

ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

Τομέας Επικοινωνιών, Ηλεκτρονικής και Συστημάτων

Έλεγχος της Επίδοσης Αλγορίθμου Κατάτμησης
Έγχρωμων Εικόνων με Βάση το Ιστόγραμμα
Ισοδύναμων Εκδοχών τους στην Κλίμακα του Γκρι

—— Διπλωματική Εργασία ——
Σταμάτης Ζαφειρόπουλος

—— Επιβλέπων Καθηγητής ——
Κωνσταντίνος Παπαοδυσσεύς
Καθηγητής ΕΜΠ





ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ
ΜΗΧΑΝΙΚΩΝ ΚΑΙ
ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

Τομέας Επικοινωνιών, Ηλεκτρονικής και
Συστημάτων

Έλεγχος της Επίδοσης Αλγορίθμου Κατάτμησης
Έγχρωμων Εικόνων με Βάση το Ιστόγραμμα
Ισοδύναμων Εκδοχών τους στην Κλίμακα του Γκρι

Διπλωματική Εργασία Σταμάτης Ζαφειρόπουλος

Επιβλέπων Καθηγητής:
Κωνσταντίνος Παπαοδυσσεύς
Καθηγητής ΕΜΠ

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 31^η Μαρτίου 2017

.....
Κωνσταντίνος Παπαοδυσσεύς
Καθηγητής ΕΜΠ

.....
Ηλίας Κουκούτσης
Επίκουρος Καθηγητής ΕΜΠ

.....
Γεώργιος Καμπουράκης
Αναπληρωτής Καθηγητής ΕΜΠ

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

Copyright © Σταμάτης Ζαφειρόπουλος, 2017
Copyright © Κωνσταντίνος Παπαοδυσσεύς, 2017
Εξώφυλλο: Θάνος Γκιώνης

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 εξετάζονται διάφορες μέθοδοι κατάτμησης εικόνας και επιλέγεται μία προταθείσα προς υλοποίηση και επέκταση.

Αρχικά γίνεται μία μικρή εισαγωγή στους χρωματικούς χώρους CIEXYZ, CIELUV και CIELAB. Έπειτα, παρουσιάζονται οι μέθοδοι κατάτμησης εικόνας K-σημείων, Διαμέσου, Κατωφλίου με τη Χρήση Ιστογράμματος, Εντοπισμού Ακμών, Ενεργών και Γεωδαιτικών Ενεργών Περιγραμμάτων και Watershed, ενώ υιοθετείται μία σύγχρονη προταθείσα μέθοδος κατάτμησης εικόνας με ανάθεση των εικονοστοιχείων σε περιοχές που περιγράφονται από κανονικές κατανομές επί του ιστογράμματος της εικόνας. Η μέθοδος αυτή επεκτείνεται με την επεξεργασία της εικόνας στο χρωματικό χώρο CIELAB, του οποίου οι ιδιότητες βοηθούν στην άμεση παραγωγή του ιστογράμματος εικόνας απόχρωσης του γκρι της εικόνας μας.

Παρουσιάζονται παραδείγματα της εύρεσης της κλίσης εικόνων στο χώρο CIELAB και παραγωγής του ιστογράμματος εικόνας απόχρωσης του γκρι, ενώ παρουσιάζεται και μία σειρά από ενδεικτικά αποτελέσματα κατάτμησης ως αποτέλεσμα κατάτμησης του ιστογράμματος, σε εικόνες απόχρωσης του γκρι καθώς και δυαδικές.

Λέξεις Κλειδιά:

Κατάτμηση εικόνας, Κανονική κατανομή, Κλίση εικόνας, Ιστόγραμμα εικόνας απόχρωσης του γκρι, χρωματικός χώρος CIELAB

Abstract

In the present Diploma Thesis, several image segmentation methods are examined and one is chosen for implementation and expansion.

At first, a small introduction is being held on color spaces CIEXYZ, CIELUV and CIELAB. Afterwards, image segmentation methods K-means, Median, Histogram Thresholding, Edge Detection, Active Contours and Geodesic Active Contours and Watershed are presented, and eventually a proposed modern image segmentation method of classifying pixels into areas described by normal distributions on the image's histogram is adopted. This method is expanded by working on the picture in CIELAB color space, the properties of which help us easily produce our image's grayscale histogram.

Examples are also presented, showcasing the calculating of images' gradient in CIELAB color space and grayscale histogram, but also a series of indicative segmentation results as a result of histogram segmenting, present in grayscale and binary pictures.

Keywords:

Image Segmentation, Normal Distribution, Image Gradient, Grayscale Image Histogram, CIELAB color space

Ευχαριστίες

Θα ήθελα να ευχαριστήσω τον κύριο Παπαοδυσσέα που μου εμπιστεύτηκε αυτή την εργασία, αλλά και που μου άνοιξε την πόρτα της αρχαιολογίας. Θα ήθελα να ευχαριστήσω επίσης τον υπεύθυνο της εργασίας μου Δημήτρη Αραμπατζή, διδακτορικό φοιτητή στη σπουδαία αυτή ερευνητική ομάδα, που ανέχτηκε ένα εκατομμύριο τηλεφωνήματα.

Τέλος, ευχαριστώ τους γονείς μου για τη στήριξή τους, καθώς και το Φοίβο και τη Φοίβη, ανθρώπους που δείχνουν εμπιστοσύνη στις γνώσεις και την αισθητική μου.

Περιεχόμενα

Εισαγωγή.....	13
Κεφάλαιο 1: Χρωματικοί Χώροι.....	15
1.1: Το μοντέλο CIEXYZ.....	15
1.2: Το μοντέλο CIELUV.....	16
1.3: Το μοντέλο CIELAB.....	17
Κεφάλαιο 2: Κατάτμηση Εικόνας.....	19
2.1: K-means.....	19
2.2: Median.....	20
2.3: Histogram Thresholding.....	20
2.4: Ανίχνευση ακμών εικόνας.....	21
2.5: Ενεργά Περιγράμματα/Παραμετρικά Ενεργά Περιγράμματα.....	23
2.6: Γεωδαιτικά Ενεργά Περιγράμματα.....	24
2.7: Watershed.....	25
Κεφάλαιο 3: Η Υιοθετηθείσα Μέθοδος Κατάτμησης Εικόνας.....	28
3.1: Παρουσίαση της μεθόδου: Ταξινόμηση των εικονοστοιχείων.....	28
3.2: Μείωση θορύβου και βελτίωση περιγραμμάτων.....	30
3.3: Υλοποίηση της μεθόδου: Εξάλειψη θορύβου με εφαρμογή μορφολογικών φίλτρων.....	32
3.4: Κλίση εικόνας στο χρωματικό χώρο LAB και υπολογισμός ισοδύναμης grayscale εικόνας.....	32
3.5: Ενδεικτικές απεικονίσεις του εντοπισμού κλίσης εικόνας στο χρωματικό χώρο LAB και παραγωγής της ισοδύναμου grayscale εικόνας.....	34
1ο παράδειγμα.....	35
2ο παράδειγμα.....	36
3ο παράδειγμα.....	38
4ο παράδειγμα.....	39
5ο παράδειγμα.....	41
6ο παράδειγμα.....	42
7ο παράδειγμα.....	43
8ο παράδειγμα.....	44
9ο παράδειγμα.....	46

10ο παράδειγμα	47
11ο παράδειγμα	49
3.6: Προσέγγιση των 2 περιοχών της εικόνας με τη χρήση κανονικών κατανομών επί του χρωματικού ιστογράμματος	50
3.7: Εξάλειψη του θορύβου	51
Κεφάλαιο 4: Αποτελέσματα Εφαρμογής της Δοκιμασθείσας Μεθόδου..	53
1ο παράδειγμα	53
2ο παράδειγμα	55
3ο παράδειγμα	57
4ο παράδειγμα	58
5ο παράδειγμα	60
6ο παράδειγμα	61
7ο παράδειγμα	63
8ο παράδειγμα	64
9ο παράδειγμα	65
10ο παράδειγμα	67
11ο παράδειγμα	68
Κεφάλαιο 5: Συμπεράσματα.....	70
Πηγές	71

Εισαγωγή

Η ραγδαία εξέλιξη της τεχνολογίας δε θα μπορούσε να αφήσει ανεπηρέαστη και την επιστήμη της αρχαιολογίας. Τα νέα και μονίμως εξελισσόμενα εργαλεία στα χέρια της **τελευταίας** ανοίγουν δρόμους αντίληψης και επαναπροσέγγισης των ευρημάτων, τέτοιους που απαλλάσσουν τους επιστήμονες του κλάδου από τη δύσκολη και πολλές φορές επισφαλής ως προς την ακρίβειά της προσέγγιση με τις παραδοσιακές μεθόδους.

Ενδεικτικά, ένας σύγχρονος αλγόριθμος θα μπορούσε να ελαχιστοποιήσει τη φθορά από μία απεικόνιση· θα μπορούσε να υποκαταστήσει τα κενά μίας απεικόνισης με τις λογικότερες αποχρώσεις· θα μπορούσε να συναρμολογήσει ένα θρυμματισμένο άγαλμα· θα μπορούσε να εντοπίσει γραμματικά στοιχεία σε μία απέραντη σειρά ιερογλυφικών συμβόλων. Ο 21^{ος} αιώνας φωτίζει την ήδη υπάρχουσα Ιστορία με τρόπο τέτοιο που είναι βέβαιο ότι θα προκύψουν πολλά νέα ιστορικά συμπεράσματα.

Στην παράδοση της συνεργασίας της ερευνητικής ομάδας του κυρίου Παπαοδυσσέα με αρχαιολογικά ιδρύματα, οργανισμούς, μουσεία και λοιπούς φορείς, η ερευνητική ομάδα παρέλαβε εικόνες από βυζαντινά χειρόγραφα αντίγραφα της Ιλιάδας. Σκοπός της ομάδας είναι, δουλεύοντας επάνω στο ψηφιοποιημένο αρχείο, να παραγάγει ένα μοντέλο το οποίο θα περιγράφει το γραφικό χαρακτήρα του καταγραφέα των κειμένων. Πρακτικά, αυτό επιτυγχάνεται ξεκινώντας με την άντληση κάθε γράμματος των κειμένων αυτών.

Εκ των πραγμάτων, μετά την αποκοπή του κάθε γράμματος και του υπόβαθρου του από την κάθε σελίδα του χειρόγραφου, δημιουργήθηκε η ανάγκη για μία ακριβή απομόνωση των γραμμάτων από το υπόβαθρό τους αυτό καθεαυτό, δηλαδή τον πάπυρο. Προκειμένου αυτό να εφαρμοστεί σε μεγάλη κλίμακα, αναζητήθηκε μία αξιόπιστη μέθοδος κατάτμησης εικόνας, που θα μας έδινε ένα πιστό μονοχρωματικό αντίγραφο του κάθε χαρακτήρα, με οποιοδήποτε ίχνος του υπόβαθρου να απουσιάζει. Οι απομονωμένοι, πλέον, χαρακτήρες, είναι έτοιμοι να εξυπηρετήσουν το σκοπό της κατασκευής τους, χρησιμοποιούμενοι στο επόμενο στάδιο της ερευνητικής εργασίας.

Εξετάστηκαν διάφορες μέθοδοι κατάτμησης εικόνας, οι οποίες και παρουσιάζονται. Στην παρούσα διπλωματική εργασία, υιοθετούμε και επεκτείνουμε μία προταθείσα μέθοδο κατάτμησης εικόνας από την ερευνητική ομάδα του κυρίου Παπαοδυσσέα. Πυρήνας αυτής της επέκτασης είναι η απευθείας παραγωγή της εικόνας μας σε αποχρώσεις του γκρι, χάρη στις ιδιότητες του χρωματικού χώρου LAB στον οποίο επιλέγουμε να εργαστούμε. Κατά συνέπεια, για την περίπτωση που θέλουμε να εργαστούμε στον ασπρόμαυρο χώρο (όπως για τους σκοπούς αυτής της ερευνητικής εργασίας), αποφεύγουμε κάποια ενδιάμεση διαδικασία μετατροπής της εικόνας σε εικόνα απόχρωσης του γκρι (grayscale), συντομεύοντας τη συνολική διαδικασία της προταθείσας μεθόδου.

Παρουσιάζονται ορισμένα αντιπροσωπευτικά παραδείγματα υλοποίησης της επεκταμένης μεθόδου μας, τόσο του εντοπισμού της κλίσης στο χρωματικό χώρο LAB, που μας βοηθάει στην απευθείας παραγωγή της εικόνας απόχρωσης του γκρι, όσο και των βημάτων της κατάτμησης αυτής με τη χρήση στατιστικών μεθόδων (κανονικών κατανομών), ενώ τέλος δίνονται και οι επιθυμητές για τους σκοπούς της εργασίας ασπρόμαυρες απεικονίσεις των εικόνων των χαρακτήρων μας.

Κεφάλαιο 1: Χρωματικοί Χώροι

Οι χρωματικοί χώροι, ως βασική επινοήση για την εξυπηρέτηση της πάγιας ανάγκης της οργάνωσης των χρωμάτων, βρίσκονται υπό διαρκή εξέλιξη. Σε γενικές γραμμές, τα μοντέλα αυτών χωρίζονται σε 4 οικογένειες:

- CIE XYZ (η πρώτη απόπειρα για οργάνωση χρωμάτων, βασισμένη σε ένα σύστημα τριών συντεταγμένων)
- RGB (μία προσέγγιση που βασίζεται στο γραμμικό συνδυασμό 3 βασικών χρωμάτων)
- Luma plus chroma (ένα μοντέλο στο οποίο η εικόνα προκύπτει από το συνδυασμό της φωτεινότητας και της κατανομής του χρώματος στις περιοχές της εικόνας)
- Hue and Saturation (ένα μοντέλο-προέκταση του Luma+Chroma, με βασικές παραμέτρους την απόχρωση, τον χρωματικό κορεσμό και τη φωτεινότητα, που προκύπτει από μετασχηματισμένο του RGB σε κυλινδρικό σύστημα συντεταγμένων)

Το βασικότερο χρωματικό μοντέλο, από το οποίο και πηγάζει και η πλειοψηφία των χρωματικών αναπαραστάσεων ως μετέπειτα μετασχηματισμοί του, είναι το RGB. Η υλοποίηση του μοντέλου βασίστηκε στην πολύ απλή σύλληψη του να δημιουργηθούν χρώματα και αποχρώσεις τους από την, ισοδύναμη των βασικών χρωμάτων, παλέτα του κόκκινου-πράσινου-μπλε.

1.1: Το μοντέλο CIE XYZ

Η βασικότερη απόπειρα για αναπαράσταση των χρωμάτων σύμφωνα με την πρόσληψή τους από το ανθρώπινο μάτι είναι το μοντέλο CIE XYZ. Η προσέγγιση αυτή, βασίζεται στην αποσύμπλεξη των χρωματικών παραμέτρων από αυτή της φωτεινότητας, με αποτέλεσμα ο κύβος της RGB αναπαράστασης να μετασχηματίζεται σε ένα επίπεδο χωρίο μέσω της κανονικοποίησης $x = \frac{R}{R+G+B}$, $y = \frac{G}{R+G+B}$.

Με συντεταγμένες στο χωρίο αυτό ενός χρώματος στο RGB τις (x_r, y_r) , (x_g, y_g) και (x_b, y_b) , και με το λευκό σημείο αναφοράς το (X_w, Y_w, Z_w) η CIE XYZ αναπαράσταση προκύπτει μέσω του κάτωθι γραμμικού μετασχηματισμού της RGB αναπαράστασης ως εξής:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = [M] \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

Όπου

$$[M] = \begin{bmatrix} S_r X_r & S_g X_g & S_b X_b \\ S_r Y_r & S_g Y_g & S_b Y_b \\ S_r Z_r & S_g Z_g & S_b Z_b \end{bmatrix}$$

$$X_r = \frac{x_r}{y_r}, Y_r = 1, Z_r = \frac{1-x_r-y_r}{y_r}, X_g = \frac{x_g}{y_g}, Y_g = 1, Z_g = \frac{1-x_g-y_g}{y_g}, X_b = \frac{x_b}{y_b}, Y_b = 1, Z_b = \frac{1-x_b-y_b}{y_b}$$

Και

$$\begin{bmatrix} S_r \\ S_g \\ S_b \end{bmatrix} = \begin{bmatrix} X_r & X_g & X_b \\ Y_r & Y_g & Y_b \\ Z_r & Z_g & Z_b \end{bmatrix}^{-1} \begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix}$$

Η ιστορία και η πράξη οδήγησε στο βασικό διαχωρισμό των χρωματικών μοντέλων-αναπαραστάσεων σε δύο κατηγορίες: τα προσθετικά και τα κυλινδρικά. Τα προσθετικά, όπως το RGB, ενδείκνυνται στην τεχνολογία και συγκεκριμένα στις οθόνες (monitor), όπου το τελικό αποτέλεσμα παράγεται με τη μετάδοση φωτός υπό τη μορφή βασικών χρωμάτων. Με αυτή την παλέτα παράγονται συνδυασμοί κι εν τέλει μία μεγάλη γκάμα χρωμάτων και αποχρώσεων. Η απεικόνιση αυτών των μοντέλων όμως δεν είναι καθόλου βολική ως προς την αντίληψη και την έρευνα, καθότι πρόκειται για κύβους με κατανομή των αποχρώσεων τέτοια που καθιστά αδύνατο τον εντοπισμό της απόχρωσης που θέλουμε. Αντίθετα, τα κυλινδρικά μοντέλα, τα οποία θα μελετήσουμε παρακάτω, δημιουργούν μία απεικόνιση που αφενός το ανθρώπινο μάτι μπορεί να αντιληφθεί καλύτερα, αφετέρου βασίζονται σε παραμέτρους αισθητές από την ανθρώπινη αντίληψη, όπως το χρώμα, η φωτεινότητα, ο χρωματικός κορεσμός κλπ. [1]

1.2: Το μοντέλο CIELUV

Το CIELUV είναι ένας μετασχηματισμός των συντεταγμένων του χώρου του μοντέλου CIE με σκοπό να προκύψει ένα μοντέλο με συνεχέστερη ομοιομορφία από αυτή του CIEXYZ. Για ένα λευκό σημείο με συντεταγμένες (X_n, Y_n, Z_n) , ο μετασχηματισμός σε συντεταγμένες (L^*, u^*, v^*) , όπου L^* η φωτεινότητα (τροποποιημένη έτσι από παλαιότερα μοντέλα ώστε να αυξάνεται με βήματα που είναι ορατά στο ανθρώπινο μάτι), και u^* και v^* οι 2 καρτεσιανές διαστάσεις χρώματος που συνδυαζόμενες παράγουν μία απόχρωση, προκύπτει από τις παρακάτω σχέσεις:

$$L^* = \begin{cases} \left(\frac{29}{3}\right)^3 \frac{Y}{Y_n} & \text{αν } \frac{Y}{Y_n} \leq \left(\frac{6}{29}\right)^3 \\ 116\left(\frac{Y}{Y_n}\right)^{\frac{1}{3}} & \text{αν } \frac{Y}{Y_n} > \left(\frac{6}{29}\right)^3 \end{cases}$$

$$u^* = 13L^* \times (u' - u'_n)$$

$$v^* = 13L^* \times (v' - v'_n)$$

Όπου

$$u' = \frac{4X}{X + 15Y + 3Z} = \frac{4x}{-2x + 12y + 3}$$

$$v' = \frac{9Y}{X + 15Y + 3Z} = \frac{9y}{-2x + 12y + 3}$$

Η κυλινδρική απεικόνιση αυτού, τέλος, έχει ως εξής:

$$C^*_{uv} = \sqrt{(u^*)^2 + (v^*)^2}$$

$$h_{uv} = \text{atan2}(u^*, v^*)$$

Οι τελευταίες δύο εξισώσεις δίνουν τις πολικές συντεταγμένες του καρτεσιανού ζεύγους (u^*, v^*) . [2]

1.3: Το μοντέλο CIELAB

Το CIELAB είναι ένας μη γραμμικός μετασχηματισμός του χώρου CIEXYZ. Ο κύριος λόγος επινόησής του ήταν η ανάγκη να προκύψει ένας πιο κατανοητός στο ανθρώπινο μάτι χώρος όσον αφορά τις αποχρώσεις και τη μετάβαση από το ένα χρώμα στο άλλο. Το CIELAB, ως τρισδιάστατος κυλινδρικός χώρος, προκύπτει ως αποτέλεσμα της πειραματικής αντιμετώπισης αυτού του προβλήματος προς αυτή την κατεύθυνση, και παρότι μη γραμμικός, τελικά αποδίδει τη διαφορά μεταξύ 2 χρωμάτων (Color Difference) ως μία γραμμική σχέση.

Με νέες συντεταγμένες τις (L^*, a^*, b^*) , όπου L^* η φωτεινότητα (ο άξονας του κυλίνδρου) και a^*, b^* 2 κάθετες διαστάσεις επί του άξονα στο ύψος μία τιμής του L^* που παράγουν τις τελικές αποχρώσεις (με το εύρος του a^* να είναι από πράσινο έως κόκκινο και αυτό του b^* από μπλε έως κίτρινο), ο μετασχηματισμός από τον χώρο CIEXYZ είναι ο εξής:

$$L^* = 116f\left(\frac{Y}{Y_n}\right) - 16$$

$$a^* = 500 \left[f\left(\frac{X}{X_n}\right) - f\left(\frac{Y}{Y_n}\right) \right]$$

$$b^* = 200 \left[f\left(\frac{Y}{Y_n}\right) - f\left(\frac{Z}{Z_n}\right) \right]$$

Όπου

$$f(t) = \begin{cases} t^{\frac{1}{3}}, & \alpha \nu t > \left(\frac{6}{29}\right)^3 \\ \frac{1}{3}\left(\frac{29}{6}\right)^2 t + \frac{4}{29}, & \alpha \nu t \leq \left(\frac{6}{29}\right)^3 \end{cases}$$

Με X_n , Y_n και Z_n το λευκό σημείο αναφοράς, συνηθίζεται βάσει του συστήματος D65 (ένα από τα πιο συνηθισμένα χρωματικά προφίλ, προσαρμοσμένο στις συνθήκες φωτισμού της Βόρειας και Κεντρικής Ευρώπης ώστε οι τιμές των χρωμάτων να δίνουν αποτελέσματα όσο το δυνατόν πιο κοντινά στον καθολικό ορισμό του χρώματος) αυτά να είναι $X_n = 95,047$, $Y_n = 100,000$ και $Z_n = 108,883$.

Με βάση τις παραπάνω σχέσεις, η διαφορά 2 χρωμάτων L^*_1 και L^*_2 ποσοτικοποιείται με την ευκλείδεια απόσταση των απεικονίσεών τους στο χώρο R^3 :

$$\Delta E^*_{ab} = \sqrt{(L^*_2 - L^*_1)^2 + (a^*_2 - a^*_1)^2 + (b^*_2 - b^*_1)^2}$$

Πειραματικά έχει προκύψει ότι μία μόλις αισθητή στο ανθρώπινο μάτι αλλαγή απόχρωσης ξεκινάει όταν η διαφορά αγγίξει το $\Delta E^*_{ab} = 2,3$, αν και για έμπειρα μάτια αυτό θεωρείται μπορεί να γίνει από το $\Delta E^*_{ab} = 1$ κιόλας. [3], [4]

Κεφάλαιο 2: Κατάτμηση Εικόνας

Η επεξεργασία εικόνας, ως αυτοτελής κλάδος της επιστήμης των υπολογιστών, δεν άργησε να αναπτυχθεί σε βαθμό υπερθετικό. Μία από τις σημαντικότερες και μονίμως εξελισσόμενες τεχνικές αυτής είναι η κατάτμηση εικόνας. Ως σύλληψη, η ιδέα της συλλογής των δεδομένων που προκύπτουν από την διαίρεση της εικόνας σε ομάδες τμημάτων με κοινά μεταξύ τους χαρακτηριστικά, απεδείχθη επαναστατική. Η ανάλυση και επεξεργασία μίας επιμέρους κατηγορίας χαρακτηριστικών μίας εικόνας διευρύνει τις δυνατότητες μελέτης και επεκτείνει την επικέντρωση σε ένα χαρακτηριστικό αυτής σε βαθμό πρακτικά ανεξάντλητο.

