

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ
ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΚΑΙ ΤΕΧΝΟΛΟΓΙΚΩΝ
ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΕΠΙΣΤΗΜΗΣ ΥΠΟΛΟΓΙΣΤΩΝ

Επαναληπτικός γραφοθεωρητικός αλγόριθμος για
τμηματοποίηση εικόνων ως προς το χρώμα με χρήση μη
Παραμετρικών Στατιστικών Δοκιμασιών

Μεταπτυχιακή Εργασία

Ηράκλειο, Φεβρουάριος 2005

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ
ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΚΑΙ ΤΕΧΝΟΛΟΓΙΚΩΝ
ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΕΠΙΣΤΗΜΗΣ ΥΠΟΛΟΓΙΣΤΩΝ

**Επαναληπτικός γραφοθεωρητικός αλγόριθμος για
τμηματοποίηση εικόνων ως προς το χρώμα με χρήση Μη
Παραμετρικών Στατιστικών Δοκιμασιών**

Εργασία που υποβλήθηκε από την
Καλλιρρόη Φλουρή
ως μερική εκπλήρωση των απαιτήσεων για την απόκτηση
ΜΕΤΑΠΤΥΧΙΑΚΟΥ ΔΙΠΛΩΜΑΤΟΣ ΕΙΔΙΚΕΥΣΗΣ

Συγγραφέας:

Καλλιρρόη Φλουρή
Τμήμα Επιστήμης Υπολογιστών
Πανεπιστήμιο Κρήτης

Εισηγητική Επιτροπή:

Στέλιος Ορφανουδάκης, Καθηγητής, Πρόεδρος

Παναγιώτης Τραχανιάς, Καθηγητής, Μέλος

Παναγιώτης Τσακαλίδης, Αναπληρωτής Καθηγητής, Μέλος

Αντώνης Αργυρός, Ερευνητής Ι.Π., Ι.Τ.Ε, Μέλος

Δεκτή:

Δημήτρης Πλεξουσάκης, Αναπληρωτής Καθηγητής
Πρόεδρος Επιτροπής Μεταπτυχιακών Σπουδών

Επαναληπτικός γραφοθεωρητικός αλγόριθμος για τμηματοποίηση εικόνων ως προς το χρώμα με χρήση μη Παραμετρικών Στατιστικών Δοκιμασιών

Φλουρή Καλλιρρόη

Μεταπτυχιακή Εργασία

Τμήμα Επιστήμης Υπολογιστών
Πανεπιστήμιο Κρήτης

ΠΕΡΙΛΗΨΗ

Η τμηματοποίηση αποτελεί ένα πολύ σημαντικό πρόβλημα στην επεξεργασία και ανάλυση εικόνων. Στόχος της τμηματοποίησης είναι ο διαχωρισμός της εικόνας σε περιοχές έτσι ώστε κάθε μία από τις αυτές να αντιστοιχεί σε ένα φυσικό αντικείμενο. Διεργασίες υψηλού επιπέδου όπως η αναγνώριση αντικειμένων και η ανάκτηση εικόνων με βάση το περιεχόμενό τους βασίζονται σε πολλές περιπτώσεις στην αξιόπιστη τμηματοποίηση εικόνων. Πολλές μέθοδοι τμηματοποίησης έχουν προταθεί οι οποίες όμως κάνουν συνήθως περιοριστικές παραδοχές για τις στατιστικές ιδιότητες των χαρακτηριστικών βάσει των οποίων γίνεται η τμηματοποίηση, το πλήθος των διαφορετικών τμημάτων της εικόνας, ή και τα δύο. Σε αυτή την εργασία προτείνεται ένας γραφοθεωρητικός αλγόριθμος για τμηματοποίηση εικόνων ως προς το χρώμα. Ο προτεινόμενος αλγόριθμος χρησιμοποιεί ένα κριτήριο ομοιότητας που καθορίζεται από μια μη παραμετρική στατιστική δοκιμασία (Mann Whitney) χωρίς να γίνεται καμία υπόθεση για το μοντέλο που ακολουθεί η κατανομή χρώματος στις εικόνες. Ο αλγόριθμος εφαρμόζεται σε έγχρωμες εικόνες, για οποιοδήποτε χρωματικό μοντέλο χωρίς να απαιτείται η εκ των προτέρων γνώση του πλήθους των τμημάτων της κάθε εικόνας. Επίσης, προτείνονται συγκεκριμένες τεχνικές που αποσκοπούν στο να μειώσουν τις υπολογιστικές απαιτήσεις της διαδικασίας τμηματοποίησης. Τα πειραματικά αποτελέσματα από την εφαρμογή του προτεινόμενου αλγορίθμου σε διάφορες εικόνες δείχνουν την αποδοτικότητά του σε μία ποικιλία περιπτώσεων.

Επόπτης: **Στέλιος Ορφανουδάκης,**
Καθηγητής Επιστήμης Υπολογιστών,
Πανεπιστήμιο Κρήτης.

Iterative graph based algorithm for colored images using non parametric statistical tests.

