεν λαμβάνει υπόψη μόνο την τοπική πληροφορία στην διαδικασία εντοπισμού των γραμμών, αλλά προσπαθεί να ενσωματώσει την καθολική πληροφορία που βρίσκεται μέσα στην εικόνα [22].

Στα πλαίσια της ταυτοποίησης του σύνθετη σε χειρόγραφα μουσικά κείμενα, οι Ng et al. [24] αναγνωρίζουν τις γραμμές κάθε πενταγράμμου αναλύοντας τα τμήματα οριζόντιων γραμμών που έχουν εντοπιστεί με τη χρήση βασικής σημειογραφίας και του ήδη υπολογισμένου πάχους των γραμμών πενταγράμμου.

Οι Sheridan και George [41] προτείνουν μια διαφορετική μέθοδο αναγνώρισης των συμβόλων ενός μουσικού κειμένου, στα πλαίσια της οποίας προσθέτουν επιπλέον γραμμές πενταγράμμου (defacing). Ακολουθώντας τον Goecke [23], οι οριζόντιες προβολές λαμβάνονται από μικρότερα κάθετα τμήματα του μουσικού κειμένου και οι διακεκομμένες γραμμές που εντοπίζονται ενώνονται στη συνέχεια.

Παρά την ύπαρξη πληθώρας μεθόδων και προσεγγίσεων που αφορούν στο θέμα του εντοπισμού των γραμμών του πενταγράμμου, καμία από τις προαναφερθείσες μεθόδους δεν εγγυάται 100% ακρίβεια, οι περισσότερες μάλιστα δεν αναφέρουν ποσοτικά τα αποτελέσματά τους. Η μη επίτευξη της

απόλυτης ακρίβειας οφείλεται στις διαφορές ποιότητας των εικόνων εισόδου, είτε αυτό αφορά στην αρχική εικόνα , στην ποιότητα σάρωσης, στην παλαιότητα του εγγράφου και στην επίδραση του χρόνου και τρόπου αποθήκευσής του, στην ποιότητα του χαρτιού και των εκτυπωμένων συμβόλων και πλήθος άλλων παραμέτρων.

Στη συνέχεια παρουσιάζεται ο προτεινόμενος αλγόριθμος αναγνώρισης των γραμμών πενταγράμμου και διαστολών που θα λειτουργήσει ως βάση του ψηφιακού μουσικού αναλογίου. Αν και βασισμένος σε κάποιες από τις μεθόδους που αναφέρθηκαν, εισάγει πρωτότυπες μεθόδους προκειμένου να εξασφαλιστεί η κατά το δυνατόν ακριβέστερη αναγνώριση γραμμών ανεξαρτήτως προέλευσης και ποιότητας εικόνας.

Κεφάλαιο 3°

Περιγραφή του αλγορίθμου

3.1 Συνοπτική περιγραφή

Η σωστή αναγνώριση των γραμμών είναι θεμελιώδους σημασίας για την υλοποίηση του ψηφιακού μουσικού αναλογίου, καθώς θα υποστηρίξει και θα διευκολύνει την περαιτέρω επεξεργασία της παρτιτούρας. Στη βάση της απλότητας και ελάχιστης επεξεργασίας, ο προτεινόμενος αλγόριθμος εντοπισμού των γραμμών πενταγράμμου εφαρμόζει μεν τη μέθοδο των οριζόντιων προβολών αλλά ακολουθεί την φιλοσοφία της σταδιακής εύρεσης και διόρθωσης των αποτελεσμάτων, προσέγγιση γνωστή και ως hill climbing [42]. Δηλαδή, βάσει μιας αρχικής εκτίμησης των γραμμών, διορθώνει και εμπλουτίζει τοπικά τα αποτελέσματα έτσι ώστε – μετά από μία σειρά επαναλήψεων και αναδρομών - να δώσει το επιθυμητό τελικό αποτέλεσμα της πλήρους και ακριβούς αναγνώρισης όλων των γραμμών του πενταγράμμου. Θα ακολουθήσει αναλυτική περιγραφή όλων των σταδίων του αλγορίθμου, προκειμένου να γίνει κατανοητή η προτεινόμενη λύση στο πρόβλημα του εντοπισμού των γραμμών, αλλά και η πολυπλοκότητα του εν λόγω προβλήματος, λόγω της ποιότητας σάρωσης των μουσικών κειμένων, αλλά και της αρχικής τους ποιότητας, της φυσικής τους φθοράς και της ποικιλομορφίας μεταξύ των διαφορετικών εκδόσεων και εκτυπώσεων (μέγεθος, τρόπος σάρωσης κλπ).

Συνοπτικά, οι επιμέρους λειτουργίες του αλγορίθμου είναι οι ακόλουθες:

- (α) αρχική εκτίμηση των θέσεων των γραμμών του πενταγράμμου με χρήση οριζόντιας προβολής ολόκληρης της εικόνας
- (β) συμπλήρωση γραμμών για πεντάγραμμο που έχουν εντοπιστεί μερικώς, δηλαδή έχει εντοπιστεί η θέση τους μέσω κάποιων αλλά όχι όλων των γραμμών τους
- (γ) περιορισμός των γραμμών του πενταγράμμου όταν οι εντοπισμένες γραμμές υπερβαίνουν τις πέντε
- (δ) αναζήτηση ολόκληρων πενταγράμμων που δεν έχουν ανιχνευτεί
- (ε) βελτιστοποίηση των θέσεων των εντοπισμένων γραμμών κάθε πενταγράμμου προκειμένου να ανταποκρίνονται ακριβέστερα στις θέσεις των πραγματικών γραμμών του πενταγράμμου
- (στ) απομάκρυνση πενταγράμμων που εντοπίστηκαν κατά τη διάρκεια της επεξεργασίας του μουσικού κειμένου αλλά δεν ανταποκρίνονται σε πραγματικά πεντάγραμμο
- (ζ) εντοπισμός των διαστολών για διαχωρισμό των μέτρων.

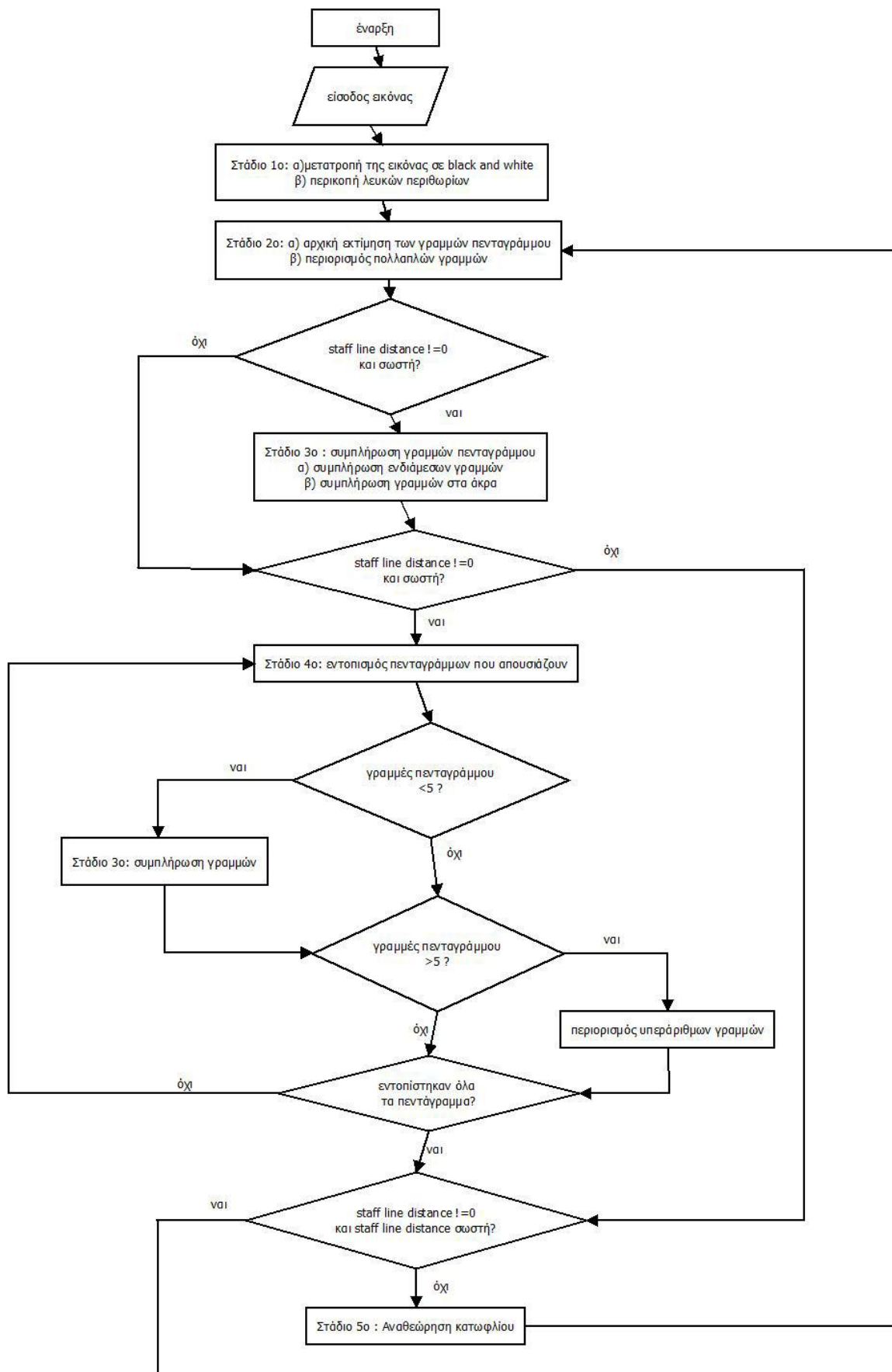
Οι λειτουργίες αυτές συνδυάζονται μεταξύ τους με έναν σύνθετο τρόπο, ο οποίος δεν είναι πλήρως καθορισμένος αλλά εξαρτάται από τις απαιτήσεις του υπό επεξεργασία μουσικού κειμένου. Για παράδειγμα, κατά τη διάρκεια της

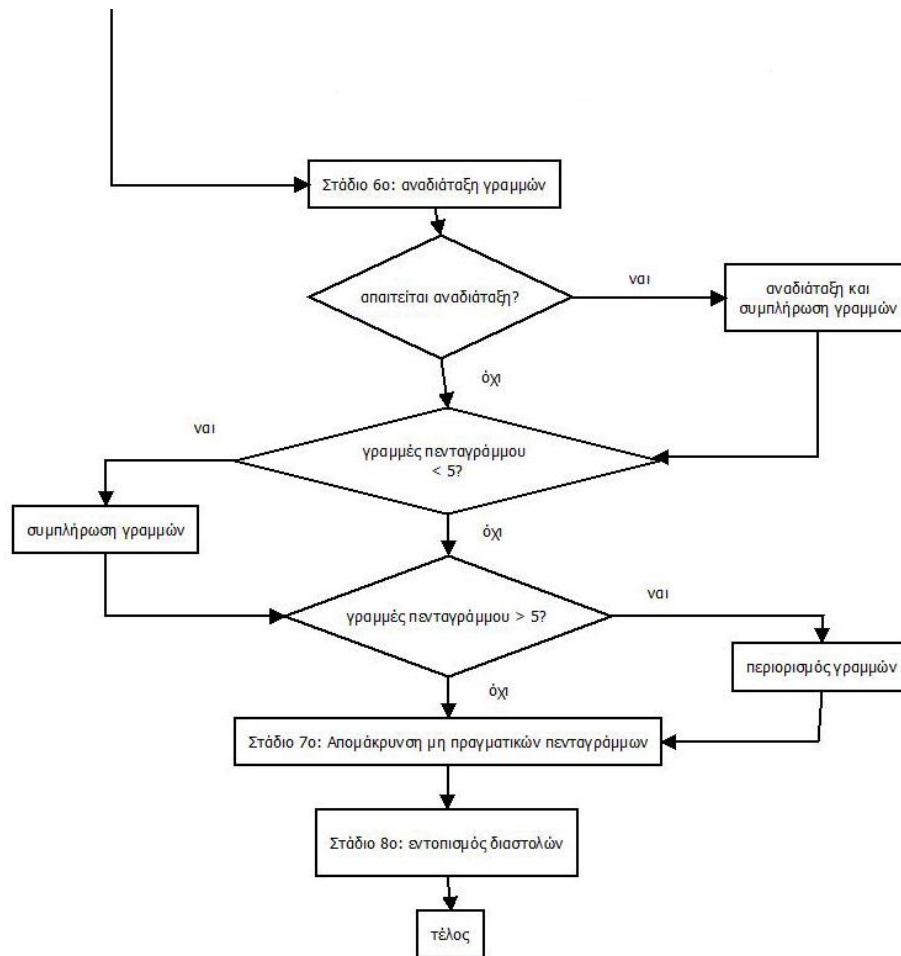
λειτουργίας (δ) γίνεται μεταπήδηση στις λειτουργίες (β) και (γ) για διόρθωση των νέων πενταγράμμων που εντοπίζονται. Το ίδιο συμβαίνει και κατά τη λειτουργία (ε), για τα νέα πεντάγραμμα που προκύπτουν από τη βελτιστοποίηση των θέσεων των ήδη εντοπισμένων γραμμών. Επιπλέον, είναι δυνατό η λειτουργία (ε) να συνδράμει στην προετοιμασία της λειτουργίας (στ), περιορίζοντας πεντάγραμμα τα οποία δεν ανταποκρίνονται σε πραγματικά. Με παρόμοιο τρόπο, η λειτουργία (γ) μπορεί να συνδράμει στην αποτελεσματικότητα της λειτουργίας (β), περιορίζοντας τον αριθμό των υπεράριθμων γραμμών ώστε να είναι μεν μικρότερος του αποδεκτού αλλά να ανταποκρίνεται σε πραγματικές γραμμές πενταγράμμου, ώστε να διευκολυνθεί η συμπλήρωσή τους σε πεντάγραμμα από τη λειτουργία (β).

Ακολουθεί διάγραμμα ροής της συνολικής διαδικασίας (βλ. Εικόνα 3.1) προκειμένου να υπάρχει καλύτερη κατανόηση της σχέσης μεταξύ των διαδοχικών σταδίων και των αναδρομών και αλμάτων που γίνονται μεταξύ τους.

Επιπλέον σε κάθε στάδιο υπάρχει διάγραμμα ροής που περιγράφει τη ροή του αλγορίθμου, για να χρησιμοποιείται - σε συνδυασμό με το συνολικό διάγραμμα ροής - ως αναφορά για τον αναγνώστη.

Σε αυτές τις βασικές λειτουργίες περιέχονται επιπλέον τεχνικές βελτιστοποίησης της λειτουργικότητας και απόδοσης του αλγορίθμου, οι οποίες θα περιγραφούν αναλυτικότερα στη συνέχεια.

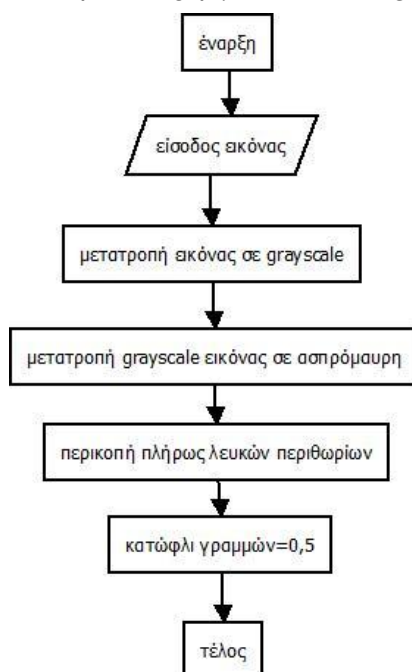




Εικόνα 3. 1: Συνοπτικό διάγραμμα ροής του προτεινόμενου αλγορίθμου εντοπισμού γραμμών (απεικονίζεται σε δύο συνεχόμενες σελίδες-η γραμμή που φαίνεται να διακόπτεται στην πρώτη σελίδα στην πραγματικότητα συνεχίζεται στην επόμενη).

3.2 Τα στάδια του αλγορίθμου

3.2.1 Στάδιο 1^ο: Προ-επεξεργασία εικόνας εισόδου



Εικόνα 3. 2: Διάγραμμα ροής σταδίου προ-επεξεργασίας εικόνας (1^ο στάδιο)- παραλείπεται η αναπαράσταση της εμφάνισης της επεξεργασμένης εικόνας πριν τον τερματισμό

Ο αλγόριθμος αναγνώρισης γραμμών υλοποιείται σε MATLAB[®] και δέχεται σαν είσοδο έγχρωμες εικόνες τύπου .bmp χωρίς περιορισμό στις διαστάσεις. Ίσως προκαλεί εντύπωση το γεγονός ότι ενώ τα μουσικά κείμενα τυπώνονται ασπρόμαυρα, οι απεικονίσεις τους είναι δυνατό να απαιτούν μεγαλύτερη χρωματική παλέτα από αυτήν του άσπρου και του μαύρου: Εξ' ορισμού, ο τύπος αποθήκευσης bitmap (.bmp) επιτρέπει τη χρήση διαφορετικών τύπων χρώματος στην αναπαράσταση της εικόνας. Δηλαδή ένα pixel της εικόνας μπορεί να χρειάζεται από 24 bit για την αναπαράσταση του χρώματός του (τρία κανάλια των 8 bit για απεικόνιση κόκκινου, πράσινου, μπλε) μέχρι 1 bit για ασπρόμαυρη εικόνα. Στις παρτιτούρες που χρησιμοποιήθηκαν πιο συχνή είναι η χρήση 24 bits για την απεικόνιση ενός pixel. Με αυτόν τον τρόπο μπορούν να απεικονιστούν με λεπτομέρεια όλες οι πιθανές αποχρώσεις που είναι δυνατό να εμφανίζονται σε μια παρτιτούρα, καθώς υπάρχουν παρτιτούρες των οποίων οι σελίδες μπορεί να εμφανίζουν μια διαφορετική απόχρωση από τη γνωστή άσπρη, λόγω παλαιότητας ή χρήσης διαφορετικού χαρτιού, κλπ. Επίσης υπάρχει διακύμανση στην ένταση του άσπρου και του μαύρου χρώματος. Γίνεται αντιληπτό ότι αυτή η ποικιλία που προσφέρει η τυποποίηση bmp μπορεί να οδηγήσει σε εικόνες διαφορετικών ποιοτήτων και διαστάσεων.

Η μη ύπαρξη περιορισμού στις διαστάσεις σημαίνει ότι η ποιότητα των εικόνων εισόδου μπορεί να είναι πολύ καλή, εξασφαλίζοντας καλύτερα αποτελέσματα (αλλά με σημαντικά αυξημένη υπολογιστική πολυπλοκότητα λόγω μεγέθους

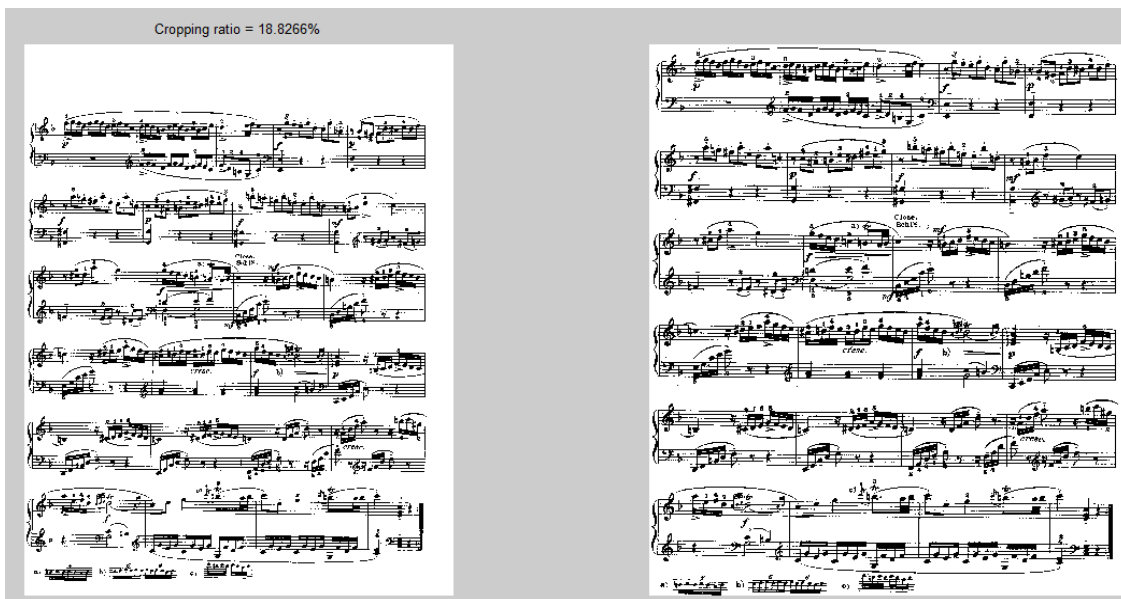
εικόνας), ή αρκετά χαμηλή, που απαιτεί λιγότερους υπολογισμούς και πόρους του συστήματος αλλά θεωρητικά μπορεί να δώσει χειρότερα αποτελέσματα. Γίνεται αντιληπτό από τα παραπάνω ότι η εικόνα εισόδου δε μπορεί να θεωρηθεί ιδανική. Καθώς το σύστημα δέχεται οποιαδήποτε εικόνα μουσικού κειμένου χωρίς περιορισμό, είναι εξαιρετικά συχνό οι γραμμές των πενταγράμμων να παρουσιάζουν μια σειρά από προβλήματα (πχ. κατακερματισμένες γραμμές, γραμμές με κλίση, γραμμές διαφορετικού πάχους), τα οποία παρουσιάζονται σταδιακά στη συνέχεια, ταυτόχρονα με την παρουσίαση και επεξήγηση του αλγορίθμου μαζί με τον προτεινόμενο τρόπο επίλυσής τους. Από την άλλη πλευρά, ακόμα και σε ιδανικές εικόνες εισόδου, παρουσιάζονται προβλήματα άλλου είδους (πχ. μεγάλες αποστάσεις μεταξύ συστημάτων). Στόχος του συστήματος είναι η εμφάνιση κατά το δυνατόν ακριβέστερων αποτελεσμάτων ανεξάρτητα από το μέγεθος και την ποιότητα της εικόνας².

Στο στάδιο αυτό, η εικόνα εισόδου μετατρέπεται σε grayscale και στη συνέχεια σε ασπρόμαυρη (black and white: όλα τα pixels της εικόνας έχουν τιμή 0 αν είναι μαύρα ή 1 αν είναι λευκά). Για την εκτέλεση αυτής της μετατροπής, χρησιμοποιείται υψηλό κατώφλι μετατροπής των grayscale pixels της εικόνας σε ασπρόμαυρα, καθώς προτιμάται τα γκριζα pixels να μετατραπούν σε μαύρα παρά σε άσπρα, ώστε να μη δημιουργηθούν πρόσθετοι κατακερματισμοί των στοιχείων του μουσικού κειμένου πέραν των ήδη υπάρχοντων. Από την άλλη πλευρά, αυτή η πρακτική πιθανόν να αφαιρεί κάποιες από τις λεπτομέρειες του περιγράμματος των στοιχείων του μουσικού κειμένου, όμως αφενός ο βαθμός απομάκρυνσης των λεπτομερειών είναι μικρός και αφετέρου η επεξεργασία των νοτών μέσα στο κείμενο ξεφεύγει από το αντικείμενο της παρούσας υλοποίησης. Η αντιμετώπιση τυχόν προβλημάτων κατά την αναγνώριση συγκεκριμένων στοιχείων μέσα στο μουσικό κείμενο θα αναφερθεί στο σχετικό κεφάλαιο.

Στη συνέχεια, γίνεται περικοπή των εντελώς λευκών περιθωρίων της εικόνας σε όλο της το περίγραμμα (επάνω, δεξιά, κάτω και αριστερή ζώνη), ώστε να μειωθούν οι διαστάσεις της και να διευκολυνθεί η περαιτέρω επεξεργασία της. Λόγω του θορύβου αλλά και των διάφορων σχολίων που μπορεί να υπάρχουν στην εικόνα (τίτλος, άλλα σχόλια) λευκά περιθώρια παραμένουν ως ένα βαθμό γύρω από το μουσικό κείμενο, ιδιότητα που θα εκμεταλλευτεί στη συνέχεια ο αλγόριθμος στη διαδικασία εύρεσης των γραμμών του πενταγράμμου. Ταυτόχρονα, καθορίζεται η αρχική, προεπιλεγμένη τιμή κατωφλίου για τον εντοπισμό των γραμμών της εικόνας που θα χρειαστεί στο αμέσως επόμενο στάδιο.

Τα αποτελέσματα του σταδίου αυτού παρουσιάζονται εποπτικά στις παρακάτω εικόνες (Εικόνες 3.3, 3.4):

² για το λόγο αυτό άλλωστε δεν υπάρχει περιορισμός στις διαστάσεις της εικόνας



Εικόνα 3. 3: Επίδειξη περικοπής της ασπρόμαυρης εικόνας του μουσικού κειμένου (Mozart sonata, σελίδα 2 [43]), επάνω αριστερά εμφανίζεται το ποσοστό της περικοπής (18,8266%). Παρατηρείται ότι τα σχόλια για την εκτέλεση στο κάτω μέρος εμποδίζουν την περαιτέρω περικοπή της εικόνας. Αλλά και στο πάνω μέρος, εμφανίζεται ικανή απόσταση από το πρώτο πεντάγραμμο ως το άκρο της εικόνας, καθώς αν εντοπιστεί έστω και ένα μαύρο pixel, η περικοπή τερματίζεται.



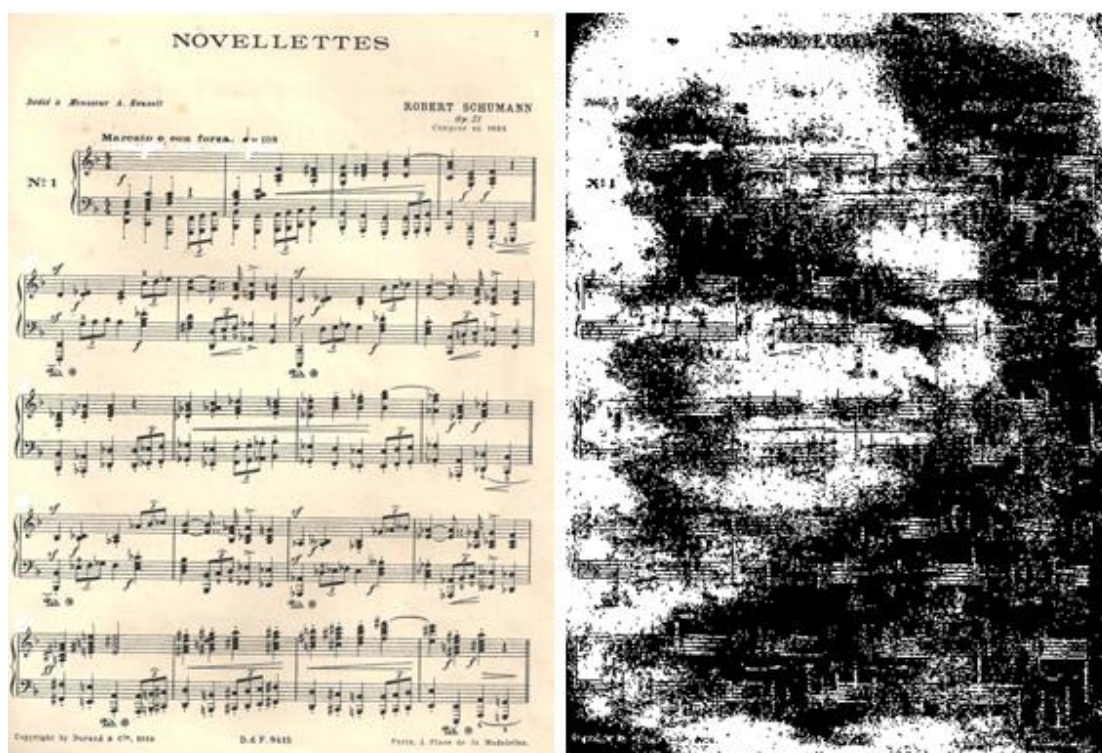
Εικόνα 3. 4: Παράδειγμα περικοπής εικόνας, (Debussy Valse, σελίδα 1, [44]). Η άνω περικοπή σταματάει στον τίτλο ενώ η κάτω στο σχόλιο που υπάρχει στο κάτω δεξί μέρος της εικόνας, το οποίο έχει προστεθεί κατά τη σάρωση του μουσικού κειμένου για τη μετατροπή του σε ηλεκτρονική μορφή. Το ποσοστό περικοπής είναι 8,773%.

Το κατώφλι μετατροπής της εικόνας σε ασπρόμαυρη

Όπως ήδη αναφέρθηκε, για τη μετατροπή της εικόνας από grayscale σε ασπρόμαυρη, απαιτείται η χρήση ενός κατωφλίου. Ο ρόλος του κατωφλίου είναι να καθορίζει το όριο της τιμής φωτεινότητας (luminance) των pixels, πάνω από το οποίο θα μετατρέπονται σε άσπρα pixels (τιμή=1) και κάτω από το οποίο

σε μαύρα pixels (τιμή=0), για την ασπρόμαυρη απεικόνιση. Το κατώφλι εκφράζεται σε κανονικοποιημένη τιμή: το όριο που χρησιμοποιεί ο προτεινόμενος αλγόριθμος είναι το 0,95 (ή 95%). Αυτό σημαίνει πως σε όλα τα pixels των οποίων η τιμή βρίσκεται πάνω από το 95% του εύρους των δυνατών τιμών (πχ.256 με χρήση 8 bit), θα ανατίθεται η τιμή 1 (άσπρο pixel) ενώ στα υπόλοιπα θα ανατίθεται η τιμή 0 (μαύρο pixel). Η χρήση ενός τόσο υψηλού κατωφλίου κρίθηκε απαραίτητη προκειμένου να εξασφαλίσει την εμφάνιση περισσότερο ενιαίων γραμμών, καθώς παρατηρήθηκε ότι η χρήση χαμηλότερου κατωφλίου παράγει περισσότερες άσπρες περιοχές, ενώ επιδεινώνει τα προβλήματα των κατακερματισμένων ή/και μη έντονων γραμμών πενταγράμμου καθιστώντας την αναγνώρισή τους ακόμα δυσκολότερη.

Παρόλα αυτά, υπάρχουν περιπτώσεις στις οποίες το υψηλό αυτό κατώφλι οδηγεί στη μετατροπή περιοχών που στην αρχική εικόνα βρίσκονται στο φόντο (background) να έρχονται στο προσκήνιο (foreground), με τον τρόπο που απεικονίζεται στην παρακάτω εικόνα (Εικόνα 3.5). Αυτό συμβαίνει λόγω της ύπαρξης ενός είδους υφής στην εικόνας της παρτιτούρας, υφή που έχει προκληθεί λόγω της παλαιότητας της έκδοσης ή/και της κακής ίσως συντήρησης.



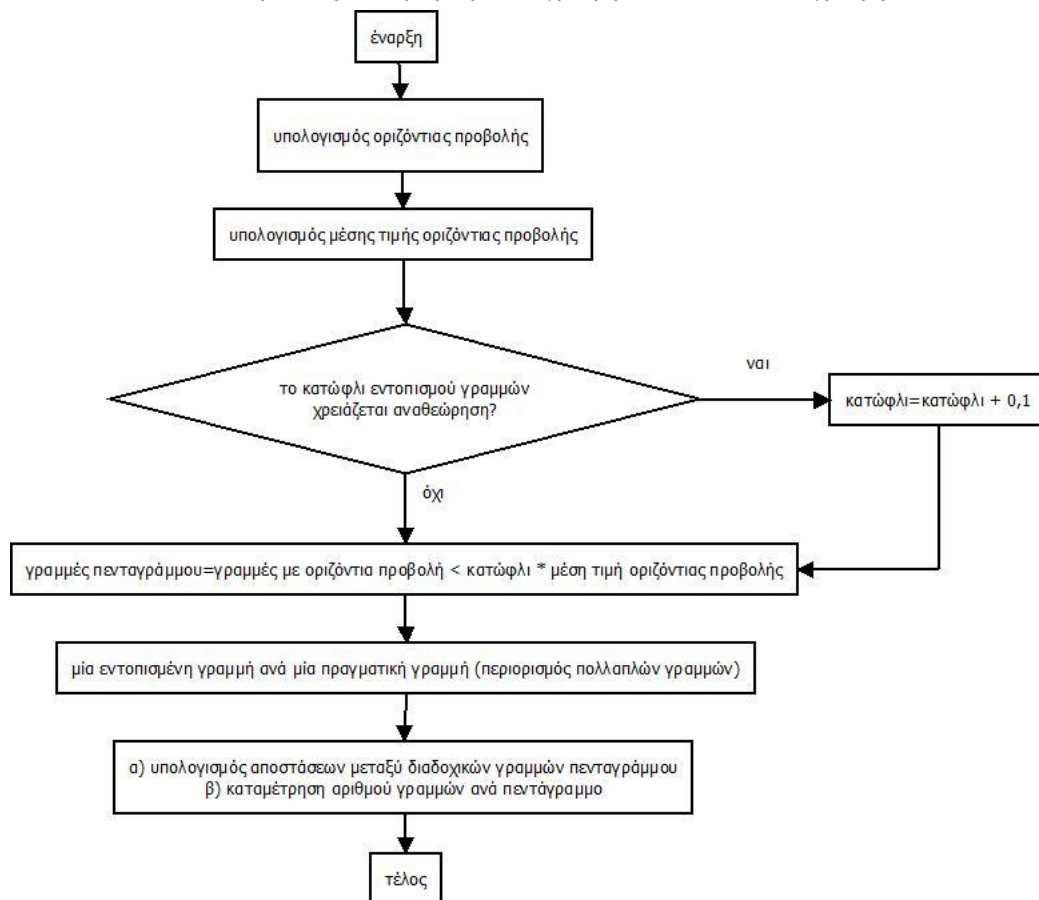
Εικόνα 3. 5: Επίδειξη προβληματικής λειτουργίας υψηλού κατωφλίου σε παρτιτούρα με μη ενιαίο υπόβαθρο. Αριστερά: Αρχική έγχρωμη εικόνα- είναι φανεροί οι διαφορετικοί τόνοι στο χρωματισμό του υποβάθρου, Δεξιά: Ασπρόμαυρη εικόνα με χρήση κατωφλίου 0,95, Schumann Novellettes, σελίδα 1 [10]

Στις περιπτώσεις αυτές (που για το σύνολο των εξεταζόμενων παρτιτούρων αφορούν δύο μόνο κομμάτια, σύνολο δεκατέσσερις εικόνες), χρησιμοποιείται χαμηλότερο κατώφλι, ίσο με 0,75 (75%). Το αποτέλεσμα για την παρτιτούρα που παρουσιάστηκε πάνω (Εικόνα 3.5)είναι το ακόλουθο (Εικόνα 3.6):



Εικόνα 3. 6: Αποτέλεσμα κατωφλίωσης με χρήση χαμηλότερου κατωφλίου για τη μετατροπή της εικόνας εισόδου σε ασπρόμαυρη την παρτιτούρα που παρουσιάστηκε στην Εικόνα 3.5. Είναι φανερή η επίδραση του κατωφλίου στο τελικό αποτέλεσμα: οι διαφορετικές περιοχές που εμφανίζονται στο υπόβαθρο έχουν αντικατασταθεί με ενιαίο άσπρο χρώμα, Schumann Novelletes, σελίδα 1, [10]

3.2.2 Στάδιο 2ο: Αρχική εκτίμηση των γραμμών του πενταγράμμου



Εικόνα 3. 7: Διάγραμμα ροής σταδίου αρχικής εκτίμησης γραμμών -παραλείπεται η αναπαράσταση της εμφάνισης της εικόνας με τις εντοπισμένες γραμμές σε υπέρθεση πριν το τέλος του σταδίου

Σκοπός του σταδίου αυτού είναι η παροχή μιας αρχικής εκτίμησης των θέσεων των γραμμών πενταγράμμου της εικόνας του μουσικού κειμένου. Με είσοδο την ασπρόμαυρη περικεκομμένη (cropped) εκδοχή της αρχικής εικόνας, υπολογίζεται η οριζόντια προβολή της, δηλαδή το άθροισμα των τιμών των pixels της εικόνας για κάθε γραμμή της εικόνας. Δημιουργείται δηλαδή, ένας πίνακας-στήλη με μήκος ίσο με το ύψος της εικόνας, όπου κάθε κελί του περιέχει το άθροισμα των τιμών των pixels της εικόνας στην αντίστοιχη γραμμή (βλ. Εικόνα 3.10 για την εμφάνιση του ιστογράμματος της οριζόντιας προβολής).