Σε γενικές γραμμές, στην κατάτμηση εικόνας χρησιμοποιούνται τεχνικές κατάτμησης και ομαδοποίησης δεδομένων που ανήκουν χονδρικά στις δυικές κατηγορίες αλγορίθμων των διαχωριστικών και των ιεραρχικών τεχνικών.

Η λογική των διαχωριστικών τεχνικών έγκειται στην υλοποίηση μίας διαδικασίας ομαδοποίησης με τελικό στόχο τον διαχωρισμό και την απομόνωση των ομαδοποιήσεων που θα δημιουργηθούν (K-means, median κλπ). Αντίθετα, οι ιεραρχικές τεχνικές λειτουργούν στη βάση της συνένωσης των ομαδοποιήσεων που θα προκύψουν (κατάτμηση ιστογράμματος, μέθοδος Watershed κλπ).

2.1: K-means

Μία ιδιαίτερα διαδεδομένη κατηγορία μεθόδων που εντάσσεται σε αυτές των διαχωριστικών τεχνικών είναι αυτή του κέντρου βάρους. Η πιο γνωστή μέθοδος αυτής της κατηγορίας είναι η K-means, η οποία δουλεύει ως εξής: χωρίζεται η εικόνα (τυχαία) σε N περιοχές με την καθεμία να διαθέτει ένα κέντρο βάρους. Κάθε σημείο γύρω από αυτό εντάσσεται στην αντίστοιχη περιοχή, δεδομένου ότι η απόστασή του (ευκλείδεια, χρωματική κλπ) δεν ξεπερνάει την ακτίνα του, δηλαδή βρίσκεται στο κοντινότερο κέντρο βάρους που είναι διαθέσιμο.

Για κάθε pixel μίας περιοχής (στην περίπτωση των γραμμάτων και της αρκετά απλής αντίθεσης του γράμματος με το υπόβαθρο, οι περιοχές είναι $N=2$), ελέγχεται το χρωματικό του περιεχόμενο κι έτσι αυτό αποδίδεται σε μία χρωματική περιοχή. Αν το στοιχείο πλησιάζει χρωματικά στην αρχική εκτίμηση του υποβάθρου, η οποία είναι επιλεγμένη ως το χρωματικό minimum της εικόνας, τότε αποδίδεται σε αυτό, ενώ το αντίστοιχο συμβαίνει αν πλησιάζει στην αρχική εκτίμηση του γράμματος, δηλαδή το χρωματικό maximum. Για κάθε επιπλέον απόδοση pixel σε μία περιοχή, το χρωματικό περιεχόμενο της περιοχής αυτής ορίζεται και πάλι ως ο χρωματικός μέσος όρος των στοιχείων που έχουν αποδοθεί σε αυτήν μέχρι στιγμής.

Ένα μεγάλο μειονέκτημα αυτής της μεθόδου, είναι το ότι δε λαμβάνει υπόψη την τοπολογία της εικόνας. Για παράδειγμα, στην περίπτωση της εικόνας ενός γράμματος, αν υπάρχει ένα φθαρμένο σημείο στο υπόβαθρο που δημιουργεί

μία σκούρα απόχρωση, αυτό είναι πολύ πιθανό να αποδοθεί στην χρωματική περιοχή του γράμματος, και αντίστροφα, αν υπάρχει ένα κενό εντός του γράμματος, αυτό να αποδοθεί στο υπόβαθρο.

2.2: Median

Μία εξίσου χρήσιμη μέθοδος είναι αυτή της διαμέσου (median). Με λογική χωρισμού της εικόνας σε N περιοχές, όπως η K -means, διαφέρει ως προς το ότι η αναφορά σε κάθε περιοχή δεν είναι το χρωματικό κέντρο βάρους αλλά η ενδιάμεση χρωματική τιμή που προκύπτει από τα ταξινομημένα ρίξει της συνολικής περιοχής. Με αναφορά λοιπόν αυτήν, η (χρωματική) ακτίνα απόστασης ενός στοιχείου από το σημείο-αντιπρόσωπο (medoid) είναι εξορισμού μικρότερη, καθότι το στοιχείο δε συγκρίνεται με μία ακραία χρωματική τιμή μίας περιοχής, αλλά με τη μέση τιμή των στοιχείων της. Τέλος, η μέση τιμή παραμένει αυτή που είναι καθ'όλη την ταξινόμηση των στοιχείων και δεν επαναπροσδιορίζεται ως ο καινούργιος μέσος όρος των στοιχείων, όπως στην K -means.

Και οι 2 παραπάνω μέθοδοι λειτουργούν καλύτερα όταν τα σχήματα των περιοχών είναι κυκλικά ή παραπλησία, με την μέθοδο της διαμέσου να είναι αρκετά πιο αξιόπιστη, όμως και οι 2 μπορούν να γίνουν προβληματικές όταν λειτουργήσουν σε περιοχές με μεγάλη διαφορά μεγεθών μεταξύ τους.

2.3: Histogram Thresholding

Μία πολύ χρήσιμη μέθοδος κατάτμησης είναι αυτή του κατωφλίου με τη χρήση ιστογράμματος. Αρχικά λαμβάνουμε ένα ιστόγραμμα από την εικόνα που σκοπεύουμε να επεξεργαστούμε, στο οποίο αποτυπώνεται σε μορφή γραφικής παράστασης η κατανομή του χρώματος στην εικόνα, στη μορφή συμπύκνωσης-χρωματικού ύψους. Το ιστόγραμμα αυτό αποτελείται από «λόφους» και «κοιλιάδες». Οι λόφοι, ως περιοχές απότομης χρωματικής συμπύκνωσης, μαρτυρούν την παρουσία κάποιου θέματος. Αν ο λόφος είναι πολύ λεπτός, τότε πιθανόν μιλάμε για ένα αντικείμενο, αντίθετα αν είναι πιο φαρδύς, πιθανότατα δηλώνει το φόντο. Αυτό που δε μπορεί να προσδιοριστεί είναι το σε ποιο χωρίο ανήκουν τα εικονοστοιχεία των κοιλιάδων.

Ουσία της μεθόδου αυτή είναι ο ορισμός μίας τιμής κατωφλίου κατά προσέγγιση ανάμεσα στους λόφους (δηλαδή μες στην κάθε κοιλιάδα) του ιστογράμματος. Στην περίπτωση 2 λόφων, η τελική εικόνα λαμβάνει μόνο 2 τιμές:

$$t(m, n) = \begin{cases} 1, & \text{αν } f(m, n) > T \\ 0, & \text{αν } f(m, n) \leq T \end{cases}$$

Η τελική εικόνα θα είναι προφανώς δυαδική. Στην περίπτωση όμως περισσότερων αντικειμένων και αποχρώσεων, έχουμε τη δυνατότητα να ορίσουμε πολλές τιμές κατωφλίου (μία για κάθε κοιλιάδα) και να παράξουμε μία εικόνα σε περισσότερες αποχρώσεις του γκρι. Δεδομένου ότι κάθε

χρωματική αλλαγή στην καινούργια εικόνα (δηλαδή μία επιπλέον απόχρωση του γκρι), δηλώνει μία αλλαγή θέματος-αντικειμένου στην εικόνα μας, εύκολα από αυτό προσδιορίζονται οι ακμές τους και αποκόπτονται από το περιβάλλον τους. [5]

2.4: Ανίχνευση ακμών εικόνας

Εδώ είναι απαραίτητο να γίνει μία αναφορά στη χρήση των ακμών μίας εικόνας για την κατάτμησή της. Η ανίχνευση ακμών αποτελεί τοπική μέθοδο επεξεργασίας μίας εικόνας, με στόχο να εντοπίσει τα εικονοστοιχεία μίας γραμμής που μπορεί να θεωρηθεί ακμή, δηλαδή μία γραμμή που οι περιοχές που χωρίζει έχουν μεγάλη διαφορά στη συνάρτηση εικόνας.

Αν μία κατά πλάτος της εικόνας (δηλαδή στον οριζόντιο άξονα) συνάρτηση φωτεινότητας (στην περίπτωση ασπρόμαυρης εικόνας) απεικονιστεί σε μία συνεχή γραφική παράσταση, θα παρατηρήσουμε ότι η μορφή $f(x)$ θα παρουσιάσει μία απότομη μεταβολή έως και να σταθεροποιηθεί σε μία νέα τιμή. Η μέθοδος ανίχνευσης της ακμής αφορά στην προσέγγιση του κεντρικού σημείου x_0 της ακμής, της οποίας η ύπαρξη μαρτυρείται στην απότομη μεταβολή της $f(x)$, και το οποίο βρίσκεται κάπου στο διάστημα στο οποίο η $f(x)$ μεταβάλλεται.

Για την προσέγγισή του κάνουμε χρήση της πρώτης και της δεύτερης παραγώγου της συνάρτησης $f(x)$, δηλαδή $f'(x)$ και $f''(x)$ αντίστοιχα. Στις νέες γραφικές παραστάσεις που θα προκύψουν, το x_0 θα δοθεί είτε ως ακρότατο της γραφική παράστασης της $f'(x)$ είτε ως σημείο διασκελισμού του μηδενός στη γραφική παράσταση της $f''(x)$. Σε κάθε περίπτωση, ακολουθούνται συγκεκριμένα βήματα ενός αλγορίθμου που μας οδηγεί στο αποτέλεσμά μας.

Στην περίπτωση της χρήσης της πρώτης παραγώγου, η γενική ιδέα έγκειται στον εντοπισμό κάθε φορά ενός τοπικού μεγίστου στο μέτρο της παραγώγου της $f(x)$ (και όχι στην παράγωγο της $f(x)$ της ίδιας, καθότι δε μας ενδιαφέρει αν η μεταβολή μαρτυρά μείωση ή αύξηση της $f(x)$). Αυτό ελέγχεται αρχικά με τη συνθήκη $|f'(x_0)| > T$, αν δηλαδή το μέτρο της κλίσης είναι ανώτερο από μία τιμή κατωφλίου, κι έπειτα με την επαλήθευση ότι η τιμή $f'(x_0)$ αποτελεί ακρότατο. Αν οι δύο συνθήκες επαληθεύονται, τότε το σημείο x_0 αποτελεί σημείο της ακμής. Στην πράξη όμως, η παραπάνω μέθοδος συναντάει γρήγορα δυσκολίες ως προς τον ορισμό του κατωφλίου T , καθώς είναι η εφαρμογή αυτή που καλείται να το ορίζει. Αυτό σημαίνει ότι με αύξηση του T , όλο και λιγότερα σημεία μπορούν να αναγνωριστούν ως σημεία της ακμής.

Το παραπάνω εμπόδιο ξεπερνιέται με τη χρήση της κλίσης της δισδιάστατης πλέον συνάρτησης της εικόνας $\nabla f(x, y)$:

$$\nabla f(x, y) = \frac{\partial f(x, y)}{\partial x} \hat{u}_x + \frac{\partial f(x, y)}{\partial y} \hat{u}_y$$

Όπου \hat{u}_x το μοναδιαίο διάνυσμα στη διεύθυνση x και \hat{u}_y το μοναδιαίο διάνυσμα στη διεύθυνση y . Η συνάρτηση $\nabla f(x, y)$ και η κλίση θ αυτής με τον άξονα x είναι συνεχείς συναρτήσεις.

Βάσει γεωμετρίας, το διάνυσμα $\nabla f(x, y)$ είναι κάθετο στην εφαπτομένη κάθε ακμής που μπορεί να εντοπιστεί σε ένα σημείο (x, y) . Ισχύει επομένως:

$$\nabla f(x, y) = G(x, y) = G_R(x, y) \cos \theta + G_C(x, y) \sin \theta$$

Όπου $G_R(x, y) = \frac{\partial f(x, y)}{\partial x}$ και $G_C(x, y) = \frac{\partial f(x, y)}{\partial y}$.

Το πλάτος της $\nabla f(x, y)$ είναι:

$$|G(x, y)| = \sqrt{G_R^2(x, y) + G_C^2(x, y)}$$

Ενώ η γωνία θ είναι:

$$\theta = \arg\left(\frac{\partial f(x, y)}{\partial x}, \frac{\partial f(x, y)}{\partial y}\right) = \tan^{-1}\left(\frac{G_C(x, y)}{G_R(x, y)}\right)$$

Θα μπορούσε να χρησιμοποιηθεί επίσης για τον προσδιορισμό της ακμής, αντί για το πλάτος της κλίσης της $f(x, y)$, το άθροισμα των απόλυτων τιμών των μερικών παραγώγων:

$$|G(x, y)| = |G_R(x, y)| + |G_C(x, y)|$$

Καθώς και η μέγιστη απόλυτη τιμή των μερικών παραγώγων:

$$|G(x, y)| = \max(|G_R(x, y)|, |G_C(x, y)|)$$

Ακολουθείται λοιπόν ο ίδιος αλγόριθμος με πριν, στον οποίο ελέγχεται πλέον η συνθήκη $|\nabla f(x_0, y_0)| > T$, με T τιμή κατωφλίου, η οποία αν ισχύει μας δίνει ένα σημείο της ακμής.

Ένα συνηθισμένο πρόβλημα που συναντάται είναι να είναι τόσα τα σημεία που ξεπερνούν την τιμή κατωφλίου, ώστε αυτά πλέον να κατανέμονται σε λωρίδες αντί για γραμμές. Σε αυτή την περίπτωση, το επόμενο βήμα του αλγορίθμου είναι ο προσδιορισμός λεπτών ακμών μέσα από τις λωρίδες αυτές, μία διαδικασία που ονομάζεται λέπτυνση ακμών (edge thinning). Αυτή υλοποιείται με τον εντοπισμό τοπικών μεγίστων και στις 2 διευθύνσεις της εικόνας. Για να οριστούν αυτά ως σημεία ακμών, λαμβάνει χώρα ο εξής έλεγχος ανάλογα με την περίπτωση:

- Αν το σημείο (x_0, y_0) είναι τοπικό μέγιστο κατά την οριζόντια διεύθυνση αλλά όχι κατά την κάθετη, τότε το σημείο αυτό είναι σημείο ακμής αν $|G_R(x_0, y_0)| > k|G_C(x_0, y_0)|$ (με τυπική τιμή k περίπου 2), δηλαδή ο ρυθμός μεταβολής της έντασης φωτεινότητας στην οριζόντια διεύθυνση είναι σημαντικά μεγαλύτερος από ότι στην κάθετη

- Αν το σημείο (x_0, y_0) είναι τοπικό μέγιστο κατά την κάθετη διεύθυνση αλλά όχι την οριζόντια, τότε το σημείο αυτό είναι σημείο ακμής αν $|G_C(x_0, y_0)| > k|G_R(x_0, y_0)|$, δηλαδή ο ρυθμός μεταβολής της έντασης φωτεινότητας στην κάθετη διεύθυνση είναι σημαντικά μεγαλύτερος από ότι στην οριζόντια.

Η υλοποίηση της κατάτμησης των εικόνων αυτής της εργασίας έχει τον πυρήνα της στη σύνδεση των 2 παραπάνω μεθόδων κατάτμησης, δηλαδή της κατάτμησης με βάση τις ακμές και με βάση το ιστόγραμμα. [6]

2.5: Ενεργά Περιγράμματα/Παραμετρικά Ενεργά Περιγράμματα

Μία μέθοδος με καθοριστική συμβολή στην κατάτμηση εικόνας ήταν τα Γεωδαιτικά Ενεργά Περιγράμματα (Geodesic Active Contours). Προερχόμενα από την κατάτμηση εικόνας με Ενεργά Περιγράμματα, η οποία υλοποιείται βάσει της ελαχιστοποίησης της ενέργειας γύρω από κάθε υποπεριοχή μίας εικόνας, στην οποία προσεγγίζεται και διαχωρίζεται ένα αντικείμενο στη βάση μιας συγκεκριμένης μεταβολής, όπως η φωτεινότητα, η απόχρωση κλπ, προσφέρουν έναν δυναμικό προσδιορισμό στην οριοθέτηση των αντικειμένων, χάρη στην παραμορφωσιμότητα των καμπυλών που το περικλείουν.

Αφετηρία αυτής της σειράς μεθόδων ήταν τα Παραμετρικά Ενεργά Περιγράμματα, όπου το περίγραμμα ενός αντικειμένου περιγράφεται με τη μορφή παραμετρικής καμπύλης σε δύο διαστάσεις:

$$\vec{C}(p) = (x(p), y(p))$$

Όπου $p \in [0,1]$ η παράμετρος της καμπύλης. Η συνάρτηση ενέργειας ορίζεται από την παρακάτω σχέση, την οποία και προσπαθούμε να ελαχιστοποιήσουμε:

$$E = \int_0^1 \{E_{int}[\vec{C}(p)] + E_{ext}[\vec{C}(p)]\} dp$$

Για κάθε τιμή της παραμέτρου p , έχουμε τις συντεταγμένες ενός σημείου στο επίπεδο και άρα ένα σημείο της καμπύλης C . Με \vec{T} και \vec{N} το εφαπτομενικό και κάθετο αντίστοιχα διάνυσμα στο σημείο αυτό και $\kappa(p)$ την καμπυλότητα της καμπύλης, που ορίζεται ως η μεταβολή της γωνίας $\theta(p)$ του \vec{T} κατά μήκος της καμπύλης, δηλαδή $\kappa(p) = \frac{d\theta}{ds}$, έχουμε τις εξής σχέσεις:

$$\vec{T}(p) = \frac{\frac{d}{dp} \vec{C}}{\left\| \frac{d}{dp} \vec{C} \right\|}$$

$$\frac{d}{dp} \vec{T}_p(p) = \kappa \vec{N}$$

$$\frac{d}{dp} \vec{N}_p(p) = -\kappa \vec{T}$$

Θεωρούμε $\vec{C} = \vec{C}(p, t)$ την εξέλιξη της καμπύλης στο χρόνο και υποθέτουμε ότι κάθε σημείο της έχει ταχύτητα 2 συνιστωσών $\vec{V}(V_T, V_N)$, συνιστωσών V_t και V_n , και άρα $\frac{\partial \vec{C}}{\partial t} = V_T \vec{T} + V_N \vec{N}$. Η συνιστώσα V_T συνεισφέρει μόνο στην ανα-παραμετροποίηση της $\vec{C}(p)$ χωρίς να μεταβάλλει τη μορφή της. Για να εξαιρεθούν οι ανα-παραμετροποιήσεις της $\vec{C}(p)$ και να ληφθούν υπ' όψιν μόνο οι μετασχηματισμοί που επιδρούν στο σχήμα της, επιλέγεται η ταχύτητα της παραμόρφωσης να εξαρτάται από την καμπυλότητα κ . Οπότε η εξίσωση κίνησης γίνεται $\frac{\partial \vec{C}}{\partial t} = \vec{V}(\kappa) \vec{N}$. Τέλος, το μέγεθος της ολικής παραμόρφωσης περιγράφεται από τη σχέση $Var(t) = \int_C |\kappa(p, t)| \left| \frac{\partial}{\partial p} \vec{C} \right| dp$, ενώ η ποσότητα $\left| \frac{\partial}{\partial p} \vec{C} \right| = \sqrt{x_p^2 + y_p^2}$ δίνει την ποσότητα της έντασης της παραμετροποίησης της καμπύλης.

Ο εντοπισμός του συνόρου στο οποίο ένα αντικείμενο διαχωρίζεται από το υπόλοιπο περιβάλλον στην εικόνα γίνεται μέσω της κίνησης μίας αρχικά δημιουργηθείσας καμπύλης προς το πραγματικό σύνορο του αντικειμένου, μέσω μίας δύναμης \vec{F} η οποία την έλκει προς αυτό.

2.6: Γεωδαιτικά Ενεργά Περιγράμματα

Για να ξεπεραστούν τα προβλήματα των διαφόρων προσεγγίσεων στα Παραμετρικά Ενεργά Περιγράμματα, με σημαντικότερο την επιρροή της παραμετροποίησης της καμπύλης στην εξέλιξη της. Στην προσπάθεια να προσδιοριστεί μία καμπύλη μόνο βάσει των γεωμετρικών της χαρακτηριστικών, δηλαδή της καμπυλότητας και του μοναδιαίου κάθετου διανύσματος \vec{N} , εξελίχθηκαν οι μέθοδοι των Γεωμετρικών Ενεργών Περιγραμμάτων.

Η κυριότερη προσέγγιση είναι αυτή των Γεωδαιτικών Ενεργών Περιγραμμάτων, η οποία συνοψίζεται ως εξής: Θεωρούμε καμπύλη $\vec{C}(p) = (x(p), y(p))$, με παράμετρο $p \in [0, 1]$ και μήκος καμπύλης $L = \int_0^1 \left\| \frac{d}{dp} \vec{C} \right\| dp$. Αν κ είναι η καμπυλότητα της καμπύλης και \vec{N} το κάθετο σε αυτή μοναδιαίο διάνυσμα, βάσει της παραμόρφωσης $\frac{\partial \vec{C}}{\partial t} = \kappa \vec{N}$, η οποία δεν εξαρτάται από την παραμετροποίηση της καμπύλης, το μήκος της καμπύλης ελαχιστοποιείται όσο το δυνατόν ταχύτερα.

Εισάγουμε τη συνάρτηση βάρους f , ώστε αντί να ελαχιστοποιήσουμε το απλό μήκος καμπύλης $L = \int_0^1 \left\| \frac{d}{dp} \vec{C} \right\| dp$, να ελαχιστοποιήσουμε το γενικευμένο μήκος καμπύλης $L = \int_0^1 \left(\left\| \frac{d}{dp} \vec{C} \right\| f \right) dp$, το οποίο και την περιλαμβάνει. Το γενικευμένο στοιχειώδες μήκος καμπύλης θα είναι $dl_f = \left\| \frac{d}{dp} \vec{C} \right\| f dp$.

Η εξίσωση της κίνησης των Γεωδαιτικών Ενεργών Περιγραμμάτων υπακούει στη σχέση:

$$\frac{\partial \vec{C}}{\partial t} = (f\kappa - \vec{N} \cdot \nabla f) \vec{N}$$

Η ποσότητα ∇f είναι αυτή που έλκει την καμπύλη προς τις ακμές της εικόνας.

2.7: Watershed

Μία ακόμα διαδομένη μέθοδος είναι η λεγόμενη Watershed. Η συλλογιστική αφετηρία αυτής της μεθόδου προκύπτει από την ιδέα ότι πλημμυρίζοντας ένα γεωγραφικό τοπίο, οι περιοχές του μπορούν ξεχωρίσουν εύκολα χάρη στο νερό που συγκεντρώνεται άνισα στην κάθε μία, εξαιτίας του διαφορετικού τους βάθους. Πρακτικά, προκύπτει ως σύνθεση των 3 βασικότερων ιδεών: της ανίχνευσης ακμών, της κατωφλιοποίησης και των της ανάπτυξης περιοχών.