Flouri Kallirro

Master of Science Thesis

Department of Computer Science

University of Crete

ABSTRACT

One very important problem in image analysis is image segmentation. The goal of image segmentation is to locate and isolate regions of the image that correspond to real objects. Higher level vision tasks such as object recognition, image retrieval based on pictorial content, etc, in many cases require a good segmentation in order to warrantee reliable results. Many techniques for image segmentation have been proposed, usually making assumptions on the statistical properties of the features on which segmentation is based and/or assumptions on the number of regions of the image after the segmentation. We propose an iterative graph based algorithm for segmenting images based on color information. The affinity measure is determined by a non parametric statistical test (Mann Whitney) making no assumptions about the distribution of color in the image. The proposed algorithm can be applied to any colored image and for every color model without a priori knowledge regarding the number of image segments. The experimental results from the application of the segmentation algorithm on several images show its performance in a variety of cases.

Supervisor: Stelios Orphanoudakis

Professor of Computer Science

University of Crete

Ευχαριστίες

... και κάπως έτσι τελειώνει αυτός ο κύκλος των μεταπτυχιακών μου σπουδών. Μετά από πολλά ξενύχτια πότε στο Πανεπιστήμιο Κρήτης (στα γνωστά "υπόγεια") πότε στο εργαστήριο Υπολογιστικής Όρασης και Ρομποτικής του Ι.Τ.Ε, περνώντας ατέλειωτες ώρες μπροστά από τον υπολογιστή πέρασαν δύο (για να είμαι ειλικρινής σχεδόν τρία) χρόνια της ζωής μου.

Αρχικά θα ήθελα να ευχαριστήσω τον επόπτη καθηγητή μου κ. Στέλιο Ορφανουδάκη για την επίβλεψη της μεταπτυχιακής μου εργασίας. Ο κ. Ορφανουδάκης αποτελεί στη συνείδηση μου πρότυπο επιστήμονα και ανθρώπου. Με τις ιδέες του και την καθοδήγηση του με βοήθησε να διευρύνω τον τρόπο σκέψης μου και να διαμορφώσω έναν ιδιαίτερο τρόπο αντίληψης που θα με συνοδεύει σε κάθε πτυχή της ζωής μου.

Περαιτέρω, θα ήθελα να ευχαριστήσω τον ερευνητή του Ιδρύματος Τεχνολογίας και Έρευνας (Ι.Τ.Ε) κ. Αντώνη Αργυρό με τον οποίο συνεργάστηκα πολύ στενά. Οι υποδείξεις και οι συμβουλές του ήταν καθοριστικές σε όλες τις αποφάσεις που πάρθηκαν στα πλαίσια της παρούσας εργασίας. Η βοήθειά του ήταν πολύτιμη για την επιτυχή ολοκλήρωση της μεταπτυχιακής μου εργασίας. Επίσης, θα ήθελα να ευχαριστήσω τον καθηγητή κ. Παναγιώτη Τσακαλίδη όπου στα πλαίσια ευρύτερης συνεργασίας μας στο Ι.Τ.Ε, μου έδειξε πως η επιμονή, υπομονή αλλά και το μεράκι είναι απαραίτητα συστατικά για μια επιτυχημένη επαγγελματική πορεία.

Στο σημείο αυτό θα ήθελα να ευχαριστήσω όλους τους μεταπτυχιακούς φοιτητές του τμήματος που με βοήθησαν να καταλάβω τον περίεργο κόσμο της επιστήμης υπολογιστών που σε εμένα ως μαθηματικός -στην αρχή τουλάχιστον, ήταν σχεδόν ακαταλαβίστικος. Απλά και μόνο η παρουσία τους στους διπλανούς υπολογιστές μου έδινε δύναμη και κουράγιο να βγάλω εις πέρας τις υποχρεώσεις μου ως μεταπτυχιακή φοιτήτρια του τμήματος. Ειδικότερα θέλω να ευχαριστήσω τους φίλους μου Γιάννη Μουστάκα (γνωστός και ως jmous), Σταμάτη Καρβουναράκη (karvouh) που με υπομονή -και τσακωμούς ενίοτε- περάσαμε πολλές ώρες παρέα κάνοντας τις εργασίες των μαθημάτων. Με ανέχτηκαν και τους ανέχτηκα σε όλες τις στιγμές άγχους, χαράς, προβληματισμού, αλλά και διασκέδασης και ανεμελιάς.

Κλείνοντας αυτό το σύντομο σημείωμα θα ήθελα να ευχαριστήσω το Ίδρυμα Τεχνολογίας και Έρευνας που μου παρείχε την οικονομική υποστήριξη και την υλικοτεχνική υποδομή κατά τη διάρκεια των μεταπτυχιακών μου σπουδών. Στον ιδιαίτερο χώρο του εργαστηρίου Υπολογιστικής Όρασης και Ρομποτικής συνάντησα άτομα με κοινά ενδιαφέροντα τα οποία επέδειξαν μεγάλη προθυμία να μοιραστούν τις ερευνητικές προκλήσεις της προσωπικής μου εργασίας. Ειδικότερα, θα ήθελα να ευχαριστήσω τα παιδιά στο εργαστήριο Γιώργο Καζαζάκη, Μιχάλη Μανιαδάκη, Μιχάλη Σφακιωτάκη, Κώστα Χατζόπουλο που έριξαν και αυτοί λίγο φως στο δρόμο των σπουδών μου όταν τους χρειαζόμουν.

Το μεγαλύτερο ευχαριστώ από τα βάθη της καρδιάς μου το οφείλω στους γονείς μου Ζωή και Μανόλη.

... και κάπως έτσι τελειώνει αυτός ο κύκλος των μεταπτυχιακών μου σπουδών.