Δεδομένου ότι τα λευκά pixels αναπαρίστανται με τη μονάδα (1) ενώ τα μαύρα με το μηδέν (0) και του ότι οι γραμμές πενταγράμμου είναι οι μόνες μαύρες ενιαίες οριζόντιες γραμμές μέσα στο μουσικό κείμενο, εύκολα συμπεραίνεται ότι όσο πιο μικρό το άθροισμα των pixels μιας γραμμής της εικόνας, τόσο πιθανότερο είναι αυτή η γραμμή να αποτελεί γραμμή πενταγράμμου. Ως υποψήφιες θέσεις γραμμών πενταγράμμου επιλέγονται οι γραμμές τις εικόνας που το άθροισμα των pixels τους είναι μικρότερο από το μισό της μέσης τιμής της οριζόντιας προβολής όλης της εικόνας:

$$\text{κατώφλι επιλογής γραμμών} = \text{κατώφλι} * \text{μέση τιμή οριζόντιας προβολής εικόνας}$$

όπου ως κατώφλι έχει προεπιλεγεί η τιμή 0,5 στο προηγούμενο στάδιο, αλλά αν αυτή θεωρηθεί ανεπαρκής (με τρόπο θα παρουσιαστεί στο 5^ο στάδιο), μπορεί να μεταβληθεί με τον τρόπο που απεικονίζεται στο διάγραμμα ροής (Εικόνα 3.7).

Στη μέση τιμή συνεισφέρουν όλες οι γραμμές τις εικόνας, οι περισσότερες από τις οποίες έχουν μεγάλη περιεκτικότητα άσπρου χρώματος (αν ληφθεί υπόψη ότι το πλήθος των κενών μεταξύ γραμμών πενταγράμμου, πενταγράμμων και συστημάτων περιέχουν μικρό αριθμό μαύρων pixels σε σχέση με το μήκος τους). Για το λόγο αυτό το κατώφλι επιλέγεται να είναι τόσο μικρό σε σχέση με την μέση τιμή της οριζόντιας προβολής.

Η αρχική εκτίμηση των γραμμών πενταγράμμου της εικόνας, δίνει ένα αποτέλεσμα της παρακάτω μορφής (οι εντοπισμένες γραμμές υπερτίθενται πάνω στην αρχική εικόνα χωρίς να τη μεταβάλλουν):



Εικόνα 3. 8: Αρχική εκτίμηση των γραμμών, απόσπασμα από Chopin nocturnes, σελίδα 1 [45]. Παρατηρείται ότι οι εντοπισμένες γραμμές είναι διαφορετικού πάχους (1. Πραγματική γραμμή στην οποία αντιστοιχούν δύο διαδοχικές ανικνευθείσες γραμμές 2. Πραγματική γραμμή στην οποία αντιστοιχεί μία ανικνευθείσα γραμμή-στο φαινόμενο αυτό οφείλεται και η εμφάνιση διαφορετικού πάχους).

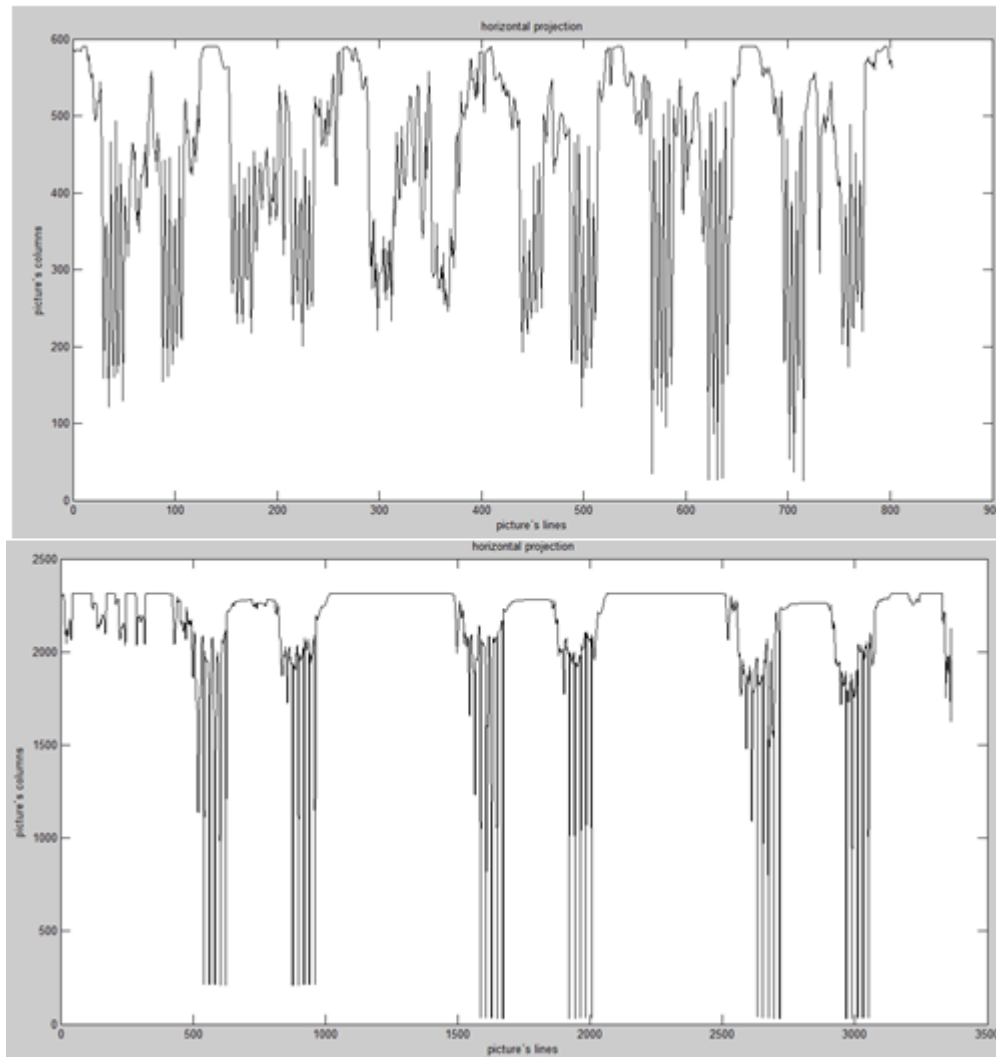
Παρατηρείται εδώ (Εικόνα 3.8) ότι υπάρχουν εντοπισμένες γραμμές που είναι παχύτερες από άλλες. Αυτό συμβαίνει γιατί οι γραμμές πενταγράμμου στις οποίες αντιστοιχούν οι εντοπισμένες γραμμές, έχουν πάχος μεγαλύτερο του ενός pixel, και η τιμή της οριζόντιας προβολής τους δεν ξεπερνά το κατώφλι σε όλο τους το πάχος, για όλες δηλαδή τις γραμμές μοναδιαίου πάχους που σχηματίζουν την πραγματική γραμμή του πενταγράμμου. Το αποτέλεσμα είναι η ύπαρξη διπλών ή τριπλών γραμμών (σε πάχος pixels) που αντιστοιχούν σε μία πραγματική γραμμή. Επιπλέον, γίνεται αντιληπτό ότι δεν εντοπίζονται όλες οι γραμμές των πενταγράμμων κατά την αρχική εκτίμηση, όπως φαίνεται και στην Εικόνα 3.8, γιατί λόγω κατακερματισμού, κλίσης ή εκτύπωσης, δε βρίσκονται όλες οι πραγματικές γραμμές κάτω από το κατώφλι που έχει τεθεί.

Ακολουθεί ένα παράδειγμα αρχικού εντοπισμού γραμμών για δύο διαφορετικές μεταξύ τους παρτιτούρες:



Εικόνα 3. 9: Από αριστερά: Beethoven Sonata, σελίδα 3, [46] και Bizet Carmen Prelude, σελίδα 1, [47], δείγματα αρχικής εκτίμησης γραμμών, όπου η αριστερή εικόνα είναι χαμηλής ποιότητας, μικρού μεγέθους και scanned, ενώ η δεξιά είναι υψηλής ποιότητας, μεγάλου μεγέθους και born digital.

Στην παραπάνω εικόνα (Εικόνα 3.9) γίνεται επίδειξη αυτής ακριβώς της διαφοράς κατά την αρχική εκτίμηση των γραμμών για τις διαφορετικές εικόνες μουσικών κομματιών. Στην αριστερή πλευρά, βρίσκεται μια εικόνα μικρών σχετικά διαστάσεων (803x590 pixels) που είναι σαρωμένη από το αρχικό πρωτότυπο (scanned), ενώ δεξιά βρίσκεται μια εικόνα πολύ μεγαλύτερων διαστάσεων (3364x2316 pixels, 16 φορές μεγαλύτερη σε σύνολο pixels από την προηγούμενη), η οποία έχει παραχθεί από πρόγραμμα συγγραφής μουσικής (born digital). Η διαφορά στην αρχική εκτίμηση των γραμμών είναι εμφανής, καθώς στη born digital εικόνα έχουν εντοπιστεί όλες οι γραμμές, με κάποιες απλώς να έχουν πάχος μεγαλύτερο του μοναδιαίου, ενώ στη scanned εικόνα, υπάρχουν ακόμα πολλές γραμμές προς εντοπισμό, ακόμα και ολόκληρα πεντάγραμμα. Η διαφορά αυτή αντικατοπτρίζεται και στα ιστογράμματα των οριζόντιων προβολών των εικόνων αυτών:



Εικόνα 3. 10: Οριζόντια προβολή των εικόνων που εμφανίζονται στην Εικόνα 3.9. Πάνω απεικονίζεται η οριζόντια προβολή της Beethoven Sonata (scanned) και κάτω του Bizet Carmen Prelude (born digital). [46] [47]

Στον οριζόντιο άξονα του ιστογράμματος της οριζόντιας προβολής, απεικονίζονται οι γραμμές της εικόνας, ενώ στον κάθετο το άθροισμα των pixels (για κάθε γραμμή της εικόνας). Τα τοπικά ελάχιστα, αντιπροσωπεύουν γραμμές του πενταγράμμου. Παρατηρείται πόσο καθαρή και ευκρινής είναι η οριζόντια προβολή του born digital μουσικού κειμένου, ειδικά σε σχέση με την αντίστοιχη του scanned. Ακόμα και στο scanned μουσικό κείμενο όμως, είναι φανερά οι θέσεις των πενταγράμμων από τον ανθρώπινο παρατηρητή, αλλά όχι αρκετά διακριτές από τις γειτονικές τους ώστε να εντοπιστούν στην αρχική εκτίμηση από τον αλγόριθμο (εξαιτίας των κεκλιμένων γραμμών, κατακερματισμένων γραμμών κτλ.)

Στη συνέχεια αντιμετωπίζεται το πρόβλημα του μη ενιαίου πάχους των γραμμών (δηλαδή των πολλαπλών αναγνώρισεων της ίδιας πραγματικής γραμμής) που παρουσιάστηκε παραπάνω. Προκειμένου να προσεγγιστεί η δυνατότητα του ανθρώπινου ματιού να αναγνωρίζει τις γραμμές πενταγράμμου ως μοναδικές, πραγματοποιείται η αντιστοίχιση μίας εντοπισμένης γραμμής

ανά πραγματική γραμμή πενταγράμμου. Δηλαδή στην περίπτωση που υπάρχει μια εντοπισμένη διπλή γραμμή ή ακόμα και τριπλή (πχ. δύο διαδοχικές γραμμές της εικόνας ή δύο γραμμές της εικόνας που απέχουν δύο pixels μεταξύ τους, κ.ά.), ο αλγόριθμος κρατάει την τελευταία από τη σειρά των διαδοχικών γραμμών.

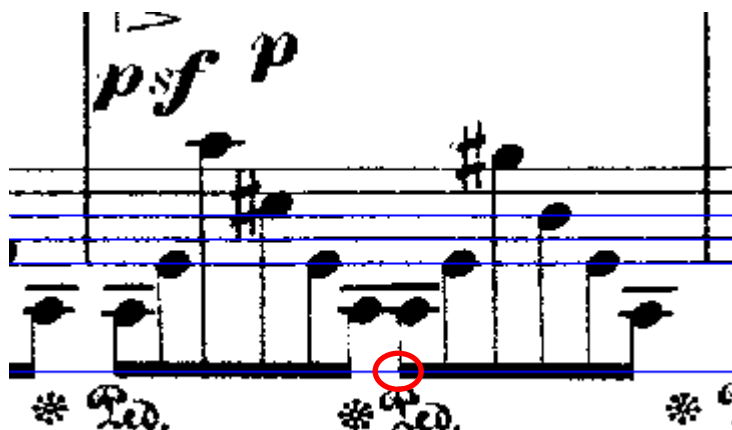
Ο λόγος που δε γίνεται περιορισμός ακόμη παχύτερων γραμμών είναι ότι λόγω των μεταβλητών διαστάσεων των εικόνων εισόδου, οι σχετικές αποστάσεις μεταξύ των γραμμών και ειδικότερα η *staff-line-distance* μπορεί να εμφανίζει διακυμάνσεις μεγέθους (σε αριθμό pixels), επομένως δεν είναι ασφαλές να περιοριστούν γραμμές που απέχουν περισσότερο από δύο pixels μεταξύ τους, καθώς μπορεί να αποτελούν διαδοχικές γραμμές πενταγράμμου σε μια εικόνα μικρών διαστάσεων και επομένως μικρής *staff-line-distance*. Γενικά, ο περιορισμός αυτός δεν αποτέλεσε εμπόδιο σε καμία παρτιτούρα από το σύνολο αυτών που εξετάστηκαν. Το αποτέλεσμα του περιορισμού των πολλαπλών γραμμών στο απόσπασμα που παρουσιάστηκε παραπάνω (βλ. Εικόνα 3.8) είναι το ακόλουθο (βλ. Εικόνα 3.11):



Εικόνα 3. 11: Αποτέλεσμα περιορισμού πολλαπλών γραμμών, απόσπασμα από *Chopin nocturnes*, σελίδα 1, [45]. Εμφανίζεται το ίδιο απόσπασμα μουσικού κειμένου με την Εικόνα 3.7 ώστε να είναι δυνατή η εποπτική σύγκριση των δύο.

Στο σημείο αυτό πρέπει να επισημανθεί ότι το κατώφλι που έχει τεθεί μπορεί να περάσουν γραμμές που δεν είναι στην πραγματικότητα γραμμές του πενταγράμμου, λόγω μεγάλης πυκνότητας μαύρων pixels, όπως φαίνεται και στην παρακάτω εικόνα (Εικόνα 3.12), δημιουργώντας με αυτό τον τρόπο το πρόβλημα των λανθασμένα αναγνωρισμένων γραμμών. Η περίπτωση αυτή θα

αντιμετωπιστεί στη συνέχεια της υλοποίησης-στο στάδιο αυτό η γραμμή αυτή θα γίνει αποδεκτή, καθώς ο αλγόριθμος δε διαθέτει (σε αυτό το αρχικό στάδιο) τρόπο να διακρίνει μια πραγματική από μια «εικονική» γραμμή πενταγράμμου και βασίζεται μόνο στην οριζόντια προβολή.



Εικόνα 3. 12: Εντοπισμός μη πραγματικής γραμμής πενταγράμμου εξαιτίας της ομαδοποίησης μεταξύ των νοτών στο κάτω μέρος του εικονιζόμενου πενταγράμμου, απόσπασμα από Chopin nocturnes, σελίδα 5 [45].

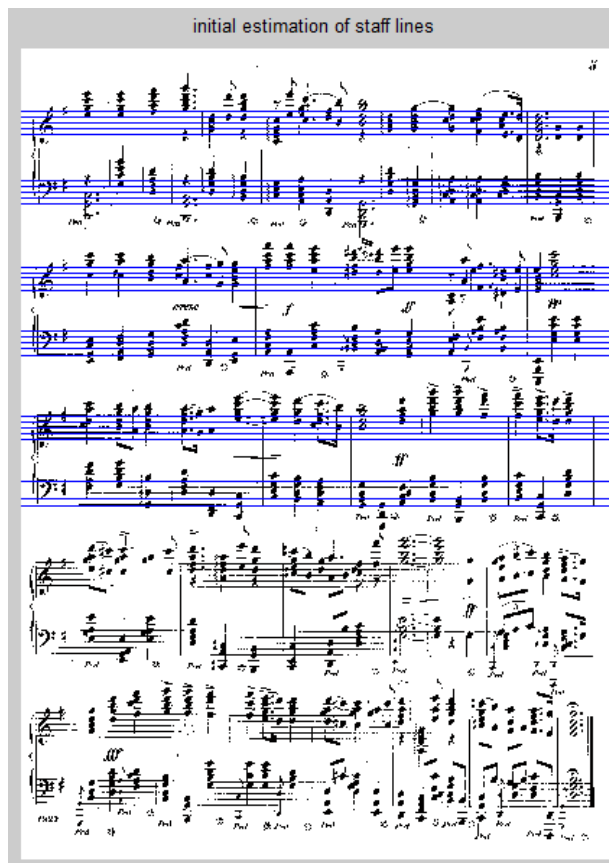
Επιπρόσθετα με την αρχική εκτίμηση των γραμμών του πενταγράμμου, σε αυτό το στάδιο υπολογίζεται και η *staff-line-distance* (η απόσταση δηλαδή μεταξύ δύο διαδοχικών γραμμών του πενταγράμμου).

Ο τρόπος υπολογισμού αυτής της απόστασης είναι ο εξής: Μετά τον περιορισμό των πολλαπλών γραμμών, υπολογίζονται οι αποστάσεις μεταξύ των διαδοχικών γραμμών πενταγράμμων. Ως *staff-line-distance* λαμβάνεται η απόσταση που συναντάται πιο συχνά στο σύνολο των υπολογισμένων αποστάσεων. Η τεχνική αυτή λειτουργεί σωστά σε όλες τις περιπτώσεις:

- Αν έχουν υπολογιστεί σωστά οι περισσότερες γραμμές πενταγράμμου γίνεται κατανοητό ότι λόγω του πλήθους τους η πιο συχνή απόσταση που εμφανίζεται μεταξύ τους θα είναι η *staff-line-distance*.
- Αν λείπουν πολλές γραμμές και από την αρχική εκτίμηση έχουν παραληφθεί μέχρι και ολόκληρα πεντάγραμμα, τότε η προσέγγιση αυτή λειτουργεί και πάλι σωστά: Λόγω της αραιότητας των γραμμών και του μικρότερου αριθμού αποστάσεων που υπολογίζεται, οι έστω και λίγες σκόρπιες διαδοχικές γραμμές πενταγράμμου που έχουν βρεθεί είναι αρκετές για να δημιουργήσουν μεγαλύτερο αριθμό ίδιων αποστάσεων (*staff-line-distance*) που θα είναι περισσότερες από τις υπόλοιπες (μεγαλύτερες) αποστάσεις που θα έχουν υπολογιστεί.
- Ακόμα κι αν έχουν εντοπιστεί μόνο δύο διαδοχικές γραμμές πενταγράμμου, και οι υπόλοιπες βρίσκονται σε διαφορετικές μεταξύ τους αποστάσεις, επειδή αν δε βρεθεί συχνότερη απόσταση μεταξύ των γραμμών επιλέγεται η μικρότερη των αποστάσεων που έχουν υπολογιστεί, αυτή προκύπτει ίση με τη *staff-line-distance* (Βλ. Εικόνα 3.45). Θα μπορούσε κανείς να υποθέσει πως ο εντοπισμός μικρού αριθμού γραμμών σε μικρότερες αποστάσεις από τη *staff-line-distance*

μπορεί να οδηγήσει σε λανθασμένη εκτίμησή της με τη μέθοδο που προτείνεται. Η υπόθεση αυτή όμως είναι λανθασμένη, καθώς το κατώφλι που χρησιμοποιείται για τον εντοπισμό γραμμών είναι αρκετά χαμηλό και έχει ήδη γίνει αντιληπτό ότι ακόμα και πραγματικές γραμμές πενταγράμμου είναι πιθανό να μην το ξεπεράσουν και να μην αναγνωριστούν. Αυτό καθιστά αδύνατο τον εντοπισμό πυκνών γραμμών που δεν αντιστοιχούν σε γραμμές πενταγράμμου. Το πρόβλημα των πυκνών γραμμών σε πολύ κοντινές αποστάσεις μπορεί να παρουσιαστεί στην περίπτωση που η εικόνα εισόδου έχει πολύ μεγάλη κλίση (0.9° και άνω), οπότε το κατώφλι αναθεωρείται και είναι πιθανό να ξεπεραστεί από μη πραγματικές γραμμές πενταγράμμου. Στην περίπτωση αυτή όμως, η εικόνα παύει να θεωρείται οριζόντια και η μέθοδος της οριζόντιας προβολής εμφανίζει περιορισμούς ως προς την ακρίβεια εντοπισμού των γραμμών της δεδομένης της μη παρέμβασης στην κλίση της εικόνας.

- Η περίπτωση στην οποία οι εντοπισμένες γραμμές δεν αρκούν για το σωστό υπολογισμό της *staff-line-distance* ή δεν υπάρχουν καθόλου γραμμές που να ξεπερνούν το κατώφλι, αντιμετωπίζεται με τρόπο που θα παρουσιαστεί στην ενότητα της αναθεώρησης κατωφλίου (βλ. Ενότητα 3.2.5).



Εικόνα 3. 13: Η αρχική εκτίμηση γραμμών είναι αρκετή για τον υπολογισμό της *staff-line-distance*, παρά την παράλειψη πολλών πενταγράμμων-η πιο συχνή απόσταση μεταξύ των εντοπισμένων γραμμών εξακολουθεί να είναι η *staff-line-distance*, απόσπασμα από Händel Largo, σελίδα 3 [48]

Πριν το τέλος του σταδίου αυτού, γίνεται υπολογισμός του αριθμού γραμμών ανά πεντάγραμμο: Αν δύο διαδοχικές γραμμές απέχουν μεταξύ τους απόσταση μικρότερη από

$$4 * staff - line - distance + staff - line - distance/4,$$

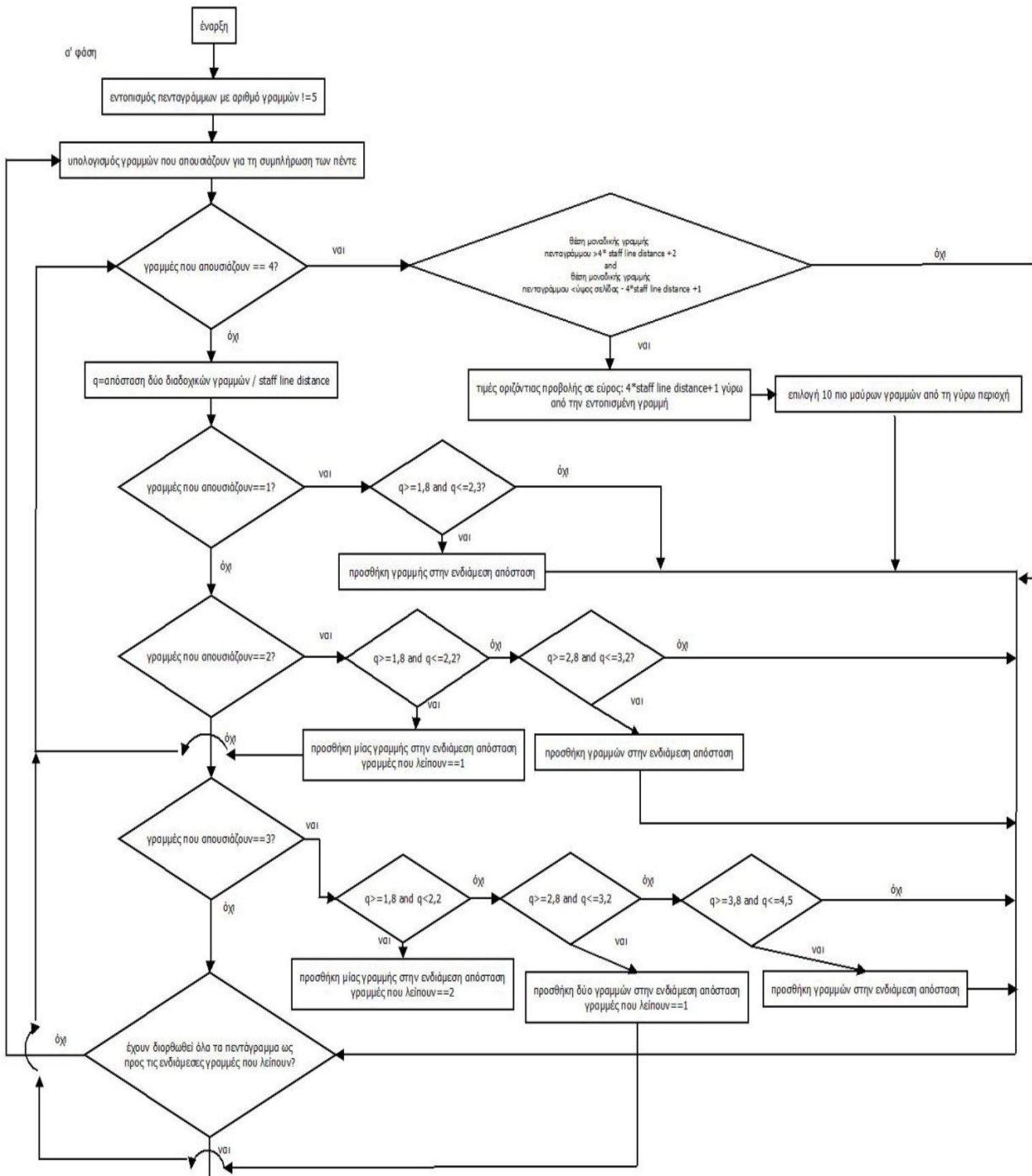
τότε θεωρείται ότι ανήκουν στο ίδιο πεντάγραμμο. Η τιμή αυτή επιλέχθηκε λόγω της γνώσης ότι το συνολικό ύψος ενός πενταγράμμου είναι περίπου ίσο με το άθροισμα των διαστημάτων μεταξύ των γραμμών- στο πεντάγραμμο τα διαστήματα αυτά είναι τέσσερα. Η τιμή είναι προσαυξημένη κατά $staff-line-distance/4$ (και όχι κατά κάποια σταθερά) προκειμένου να συμπεριλάβει το πάχος των γραμμών πενταγράμμου στον υπολογισμό καθώς και τυχόν αποκλίσεις των θέσεων των εντοπισμένων γραμμών από τις πραγματικές γραμμές.

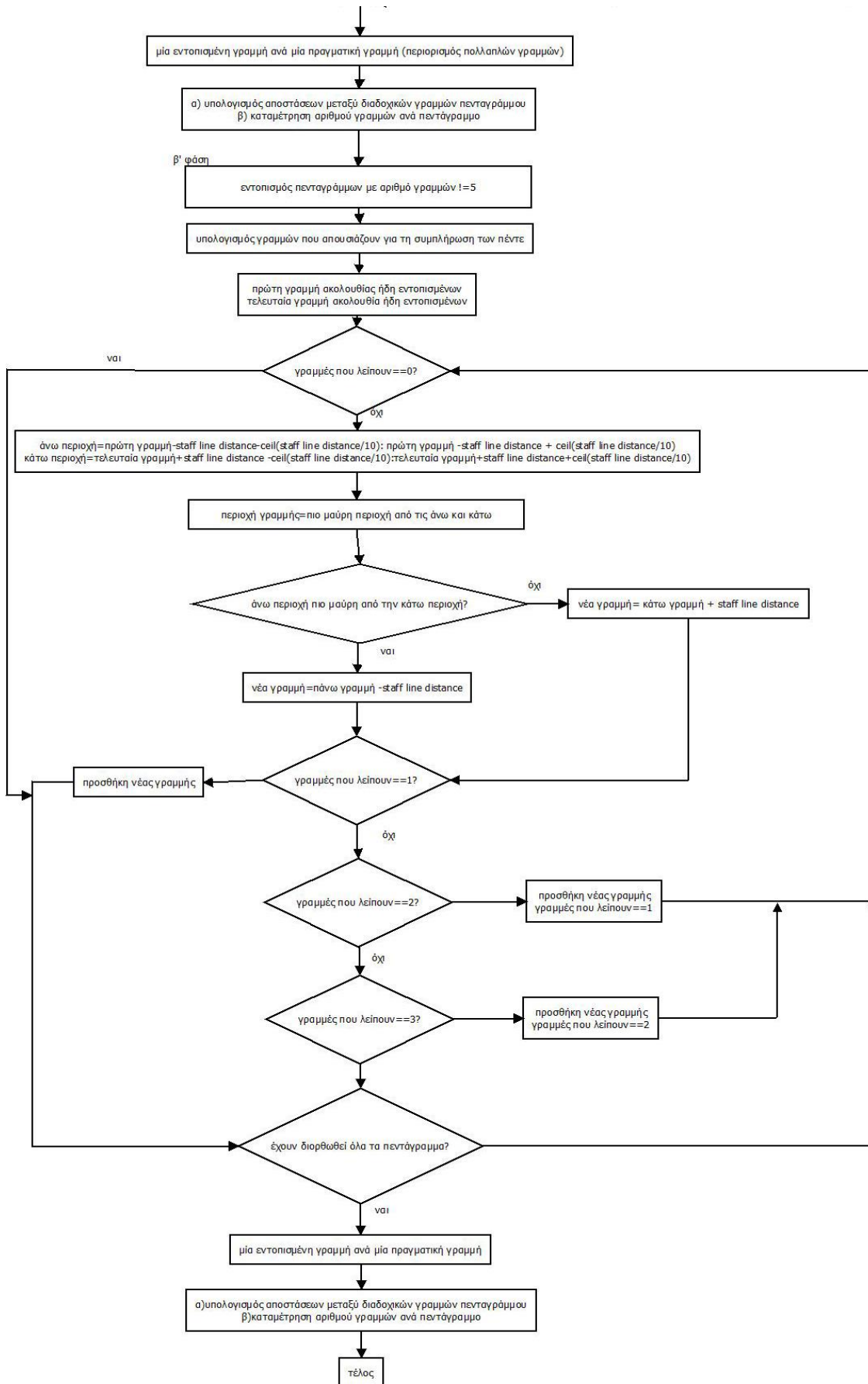
Αξίζει να παρατηρηθεί εδώ ότι σε όλα τα στάδια έχει γίνει προσπάθεια να αποφευχθεί η χρήση σταθερών προκειμένου ο αλγόριθμος να προσαρμόζεται στα δεδομένα της κάθε εικόνας ξεχωριστά, ώστε τα κατώφλια που χρησιμοποιούνται να ανταποκρίνονται κατά το δυνατόν στην υπό επεξεργασία εικόνα ώστε να επιτυγχάνουν καλύτερα αποτελέσματα.

3.2.3: Στάδιο 3^ο: Συμπλήρωση γραμμών εντοπισμένων πενταγράμμων

Στο σημείο αυτό, οι εντοπισμένες γραμμές είναι προς το παρόν εκείνες που ξεπέρασαν το κατώφλι που είχε τεθεί, με μία εντοπισμένη γραμμή ανά πραγματική γραμμή πενταγράμμου, ενώ είναι γνωστή και η $staff-line-distance$. Είναι επίσης γνωστό πόσες γραμμές έχουν εντοπιστεί ανά πεντάγραμμο.

Ακολουθεί το διάγραμμα ροής του σταδίου (Εικόνα 3.14), χωρισμένο σε δύο σελίδες (το βέλος που φαίνεται να διακόπτεται στην πρώτη σελίδα, στην πραγματικότητα συνεχίζει στην επόμενη – η ίδια λογική ακολουθείται και για όλα τα διαγράμματα που καταλαμβάνουν χώρο μεγαλύτερο της μίας σελίδας):



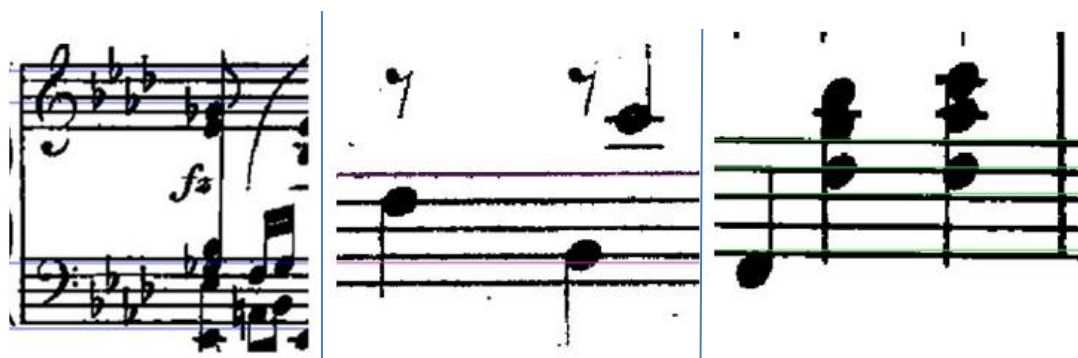


Εικόνα 3. 14: Διάγραμμα ροής σταδίου συμπλήρωσης γραμμών

Σκοπός του σταδίου αυτού είναι να συμπληρώσει τις πέντε γραμμές στα ελλειπή πεντάγραμμα, δηλαδή στα πεντάγραμμα εκείνα στα οποία δεν έχει συμπληρωθεί αριθμός πέντε εντοπισμένων γραμμών (πρόβλημα μη αναγνωρισμένων γραμμών). Επισημαίνεται ότι στο στάδιο αυτό αγνοούνται τόσο πεντάγραμμα με περισσότερες από πέντε εντοπισμένες γραμμές, όσο και πεντάγραμμα που δεν έχουν εντοπιστεί καθόλου.