Ανοίγοντας οπές στη θέση του κάθε ελάχιστου της εικόνας, ξεκινούμε την πλημμυρίδα από αυτά (δηλαδή τις πιο σκοτεινές περιοχές) και αφήνουμε τη στάθμη να ανεβαίνει από κάτω προς τα πάνω με σταθερό ρυθμό. Σε κάποια χρονική στιγμή, η στάθμη θα έχει ανέβει τόσο σε κάθε λεκάνη απορροής (ο χώρος που δημιουργείται γύρω από την οπή) που θα είναι κοντά στο να υπερχειλίσει σε μία διπλανή. Όταν αυτό τείνει να συμβεί, ο αλγόριθμος Watershed κατασκευάζει ένα φράγμα πάνω στην ακμή που εμποδίζει οριακά την υπερχειλίση, κι έτσι δημιουργείται ένα κομμάτι του περιγράμματος της περιοχής. Η εικόνα συνεχίζει να πλημμυρίζεται και η ίδια διαδικασία λαμβάνει χώρα σε κάθε ακμή κάθε περιοχής που τείνει να εξαφανιστεί από την υπερχειλίση της πλημμυρίδας προς διπλανή περιοχή.

Συνήθως ο αλγόριθμος Watershed χρησιμοποιείται στο μέτρο της κλίσης της προς κατάτμηση εικόνας και όχι στην εικόνα αυτή καθαυτή, καθώς με βάση τις ακμές, η κατάτμηση ορίζεται από τις περιοχές μονοτονίας του μέτρου της κλίσης της εικόνας. Η υλοποίηση του είναι η παρακάτω:

Συμβολίζουμε $M_1, M_2 \dots M_R$ τις συντεταγμένες των σημείων των ελαχίστων μίας εικόνας $g(x, y)$, καθώς και $C(M_i)$ το σύνολο των συντεταγμένων της αντίστοιχης της οπής λεκάνης απορροής που σχηματίζεται και min και max για το ελάχιστο και το μέγιστο της εικόνας αντίστοιχα. Με $T[n]$ το σύνολο των συντεταγμένων (s, t) για τις οποίες θα ισχύει $g(s, t) < n$, έχουμε:

$$T[n] = \{(s, t) \mid g(s, t) < n\}$$

Δηλαδή, το σύνολο $T[n]$ είναι το σύνολο των σημείων της εικόνας $g(x, y)$ που βρίσκονται κάτω από το επίπεδο $g(x, y) = n$.

Θεωρούμε ότι η τοπογραφία της εικόνας μας πλημμυρίζει με νερό και το ύψος της στάθμης αυξάνεται σταθερά και ακέραια, ξεκινώντας από $n = min + 1$ και φτάνοντας στην τιμή $n = max + 1$. Για να υλοποιηθεί ο αλγόριθμος, πρέπει να

γνωρίζει σε κάθε βήμα της διαδικασίας το πλήθος των σημείων που είναι κάτω από τη στάθμη της πλημμυρίδας, γι αυτό και οι συντεταγμένες $T[n]$ του συνόλου αυτού που βρίσκονται κάτω από το επίπεδο $g(x, y) = n$ ζωγραφίζονται μαύρες. Έτσι, μπορούμε και βλέπουμε σε κάθε βήμα n σε ποια στάθμη βρίσκεται η πλημμύρα.

Με $C_n(M_i)$ το σύνολο των συντεταγμένων μίας λεκάνης απορροής που βρίσκονται υπό την πλημμύρα στο στάδιο n , ισχύει:

$$C_n(M_i) = C(M_i) \cap T[n]$$

Η σχέση αυτή θα ισούται με 1 στη θέση (x, y) μόνο εάν ισχύει ταυτόχρονα $(x, y) \in C(M_i)$ και $(x, y) \in T[n]$. Επιτυγχάνουμε μέσω αυτής της σύζευξης την απομόνωση των σημείων που ανήκουν σε μία λεκάνη απορροής βάσει του αν βρίσκονται υπό τη στάθμη της πλημμυρίδας.

Συμβολίζουμε με $C[n]$ έπειτα την ένωση των πλημμυρισμένων λεκανών απορροής έως και το στάδιο n :

$$C[n] = \bigcup_{i=1}^R C_n(M_i)$$

Στο δε τελικό στάδιο της πλημμυρίδας, η ένωση όλων των λεκανών απορροής θα είναι:

$$C[\max + 1] = \bigcup_{i=1}^R C_n(M_i)$$

Έχοντας ως δεδομένο ότι τα στοιχεία $C_n(M_i)$ ενός συνόλου δεν αλλάζουν ποτέ, παρά μόνο ίσως αυξάνονται σε κάθε βήμα n της πλημμύρας, ισχύει ότι το $C[n - 1]$ είναι υποσύνολο του $C[n]$, το οποίο με τη σειρά του είναι υποσύνολο του $T[n]$ (το σύνολο όλων των σημείων κάτω από τη στάθμη της πλημμυρίδας). Συμπεραίνουμε έτσι, ότι κάθε συνδεδεμένη συνιστώσα του $C[n - 1]$ (ένα συμπαγές σύνολο εικονοστοιχείων που συνδέονται με ένα τυχαίο εικονοστοιχείο εντός ενός συγκεκριμένου συνόλου), ή γεωγραφικά μία λεκάνη απορροής, είναι υποσύνολο της συνδεδεμένης συνιστώσας του συνόλου $T[n]$.

Για την υλοποίηση του αλγορίθμου Watershed, γίνεται η αρχικοποίηση $C[\min + 1] = T[\min + 1]$, και συνεχίζεται με τον υπολογισμό του $C[n]$ αναδρομικά από το $C[n - 1]$. Αν ορίσουμε ως Q το σύνολο των συνδεδεμένων συνιστωσών (οι αλγεβρικές απεικονίσεις των λεκανών απορροής) του συνόλου $T[n]$, υπάρχουν τρία ενδεχόμενα για την κάθε συνδεδεμένη συνιστώσα $q \in Q[n]$:

- Το σύνολο $q \cap C[n - 1]$ είναι κενό.
- Το σύνολο $q \cap C[n - 1]$ περιέχει μόνο μία συνδεδεμένη συνιστώσα του συνόλου $C[n - 1]$.

- Το σύνολο $q \cap C[n - 1]$ περιέχει περισσότερες από μία συνδεδεμένες συνιστώσες του συνόλου $C[n - 1]$.

Η πρώτη συνθήκη ισχύει όταν συναντάται ένα νέο ελάχιστο, οπότε και η συνδεδεμένη συνιστώσα q εντάσσεται στο σύνολο $C[n - 1]$ κι έτσι προκύπτει το σύνολο $C[n]$, στην καινούργια πλέον στάθμη.

Η δεύτερη συνθήκη ισχύει όταν το q βρίσκεται ήδη εντός μίας σχηματισμένης λεκάνης απορροής, οπότε και πάλι εντάσσεται στο σύνολο $C[n - 1]$.

Η τρίτη συνθήκη ισχύει όταν μέσα στο q υπάρχει μία καμπύλη που διαχωρίζει τις υπάρχουσες λεκάνες απορροής, οι οποίες με επιπλέον πλημμύρα θα συγχωνευτούν, οπότε και προκύπτει η ανάγκη κατασκευής φραγμάτων που θα το εμποδίσουν. [7]

Πρακτικά, προκύπτει η ανάγκη οι περιοχές που δημιουργήθηκαν από τον αλγόριθμο Watershed να συγχωνευτούν για περαιτέρω ανάλυση. Συνηθίζεται σε αυτή την περίπτωση η χρήση γράφων για να επιτευχθεί αυτό. Συνοπτικά, έχοντας αντιστοιχίσει κάθε εικονοστοιχείο της εικόνας σε ένα κόμβο ενός γράφου, καθώς και τη σύνδεσή του με τα 4 εικονοστοιχεία γύρω του με τις ακμές του κόμβου αυτού, ορίζουμε ως σχετικότητα των μεταξύ τους εικονοστοιχείων το βάρος της κάθε ακμής (που συνήθως προκύπτει από τη διαφορά πυκνότητας των δύο εικονοστοιχείων που ενώνονται με την ακμή αυτή). Με αυτόν τον τρόπο, κάθε εικονοστοιχείο της κάθε απομονωμένης εικόνας πάνω στην οποία δουλεύουμε, περιλαμβάνει τις πληροφορίες που θα χρησιμοποιηθούν για να προχωρήσει σε ένωση με το σχετικότερο μίας άλλης εικόνας (όπου πάλι θα έχει αντιστοιχηθεί σε κόμβο γράφου). [8], [9]

Κεφάλαιο 3: Η Υιοθετηθείσα Μέθοδος Κατάτμησης Εικόνας

Στην παρούσα διπλωματική, εξετάζουμε την εφαρμογή και την επίδοση μίας πρωτότυπης μεθόδου κατάτμησης έγχρωμης εικόνας. Η βασική της ιδέα εντοπίζεται στην ταξινόμηση των εικονοστοιχείων σε περιοχές που η καθεμία περιγράφεται από μία κανονική κατανομή επί του ιστογράμματος της εικόνας. Η χρησιμότητα αυτής της μεθόδου έγκειται στην ικανότητά της να εντοπίζει σαφή περιγράμματα στις περιοχές μίας εικόνας με κατά τα άλλα ανομοιογένεια στην κατανομή των αποχρώσεων, εξασθένηση των χρωμάτων λόγω της φθοράς του χρόνου, υλικές ζημιές όπως ρωγμές, αποσύνθεση κλπ.

Η μέθοδος αναφοράς πάνω στην οποία βασίζεται η μεθοδολογία κατάτμησης είναι αυτή του άρθρου [10].

3.1: Παρουσίαση της μεθόδου: Ταξινόμηση των εικονοστοιχείων

Ορίζουμε αρχικά το θόρυβο λόγω φθοράς, μία έννοια για την περιγραφή της φθοράς που έχει προκληθεί στα χρώματα της εικόνας λόγω της παλαιότητας του υλικού (fragment decay). Τα βήματα αυτής της μεθόδου για τον υπολογισμό της κατανομής του θορύβου φθοράς είναι τα εξής:

1. Μετατρέπουμε την έγχρωμη εικόνα μας σε ασπρόμαυρης κλίμακας.
2. Υπολογίζουμε την ένταση της κλίσης σε κάθε εικονοστοιχείο (i,j) της εικόνας:

$$\begin{aligned} G_f(i,j) &\equiv \sqrt{(\Delta_x[I(i,j)])^2 + (\Delta_y[I(i,j)])^2} \\ &= \sqrt{(I(i,j+1) - I(i,j))^2 + (I(i+1,j) - I(i,j))^2} \end{aligned}$$

3. Υπολογίζουμε τη μέση τιμή E_G και διασπορά R_G της κλίσης $G_f(i,j)$.
4. Ελέγχουμε τον αριθμό των εικονοστοιχείων N_G εντός του διαστήματος $(E_G - R_G, E_G + R_G)$

Αν ο αριθμός N_G είναι μικρότερος ενός σημαντικού ποσοστού P_g (εν προκειμένω $P_g < 0,8$) του συνολικού αριθμού N εικονοστοιχείων της εικόνας, τότε ακολουθούμε μία ειδική διαδικασία εξομάλυνσης του χρώματος, καθότι θεωρούμε πως ο θόρυβος λόγω της φθοράς είναι τόσοσ που αυτό απαιτείται. Αυτό το επιτυγχάνουμε μεταθέτοντας μία μάσκα (κατά προσέγγιση, ιδανικών διαστάσεων 9×9) στην έγχρωμη εικόνα μας, ώστε κάθε φορά, ένα εικονοστοιχείο της εικόνας να είναι το κέντρο της μάσκας. Έπειτα αποδίδουμε τη μέση τιμή του κόκκινου (R), πράσινου (G)

και μπλε (B) όλων των εικονοστοιχείων της εικόνας στο κεντρικό εικονοστοιχείο, επιτυγχάνοντας έτσι τη χρωματική εξομάλυνση αυτής.

Αν ο αριθμός N_G είναι μεγαλύτερος ή ίσος της τιμής $P_G * N$, τότε δε χρειάζεται να προχωρήσουμε σε εξομάλυνση.

5. Δημιουργούμε ένα χρωματικό ιστόγραμμα και αντλούμε το μέγιστό του με την ακόλουθη διαδικασία:

a) Υπολογίζουμε το ελάχιστο και το μέγιστο κάθε χρωματικής συνιστώσας R, G, B , ως $R_m, R_M, G_m, G_M, B_m, B_M$.

b) Διαιρούμε τα διαστήματα $[R_m, R_M], [G_m, G_M], [B_m, B_M]$ σε p υποδιαστήματα $S_X^i, i = 1, 2, \dots, p, X = R, G, B$, μήκους $l_X = \frac{X_M - X_m}{p}$.

c) Θεωρούμε το καρτεσιανό γινόμενο των μονοδιάστατων αυτών διαστημάτων κι έτσι προκύπτει μία τρισδιάστατη διαμέριση του παραλληλεπιπέδου $[R_m, R_M] \times [G_m, G_M] \times [B_m, B_M]$.

d) Ταξινομούμε τα εικονοστοιχεία της εικόνας μέσα σε αυτά τα τρισδιάστατα διαστήματα με τον εξής τρόπο: έστω ένα εικονοστοιχείο χρώματος R, G, B , τότε αυτό ταξινομείται στο τρισδιάστατο διάστημα $I_{i,j,k}$ αν και μόνον αν $R \in S_R^i, G \in S_G^i$ και $B \in S_B^i$. Πειραματικά, διαπιστώθηκε ότι η τιμή $p = 32$ έδωσε καλά αποτελέσματα για τη συγκεκριμένη εργασία.

6. Μετρούμε τον αριθμό των εικονοστοιχείων $M_{i,j,k}$ που ανήκουν στην αντίστοιχη κυψέλη $I_{i,j,k}$, και εντοπίζουμε το τοπικό μέγιστο της ακολουθίας $M_{i,j,k}, i, j, k = 0, 1, 2 \dots, p$, που αντιστοιχεί όμως σε έναν αριθμό εικονοστοιχείων άνω ενός κατωφλίου $TH_M = 100$. Για κάθε τέτοιο τοπικό μέγιστο (a, b, c) θεωρούμε την «επεκταμένη» κυψέλη:

$$E_{C(i,j,k)} = \bigcup_{i=a-1}^{a+1} \bigcup_{j=b-1}^{b+1} \bigcup_{k=c-1}^{c+1} I_{i,j,k}$$

Μέσω των εικονοστοιχείων που ανήκουν σε αυτήν την «επεκταμένη» κυψέλη, κάνουμε μία πρώτη προσέγγιση μίας περιοχής της εικόνας. Ενδέχεται να λάβουμε μία κλάση περιοχών $E_{C(i,j,k)}$ που αλληλοεπικαλύπτονται, παρόλα αυτά δεν επικαλύπτουν τη συνολική εικόνα.

7. Θεωρώντας ότι σε μία μονοχρωματική περιοχή, οι χρωματικές διαβαθμίσεις είναι τυχαίες μεταβλητές που ακολουθούν κανονική κατανομή, μπορούμε να επεκτείνουμε τις ορισμένες περιοχές μέσω της ακόλουθης διαδικασίας:

- a) Υπολογίζουμε τη μέση τιμή μ_x και τη διασπορά σ_x των χρωματικών συντεταγμένων $x = R, G, B$, στα ορισμένα σύνολα εικονοστοιχείων $E_{C(i,j,k)}$.
- b) Λόγω της κανονικής κατανομής, θεωρούμε ότι το μεγαλύτερο μέρος του πληθυσμού των εικονοστοιχείων ανήκει στο τρισδιάστατο διάστημα:

$$J = [\mu_R - 3,12\sigma_R, \mu_R + 3,12\sigma_R] \times [\mu_G - 3,12\sigma_G, \mu_G + 3,12\sigma_G] \\ \times [\mu_B - 3,12\sigma_B, \mu_B + 3,12\sigma_B]$$

Προκύπτει λοιπόν, ότι αφού τα περισσότερα εικονοστοιχεία που ανήκουν στην περιοχή $E_{C(i,j,k)}$ ανήκουν σε αυτό το διάστημα, τότε έχουν και χρωματικές συντεταγμένες σε ένα αντίστοιχο διάστημα $J_{(i,j,k)}$. Έτσι, επεκτείνουμε την περιοχή $E_{C(i,j,k)}$ σε μία ευρύτερη, στην οποία περιέχονται όλα τα εικονοστοιχεία με συντεταγμένες στο διάστημα $J_{(i,j,k)}$. Με αυτόν τον τρόπο παράγεται μία κλάση ευρύτερων περιοχών $J_{(i,j,k)}$, με ακόμα μεγαλύτερη πιθανότητα αλληλοκάλυψης.

Τελικός στόχος είναι να ταξινομήσουμε κάθε εικονοστοιχείο σε μία μοναδική περιοχή. Για να το επιτύχουμε, εργαζόμαστε ως εξής:

Ορίζουμε την απόσταση d από μία χρωματική περιοχή ενός εικονοστοιχείου P με χρωματικές συντεταγμένες R_P, G_P, B_P , από μία τυχαία περιοχή U με μέσες τιμές χρωματικών συντεταγμένων των εικονοστοιχείων της $\mu_X, X = R, G, B$, ως

$$d_{(p,u)} = |\mu_R - R_P| + |\mu_G - G_P| + |\mu_B - B_P|$$

Αν ένα εικονοστοιχείο ανήκει σε περισσότερες από μία περιοχές U_1, U_2, \dots, U_n , ή αν ο παραπάνω αλγόριθμος δεν κατάφερε να το κατατάξει σε καμία περιοχή, τότε το τοποθετούμε στην περιοχή από την οποία έχει τη μικρότερη απόσταση. Τέλος, αποδίδουμε στο κάθε εικονοστοιχείο της εικόνας μας χρώμα με συντεταγμένες ίσες με της μέσης τιμής $\mu_X, X = R, G, B$ της περιοχής στην οποία αυτό ανήκει.

3.2: Μείωση θορύβου και βελτίωση περιγραμμάτων

Στη συνέχεια, δουλεύουμε πάνω στη μείωση του θορύβου στην εικόνα λόγω φθοράς και στη βελτίωση περιγραμμάτων. Στην «ομογενοποιημένη» εικόνα που προέκυψε με την παραπάνω διαδικασία, μειώνουμε το θόρυβο με την εξής διαδικασία:

Μία μάσκα $m \times n$ διαστάσεων, με m, n περιπτώσεις, ολισθαίνει σε όλη την εικόνα με κέντρο ένα εικονοστοιχείο P αυτής κάθε φορά. Για κάθε θέση της μάσκας, καταγράφουμε τον αριθμό των εικονοστοιχείων που βρίσκονται στην περίμετρό της, τα οποία και εξετάζουμε ως προς το σε ποια περιοχή U από τις προηγουμένως ορισμένες ανήκουν. Με U_M την περιοχή με τα περισσότερα

εικονοστοιχεία περιμέτρου, και N_M τον αριθμό αυτών, εξετάζουμε αν N_M είναι μεγαλύτερος ενός ποσοστού $\Pi_{m,n}$ του αριθμού όλων των εικονοστοιχείων της περιμέτρου. Αν αυτό ισχύει, όλα τα εικονοστοιχεία της μάσκας αποδίδονται στην περιοχή U_M . Αυτή η διαδικασία επαναλαμβάνεται έως ότου δεν έχει μείνει κάποιο εικονοστοιχείο προς απόδοση σε περιοχή. Τέλος, κάθε μία από τις διαστάσεις της μάσκας μειώνεται κατά δύο, και η διαδικασία έρχεται σε πέρας όταν $m = n = 3$.

Παρά την εκτέλεση του παραπάνω αλγορίθμου, είναι δυνατόν να μείνει ακόμα θόρυβος λόγω φθοράς στα σύνορα της κάθε περιοχής. Για να μειωθεί κι αυτός ο θόρυβος, υλοποιούμε την παρακάτω διαδικασία:

Με μία μάσκα και πάλι $m \times n$ διαστάσεων, με m, n περιπτώσεις, ολισθαίνει σε όλη την εικόνα με κέντρο ένα εικονοστοιχείο P αυτής κάθε φορά. Για κάθε θέση της μάσκας, καταγράφουμε τον αριθμό των εικονοστοιχείων που ανήκουν σε κάθε περιοχή. Έστω U_{max} η περιοχή με τα περισσότερα εικονοστοιχεία, με πλήθος N_{max} , και U_{min} η περιοχή με τα λιγότερα, με πλήθος N_{min} . Ελέγχουμε αν ικανοποιούνται οι παρακάτω συνθήκες:

1. Το κέντρο της μάσκας να ανήκει στην περιοχή U_{min} .
2. Το N_{min} να είναι μικρότερο από ένα ποσοστό $\Pi_{\min(m,n)}$ του αριθμού των εικονοστοιχείων της εικόνας.
3. Το N_{max} να είναι μεγαλύτερο από ένα ποσοστό $\Pi_{\max(m,n)}$ του αριθμού των εικονοστοιχείων της εικόνας.

Αν αυτό συμβαίνει, αποδίδουμε το κεντρικό εικονοστοιχείο της μάσκας στην περιοχή U_{max} . Αυτή η διαδικασία επαναλαμβάνεται έως ότου δεν έχει μείνει κάποιο εικονοστοιχείο προς απόδοση σε περιοχή. Τέλος, κάθε μία από τις διαστάσεις της μάσκας μειώνεται κατά δύο, και η διαδικασία έρχεται σε πέρας όταν $m = n = 3$.

Στην περίπτωση που επιθυμούμε περισσότερη «ομογενοποίηση» της εικόνας μας, επαναλαμβάνουμε την ίδια διαδικασία σε κάθε μία από τις περιοχές που αντήσαμε ξεχωριστά, αλλά με χρήση του αρχικού περιεχομένου του χρώματος των εικονοστοιχείων που εξετάζουμε. Έτσι, η περιοχή που είχε προκύψει ως αποτέλεσμα της παραπάνω διαδικασίας, υποβάλλεται εκ νέου σε αυτήν, με στόχο την παραγωγή περισσότερων περιοχών.

Δεδομένου ότι θέλουμε σε αυτή την εργασία το τελικό μας αποτέλεσμα να είναι μία δίτιμη (binary) ασπρόμαυρη εικόνα, αρχικά μετατρέπουμε την κάθε εικόνα σε εικόνα διαβάθμισης του γκρι (grayscale), και έπειτα εφαρμόζουμε την παραπάνω μεθοδολογία σε μονοδιάστατη διαμέριση. Με αυτόν τον τρόπο, έχουμε τελικώς μόνο 2 περιοχές αποχρώσεων του γκρι, στις οποίες αποδίδουμε μία τιμή (0 ή 1) αναλόγως του κατωφλίου που ορίζουμε, και εν τέλει παράγεται η ασπρόμαυρη εικόνα.

3.3: Υλοποίηση της μεθόδου: Εξάλειψη θορύβου με εφαρμογή μορφολογικών φίλτρων

Ως επέκταση της μεθόδου που παρουσιάσαμε, θα χρησιμοποιήσουμε την παρακάτω μέθοδο για το διαχωρισμό της κάθε εικόνας σε 2 θεματικές περιοχές, αυτές του υποβάθρου της εικόνας (πάπυρος) και αυτές του κύριου θέματος (γράμμα). Η διαδικασία αυτή λαμβάνει χώρα σε ένα μεγάλο πλήθος αποκομμένων εικόνων από τη συνολική σαρωμένη σελίδα των χειρόγραφων. Κάθε αποκομμένη εικόνα περιλαμβάνει ένα και μόνο γράμμα, το οποίο επιδιώκουμε να απομονώσουμε από το υπόβαθρό του. Σε αυτή τη διαδικασία, όπως αναπτύχθηκε από την ερευνητική ομάδα του κ. Παπαοδυσσέα, η έγχρωμη εικόνα μας μετατρέπεται σε εικόνα κλίμακας του γκρι, με τρόπο τέτοιο ώστε οι στοιχειώδεις χρωματικές διαφοροποιήσεις της έγχρωμης εικόνας και οι αντίστοιχες διαφοροποιήσεις της απόχρωσης του γκρι να είναι ισοδύναμες. Αυτό επιτεύχθη με την αναπαράσταση της χρωματικής διαφοροποίησης της έγχρωμης εικόνας στο χρωματικό χώρο Lab, όπου οι χρωματικές μεταβολές λογίζονται Ευκλείδειες. Με την ανάλυση που έγινε από την ερευνητική ομάδα, και παρατίθεται στην παρ. 3.4, αυτές οι στοιχειώδεις Ευκλείδειες χρωματικές μεταβολές μπορούν να αναπαρασταθούν με ένα μόνο διανυσματικό πεδίο, η ολοκλήρωση του οποίου δίνει την «ισοδύναμη» grayscale εικόνα. Υπό τη μετατροπή αυτή, η υιοθετηθείσα μεθοδολογία κατάτμησης έγχρωμων εικόνων που περιγράφηκε στις προηγούμενες παραγράφους, απλοποιείται σε μεθοδολογία μονοδιάστατης κατάτμησης του ιστογράμματος της ισοδύναμης Grayscale εικόνας.