1	ΕΙΣΑΓΩΓΗ	7
1.1	Ορισμός προβλήματος	7
1.2	Είδη τμηματοποίησης	10
1.2.1	Τεχνικές ιστογράμματος	10
1.2.2	Τεχνικές εκτίμησης μοντέλου Markov ή Gibbs	12
1.2.3	Τεχνικές ακμών	13
1.2.4	Τεχνικές περιοχών	14
1.2.5	Τεχνικές βασισμένες σε γράφους	17
1.2.6	Υβριδικές τεχνικές	22
1.3	Η προτεινόμενη μέθοδος τμηματοποίησης	25
2	ΒΑΣΙΚΕΣ ΕΝΝΟΙΕΣ	28
2.1	Αναπαράσταση του χρώματος στην επεξεργασία εικόνων	28
2.1.1	RGB	29
2.1.2	HSV	29
2.1.3	YUV	30
2.2	Στατιστικές δοκιμασίες	32
2.2.1	Παραμετρικές δοκιμασίες	32
2.2.2	Μη παραμετρικές δοκιμασίες	35
3	ΑΝΑΛΥΣΗ ΜΕΘΟΔΟΥ	39
3.1	Προεπεξεργασία εικόνας	39
3.1.1	Ανισοτροπική Εξομάλυνση (Anisotropic diffusion)	40
3.1.2	Υπερτμηματοποίηση (Over-segmentation)	44
3.1.3	Εφαρμογή κατωφλιού (size thresholding)	49
3.2	Επεξεργασία δεδομένων της εικόνας με τη δομή γράφου	54
3.2.1	Κατασκευή γράφου γειτονικών περιοχών	54
3.2.2	Καθορισμός μεγέθους δείγματος σε κάθε περιοχή	56
3.2.3	Υπολογισμός βαρών του γράφου με την δοκιμασία Mann-Whitney	57
3.2.4	Επαναληπτικός αλγόριθμος ένωσης κόμβων του γράφου	58
3.2.5	Ανανέωση της δομής του γράφου μετά την ένωση δύο περιοχών	61
4	ΑΠΟΤΕΛΕΣΜΑΤΑ	65
4.1	Σύγκριση αποτελεσμάτων ως προς το χρωματικό μοντέλο που χρησιμοποιήθηκε	66
4.2	Σύγκριση αποτελεσμάτων για διάφορα μεγέθη του δείγματος	72
5	ΣΥΜΠΕΡΑΣΜΑΤΑ	90
6	ΜΕΛΛΟΝΤΙΚΕΣ ΕΠΕΚΤΑΣΕΙΣ	92
7	ΑΝΑΦΟΡΕΣ	93

1 Εισαγωγή

1.1 Ορισμός προβλήματος

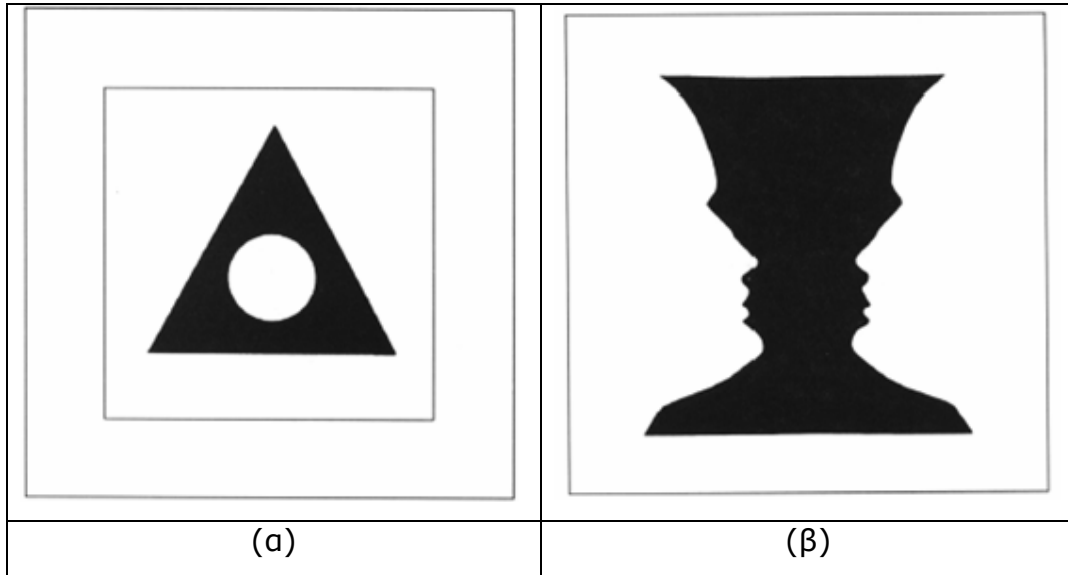
Η τμηματοποίηση αποτελεί ένα πολύ σημαντικό πρόβλημα στην επεξεργασία και ανάλυση εικόνων. Με τον όρο τμηματοποίηση εικόνων ονομάζουμε τον διαχωρισμό μιας εικόνας σε περιοχές, έτσι ώστε κάθε μια από αυτές να αντιστοιχεί σε ένα φυσικό αντικείμενο. Η τμηματοποίηση της σε k περιοχές R_1, R_2, \dots, R_k πρέπει να είναι τέτοια ώστε η ένωση όλων των τμημάτων να αποτελούν ολόκληρη την εικόνα όπως φαίνεται στη σχέση (1), αλλά και κάθε σημείο της εικόνας να ανήκει σε μια μόνο περιοχή, σχέση (2).