Το στάδιο συμπλήρωσης γραμμών χωρίζεται σε δύο φάσεις: Τη φάση της συμπλήρωσης των γραμμών που απουσιάζουν ενδιάμεσα ήδη εντοπισμένων γραμμών και τη φάση της συμπλήρωσης των γραμμών που απουσιάζουν από τα άκρα του πενταγράμμου.

Στην πρώτη φάση, συμπληρώνονται γραμμές πενταγράμμου που απουσιάζουν από ενδιάμεσες περιοχές του πενταγράμμου (πχ. έχει βρεθεί η πρώτη και η πέμπτη γραμμή του πενταγράμμου και απουσιάζουν οι ενδιάμεσες ή έχουν βρεθεί η δεύτερη και η τέταρτη γραμμή και απουσιάζει η τρίτη-στην περίπτωση αυτή απουσιάζουν επιπλέον και η πρώτη με την πέμπτη γραμμή αλλά επειδή αυτές βρίσκονται εκτός των ορίων των εντοπισμένων γραμμών θα αντιμετωπιστούν στην επόμενη φάση, βλ. Εικόνα 3.15 για παραδείγματα έλλειψης ενδιάμεσων γραμμών). Επιπλέον, στη φάση αυτή συμπληρώνονται όλες οι γραμμές πενταγράμμου στην περίπτωση που έχει εντοπιστεί μία μόνο γραμμή στο πεντάγραμμα, ανεξάρτητα από τη θέση της γραμμής αυτής (δηλαδή αν η μοναδική γραμμή βρίσκεται στο άκρο ή στο κέντρο του πενταγράμμου, βλ. Εικόνα 3.16). Στις παρακάτω εικόνες απεικονίζονται παραδείγματα τέτοιων περιπτώσεων:



Εικόνα 3. 15: Παραδείγματα γραμμών που λείπουν ενδιάμεσα ήδη εντοπισμένων γραμμών. Από αριστερά: Chopin Concerto no2, σελίδα 11 [49]/Schumann Novelletes, σελίδα 3 [10]/ Schubert Waltzes σελίδα 4 [50].

Επισημαίνεται εδώ ότι οι διαφορετικές διαστάσεις και ποιότητα στις παρουσιαζόμενες εικόνες οφείλονται στη διακύμανση της ποιότητας στις εικόνες εισόδου αλλά και στο μέγεθος της μεγέθυνσης (zoom) που έχει επιλεγεί για την καλύτερη απεικόνιση των αποτελεσμάτων του αλγορίθμου. Όπως έχει ήδη αναφερθεί, κάποιες εικόνες εμφανίζουν αυξημένη ποιότητα αλλά και μέγεθος (με αποτέλεσμα τη μεγαλύτερη καθαρότητα και λεπτομέρεια) ενώ άλλες έχουν μειωμένη ποιότητα αλλά και μικρότερο μέγεθος. Παρόλα αυτά, λόγω της αρχικής εκτύπωσης είναι πιθανό μια μεγάλη εικόνα με λεπτομέρεια να μην είναι εξίσου καθαρή με άλλες αντίστοιχου μεγέθους.



Εικόνα 3. 16: Δύο διαφορετικά παραδείγματα μοναδικής εντοπισμένης γραμμής σε πεντάγραμμο, στη δεύτερη και στην πρώτη γραμμή αντίστοιχα (από αριστερά). Από αριστερά: Brahms Cadenzas, σελίδα 3 [12], Schubert Waltzes, σελίδα 4 [50].

Στη δεύτερη φάση, και αφού έχουν συμπληρωθεί όλες οι «εσωτερικές» γραμμές, συμπληρώνονται γραμμές που λείπουν από τα άκρα των πενταγράμμων (πχ. απουσιάζει μόνο η πέμπτη γραμμή ή απουσιάζουν η πρώτη, η δεύτερη και η πέμπτη γραμμή, ενώ οι «εσωτερικές» τρίτη και τέταρτη έχουν εντοπιστεί, βλ. Εικόνες 3.17, 3.18). Ακολουθούν παραδείγματα τέτοιων περιπτώσεων:



Εικόνα 3. 17: Από αριστερά: Έλλειψη μιας και τριών γραμμών αντίστοιχα από το άνω άκρο του πενταγράμμου, Chopin Concerto no 2, σελίδα 1, [49]



Εικόνα 3. 18: Έλλειψη τριών γραμμών (δύο από το άνω άκρο και μίας από το κάτω) και δύο γραμμών από το κάτω άκρο του πενταγράμμου, Mozart Sonata σελίδα 5 [43]

Η διαδικασία με την οποία προστίθενται οι γραμμές που λείπουν κατά την πρώτη φάση είναι (για καλύτερη κατανόηση περισσότερων λεπτομερειών της υλοποίησης, μπορεί να γίνει αντιπαραβολή με το διάγραμμα ροής - Εικόνα 3.14): Αρχικά υπολογίζεται ο αριθμός των γραμμών που λείπουν για τη συμπλήρωση του πενταγράμμου και γίνεται διάκριση σε δύο περιπτώσεις, αυτήν της έλλειψης τριών ή λιγότερων γραμμών και αυτή της έλλειψης τεσσάρων γραμμών (που είναι η περίπτωση που έχει βρεθεί μόνο μία γραμμή από ολόκληρο το πεντάγραμμο).

Στην πρώτη περίπτωση (έλλειψη τριών ή λιγότερων γραμμών), υπολογίζονται οι αποστάσεις μεταξύ των διαδοχικών εντοπισμένων γραμμών του πενταγράμμου. Η σύγκριση των αποστάσεων αυτών με τη *staff-line-distance* υποδηλώνει εάν ο λόγος τους είναι τέτοιος ώστε να θεωρηθεί ότι υπάρχει χώρος για επιπλέον γραμμή (ή γραμμές, ανάλογα με την απόσταση) στο ενδιαμέσο διάστημα. Εάν αυτό ισχύει, τότε η γραμμή (ή οι γραμμές) προστίθεται μεταξύ των ήδη εντοπισμένων γραμμών σε απόσταση *staff-line-distance* από την πρώτη γραμμή μεταξύ των δύο διαδοχικών που θεωρήθηκε ότι υπάρχει χώρος.

Στη συνέχεια, ο αριθμός των γραμμών που λείπουν για τη συμπλήρωση του πενταγράμμου ανανεώνεται και η διαδικασία επαναλαμβάνεται μέχρι να μην υπάρχουν άλλες ενδιαμέσες γραμμές προς εύρεση, δηλαδή μέχρι ο λόγος των ενδιαμέσων αποστάσεων με τη *staff-line-distance* να μην υπερβαίνει το κατώφλι που έχει τεθεί.

Στη δεύτερη περίπτωση που έχει βρεθεί μόνο μία γραμμή από ολόκληρο το πεντάγραμμο, η διαδικασία που ακολουθεί ο αλγόριθμος είναι διαφορετική: Επειδή είναι άγνωστη η θέση της εντοπισμένης γραμμής μέσα στο πραγματικό πεντάγραμμο, θεωρείται ότι η γραμμή είναι είτε η πρώτη είναι η πέμπτη του πενταγράμμου, και ο αλγόριθμος συλλέγει 4 γραμμές πάνω από την εντοπισμένη σε 4 γραμμές κάτω από αυτήν τοποθετημένες μεταξύ τους σε απόσταση ίση με *staff-line-distance*. Με αυτό τον τρόπο είναι σίγουρο ότι οι υπόλοιπες γραμμές του πενταγράμμου θα περιλαμβάνονται σίγουρα μέσα στις γραμμές που συλλέχθηκαν.

Προκειμένου να αποφευχθούν τυχόν λάθη λόγω πιθανής ελαφρά μετατοπισμένης θέσης της μοναδικής εντοπισμένης γραμμής, το πρότυπο των οκτώ (8) γραμμών, επαναλαμβάνεται μετατοπισμένο κατά ένα pixel πάνω και ένα pixel κάτω. Συνολικά δηλαδή, υπάρχουν εικοσιτέσσερις (24) υποψήφιες γραμμές πενταγράμμου γύρω από την ήδη εντοπισμένη. Από τις γραμμές αυτές επιλέγονται οι δέκα (10) με τη μικρότερη τιμή οριζόντιας προβολής (που έχει ήδη υπολογιστεί), δηλαδή με το μεγαλύτερο άθροισμα μαύρων pixels. Ο λόγος που επιλέγονται οι δέκα πρώτες (αντί για τις τέσσερις, που είναι ο αριθμός γραμμών που λείπουν) είναι αυτός που αναφέρθηκε στο στάδιο της αρχικής εκτίμησης των γραμμών: μία πολύ μαύρη, ενιαία γραμμή πάχους μεγαλύτερου του ενός pixel μπορεί να αντιπροσωπεύεται από δύο ή και τρεις εντοπισμένες γραμμές. Έτσι αν επιλεγούν μόνο τέσσερις γραμμές είναι εξαιρετικά πιθανό να περιλαμβάνονται σε αυτές τέτοιες περιπτώσεις.

Για το λόγο αυτό είναι προτιμότερο να επιλεγούν περισσότερες γραμμές οι οποίες θα περιοριστούν στη συνέχεια, τόσο με τρόπο που αναφέρθηκε στο 1^ο στάδιο (περιορισμός πολλαπλών γραμμών) όσο και με τη μέθοδο περιορισμού γραμμών που υπερβαίνουν τις πέντε για ένα πεντάγραμμο, που θα αναφερθεί στη συνέχεια. Να σημειωθεί εδώ ότι στο τέλος κάθε φάσης του σταδίου ανανεώνεται ο αριθμός των εντοπισμένων γραμμών πενταγράμμου, των μεταξύ τους αποστάσεων, καθώς και ο αριθμός των γραμμών ανά εντοπισμένο πεντάγραμμο.

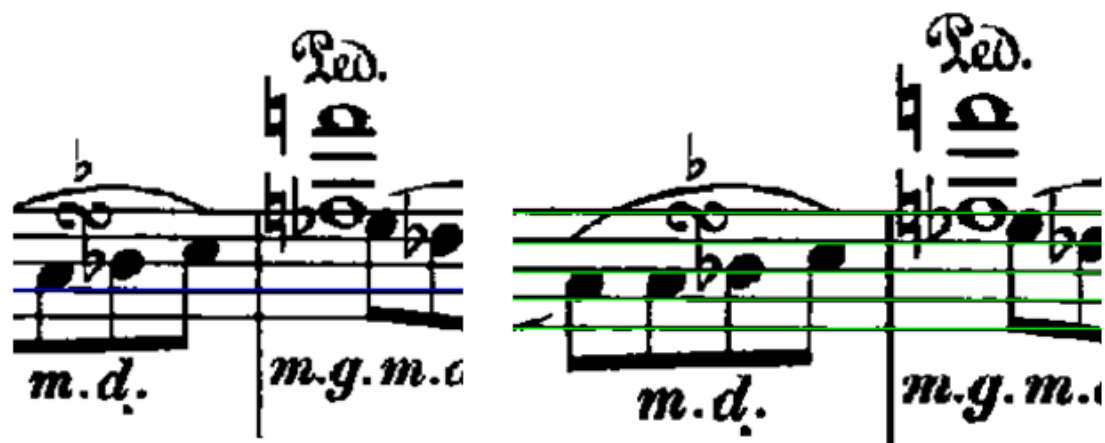
Τα αποτελέσματα του πρώτης φάσης του σταδίου παρουσιάζονται εποπτικά στις παρακάτω εικόνες σε ζεύγη (πριν και μετά τον εντοπισμό των γραμμών που απουσιάζουν):



Εικόνα 3. 19: Παράδειγμα συμπλήρωσης μίας ενδιάμεσης γραμμής (πάνω πεντάγραμμο) και τριών ενδιάμεσων γραμμών (κάτω πεντάγραμμο). Οι δύο γραμμές που εξακολουθούν να λείπουν από το πάνω πεντάγραμμο μετά τη συμπλήρωσή του θα προστεθούν στην επόμενη φάση, καθώς βρίσκονται στο άκρο του πάνω πενταγράμμου, Chopin Concerto no2, σελίδα 11 [49].



Εικόνα 3. 20: Συμπλήρωση μίας ενδιάμεσης γραμμής (αριστερά) – οι δύο κάτω γραμμές που δεν έχουν εντοπιστεί βρίσκονται σε άκρο του πενταγράμμου και θα βρεθούν στην επόμενη φάση, συμπλήρωση τεσσάρων γραμμών από μοναδική γραμμή του πενταγράμμου (δεξιά). Από αριστερά: Mozart Sonata, σελίδα 5 [43], Schubert Waltzes, σελίδα 4 [50].



Εικόνα 3. 21: Συμπλήρωση τεσσάρων γραμμών από μοναδική γραμμή του πενταγράμμου Brahms Cadenzas, σελίδα 3, [12].



Εικόνα 3. 22: Συμπλήρωση δύο συνεχόμενων ενδιάμεσων γραμμών πενταγράμμου, Chopin Concerto no 2, σελίδα 12 [49].

Στη δεύτερη φάση του σταδίου, συμπληρώνονται γραμμές που απουσιάζουν από τα άκρα του πενταγράμμου. Στο σημείο αυτό, και καθώς έχει ήδη αντιμετωπιστεί η περίπτωση της ύπαρξης μοναδικής εντοπισμένης γραμμής στο πεντάγραμμο, είναι βέβαιο ότι θα υπάρχουν τουλάχιστον δύο γραμμές ανά εντοπισμένο πεντάγραμμο και ότι ενδιάμεσως των εντοπισμένων γραμμών δε θα απουσιάζει καμιά γραμμή. Ο αλγόριθμος βρίσκει την πρώτη και την τελευταία σε σειρά εντοπισμένη γραμμή του πενταγράμμου και επιλέγει αν θα προσθέσει γραμμές πάνω ή κάτω από την περιοχή που ορίζουν, βάσει του ποια περιοχή από τις δύο έχει τη μικρότερη μέση τιμή οριζόντιας προβολής, δηλαδή τις πιο μαύρες γραμμές. Η διαδικασία που ακολουθείται, με περισσότερη λεπτομέρεια, είναι η εξής: Σε εύρος:

$staff - line - distance - ceil (staff - line - distance / 10)$ ως $staff - line - distance + ceil (staff - line - distance / 10)$,

πάνω από την πρώτη γραμμή της σειράς των διαδοχικών εσωτερικών γραμμών που έχουν ήδη εντοπιστεί, και σε αντίστοιχο εύρος κάτω από την τελευταία γραμμή της σειράς των εντοπισμένων γραμμών, λαμβάνεται η μέση τιμή της οριζόντιας προβολής των αντίστοιχων γραμμών. Η περιοχή στην οποία θα προστεθεί κάποια γραμμή είναι αυτή με τη μικρότερη μέση τιμή.

Μετά την προσθήκη της γραμμής (είτε πάνω είτε κάτω) ανανεώνονται οι τιμές της πρώτης και τελευταίας γραμμής στη σειρά των εντοπισμένων γραμμών και η διαδικασία επαναλαμβάνεται μέχρι να ολοκληρωθεί η συμπλήρωση όλων των γραμμών.

Τα αποτελέσματα αυτής της φάσης παρουσιάζονται εποπτικά στις παρακάτω εικόνες (σε ζεύγη πριν και μετά, όπως προηγουμένως):



Εικόνα 3. 23: Συμπλήρωση μοναδικής γραμμής από το άνω άκρο του πενταγράμμου, Beethoven Sonata, σελίδα 12 [46].



Εικόνα 3. 24: Συμπλήρωση τριών γραμμών στα άκρα του πενταγράμμου, δύο πάνω και ένα κάτω (άνω πεντάγραμμο) και συμπλήρωση δύο γραμμών στο κάτω άκρο του πενταγράμμου (κάτω πεντάγραμμο), Mozart Sonata, σελίδα 5, [43].



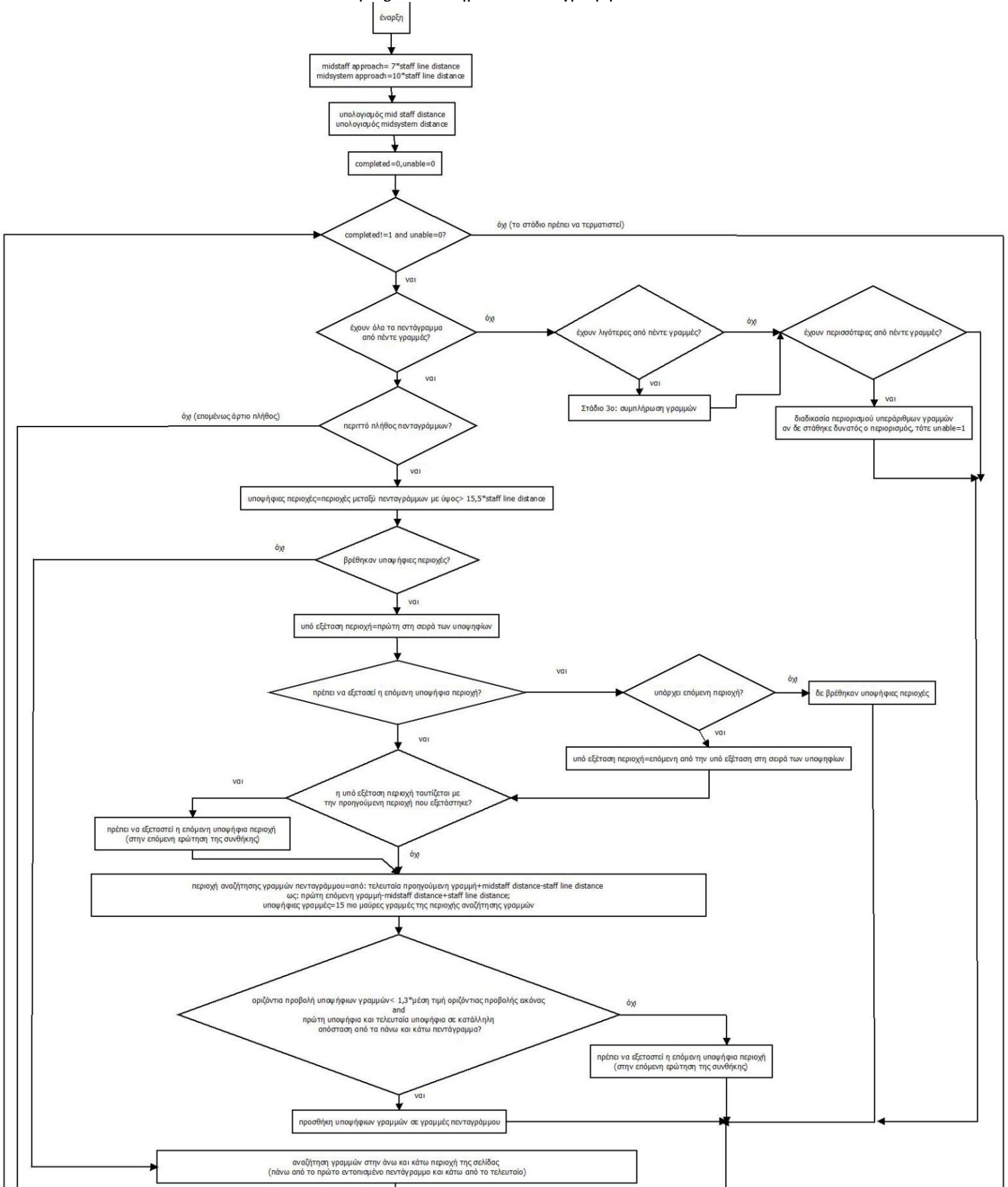
Εικόνα 3. 25: Συμπλήρωση πρώτης και πέμπτης γραμμής του πενταγράμμου, Beethoven Sonata, σελίδα 15, [46]

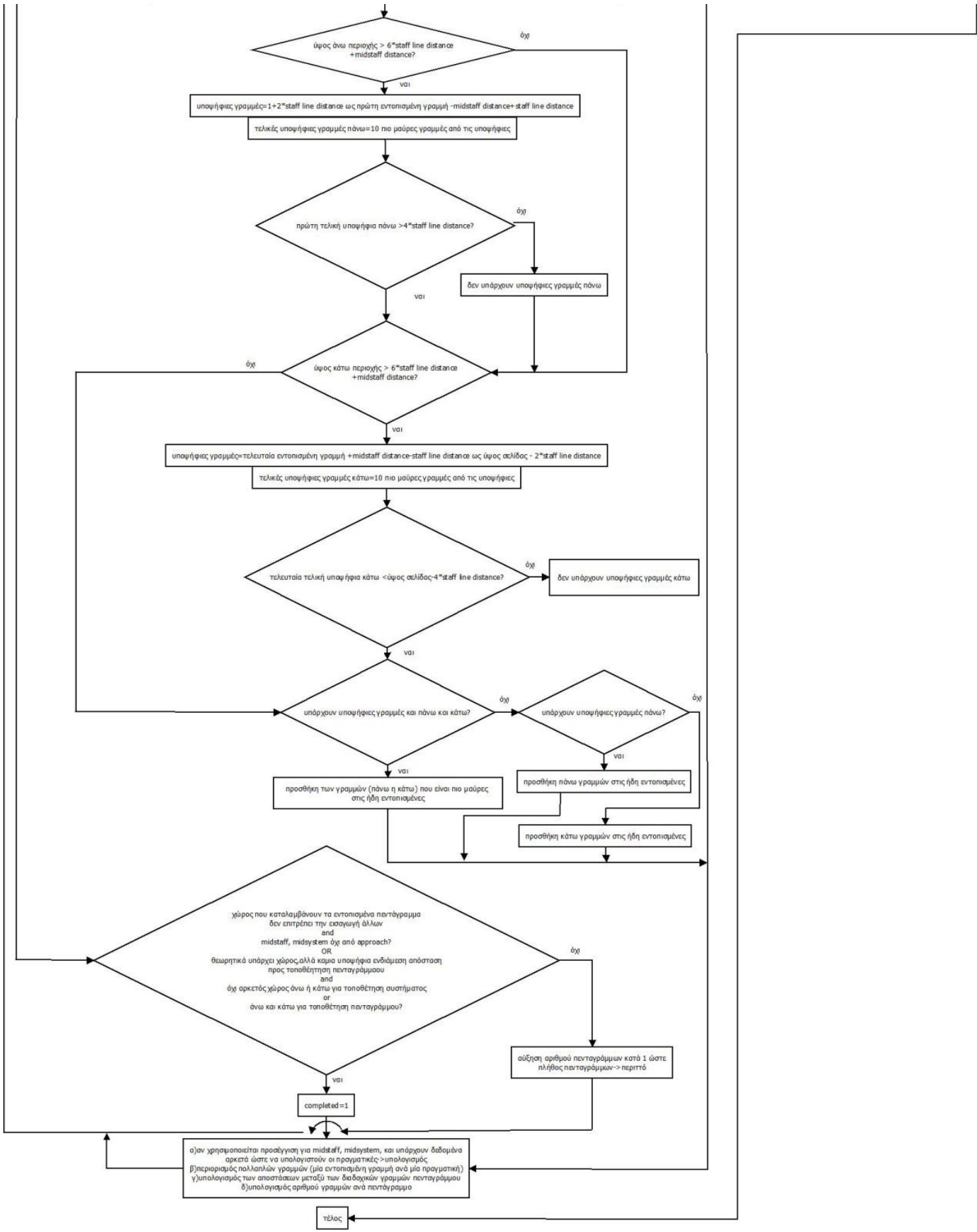


Εικόνα 3. 26: Συμπλήρωση τριών διαδοχικών γραμμών στο πάνω μέρος του πενταγράμμου, Chopin Concerto no 2 [49]

Με το πέρας αυτού του σταδίου, όλα τα εντοπισμένα πεντάγραμμα πρέπει να είναι πλήρη από γραμμές (μπορεί να υπάρχουν και κάποια με παραπάνω από πέντε γραμμές) και απομένει, για την ολοκλήρωση του βασικού τμήματος του αλγορίθμου, να εντοπιστούν τυχόν πεντάγραμμα που έχουν διαφύγει της μέχρι τώρα διαδικασίας και να περιοριστούν οι γραμμές σε πεντάγραμμα με υπεράριθμες γραμμές.

3.2.4: Στάδιο 4ο: Εντοπισμός ολόκληρων πενταγράμμων





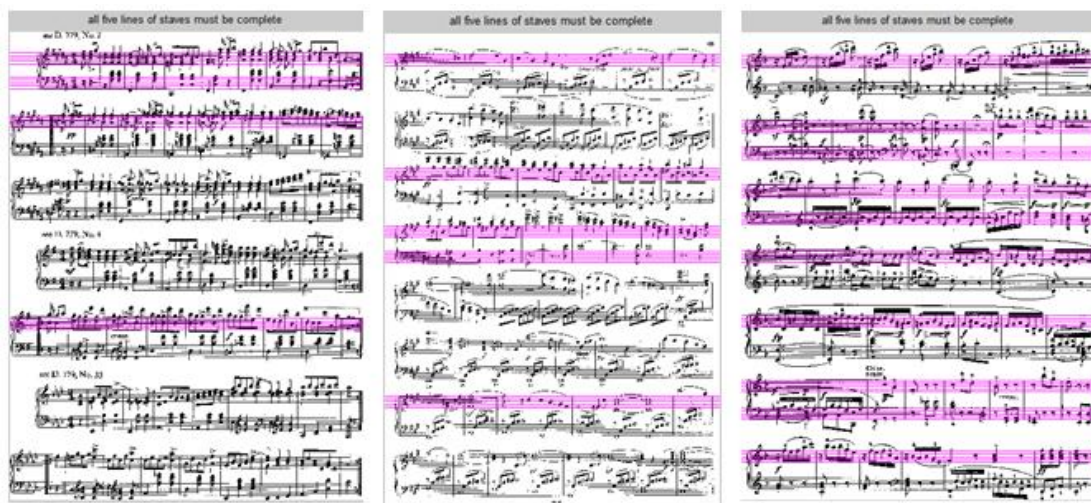
Εικόνα 3. 27: Διάγραμμα ροής σταδίου εντοπισμού πενταγράμμων

Σκοπός του σταδίου αυτού είναι να ολοκληρώσει τον εντοπισμό των γραμμών του μουσικού κειμένου, προσθέτοντας πεντάγραμμα που διέφυγαν της αρχικής εκτίμησης. Ο αριθμός των πενταγράμμων αυτών μπορεί να είναι από μηδενικός (πχ. για εικόνες εισόδου υψηλής ποιότητας με έντονες γραμμές) έως αρκετά μεγάλος (πχ. για εικόνες με κλίση, διακεκομμένες γραμμές, χαμηλή ποιότητα σάρωσης, κλπ.). Οι θέσεις των πενταγράμμων αυτών μπορεί να είναι οποιεσδήποτε, καθώς είναι δυνατό να απουσιάζουν μεμονωμένα πεντάγραμμα, διαδοχικά πεντάγραμμα που δεν ανήκουν στο ίδιο σύστημα ή ακόμα και ολόκληρα συστήματα, από όλες τις θέσεις της εικόνας, δηλαδή είτε από το πάνω και κάτω μέρος, είτε ενδιάμεσα εντοπισμένων πενταγράμμων, είτε ένας συνδυασμός των προαναφερθέντων.

Επίσης, τα νέα πεντάγραμμα που εντοπίζονται από τον αλγόριθμο εμφανίζουν τις πιθανές επιπλοκές που συναντήσαμε ως τώρα, δηλαδή χρειάζονται συμπλήρωση ή περιορισμό των γραμμών τους σε πέντε, ώστε να επιτευχθεί η μέγιστη ακρίβεια στον εντοπισμό των γραμμών του μουσικού κειμένου. Παρακάτω, παρουσιάζονται ορισμένες περιπτώσεις μουσικών κειμένων μετά την ολοκλήρωση του προηγούμενου σταδίου, προκειμένου να γίνει κατανοητό το σημαντικό εύρος των περιπτώσεων έλλειψης πενταγράμμων (Εικόνες 3.28, 3.29).



Εικόνα 3. 28: Από αριστερά: Έλλειψη ενός πενταγράμμου από το κάτω μέρος της εικόνας. Επιπλέον δύο πεντάγραμμα (2^ο και 8^ο ξεκινώντας από πάνω) εμφανίζουν έξι εντοπισμένες γραμμές λόγω γραμμών που υπάρχουν στην αρχική εικόνα αλλά δεν αποτελούν γραμμές πενταγράμμου/ Έλλειψη του δεύτερου πενταγράμμου του πρώτου συστήματος-το πεντάγραμμα που απουσιάζει δεν καταλαμβάνει όλο το πλάτος της εικόνας και επιπλέον οι γραμμές του είναι κατακερματισμένες, για αυτό και δεν εντοπίστηκε κατά το 2^ο στάδιο. Beethoven Sonata σελίδα 15 [46], Chopin Concerto no2, σελίδα 1, [49].



Εικόνα 3. 29: Έλλειψη μεγάλου αριθμού πενταγράμμων (και ολόκληρων συστημάτων). Από αριστερά: Schubert Waltzes, σελίδα 4, [50], Mendelssohn Retour, σελίδα 4 [51], Mozart Sonata, σελίδα 12, [43].

Δεδομένου ότι τα πεντάγραμμα στην παρτιτούρα του πιάνου εμφανίζονται πάντοτε σε συστήματα των δύο πενταγράμμων (με σπάνιες εξαιρέσεις στις οποίες γίνεται αναφορά στη συνέχεια - βλ. Εικόνα 3.39) και επομένως σε ζεύγη, ο αριθμός των εντοπισμένων πενταγράμμων αποτελεί ισχυρή ένδειξη για την απουσία πενταγράμμων: Αν ο αριθμός αυτός είναι περιττός, τότε είναι βέβαιο ότι απουσιάζει είτε ένα πεντάγραμμα, είτε περιττός αριθμός πενταγράμμων. Από την άλλη πλευρά, αν ο αριθμός είναι άρτιος, αυτό δε σημαίνει ότι έχουν εντοπιστεί όλα τα πεντάγραμμα, καθώς μπορεί να απουσιάζει άρτιος αριθμός πενταγράμμων (βλ. Εικόνα 3.29, πρώτη παρτιτούρα από αριστερά: ο αριθμός των εντοπισμένων πενταγράμμων είναι άρτιος και ίσος με τέσσερα, απουσιάζουν όμως δέκα πεντάγραμμα).

Επιπλέον, επειδή δεν υπάρχει κάποιο ανώτατο όριο στον αριθμό των συστημάτων που είναι δυνατό να παρουσιάζονται σε κάθε σελίδα του μουσικού κειμένου, και δεν είναι υποχρεωτικό όλες οι σελίδες του ίδιου μουσικού κειμένου να έχουν τον ίδιο αριθμό συστημάτων, ο αλγόριθμος δεν έχει τρόπο να αποφασίσει τον αριθμό των πενταγράμμων που απουσιάζουν. Ακριβώς για το λόγο αυτό, χρησιμοποιεί τη μέθοδο hill climbing (το πλαίσιο της οποίας έχει ήδη αναφερθεί ενότητα 3.1) και στο στάδιο αυτό θα παρουσιαστεί αναλυτικά ο τρόπος εφαρμογής της στον προτεινόμενο αλγόριθμο.

Το κατεξοχήν εργαλείο που κατέχει ο αλγόριθμος προκειμένου να αποφασίσει εάν μεταξύ δύο εντοπισμένων διαδοχικών πενταγράμμων είναι δυνατό να υπάρχουν ένα ή περισσότερα πεντάγραμμα (που παραμένουν ανεξικνίαστα) είναι η απόσταση μεταξύ τους, συγκεκριμένα η σχέση της απόστασης αυτής με τις mid-staff-distance και mid-system-distance. Γίνεται κατανοητό ότι αν η απόσταση μεταξύ δύο διαδοχικών πενταγράμμων ξεπερνάει κατά πολύ τη mid-staff-distance κατά περίπτωση αλλά και τη mid-system-distance σε όλες τις περιπτώσεις, αυτό αποτελεί ένδειξη ότι στην περιοχή που παρεμβάλλεται μεταξύ των εντοπισμένων πενταγράμμων υπάρχει κάποιο πεντάγραμμα που δεν έχει εντοπιστεί ακόμα.

Για το λόγο αυτό, στο παρόν στάδιο γίνεται αρχικά εκτίμηση των αποστάσεων *mid-staff-distance* και *mid-system-distance*. Προκειμένου να υπολογιστούν αυτές οι αποστάσεις, χρησιμοποιούνται οι προσεγγίσεις:

$$\textit{mid-staff-distance} = 7 * \textit{staff-line-distance}$$

και

$$\textit{mid-system-distance} = 10 * \textit{staff-line-distance}.$$

Καθώς δεν υπάρχει κάποιος κανόνας που να συνδέει τις αποστάσεις αυτές με τη *staff-line-distance*, οι τιμές αυτές υπολογίστηκαν με εμπειρικό τρόπο από τη μέτρηση των αναλογιών διαφορετικών παρτιτούρων και δοκιμάστηκαν με επιτυχία σε όλες τις παρτιτούρες που χρησιμοποιηθήκαν για τη δοκιμή της αποτελεσματικότητας του αλγορίθμου.