Η μέθοδός μας ξεκινάει με την εφαρμογή μορφολογικών φίλτρων με σκοπό την εξάλειψη του θορύβου. Θεωρούμε αρχικά μία περιοχή με κέντρο ένα τυχαίο σημείο M της εικόνας και του αποδίδουμε την ελάχιστη χρωματική απόχρωση των εικονοστοιχείων αυτής της περιοχής. Στη συνέχεια, επαναλαμβάνουμε τη σάρωση της πλέον μικρυσμένης εικόνας, αλλά αυτή τη φορά αποδίδουμε στο κέντρο κάθε περιοχής M τη μέγιστη χρωματική απόχρωση των εικονοστοιχείων αυτής της περιοχής, καταλήγοντας έτσι στην ανοιχτή εκδοχή της εικόνας (opened image). Στη συνέχεια, με κέντρο κάθε σημείο M της εικόνας και ακτίνα ένα εικονοστοιχείο γύρω από αυτό, αποδίδουμε στο M τη μέγιστη απόχρωση αυτής της περιοχής, μία διαδικασία που ονομάζεται Dilation. Αυξάνουμε συνεχώς την ακτίνα έως ότου υπάρξει μία περιοχή με εικονοστοιχείο με απόχρωση μεγαλύτερη αυτής του αρχικού μας M . Παράγεται έτσι, εν τέλει, μία ομαλοποιημένη εκδοχή της εικόνας για τα επόμενα βήματα της μεθόδου.

3.4: Κλίση εικόνας στο χρωματικό χώρο LAB και υπολογισμός ισοδύναμης grayscale εικόνας

Στο επόμενο βήμα, υπολογίζουμε το μέτρο της κλίσης της εικόνας μας στο χρωματικό χώρο LAB. Δεδομένου ότι η χρωματική απόσταση στο χώρο LAB είναι ευκλείδεια, όπως αυτή στον grayscale χώρο, μπορεί να υποκατασταθεί η τοπική μεταβολή των συνιστωσών της LAB αναπαράστασης από μία χρωματικά ισοδύναμη τοπική μεταβολή. Αυτή με τη σειρά της μπορεί να αντιστοιχηθεί στην κλίση (gradient) της υποθετικής αντίστοιχης grayscale

εικόνας. Χρησιμοποιούμε λοιπόν την ισοδύναμη του ιστογράμματος κατανομή για να εντοπίσουμε τις περιοχές της εικόνας μας. Η κατανομή αυτή προκύπτει από την παρακάτω διαδικασία:

Ορίζουμε το διάνυσμα της αναπαράστασης LAB μίας εικόνας ως εξής:

$$F(x, y) = \begin{bmatrix} L(x, y) \\ a(x, y) \\ b(x, y) \end{bmatrix}$$

Ορίζουμε επίσης τον τελεστή:

$$\partial_r \equiv \begin{bmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \end{bmatrix}$$

Η δράση του τελεστή ∂_r στο διάνυσμα F θα είναι:

$$\partial_r F^T = \begin{bmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \end{bmatrix} [L \quad a \quad b] = \begin{bmatrix} L_x & a_x & b_x \\ L_y & a_y & b_y \end{bmatrix}$$

Θεωρούμε ότι η στοιχειώδης χρωματική διαφορά δίνεται από τη διαφορική ποσότητα $\|dF\|$, οπότε και ισχύει λοιπόν $\|dF\|^2 = dr^T \partial_r F^T (\partial_r F^T)^T dr$. Η λογική της διαδικασίας αυτής έγκειται στον εντοπισμό μίας grayscale εικόνας $f(x, y)$, τέτοιας που να ικανοποιεί τη σχέση $dr^T \partial_r F^T (\partial_r F^T)^T dr = dr^T \nabla f (\nabla f)^T dr$, ώστε ο εντοπισμός των ακμών της έγχρωμης εικόνας μας να ταυτίζεται ως διαδικασία με τον εντοπισμό των ακμών της $f(x, y)$. Ο πυρήνας του τετραγωνικού τύπου $dr^T \partial_r F^T (\partial_r F^T)^T dr$ θα είναι ο συμμετρικός θετικά ορισμένος πίνακας $\partial_r F^T (\partial_r F^T)^T$, ο οποίος μέσω του Spectral Theorem γράφεται με μοναδικό τρόπο $\partial_r F^T (\partial_r F^T)^T = \lambda_1^2 v_1 v_1^T + \lambda_2^2 v_2 v_2^T$. Ισοδύναμα, το διανυσματικό πεδίο που παράγει τη διαφορική μεταβολή $\|dF\|^2$ είναι το $\lambda_1 v_1 + \lambda_2 v_2$, το οποίο είναι και το ζητούμενο ∇f . Ορίζουμε:

$$\nabla f = \lambda_1 v_1 + \lambda_2 v_2$$

Όπου $(\lambda_1^2, \lambda_2^2)$ ιδιοτιμές και (v_1, v_2) ιδιοδιανύσματα του πίνακα $(\partial_r F^T (\partial_r F^T)^T)$. Αυτά υπολογίζονται αναλυτικά από τη σχέση $\partial_r F^T (\partial_r F^T)^T = \begin{bmatrix} \|F_x\|^2 & F_x \cdot F_y \\ F_x \cdot F_y & \|F_y\|^2 \end{bmatrix}$ ως εξής:

$$\lambda_1^2, \lambda_2^2 = \frac{1}{2} \left(\|F_x\|^2 + \|F_y\|^2 \pm \sqrt{\left((\|F_x\|^2 - \|F_y\|^2)^2 + 4(F_x \cdot F_y)^2 \right)} \right)$$

$$v_1 = \begin{bmatrix} \cos \theta \\ \sin \theta \end{bmatrix}, \quad v_2 = \begin{bmatrix} -\sin \theta \\ \cos \theta \end{bmatrix}, \quad \theta = \frac{1}{2} \operatorname{atan2} \left(2F_x \cdot F_y, \|F_x\|^2 - \|F_y\|^2 \right)$$

Ενώ θα ισχύει:

$$\|\nabla f\|^2 = \lambda_1^2 + \lambda_2^2 = \|F_x\|^2 + \|F_y\|^2$$

Για τη λύση της διαφορικής εξίσωσης $\nabla f = \lambda_1 v_1 + \lambda_2 v_2$ χρησιμοποιούμε τον Poisson Solver με βάση το μετασχηματισμό wavelet, όπως προτείνεται στις αναφορές [11], [12], [13], έτσι ώστε να επιλυθεί αριθμητικά η διαφορική εξίσωση:

$$\nabla^2 f = \operatorname{div}(\lambda_1 v_1 + \lambda_2 v_2)$$

Η νέα grayscale εικόνα f που προκύπτει είναι αυτή στην οποία θα κάνουμε την ανάλυση ιστογράμματος ώστε να επιτύχουμε τη δυαδική κατάτμησή της.

3.5: Ενδεικτικές απεικονίσεις του εντοπισμού κλίσης εικόνας στο χρωματικό χώρο LAB και παραγωγής της ισοδύναμου grayscale εικόνας

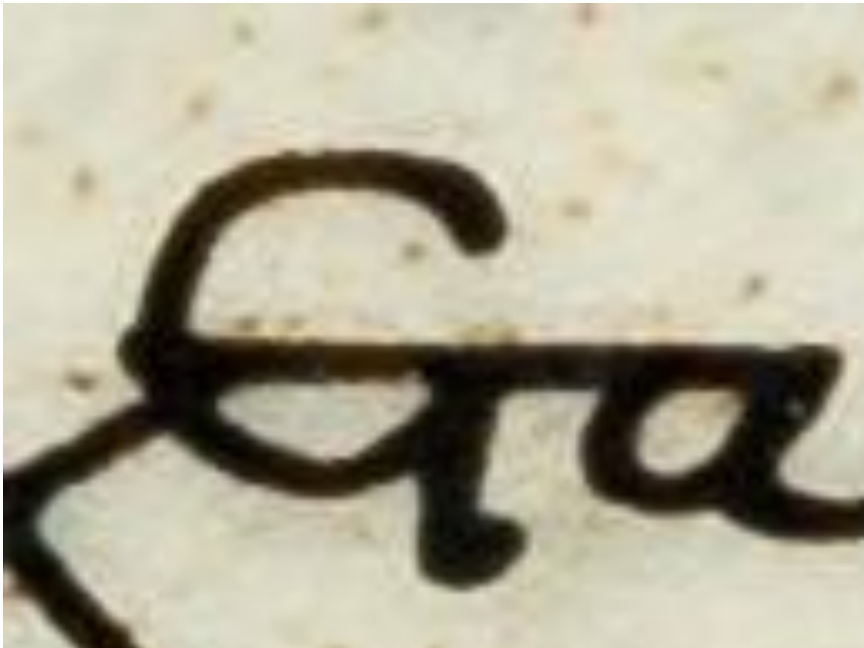
Σε αυτό το σημείο θα παρουσιάσουμε ορισμένα αντιπροσωπευτικά αποτελέσματα του αλγορίθμου της παραπάνω διαδικασίας, ο οποίος επιλέχθηκε να υλοποιηθεί σε περιβάλλον MATLAB, κυρίως για την ταχύτητα που αυτό προσφέρει, καθώς και για τα ποικίλα εργαλεία (μαθηματικές συναρτήσεις, διαδραστικά εργαλεία, γραφικές διεπαφές κλπ) του.

Βασικός παράγοντας ο οποίος ρυθμίστηκε κατά την συγγραφή του κώδικα στο εργαστήριο του κυρίου Παπαοδυσσέα, είναι ο βαθμός της πολυωνυμικής προσαρμογής στην επεξεργασία του ιστογράμματος της εικόνας. Προσεγγίστηκε ο ελάχιστος δυνατός βαθμός πολυωνύμου, από τον οποίο θα πήγαζε μία ανώτερης ποιότητας κατάτμηση.

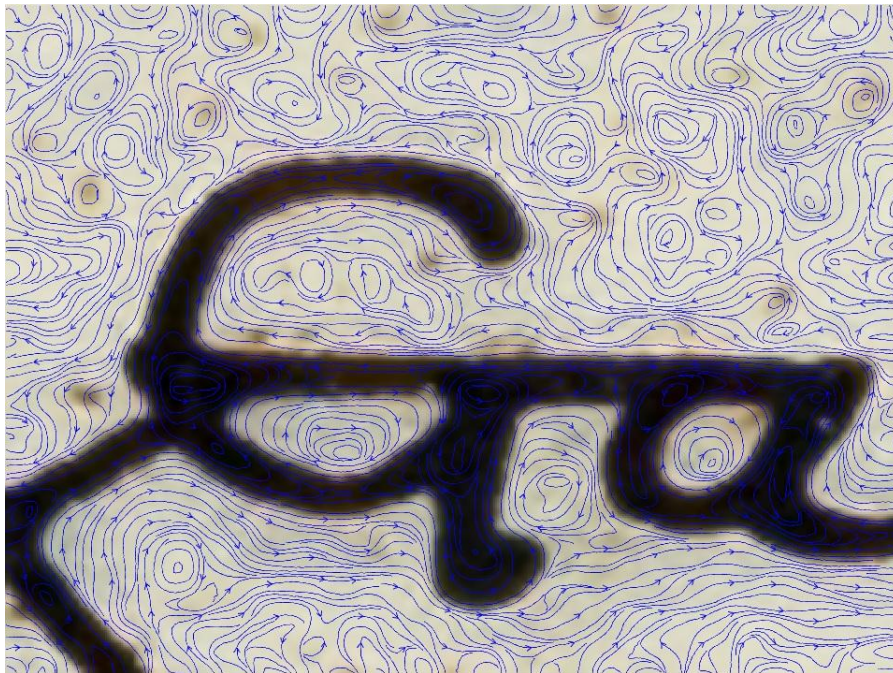
Παρακάτω παρουσιάζουμε μερικά παραδείγματα υλοποίησης αυτής της διαδικασίας. Με τη σειρά απεικονίζονται για κάθε γράμμα:

- Η πηγαία RGB εικόνα, δηλαδή η εικόνα του γράμματος σε ένα κομμάτι παπύρου.
- Η ίδια εικόνα με εκτυπωμένη τη ροή που προκύπτει από την επικαμπύλια ολοκλήρωση του πεδίου των κλίσεων που παρήχθη με βάση την LAB αναπαράσταση της εικόνας και την μέθοδο της προηγούμενης παραγράφου.
- Η grayscale εικόνα που παράγεται από την επίλυση της εξίσωσης Poisson που ορίζεται με βάση το προσδιορισθέν πεδίο των κλίσεων.

1ο παράδειγμα:



Εικόνα 3.5-1: Λεπτομέρεια ενός συμπλέγματος των χαρακτήρων «Ετα» σε RGB

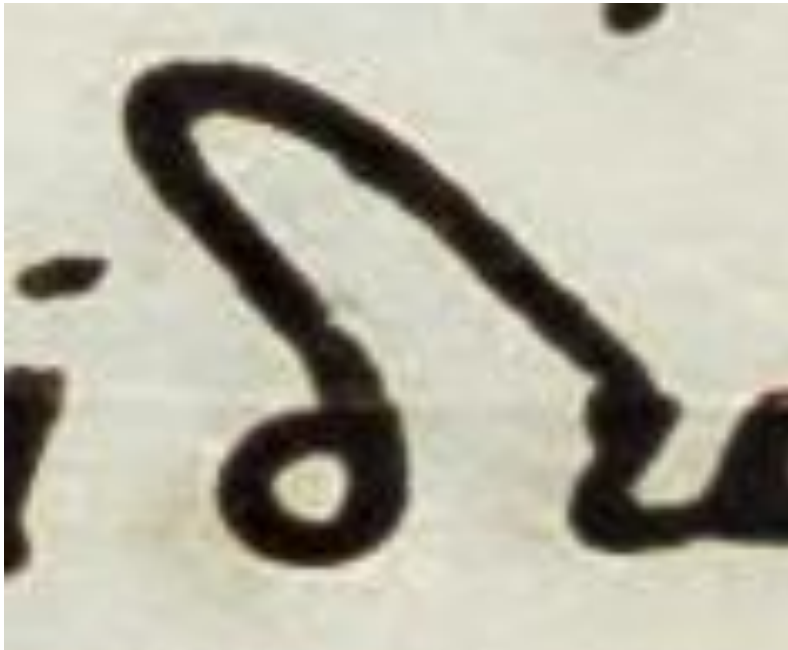


Εικόνα 3.5-2: Αποτύπωση της ροής που επάγει το πεδίο των κλίσεων, το οποίο υπολογίστηκε ως ισοδύναμο των τοπικών χρωματικών μεταβολών της εικόνας σύμφωνα με την LAB αναπαράστασή της.



Εικόνα 3.5-3: Η grayscale εικόνα που αντιστοιχεί στο υπολογισθέν πεδίο κλίσεων, όπως υπολογίστηκε με αριθμητική επίλυση κατάλληλα ορισμένου προβλήματος συνοριακών τιμών και όπως περιγράφηκε στην παράγραφο 3.4.

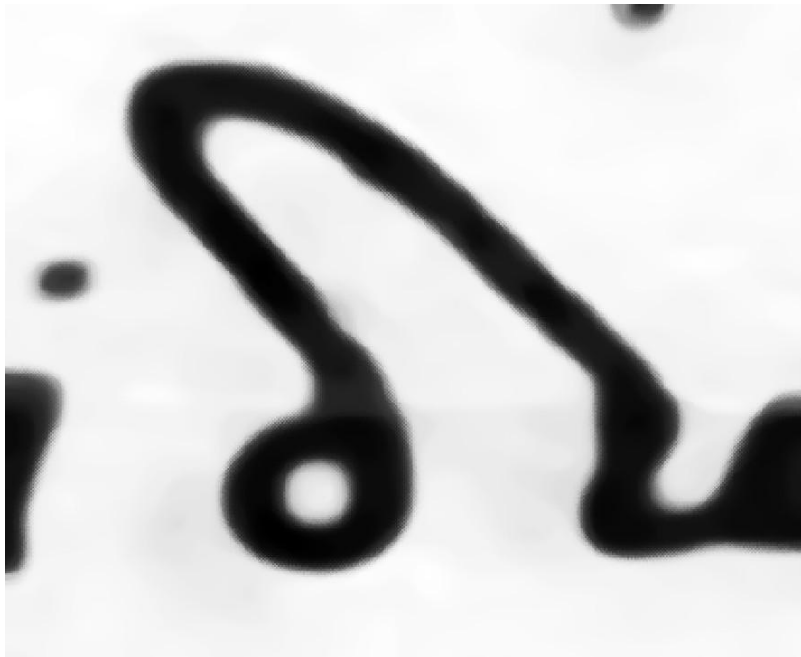
2ο παράδειγμα:



Εικόνα 3.5-4: Λεπτομέρεια του χαρακτήρα «δ» σε RGB

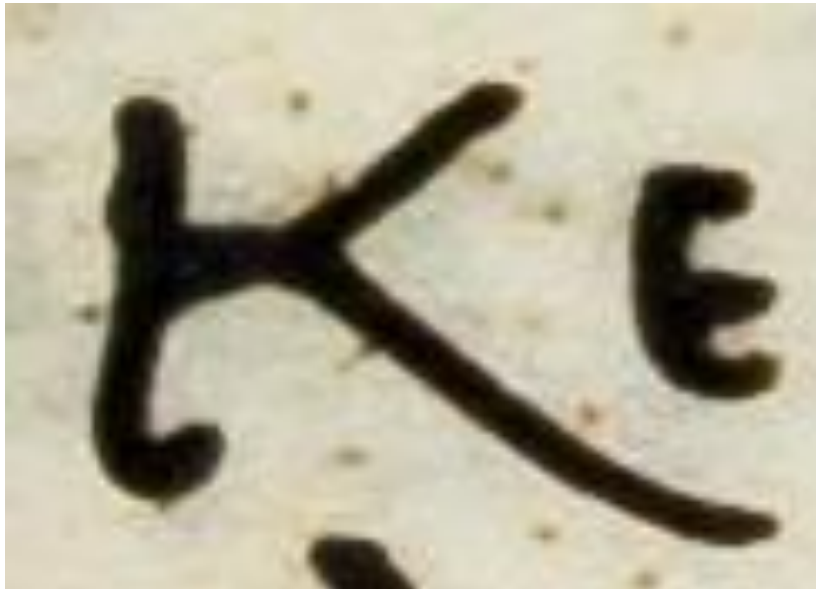


Εικόνα 3.5-5: Αποτύπωση της ροής που επάγει το πεδίο των κλίσεων, το οποίο υπολογίστηκε ως ισοδύναμο των τοπικών χρωματικών μεταβολών της εικόνας σύμφωνα με την LAB αναπαράστασή της.

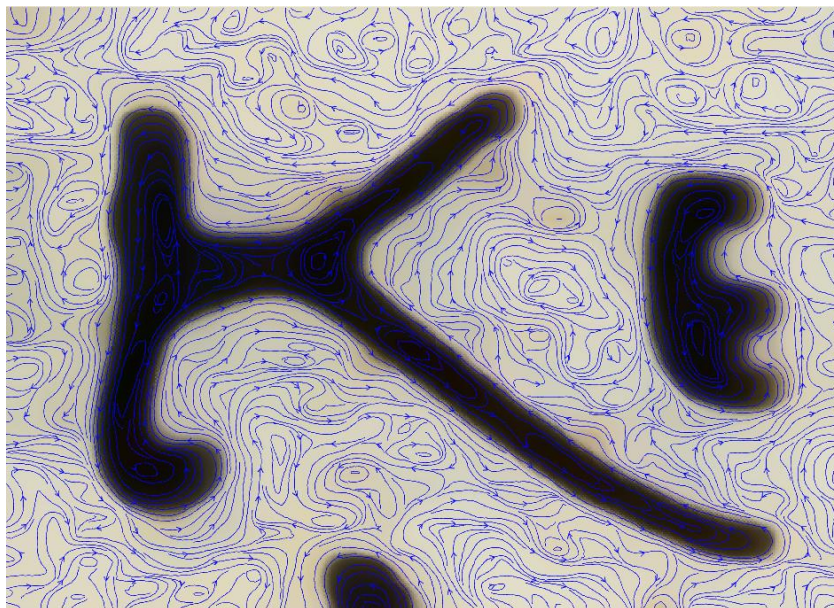


Εικόνα 3.5-6: Η grayscale εικόνα που αντιστοιχεί στο υπολογισθέν πεδίο κλίσεων, όπως υπολογίστηκε με αριθμητική επίλυση κατάλληλα ορισμένου προβλήματος συνοριακών τιμών και όπως περιγράφηκε στην παράγραφο 3.4.

3^ο παράδειγμα:



Εικόνα 3.5-7: Λεπτομέρεια των χαρακτήρων «Κ» και «ε» σε RGB

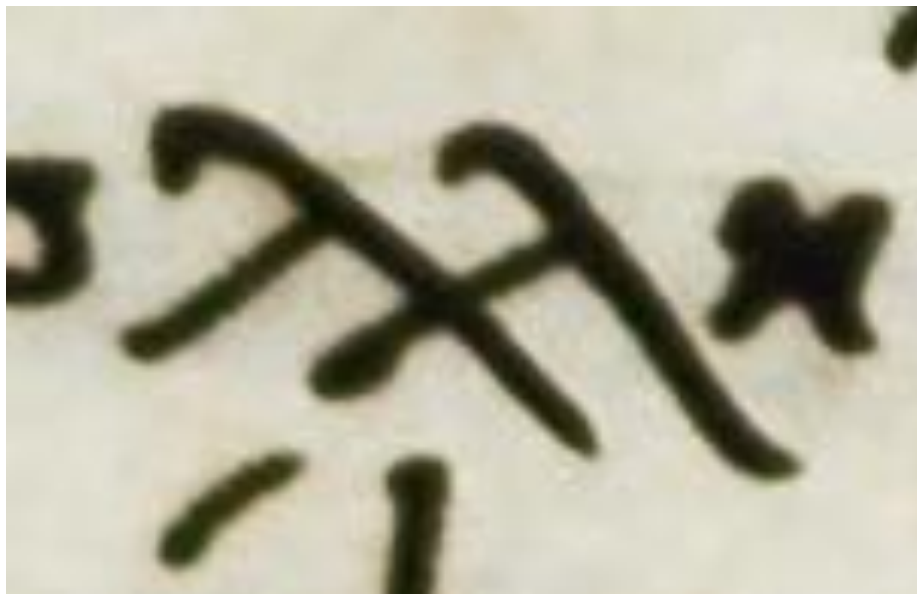


Εικόνα 3.5-8: Αποτύπωση της ροής που επάγει το πεδίο των κλίσεων, το οποίο υπολογίστηκε ως ισοδύναμο των τοπικών χρωματικών μεταβολών της εικόνας σύμφωνα με την LAB αναπαράστασή της.