$$\bigcup_{k=1}^K R_k = R \quad (1)$$

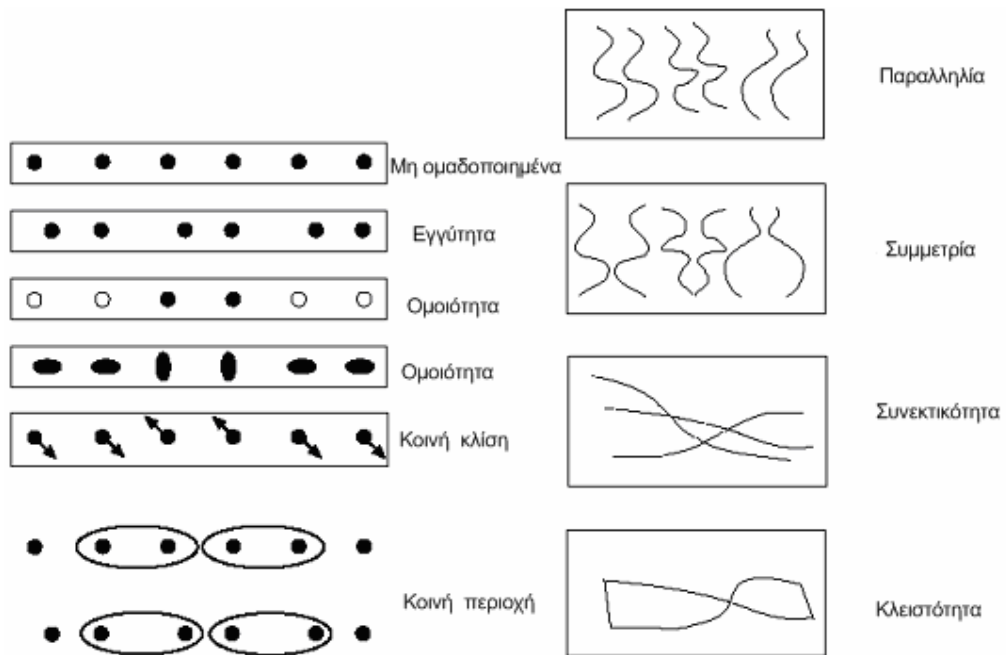
$$R_i \cap R_j = \emptyset, \quad i \neq j \quad (2)$$

Η τμηματοποίηση μιας εικόνας μπορεί να γίνει με βάση διάφορα κριτήρια όπως το χρώμα, την υφή, την φωτεινότητα, την κίνηση όταν πρόκειται για *video*, ή κάποιο συνδυασμό των παραπάνω χαρακτηριστικών.

Στο σημείο αυτό πρέπει να αναφερθεί η υποκειμενικότητα στο θέμα της τμηματοποίησης. Στο κλάδο της Ψυχολογίας έχουν γίνει πολλές μελέτες η επικρατέστερη των οποίων (θεωρία Gestalt) βασίζεται στον διαχωρισμό μιας εικόνας στις επικρατέστερες φιγούρες ανάλογα με το προσδιορισμό του φόντου [1]. Είναι δυνατό όπως φαίνεται στο Σχήμα 1 ανάλογα με το τι θεωρεί το ανθρώπινο μάτι φόντο, η εξαγωγή των πιο σημαντικών αντικειμένων της εικόνας να διαφέρει. Εάν θεωρηθεί το άσπρο χρώμα ως φόντο, τότε περιοχή ενδιαφέροντος στην (α) εικόνα αποτελεί το μαύρο τρίγωνο και στην (β) εικόνα το βάζο, ενώ εάν θεωρηθεί το μαύρο χρώμα ως φόντο τότε στην (α) περιοχή ενδιαφέροντος αποτελεί ο άσπρος κύκλος και στην (β) εικόνα τα δύο πρόσωπο. Κάθε άνθρωπος επομένως διαχωρίζει ή ομαδοποιεί αντικείμενα στο χώρο με διαφορετικό τρόπο. Επομένως, δεν υπάρχει αντικειμενική αλήθεια για τμηματοποίηση ή διαχωρισμό αντικειμένων. Το γεγονός αυτό συντελεί στο να χαρακτηριστεί η τμηματοποίηση εικόνων ως ένα δύσκολο πρόβλημα.



Σχήμα 1: Φόντο έναντι αντικειμένου



Σχήμα 2: Κριτήρια ομοιότητας Gestalt [2]