Ο λόγος για τον οποίο κρίθηκε απαραίτητη η χρήση προσεγγίσεων είναι ότι, για παρτιτούρες στις οποίες έχει βρεθεί πολύ μικρός αριθμός πενταγράμμων σε σχέση με τον πραγματικό, η απευθείας μέτρηση από την εικόνα θα έδινε πλήρως λανθασμένες τιμές για αυτές τις αποστάσεις, επομένως ήταν σημαντικό να υπάρχει μια προσέγγιση ως μέτρο του μεγέθους των αποστάσεων που έπρεπε να υπολογιστούν. Άλλωστε, ο κύριος σκοπός της χρήσης των προσεγγίσεων είναι η συνδρομή τους στον υπολογισμό των πραγματικών αποστάσεων βάσει των δεδομένων σε κάθε εικόνα μουσικού κειμένου. Αυτό γίνεται με τον ακόλουθο τρόπο:

Μετά τον υπολογισμό των προσεγγίσεων, ο αλγόριθμος επιχειρεί να υπολογίσει τις πραγματικές τιμές των αποστάσεων *mid-staff-distance* και *mid-system-distance*. Η λεπτομερής περιγραφή της διαδικασίας έχει παραληφθεί από το διάγραμμα ροής και παρουσιάζεται αναλυτικά εδώ:

Ο αλγόριθμος συλλέγει τις τιμές των αποστάσεων μεταξύ των διαδοχικών εντοπισμένων πενταγράμμων (ακόμα και αν αυτά περιέχουν παραπάνω από πέντε γραμμές) και συγκρίνοντας τις τιμές αυτές με τις προσεγγίσεις των *mid-staff-distance* και *mid-system-distance* τις διατηρεί ως σχετικές ή τις απορρίπτει (επιτρέποντας απόκλιση για το συσχετισμό τους με τις προσεγγίσεις, ίση με $1.5 * \textit{staff-line-distance}$ προς τα πάνω και προς τα κάτω από τις τιμές προσέγγισης). Έτσι δίνεται και ένα περιθώριο απόκλισης των πραγματικών τιμών από τις προσεγγίσεις κατά συνολικό διάστημα $3 * \textit{staff-line-distance}$. Στο σημείο αυτό προκύπτει το ακόλουθο πρόβλημα: αν οι πραγματικές αποστάσεις απέχουν περισσότερο από αυτή την τιμή $3 * \textit{staff-line-distance}$ από τις προσεγγίσεις, τότε απορρίπτονται ως μη σχετικές, και ως πραγματικές τιμές των *mid-staff-distance* και *mid-system-distance* χρησιμοποιούνται αυτές των προσεγγίσεων. Αυτό μπορεί να έχει σαν αποτέλεσμα τον εντοπισμό μη πραγματικών πενταγράμμων, πρόβλημα για το οποίο ο αλγόριθμος προτείνει μια μέθοδο αντιμετώπισης (βλ. 7^ο στάδιο).

Συνεχίζοντας, αν δεν έχουν απορριφθεί όλες οι υποψήφιες αποστάσεις, αυτές που κρατήθηκαν κατηγοριοποιούνται ως *mid-staff-distance* και *mid-system-*

distance και οι τελικές πραγματικές τιμές για τις αποστάσεις mid-staff-distance και mid-system-distance προκύπτουν από τη μέση τιμή των τιμών που συλλέχθηκαν για κάθε κατηγορία απόστασης. Με αυτό τον τρόπο, ο αλγόριθμος υπολογίζει τις mid-staff-distance και mid-system-distance από τα δεδομένα της κάθε εικόνας ξεχωριστά. Στην περίπτωση που στο βήμα αυτό οι αποστάσεις μεταξύ των πενταγράμμων είναι τέτοιες που δεν είναι δυνατό να υπολογιστούν σωστές τιμές για τις mid-staff και mid-system-distance (οι υπάρχουσες αποστάσεις αποκλίνουν σε μεγάλο βαθμό από τις προσεγγίσεις-ένδειξη ότι τα υπάρχοντα δεδομένα δεν είναι αρκετά για την εκτέλεση σωστού υπολογισμού), χρησιμοποιείται η τιμή της προσέγγισης, μέχρι να συλλεχθούν αρκετά δεδομένα ώστε να είναι ασφαλής η εκ νέου εκτίμηση των αποστάσεων. Το πρόβλημα αυτό είναι διαφορετικό από το πρόβλημα της μη επάρκειας των προσεγγίσεων που παρουσιάστηκε προηγουμένως, παρόλο που και στις δύο περιπτώσεις τελικά χρησιμοποιείται η τιμή της προσέγγισης. Η διαφορά τους έγκειται στο γεγονός ότι αν τα δεδομένα δεν είναι αρκετά, μόλις υπάρξει ανανέωσή τους σε βαθμό που να καταστήσει τη διάκριση μεταξύ mid-staff-distance και mid-system-distance δυνατή, οι προσεγγίσεις θα αντικατασταθούν από τις πραγματικές τιμές. Αν όμως τα δεδομένα παρότι αρκετά, απέχουν από τις προσεγγίσεις σε βαθμό μεγαλύτερο του αποδεκτού, τότε θα εξακολουθήσει να γίνεται χρήση της προσέγγισης.

Με γνωστές πλέον τις τιμές των mid-staff distance και mid-system-distance (αλλά και staff-line-distance ήδη από το 2^ο στάδιο), ο αλγόριθμος μπορεί πλέον να προχωρήσει στην αναζήτηση μη εντοπισμένων πενταγράμμων. Η γενική προσέγγιση που ακολουθείται είναι η επαναληπτική διερεύνηση και ανανέωση των αποστάσεων μεταξύ των πενταγράμμων, η χρήση τοπικών ελαχίστων για τον εντοπισμό νέων πενταγράμμων στις πιθανές τους θέσεις και η διόρθωση των γραμμών των νέων πενταγράμμων που εντοπίζονται μέχρι να θεωρηθεί ότι ολοκληρώθηκε ο εντοπισμός όλων των πενταγράμμων. Το γενικό πλαίσιο δεν είναι αρκετό για την κατανόηση των διαδοχικών βημάτων του αλγορίθμου και των δυσκολιών που καλείται να αντιμετωπίσει λόγω της πληθώρας των διαφορετικών περιπτώσεων που πρέπει να αντιμετωπιστούν. Για αυτό το λόγο, ακολουθεί λεπτομερέστερη περιγραφή:

Ο αλγόριθμος χρησιμοποιεί μια επαναληπτική διαδικασία: Αρχικά γίνεται έλεγχος αν όλα τα πεντάγραμμα αποτελούνται από πέντε γραμμές. Για τα πεντάγραμμα που έχουν λιγότερες γραμμές, γίνεται αναδρομή στο 3^ο στάδιο προκειμένου να συμπληρωθούν. Για τα πεντάγραμμα που έχουν περισσότερες γραμμές, καλείται η διαδικασία περιορισμού των γραμμών σε πέντε, η οποία θα περιγραφεί με λεπτομέρεια στο τέλος του σταδίου αυτού. Προς το παρόν, αρκεί η γνώση ότι η διαδικασία περιορίζει τις υπεράριθμες γραμμές του πενταγράμμου, κρατώντας μόνο εκείνες που θεωρεί ότι ανταποκρίνονται σε πραγματικές γραμμές. Αυτό σημαίνει ότι οι γραμμές που τελικά διατηρούνται μπορεί να είναι ακόμα και λιγότερες από πέντε σε αριθμό, αλλά με την αναδρομή στο 3^ο στάδιο οι γραμμές συμπληρώνονται σε πέντε. Όταν όλα τα ήδη εντοπισμένα πεντάγραμμα έχουν πλήρη αριθμό γραμμών, γίνεται έλεγχος του πλήθους των πενταγράμμων.

Εδώ ακολουθούνται δύο κατευθύνσεις:

- Αν ο αριθμός του πλήθους είναι περιττός, τότε υπάρχει οπωσδήποτε τουλάχιστον ένα πεντάγραμμο (ή περιττός αριθμός πενταγράμμων) που δεν έχει αναγνωριστεί. Ο αλγόριθμος ελέγχει τις αποστάσεις μεταξύ των διαδοχικών πενταγράμμων και καταγράφει τις αποστάσεις εκείνες που υπερβαίνουν την τιμή $15,5 * staff - line - distance$, θεωρώντας ότι σε μια τέτοια απόσταση (που υπερβαίνει κατά το ύψος ενός πενταγράμμου περίπου, τη mid-system-distance) είναι πολύ πιθανό να υπάρχει ένα μη εντοπισμένο πεντάγραμμο.

Αν βρεθεί τέτοια απόσταση (ή περισσότερες από μία τέτοιες αποστάσεις), εξετάζεται η πρώτη σε σειρά από αυτές (ξεκινώντας δηλαδή από την αρχή της σελίδας του μουσικού κειμένου και με κίνηση προς τα κάτω).

Ο αλγόριθμος καταγράφει τη θέση της τελευταίας γραμμής πενταγράμμου που έχει εντοπιστεί πριν την υπό εξέταση περιοχή καθώς και της πρώτης γραμμής πενταγράμμου μετά την υπό εξέταση περιοχή.

Η περιοχή κάτω από την τελευταία εντοπισμένη γραμμή και πάνω από την πρώτη επόμενη έχει επιλεχθεί να έχει τέτοιο εύρος, ώστε να μπορεί να εντοπιστεί το πεντάγραμμο που λείπει, είτε αυτό σχηματίζει σύστημα με το προηγούμενο εντοπισμένο (βρίσκεται δηλαδή πιο κοντά στο πάνω πεντάγραμμο), είτε σχηματίζει σύστημα με το επόμενο εντοπισμένο πεντάγραμμο (και επομένως βρίσκεται πιο κοντά στο κάτω πεντάγραμμο).

Στην περιοχή αυτή επιλέγονται οι δεκαπέντε (15) πιο μαύρες γραμμές, βάσει της τιμής της οριζόντιας προβολής τους, που έχει ήδη υπολογιστεί. Ο λόγος που επιλέγονται δεκαπέντε και όχι πέντε γραμμές είναι αυτός που περιγράφηκε στο 3^ο στάδιο κατά τη συμπλήρωση των γραμμών: Είναι προτιμότερο να επιλεγούν κάποιες γραμμές παραπάνω προκειμένου να εντοπιστούν όλες οι γραμμές του πενταγράμμου ακόμα κι αν περιέχονται σε αυτές διπλές ή και τριπλές αναγνωρίσεις της ίδιας πραγματικής γραμμής, παρά να αγνοηθεί μια πραγματική γραμμή του πενταγράμμου.

Αν η μέση τιμή οριζόντιας προβολής αυτού του συνόλου των γραμμών που επιλέχθηκαν είναι μεγαλύτερη από την τιμή $1,3 * (\text{μέση τιμή της οριζόντιας προβολής ολόκληρης της εικόνας})$ και επιπλέον η πρώτη σε σειρά από τις γραμμές που επιλέχθηκαν βρίσκεται σε απόσταση μεγαλύτερη της mid-staff-distance - staff-line-distance (το ίδιο ισχύει για την τελευταία γραμμή από τις επιλεγμένες σε σχέση με την πρώτη επόμενη ήδη εντοπισμένη γραμμή) τότε το σύνολο των γραμμών γίνεται αποδεκτό ως νέο πεντάγραμμο και οι γραμμές που το αποτελούν προστίθενται στις ήδη εντοπισμένες γραμμές πενταγράμμου.

Η τιμή $1,3 * (\text{μέση τιμή της οριζόντιας προβολής ολόκληρης της εικόνας})$ πιθανόν να φαίνεται μεγάλη σε πρώτη ανάγνωση και ειδικά σε σύγκριση με το αρχικό κατώφλι εντοπισμού γραμμών που είναι ίσο με $0,5 * (\text{μέση τιμή της οριζόντιας προβολής})$, ο λόγος όμως που χρησιμοποιείται είναι ότι τα πεντάγραμμο που δεν έχουν εντοπιστεί από την αρχική εκτίμηση, είναι πεντάγραμμο που έχουν γενικά μεγαλύτερο ποσοστό λευκών pixels από τα ήδη εντοπισμένα. Επιπλέον, η δοκιμή σε πολλαπλές παρτιτούρες απέδειξε ότι η χρήση χαμηλότερου κατωφλίου αποτύγχανε στον εντοπισμό των αγνοούμενων πενταγράμμων.

Παρόλα αυτά, είναι σωστή η εκτίμηση ότι αυτό το αρκετά υψηλό κατώφλι μπορούν να ξεπεράσουν και γραμμές που δεν αποτελούν γραμμές πενταγράμμου, αλλά έχουν σημαντικό αριθμό μαύρων pixels επειδή περιλαμβάνουν κείμενο, σχόλια ερμηνείας κλπ. Εξαιτίας

ακριβώς αυτής της διαπίστωσης υπάρχει το δεύτερο σκέλος της συνθήκης, ότι οι γραμμές που γίνονται αποδεκτές πρέπει να βρίσκονται σε τέτοιες αποστάσεις από τις ήδη εντοπισμένες γραμμές ώστε να μπορούν να αποτελέσουν σύστημα (είτε με το πάνω είτε με το κάτω εντοπισμένο πεντάγραμμο) και με αυτό τον τρόπο περιορίζονται οι περιπτώσεις αναγνώρισης λανθασμένων γραμμών που περιέχουν σχόλια και βρίσκονται πολύ κοντά στα πεντάγραμμα.

Μια άλλη προσέγγιση θα ήταν να γίνεται καταμέτρηση των πενταγράμμων που έχουν ήδη εντοπιστεί προκειμένου να είναι γνωστό αν το υπό αναζήτηση πεντάγραμμο είναι το άνω ή το κάτω πεντάγραμμο ενός συστήματος: Η προσέγγιση αυτή δεν είναι καθόλου ασφαλής όμως, καθώς είναι πιθανό να απουσιάζουν πεντάγραμμο από την αρχή του μουσικού κειμένου, και καθώς αυτή η απουσία δεν αντικατοπτρίζεται στις αποστάσεις μεταξύ διαδοχικών πενταγράμμων, είναι μεγάλος ο κίνδυνος για λάθος εκτίμηση. Ένα επιπλέον προτέρημα του τρόπου με τον οποίο γίνεται η αναζήτηση πενταγράμμων είναι ότι αν στην ενδιάμεση απόσταση (την υπό εξέταση) απουσιάζει αριθμός πενταγράμμων μεγαλύτερος του ενός, η περιοχή στην οποία γίνεται η αναζήτηση και ο αριθμός των γραμμών που διατηρούνται ως γραμμές πενταγράμμου είναι τέτοια ώστε να είναι δυνατός ο ταυτόχρονος εντοπισμός πολλαπλών πενταγράμμων (βλ. Εικόνα 3.32).

Στη συνέχεια, ανανεώνονται οι αποστάσεις μεταξύ των γραμμών και των πενταγράμμων, ο αριθμός γραμμών ανά πεντάγραμμο και η διαδικασία επαναλαμβάνεται, με διόρθωση των υπεράριθμων ή ελλειπών πενταγράμμων με τους τρόπους που προαναφέρθηκαν, έλεγχο του πλήθους πενταγράμμων κοκ. Μετά τη συμπλήρωση του ενός πενταγράμμου ο αλγόριθμος να φτάνει στην περίπτωση του άρτιου αριθμού πενταγράμμων, όπου γίνεται έλεγχος της συνθήκης τερματισμού του σταδίου.

- Αν το πλήθος των πενταγράμμων είναι άρτιο τότε είτε έχουν εντοπιστεί όλα τα πεντάγραμμο είτε λείπει άρτιος αριθμός πενταγράμμων. Στην περίπτωση αυτή, ελέγχεται ο χώρος που καταλαμβάνουν τα ήδη εντοπισμένα πεντάγραμμο (των οποίων ο αριθμός είναι γνωστός) μαζί με τις mid-staff και mid-system αποστάσεις που τους αντιστοιχούν. Αυτός ο έλεγχος αποτελεί και τη συνθήκη τερματισμού του σταδίου: Αν ο χώρος αυτός είναι τέτοιος ώστε να θεωρηθεί ότι δε χωράει τουλάχιστον ένα ολόκληρο ακόμα σύστημα (σε σχέση με το ύψος της εικόνας του μουσικού κειμένου) και επιπλέον οι τιμές που χρησιμοποιούνται για τη mid-staff-distance και τη mid-system distance δεν είναι οι προσεγγίσεις (αρκεί αυτό να συμβαίνει για μία από τις δύο), τότε η αναζήτηση τερματίζεται.

Εδώ αναδεικνύεται και η σημασία του υπολογισμού των mid-staff-distance και mid-system-distance καθώς η χρήση της σωστής τους τιμής είναι αυτή που καθορίζει αν θα τερματιστεί ή όχι το στάδιο, καθώς η διαφορά τους από την τιμή προσέγγισης ακόμα και για μερικά pixels μπορεί να επηρεάσει το συνολικό αποτέλεσμα της αναζήτησης των πενταγράμμων.

Η συνθήκη για τις προσεγγίσεις εισάγεται για να γίνεται ακριβέστερη προσαρμογή του αλγορίθμου στα δεδομένα της υπό επεξεργασία εικόνας μουσικού κειμένου και για να βεβαιώνεται ότι πριν τερματιστεί ο αλγόριθμος, οι τιμές που έχουν χρησιμοποιηθεί για τον υπολογισμό του διαστήματος που καταλαμβάνουν τα συστήματα στη σελίδα του

μουσικού κειμένου ανταποκρίνονται στην πραγματικότητα. Σκοπός δηλαδή είναι να αποφευχθεί ο πρόωρος τερματισμός της διαδικασίας αναζήτησης πεντάγραμμων.

Εκτός από αυτή την περίπτωση, το στάδιο τερματίζεται αν, παρά το ότι θεωρείται ότι υπάρχει επιπλέον χώρος για πεντάγραμμο, καμία απόσταση ενδιάμεσα των πεντάγραμμων δεν υπερβαίνει το κατώφλι διερεύνησης για νέο πεντάγραμμο και επιπλέον δεν υπάρχει αρκετός χώρος για ολόκληρο σύστημα στο πάνω μέρος της σελίδας ή στο κάτω μέρος της σελίδας ή ταυτόχρονα χώρος για ένα πεντάγραμμο στο άνω και ένα πεντάγραμμο στο κάτω μέρος της σελίδας (δεδομένου ότι αν λείπουν επιπλέον πεντάγραμμο θα πρέπει αυτά να είναι δύο).

Αν το στάδιο δεν τερματιστεί, η διαδικασία συνεχίζεται με την εικονική αύξηση του αριθμού των πεντάγραμμων κατά ένα ώστε το πλήθος τους να θεωρηθεί περιττό και να συνεχιστεί η αναζήτηση. Γίνεται επανεκτίμηση των *mid-staff-distance* και *mid-system-distance* στην περίπτωση που χρησιμοποιείται ακόμα η προσέγγιση για τις αποστάσεις αυτές. Δεδομένου ότι τώρα ο αριθμός των πεντάγραμμων θεωρείται περιττός, γίνεται αναζήτηση των αποστάσεων μεταξύ των πεντάγραμμων, όπως περιγράφηκε στην περίπτωση του περιττού αριθμού πεντάγραμμων.

Στο σημείο αυτό θα περιγραφεί η εξέλιξη του αλγορίθμου στην περίπτωση που δε βρεθεί απόσταση τέτοια που να θεωρηθεί ότι περικλείει μη εντοπισμένο πεντάγραμμο (συστήνεται η αναφορά και στο διάγραμμα ροής για καλύτερη κατανόηση - Εικόνα 3.27): Γίνεται αναζήτηση στην περιοχή πάνω από την πρώτη εντοπισμένη γραμμή όλης της σελίδας και στην περιοχή κάτω από την τελευταία εντοπισμένη γραμμή όλης της σελίδας, δηλαδή στην άνω και κάτω περιοχή της εικόνας του μουσικού κειμένου. Αν αυτή η περιοχή θεωρηθεί ικανή ώστε να χωράει ένα πεντάγραμμο (η απόσταση από την πρώτη γραμμή της εικόνας για την άνω περιοχή ή από την τελευταία γραμμή για την κάτω περιοχή είναι μεγαλύτερη από $6 * staff - line - distance + mid - staff - distance$, τότε επιλέγονται οι δέκα πιο μαύρες γραμμές σύμφωνα με την τιμή της οριζόντιας προβολής τους, ως υποψήφιας γραμμές πεντάγραμμου.

Αν η πρώτη υποψήφια γραμμή για την άνω περιοχή (ή η τελευταία για την κάτω περιοχή) βρίσκεται σε απόσταση μεγαλύτερη από $4 * staff - line - distance$ από την αρχή της εικόνας (ή από το τέλος της εικόνας για την κάτω περιοχή) τότε οι υποψήφιας γραμμές γίνονται δεκτές. Το κατώφλι αυτό θεσπίζεται καθώς θεωρείται αδύνατο να βρεθεί πεντάγραμμο τόσο κοντά στην άκρη της εικόνας, λόγω των στοιχείων που προεξέχουν (πχ. κλειδί του σολ, γραμμές μουσικής σημειογραφίας) και αν εντοπιστεί «πεντάγραμμο» τόσο κοντά στην άκρη το πιθανότερο είναι να αντιστοιχεί είτε σε κείμενο στην αρχή της εικόνας, είτε σε συνδεδεμένες νότες που βρίσκονται πάνω από το πεντάγραμμο είτε σε διάφορες γραμμές σχετικές με το ύφος εκτέλεσης του κομματιού, είτε ακόμα και σε θόρυβο και είναι εν γένει εξαιρετικά σπάνιο να ανήκει σε πεντάγραμμο. Εδώ ακριβώς έγκειται η εκμετάλλευση του χαρακτηριστικού αυτού των παρτιτούρων που αναφέρθηκε στο 1^ο στάδιο. Αν βρεθεί και στις δύο περιοχές (άνω και κάτω) υποψήφιο πεντάγραμμο, τότε επιλέγεται να κρατηθεί το πεντάγραμμο με την πιο μικρή μέση τιμή οριζόντιας προβολής, δηλαδή το πεντάγραμμο με τα πιο μαύρα pixels. Και πάλι, δεν επιλέγονται μόνο πέντε γραμμές, εξαιτίας των λόγων που έχουν προαναφερθεί.

Ακολουθούν κάποια παραδείγματα εντοπισμού πενταγράμμων (οι εικόνες εμφανίζονται σε ζεύγη πριν τον εντοπισμό και μετά τον εντοπισμό ώστε να γίνεται αντιληπτή η συμπλήρωση των πενταγράμμων):



Εικόνα 3. 30: Διαδοχικά στιγμιότυπα του τελευταίου συστήματος της Beethoven Sonata, σελίδα 15 [46]. Πάνω απουσιάζει ολόκληρο το δεύτερο πεντάγραμμο, αλλά με αναζήτηση στην κάτω περιοχή της εικόνας, και μετά τον περιορισμό των διπλών/τριπλών γραμμών όλες οι γραμμές του πενταγράμμου, οι γραμμές του δεύτερου πενταγράμμου (κάτω) έχουν εντοπιστεί σωστά και δε χρειάζονται καμία συμπλήρωση.



Εικόνα 3. 31: Εμφάνιση διαδοχικών στιγμιότυπων του ίδιου αποσπάσματος της παρτιτούρας. Τοποθέτηση επιπλέον πενταγράμμου (2^ο πεντάγραμμο από πάνω) χωρίς να είναι απαραίτητη η συμπλήρωση ή η ελάττωση των γραμμών του, Chopin Concerto no 2 [49].



Εικόνα 3. 32: Διαδοχικά στιγμιότυπα του κάτω μέρους της σελίδας 3 του Händel Largo [48]. Πάνω αριστερά: Απουσιάζουν 4 διαδοχικά πεντάγραμμα (δύο συστήματα) από το κάτω μέρος του πενταγράμμου, πάνω δεξιά: εύρεση κάποιων γραμμών τριών από τα τέσσερα πεντάγραμμα ταυτόχρονα, κάτω αριστερά: συμπλήρωση των γραμμών- παρατηρείται ότι οι θέσεις των εντοπισμένων γραμμών δεν ανταποκρίνονται ακριβώς στις πραγματικές, κάτω δεξιά: Εύρεση και του πενταγράμμου που απουσίαζε ως τώρα. Οι θέσεις των γραμμών θα διορθωθούν σε επόμενο στάδιο.



Εικόνα 3. 33: Δύο στιγμιότυπα της πάνω πλευράς της σελίδας 5, Schumann Novellettes [10], Αριστερά φαίνεται η έλλειψη του πρώτου πενταγράμμου της σελίδας, δεξιά το πρώτο πεντάγραμμα έχει πλέον εντοπιστεί με όλες τις γραμμές του (έχει παραληφθεί η απεικόνιση του ενδιάμεσου σταδίου που βρίσκει τις τέσσερις από τις πέντε γραμμές του πρώτου πενταγράμμου).



Εικόνα 3. 34: Διαδοχικά στιγμιότυπα εντοπισμού γραμμών πενταγράμμου, με κύκλο έχει επισημανθεί η περιοχή που μεταβάλλεται ή αναμένεται να μεταβληθεί. Αριστερά: Ο αλγόριθμος θα διερευνήσει τη μοναδική μεγάλη απόσταση μεταξύ διαδοχικών πενταγράμμων, Δεξιά: Εντοπίζονται ταυτόχρονα δύο πεντάγραμμα και συμπληρώνονται οι γραμμές των νέων πενταγράμμων, η επόμενη προς διερεύνηση απόσταση φαίνεται κυκλωμένη, Schubert Waltzes, σελίδα 4 [50].



Εικόνα 3. 35: Συνέχεια από Εικόνα 3.34. Αριστερά: Εντοπίζονται ταυτόχρονα δύο πεντάγραμμα σε σχέση με την προηγούμενη εικόνα, Δεξιά: Οι γραμμές τους συμπληρώνονται ενώ η επόμενη υποψήφια περιοχή φαίνεται κυκλωμένη- στη συνέχεια συμπληρώνεται το πεντάγραμμα που έλειπε, και γίνεται αναζήτηση στην άνω και κάτω περιοχή της σελίδας, για την κάτω περιοχή εντοπίζονται ταυτόχρονα τρία πεντάγραμμα, Schubert Waltzes, σελίδα 4 [50].

all five lines of staves must be complete

Op. 9, No. 2

Op. 9, No. 4

Op. 9, No. 13

Εικόνα 3. 36: Συνέχεια από Εικόνα 3.35. Οι γραμμές των νέων πενταγράμμων συμπληρώνονται, η επόμενη (και μοναδική) υποψήφια περιοχή φαίνεται με κόκκινο, Δεξιά: Συμπληρώνεται και το τελευταίο πεντάγραμμο, Schubert Waltzes, σελίδα 4 [50]

Αφού περιγράφηκε ο βασικός σκελετός του αλγορίθμου στο στάδιο αυτό, θα ακολουθήσει περιγραφή των επιπλέον ιδιοτήτων και διαδικασιών που έχουν αναπτυχθεί στο κομμάτι αυτό του αλγορίθμου προκειμένου να επιτυγχάνεται καλύτερη αναγνώριση των γραμμών πενταγράμμου και να εξαλειφεται πλήρως ο κίνδυνος να εγκλωβίζεται ο αλγόριθμος τοπικά σε κάποια σημεία της εικόνας.

Όπως έχει ήδη παρουσιαστεί, στην περίπτωση που ο αλγόριθμος εντοπίζει περιοχές μεταξύ των ήδη εντοπισμένων πενταγράμμων όπου θεωρεί ότι υπάρχει χώρος για ένα επιπλέον πεντάγραμμο ασχολείται με την πρώτη σε σειρά από αυτές τις περιοχές. Διορθώνοντας το πρόβλημα, εντοπίζοντας δηλαδή το αγνοούμενο πεντάγραμμο, κατά την επόμενη επανάληψη όπου ξαναγίνεται αναζήτηση των αποστάσεων μεταξύ των πενταγράμμων, η προηγούμενη απόσταση θα έχει εξαλειφθεί λόγω της τοποθέτησης του νέου πενταγράμμου και λογικά πρώτη σε σειρά θα βρίσκεται τώρα εκείνη η απόσταση που προηγουμένως ήταν δεύτερη.

Τι θα συμβεί όμως στην περίπτωση που στην προηγούμενη υποψήφια περιοχή δεν έγινε αποδεκτό το υποψήφιο πεντάγραμμο επειδή δεν πληρούσαν οι προϋποθέσεις που έχουν τεθεί (σχετικά με τη μέση τιμή οριζόντιας προβολής και τις αποστάσεις από τα γειτονικά εντοπισμένα πεντάγραμματα);

Στην περίπτωση αυτή, η υπό εξέταση (πρώτη σε σειρά) απόσταση είναι και πάλι η ίδια που απορρίφθηκε κατά την προηγούμενη επανάληψη του αλγορίθμου. Αν αυτή η διαδικασία συνεχιζόταν ως έχει, τότε ο αλγόριθμος θα εγκλωβιζόταν στην περιοχή αυτή, αγνοώντας τις επόμενες υποψήφιες περιοχές που πιθανόν να περιέχουν μη εντοπισμένα πεντάγραμματα.

Αυτό που συμβαίνει είναι το εξής: πριν τον έλεγχο μιας περιοχής υποψήφιας προς διερεύνηση, εξετάζεται αν η τρέχουσα περιοχή βρίσκεται στην ίδια θέση στον πίνακα των αποστάσεων και έχει την ίδια τιμή εύρους με την περιοχή που ελέγχθηκε στην προηγούμενη επανάληψη. Αν αυτό συμβαίνει, τότε έχουμε την αύξηση ενός μετρητή, ο οποίος στην επόμενη επανάληψη θα οδηγήσει τον αλγόριθμο στην εξέταση της επόμενης από την πρώτη απόστασης στον πίνακα των υποψήφιων αποστάσεων. Με αυτό τον τρόπο, αν αυτό συμβεί και για τη δεύτερη απόσταση, ο αλγόριθμος θα προχωρήσει στην τρίτη κοκ.

Αν εξαντληθούν όλες οι υποψήφιες αποστάσεις, τότε τίθεται μια σημαία που εξισώνει τον πίνακα των υποψήφιων αποστάσεων με το κενό. Να υπενθυμιστεί εδώ ότι έλεγχος αυτού του πίνακα γίνεται και στη συνθήκη τερματισμού του αλγορίθμου. Γίνεται κατανοητό λοιπόν, ότι για να προσπεράσει ο αλγόριθμος μια περιοχή που αποδεικνύεται ότι δεν περιέχει πεντάγραμμα, πρέπει αυτή να προσπελαστεί δύο φορές, την πρώτη που απορρίπτεται και τη δεύτερη που γίνεται αντιληπτό ότι έχει επανελεγχθεί η ίδια περιοχή, οπότε και τίθεται η σημαία (βλ. Εικόνα 3.37).

Μία ακριβώς αντίστοιχη διαδικασία υπάρχει και για τον έλεγχο της άνω και κάτω περιοχής του πενταγράμμου, καθώς είναι συχνό το φαινόμενο να υπάρχει αρκετός κενός χώρος πάνω και κάτω από το μουσικό κείμενο, επομένως αφού οι υποψήφιες περιοχές θα απορρίπτονταν αφού δε θα εντοπιζόνταν πεντάγραμμα στο εύρος τους- ο αλγόριθμος και πάλι θα εγκλωβιζόνταν εκεί αν δεν υπήρχε τρόπος να προχωρήσει.

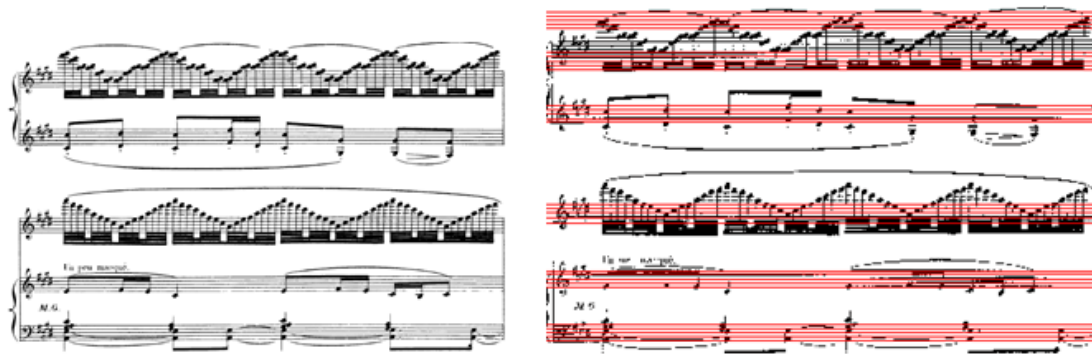
Εικόνα 3. 37: Διαδοχικά στιγμιότυπα εντοπισμού πενταγράμμων. Παρά το ότι όλα τα πραγματικά πεντάγραμμα έχουν εντοπιστεί (πάνω αριστερά), ο αλγόριθμος, λόγω της μεγάλης απόστασης μεταξύ των δύο πρώτων συστημάτων θεωρεί ότι παρεμβάλλεται πεντάγραμμα (πάνω κέντρο), αλλά στη συνέχεια διορθώνει την εκτίμηση του (πάνω δεξιά), η ακολουθία της πάνω σειράς επαναλαμβάνεται μία ακόμη φορά μέχρι να τεθεί η σημαία ότι πρέπει να εξεταστεί η επόμενη περιοχή και έτσι επειδή και η επόμενη απόσταση μεταξύ των συστημάτων θεωρείται μεγάλη, ο αλγόριθμος εντοπίζει και εκεί μη πραγματικό πεντάγραμμα (κάτω αριστερά), κίνηση που στη συνέχεια διορθώνεται (κάτω κέντρο), Beethoven Sonata, σελίδα 10 [46]

Κατά τη διάρκεια εντοπισμού των πενταγράμμων, είναι πιθανό, λόγω των επιλεγμένων κατωφλίων, να εντοπιστούν γραμμές οι οποίες δεν αντιστοιχούν σε πεντάγραμμα, καθώς όπως αναφέρθηκε θεωρείται προτιμότερο να εντοπιστούν λανθασμένα πεντάγραμμα παρά να αγνοηθούν πραγματικά. Τέτοιες γραμμές

μπορεί να αντιστοιχούν σ-ε κάποιο κείμενο που υπάρχει μέσα στην παρτιτούρα, πχ. στον τίτλο ή σε κάποια ενδιάμεσα σχόλια κτλ. Επειδή στο παρόν στάδιο ο αλγόριθμος δεν έχει τρόπο να διαχωρίσει τα πραγματικά από τα μη πραγματικά πεντάγραμμα, τα επεξεργάζεται όλα σαν πραγματικά (πχ. συμπληρώνει τις γραμμές τους, ή τις περιορίζει αν υπερβαίνουν τις πέντε). Ειδικά στις born digital παρτιτούρες, καθώς δεν αποτελούν υποχρεωτικά επίσημες εκδόσεις (οι οποίες είναι προσεγμένες και ακολουθούν όλους τους κανόνες σωστής και ενιαίας διάταξης στα συστήματά τους) τα κενά μεταξύ των πενταγράμμων και των συστημάτων μπορεί να είναι πολύ μεγαλύτερα των προσεγγίσεων, ενδιάμεσως εντοπίζονται αρκετά μη πραγματικά πεντάγραμμα (βλ. Εικόνα 3.38). Τα μη πραγματικά πεντάγραμμα αντιμετωπίζονται σε μεταγενέστερο στάδιο του αλγορίθμου. Παρόλα αυτά η εισαγωγή τους στο παρόν στάδιο μπορεί να αποδειχτεί χρήσιμη, καθώς στη σπάνια περίπτωση της ύπαρξης περιττού αριθμού πενταγράμμων, η εισαγωγή ενός επιπλέον μη υπαρκτού πενταγράμμου επιτρέπει στο στάδιο αναζήτησης πενταγράμμων να τερματιστεί έτσι ώστε ο αλγόριθμος να προχωρήσει στο επόμενο στάδιο (βλ. Εικόνα 3.39).