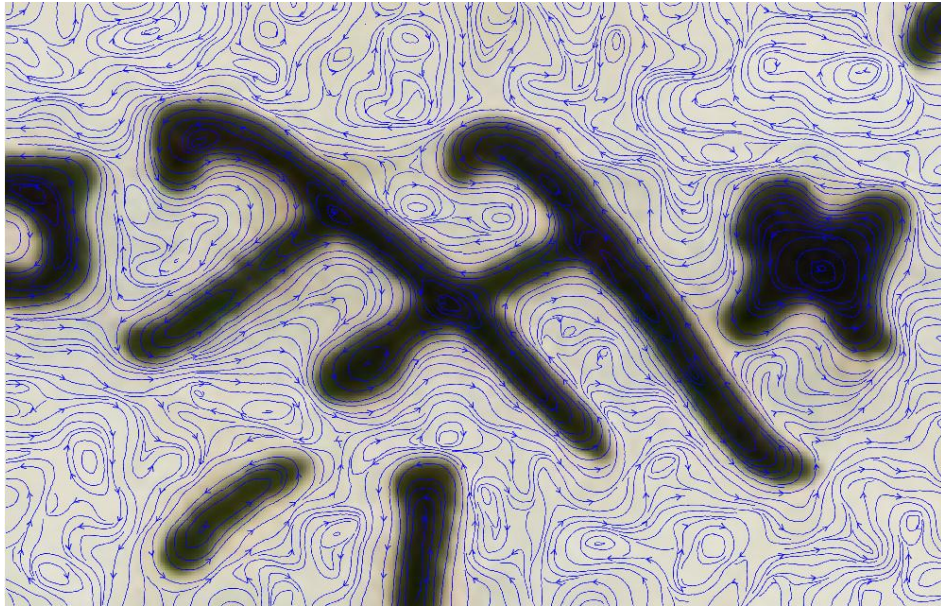


Εικόνα 3.5-9: Η grayscale εικόνα που αντιστοιχεί στο υπολογισθέν πεδίο κλίσεων, όπως υπολογίστηκε με αριθμητική επίλυση κατάλληλα ορισμένου προβλήματος συνοριακών τιμών και όπως περιγράφηκε στην παράγραφο 3.4.

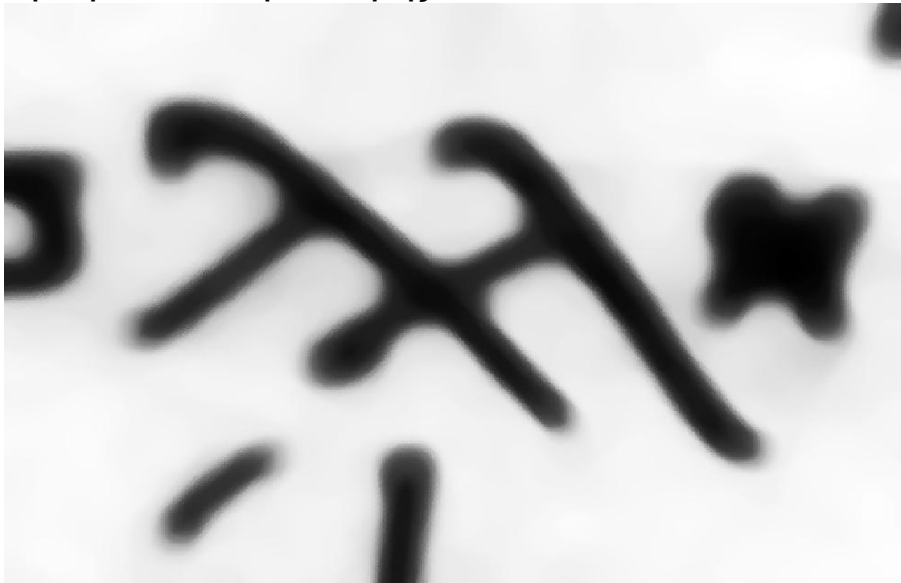
4ο παράδειγμα:



Εικόνα 3.5-7: Λεπτομέρεια του συμπλέγματος χαρακτήρων «λλ» σε RGB

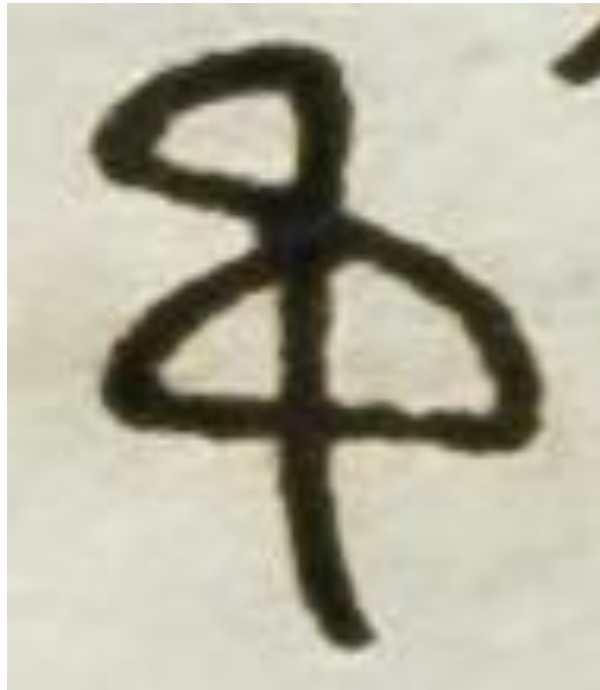


Εικόνα 3.5-11: Αποτύπωση της ροής που επάγει το πεδίο των κλίσεων, το οποίο υπολογίστηκε ως ισοδύναμο των τοπικών χρωματικών μεταβολών της εικόνας σύμφωνα με την LAB αναπαράστασή της.

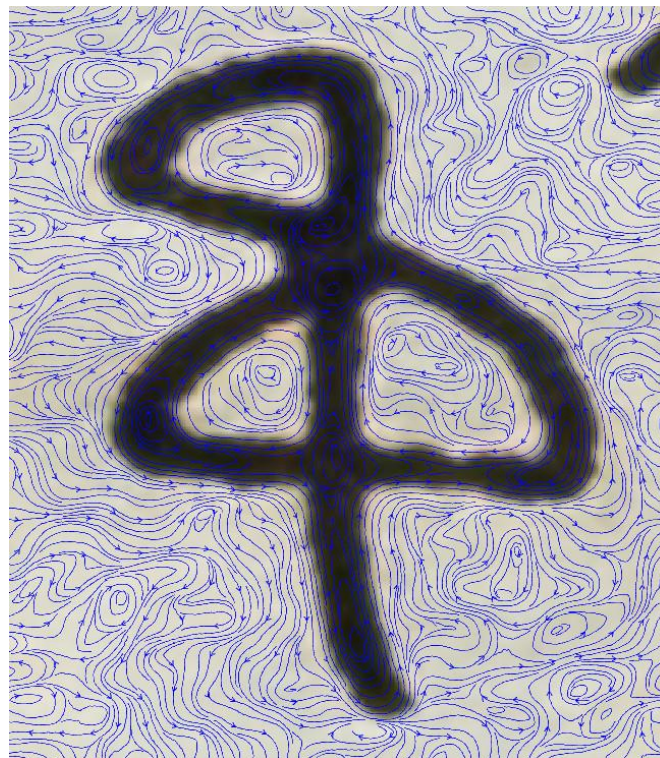


Εικόνα 3.5-12: Η grayscale εικόνα που αντιστοιχεί στο υπολογισθέν πεδίο κλίσεων, όπως υπολογίστηκε με αριθμητική επίλυση κατάλληλα ορισμένου προβλήματος συνοριακών τιμών και όπως περιγράφηκε στην παράγραφο 3.4.

5ο παράδειγμα:



Εικόνα 3.5-13: Λεπτομέρεια του χαρακτήρα «Φ» σε RGB

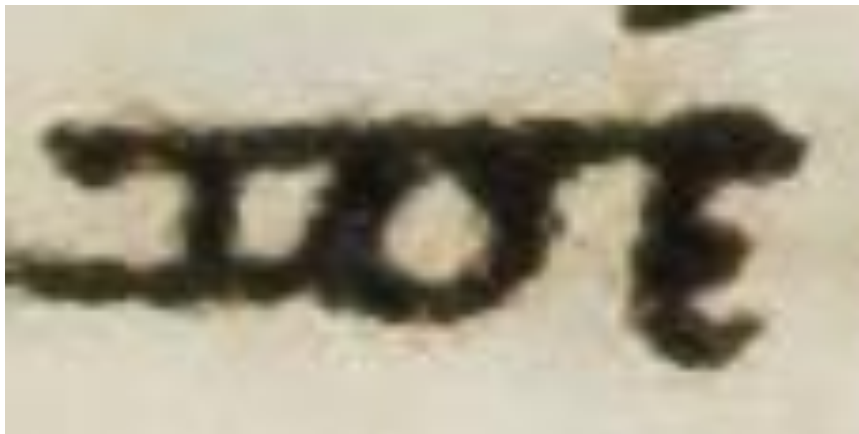


Εικόνα 3.5-14: Αποτύπωση της ροής που επάγει το πεδίο των κλίσεων, το οποίο υπολογίστηκε ως ισοδύναμο των τοπικών χρωματικών μεταβολών της εικόνας σύμφωνα με την LAB αναπαράστασή της.

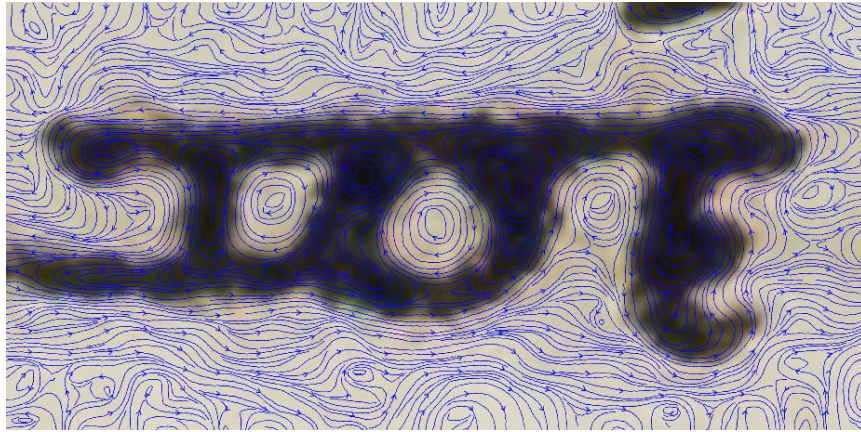


Εικόνα 3.5-15: Η grayscale εικόνα που αντιστοιχεί στο υπολογισθέν πεδίο κλίσεων, όπως υπολογίστηκε με αριθμητική επίλυση κατάλληλα ορισμένου προβλήματος συνοριακών τιμών και όπως περιγράφηκε στην παράγραφο 3.4.

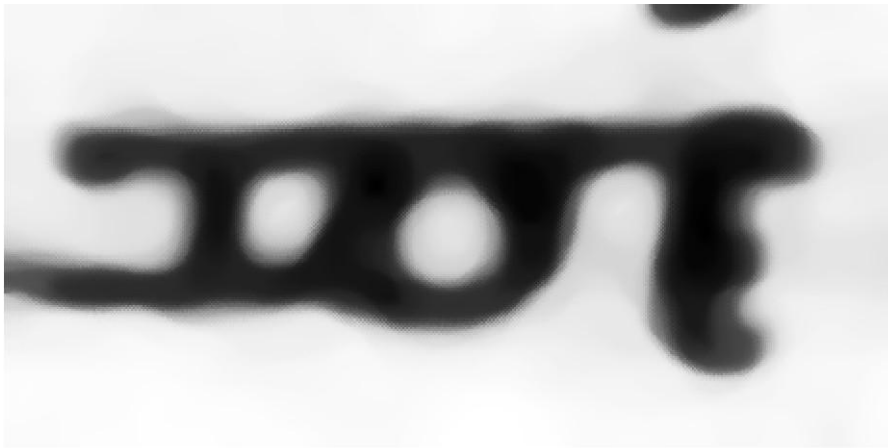
6ο παράδειγμα:



Εικόνα 3.5-16: Λεπτομέρεια του συμπλέγματος χαρακτήρων «πε» σε RGB



Εικόνα 3.5-17: Αποτύπωση της ροής που επάγει το πεδίο των κλίσεων, το οποίο υπολογίστηκε ως ισοδύναμο των τοπικών χρωματικών μεταβολών της εικόνας σύμφωνα με την LAB αναπαράστασή της.

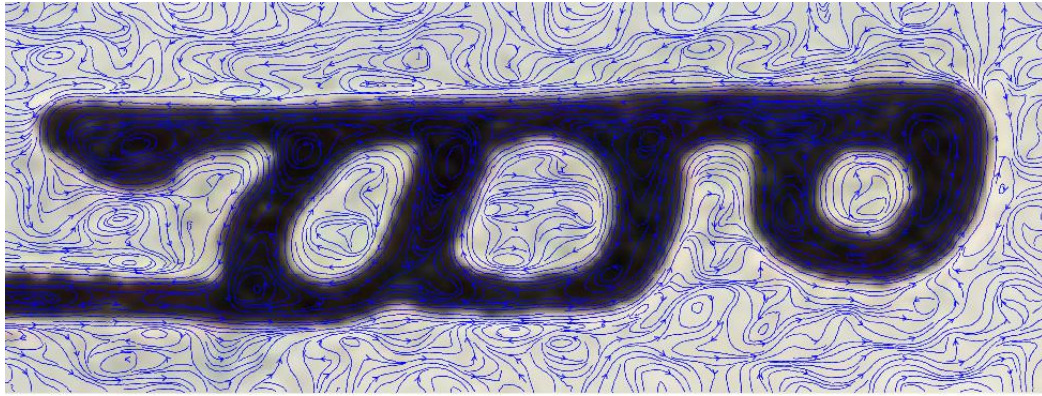


Εικόνα 3.5-18: Η grayscale εικόνα που αντιστοιχεί στο υπολογισθέν πεδίο κλίσεων, όπως υπολογίστηκε με αριθμητική επίλυση κατάλληλα ορισμένου προβλήματος συνοριακών τιμών και όπως περιγράφηκε στην παράγραφο 3.4.

7ο παράδειγμα:



Εικόνα 3.5-19: Λεπτομέρεια του συμπλέγματος χαρακτήρων «πο» σε RGB

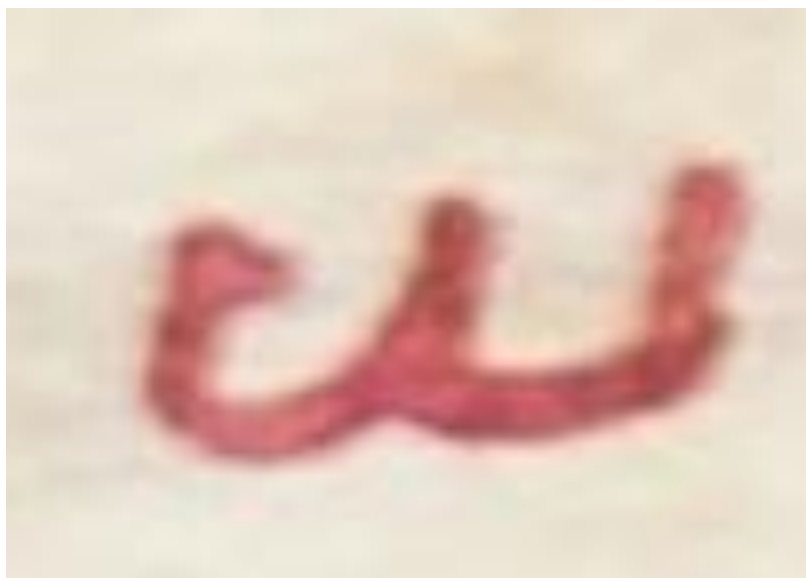


Εικόνα 3.5-20: Αποτύπωση της ροής που επάγει το πεδίο των κλίσεων, το οποίο υπολογίστηκε ως ισοδύναμο των τοπικών χρωματικών μεταβολών της εικόνας σύμφωνα με την LAB αναπαράστασή της.

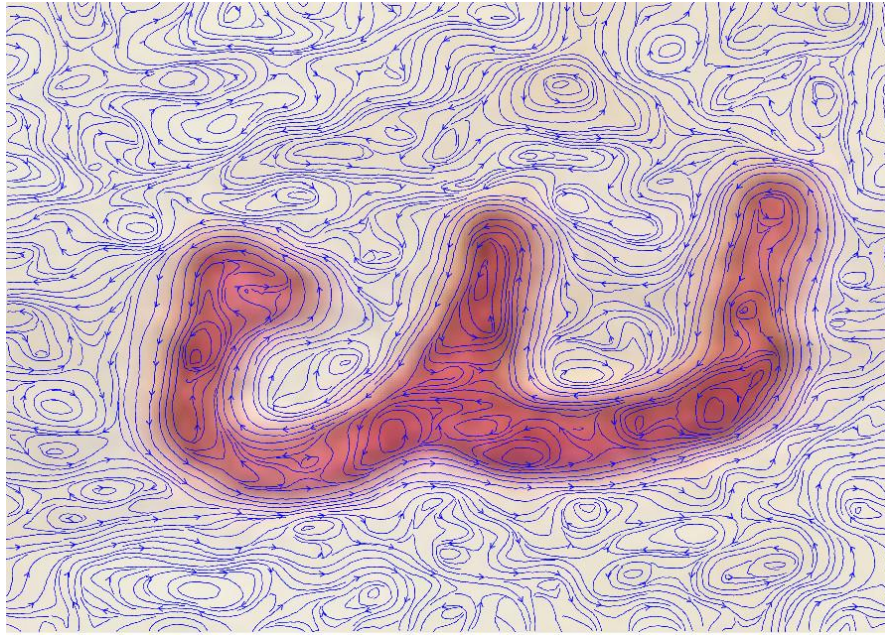


Εικόνα 3.5-21: Η grayscale εικόνα που αντιστοιχεί στο υπολογισθέν πεδίο κλίσεων, όπως υπολογίστηκε με αριθμητική επίλυση κατάλληλα ορισμένου προβλήματος συνοριακών τιμών και όπως περιγράφηκε στην παράγραφο 3.4.

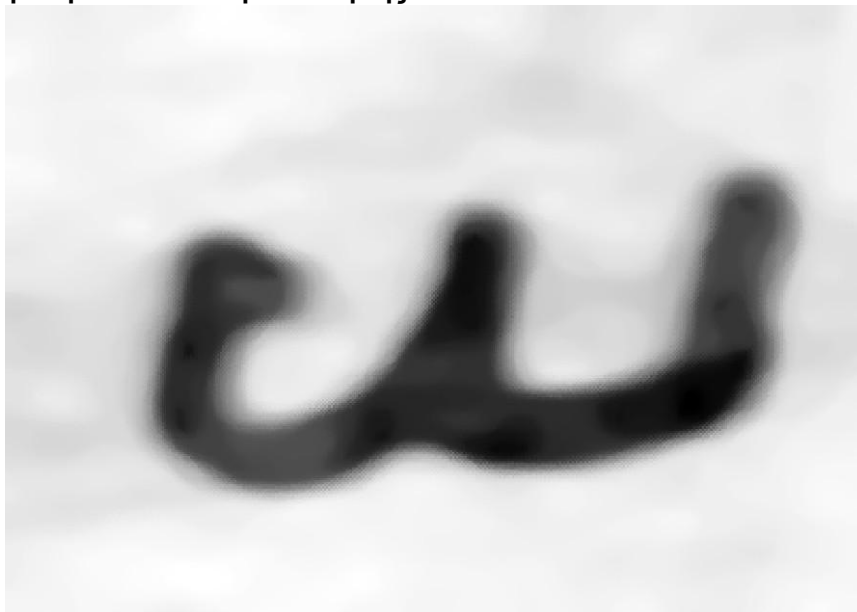
8ο παράδειγμα:



Εικόνα 3.5-22: Λεπτομέρεια του χαρακτήρα «ω» σε RGB



Εικόνα 3.5-23: Αποτύπωση της ροής που επάγει το πεδίο των κλίσεων, το οποίο υπολογίστηκε ως ισοδύναμο των τοπικών χρωματικών μεταβολών της εικόνας σύμφωνα με την LAB αναπαράστασή της.

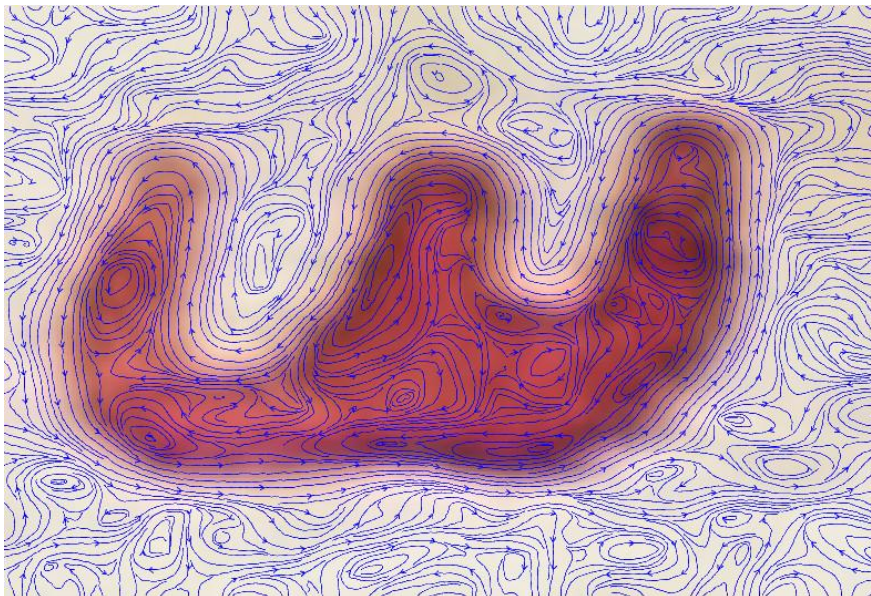


Εικόνα 3.5-24: Η grayscale εικόνα που αντιστοιχεί στο υπολογισθέν πεδίο κλίσεων, όπως υπολογίστηκε με αριθμητική επίλυση κατάλληλα ορισμένου προβλήματος συνοριακών τιμών και όπως περιγράφηκε στην παράγραφο 3.4.

9ο παράδειγμα:



Εικόνα 3.5-25: Λεπτομέρεια του χαρακτήρα «ω» σε RGB

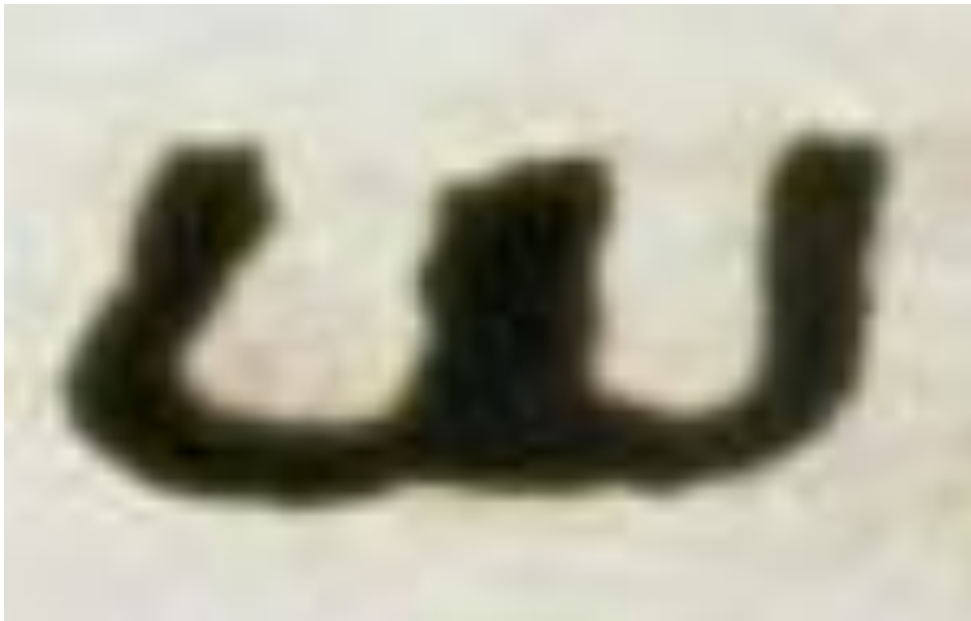


Εικόνα 3.5-26: Αποτύπωση της ροής που επάγει το πεδίο των κλίσεων, το οποίο υπολογίστηκε ως ισοδύναμο των τοπικών χρωματικών μεταβολών της εικόνας σύμφωνα με την LAB αναπαράστασή της.