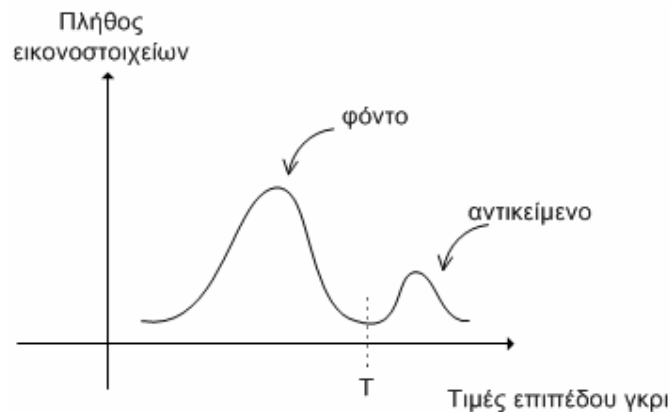
Από την άλλη πλευρά, οι εφαρμογές της τμηματοποίησης εικόνων είναι ιδιαίτερα χρήσιμες όχι μόνο στον χώρο της επεξεργασίας εικόνων για την εξαγωγή συγκεκριμένων αντικειμένων από μια εικόνα αλλά και σε άλλους τομείς. Έτσι, σε συστήματα διαχείρισης εικόνων, το αποτέλεσμα της ανάκτησης εικόνων με βάση το περιεχόμενό τους επηρεάζεται από την τμηματοποίηση κάθε εικόνας του συστήματος σε ένα ενδιάμεσο στάδιο πριν την ανάκτηση. Επιπλέον, ακόμα και στον χώρο της ιατρικής, με την τμηματοποίηση ιατρικών εικόνων γίνεται καλύτερη ανάλυση εικόνων από μικροσκόπιο, υπερηχογραφήματων, ακόμα και διάγνωση ασθενειών [2]. Γενικά, μπορεί κανείς να διαπιστώσει ότι διεργασίες υψηλού επιπέδου βασίζονται σε πολλές περιπτώσεις στην αξιόπιστη τμηματοποίηση εικόνων.

Εν κατακλείδι, παρόλο που η τμηματοποίηση εικόνων χαρακτηρίζεται ως ένα δύσκολο πρόβλημα, η χρησιμότητα των εφαρμογών της συντέλεσαν στην εκτενή μελέτη του προβλήματος με αποτέλεσμα την ύπαρξη πολλών προσεγγίσεων για τμηματοποίηση εικόνων. Στην επόμενη ενότητα θα αναφερθούν και θα περιγραφούν σύντομα οι πιο χαρακτηριστικές τεχνικές.

1.2 Είδη τμηματοποίησης

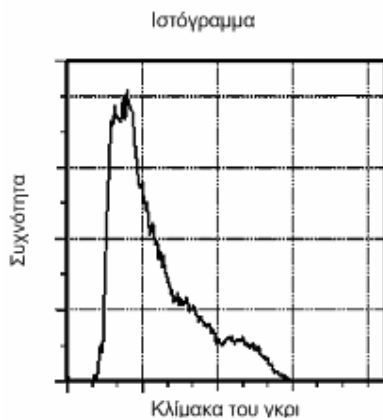
1.2.1 Τεχνικές ιστογράμματος

Οι τεχνικές ιστογράμματος (ή κατωφλιού) λόγω της απλότητάς τους χρησιμοποιήθηκαν πρώτες και αποδείχθηκαν αποτελεσματικές κυρίως για απλές εικόνες. Τέτοιες τεχνικές βασίζονται στην υπόθεση ότι γειτονικά εικονοστοιχεία μιας εικόνας με παρόμοια τιμή (είτε φωτεινότητας, είτε χρώματος, είτε υψής) ανήκουν στο ίδιο αντικείμενο μιας εικόνας [3]. Για παράδειγμα, εάν μια εικόνα αποτελείται από ένα αντικείμενο το οποίο παρουσιάζει ομοιογένεια στα επίπεδα του γκρι είναι πολύ απλό να διαχωριστεί από το φόντο διαφορετικής απόχρωσης του γκρι όπως φαίνεται στο Σχήμα 3.



Σχήμα 3: Ιστόγραμμα

Η εικόνα που αντιστοιχεί στο ιστογράμμο του Σχήματος 3 μπορεί να χωριστεί σε δύο περιοχές πολύ εύκολα με τη μέθοδο αυτή. Τα εικονοστοιχεία τα οποία έχουν τιμή επιπέδου γκρι μεγαλύτερη από το T αποτελούν το αντικείμενο, ενώ τα αυτά με τιμή επιπέδου γκρι μικρότερη από το T ανήκουν στο φόντο (background) της εικόνας. Η μέθοδος αυτή μπορεί να επεκταθεί και για περισσότερες από δύο τελικές αποχρώσεις όπως φαίνεται στο Σχήμα 4, [4].



Σχήμα 4: Το ιστόγραμμα της αρχικής εικόνας και δεξιά το αποτέλεσμα της τμηματοποίησης με χρήση 3 αποχρώσεων

Το ιστόγραμμα πρώτης τάξης αποτελεί την δειγματική πυκνότητα πιθανότητας μίξης κατανομών. Επιδιώκεται η εκτίμηση των παραμέτρων της μίξης. Δοσμένων των παραμέτρων αυτών, ορίζονται κάποια κατώφλια (Bayes classifier) με βάση τα οποία τα στοιχεία της εικόνας κατατάσσονται στις κλάσεις μίξης. Το ιστόγραμμα είναι δυνατό να κατασκευαστεί με βάση όλη την εικόνα ή κάποιες υποπεριοχές αυτής. Στην τελευταία περίπτωση για κάθε υποπεριοχή κατασκευάζεται το ιστόγραμμα και ακολούθως υπολογίζονται τα κατώφλια.