Εικόνα 3. 38: Εντοπισμός μη πραγματικού πενταγράμμου (κέντρο) και στη συνέχεια διόρθωση (δεξιά) λόγω της πολύ μεγάλης απόστασης μεταξύ των διαδοχικών συστημάτων σε born digital παρτιτούρα. Ο εντοπισμός των επιπλέον γραμμών θα είχε αποφευχθεί αν γίνονταν χρήση ενός χαμηλότερου κατωφλίου αποδοχής (πιο μαύρες γραμμές), αλλά εξαιτίας της χρήσης ενιαίου μεγαλύτερου κατωφλίου τόσο για scanned όσο και για born digital παρτιτούρες, εμφανίζεται το παραπάνω φαινόμενο, Bizet Carmen Prelude, [47]

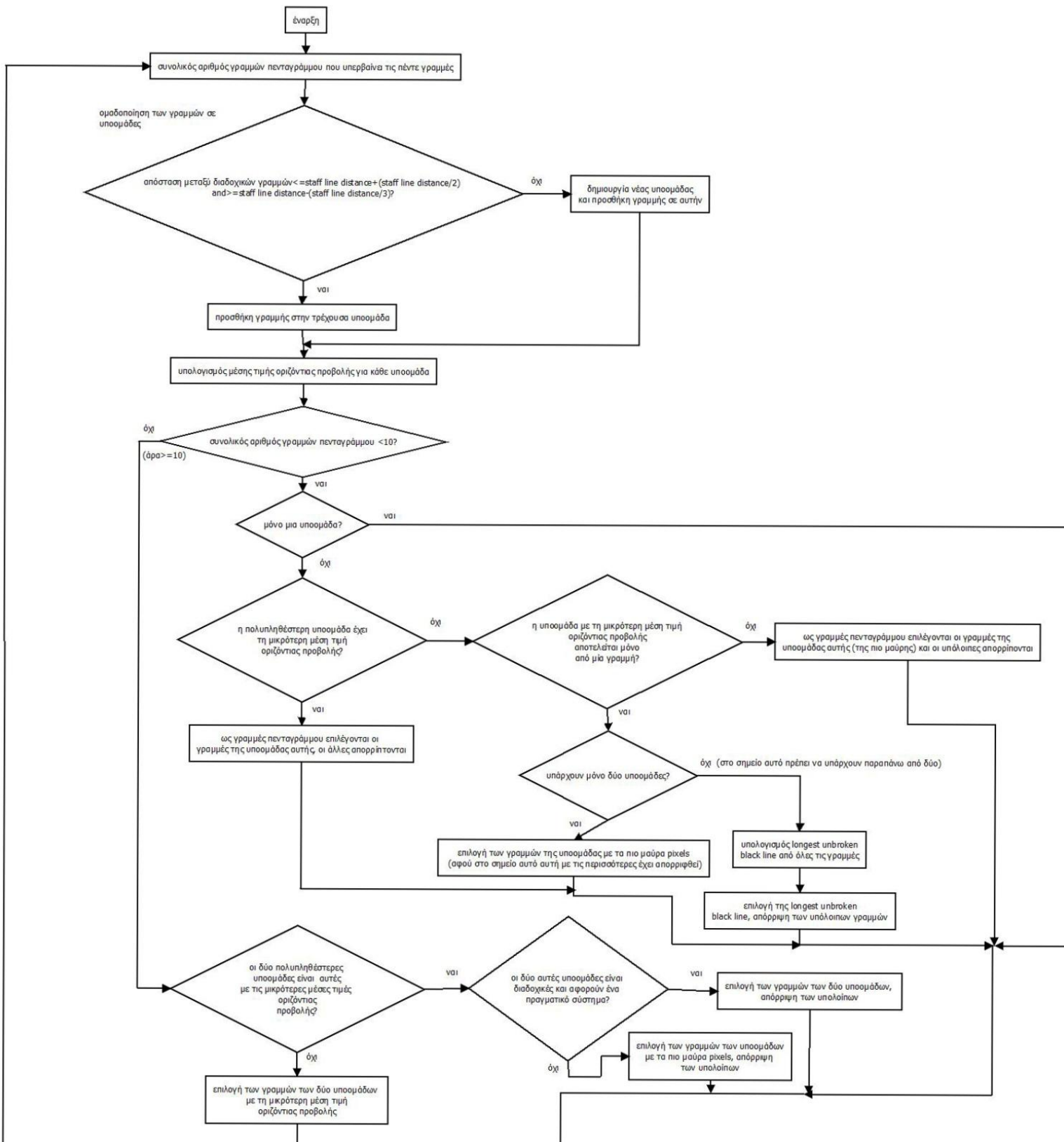


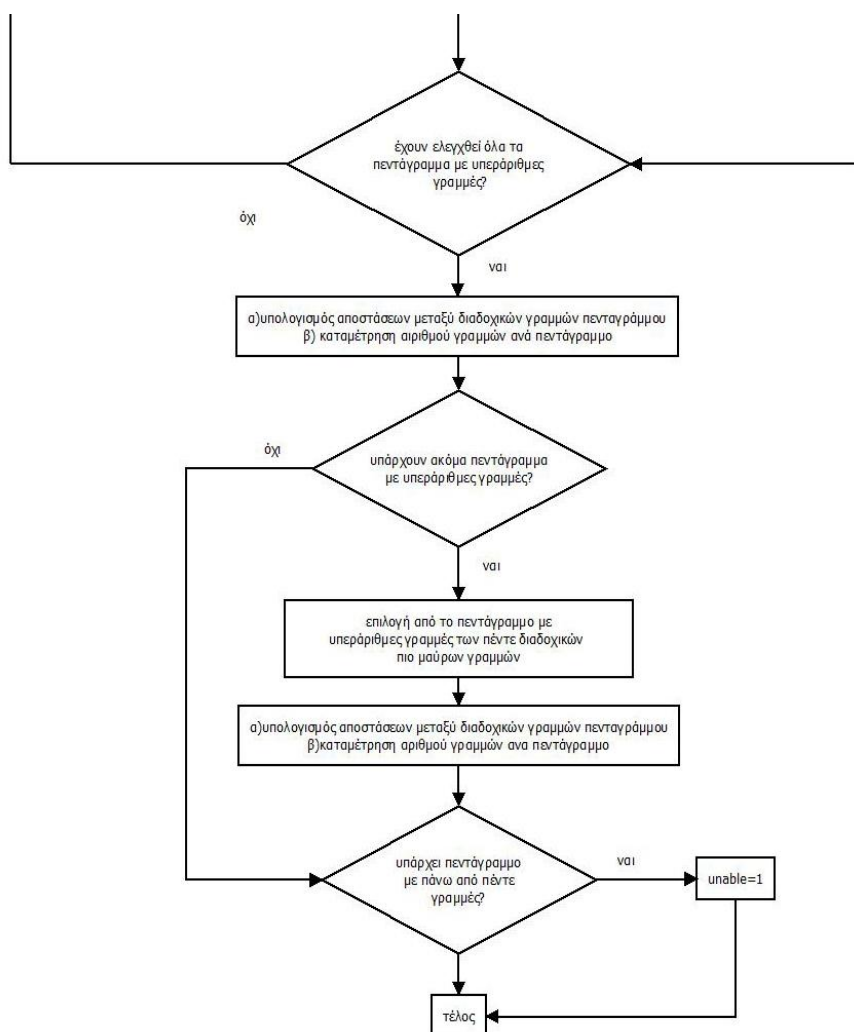
Εικόνα 3. 39: Ravel *Jeux d' eau*, απόσπασμα από σελίδα 12 [52], Παρατηρείται ότι υπάρχει περιττός αριθμός πεντάγραμμων. Στο κάτω «ευρύτερο» σύστημα περιλαμβάνονται τρία πεντάγραμμα. Στην περίπτωση αυτή, τα δύο πεντάγραμμα που συνδέονται με αγκύλη, πρέπει να παιχτούν με το αριστερό χέρι του εκτελεστή (εμφανίζεται η σχετική ένδειξη M.G.=main gauche=αριστερό χέρι στα γαλλικά), ενώ το πάνω πεντάγραμμα πρέπει να παιχτεί με το δεξί χέρι. Κατά τη διαδικασία εντοπισμού των πεντάγραμμων, η τοποθέτηση ενός μη υπαρκτού στην πραγματικότητα πενταγράμμου στο πάνω μέρος του εικονιζόμενου αποσπάσματος ευνοεί τον τερματισμό του σταδίου.

Τι θα συμβεί όμως στην περίπτωση που δεν εντοπιστεί έστω και λανθασμένα ένα τέτοιο μη πραγματικό πεντάγραμμα; Στην περίπτωση αυτή, ο αλγόριθμος, μη βρίσκοντας κάποια υποψήφια απόσταση για ύπαρξη πενταγράμμου ενδιαμέσως των ήδη εντοπισμένων ή έχοντας απορρίψει τις ενδιαμέσως αποστάσεις (και δεδομένου του περιττού αριθμού εντοπισμένων πενταγράμμων), εγκλωβίζεται στην αναζήτηση πενταγράμμων στο άνω και κάτω μέρος. Για το λόγο αυτό εισάγεται η πρόβλεψη για τερματισμό του σταδίου αν η ίδια άνω ή ίδια κάτω περιοχή προσπελαστεί πάνω από τρεις φορές, το στάδιο τερματίζεται. Όταν γίνεται αναφορά στην «ίδια» άνω ή κάτω περιοχή, αυτό σημαίνει ότι πρέπει η κατάσταση της άνω ή κάτω περιοχής να παραμένει αμετάβλητη, δηλαδή να μην εισάγονται νέα πεντάγραμμα για να θεωρεί ότι γίνεται ξανά προσπέλαση της ίδιας περιοχής. Η συνθήκη αυτή έχει παραληφθεί από το διάγραμμα ροής.

Υπάρχει μία ακόμα περίπτωση κατά την οποία είναι δυνατός ο τερματισμός του σταδίου ακόμα και αν έχει εντοπιστεί περιττός αριθμός πενταγράμμων: αν υποχρεωθεί από τη διαδικασία περιορισμού των υπεράριθμων πενταγράμμων που περιγράφεται ακολούθως.

Περιορισμός των υπεράριθμων πενταγράμμων





Εικόνα 3. 40: Διάγραμμα ροής διαδικασίας περιορισμού υπεράριθμων γραμμών, οι τρεις γραμμές που φαίνονται να διακόπτονται στην πρώτη σελίδα, στην πραγματικότητα συνεχίζονται στην επόμενη και είναι οι τρεις γραμμές που φαίνονται στο άνω μέρος της εικόνας.

Στην περίπτωση που σε κάποιο πεντάγραμμα διαπιστωθεί η ύπαρξη παραπάνω από πέντε γραμμών (χωρίς να περιλαμβάνονται οι διπλές ή τριπλές γραμμές οι οποίες αντιμετωπίζονται στο τέλος κάθε σταδίου και σε κάθε επανάληψη του τρέχοντος σταδίου) τότε ο αλγόριθμος μεταπηδά στη διαδικασία περιορισμού των υπεράριθμων γραμμών. Η διαδικασία αντιμετωπίζει διαδοχικά όλα τα υπεράριθμα σε γραμμές πεντάγραμμα, με τον ακόλουθο τρόπο:

Για το τρέχον πεντάγραμμα που εξετάζεται, γίνεται ομαδοποίηση των γραμμών που περιέχει σύμφωνα με τις μεταξύ τους αποστάσεις. Συγκεκριμένα, αν η απόσταση μεταξύ δύο διαδοχικών γραμμών του εξεταζόμενου (υπεράριθμου ως προς τις γραμμές) πενταγράμμου είναι μικρότερη από την τιμή $staff - line - distance + (staff - line - distance/2)$ και μεγαλύτερη από την τιμή $staff - line - distance - (staff - line - distance/3)$, τότε οι γραμμές αυτές τοποθετούνται στην ίδια υποομάδα του πενταγράμμου.

Να επισημανθεί εδώ, ότι με αυτό τον τρόπο οι γραμμές που βρίσκονται μεταξύ τους πιο κοντά από το επιτρεπόμενο όριο, δεν τοποθετούνται στην ίδια υποομάδα. Αφού όλες οι γραμμές του πενταγράμμου αντιστοιχιστούν σε μια υποομάδα, υπολογίζεται η μέση τιμή της οριζόντιας προβολής κάθε υποομάδας.

Στο σημείο αυτό, η εκτέλεση ακολουθεί δύο ξεχωριστά μονοπάτια, ένα για την περίπτωση που ο αριθμός των γραμμών στο πεντάγραμμο είναι μικρότερος από δέκα (10) και ένα στην για την περίπτωση που ο αριθμός αυτός είναι μεγαλύτερος ή ίσος του δέκα (10). Ο διαχωρισμός αυτός συμβαίνει γιατί όταν ο αριθμός των γραμμών είναι μεγαλύτερος του δέκα, είναι εξαιρετικά πιθανό να αφορά δύο διαδοχικά πεντάγραμμο, τα οποία βρίσκονται μεταξύ τους σε τόσο κοντινή απόσταση ώστε η διάκριση μεταξύ των πενταγράμμων να τα τοποθετεί στο ίδιο πεντάγραμμο και να θεωρείται έτσι ότι υπερβαίνουν το όριο των πέντε γραμμών.

Στην περίπτωση που οι εντοπισμένες γραμμές στο πεντάγραμμο είναι λιγότερες από δέκα και δεν έχουν κατηγοριοποιηθεί σε μία μόνο υποομάδα (δηλαδή οι αποστάσεις μεταξύ όλων των γραμμών θεωρήθηκαν αποδεκτές):

Αν η πολυπληθέστερη υποομάδα είναι αυτή με τη μικρότερη μέση τιμή οριζόντιας προβολής τότε διατηρούνται ως γραμμές πενταγράμμου μόνο οι γραμμές της υποομάδας αυτής.

Η συνθήκη αυτή χρησιμοποιείται γιατί όταν υπάρχουν παραπάνω από πέντε εντοπισμένες γραμμές ανά πεντάγραμμο, το πιθανότερο είναι οι πέντε εξ αυτών να ανταποκρίνονται σε πραγματικές γραμμές πενταγράμμου και οι υπόλοιπες να τοποθετήθηκαν λόγω γραμμών γύρω από το πεντάγραμμο ή ομαδοποιημένων νοτών, και επομένως έχουν κάποια απόσταση από το πεντάγραμμο. Στο σημείο αυτό δεν εξετάζεται πόσες είναι οι γραμμές που τελικά διατηρούνται: αν εξακολουθούν να υπερβαίνουν τις πέντε, αυτό θα αντιμετωπιστεί στη συνέχεια της διαδικασίας.

Αν η πολυπληθέστερη ομάδα δεν είναι αυτή που έχει τη μικρότερη μέση τιμή, τότε ως γραμμές πενταγράμμου διατηρούνται εκείνες οι οποίες ανήκουν στην ομάδα με τη μικρότερη μέση τιμή της οριζόντιας προβολής, δηλαδή η ομάδα με τα περισσότερα μαύρα pixels, με μόνη προϋπόθεση η ομάδα αυτή να μην αποτελείται μόνο από μία γραμμή. Αυτός ο περιορισμός εισάγεται γιατί αν η τιμή της οριζόντιας προβολής προέρχεται μόνο από μία γραμμή η σύγκριση με τις άλλες υποομάδες δεν είναι δίκαιη, ενώ συχνά η πιο μαύρη γραμμή μπορεί να αντιστοιχεί σε ομαδοποιημένες νότες και να μην ανήκει στο πεντάγραμμο.

Στην περίπτωση που η ομάδα με τα πιο μαύρα pixels αποτελείται από μόνο μια γραμμή, έχουν δηλαδή απορριφθεί οι δύο προηγούμενες περιπτώσεις, αλλά το εξεταζόμενο πεντάγραμμο είναι χωρισμένο μόνο σε δύο υποομάδες, τότε υποχρεωτικά επιλέγονται ως γραμμές πενταγράμμου εκείνες που ανήκουν στη μη μοναδιαία υποομάδα, ανεξαρτήτως του αριθμού τους. Σημειώνεται ότι στο τέλος ολόκληρης της διαδικασίας, υπάρχει πρόβλεψη στην περίπτωση που ο αριθμός των γραμμών πενταγράμμου εξακολουθεί να υπερβαίνει τις πέντε.

Αν ακόμα δεν έχει επιλεγεί κάποια υποομάδα, και επιπλέον όλες οι γραμμές έχουν τοποθετηθεί σε παραπάνω από δύο υποομάδες (δεδομένου ότι η περίπτωση των δύο υποομάδων ακριβώς εξετάστηκε ήδη), ο αλγόριθμος επιστρατεύει τη χρήση μιας τεχνικής που μετράει το μέγιστο αριθμό συνεχόμενων μαύρων pixels (longest unbroken black line) για κάθε γραμμή του εξεταζόμενου πενταγράμμου και επιλέγει ως μοναδική γραμμή του πενταγράμμου αυτή που εμφανίζει το μεγαλύτερο άθροισμα συνεχόμενων pixels.

Στο σημείο αυτό, και αντίθετα από τις προηγούμενες προσεγγίσεις όπου θεωρούνταν προτιμότερο να εντοπιστούν περισσότερες γραμμές προκειμένου να μην αγνοηθούν πραγματικές γραμμές, εδώ θεωρείται προτιμότερο να εντοπιστεί έστω μία αλλά πραγματική γραμμή πενταγράμμου. Οι γραμμές που λείπουν θα συμπληρωθούν σε νέα, σωστή βάση κατά την αναδρομή που εκτελείται στο 3^ο στάδιο από στάδιο εντοπισμού πενταγράμμων, όταν τα πεντάγραμμο δεν είναι συμπληρωμένα. Υπενθυμίζεται ότι στη διαδικασία περιορισμού γραμμών ο αλγόριθμος μπορεί να βρεθεί μόνο μεταπηδώντας από το στάδιο εντοπισμού πενταγράμμων.

Αν τίποτα από τα παραπάνω δεν οδήγησε σε επιλογή γραμμών, τότε η επιλογή θα γίνει στη συνέχεια της διαδικασίας ανεξαρτήτων του πλήθους των γραμμών του εξεταζόμενου πενταγράμμου. Το ίδιο ισχύει και για την περίπτωση που όλες οι γραμμές τοποθετήθηκαν εξ αρχής σε μία μόνο υποομάδα.

Στην περίπτωση που οι εντοπισμένες γραμμές στο πεντάγραμμο είναι περισσότερες από δέκα:

Όπως ήδη επισημάνθηκε, σε αυτή την περίπτωση είναι πολύ πιθανό δύο διαδοχικά πεντάγραμμο να βρίσκονται σε απόσταση τόσο μικρή ή να παρεμβάλλονται επιπλέον γραμμές ανάμεσα σε δύο διαδοχικά πεντάγραμμο ώστε θεωρηθούν ως ένα πεντάγραμμο. Τότε, ο αλγόριθμος αναζητά τις δύο μεγαλύτερες (σε πλήθος γραμμών) υποομάδες που έχουν δημιουργηθεί. Αν αυτές οι υποομάδες είναι και αυτές που έχουν τις πιο μικρές μέσες τιμές οριζόντιας προβολής και διαπιστωθεί από τη μεταξύ τους απόσταση ότι στην πραγματικότητα αποτελούν δύο διαφορετικά πεντάγραμμο πιθανώς του ίδιου συστήματος (για να βρίσκονται τόσο κοντά μεταξύ τους), τότε διατηρούνται και οι δύο υποομάδες ως γραμμές πενταγράμμου ενώ οι υπόλοιπες αγνοούνται. Η απόσταση μεταξύ των δύο υποομάδων επανεξετάστηκε με χαμηλότερο κριτήριο ($3 * staff - line - distance$) σε σχέση με αυτό που χρησιμοποιείται στα προηγούμενα στάδια του αλγορίθμου. Αν διαπιστώθηκε ότι πράγματι οι υποομάδες ανήκουν σε δύο διαφορετικά πεντάγραμμο τότε τίθεται μια σημαία που θα εμποδίσει τον περαιτέρω περιορισμό των γραμμών του αυτού του πενταγράμμου στη συνέχεια της διαδικασίας.

Αν οι δύο πολυπληθέστερες υποομάδες δεν είναι αυτές με τα περισσότερο μαύρα pixels, τότε επιλέγονται ως γραμμές πενταγράμμου αυτές που ανήκουν στις δύο ομάδες με τα περισσότερο μαύρα pixels.

Στη συνέχεια της διαδικασίας και ανεξάρτητα από τον αρχικό αριθμό των γραμμών του εξεταζόμενου πενταγράμμου (αν υπερέβαινε ή όχι το δέκα), ακολουθείται η παρακάτω πρακτική, η οποία περιορίζει τον αριθμό των γραμμών σε πεντάγραμμο στα οποία είτε δεν έγινε περιορισμός των γραμμών είτε ο περιορισμός που έγινε δε μείωσε τον αριθμό των γραμμών σε αριθμό μικρότερο ή ίσο του πέντε. Όπως γίνεται φανερό, αν ο αριθμός των γραμμών του εξεταζόμενου πενταγράμμου είναι μικρότερος του πέντε, αγνοείται από την πρακτική που περιγράφεται ακολούθως:

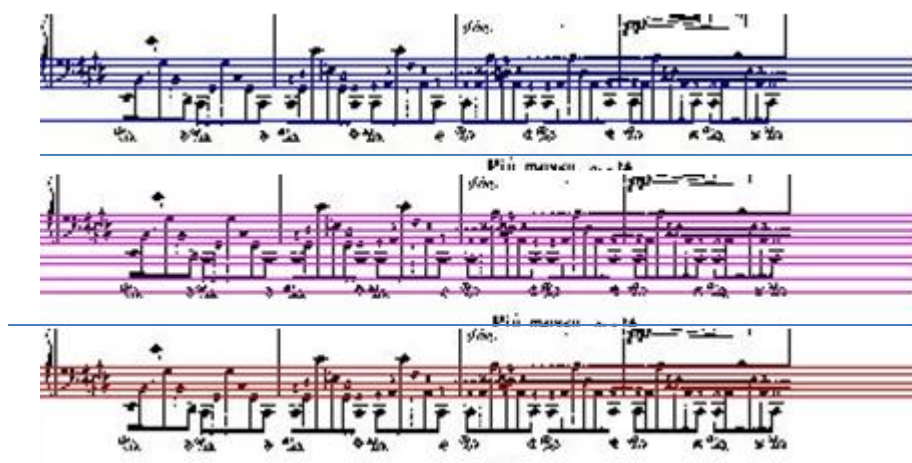
Η πρακτική αυτή, επιλέγει, από το σύνολο των νέων γραμμών του πενταγράμμου (ή των αρχικών αν δεν έγινε καμία αλλαγή) τις πέντε διαδοχικές

γραμμές πενταγράμμου με τη μικρότερη μέση τιμή οριζόντιας προβολής. Δηλαδή, για όλους τους συνδυασμούς διαδοχικών πεντάδων των γραμμών του πενταγράμμου, επιλέγονται εκείνες οι πέντε που βρίσκονται η μία μετά την άλλη σε αποδεκτές αποστάσεις και εμφανίζουν τα περισσότερα μαύρα ρίξελς σε σχέση με τις άλλες διαδοχικές πεντάδες που μπορούν να σχηματιστούν από τις υπάρχουσες γραμμές του πενταγράμμου.

Στην εξαιρετική περίπτωση που μετά από όλη τη διαδικασία, παραμένουν πεντάγραμμα με περισσότερες από πέντε γραμμές, τότε τίθεται μια σημαία που σταματά την επαναληπτική εκτέλεση του εντοπισμού πενταγράμμων (και αυτή είναι η περίπτωση στην οποία όπως αναφέρθηκε πριν την παρουσίαση της διαδικασίας περιορισμού υπεράριθμων γραμμών, αυτή μπορεί να σταματήσει την εκτέλεση του σταδίου).

Η διαδικασία του περιορισμού των πενταγράμμων μπορεί να έχει εξαιρετικά θετικές επιπτώσεις στο σωστό εντοπισμό των γραμμών του πενταγράμμου, όπως φαίνεται στις εικόνες που ακολουθούν (βλ. Εικόνες 3.41- 3.43) και σε κάποιες που προηγήθηκαν (βλ. Εικόνες 3.37- 3.38) σε συνδυασμό και με την υπόλοιπη διαδικασία του αλγορίθμου.

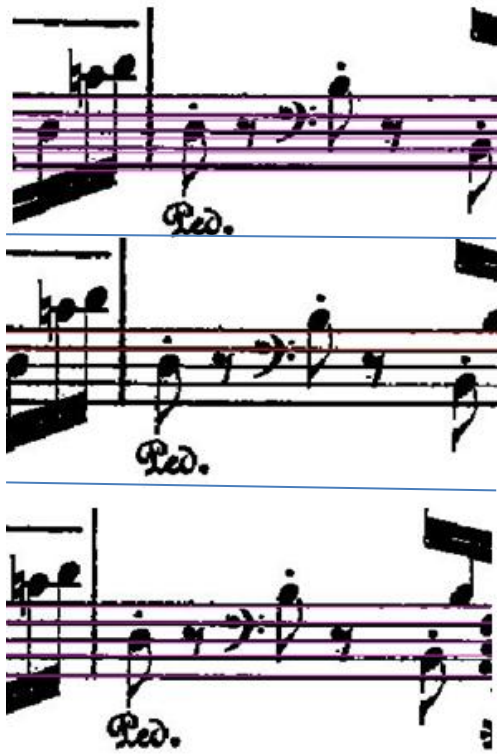
Ειδικά στην περίπτωση όπου οι εντοπισμένες γραμμές ενός πενταγράμμου δεν αντιστοιχούν σε πραγματικό πεντάγραμμα, αλλά σε μια περιοχή κάτω ή πάνω από ένα πεντάγραμμα όπου έχει θεωρηθεί από τον αλγόριθμο ότι υπάρχει αρκετός χώρος για ένα ακόμα πεντάγραμμα, οι νέες αυτές γραμμές βρίσκονται κοντά στο πραγματικό πεντάγραμμα. Έτσι, είναι πιθανό να ομαδοποιηθούν ως ενιαίο πεντάγραμμα και η διαδικασία περιορισμού υπεράριθμων να κρατήσει μόνο το πραγματικό πεντάγραμμα, κίνηση που θα αποτρέψει τον εντοπισμό ενός μη υπαρκτού πενταγράμμου (βλ. Εικόνα 3.44).



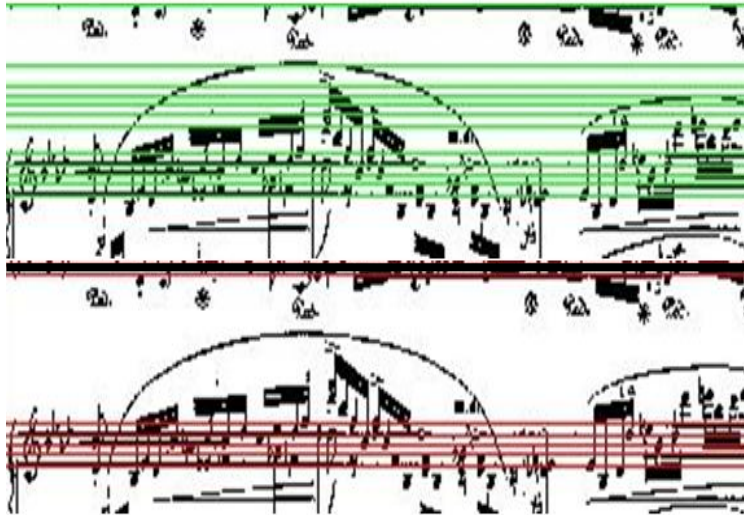
Εικόνα 3. 41: Από πάνω προς τα κάτω παρουσιάζονται τρία διαδοχικά στιγμιότυπα του ίδιου πενταγράμμου, καθώς προχωράει η εκτέλεση του αλγορίθμου. Παρατηρείται ότι έχει εντοπιστεί στην αρχική εκτίμηση (πάνω) μια επιπλέον γραμμή πενταγράμμου λόγω των ομαδοποιημένων ογδόων. Στο επόμενο βήμα (κέντρο) θεωρείται ότι αυτή η επιπλέον γραμμή ανήκει σε ξεχωριστό πεντάγραμμα και συμπληρώνεται. Η συμπλήρωση αυτή όμως έχει σαν αποτέλεσμα το πεντάγραμμα να αντιμετωπιστεί σαν ενιαίο υπεράριθμο σε γραμμές πεντάγραμμα, οπότε βγαίνοντας από τη διαδικασία περιορισμού των γραμμών του πενταγράμμου (κάτω), ο εντοπισμός των γραμμών είναι σωστός. Chopin nocturnes, σελίδα 2, [45].



Εικόνα 3. 42: Από πάνω προς τα κάτω εμφανίζονται τέσσερα διαδοχικά στιγμιότυπα των βημάτων του αλγορίθμου. Αρχική εκτίμηση με εντοπισμό τριών από τις πέντε πραγματικών γραμμών του πενταγράμμου και μια μη πραγματική λόγω ομαδοποιημένων ογδών (πάνω), Συμπλήρωση των γραμμών θεωρώντας ότι πρόκειται για ελλιπές πεντάγραμμο (δεύτερη από πάνω), αποτέλεσμα από τον περιορισμό των υπεράριθμων γραμμών (τρίτη από πάνω), συμπλήρωση των σωστών γραμμών του πενταγράμμου με αναδρομή στα στάδια συμπλήρωσης (κάτω), Chopin nocturnes, σελίδα 5, [45].



Εικόνα 3. 43: Τρία διαδοχικά στιγμιότυπα του ίδιου πενταγράμμου: Υπεράριθμες γραμμές σε ένα πεντάγραμμο (πάνω), περιορισμός των γραμμών σε δύο πραγματικές γραμμές μετά τη διαδικασία περιορισμού (κέντρο), συμπλήρωση σωστών γραμμών πενταγράμμου με αναδρομή στο στάδιο συμπλήρωσης (κάτω). Chopin Concerto no2, σελίδα 11 [49].



Εικόνα 3. 44: Δύο διαδοχικά στιγμιότυπα του ίδιου πενταγράμμου. Πάνω: Οι γραμμές στο επάνω μέρος που δεν αντιστοιχούν σε πραγματικό πεντάγραμμο, τοποθετήθηκαν επειδή ο αλγόριθμος θεώρησε αρκετή την απόσταση από το προηγούμενο πεντάγραμμο. Κάτω: Μετά τη διαδικασία περιορισμού των γραμμών, καθώς το μη πραγματικό με το πραγματικό πεντάγραμμο θεωρήθηκε ότι ανήκουν στην ίδια ομάδα, παραμένουν μόνο οι σωστές γραμμές. Chopin Concerto no2, σελίδα 11 [49].

3.2.5 Στάδιο 5^ο: Αναθεώρηση κατώφλιου

Το στάδιο αυτό εκτελείται μόνο στην περίπτωση που το αρχικό κατώφλι που τέθηκε για τον εντοπισμό των γραμμών πενταγράμμου και περιγράφηκε στο 2^ο στάδιο ($0,5 * \text{μέση τιμή οριζόντιας προβολής ολόκληρης της εικόνας}$) αποδείχθηκε πολύ αυστηρό για τα δεδομένα της εικόνας εισόδου και είτε εντοπίστηκε μηδενικός αριθμός γραμμών πενταγράμμων είτε ο αριθμός που εντοπίστηκε κρίθηκε ακατάλληλος για την ορθή εκτίμηση της staff-line-distance. Καθώς η διαδικασία που ακολουθεί είναι εξαιρετικά απλή, το στάδιο δε συνοδεύεται από διάγραμμα ροής. Η διαδικασία είναι η ακόλουθη:

Στην περίπτωση που το κατώφλι αρχικού εντοπισμού γραμμών κριθεί ακατάλληλο, το στάδιο αυτό υποχρεώνει τον αλγόριθμο να επιστρέψει στο 2^ο στάδιο και να χρησιμοποιήσει ένα λιγότερο αυστηρό κατώφλι εντοπισμού γραμμών (όπως φαίνεται στο διάγραμμα ροής του 2^{ου} σταδίου, βλ. Εικόνα 3.7): Επιβάλλει την αύξηση τιμής με την οποία πολλαπλασιάζεται η μέση τιμή της οριζόντιας προβολής, κατά 0.1 ($0,6 * \text{μέση τιμή οριζόντιας προβολής κατά την πρώτη αναδρομή}$). Στην περίπτωση που και το νέο κατώφλι κριθεί ακατάλληλο τότε επιβάλλεται επιστροφή στο 2^ο στάδιο και αυξάνεται και πάλι το κατώφλι κατά 0.1 ($0.7 * \text{μέση τιμή οριζόντιας προβολής}$). Το κατώφλι εξακολουθεί να αυξάνεται κατά 0.1 μέχρι να κριθεί ικανοποιητικό, δηλαδή μέχρι να εντοπιστεί μη μηδενικός αριθμός γραμμών πενταγράμμου και παράλληλα η εκτίμηση της staff-line-distance θεωρηθεί σωστή.

Ο τρόπος με τον οποίο αποφασίζεται αν η εκτίμηση της staff-line-distance είναι ορθή είναι:

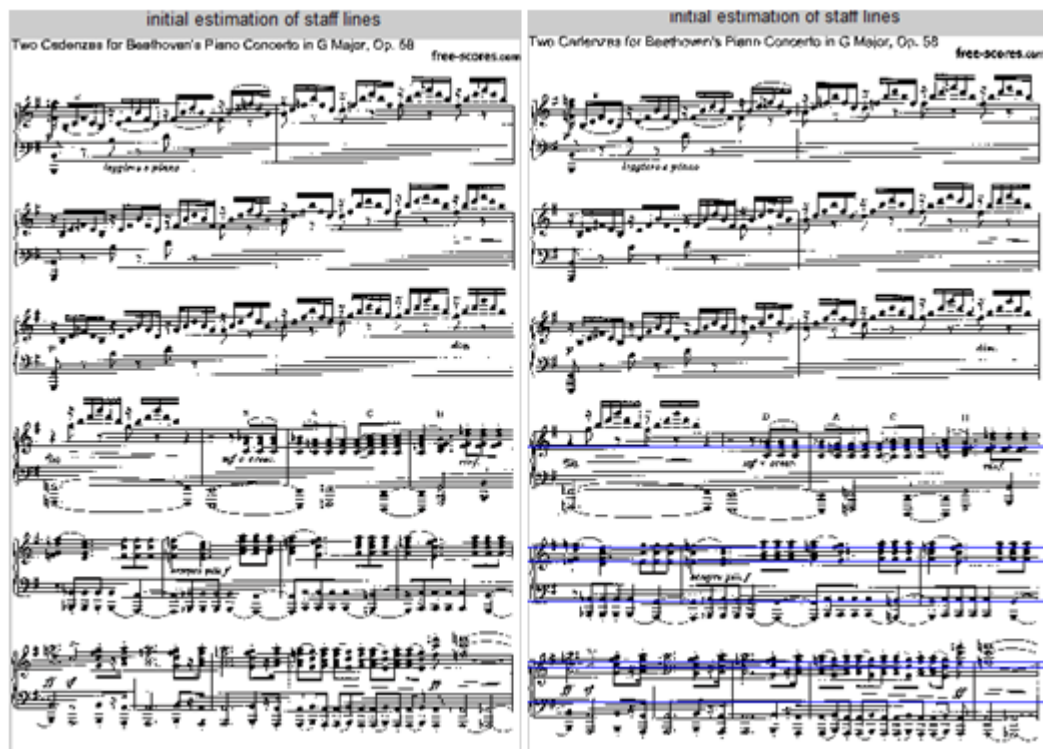
Στην περίπτωση που δεν έχουν εντοπιστεί καθόλου γραμμές από τον αλγόριθμο είναι προφανές ότι η *staff-line-distance* είναι μηδενική και ο αλγόριθμος προσπερνά όλα τα ενδιάμεσα στάδια (συμπλήρωσης γραμμών, εντοπισμού πενταγράμμων κτλ) για να φτάσει στο τρέχον που θα του επιτρέψει την αναδρομή με αναθεωρημένο κατώφλι.