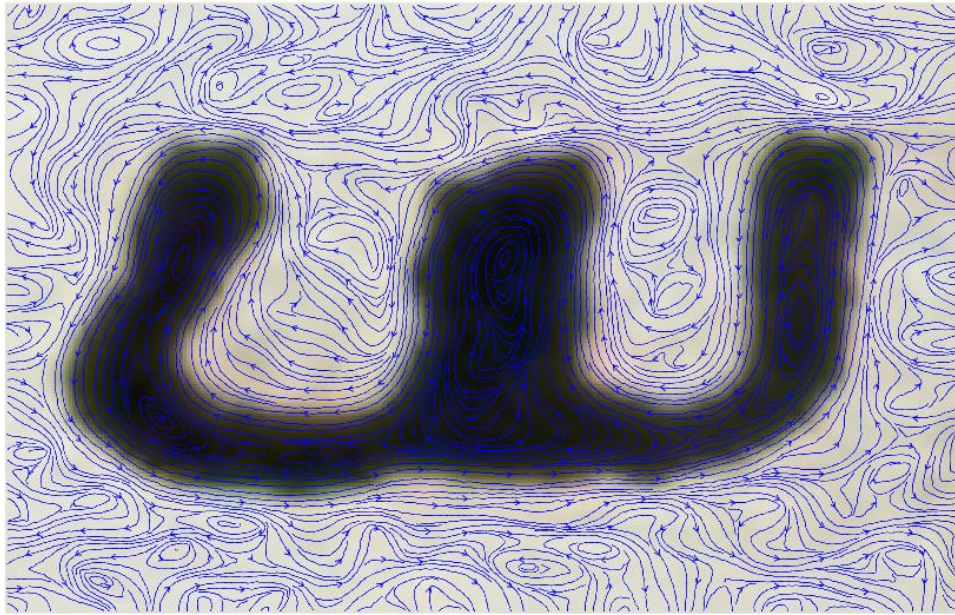


Εικόνα 3.5-27: Η grayscale εικόνα που αντιστοιχεί στο υπολογισθέν πεδίο κλίσεων, όπως υπολογίστηκε με αριθμητική επίλυση κατάλληλα ορισμένου προβλήματος συνοριακών τιμών και όπως περιγράφηκε στην παράγραφο 3.4.

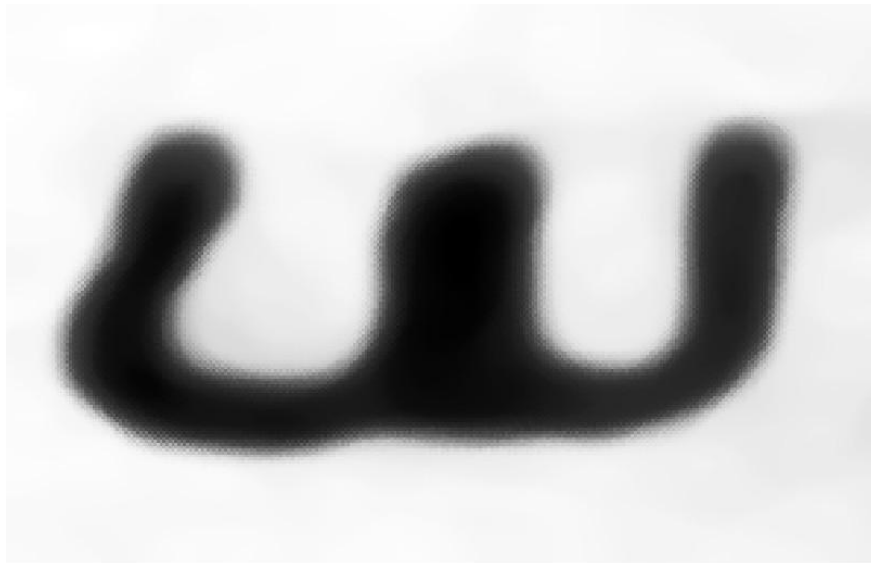
10ο παράδειγμα:



Εικόνα 3.5-28: Λεπτομέρεια του χαρακτήρα «ω» σε RGB

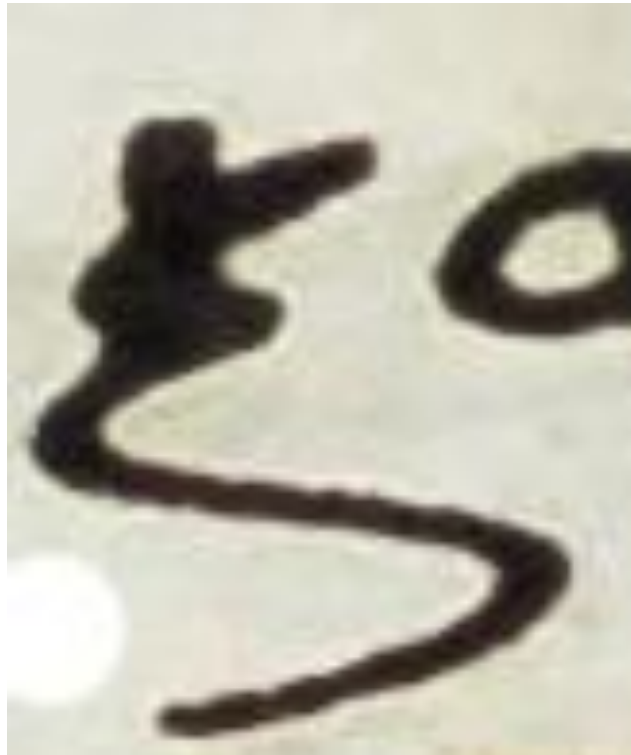


Εικόνα 3.5-29: Αποτύπωση της ροής που επάγει το πεδίο των κλίσεων, το οποίο υπολογίστηκε ως ισοδύναμο των τοπικών χρωματικών μεταβολών της εικόνας σύμφωνα με την LAB αναπαράστασή της.



Εικόνα 3.5-30: Η grayscale εικόνα που αντιστοιχεί στο υπολογισθέν πεδίο κλίσεων, όπως υπολογίστηκε με αριθμητική επίλυση κατάλληλα ορισμένου προβλήματος συνοριακών τιμών και όπως περιγράφηκε στην παράγραφο 3.4.

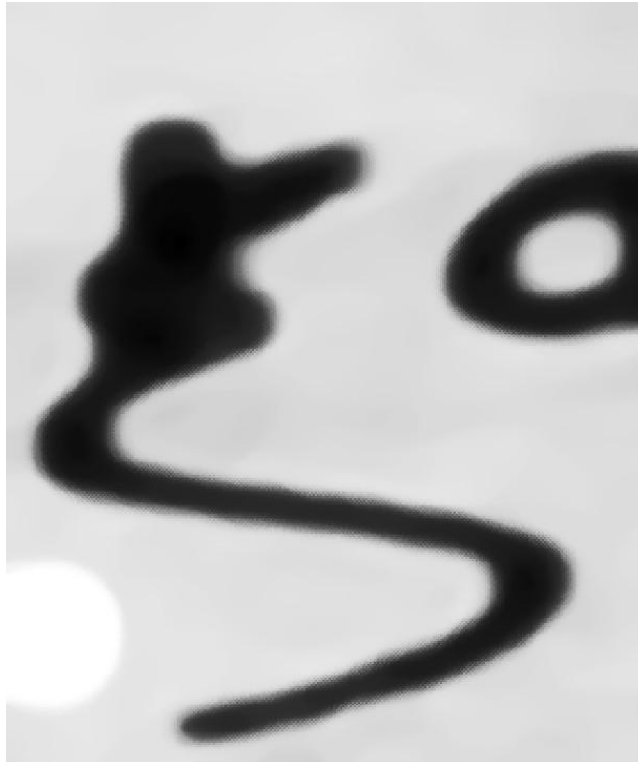
11ο παράδειγμα:



Εικόνα 3.5-31: Λεπτομέρεια του χαρακτήρα «ξ» σε RGB



Εικόνα 3.5-32: Η grayscale εικόνα που αντιστοιχεί στο υπολογισθέν πεδίο κλίσεων, όπως υπολογίστηκε με αριθμητική επίλυση κατάλληλα ορισμένου προβλήματος συνοριακών τιμών και όπως περιγράφηκε στην παράγραφο 3.4.



Εικόνα 3.5-33: Η grayscale εικόνα που αντιστοιχεί στο υπολογισθέν πεδίο κλίσεων, όπως υπολογίστηκε με αριθμητική επίλυση κατάλληλα ορισμένου προβλήματος συνοριακών τιμών και όπως περιγράφηκε στην παράγραφο 3.4.

3.6: Προσέγγιση των 2 περιοχών της εικόνας με τη χρήση κανονικών κατανομών επί του χρωματικού ιστογράμματος

Στην πορεία, προσεγγίζουμε το ιστόγραμμα της εικόνας βάσει 2 κανονικών κατανομών, μία για κάθε περιοχή της εικόνας, με τη μέθοδο των ελαχίστων τετραγώνων. Τα σημεία καμπής των δύο αυτών κανονικών κατανομών, με μέση τιμή μ και τυπική απόκλιση σ , προκύπτουν από τις βασικές ιδιότητες της κανονικής κατανομής:

Με συνάρτηση πυκνότητας πιθανότητας: $f(x) = \frac{1}{s\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

Πρώτη παράγωγο: $f'(x) = -\frac{x-\mu}{s^3\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

Δεύτερη παράγωγο: $f''(x) = -\frac{1}{s^3\sqrt{2\pi}} \left[1 - \frac{(x-\mu)^2}{\sigma^2} \right] e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

Άρα, τα σημεία καμπής εντοπίζονται από τη λύση της $f''(x) = 0$, και είναι:

$$\begin{aligned} x_k^1 &= \mu + \sigma \\ x_k^2 &= \mu - \sigma \end{aligned}$$

Αντί να αναζητήσουμε τις ακμές στην εικόνα μας $I(x, y)$ με την εύρεση των σημείων καμπής της, δηλαδή με τη λύση της εξίσωσης $\nabla(|\nabla I(x, y)|) = 0$, χρησιμοποιούμε τις δύο κανονικές κατανομές στο ιστογράμμα μας, οι οποίες και αντιστοιχούν στις δύο περιοχές της εικόνας (γράμμα και υπόβαθρο). Συμπεραίνουμε άμεσα ότι η κανονική κατανομή με μέγιστο κοντινότερο στο μηδέν (δηλαδή μαύρο χρώμα) περιλαμβάνει τα εικονοστοιχεία του γράμματος, ενώ η κανονική κατανομή με μέγιστο κοντινότερο στο 255 (λευκό) περιλαμβάνει εικονοστοιχεία του υποβάθρου. Αυτό βέβαια, δε μπορεί να ισχύει για το σύνολο των στοιχείων της κάθε κατανομής, καθώς το υπόβαθρο είναι πολύ πιθανό να περιλαμβάνει φθορές με σκούρο χρώμα, όπως και το γράμμα γίνεται να έχει λευκά στίγματα. Παρόλα αυτά, θεωρούμε πως ο γενικός κανόνας υπακούει σε αυτό το διαχωρισμό.

Έτσι λοιπόν, όσα εικονοστοιχεία βρίσκονται στο ίδιο κυτίο του ιστογράμματος με την κορυφή της κανονικής κατανομής K^A , η οποία βρίσκεται κοντύτερα στο μαύρο χρώμα, αποδίδονται στο σώμα του γράμματος, ενώ αντίστοιχα όσο εικονοστοιχεία βρίσκονται στο ίδιο κυτίο του ιστογράμματος με την κορυφή της κανονικής κατανομής K^B αποδίδονται στο υπόβαθρο. Ονομάζουμε τη θεματική περιοχή του γράμματος Σ και τη θεματική περιοχή του υποβάθρου Υ .

Έπειτα, βρίσκουμε το σημείο τομής I^T μεταξύ των 2 κανονικών κατανομών, καθώς και το δεξί σημείο καμπής της αριστερής κατανομής I^K , με τη μέθοδο που παρουσιάστηκε παραπάνω. Επιλέγουμε μεταξύ των I^T και I^K αυτό το οποίο έχει τη μικρότερη απόσταση από το μέγιστο της αριστερής κανονικής κατανομής, και το ονομάζουμε I^H .

Σε κάθε μία από τις κορυφές I^{KA} (κορυφή της κατανομής K^A) και I^{KB} (κορυφή της κατανομής K^B) κινούμαστε εκατέρωθεν έως ότου, από τα εικονοστοιχεία που έχουν συμπεριληφθεί, να φτάσουμε σε μία απόχρωση του γκρι τιμής $I^C = I(I^H) + 0,20(I(I^{KA}) - I(I^H))$.

Στην περίπτωση όμως που οι δύο κορυφές είναι δύσκολο να διαχωριστούν, σε σημείο που οι 2 κανονικές κατανομές πρακτικά ταυτίζονται, η κυρίαρχη κατανομή που προσφέρεται είναι αυτή με την δεξιά κορυφή, η οποία αντιστοιχεί στο υπόβαθρο. Θεωρούμε το αριστερό σημείο καμπής X^{AK} της μοναδικής κατανομής που βλέπουμε ως το σημείο στο οποίο τα εικονοστοιχεία του υποβάθρου διαχωρίζονται από αυτά του γράμματος, ενώ το κυτίο του ιστογράμματος με κέντρο X^{AK} περιλαμβάνει τα εικονοστοιχεία του περιγράμματος του γράμματος.

3.7: Εξάλειψη του θορύβου

Στο επόμενο βήμα της επεξεργασίας μας, θα προσπαθήσουμε να εξαλείψουμε το θόρυβο, δια της ομογενοποίησης των θεματικών περιοχών (αποχρώσεις έως και την τιμή I^C γύρω από τις τιμές I^{KA} και I^{KB}) που έχουμε παράξει. Για να υπολογίσουμε την κατανομή του θορύβου λόγω φθοράς που δημιουργήθηκε από την πάροδο του χρόνου επάνω στο υλικό του χαρτιού,

υπολογίζουμε την ένταση της κλίσης σε κάθε εικονοστοιχείο (i, j) της ασπρόμαυρης, όπως παράξαμε παραπάνω, εικόνας μας:

$$G_f(i, j) = \sqrt{\left(\frac{\partial I}{\partial x}\right)^2 + \left(\frac{\partial I}{\partial y}\right)^2} \cong \sqrt{(I(i+1, j) - I(i, j))^2 + (I(i, j+1) - I(i, j))^2}$$

Υπολογίζουμε τη μέση τιμή E_G και τη διασπορά R_G του μέτρου της κλίσης της του $G_f(i, j)$ και ελέγχουμε τον αριθμό N_G των εικονοστοιχείων που βρίσκονται στο διάστημα $(E_G - R_G, E_G + R_G)$. Αν ο αριθμός N_G είναι μικρότερος ενός ποσοστού P_G των συνολικών εικονοστοιχείων N , όπου σε αυτή την εφαρμογή επιλέχθηκε $P_G = 0,8$, τότε θεωρούμε ότι υφίσταται μεγάλη ποσότητα θορύβου λόγω φθοράς. Εφαρμόζουμε λοιπόν τη διαδικασία με την ολίσθηση της μάσκας περιττών διαστάσεων, η οποία παρουσιάστηκε στην παράγραφο 3.2, και λαμβάνουμε έτσι την τελικώς ομογενοποιημένη εικόνα μας.

Αφού δημιουργήσουμε τις ομογενοποιημένες περιοχές μέσω των παραπάνω διαδικασιών, μένει το τελικό βήμα του να αποδώσουμε όσο εικονοστοιχεία απέμειναν «ορφανά» σε αυτές. Έστω Σ^0 η περιοχή που αντιστοιχεί στο γράμμα και Υ^0 η περιοχή που ανήκει στο υπόβαθρο. Με $I_\mu^{\Sigma^0}$ και $I_\mu^{\Upsilon^0}$ τους μέσους όρους των αποχρώσεων του γκρι των εικονοστοιχείων των περιοχών Σ^0 και Υ^0 αντίστοιχα, και I^X την απόχρωση ενός τυχαίου αταξινόμητου εικονοστοιχείου, ορίζουμε το εξής μέτρο απόστασης αποχρώσεως του εικονοστοιχείου από την περιοχή Σ^0 :

$$d(I^X, \Sigma^0) = (\text{απόσταση } I^X \text{ από } I_\mu^{\Sigma^0}) + \left(\begin{array}{l} \text{ελάχιστη απόσταση } I^X \text{ από σύνολο} \\ \text{αποχρώσεων των εικονοστοιχείων της } \Sigma^0 \end{array} \right)$$

Ενώ αντίστοιχα, ορίζουμε το μέτρο απόστασης απόχρωσης του εικονοστοιχείου από την περιοχή Υ^0 :

$$d(I^X, \Upsilon^0) = (\text{απόσταση } I^X \text{ από } I_\mu^{\Upsilon^0}) + \left(\begin{array}{l} \text{ελάχιστη απόσταση } I^X \text{ από σύνολο} \\ \text{αποχρώσεων των εικονοστοιχείων της } \Upsilon^0 \end{array} \right)$$

Κάνουμε την αφαίρεση $d(I^X, \Sigma^0) - d(I^X, \Upsilon^0)$ για κάθε εικονοστοιχείο I^X , και το εικονοστοιχείο με τη μικρότερη τιμή αυτής τη διαφοράς το αποδίδουμε στην περιοχή Σ^0 , ενώ αυτό με τη μεγαλύτερη τιμή διαφοράς το αποδίδουμε στην περιοχή Υ^0 . Επαναλαμβάνουμε τη διαδικασία έχοντας πλέον επανορίσει τις περιοχές Σ^0 και Υ^0 και τους μέσους όρους των αποχρώσεών τους $I_\mu^{\Sigma^0}$ και $I_\mu^{\Upsilon^0}$.

Για να λάβουμε, τέλος, μία binary εκδοχή της τελικής μας εικόνας, δηλαδή μία αποκλειστικά μαύρου και άσπρου χρώματος, αποδίδουμε στην περιοχή Σ^0 την τιμή 0, δηλαδή μαύρο χρώμα και στην περιοχή Υ^0 την τιμή 1 ή 255, δηλαδή άσπρο χρώμα, αναλόγως του συστήματος που χρησιμοποιούμε.

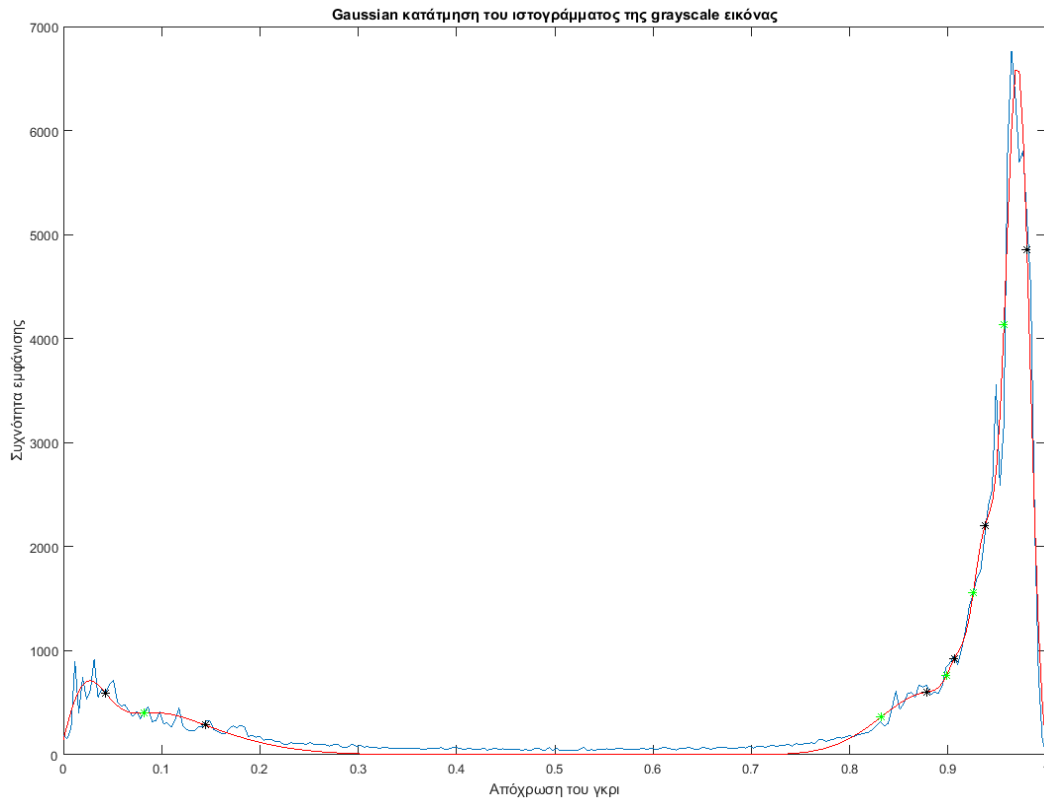
Κεφάλαιο 4: Αποτελέσματα Εφαρμογής της Δοκιμασθείσας Μεθόδου

Τέλος, σε αυτό το κεφάλαιο θα παρουσιάσουμε τα αποτελέσματα του αλγορίθμου πάνω στη σειρά των γραμμάτων που επιλέχθηκαν ως αντιπροσωπευτικά για την απεικόνιση του μετασχηματισμού σε χρωματικό χώρο LAB και κατόπιν μέσω αυτού σε grayscale. Παρουσιάζεται για κάθε γράμμα η αυθεντική του απεικόνιση, όπως εμφανίζεται στο κείμενο του πατύρου και σε παράθεσή η grayscale εκδοχή της, όπως υπολογίστηκε με βάση τη μεθοδολογία της παραγράφου 3.4. Έπειτα, παρουσιάζεται το ιστόγραμμα της grayscale εκδοχής και η προσέγγισή του με κανονικές κατανομές. Επί της προσέγγισης αυτής, τυπώνεται η αλληλουχία των σημείων που ανά τρία ορίζουν διαστήματα αποχρώσεων του γκρι που αντιστοιχούν σε περιοχές της εικόνας διαχωρισμένες μεταξύ τους με ισχυρές ακμές. Στη συνέχεια παρατίθενται τόσο η εικόνα που προκύπτει διακρίνοντας και ομογενοποιώντας τις περιοχές αυτές, όσο και ομαδοποιώντας τες ώστε να προκύψει η μεγιστοπιθήνη δυαδική (ασπρόμαυρη) κατάτμηση εικόνας. Τέλος παρατίθεται και η ιδανική για τους σκοπούς της εργασίας μας απεικόνιση, έπειτα από εξεργασία ατελειών στο χέρι (αφαίρεση κόκκων, σημαδιών κλπ και απομόνωση του χαρακτήρα ή του συμπλέγματος χαρακτήρων που επιθυμούμε) καθώς και αλλαγές στις τιμές του thresholding και λοιπόν παραμέτρων, είτε στην τελική ασπρόμαυρη εικόνα είτε στην αμέσως προηγούμενη αυτής grayscale.