Για την καλή εκτίμηση των παραμέτρων της μίξης απαιτούνται συνθήκες που σπάνια ικανοποιούνται από πραγματικές εικόνες (αξιόλογο μέγεθος αντικειμένων και μεγάλος σηματοθορυβικός λόγος). Ένα ακόμα μειονέκτημα των τεχνικών αυτών είναι ότι το πλήθος των αντικειμένων πρέπει να είναι γνωστό. Πρέπει δηλαδή ο χρήστης να δώσει μια τιμή για την μεταβλητή που χαρακτηρίζει

το πλήθος των τελικών τμημάτων της εικόνας, πριν ξεκινήσει η διαδικασία της τμηματοποίησης. Τέλος, ένα ακόμα χαρακτηριστικό των τεχνικών αυτών είναι ότι δεν χρησιμοποιείται η διάταξη των αντικειμένων στο χώρο. Έτσι, είναι δυνατό αντικείμενα που δεν έχουν καμία σχέση όσον αφορά την διάταξή τους στο χώρο αλλά έχουν παρόμοιες φωτεινότητες να εκληφθούν ως μια κλάση και έτσι να μεταβληθούν σημαντικά τα αποτελέσματα. Τεχνικές οι οποίες προσπαθούν να εκμεταλλευτούν τη διάταξη στο χώρο έχουν προταθεί [5] με σχετικά ικανοποιητικά αποτελέσματα, αλλά σε περιορισμένο πεδίο εφαρμογής.

1.2.2 Τεχνικές εκτίμησης μοντέλου Markov ή Gibbs

Σε αντιδιαστολή με την απλή ανάλυση του ιστογράμματος μιας εικόνας, η ανάλυση με βάση ένα μοντέλο Markov για το πεδίο των αποφάσεων κατάταξης των σημείων σε τμήματα χρησιμοποιεί τις αλληλεξαρτήσεις που υπάρχουν τοπικά σε γειτονικά σημεία. Για κάθε μοντέλο Markov που δίνει τοπικές αλληλεξαρτήσεις ορίζεται ισοδύναμα μια κατανομή Gibbs που δίνει συνολικά όλες τις εξαρτήσεις του πεδίου των αποφάσεων. Αν $L = \{(m, n) : 0 \leq m \leq M, 0 \leq n \leq N\}$ είναι το σύνολο των σημείων της εικόνας τότε το πεδίο Gibbs ορίζεται ως εξής :

$$P_r \{X(m, n) = x(m, n) : (m, n) \in L\} = \frac{e^{-U(x)}}{\psi} \quad (3)$$

όπου $U(x)$ ονομάζεται συνάρτηση ενέργειας, και είναι συνάρτηση MN μεταβλητών και $\psi = \sum_{c \in C} e^{-U(x)}$ σταθερά.

Οι τεχνικές αυτές υποθέτουν ότι η πραγματική εικόνα αποτελεί την πραγματοποίηση ενός Μαρκοβιανού πεδίου το οποίο παραμορφώνεται από θόρυβο και αναζητείται η εικόνα με τη μέγιστη πιθανότητα δοσμένης της παρατηρούμενης. Το πρόβλημα μεγιστοποίησης που προκύπτει είναι συνήθως μη γραμμικό και το πλήθος των τοπικών μεγίστων μεγάλο. Το κυριότερο μειονέκτημα τέτοιων προβλημάτων μεγιστοποίησης είναι η αργή σύγκλιση των αλγορίθμων. Για το λόγο αυτό έχουν προταθεί διάφορες μέθοδοι εκτίμησης όπως η μέθοδος εκτίμησης του μέσου πεδίου (mean field) που αποτελεί εκτιμητή ελάχιστης διασποράς [6,7] , και μια μέθοδος που μεγιστοποιεί την περιθωριακή πιθανότητα των στοιχείων, δοσμένων της εκτίμησης των γειτόνων τους και των παρατηρούμενων τιμών.

1.2.3 Τεχνικές ακμών

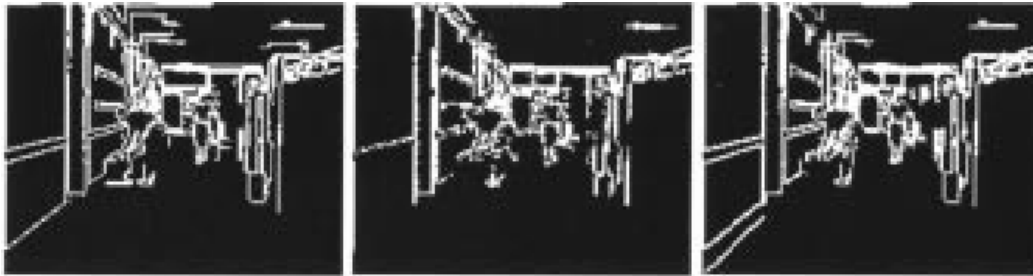
Ως σημείο ακμής ορίζεται εκείνο το σημείο όπου η μεταβολή της φωτεινής έντασης είναι τοπικά μέγιστη. Ένα σύνολο συνδεδεμένων σημείων ακμής αποτελεί ένα περίγραμμα. Σε κάθε σημείο ενός περιγράμματος μπορεί να ορισθεί ο προσανατολισμός του σε σχέση με το σύστημα συντεταγμένων του επιπέδου της εικόνας. Η φωτεινή ένταση παρουσιάζει την μέγιστη μεταβολή στην κάθετη στο περίγραμμα κατεύθυνση [8,9]. Η ανίχνευση των σημείων ακμής μπορεί να γίνει με τη βοήθεια είτε του διανύσματος των δύο πρώτων μερικών παραγώγων της φωτεινής έντασης (κλίση), είτε του αθροίσματος των δύο δεύτερων μερικών παραγώγων της φωτεινής έντασης (Laplacian τελεστής).