Αν όμως υπάρχει μη μηδενικός αριθμός εντοπισμένων γραμμών, ο αλγόριθμος πρέπει να αποφασίσει αν η εκτίμηση της *staff-line-distance* είναι ορθή ή αν ο αριθμός των εντοπισμένων γραμμών αλλά και οι θέσεις τους είναι τέτοιες ώστε η εκτίμηση είναι λανθασμένη και πρέπει να αναθεωρηθεί το κατώφλι.

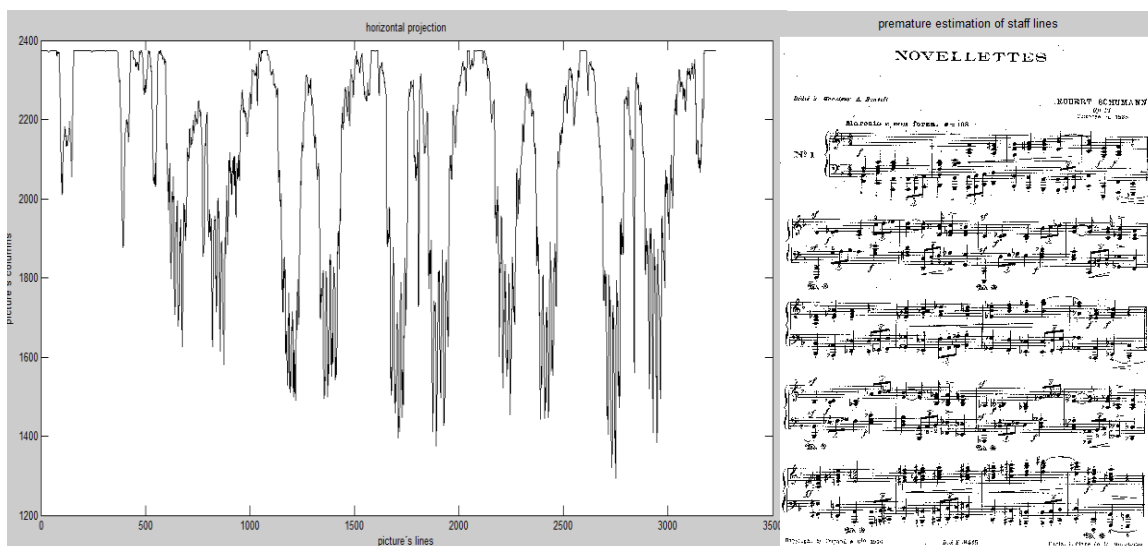
Αν η τιμή της υπολογισμένης *staff-line-distance* υπερβαίνει σε μέγεθος το ένα τεσσαρακοστό ($1/40$) του συνολικού ύψους της τρέχουσας σελίδας του μουσικού κειμένου, τότε η εκτίμηση θεωρείται λανθασμένη και ο αλγόριθμος αναγκάζεται να αναθεωρήσει το κατώφλι που χρησιμοποιεί για τον εντοπισμό των γραμμών.

Η επιλογή της τιμής αυτής βασίζεται στον ακόλουθο συλλογισμό: Στην αμέσως χειρότερη περίπτωση από αυτήν του μηδενικού εντοπισμού γραμμών και του εντοπισμού μοναδικής γραμμής, όπου η *staff-line-distance* είναι μηδενική (καθώς δεν υπάρχουν αποστάσεις μεταξύ των γραμμών προκειμένου να προκύψει μια εκτίμηση) βρίσκεται η περίπτωση του εντοπισμού δύο γραμμών. Σε όποια απόσταση και αν βρίσκονται οι γραμμές αυτές (εκτός από την περίπτωση να είναι διαδοχικές γραμμές πενταγράμμου που είναι αδύνατο να υποθεθεί ότι θα ισχύει σε κάθε περίπτωση) η εκτιμώμενη *staff-line-distance* θα είναι αρκετά μεγάλη (αφού θα είναι η απόσταση μεταξύ των δύο εντοπισμένων γραμμών) και κατά πάσα πιθανότητα μεγαλύτερη από το μισό του ύψους της συνολικής εικόνας. Αντίστοιχα, πχ. για πέντε εντοπισμένες γραμμές, στη χειρότερη περίπτωση όπου αυτές βρίσκονται σε διαφορετικά συστήματα ή πεντάγραμμα μεταξύ τους, η απόσταση που υπολογίζεται ως *staff-line-distance* θα είναι περίπου ίση με το $1/5$ του συνολικού ύψους της σελίδας. Ακολουθώντας αυτή τη λογική και σε συνδυασμό με μετρήσεις που έγιναν στο σύνολο των υπό εξέταση παρτιτούρων, αποφασίστηκε η χρήση της τιμής $1/40$ ως κατωφλίου.

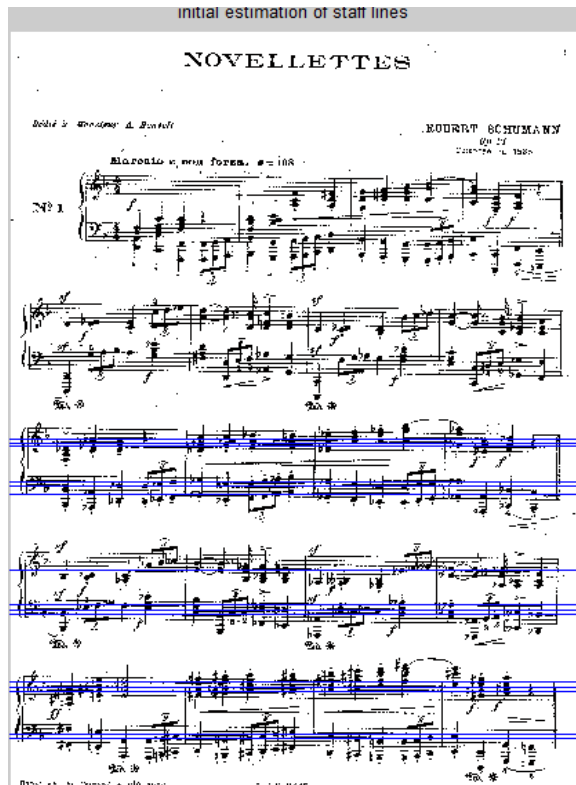
Να επισημανθεί ότι η αναθεώρηση του κατωφλίου συνήθως είναι απαραίτητη σε εικόνες μουσικών κειμένων με μεγάλη κλίση, όπου οι γραμμές δεν είναι κάθετες και επομένως το άθροισμα των μαύρων pixels των γραμμών πενταγράμμου διασπείρεται μεταξύ των οριζόντιων γραμμών που καταλαμβάνει η γραμμή αυτή λόγω κλίσης. Το ίδιο ισχύει και για εικόνες με κατακερματισμένες γραμμές ή όχι έντονα μαύρες γραμμές, όπου οι γραμμές πενταγράμμου δε θα είναι τόσο έντονα μαύρες σε σχέση με τις υπόλοιπες ώστε να καταφέρουν να περάσουν το αρχικό κατώφλι. Ακολουθούν ορισμένα παραδείγματα απεικόνισης των αποτελεσμάτων της αλλαγής του κατωφλίου (βλ. Εικόνες 3.45-3.47):



Εικόνα 3. 45: Επίδειξη σταδίου αναθεώρησης κατωφλίου, για διαφορετικά στιγμιότυπα του ίδιου μουσικού κειμένου Από αριστερά: Η πρώτη εκτίμηση γραμμών πενταγράμμου δεν επιστρέφει κανένα αποτέλεσμα λόγω της κλίσης της εικόνας (0.5°). Η δεύτερη εκτίμηση βρίσκει μικρό αριθμό γραμμών, ο οποίος είναι αρκετός για το σωστό υπολογισμό της staff-line-distance καθώς αν δεν υπάρχει απόσταση μεταξύ των γραμμών που να επαναλαμβάνεται ώστε να επιλεγεί ως staff-line-distance, επιλέγεται η πιο μικρή απόσταση, η οποία είναι πράγματι η staff-line-distance, λόγω του εντοπισμού δύο διαδοχικών γραμμών. Brahms Cadenzas, σελίδα 2, [12]



Εικόνα 3. 46: Συνοπτική επίδειξη αναθεώρησης κατωφλίου: Αριστερά φαίνεται το ιστόγραμμα της οριζόντιας προβολής της εικόνας που φαίνεται δεξιά. Παρατηρείται πως ήδη από το ιστόγραμμα (και σε σύγκριση με τα ιστογράμματα της Εικόνας 3.10) είναι φανερό ότι τα πεντάγραμμα έχουν κλίση (λόγω της διαταραχής, μη ομοιόμορφης μορφής του ιστογράμματος). Η κλίση αυτή είναι ίση με 0.8 μοίρες (0.8°). Αμέσως δεξιά, η πρώτη (αλλά και η δεύτερη) αναθεώρηση του κατωφλίου επιστρέφουν μηδενικό αριθμό εντοπισμένων γραμμών. Schumann Novelletes, σελίδα 1, [10]



Εικόνα 3. 47: Το αποτέλεσμα της τρίτης αναθεώρησης του κατώφλιου δίνει τελικά μια αρχική εκτίμηση γραμμών τέτοια που να επιτρέπει το σωστό υπολογισμό της staff-line-distance. Novellettes, σελίδα 1, [10].

Η κλίση

Στο σημείο αυτό, θα ήταν χρήσιμο να παρουσιαστεί η επίδραση της κλίσης στο αποτέλεσμα του εντοπισμού των γραμμών του πενταγράμμου. Η κλίση στις εικόνες των μουσικών κειμένων δημιουργείται κατά τη σάρωση των μουσικών κειμένων από την πρωτότυπη μορφή τους (χαρτί) σε ηλεκτρονική μορφή (εικόνα). Παρουσιάζονται ορισμένα παραδείγματα κεκλιμένων πενταγράμμων μαζί με τις εντοπισμένες γραμμές στα πεντάγραμμα αυτά (Εικόνες 3.47, 3.48):



Εικόνα 3. 48: Επίδειξη επίδρασης κλίσης στην τοποθέτηση των γραμμών: Αριστερά, φαίνεται η αρχή του τέταρτου συστήματος της Beethoven Sonata, σελίδα 9, [46]. Η τοποθέτηση των γραμμών είναι τέτοια που θα μπορούσε κανείς να πει ότι ο αλγόριθμος έχει «χάσει» την πρώτη γραμμή τόσο από το άνω όσο και από των κάτω πεντάγραμμο. Κοιτώντας όμως το τέλος του ίδιου συστήματος (δεξιά) παρατηρείται ότι όλες οι γραμμές είναι τοποθετημένες στις σωστές τους θέσεις. Η διαφορά αυτή μεταξύ των τέλεια οριζόντιων εντοπισμένων γραμμών οφείλεται στο ότι οι πραγματικές γραμμές πενταγράμμου δεν είναι οριζόντιες, αλλά εμφανίζουν κλίση, ίση με 0.5° . Παρά την κλίση όμως, ο αλγόριθμος επιτυγχάνει τη σωστή αναγνώριση των γραμμών.

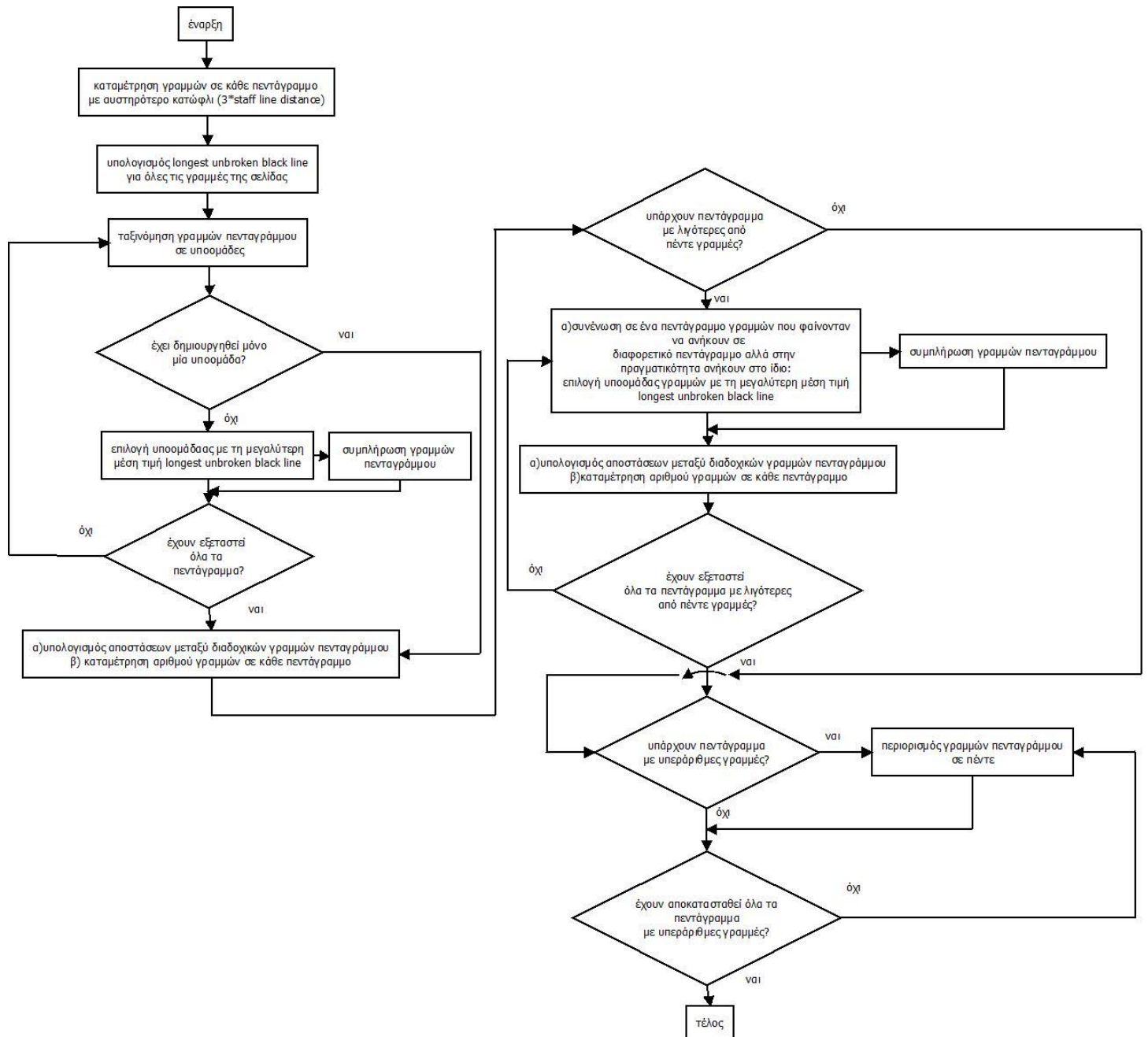


Εικόνα 3. 49: Παράδειγμα αντίστοιχο με το προηγούμενο (Εικόνας 3.48). Πάνω φαίνεται η αρχή του τελευταίου πενταγράμμου από Brahms Cadenzas [12], σελίδα 2. Παρατηρείται πόσο φαίνεται να απέχουν οι εντοπισμένες γραμμές από τις πραγματικές. Κάτω, στο τέλος του ίδιου πενταγράμμου, οι εντοπισμένες γραμμές του πενταγράμμου ταυτίζονται ακριβώς με τις πραγματικές. Η κλίση της εικόνας στο σημείο αυτό μετρήθηκε ίση με 0.5° .

Ολοκληρώνοντας το παραπάνω σχόλιο για την κλίση, αξίζει να παρατηρηθεί ότι η κλίση των πεντάγραμμων σε ένα μουσικό κείμενο, που έχει προκληθεί από τον τρόπο σάρωσης της εικόνας, δεν είναι ενιαία σε ολόκληρη τη σελίδα και τείνει να εμφανίζεται αυξημένη στα ακραία πεντάγραμμο της εικόνας.

Περισσότερα για την επίδραση της κλίσης στην ακρίβεια του αλγορίθμου θα παρουσιαστούν στο 4ο κεφάλαιο, κατά την παρουσίαση των αποτελεσμάτων του αλγορίθμου.

3.2.6 Στάδιο 6ο : Αναδιάταξη εντοπισμένων γραμμών



Εικόνα 3. 50: Διάγραμμα ροής σταδίου αναδιάταξης γραμμών, παραλείπεται η αναπαράσταση της εικόνας εξόδου

Στο σημείο αυτό έχουν εντοπιστεί όλες οι γραμμές πενταγράμμου της υπό επεξεργασία εικόνας. Σκοπός του σταδίου αυτού είναι η αναδιάταξη των εντοπισμένων γραμμών προκειμένου να επιτυγχάνεται καλύτερη ταύτιση των εντοπισμένων γραμμών με τις πραγματικές. Συγκεκριμένα, το στάδιο καλείται να ασχοληθεί με πεντάγραμμα στα οποία ο αριθμός των γραμμών που εντοπίστηκε ήταν ίσος με πέντε από την αρχή και δε χρειάστηκε να συμπληρωθούν ή να μειωθούν σε αριθμό. Αυτό σημαίνει ότι δεν έχει γίνει καμία διόρθωση των θέσεων των εντοπισμένων γραμμών – είναι πιθανό αυτές να βρίσκονται σε πολύ κοντινές ή πολύ μακρινές μεταξύ τους αποστάσεις και παρά το ότι είναι πέντε στον αριθμό, να μην ανταποκρίνονται στις πραγματικές θέσεις των γραμμών (βλ. Εικόνα 3.51):



Εικόνα 3. 51: Στο πρώτο και τρίτο από τα απεικονιζόμενα πεντάγραμμα, ενώ έχουν συμπληρωθεί πέντε γραμμές, αυτές δεν είναι συνεχόμενες, αλλά σχηματίζουν δύο διακριτές ομάδες. *Händel Largo*, σελίδα 3 [48]

Η διαδικασία που ακολουθείται είναι η εξής:

Δεδομένου ότι μέχρι το σημείο αυτό για την εκτίμηση και διόρθωση των γραμμών του πενταγράμμου έχει χρησιμοποιηθεί κυρίως η οριζόντια προβολή, στο στάδιο αυτό επιστρατεύεται η χρήση της τεχνικής της μακρύτερης συνεχόμενης μαύρης γραμμής, που χρησιμοποιήθηκε μερικώς και στη διαδικασία περιορισμού των υπεράριθμων γραμμών στα πεντάγραμμα.

Η τιμή της *longest unbroken black line* υπολογίζεται για όλες τις γραμμές της εικόνας. Στη συνέχεια, διατρέχονται όλα τα εντοπισμένα πεντάγραμμα της εικόνας και για κάθε πεντάγραμμα, γίνεται χωρισμός των γραμμών που το αποτελούν σε υποομάδες, με τον ίδιο τρόπο που περιγράφηκε στη διαδικασία περιορισμού υπεράριθμων γραμμών. Σημειώνεται ότι για το διαχωρισμό των πενταγράμμων χρησιμοποιείται το αυστηρότερο κατώφλι που χρησιμοποιήθηκε σε κάποιο κομμάτι της διαδικασίας περιορισμού υπεράριθμων γραμμών ($3 * staff - line - distance$).

Αν οι γραμμές του πενταγράμμου έχουν τοποθετηθεί σε περισσότερες από μία υποομάδες, το οποίο σημαίνει ότι οι αποστάσεις μεταξύ τους δεν είναι αυτές

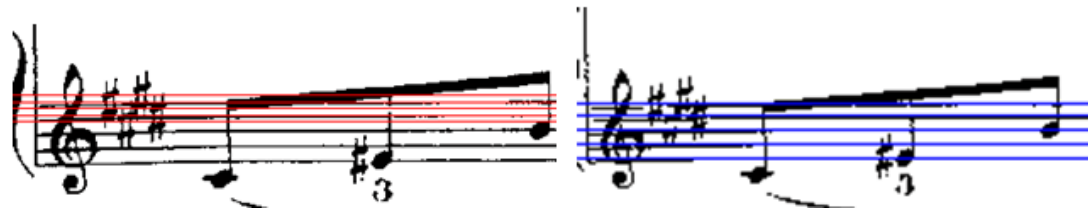
που θα έπρεπε, τότε ο αλγόριθμος θεωρεί ότι η μόνη υποομάδα που περιέχει πραγματικές γραμμές πενταγράμμου είναι αυτή με την υψηλότερη μέση τιμή longest unbroken black line. Οι υπόλοιπες γραμμές του πενταγράμμου απορρίπτονται και ο αλγόριθμος μεταπηδά σε μια διαδικασία συμπλήρωσης γραμμών που διαφέρει από αυτήν που παρουσιάστηκε στο 3^ο στάδιο στο ότι είναι προσαρμοσμένη για επεξεργασία ενός μόνο πενταγράμμου κάθε φορά. Στην ουσία πρόκειται για μια απλούστερη εκδοχή του 3^{ου} σταδίου, καθώς περιλαμβάνει μόνο τη δεύτερη φάση του. Αυτό συμβαίνει, γιατί δεδομένου ότι οι επιλεχθείσες γραμμές συνιστούν μια υποομάδα, είναι αδύνατο να λείπουν γραμμές στο ενδιάμεσό τους, επομένως οι μόνες γραμμές που πρέπει να συμπληρωθούν βρίσκονται είτε πάνω είτε κάτω είτε εκατέρωθεν της επιλεγμένης υποομάδας.



Εικόνα 3. 52: Σε αντιστοιχία με την Εικόνα 3.51, το αποτέλεσμα της επιδιόρθωσης των θέσεων των γραμμών του πρώτου και τρίτου πενταγράμμου. Και στα δύο πεντάγραμμα υπάρχει μια γραμμή που έχει διαφύγει από τον εντοπισμό, αλλά το αποτέλεσμα είναι καλύτερο από το προηγούμενο, οι αποστάσεις μεταξύ των γραμμών σωστές και αυτή η μικρή ολίσθηση των εντοπισμένων πενταγράμμων σε σχέση με τα πραγματικά δε θα επηρεάσει την περεταίρω εξέλιξη της αναγνώρισης. Händel Largo, σελίδα 3 [48]



Εικόνα 3. 53: Επίδειξη επιδιόρθωσης θέσης γραμμών πενταγράμμων, απόσπασμα πενταγράμμου από Mozart Sonata, σελίδα 11 [43]



Εικόνα 3. 54: Αναδιάταξη γραμμών πενταγράμμου: Αριστερά φαίνονται οι εντοπισμένες γραμμές του πενταγράμμου-αν και πέντε στον αριθμό οι θέσεις τους δεν ανταποκρίνονται στις πραγματικές γραμμές. Δεξιά φαίνεται το αποτέλεσμα μετά από την αναδιάταξη των γραμμών, Ravel Jeux d' eau, σελίδα 5 [52]

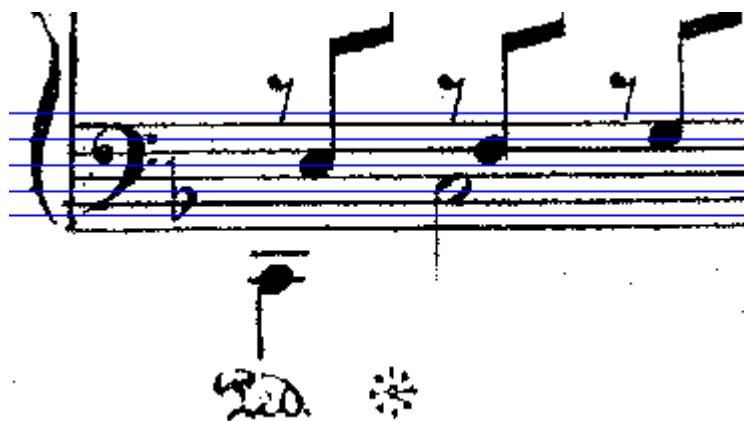
Στη συνέχεια, και αφού έχουν συμπληρωθεί οι γραμμές στις σωστές τους θέσεις, γίνεται επανεκτίμηση του αριθμού γραμμών ανά πεντάγραμμο (δηλαδή το πλήθος των γραμμών που περιέχεται σε κάθε πεντάγραμμο): χρησιμοποιείται και πάλι το αυστηρότερο κατώφλι. Το κατώφλι αυτό θα χρησιμοποιηθεί και σε όλα τα επόμενα στάδια που θα χρειαστεί να γίνει διαχωρισμός των πενταγράμμων. Συγκεκριμένα, αν δύο διαδοχικές γραμμές απέχουν μεταξύ τους λιγότερο από $3 * staff - line - distance$ τότε κατηγοριοποιούνται στο ίδιο πεντάγραμμο. Ο λόγος που χρησιμοποιείται χαμηλότερο κατώφλι είναι για να αντιμετωπίσει τις περιπτώσεις όπου οι εντοπισμένες γραμμές του πενταγράμμου εξακολουθούν να βρίσκονται σε διακριτές ομάδες αλλά στάθηκε αδύνατη η επιδιόρθωση τους κατά τη διάρκεια της προηγούμενης διαδικασίας.

Στη συνέχεια εξετάζονται τα πεντάγραμμα ως προς τον αριθμό των γραμμών τους, καθώς κατά την προηγούμενη διαδικασία ελέγχθηκαν οι αποστάσεις μεταξύ των γραμμών των πενταγράμμων. Στην περίπτωση που στο σημείο αυτό βρεθούν μη συμπληρωμένα πεντάγραμμα, είναι διαδοχικά, και εμφανίζονται να έχουν λιγότερες από πέντε γραμμές, αλλά το άθροισμα των γραμμών που τα αποτελούν είναι ίσο με το πέντε τότε αποφασίζεται ότι οι γραμμές που τα αποτελούν ανήκουν στο ίδιο πραγματικό πεντάγραμμο και επιλέγονται ως γραμμές του ενιαίου πλέον πενταγράμμου οι γραμμές που εμφανίζουν μεγαλύτερη μέση τιμή *longest unbroken black line*. Στη συνέχεια γίνεται μεταπήδηση στο στάδιο συμπλήρωσης γραμμών που αναφέρθηκε στην προηγούμενη διαδικασία.

Ο ρόλος της παραπάνω διαδικασίας δηλαδή, είναι να εντοπίζει πεντάγραμμα των οποίων οι διαδοχικές υποομάδες τυχαίνει να βρίσκονται μεταξύ τους σε απόσταση μεγαλύτερη της $3 * staff - line - distance$ (βλ. Εικόνες 3.50, 3.51) και κατά την προηγούμενη διαδικασία δεν αντιμετωπίστηκαν ως ενιαίο πεντάγραμμο ώστε να διορθωθούν οι θέσεις τους. Το κομμάτι αυτό δηλαδή λειτουργεί προσθετικά και βοηθητικά προς την πρώτη διαδικασία του σταδίου αυτού, με σκοπό να αντιμετωπιστούν όλες οι περιπτώσεις της λανθασμένης τοποθέτησης γραμμών.



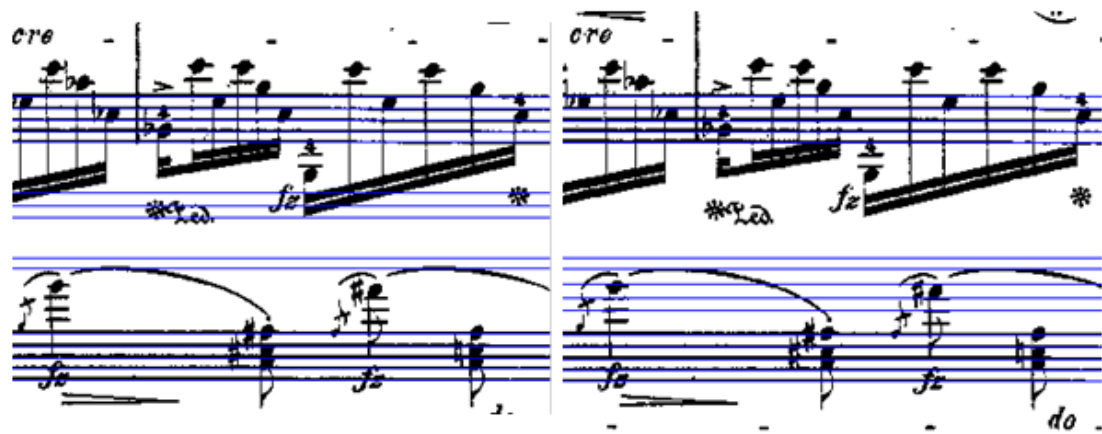
Εικόνα 3. 55: Το πρόβλημα της μη κατάταξης των γραμμών του ίδιου πενταγράμμου στο ίδιο πεντάγραμμο, παρά το αυστηρότερο κατώφλι. Αριστερά φαίνεται το αποτέλεσμα μετά την εκτίμηση γραμμών και δεξιά το πρόβλημα που παραμένει, μετά την πρώτη διαδικασία του σταδίου αυτού, εξαιτίας της απόστασης μεταξύ των ομάδων γραμμών που είναι μεγαλύτερη από το κατώφλι που χρησιμοποιείται ($3 \cdot \text{staff-line-distance}$). Schumann Novelletes, σελίδα 3 [10]



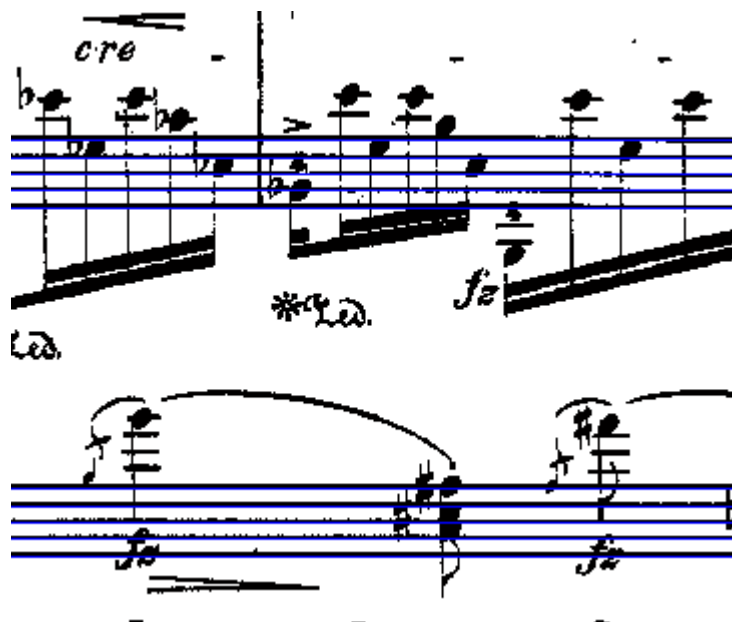
Εικόνα 3. 56: Η επίλυση του προβλήματος μέσω του ελέγχου αν διαδοχικά ελλειπή «πεντάγραμμο» παρουσιάζουν άθροισμα γραμμών ίσο με πέντε, Schumann Novelletes, σελίδα 3 [10].

Τέλος ακολουθεί μια διαδικασία εντοπισμού και περαιτέρω διόρθωσης των πενταγράμμων στα οποία περιλαμβάνονται, μετά τις παραπάνω διαδικασίες, πάνω από πέντε γραμμές: η απομάκρυνση των περιττών γραμμών γίνεται με μεταπήδηση σε μια απλούστερη εκδοχή της διαδικασίας περιορισμού των υπεράριθμων γραμμών στα πεντάγραμμο (που περιγράφηκε στο 4^ο στάδιο). Η απλή αυτή εκδοχή είναι σχεδιασμένη ώστε να αντιμετωπίζει ένα μόνο πεντάγραμμο κάθε φορά, ενώ για τον περιορισμό των γραμμών χρησιμοποιεί μόνο τη μέθοδο της διατήρησης των πέντε διαδοχικών γραμμών του πενταγράμμου που εμφανίζουν τη μικρότερη μέση τιμή οριζόντιας προβολής.

Η τελευταία αυτή διαδικασία μπορεί να λειτουργήσει θετικά και για τον περιορισμό τυχόν εντοπισμένων πενταγράμμων που δεν αντιστοιχούν σε πραγματικά πεντάγραμμο. Αν και αυτή η εργασία έχει ανατεθεί στο ακόλουθο στάδιο του αλγορίθμου, υπάρχει περίπτωση να συμβεί και σε αυτό, αν το επιπλέον πεντάγραμμο βρίσκεται τόσο κοντά σε κάποιο πραγματικό ώστε κατά τη διάρκεια του σταδίου αυτού συνενώνονται σε ένα πεντάγραμμο και τελικά απομακρύνεται με την αφαίρεση των υπεράριθμων γραμμών (βλ. Εικόνες 3.57, 3.58).