1ο παράδειγμα:



Εικόνα 4-1: Αυθεντική απεικόνιση ενός συμπλέγματος των χαρακτήρων «Ετα» (αριστερά) και η grayscale εκδοχή του (δεξιά)



Εικόνα 4-2: Η κατάτμηση του ιστογράμματος της grayscale εκδοχής της εικόνας του συμπλέγματος των χαρακτήρων «Ετα»

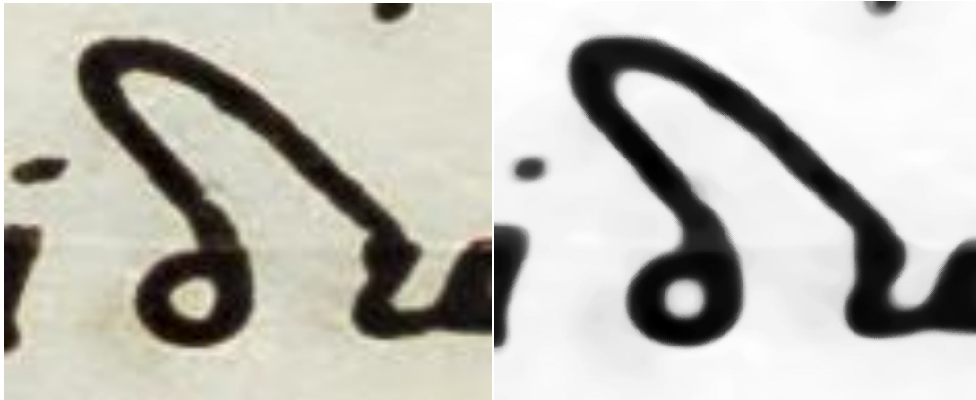


Εικόνα 4-3: Η απεικόνιση της κατάτμησης του ιστογράμματος ως κατάτμηση της grayscale εικόνας (αριστερά) και η μεγιστοποιήθη αντιστοιχη δυαδική κατάτμησης της (δεξιά)

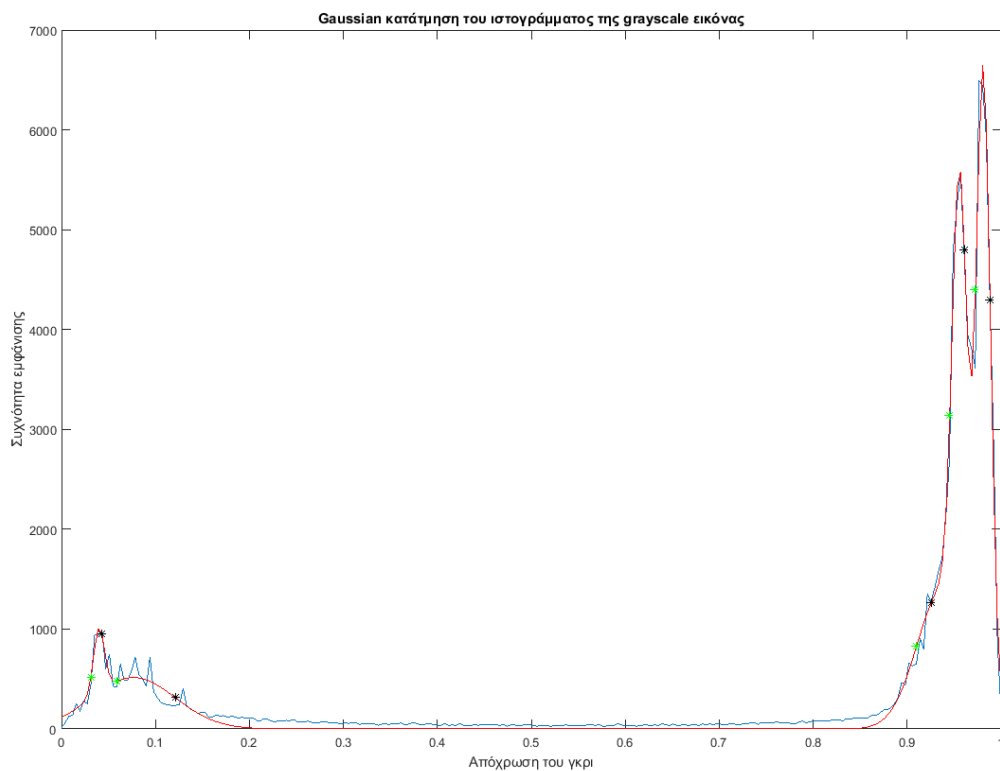


Εικόνα 4-4: Επιθυμητή κατάτμηση του συμπλέγματος των χαρακτήρων «Ετ» από χρήστη

2ο παράδειγμα:



Εικόνα 4-5: Αυθεντική απεικόνιση του χαρακτήρα «δ» (αριστερά) και η grayscale εκδοχή του (δεξιά)



Εικόνα 4-6: Η κατάτμηση του ιστογράμματος της grayscale εκδοχής της εικόνας του χαρακτήρα «δ»



Εικόνα 4-7: Η απεικόνιση της κατάτμησης του ιστογράμματος ως κατάτμηση της grayscale εικόνας (αριστερά) και η μεγιστοποιήσιμη αντίστοιχη δυαδική κατάτμησης της (δεξιά)

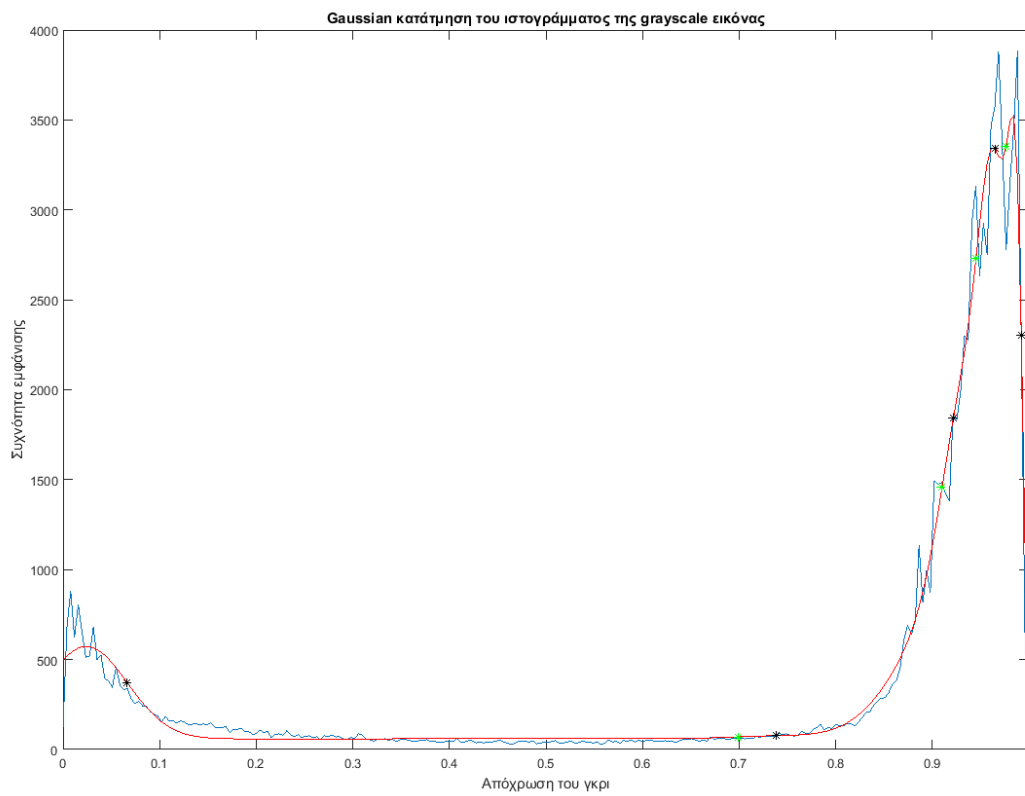


Εικόνα 4-8: Επιθυμητή κατάτμηση του χαρακτήρα «δ» από χρήστη

3ο παράδειγμα:



Εικόνα 4-9: Αυθεντική απεικόνιση του χαρακτήρα «Κ» (αριστερά) και η grayscale εκδοχή του (δεξιά)



Εικόνα 4-10: Η κατάτμηση του ιστογράμματος της grayscale εκδοχής της εικόνας του χαρακτήρα «Κ»

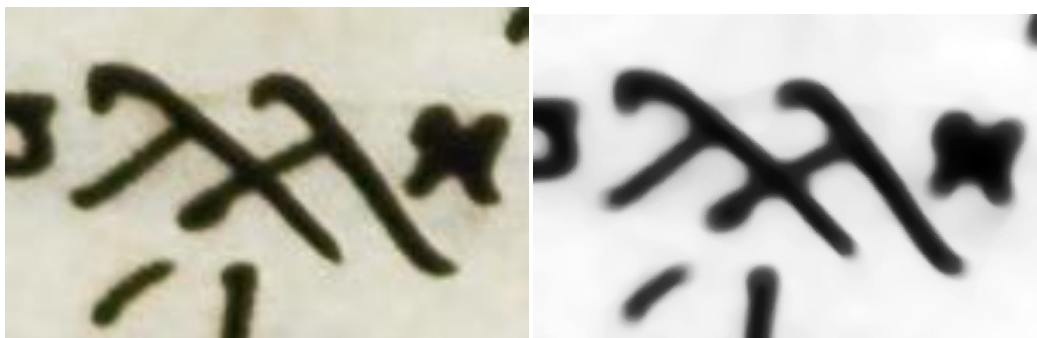


Εικόνα 4-11: Η απεικόνιση της κατάτμησης του ιστογράμματος ως κατάτμηση της grayscale εικόνας (αριστερά) και η μεγιστοποιημένη αντίστοιχη δυαδική κατάτμησης της (δεξιά)

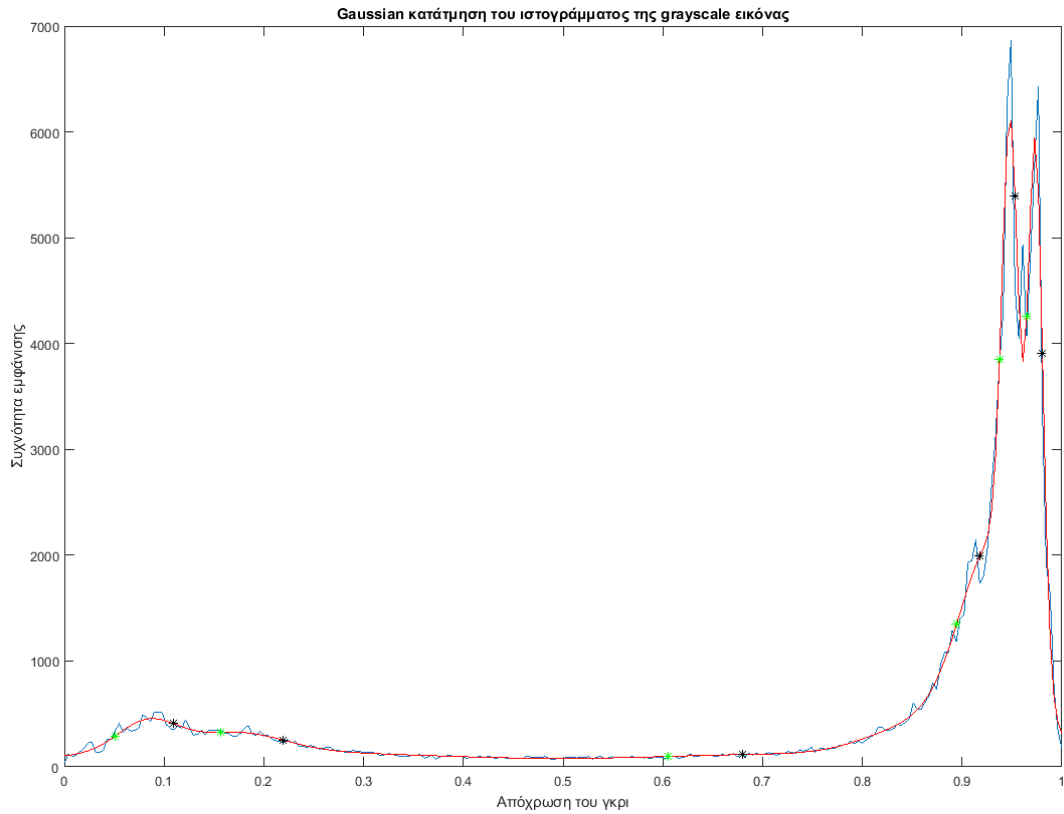


Εικόνα 4-12: Επιθυμητή κατάτμηση του χαρακτήρα «Κ» από χρήστη

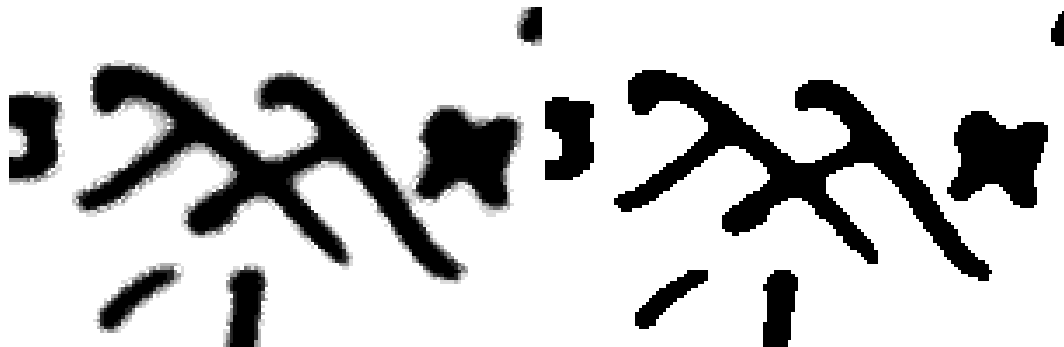
4ο παράδειγμα:



Εικόνα 4-9: Αυθεντική απεικόνιση του συμπλέγματος χαρακτήρων «λλ» (αριστερά) και η grayscale εκδοχή του (δεξιά)



Εικόνα 4-10: Η κατάτμηση του ιστογράμματος της grayscale εκδοχής της εικόνας του συμπλέγματος χαρακτήρων «λλ»



Εικόνα 4-11: Η απεικόνιση της κατάτμησης του ιστογράμματος ως κατάτμηση της grayscale εικόνας (αριστερά) και η μεγιστοποιήσιμη αντίστοιχη δυαδική κατάτμησης της (δεξιά)

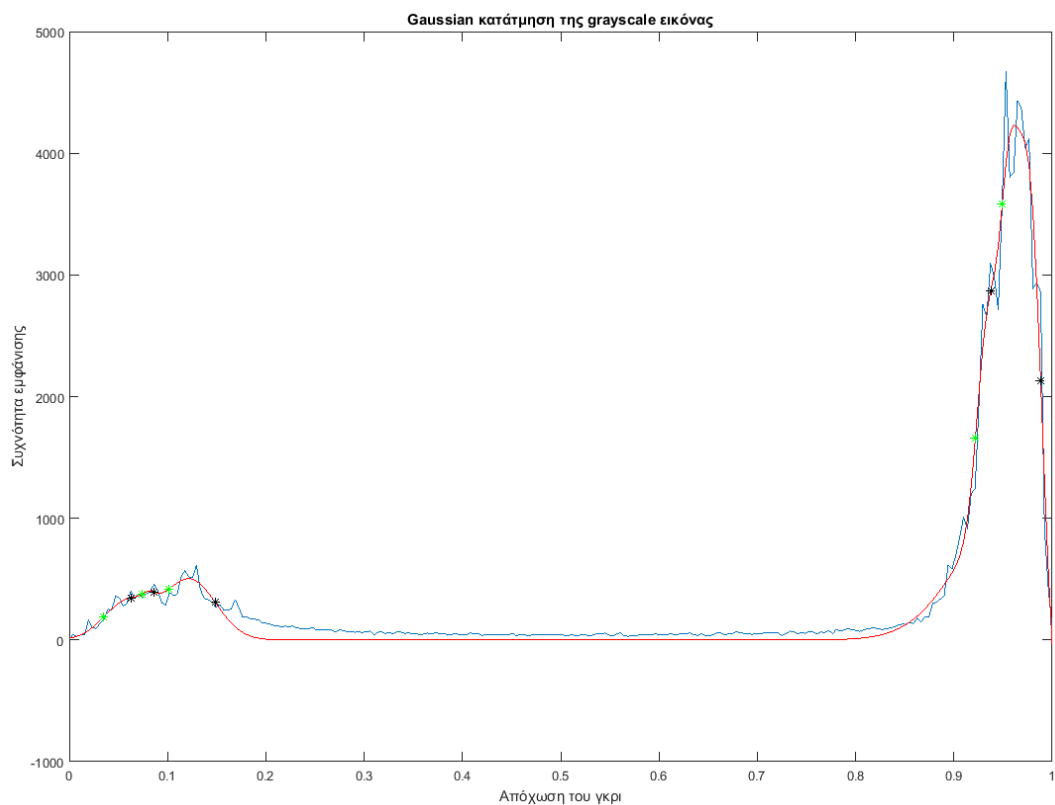


Εικόνα 4-12: Επιθυμητή κατάτμηση του χαρακτήρα «λ» από χρήστη (απομονωμένου από το σύμπλεγμα για τους σκοπούς της εργασίας)

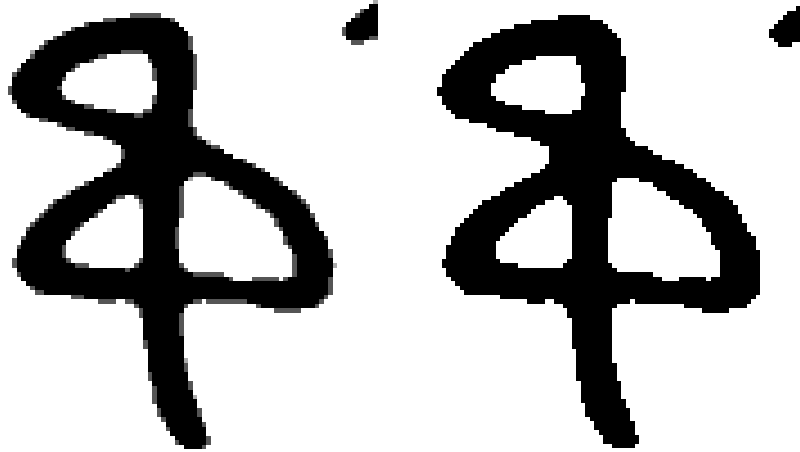
5ο παράδειγμα:



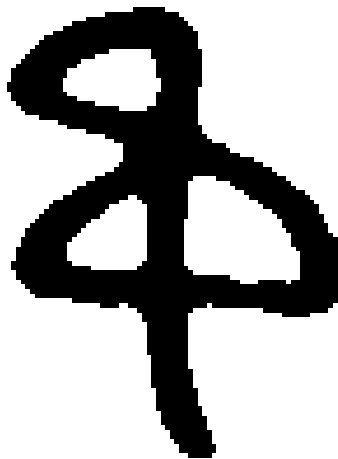
Εικόνα 4-13: Αυθεντική απεικόνιση του χαρακτήρα «Φ» (αριστερά) και η grayscale εκδοχή του (δεξιά)



Εικόνα 4-14: Η κατάτμηση του ιστογράμματος της grayscale εκδοχής της εικόνας του χαρακτήρα «Φ»



Εικόνα 4-15: Η απεικόνιση της κατάτμησης του ιστογράμματος ως κατάτμηση της grayscale εικόνας (αριστερά) και η μεγιστοποιήθη αντιστοιχη δυαδική κατάτμησης της (δεξιά)

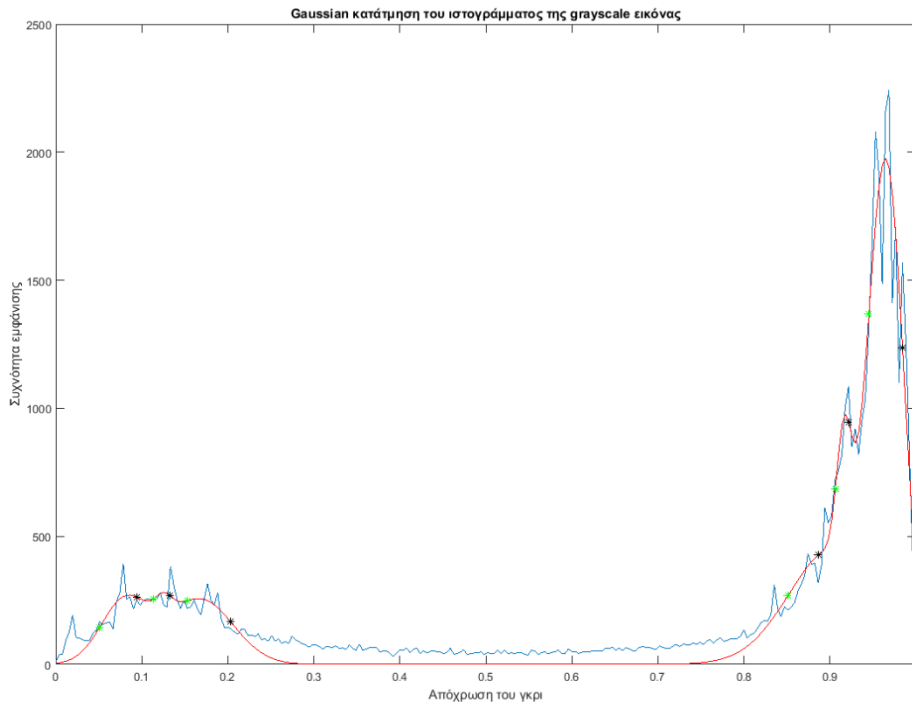


Εικόνα 4-16: Επιθυμητή κατάτμηση του χαρακτήρα «Φ» από χρήστη

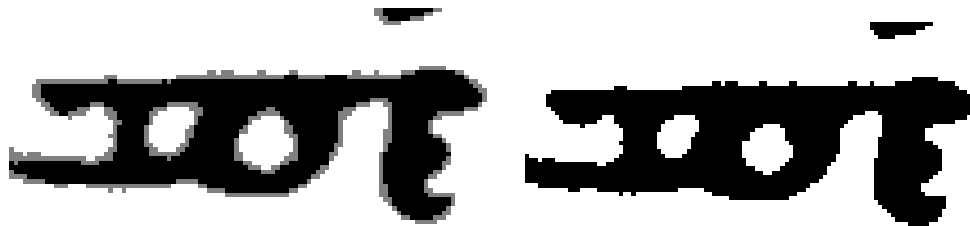
6ο παράδειγμα:



Εικόνα 4-17: Αυθεντική απεικόνιση του συμπλέγματος χαρακτήρων «πε» (αριστερά) και η grayscale εκδοχή του (δεξιά)



Εικόνα 4-18: Η κατάτμηση του ιστογράμματος της grayscale εκδοχής της εικόνας του συμπλέγματος χαρακτήρων «πε»



Εικόνα 4-19: Η απεικόνιση της κατάτμησης του ιστογράμματος ως κατάτμηση της grayscale εικόνας (αριστερά) και η μεγιστοπίθανη αντίστοιχη δυαδική κατάτμησης της (δεξιά)