Με τις τεχνικές αυτές αρχικά ανιχνεύονται οι ακμές της εικόνας και ακολούθως ομαδοποιούνται σε συνεχή και κλειστά περιγράμματα. Η ανίχνευση ακμών βασίζεται στην προσέγγιση της παραγώγου όπως αναφέρθηκε παραπάνω και πραγματοποιείται με διάφορα φίλτρα (Roberts, Sobel, Canny). Μερικά παραδείγματα ανίχνευσης ακμών με τη βοήθεια κατάλληλων φίλτρων φαίνονται στο Σχήμα 5. Ένα απλό φίλτρο Roberts όμως δεν θεωρείται κατάλληλο διότι ενισχύει τον θόρυβο (αύξηση διασποράς με αποτέλεσμα την μείωση του σηματοθορυβικού λόγου) [10]. Επομένως ένα «καλό» φίλτρο υπολογισμού της παραγώγου θα πρέπει να αυξάνει τον σηματοθορυβικό λόγο ώστε να γίνει πιο ακριβής η ανίχνευση. Επίσης, η υπολογιζόμενη παράγωγος θα πρέπει να λαμβάνει τη μέγιστη τιμή της κοντά στη θέση της ασυνέχειας, ενώ ταυτόχρονα δεν θα πρέπει να ανιχνεύονται πολλά μέγιστα στην περιοχή της ασυνέχειας. Με βάση τα παραπάνω κριτήρια, αποδεικνύεται πως το βέλτιστο φίλτρο προσεγγίζεται καλά από τις παραγώγους του φίλτρου Gauss.

Τέλος, παρόλο που οι τεχνικές αυτές επιτυγχάνουν την ορθή ανίχνευση μεγάλου ποσοστού πραγματικών ακμών, αδυνατούν να επιτύχουν ικανοποιητική ομαδοποίηση των ακμών σε περιγράμματα υπό την έννοια ότι δεν εγγυώνται συνεχή και κλειστά περιγράμματα. Υπάρχουν βέβαια διεργασίες χαμηλού επιπέδου (edge tracking, gap filling, smoothing, thinning) που πραγματοποιούν τέτοια περιγράμματα, αλλά απαιτούν πολύ υπολογιστικό χρόνο σε σχέση με το είδος της διεργασίας που εκτελούν.



(α)



(β)

(γ)

(γ)

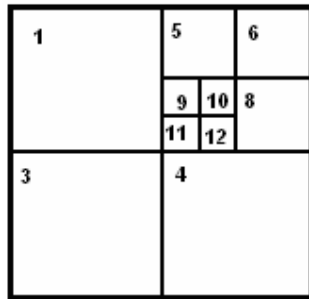
Σχήμα 5: (α) Αρχική εικόνα, (β) Ανίχνευση ακμών με φίλτρο Sobel, (γ) Ανίχνευση ακμών με φίλτρο Laplacian, και (γ) Ανίχνευση ακμών με φίλτρο Roberts.

1.2.4 Τεχνικές περιοχών

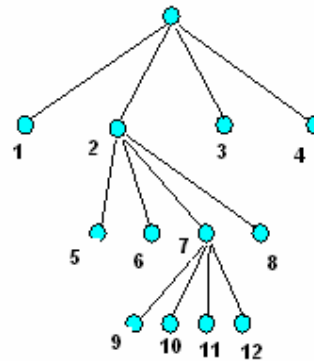
Οι τεχνικές αυτές βασίζονται στην ομοιότητα (ως προς την φωτεινότητα, ή το χρώμα ή την υφή ή κάποιο άλλο κριτήριο) γειτονικών εικονοστοιχείων που ανήκουν στην ίδια περιοχή [5]. Αν R είναι ολόκληρη η εικόνα ως σύνολο σημείων, ζητείται η τμηματοποίηση σε συνεκτικές περιοχές R_1, R_2, \dots, R_k ώστε $\bigcup_{k=1}^K R_k = R$ και $R_i \cap R_j = \emptyset, i \neq j$. Έχουν προταθεί τρεις σημαντικότερες τεχνικές [5], διάσπαση και ένωση περιοχών (split and merge), ιεραρχική ένωση περιοχών (bottom up) και η τεχνική της συσσώρευσης γύρω από κάποια περιοχή (region growing).

Στην πρώτη τεχνική που αναφέρθηκε (split and merge [5]) όλη η εικόνα αρχικά θεωρείται ως μια περιοχή. Καθορίζεται επίσης ένα κριτήριο ομοιότητας σύμφωνα με το οποίο γίνεται η σύγκριση των περιοχών. Κάθε τμήμα χωρίζεται συνήθως σε τέσσερα ορθογώνια τμήματα χωρίς να λαμβάνεται υπόψη η δομή της εικόνας στο τμήμα αυτό. Η δομή που χρησιμεύει σε τέτοιου είδους αλγορίθμους

είναι τα δέντρα, όπου κάθε φύλλο του δέντρου αποτελεί ένα ορθογώνιο τμήμα της εικόνας (Σχήμα 6).