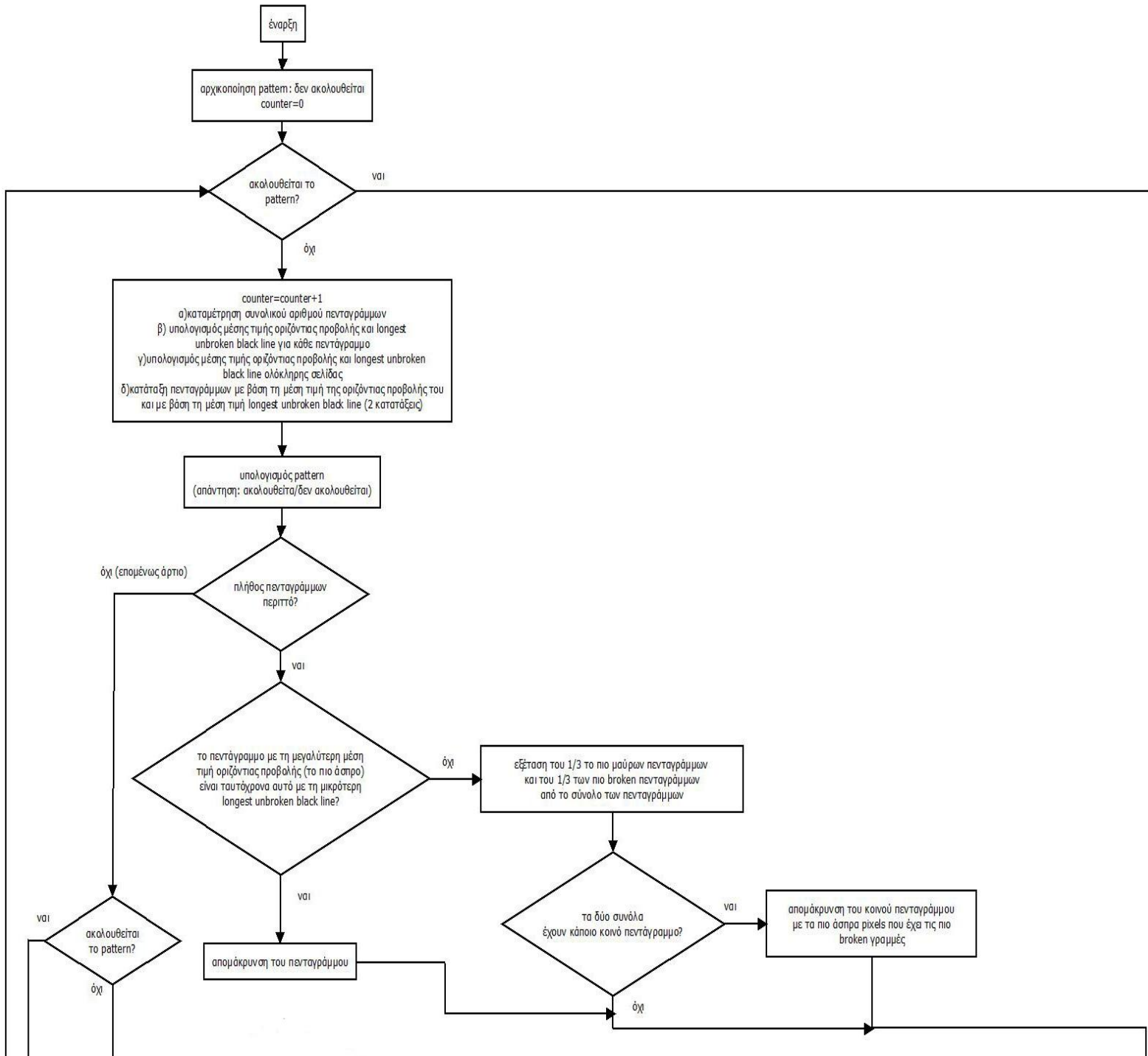


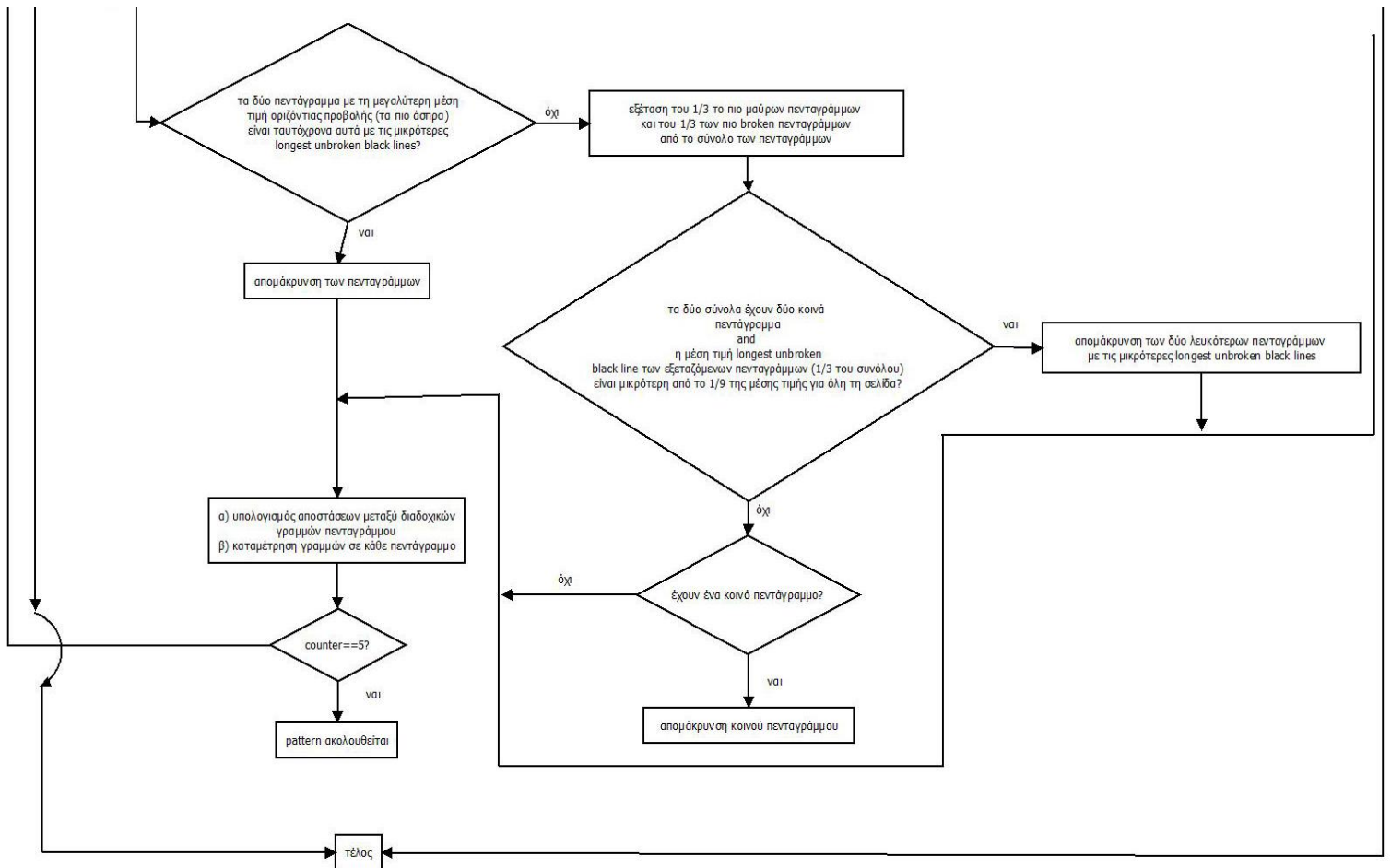
Εικόνα 3. 57: Δημιουργία μη πραγματικού πενταγράμμου. Αριστερά: Οι επιπλέον γραμμές (τρεις για το πάνω, δύο για το κάτω πεντάγραμμο) των πενταγράμμων, λόγω της μεγάλης μεταξύ τους απόστασης, δε μεταβάλλονται κατά την πρώτη διαδικασία του σταδίου. Δεξιά, κατά τη δεύτερη διαδικασία, επειδή εμπίπτουν στην περίπτωση των δύο διαδοχικών «πενταγράμμων» που σχηματίζουν στην πραγματικότητα ένα πεντάγραμμο, συνενώνονται., σχηματίζοντας ένα μη πραγματικό πεντάγραμμο Chopin nocturnes, σελίδα 9, [45].



Εικόνα 3. 58: Ο περιορισμός των υπεράριθμων γραμμών του πενταγράμμου απομακρύνει το μη πραγματικό πεντάγραμμο, η δημιουργία του οποίου απεικονίστηκε στην Εικόνα 3.56. Καθώς το μη πραγματικό πεντάγραμμο βρέθηκε σε τέτοια απόσταση από το κάτω πεντάγραμμο της εικόνας, ώστε να θεωρηθεί ότι σχηματίζουν ένα επιπλέον πεντάγραμμο, οι γραμμές απομακρύνθηκαν ως υπεράριθμες. Chopin nocturnes, σελίδα 9, [45].

3.2.7 Στάδιο 7ο: Απομάκρυνση μη πραγματικών πενταγράμμων





Εικόνα 3. 59: Διάγραμμα ροής σταδίου απομάκρυνσης μη πραγματικών πενταγράμμων. Οι γραμμές που φαίνονται να διακόπτονται συνεχίζουν στην επόμενη σελίδα του διαγράμματος.

Σκοπός του σταδίου αυτού είναι να απομακρύνει τυχόν εντοπισμένα πεντάγραμμα που δεν ανταποκρίνονται σε πραγματικά πεντάγραμμα. Τα πεντάγραμμα αυτά αναγνωρίστηκαν κατά τη διαδικασία εντοπισμού γραμμών εξαιτίας είτε μεγάλων αποστάσεων μεταξύ των συστημάτων (αποστάσεων που ξεπερνούσαν τη mid-system-distance σε τέτοιο βαθμό που θεωρήθηκε από τον αλγόριθμο ότι πρέπει να παρεμβάλλεται κάποιο μη εντοπισμένο πεντάγραμμα σε αυτή την απόσταση) είτε νοτών/κειμένου/σχολίων στις ενδιάμεσες περιοχές των πενταγράμμων που ήταν αρκετά έντονα ώστε να θεωρηθούν γραμμές πενταγράμμου.

Όπως έχει ήδη αναφερθεί κατά τη διάρκεια των προηγούμενων σταδίων, ο αλγόριθμος δε διαθέτει τρόπο να κατανοήσει αν τα πεντάγραμμα των οποίων τις γραμμές συμπληρώνει, περιορίζει ή αναδιατάσσει ανήκουν σε πραγματικές γραμμές πενταγράμμου-το υποθέτει, δεδομένης της αρχικής αναγνώρισής τους. Έτσι οι γραμμές που δεν αντιστοιχούν σε πραγματικά πεντάγραμμα αλλά στον τίτλο, σε ενδιάμεσα σχόλια, σε κείμενο στο κάτω μέρος του πενταγράμμου ή

ακόμα και σε νότες πάνω και κάτω από τα πεντάγραμμα ή σε συνδυασμό αυτών, έχουν σχηματίσει πλήρη πεντάγραμμα μέχρι να φτάσει ο αλγόριθμος στο παρόν στάδιο. Παρουσιάζονται κάποια παραδείγματα, προκειμένου να γίνει κατανοητή η έννοια του μη πραγματικού πενταγράμμου (βλ. Εικόνες 3.60-3.63):



Εικόνα 3. 60: Μη πραγματικό πεντάγραμμα, μετά την ολοκλήρωση του σταδίου αναδιάταξης των γραμμών, Beethoven Sonata, σελίδα 8, [46]



Εικόνα 3. 61: Δύο μη πραγματικά πεντάγραμμα, εξαιτίας του τίτλου και των σχολίων στο πάνω μέρος της σελίδας, Schumann Carnival, σελίδα 1, [9].



Εικόνα 3. 62: Ένα μη πραγματικό πεντάγραμμα ενδιάμεσως αναγνωρισμένων πενταγράμμων, εξαιτίας του κειμένου και της μεγάλης απόστασης μεταξύ των αναγνωρισμένων πενταγράμμων, Schumann Carnival, σελίδα 18, [9].



Εικόνα 3. 63: Δύο μη πραγματικά πεντάγραμμα (φαίνονται κυκλωμένα) σε τυχαίες θέσεις εξαιτίας νοτών και μεγάλων αποστάσεων μεταξύ των πενταγράμμων, Ravel, *Jeux d' eau*, σελίδα 5 [52]

Ο κυρίως τρόπος που διαθέτει ο αλγόριθμος για να αντιμετωπίσει τέτοιου είδους πεντάγραμμα είναι για μία ακόμα φορά (και κατ' αντιστοιχία με το 3^ο στάδιο) το πλήθος τους. Δηλαδή αν το πλήθος των εντοπισμένων πενταγράμμων είναι περιττό, αυτό αποτελεί ένδειξη ότι υπάρχει ένα εντοπισμένο πεντάγραμμα (ή περιττός αριθμός πενταγράμμων) που δεν αντιστοιχεί σε πραγματικό πεντάγραμμα. Από την άλλη πλευρά, αν το πλήθος των εντοπισμένων πενταγράμμων είναι άρτιο, παραμένει η πιθανότητα να υπάρχουν δύο πεντάγραμμα (ή άρτιος αριθμός πενταγράμμων) που να μην ανταποκρίνονται σε πραγματικά.

Ένα ακόμα εργαλείο που διαθέτει ο αλγόριθμος σχετίζεται άμεσα με τις *mid-staff-distance* και *mid-system distance*. Καθώς σε όλες τις παρτιτούρες ανεξαιρέτως τα συστήματα εμφανίζονται σε σειρά το ένα μετά το άλλο και καθώς η απόσταση από την αρχή της σελίδας ως το πρώτο πεντάγραμμα και από το τέλος της σελίδας ως το τελευταίο πεντάγραμμα δεν υπολογίζονται, οι αποστάσεις μεταξύ των πενταγράμμων πρέπει να ακολουθούν το πρότυπο: *midstaff-midsystem-midstaff-midsystem-midstaff*, κοκ., αρχίζοντας και τερματίζοντας απαραίτητα με *mid-staff distance*. Η δομή του προτύπου (*pattern*) είναι: *mid-staff *(midsystem-midstaff)*midsystem-midstaff*. Η ακολουθία μέσα στην παρένθεση έχει επαναληπτικό χαρακτήρα (βλ. Εικόνα 3.64).

PIANOFORTE.

fr

Maestoso. $\text{♩} = 125.$
Tutti.

CONCERTO.

Εικόνα 3. 64: Η ακολουθία των αποστάσεων *mid-staff-distance* και *mid-system-distance* σε μια παρτιτούρα (απεικονίζονται τα τρία πρώτα συστήματα). Με *m* χαρακτηρίζεται η *mid-staff-distance* και με *M* η *mid-system-distance*. Είναι φανερό ότι όλες οι παρτιτούρες ακολουθούν το παρουσιαζόμενο πρότυπο στη διαδοχή των αποστάσεων. Chopin Concerto no2, σελίδα 1, [49].

Δεδομένων των εργαλείων που είναι διαθέσιμα, ο αλγόριθμος ακολουθεί για τον εντοπισμό και την απομάκρυνση των μη πραγματικών πενταγράμμων, την παρακάτω διαδικασία:

Για κάθε εντοπισμένο πεντάγραμμο, υπολογίζει τη μέση τιμή της οριζόντιας προβολής του καθώς και τη μέση τιμή της *longest unbroken black line*. Στη συνέχεια, καταγράφει σε σειρά τις αποστάσεις μεταξύ των διαδοχικών πενταγράμμων. Οι αποστάσεις αυτές πρέπει να ακολουθούν το πρότυπο που προαναφέρθηκε. Ο αλγόριθμος ελέγχει την εφαρμογή του προτύπου, με δεδομένο ότι η *mid-staff distance* είναι πάντοτε μικρότερη της *mid-system distance* ξεκινώντας από την πρώτη καταγεγραμμένη απόσταση, η οποία πρέπει να είναι *mid-staff*. Συγκεκριμένα, εξετάζει τις καταγεγραμμένες αποστάσεις ανά δύο, ελέγχοντας αν η επόμενη σε σειρά απόσταση είναι μεγαλύτερη της τρέχουσας και αποφασίζει αν το pattern ακολουθείται ή όχι.

Στη συνέχεια, γίνεται έλεγχος του πλήθους των πενταγράμμων. Αν είναι περιττά σε πλήθος, τότε υπάρχουν οι ακόλουθες επιλογές:

Αν το πεντάγραμμο με τη χαμηλότερη μέση τιμή *longest unbroken black line* (με το μικρότερο συνεχόμενο μήκος γραμμής) ταυτίζεται με το πεντάγραμμο με τη μεγαλύτερη μέση τιμή οριζόντιας προβολής (με το μεγαλύτερο αριθμό άσπρων pixels), τότε το πεντάγραμμο αυτό θεωρείται μη πραγματικό και κατά συνέπεια απομακρύνεται.

Στη σπάνια περίπτωση που δεν επιτευχθεί η παραπάνω ταύτιση, από το σύνολο των πενταγράμμων επιλέγονται δύο ομάδες πενταγράμμων: μία ομάδα που αποτελείται από τα πιο κατακερματισμένα πεντάγραμμο (συνολικού πλήθους ίσου με το ένα τρίτο του συνολικού αριθμού των πενταγράμμων) και μια ομάδα αντίστοιχου πλήθους που αποτελείται από τα πεντάγραμμο με τα περισσότερα άσπρα pixels. Αν αυτές οι δύο ομάδες έχουν ένα τουλάχιστον κοινό πεντάγραμμο τότε απομακρύνεται το πεντάγραμμο με τη μεγαλύτερη μέση τιμή οριζόντιας προβολής που ταυτόχρονα έχει τη μικρότερη μέση τιμή longest unbroken black line.

Αν δε βρεθεί κοινό πεντάγραμμο τότε κανένα πεντάγραμμο δεν υφίσταται απομάκρυνση.

Αν τα πεντάγραμμο είναι άρτια σε πλήθος, τότε είτε όλες οι γραμμές έχουν εντοπιστεί σωστά, είτε υπάρχει άρτιο πλήθος μη πραγματικών πενταγράμμων:

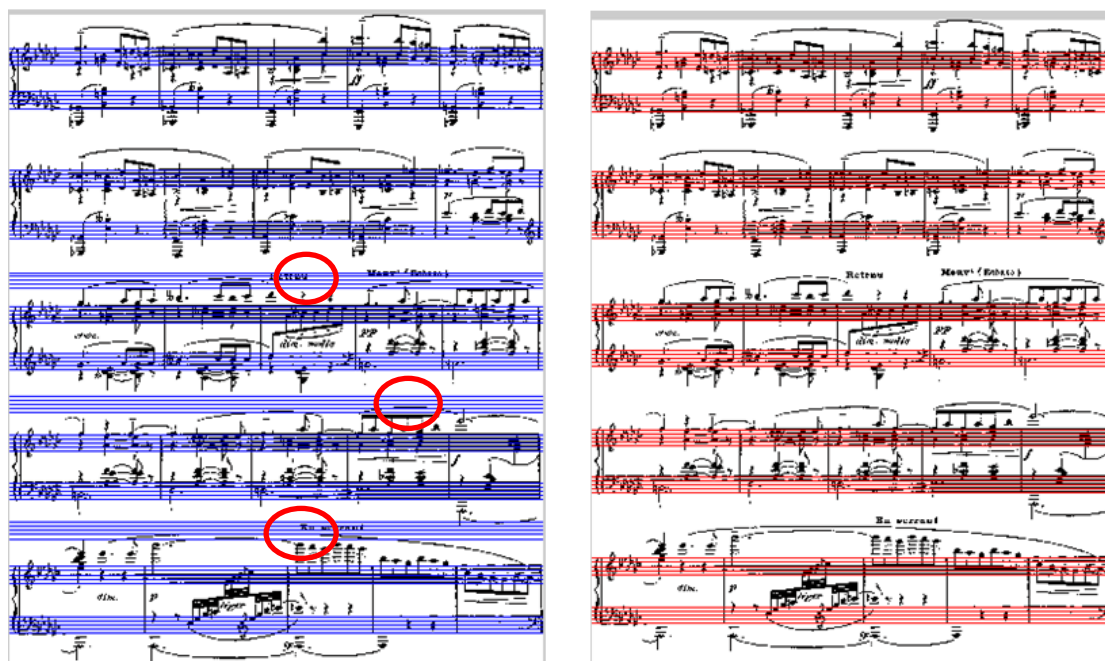
Αν το pattern ακολουθείται, τότε θεωρείται ότι όλα τα πεντάγραμμο είναι σωστά τοποθετημένα και το στάδιο τερματίζεται.

Διαφορετικά, η διαδικασία που ακολουθείται είναι παρόμοια με αυτή της περίπτωσης του περιττού πλήθους πενταγράμμων: αν τα δύο πεντάγραμμο με τα περισσότερα άσπρα pixels είναι ταυτόχρονα αυτά με τις πιο κατακερματισμένες γραμμές, τότε θεωρούνται μη πραγματικά και απομακρύνονται. Αν δεν υπάρξει τέτοια ταύτιση, τότε επιλέγονται δύο ομάδες πενταγράμμων και ακολουθείται η διαδικασία που ήδη αναφέρθηκε, με τη διαφορά ότι στην περίπτωση αυτή πρέπει η μέση τιμή της longest unbroken black line της ομάδας που επιλέγεται προς εξέταση να είναι μικρότερη από το $1/9$ της μέση τιμής για όλα τα πεντάγραμμο (βλ. διάγραμμα ροής, Εικόνα 3.59 για καλύτερη κατανόηση). Αυτός ο περιορισμός τοποθετείται προκειμένου να μην περιοριστούν πραγματικά πεντάγραμμο- μια τέτοια διαφορά από τη μέση τιμή όλων των πενταγράμμων περιορίζει στο ελάχιστο περιπτώσεις απομάκρυνσης πραγματικών πενταγράμμων. Η προσθήκη του περιορισμού ακολουθεί το σκεπτικό ότι είναι προτιμότερο να υπάρχουν κάποια μη πραγματικά πεντάγραμμο από το να απομακρυνθούν πραγματικά πεντάγραμμο.

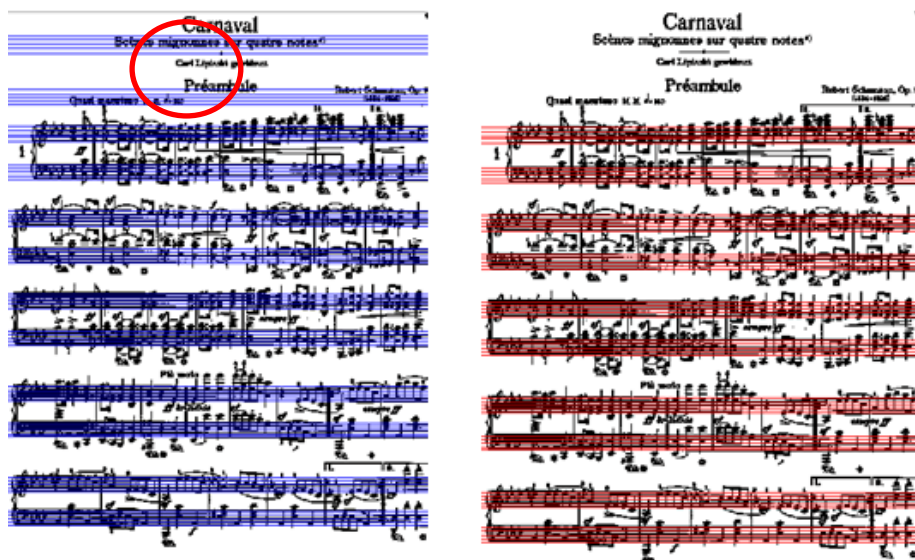
Η συνολική διαδικασία είναι επαναληπτική και τερματίζεται όταν ο αλγόριθμος θεωρήσει ότι όλα τα περιττά πεντάγραμμο έχουν απομακρυνθεί (δηλαδή ακολουθείται το pattern και υπάρχει άρτιο πλήθος εντοπισμένων πενταγράμμων).

Σε περίπτωση εγκλωβισμού του αλγορίθμου αν δεν ακολουθείται το pattern αλλά δεν ταυτόχρονα δεν περιορίζεται και κάποιο πεντάγραμμο (όπως είναι πιθανό να συμβεί όπως φάνηκε), δηλαδή τα δεδομένα παραμένουν ίδια σε κάθε επανάληψη, υπάρχει ένας μετρητής που αναγκάζει το στάδιο να τερματίσει μετά την πάροδο πέντε επαναλήψεων.

Παρακάτω εμφανίζονται παραδείγματα της αποτελεσματικότητας του σταδίου απομάκρυνση μη πραγματικών πενταγράμμων (βλ. Εικόνες 3.65-3.67). Οι εικόνες εμφανίζονται σε ζεύγη πριν και μετά την απομάκρυνση:



Εικόνα 3. 65: Παράδειγμα περιορισμού τριών μη πραγματικών πενταγράμμων- οι θέσεις τους φαίνονται κυκλωμένες στην αριστερή εικόνα. Δεξιά, φαίνεται το αποτέλεσμα μετά τον περιορισμό των μη πραγματικών πενταγράμμων. Debussy Valse, σελίδα 3, [44].

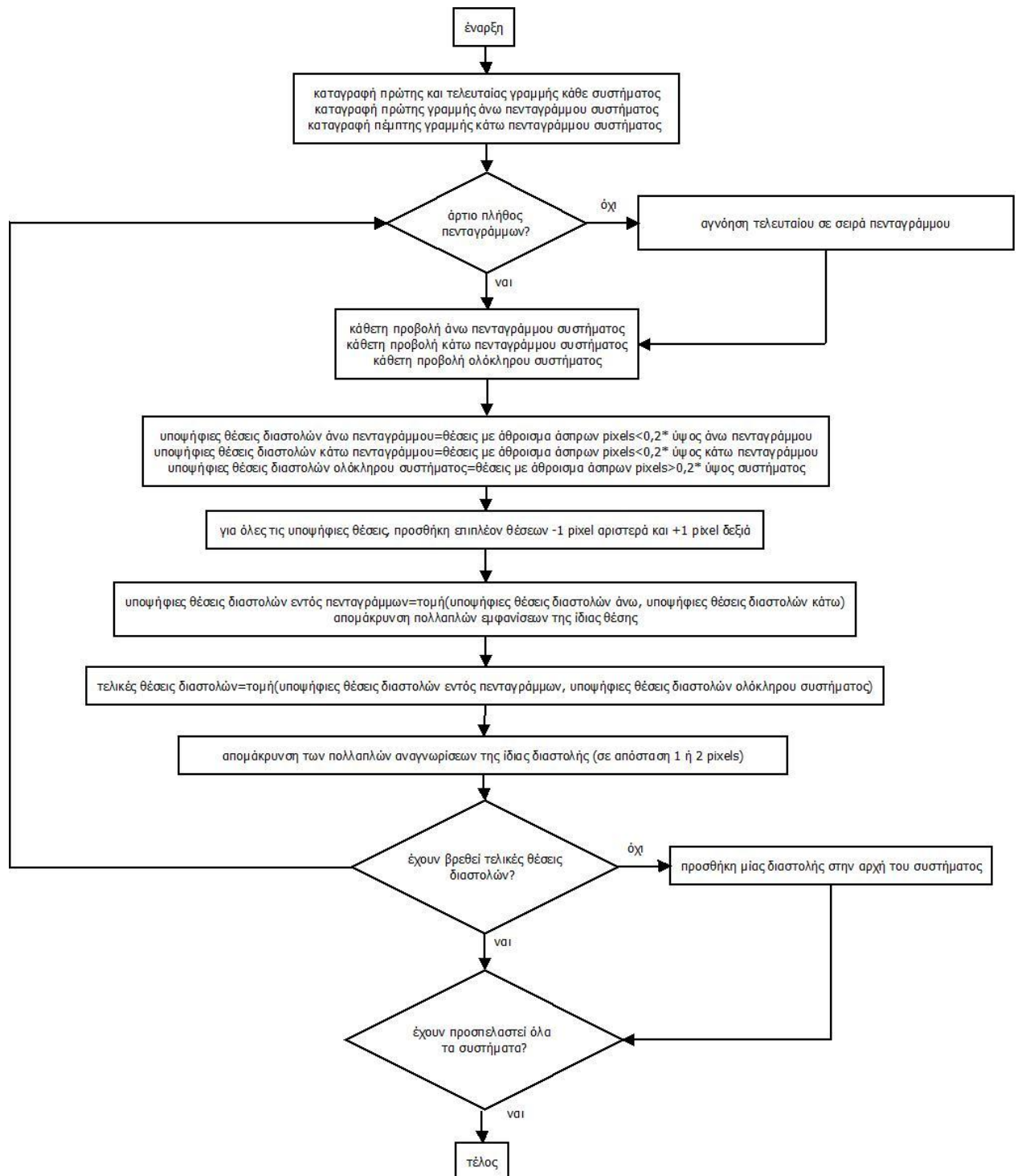


Εικόνα 3. 66: Παράδειγμα περιορισμού δύο πενταγράμμων από την άνω πλευρά της εικόνας- οι θέσεις τους φαίνονται κυκλωμένες στην αριστερή πλευρά, Schumann Carnival, σελίδα 1, [9].

Εικόνα 3. 67: Παράδειγμα περιορισμού δύο πενταγράμμων σε τυχαίες θέσεις- οι θέσεις τους φαίνονται κυκλωμένες στην αριστερή πλευρά, *Chopin Nocturnes*, σελίδα 8, [45].

3.2.8: Στάδιο 8^ο : Εντοπισμός διαστολών- Χωρισμός μέτρων

Σκοπός του σταδίου αυτού, με δεδομένη την τελική εκτίμηση των θέσεων των εντοπισμένων γραμμών του πενταγράμμου είναι να εντοπίσει τις θέσεις των διαστολών (και επομένως των μέτρων) μέσα στα συστήματα. Καθώς οι διαστολές είναι οι μοναδικές συνεχείς κάθετες γραμμές εντός των ορίων ενός συστήματος, ο αλγόριθμος χρησιμοποιεί για την εύρεση τους κάθετες προβολές.



Εικόνα 3. 68: Διάγραμμα ροής σταδίου εντοπισμού διαστολών.

Παρόλα αυτά, επειδή η χρήση απλής κάθετης προβολής μπορεί να είναι παραπλανητική στις περιπτώσεις που πολλές νότες πρέπει να παιχτούν ταυτόχρονα και παράλληλα και από τα δύο χέρια (οι γραμμές τους έχουν την ίδια εμφάνιση με τις διαστολές (βλ. Εικόνα 3.69), ο αλγόριθμος εφαρμόζει την ακόλουθη μέθοδο για τον εντοπισμό των διαστολών:

Στα όρια του κάθε συστήματος (τα πεντάγραμμο συνδυάζονται από την αρχή της εικόνας ανά δύο) δηλαδή από την πέμπτη γραμμή του άνω πενταγράμμου

ως την πρώτη γραμμή του κάτω πενταγράμμου και λαμβάνεται η τιμή της κάθετης προβολής, για όλο το πλάτος της εικόνας του μουσικού κειμένου.

Η κάθετη προβολή είναι αντίστοιχη της οριζόντιας με τη διαφορά ότι αντί για το άθροισμα των τιμών των pixels στις γραμμές, δίνει σαν αποτέλεσμα το άθροισμα των τιμών των pixels στις στήλες της εικόνας.

Επιπλέον, η τιμή της κάθετης προβολής λαμβάνεται μεμονωμένα εντός των ορίων του άνω πενταγράμμου καθώς και εντός των ορίων του κάτω πενταγράμμου. Συνολικά δηλαδή υπάρχουν τρεις μετρήσεις για την κάθετη προβολή. Και για τις τρεις περιπτώσεις ως υποψήφια θέσεις διαστολών λαμβάνονται οι θέσεις εκείνες όπου το άθροισμα των άσπρων pixels της κάθετης προβολής είναι μικρότερο από το 20% (ισοδύναμα, το άθροισμα των μαύρων pixels της κάθετης προβολής είναι μεγαλύτερο από το 80%) του συνολικού ύψους του πενταγράμμου ή του συστήματος, ανάλογα με τα όρια στα οποία λαμβάνεται κάθε φορά η κάθετη προβολή.



Εικόνα 3. 69: Δύο παραδείγματα γραμμών από νότες με ίδια εμφάνιση με τις διαστολές, καθώς καταλαμβάνουν όλο το ύψος των πενταγράμμων και μεγάλο μέρος του ύψους του συστήματος. Από αριστερά: Brahms Cadenzas, σελίδα 7, [12], Debussy Piece pour Piano, [53].

Στη συνέχεια, για κάθε υποψήφια θέση διαστολής προστίθενται ως πραγματικές θέσεις και οι δύο γειτονικές της θέσεις, αυτές δηλαδή που βρίσκονται ένα pixel δεξιά ένα pixel αριστερά της πραγματικής θέσης. Αυτό εξυπηρετεί τη διαδικασία τελικής επιλογής των πραγματικών θέσεων: Ένα πρώτο σύνολο τελικών υποψήφια θέσεων προκύπτει από την τομή των θέσεων των διαστολών που βρέθηκαν εντός μόνο του άνω και εντός μόνο του κάτω πενταγράμμου. Οι θέσεις αυξήθηκαν με τον τρόπο που περιγράφηκε, ακριβώς για να μη χάνονται πραγματικές θέσεις πενταγράμμου λόγω οριζόντιας μετατόπισης ενός ή δύο pixels από το πάνω στο κάτω πεντάγραμμο (πχ. λόγω κλίσης της παρτιτούρας).

Οι τελικές θέσεις προκύπτουν από την τομή του πρώτου συνόλου τελικών υποψήφια θέσεων (που προέκυψαν από το άνω και κάτω πεντάγραμμο) με το

σύνολο των θέσεων που έχουν προκύψει από την κάθετη προβολή σε όλο το ύψος του συστήματος. Με αυτό τον τρόπο οι θέσεις που αντιστοιχούν απλώς σε γραμμές από νότες που βρίσκονται στο ενδιάμεσο των δύο πενταγράμμων, ή μόνο στα πεντάγραμμο αλλά όχι στο ενδιάμεσο (βλ. Εικόνα 3.69) περιορίζονται. Φυσικά, οι θέσεις του τελικού συνόλου διαστολών που έχουν καταχωρηθεί δύο φορές για το ίδιο σύστημα, ή απέχουν μεταξύ τους ένα ή δύο pixels απομακρύνονται, προκειμένου να μην υπάρχει πολλαπλός εντοπισμός της ίδιας διαστολής.

Αν στην αρχή του σταδίου αυτού, έχει αναγνωριστεί περιττός αριθμός πενταγράμμων, τότε τα αναγνωρισμένα πεντάγραμμο τοποθετούνται σε ζεύγη διαδοχικά ξεκινώντας από την αρχή, όπως συμβαίνει σε όλες τις περιπτώσεις, αλλά το τελευταίο πεντάγραμμο (που δε διαθέτει ζεύγος) αγνοείται. Με αυτόν τον τρόπο εξασφαλίζεται μια εκτίμηση των γραμμών πενταγράμμου ακόμα και αν τα τελευταία έχουν περιττός πλήθος. Ανάλογα με τη θέση του «περιττού» πενταγράμμου (είναι πιθανό να μην υπάρχει μη πραγματικό πεντάγραμμο, αλλά να μην έχει αναγνωριστεί ένα πραγματικό πεντάγραμμο) κάθε φορά, η εκτίμηση των θέσεων των διαστολών εμφανίζει διακυμάνσεις ως προς την ακρίβειά της, καθώς αν το πεντάγραμμο αυτό βρίσκεται στο τέλος τότε δεν επηρεάζονται οι προηγούμενες διαστολές. Αν όμως το πεντάγραμμο (ή τα πεντάγραμμο) βρίσκονται στην αρχή τότε επηρεάζουν την εκτίμηση των πενταγράμμων που θεωρείται ότι σχηματίζουν ένα σύστημα.

Φυσικά το ίδιο μπορεί να συμβεί και όταν υπάρχει άρτιος αριθμός μη πραγματικών πενταγράμμων-αν αυτά βρίσκονται σε ζεύγη δε δημιουργείται κανένα πρόβλημα-αν όμως δε συμβαίνει αυτό, τότε εμφανίζεται και πάλι το προαναφερθέν πρόβλημα. Για το λόγο αυτό, γίνεται μεγάλη προσπάθεια στην υλοποίηση για την επίτευξη της αρτιότερης δυνατής αναγνώρισης των πενταγράμμων.