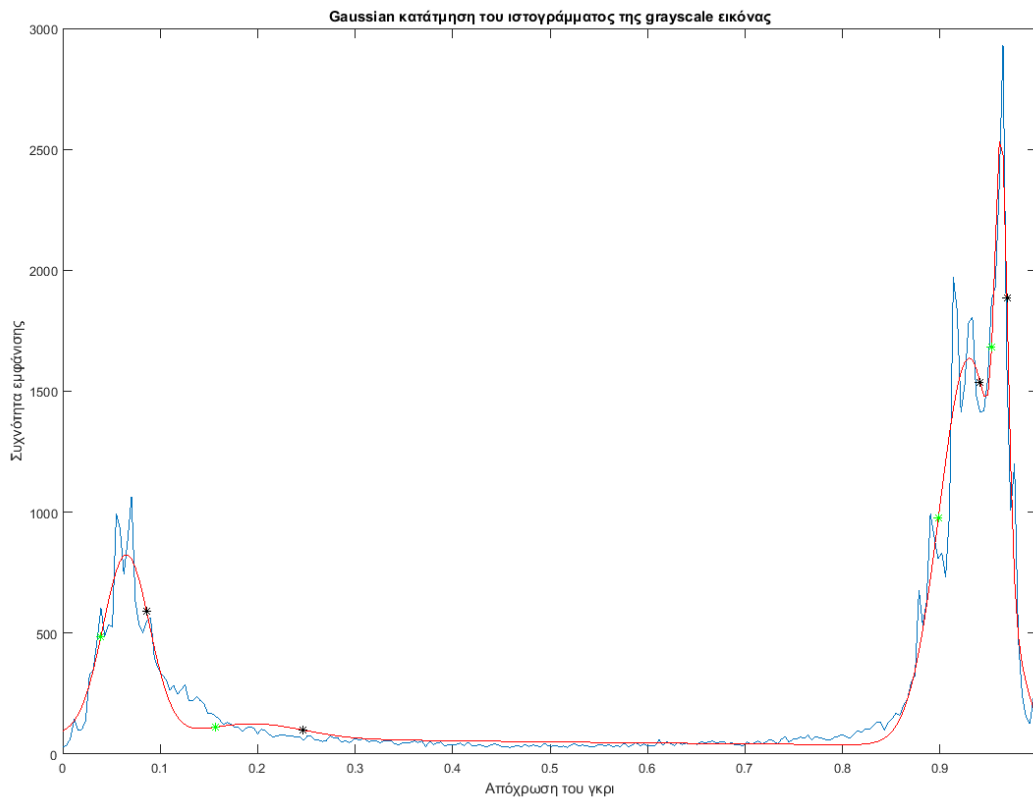


Εικόνα 4-20: Επιθυμητή κατάτμηση του (απομονωμένου) χαρακτήρα «π» από χρήστη

7ο παράδειγμα:



Εικόνα 4-21: Αυθεντική απεικόνιση του συμπλέγματος χαρακτήρων «πο» (αριστερά) και η grayscale εκδοχή του (δεξιά)



Εικόνα 4-22: Η κατάτμηση του ιστογράμματος της grayscale εκδοχής της εικόνας του συμπλέγματος χαρακτήρων «πο»



Εικόνα 4-23: Η απεικόνιση της κατάτμησης του ιστογράμματος ως κατάτμηση της grayscale εικόνας (αριστερά) και η μεγιστοποιήσιμη αντίστοιχη δυαδική κατάτμησης της (δεξιά)

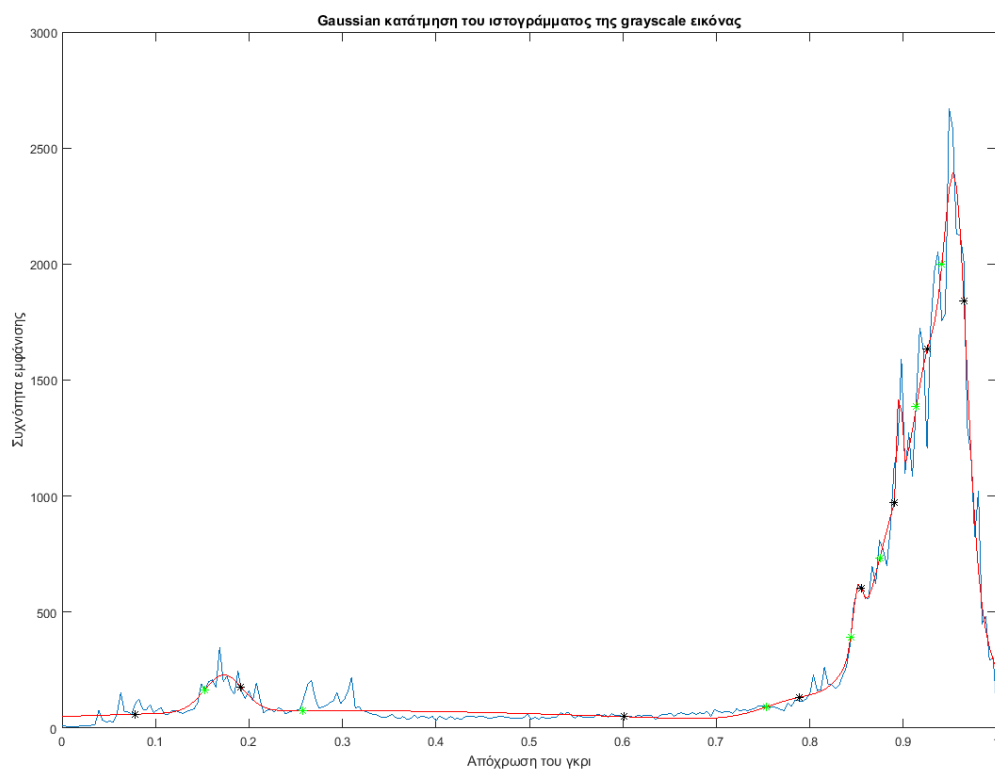


Εικόνα 4-24: Επιθυμητή κατάτμηση του χαρακτήρα «π» από χρήστη

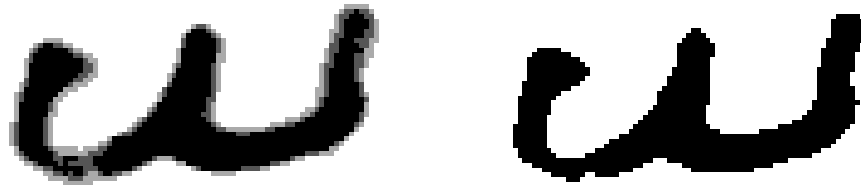
8ο παράδειγμα:



Εικόνα 4-25: Αυθεντική απεικόνιση του χαρακτήρα «ω» (αριστερά) και η grayscale εκδοχή της (δεξιά)



Εικόνα 4-26: Η κατάτμηση του ιστογράμματος της grayscale εκδοχής της εικόνας του χαρακτήρα «ω»



Εικόνα 4-27: Η απεικόνιση της κατάτμησης του ιστογράμματος ως κατάτμηση της grayscale εικόνας (αριστερά) και η μεγιστοποιήθη αντιστοιχη δυαδική κατάτμησης της (δεξιά)

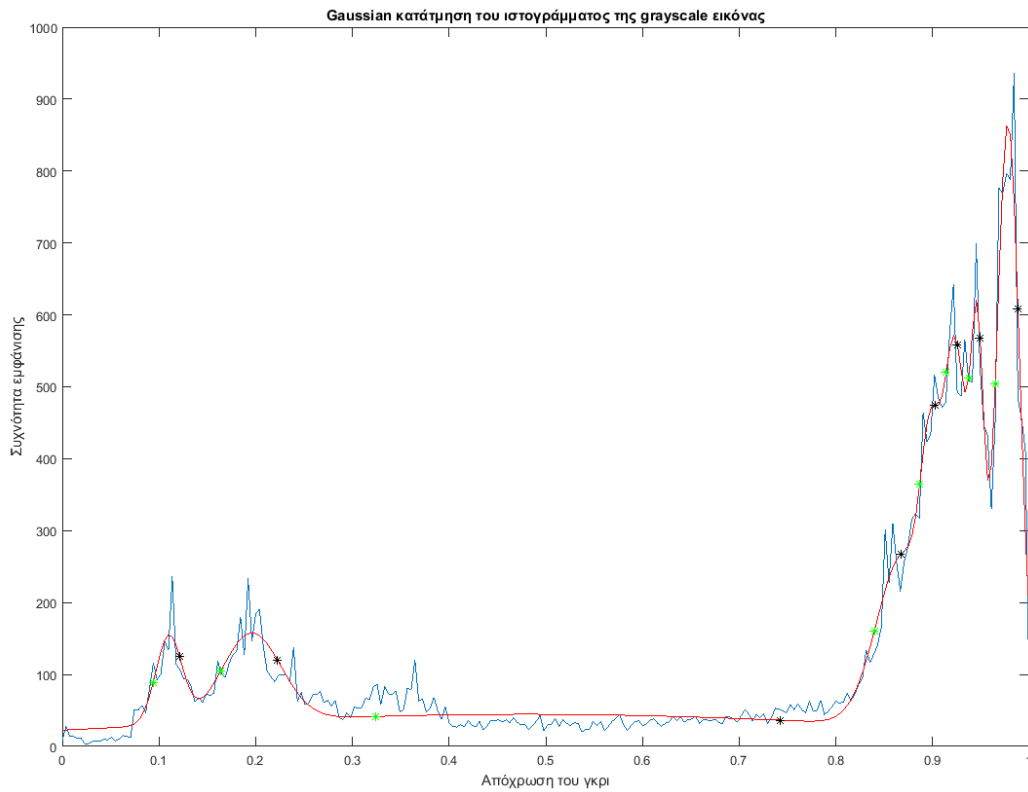


Εικόνα 4-28: Επιθυμητή κατάτμηση του χαρακτήρα «ξ» από χρήστη

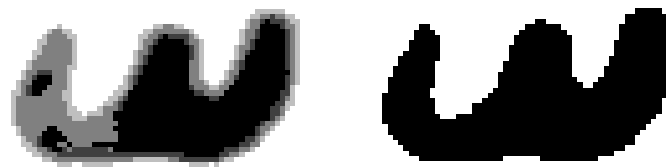
9ο παράδειγμα:



Εικόνα 4-29: Αυθεντική απεικόνιση του χαρακτήρα «ω» (αριστερά) και η grayscale εκδοχή της (δεξιά)



Εικόνα 4-30: Η κατάτμηση του ιστογράμματος της grayscale εκδοχής της εικόνας του χαρακτήρα «ω»



Εικόνα 4-31: Η απεικόνιση της κατάτμησης του ιστογράμματος ως κατάτμηση της grayscale εικόνας (αριστερά) και η μεγιστοποιήσιμη αντίστοιχη δυαδική κατάτμησης της (δεξιά)

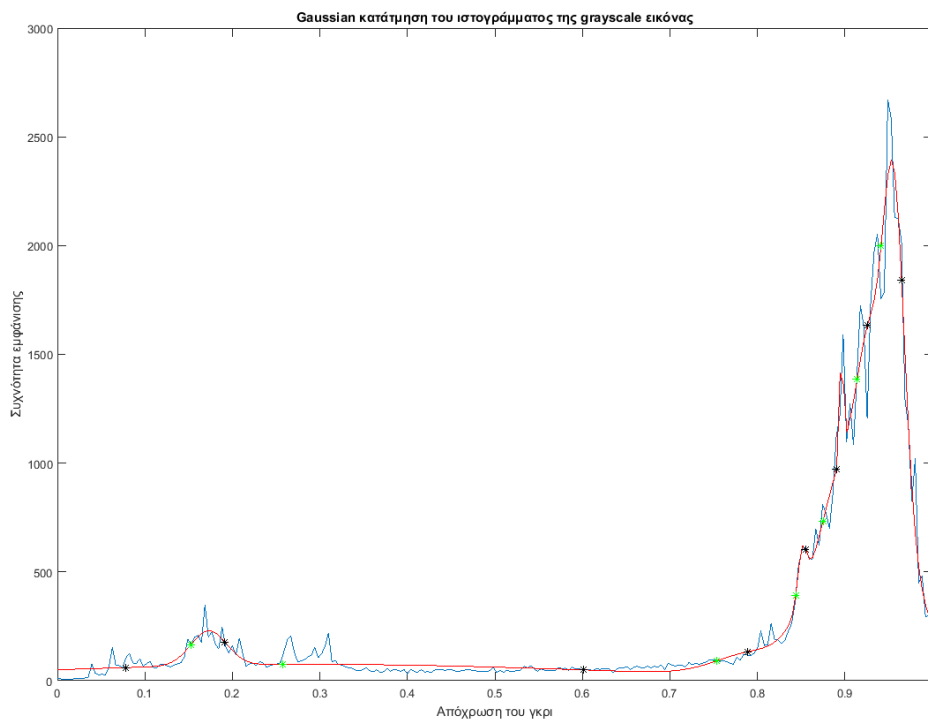


Εικόνα 4-32: Επιθυμητή κατάτμηση του χαρακτήρα «ξ» από χρήστη

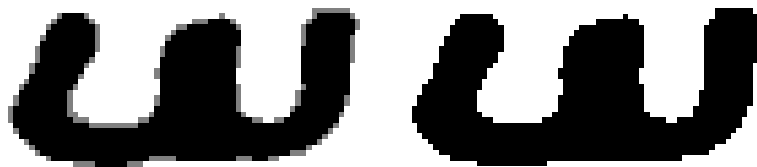
10ο παράδειγμα:



Εικόνα 4-33: Αυθεντική απεικόνιση του χαρακτήρα «ω» (αριστερά) και η grayscale εκδοχή της (δεξιά)



Εικόνα 4-34: Η κατάτμηση του ιστογράμματος της grayscale εκδοχής της εικόνας του χαρακτήρα «ω»

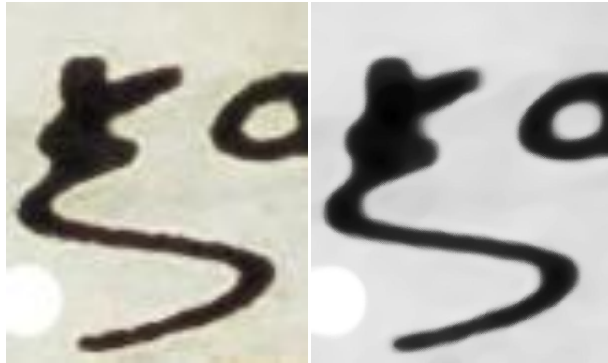


Εικόνα 4-35: Η απεικόνιση της κατάτμησης του ιστογράμματος ως κατάτμηση της grayscale εικόνας (αριστερά) και η μεγιστοποιήσιμη αντίστοιχη δυαδική κατάτμησης της (δεξιά)

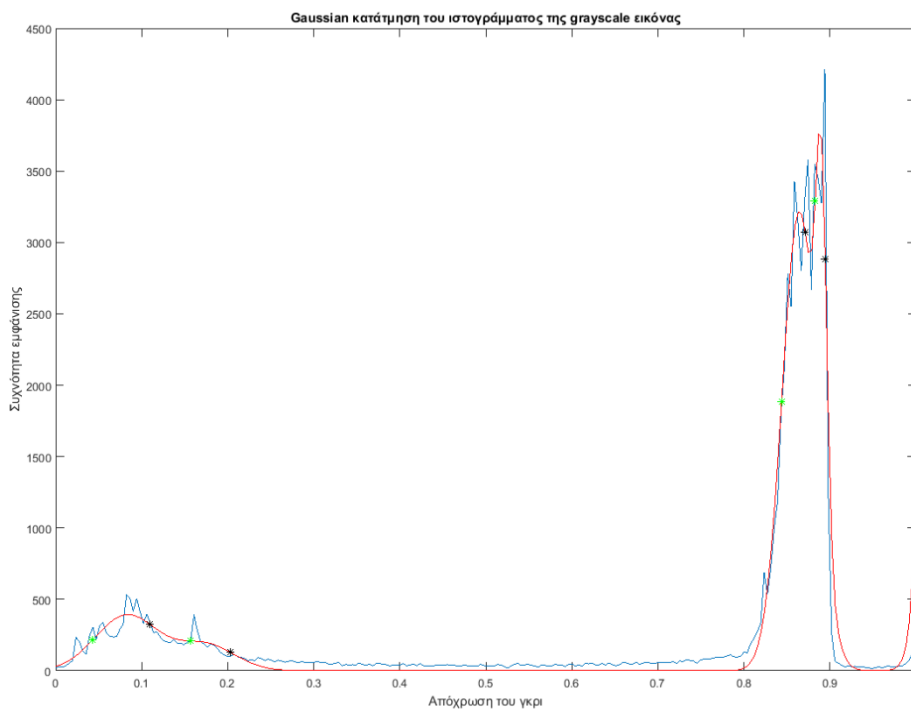


Εικόνα 4-36: Επιθυμητή κατάτμηση του χαρακτήρα «ξ» από χρήστη

11ο παράδειγμα:



Εικόνα 4-37: Αυθεντική απεικόνιση ενός χαρακτήρα «ξ» (αριστερά) και η grayscale εκδοχή της (δεξιά)



Εικόνα 4-38: Η κατάτμηση του ιστογράμματος της grayscale εκδοχής της εικόνας του χαρακτήρα «ξ»



Εικόνα 4-39: Η απεικόνιση της κατάτμησης του ιστογράμματος ως κατάτμηση της grayscale εικόνας (αριστερά) και η μεγιστοποιημένη αντίστοιχη δυαδική κατάτμησης της (δεξιά)



Εικόνα 4-40: Επιθυμητή κατάτμηση του χαρακτήρα «ξ» από χρήστη

Κεφάλαιο 5: Συμπεράσματα

Έπειτα από μία ιδιαίτερα εκτεταμένη χρήση της υιοθετηθείσας μεθόδου κατάτμησης εικόνας, σε ένα σύνολο 2.261 χαρακτήρων, μπορούμε και αντλούμε τα εξής συμπεράσματα:

Η ακρίβεια της μεθόδου επιβεβαιώνεται από την πλήρη αξιοποίηση των τελικών ασπρόμαυρων χαρακτήρων από την ερευνητική ομάδα για τους σκοπούς των επόμενων σταδίων της συνολικής ερευνητικής εργασίας. Η δημιουργηθείσα βάση των χαρακτήρων αποδεικνύεται πολύ χρήσιμη στον τελικό στόχο της υλοποίησης ενός αλγορίθμου εντοπισμού του γραφικού χαρακτήρα του καταγραφέα σε άλλα κείμενα.

Επιπλέον, στην πράξη επιβεβαιώνεται η απλοποίηση της αρχικής προταθείσας από την ερευνητική ομάδα του κυρίου Παπαοδυσσέα μεθόδου. Αποφεύγοντας τη μετατροπή της εικόνας σε grayscale και τη χρήση εργαλείων γι αυτό, και αντλώντας μέσω της κλίσης της εικόνας στο χρωματικό χώρο LAB το ιστόγραμμα της αντίστοιχης grayscale, έχουμε σημαντική διαφορά στο χρόνο υλοποίησης της διαδικασίας της τροποποιημένης μεθόδου συγκριτικά με της αρχικής προταθείσας, κάτι που αποτυπώνεται εντονότερα στην περίπτωση αυτής της διπλωματικής εργασίας, εξαιτίας του μεγάλου όγκου εικόνων που χρειάστηκε να υποστούν επεξεργασία.

Ένα σημαντικό μειονέκτημα της μεθόδου αυτής, στα πλαίσια της συνολικής εργασίας, είναι η αδυναμία της να εντοπίσει το κύριο σώμα του γράμματος μέσα σε μία εικόνα με ποικίλα σημάδια φθοράς, μελανιού, ίχνη διπλανών ή και ενωμένων για χάρη καλλιγραφίας γραμμάτων, πνευμάτων και λοιπών στοιχείων που κατά την κατάτμηση δεν θεωρήθηκαν υπόβαθρο ώστε να απαλειφθούν. Σε ένα μεγάλο αριθμό χαρακτήρων αναγκαστήκαμε να προχωρήσουμε σε επεξεργασία με το χέρι πάνω στην τελική δυαδική κατάτμηση προκειμένου να απαλείψουμε τα παραπάνω και να μείνει μόνο το πραγματικό σώμα τους.

Ένα δεύτερο μειονέκτημα της μεθόδου αυτής το συναντήσαμε σε μία μικρή γκάμα χαρακτήρων με κοινό χαρακτηριστικό τη φθορά περιμετρικά αυτών, λόγω της παλαιότητας του παπύρου που εξασθενούσε την ένταση του μελανιού. Σε αυτές τις περιπτώσεις, επειδή ο αλγόριθμος αντιλαμβανόταν τη φθορά αυτή εν μέρει ως μέρος του υπόβαθρου κι εν μέρει του γράμματος, έδινε μία κατάτμηση με αρκετά «μυτερή» και ακαλαίσθητη περίμετρο (εσωτερική ή εξωτερική) στο κύριο σώμα του γράμματος. Λογικό βήμα σε αυτά τα εμπόδια ήταν ξανά η επεξεργασία με το χέρι, αυτή τη φορά όμως πάνω στην grayscale απεικόνιση της κατάτμησης του grayscale ιστογράμματος, με στόχο να δοκιμάσουμε διαφορετικές τιμές στο thresholding και να επιτύχουμε μία πιο συμπαγή μορφή του χαρακτήρα, κι έπειτα η χειροκίνητη αποθήκευση του αποτελέσματος σε μία νέα δυαδική κατάτμηση.

Πηγές

- [1] Bruce Justin Lindbloom, on-line κείμενο http://www.brucelindbloom.com/index.html?Eqn_RGB_XYZ_Matrix.html
- [2] <https://en.wikipedia.org/wiki/CIELUV>
- [3] https://en.wikipedia.org/wiki/Lab_color_space
- [4] Gernot Hoffmann, " CIELab Color Space", on-line κείμενο <http://docs-hoffmann.de/cielab03022003.pdf>
- [5] Νικόλαος Η. Παπαμάρκος, «Ψηφιακή επεξεργασία και ανάλυση εικόνας», Εκδόσεις Γκιούρδας, σελ. 5.12
- [6] Νικόλαος Η. Παπαμάρκος, «Ψηφιακή επεξεργασία και ανάλυση εικόνας», Εκδόσεις Γκιούρδας, σελ. 6.3
- [7] Rafael Gonzalez, Richard Woods, «Ψηφιακή Επεξεργασία Εικόνας», Εκδόσεις Τζιόλας, σελ. 1100
- [8] https://en.wikipedia.org/wiki/Minimum_spanning_tree-based_segmentation
- [9] Jan Wassenberg , Wolfgang Middelmann, Peter Sanders, "An Efficient Parallel Algorithm for Graph-Based Image Segmentation", Computer Analysis of Image Patterns, Lecture Notes in Computer Science, vol. 5702, pp. 1003-1010, 2009
- [10] Constantin Papaodysseus, Mihalis Exarhos, Mihalis Panagopoulos, Panayiotis Rousopoulos, Constantin Triantafillou, and Thanasis Panagopoulos, "Image and Pattern Analysis of 1650 B.C. Wall Paintings and Reconstruction", IEEE Trans. on Systems, Man and Cybernetics - Part A: Systems and Humans ,vol. 38, no. 4, pp. 958-965, 2008.
- [11] I.S. Sevcenco, P.J. Hampton, P. Agathoklis, "A wavelet based method for image reconstruction from gradient data with applications", Multidimensional Systems and Signal Processing, November 2013
- [12] P.J. Hampton, P. Agathoklis, C. Bradley, "A New Wave-Front Reconstruction Method for Adaptive Optics Systems Using Wavelets", IEEE Journal of Selected Topics in Signal Processing, vol. 2, no. 5, October 2008
- [13] P.J. Hampton, P. Agathoklis, "Comparison of Haar Wavelet-based and Poisson-based Numerical Integration Techniques", Proceedings of Circuits and Systems (ISCAS), pp.1623-1626, 2010.