Τα αποτελέσματα του σταδίου αυτού, που είναι και τα τελικά αποτελέσματα του αλγορίθμου παρουσιάζονται στα ακόλουθα παραδείγματα (βλ. Εικόνες 3.70-3.71):

The image shows a page of musical notation for 'Deux Nocturnes' by Frédéric Chopin. The page is divided into two columns of music. The left column contains the first Nocturne, and the right column contains the second. The music is written for piano and includes various dynamics such as *pp*, *f*, and *ff*. The tempo is marked 'Andante'. The page number '16 790' is visible at the bottom left.

Εικόνα 3. 70: Τελικό αποτέλεσμα αλγορίθμου, απεικονίζονται και θέσεις των διαστολών. Από αριστερά: Beethoven Sonata, σελίδα 8, [46], Chopin Nocturnes, σελίδα 1, [45].

The image shows a page of musical notation for 'LARGO' by George Frideric Handel. The page is divided into two columns of music. The left column contains the first part of the score, and the right column contains the second part. The music is written for piano and includes various dynamics such as *mf* and *f*. The tempo is marked 'LARGO'. The page number '93 27' is visible at the bottom left.

Εικόνα 3. 71: Τελικό αποτέλεσμα αλγορίθμου, με εμφάνιση των θέσεων των διαστολών αλλά και των εντοπισμένων γραμμών. Στη δεξιά εικόνα, εμφανίζονται και δύο περιπτώσεις εντοπισμού μη πραγματικών διαστολών. Από αριστερά: Händel Largo, σελίδα 1, [48], Mozart Sonata, σελίδα 11, [43].

Κεφάλαιο 4°

Αποτελέσματα, Συμπεράσματα και Μελλοντική εξέλιξη

4.1 Αποτελέσματα και παρατηρήσεις

Για τον έλεγχο της λειτουργικότητας και της ακρίβειας του αλγορίθμου, χρησιμοποιήθηκε ένα σύνολο 17 μουσικών κομματιών αποτελούμενο από 175 διαφορετικές μεταξύ τους σελίδες.

Στο σύνολο των 17 μουσικών κομματιών:

- 13 κομμάτια είναι τύπου scanned, σύνολο 153 σελίδες
- 4 κομμάτια είναι τύπου born digital, σύνολο 22 σελίδες

Όπως θα γίνει αντιληπτό και στη συνέχεια, παρουσιάζεται ένα σημαντικό εύρος στο πλήθος των σελίδων του κάθε κομματιού, στον αριθμό πενταγράμμων ανά σελίδα καθώς και στον αριθμό διαστολών ανά σελίδα. Επιπλέον, κάθε διαφορετικό μουσικό κείμενο έχει διαφορετική εμφάνιση από τα υπόλοιπα, σε σχέση με την ποιότητα εκτύπωσης, την κλίση, το μέγεθος των μουσικών στοιχείων, κλπ.

Συνοπτικά, οι παρτιτούρες που χρησιμοποιήθηκαν, το πλήθος των σελίδων του κάθε μουσικού κειμένου, ο τύπος του, και οι μέσες διαστάσεις κάθε σελίδας, παρατίθενται στον παρακάτω πίνακα (Πίνακας 4.1):

α/α	Αναγνωριστικό μουσικού κειμένου	Αριθμός σελίδων	τύπος	Μέσο ύψος (pixels)	Μέσο πλάτος σελίδας (pixels)
1	Beethoven Sonata [46]	17	scanned	812	589
2	Chopin Nocturnes [45]	11	scanned	2965	2281
3	Brahms Cadenzas [12]	8	scanned	3265	2172
4	Debussy Piece [53]	1	scanned	3264	2404
5	Debussy Valse [44]	6	scanned	3121	2400
6	Händel Largo [48]	3	scanned	3138	2361
7	Mozart Sonata [43]	12	scanned	3083	2371
8	Ravel Jeux [52]	13	scanned	2853	2273
9	Schubert Waltzes [50]	4	scanned	2799	2147
10	Schumann Carnaval [9]	29	scanned	768	591
11	Chopin Concerto [45]	35	scanned	1917	1416
12	Schumann Novelletes [10]	7	scanned	3167	2360
13	Mendelssohn Retour [51]	7	scanned	3373	2368
14	Beethoven Elise [11]	5	born digital	2918	2284
15	Bizet Carmen [47]	7	born digital	2941	2317
16	Brendon Piano Solo [54]	8	born digital	3032	2331
17	Tchaikovsy SwanLake Scene [55]	2	born digital	3071	2209

Πίνακας 4. 1: Παρουσίαση χαρακτηριστικών συνόλου παρτιτούρων.

Σημειώνεται ότι οι διαστάσεις που χρησιμοποιήθηκαν για την εξαγωγή του μέσου μήκους και πλάτους κάθε σελίδας, αναφέρονται στις τιμές που προέκυψαν μετά την περικοπή των λευκών περιθωρίων στο 1^ο στάδιο του αλγορίθμου.

Αναλύοντας περισσότερο τα επιμέρους χαρακτηριστικά κάθε σελίδας, και προκειμένου να γίνει αντιληπτό το εύρος των διαφορετικών περιπτώσεων που εμφανίζονται στο εξεταζόμενο σύνολο μουσικών κειμένων, παρατίθενται οι μέσες τιμές του πλήθους γραμμών και του πλήθους διαστολών ανά σελίδα για κάθε μουσικό κείμενο, μαζί με τις αντίστοιχες τυπικές αποκλίσεις. Σημειώνεται ότι για τις μέσες τιμές έχει γίνει στρογγυλοποίηση του αποτελέσματος στις μονάδες (καθώς οι γραμμές και οι διαστολές είναι διακριτά στοιχεία):

α/α	Αναγνωριστικό μουσικού κειμένου	Αριθμός σελίδων	Μέση τιμή πλήθους γραμμών ανά σελίδα	Τυπική απόκλιση	Μέση τιμή πλήθους διαστολών ανά σελίδα	Τυπική απόκλιση
1	Beethoven Sonata [46]	17	55	4,9	34	7,73
2	Chopin Nocturnes [45]	11	50	0	22	5,27
3	Brahms Cadenzas [12]	8	54	4,84	26	7,23
4	Debussy Piece [53]	1	60	0	33	0
5	Debussy Valse [44]	6	48	3,73	29	1,97
6	Händel Largo [48]	3	47	4,71	35	2,94
7	Mozart Sonata [43]	12	62	7,99	39	13,41
8	Ravel Jeux [52]	13	48	3,73	11	1,86
9	Schubert Waltzes [50]	4	63	4,33	65	8,62
10	Schumann Carnaval [9]	29	58	4,27	48	8,94
11	Chopin Concerto [45]	35	56	4,83	33	10,97
12	Schumann Novellettes [10]	7	50	0	25	0,72
13	Mendelssohn Retour [51]	7	77	4,52	60	4,59
14	Beethoven Elise [11]	5	42	11,66	25	6,97
15	Bizet Carmen [47]	7	34	10,5	25	7,15
16	Brendon Piano Solo [54]	8	48	4,33	23	2,17
17	Tchaikovsky SwanLake Scene [55]	2	45	5	19	2,5

Πίνακας 4. 2: Στατιστικά συνόλου παρτιτούρων σχετικά με πλήθος γραμμών και διαστολών ανά σελίδα

Ο αριθμός των σελίδων εμφανίζεται στον πίνακα ώστε να παρέχει την προοπτική και το σωστό πλαίσιο μέσα στο οποίο πρέπει να γίνονται κατανοητές οι τιμές που παρατίθενται.

Η μέση τιμή του πλήθους των γραμμών ανά σελίδα προτιμήθηκε από τη μέση τιμή των πενταγράμμων ανά σελίδα, ώστε να παρέχει καλύτερο μέτρο σύγκρισης κατά την παρουσίαση της ακρίβειας εντοπισμού των γραμμών του αλγορίθμου: Καθώς οι γραμμές εντοπίζονται μία προς μία και είναι πιθανό σε ένα πεντάγραμμο να έχει γίνει σωστός εντοπισμός ορισμένων μόνο από τις γραμμές του, η καταμέτρηση των σωστά εντοπισμένων γραμμών δίνει πολύ πιο αντιπροσωπευτικό ποσοστό σε σχέση με την καταμέτρηση των σωστά εντοπισμένων πενταγράμμων.

Από τον Πίνακα 4.2, παρατηρείται ότι η μέση τιμή αριθμού γραμμών ανά σελίδα κυμαίνεται από 34 έως 77 γραμμές (υπενθυμίζεται ότι οι τιμές αυτές αφορούν εκτίμηση μέσου όρου και οι γραμμές σε μια παρτιτούρα πιάνου εμφανίζονται πάντα σε άρτια πολλαπλάσια του πέντε). Από τις τυπικές αποκλίσεις γίνεται επίσης σαφές ότι ακόμα και μέσα στο ίδιο μουσικό κείμενο είναι δυνατό να υπάρχει μεγάλη διακύμανση στον αριθμό γραμμών ανά σελίδα (πχ. στο Beethoven Elise [11] η τυπική απόκλιση είναι ίση με 11,66

αντικατοπτρίζοντας το γεγονός ότι από τις πέντε συνολικά σελίδες της παρτιτούρας οι τρεις έχουν από 50 γραμμές, μία έχει 40 γραμμές, ενώ η τελευταία σελίδα αποτελείται μόνο από 20 γραμμές). Το αντίστοιχο ισχύει και για τη μέση τιμή αριθμού διαστολών ανά σελίδα: Κυμαίνεται από 11 διαστολές (σε ένα μουσικό κείμενο με αραιά μεγάλα σε εύρος μέτρα) σε 60 διαστολές (σε ένα πυκνό κείμενο με στενά σε εύρος μέτρα).

Γίνεται κατανοητό ότι ο αριθμός των γραμμών (και άρα των πενταγράμμων και συστημάτων) ανά σελίδα επηρεάζει και τον αριθμό των διαστολών μέσα στη σελίδα, καθώς είναι αναμενόμενο σε μια σελίδα με πολλά συστήματα να υπάρχει αντίστοιχος αριθμός μέτρων. Παρόλα αυτά, η αύξηση του αριθμού των διαστολών συναρτήσει του αριθμού των συστημάτων δεν είναι απαραίτητα ανάλογη, καθώς δεν υπάρχει κάποιος κανόνας που να υπαγορεύει ενιαίο εύρος στα μέτρα ενός κομματιού- το εύρος εξαρτάται από το μέγεθος των αξιών και άρα το πλήθος των νοτών μέσα σε κάθε μέτρο προκειμένου να υπολογιστεί ο χώρος που καταλαμβάνει.

Τα αποτελέσματα που προέκυψαν ξεχωριστά για κάθε μουσικό κομμάτι, παρουσιάζονται στον ακόλουθο πίνακα (Πίνακας 4.3 - διατηρείται και πάλι το πλήθος των σελίδων ανά μουσικό κομμάτι και αναφέρεται ο τύπος της κάθε παρτιτούρας- η σειρά εμφάνισης των παρτιτούρων ταυτίζεται με τη σειρά εμφάνισης στους προηγούμενους πίνακες):

α/ α	Τύπος μουσικού κειμένου	Αριθμός σελίδων	κατώφλι black and white	Σύνολο γραμμών σε όλες τις σελίδες	Σύνολο διαστολών σε όλες τις σελίδες	Ποσοστό επιτυχίας-γραμμές	Τυπική απόκλιση	Ποσοστό επιτυχίας-διαστολές	Τυπική απόκλιση
1	scanned	17	0,95	940	586	99,79%	0,01	97,10%	0,07
2	scanned	11	0,95	550	242	99,09%	0,03	97,93%	0,04
3	scanned	8	0,95	430	205	96,74%	0,09	96,10%	0,08
4	scanned	1	0,95	60	33	100,00%	0,00	100,00%	0
5	scanned	6	0,95	290	176	100,00%	0,00	100,00%	0
6	scanned	3	0,95	140	105	62,86%	0,46	63,81%	0,47
7	scanned	12	0,95	740	472	95,14%	0,14	93,86%	0,24
8	scanned	13	0,95	620	148	83,87%	0,27	69,59%	0,35
9	scanned	4	0,95	250	258	100,00%	0,00	100,00%	0
10	scanned	29	0,95	1670	1384	98,74%	0,04	94,94%	0,17
11	scanned	35	0,95	1970	1150	100,00%	0,00	99,48%	0,02
12	scanned	7	0,75	350	172	80,00%	0,35	73,26%	0,36
13	scanned	7	0,75	540	423	96,30%	0,06	81,80%	0,35
14	digital	5	0,95	210	127	100,00%	0	100,00%	0,00
15	digital	7	0,95	234	146	97,50%	0,21	92,49%	0,45
16	digital	8	0,95	380	186	100,00%	0	100,00%	0,00
17	digital	2	0,95	90	37	100,00%	0	67,57%	0,00

Πίνακας 4. 3: Αποτελέσματα εντοπισμού γραμμών και πενταγράμμων ανά μουσικό κομμάτι

Το **συνολικά ποσοστά επιτυχίας** είναι:

Στην αναγνώριση των γραμμών:

96,26% με τυπική απόκλιση 0,15 για τις scanned παρτιτούρες

99,35% με τυπική απόκλιση 0,12 για τις born digital παρτιτούρες

Στην αναγνώριση των διαστολών:

93,63% με τυπική απόκλιση 0,22 για τις scanned παρτιτούρες

91,84% με τυπική απόκλιση 0,32 για τις born digital παρτιτούρες

Το χαμηλό ποσοστό αναγνώρισης γραμμών που παρατηρείται σε κάποιες παρτιτούρες (πχ. 6, 8, 12 στον παραπάνω πίνακα) οφείλεται σε μεμονωμένες παρτιτούρες μέσα στο σύνολο, που δεν έδωσαν καλά αποτελέσματα και όχι στο χαμηλό επίπεδο αναγνώρισης όλων των παρτιτούρων-αυτό αντικατοπτρίζεται και από τις τιμές της τυπικής απόκλισης. Όσον αφορά στην αναγνώριση των διαστολών, συχνά επηρεάζεται από τυχόν μη πραγματικά πεντάγραμμα που δεν περιορίστηκαν στο στάδιο απομάκρυνσης μη πραγματικών πενταγράμμων και οι θέσεις τους είναι τέτοιες ώστε τα ζεύγη πενταγράμμων που δημιουργούνται να μην ανταποκρίνονται σε πραγματικά συστήματα και επομένως να μη γίνεται ορθή αναγνώριση διαστολών στη σελίδα. Το ίδιο μπορεί να συμβεί και αν ένα πεντάγραμμα διαφύγει πλήρως της αναγνώρισης: αν το πεντάγραμμα αυτό δε βρίσκεται στο τέλος της σελίδας ώστε να μην επηρεάσει την ομαδοποίηση των υπόλοιπων πενταγράμμων σε συστήματα, τότε τα αποτελέσματα εξαρτώνται από το που βρίσκεται το πεντάγραμμα που απουσιάζει σε σχέση με τα υπόλοιπα, αν δηλαδή θα επιτρέψει το σωστό υπολογισμό των διαστολών στα υπόλοιπα συστήματα.

Για παράδειγμα στο 6 (από τον πίνακα 4.3) που αναφέρεται στο Händel Largo [48] και αποτελείται από τρεις συνολικά σελίδες, ενώ οι δύο από αυτές εμφάνισαν υψηλά ποσοστά αναγνώρισης γραμμών 100% και 96%, η τρίτη σελίδα είχε τόσο μεγάλη κλίση(1°, που μεταφράζεται σε ολίσθηση κατά ύψος ίσο με δύο γραμμές πενταγράμμου από την οριζόντια θέση) που αν και υπήρχαν εντοπισμένες γραμμές, ο αριθμός των αναγνωρισμένων θεωρήθηκε μηδενικός εξαιτίας του μεγάλου πλήθους μη πραγματικών γραμμών και της πυκνότητάς τους. Αυτή είναι και η αιτία του χαμηλού ποσοστού αναγνώρισης διαστολών, καθώς για τις δύο σελίδες ο αλγόριθμος αναγνώρισε επιτυχώς το 100% των διαστολών, στην τρίτη σελίδα όμως το ποσοστό αυτό είναι και πάλι μηδενικό.

Στο 8, που αναφέρεται στο Ravel Jeux [52], από τις 13 στο σύνολο παρτιτούρες οι 6 εμφανίζουν ποσοστό επιτυχίας αναγνώρισης ίσο με 100%, ενώ ακόμα 3 έχουν ποσοστό μεγαλύτερο από 90%. Υπάρχει όμως στο σύνολο και μία σελίδα με τόσο μεγάλη κλίση (0,9°, που μεταφράζεται σε ολίσθηση κατά ύψος ίσο με δύο γραμμές πενταγράμμου από την οριζόντια θέση) που και πάλι το ποσοστό αναγνώρισης γραμμών θεωρήθηκε μηδενικό δημιουργώντας το συνολικό αποτέλεσμα ακρίβειας που παρουσιάζεται. Η κλίση σε συνδυασμό με κάποια

μη πραγματικά πεντάγραμμα που δεν απομακρύνθηκαν στο αντίστοιχο στάδιο οδήγησε και στο χαμηλό σχετικά ποσοστό της αναγνώρισης διαστολών.

Στο 12, που αναφέρεται στο Schumann Novelletes, οι 5 από τις 7 παρτιτούρες έχουν ποσοστό επιτυχίας αναγνώρισης γραμμών ίσο με 100%, μία παρτιτούρα εμφανίζει ένα μέτριο ποσοστό 60%, αλλά η τελευταία παρτιτούρα εμφανίζει λόγω κλίσης ($0,85^\circ$) μηδενικό ποσοστό αναγνώρισης. Η κλίση είναι αυτή στην οποία οφείλεται και το χαμηλό ποσοστό αναγνώρισης διαστολών.

Όπως γίνεται αντιληπτό η μεγάλη κλίση των πενταγράμμων δημιουργεί προβλήματα αναγνώρισης στον αλγόριθμο δεδομένου ότι η κύρια μέθοδος που χρησιμοποιείται είναι η οριζόντια προβολή, η αποτελεσματικότητα της οποίας ελαττώνεται με τη μεγάλη αύξηση της κλίσης, καθώς τα πεντάγραμμα παύουν να είναι οριζόντια.

Παρόλα αυτά, υπάρχουν περιπτώσεις όπου παρά την κλίση της παρτιτούρας ο αλγόριθμος δίνει πολύ καλά αποτελέσματα: Στο 1, που αναφέρεται στη Beethoven Sonata [46], δύο σελίδες εμφανίζουν κλίση $0,5^\circ$ και $0,6^\circ$ αλλά τα ποσοστά επιτυχίας στην αναγνώριση των γραμμών τους είναι 100% και 97% αντίστοιχα. Στο 2, που αναφέρεται στο Chopin Nocturnes [45], υπάρχει μία σελίδα με κλίση $0,4^\circ$ στην οποία το ποσοστό επιτυχίας για την αναγνώριση των γραμμών είναι 100%. Το ίδιο ισχύει και για το 3, που αναφέρεται στο Brahms Cadenzas, με την κλίση αυτή τη φορά να είναι ίση με $0,4^\circ$, καθώς και για το 7, που αναφέρεται στη Mozart Sonata, με κλίση ίση με $0,5^\circ$. Αντίστοιχα, για κλίση ίση με $0,7^\circ$, $0,6^\circ$ και $0,85^\circ$ στο 10 (Schumann Carnaval [9]) με επιτυχία 100%.

Φαίνεται δηλαδή ότι η λειτουργία του αλγορίθμου περιορίζεται σημαντικά για κλίση σελίδας μεγαλύτερη από περίπου $0,85^\circ$.

Γενικά παρατηρείται ότι οι born digital παρτιτούρες εμφανίζουν μεγαλύτερο ποσοστό επιτυχίας στην αναγνώριση των γραμμών. Αν και οι ομάδες παρτιτούρων scanned και born digital δεν είναι ισάριθμες, η παρατήρηση αυτή μπορεί να υποτεθεί πως ισχύει γενικά καθώς στις born digital παρτιτούρες ο θόρυβος είναι σχεδόν μηδενικός και οι παρτιτούρες πολύ καθαρές και χωρίς κλίση. Από την άλλη πλευρά, στην αναγνώριση των διαστολών, οι scanned παρτιτούρες εμφανίζουν μεγαλύτερο ποσοστό επιτυχίας. Ξανά, οι ομάδες δεν είναι ισάριθμες, αλλά τα χαμηλότερα ποσοστά των born digital οφείλονται κυρίως στο γεγονός του μη περιορισμού μη πραγματικών πενταγράμμων που παρεμβαίνουν στη διαδικασία του ταιριάσματος των πενταγράμμων. Τα πεντάγραμμα αυτά συντίθενται κυρίως λόγω του ενιαίου κατωφλίου που χρησιμοποιείται για born digital και scanned παρτιτούρες: προκειμένου να γίνουν αποδεκτά υπό εντοπισμό πεντάγραμμα στις scanned το κατώφλι πρέπει να είναι αρκετά μεγάλο, γεγονός που επιτρέπει στις born digital να κάνουν δεκτά μη πραγματικά πεντάγραμμα.

Για παράδειγμα το 17, που αναφέρεται στο SwanLake Scene [55], καθώς υπάρχουν δύο μόνο σελίδες το χαμηλό ποσοστό αναγνώρισης διαστολών οφείλεται στην παρουσία επιπλέον πενταγράμμων που δεν απομακρύνθηκαν

στο αντίστοιχο στάδιο, μην επιτρέποντας το σωστό ταίριασμα των πενταγράμμων στα συστήματα.

Η χρήση διαφορετικού κατωφλίου αποδοχής πενταγράμμων για born digital και scanned παρτιτούρες θα μπορούσε να αποτελέσει λύση στο πρόβλημα. Λύση επίσης φαίνεται να έχει και το πρόβλημα των κεκλιμένων παρτιτούρων: ένα είδος προεπεξεργασίας που θα ελάττωνε την κλίση θα περιόριζε κατά πολύ το πρόβλημα της κλίσης, αυξάνοντας ταυτόχρονα τα ποσοστά επιτυχίας του αλγορίθμου.

Καθώς κατά τη βιβλιογραφική μελέτη διαπιστώθηκε ότι οι περισσότερες μέθοδοι δεν δίνουν ποσοστά επιτυχίας στην ακρίβεια του εντοπισμού γραμμών, ούτε εργάζονται σε ενιαίο σύνολο παρτιτούρων ή έστω πλήθος παρτιτούρων, η σύγκριση με αυτές δε μπορεί να προσφέρει αντικειμενική εποπτεία της θέσης του προτεινόμενου αλγορίθμου στο πλαίσιο της υπάρχουσας έρευνας.

4.2 Συμπεράσματα

Συνολικά, προτάθηκε και παρουσιάστηκε ένας αλγόριθμος για τον εντοπισμό των γραμμών και διαστολών ενός μουσικού κειμένου για πιάνο, ανεξάρτητα από το είδος του, τον τρόπο συγγραφής, την ποιότητα, το μέγεθος και τα επιμέρους χαρακτηριστικά του. Ο αλγόριθμος ακολούθησε την κατεύθυνση της ελάχιστης επεξεργασίας, καθώς έγινε χρήση απλών μεθόδων, οι οποίες σε συνδυασμό με την προσέγγιση hill climbing που υιοθετήθηκε έδωσαν ένα πολλά υποσχόμενο αποτέλεσμα στην αναγνώριση ενός συνόλου διαφορετικών μεταξύ τους παρτιτούρων.

4.3 Μελλοντικές εξελίξεις

Η περαιτέρω βελτιστοποίηση του αλγορίθμου και η προσθήκη ενός επιπλέον σταδίου για την αναγνώριση των θέσεων κάποιων βασικών μουσικών στοιχείων της παρτιτούρας του πιάνου (κλειδιά, οπλισμός, χρόνος) αναμένεται να καταστήσει δυνατή την υλοποίηση ενός πλήρους συστήματος ψηφιακού μουσικού αναλογίου που θα δέχεται σαν είσοδο οποιαδήποτε παρτιτούρα ενώ ταυτόχρονα θα παρέχει διευρυμένες δυνατότητες απεικόνισης της παρτιτούρας σύμφωνα με τις απαιτήσεις του χρήστη.

Η υιοθέτηση της προτεινόμενης κατεύθυνσης, μπορεί να επεκταθεί με μικρές ή μεγαλύτερες τροποποιήσεις και σε παρτιτούρες διαφορετικών μουσικών οργάνων που εμφανίζουν απλούστερη μορφή από το πιάνο, όπως η κιθάρα και το βιολί, αλλά και σε παρτιτούρες πιο περίπλοκες σε μορφή από αυτή του πιάνου, όπως του μαέστρου μιας συμφωνικής ορχήστρας.

Βιβλιογραφία

- [1] Α. Αμαραντίδης, Το Τονικό Μουσικό Σύστημα, η θεωρία της μουσικής, Αθήνα: Κ. Παπαρηγορίου, Χ. Νάκας, 1990.
- [2] [Ηλεκτρονικό]. Available: <http://openclipart.org/detail/118225>. [Πρόσβαση Ιούλιος 2012].
- [3] «Wikipedia, the free encyclopedia,» [Ηλεκτρονικό]. Available: http://en.wikipedia.org/wiki/Modern_musical_symbols#Piano. [Πρόσβαση Ιούλιος 2012].
- [4] [Ηλεκτρονικό]. Available: erianostudio.com. [Πρόσβαση Ιούλιος 2012].
- [5] P. Nickol, Learning to Read Music, 3rd Edition, Oxford, United Kingdom: How to Books, 2008.
- [6] «Sibelius 7, trial version,» Avid Technology, 2012. [Ηλεκτρονικό]. Available: <http://www.sibelius.com/cgi-bin/download/get.pl?com=sh&prod=sibdemo>.
- [7] [Ηλεκτρονικό]. Available: <http://en.wikipedia.org/wiki/File:Commontime.svg>.
- [8] [Ηλεκτρονικό]. Available: <http://en.wikipedia.org/wiki/File:Allabreve.svg>.
- [9] R. Schumann, «Carnaval[Op 9.],» Edwin F. Kalmus, New York, NY, Indiana University Online Musical Scores, Available: <http://www.dlib.indiana.edu/variations/scores/>.
- [10] R. Schumann, «Novellettes [Op.21 N°1],» Available: [free-scores.com](http://www.free-scores.com).
- [11] L. v. Beethoven, «For Elise,» Available: [free-scores.com](http://www.free-scores.com).
- [12] J. Brahms, «Two Cadenzas for Beethoven's piano concerto in G major [Op. 58],» Available: [free-scores.com](http://www.free-scores.com).
- [13] [Ηλεκτρονικό]. Available: http://www.sibelius.com/home/index_flash.html. [Πρόσβαση Ιούλιος 2012].
- [14] [Ηλεκτρονικό]. Available: <http://www.finalemusic.com/default.aspx>. [Πρόσβαση Ιούλιος 2012].
- [15] [Ηλεκτρονικό]. Available: <http://lilypond.org/>. [Πρόσβαση Ιούλιος 2012].
- [16] [Ηλεκτρονικό]. Available: <http://www.sibelius.com/products/scorch/index.html>. [Πρόσβαση Ιούλιος 2012].

- [17] [Ηλεκτρονικό]. Available: <http://www.forscoreapp.com/>. [Πρόσβαση Ιούλιος 2012].
- [18] [Ηλεκτρονικό]. Available: <http://www.sibelius.com/products/photoscore/lite.html>. [Πρόσβαση Ιούλιος 2012].
- [19] T. Tambouratzis, «The digital music rest e-service,» σε *Proceedings of the “16th International Conference on Computing CIC-2007”, Mexico City, Mexico, November 4th-9th*, A.Gelbukh, S.S.Cuerra, H.C.Castro (eds) in “Research in Computing Science – Special Issue: Advances in Computer Science and Engineering” (Vol. 29)., 2007, pp. 138-146.
- [20] T. Tambouratzis, « The digital music stand: a multi-purpose music-manuscript on-demand e-service,» σε *Proceedings of the “2008 IEEE SMC International Conference on Distributed Human-Machine Systems”, Athens, Greece, V.Marik, J.M.Bradshaw, J.Mayer, W.A.Gruver, P.Benda, (eds)., March 9th-12th, 2008*, pp. 299-304.
- [21] T. Tambouratzis, «Identification of Key Music Symbols for Optical Music Recognition and On-Screen Presentation,» σε *Proceedings of the “IJCNN 2011 - International Joint Conference on Neural Networks”, San Jose, California, U.S.A, July 31st-August 6th, 2011*, pp. 1935-1942 (Best Paper Award nomination).
- [22] Ana Rebelo, Ichiro Fujinaga, Filipe Paszkiewicz, Andre R.S. Marcal, Carlos Guedes, Jaime S. Cardoso, *Optical music recognition: state-of-the-art and open issues*, London: Springer, 2012.
- [23] R. Goecke, «Building a system for writer identification on handwritten musc scores,» σε *Proceedings of IASTED international conference on signal processing, pattern recognition, and applications*, Anaheim, Acta Press, 2003, pp. 205-225.
- [24] Kia Ng, David Cooper, Ewan Stefani, Roger Boyle, Nick Bailey, «Embracing the composer: optical recognition of handwritten manuscripts,» σε *Proceedings of the international computer music conference*, 1999, pp. 500-503.
- [25] David Bainbridge, Tim Bell, «Identifying music documents in a collection of images,» σε *International Conference on Music Information Retrieval and Related Activities (ISMIR 2006)*, 8-12 , Victoria, Canada, Oct 2006.
- [26] Davin Bainbridge, Tim Bell, «An Extensible Optical Music Recognition System,» σε *19th Australian Computer Science Conference*, 1997.
- [27] Hidetoshi Miyao, Masayuki Okamoto, «Stave Extraction for Printed Music Scores Using DP Matching,» *Journal of Advanced Computational Intelligence and Intelligent Informatics*, p. Vol.8 No.2, 2004.
- [28] L. S. Johansen, *Optical Music Recognition*, Thesis for the Degree of Master of Science in Modelling and Data Analysis, University of Oslo, 2009, p. 16.

- [29] Christoph Dalitz, Michael Droettboom, Bastian Pranzas, Ichiro Fujinaga, «A Comparative Study of Staff Removal Algorithms,» σε *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.30, No.5, May 2008, pp. 753-766.
- [30] David Bainbridge, Tim Bell, «The Challenge of Optical Music Recognition,» σε *Computers and Humanities 35*, Printed in the Netherlands, Kluwer Academic Publishers, 2001, pp. 95-121.
- [31] R.O. Duda, P.E. Hart, «Use of the Hough Transform to detect lines and curves in pictures,» σε *Communication of the ACM*, January 1972, pp. 11-15.
- [32] G. Sayeed Choudhury, Michael Droettboom, Tim DiLauro, Ichiro Fujinaga, Brian Harrington, «Optical music recognition system within a large-scale digitization,» σε *Proceedings of the International Society of Music Information retrieval*, 2000.
- [33] I. Fujinaga, «Staff detection and removal,» σε *George S(ed) Visual perception of music notation: on-line and off-line recognition*, Idea Group Inc, Hershey, pp. 1-39.
- [34] S. Marinai, P. Nesi, «Projection Based Segmentation of Musical Sheets,» σε *Proceedings of the Fifth International Conference on Document Analysis and Recognition, ICDAR '99*, 20-22 Sep 1999, pp. 515-518.
- [35] Irwin Sobel, G.Feldman, «A 3x3 isotropic gradient operator for image processing,» Presented at a talk at the Stanford Artificial project, 1968.
- [36] P. Bellini, I. Bruno, P. Nesi, «Optical music sheet segmentation,» σε *Proceedings of the first international conference on web deliberating of music*, 2001, pp. 183-190.
- [37] M. Szwoch, «A Robust Detector for Distorted Music Staves,» σε *Lecture notes in computer science*, Vol. 3691, Springer, 2005, pp. 701-708.
- [38] L. Pugin, «Optical Music Recognition of Early Typographic Prints using Hidden Markov Models,» σε *Proc. Seventh Int'l Conf. Music Information Retrieval*, 2006, pp. 53-56.
- [39] G.E. Kopec, P.A. Chou, «Markov Source Model for Printed Music Decoding,» σε *Journal of Electronic Imaging*, vol.5, no.1, 1996.
- [40] Jaime dos Santos Cardoso, Artur Capela, Carlos Guedes, Joaquim Pinto da Costa, «Staff detection with Stable Paths,» σε *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.31, No. 6, June 2009.
- [41] Scott Sheridan, Susan E. George, «Defacing music score for improved recognition,» σε *Abraham G., Rubinstein BIP (eds) Proceedings of the second Australian undergraduate students' computing conference.*, Australian undergraduate students' conference, 2004, pp. 1-7.

- [42] Stuart J. Russell, Peter Norvig, «Artificial Intelligence: A modern approach (2nd ed.),» Upper Saddle River, New Jersey, Prentice Hall, pp. 111-114.
- [43] W. A. Mozart, « Sonata No. 6 in F-major,[K280],» Available: free-scores.com.
- [44] C. Debussy, «La plus que Lente (Valse),» Available: Project Gutenberg: Gutenberg.org compiled in pdf by free-scores.com.
- [45] F. Chopin, «Nocturnes [Op.27,» London, Augener, Ed. 6094, Plate 8081, 1882, Available: free-scores.com.
- [46] L. v. Beethoven, «Sonata dedicata a Joseph Haydn [Op.2.N.1],» Alfredo Casella, Indiana University Online Musical Scores Available: <http://www.dlib.indiana.edu/variations/scores/>.
- [47] G. Bizet, «Prelude from Carmen,» Mutoxia project- 2005/11/22-635, Available: free-scores.com.
- [48] G.F.Händel, «Largo,» New York, G.Schirmer, arranged by A.R. Parsons, No.1 Original Edition, Available: free-scores.com.
- [49] F. Chopin, «Concerto No.2 [Op. 21],» Leipzig, Breitkopf & Härtel, n.d.(April 1836), Plate 5654, Available: free-scores.com.
- [50] F. P. Schubert, «Waltzes Early Versions of Dances [Dances],» Available: free-scores.com.
- [51] F. B. Mendelssohn, « Retour dans la Patrie [Op.89]».
- [52] M. Ravel, «Jeux d' eau,» Courtesy of The Sheet Music Archive, www.sheetmusicarchive.com, Available: free-scores.com.
- [53] C. Debussy, « Pièce pour piano (Morceau de Concours),» Gutenberg.org compiled in pdf by free-scores.com.
- [54] L. Brendon, «Piano Solo No.3, Op.1 No 3,,» Available:free-scores.com, 2011.
- [55] P. I. Tchaikovsky, «Scene from Swan Lake,» *œ simplified by Clemen Carl de Guzman*, Available:free-scores.com, 2009.