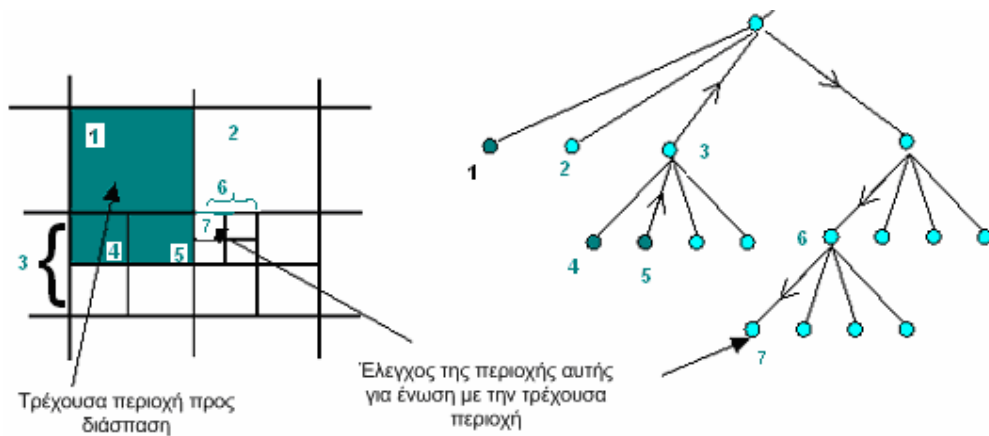
(α)



(β)

Σχήμα 6:(α) Η διαδοχική διάσπαση εικόνας σε τέσσερα τμήματα κατά την μέθοδο split and merge, (β) αναπαράσταση της διάσπασης με δέντρο όπου κάθε φύλλο του δέντρου αποτελεί το αντίστοιχο τμήμα της εικόνας

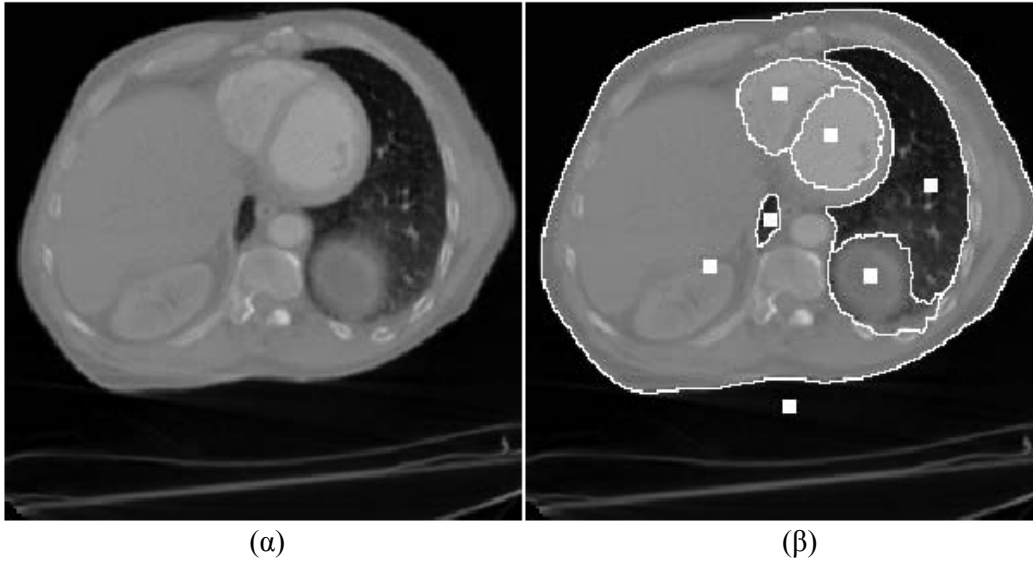
Όσο το κριτήριο ομοιογένειας δεν ικανοποιείται η εικόνα διαμερίζεται και στα νέα τμήματα που δημιουργήθηκαν εφαρμόζεται η ίδια τεχνική. Μόλις τελειώσει η διαίρεση αυτή αρχίζει μια διαδικασία ένωσης γειτονικών τμημάτων με βάση την ομοιότητά τους. Εφόσον η διαίρεση όπως ήδη αναφέρθηκε πραγματοποιείται χωρίς να λαμβάνεται υπόψη η δομή (πέραν της ομοιογένειας ή ανομοιογένειας), παραμορφώνονται σε μεγάλο βαθμό τα περιγράμματα των αντικειμένων της εικόνας. Στο Σχήμα 7 φαίνεται η συγκεκριμένη τεχνική για ένα μέρος της εικόνας. Η εικόνα αρχικά διασπάστηκε (split) σε τέσσερα τμήματα όπως περιγράφηκε παραπάνω. Εφαρμόζεται το κριτήριο ομοιογένειας για κάθε τέτοιο τμήμα. Όταν δυο τμήματα μοιάζουν τότε αρχίζει η ένωση τους (merge). Για παράδειγμα, έστω ότι το τμήμα με την ετικέτα 5 μοιάζει με το τμήμα 1, τότε δεν διασπάται το τμήμα σε μικρότερα τμήματα και ενώνεται με αυτό (merge). Όμως το τμήμα 6 δεν μοιάζει με το 1 οπότε διασπάται πάλι σε τέσσερα τμήματα (split).



Σχήμα 7: Η τεχνική της διάσπασης και ένωσης εικόνας (split and merge) με αναπαράσταση δέντρου. Εάν το τμήμα 7 μοιάζει με το 1 τα 1,4,5,7 θα αποτελούν πλέον μια περιοχή, διαφορετικά θα διασπαστεί σε άλλα τέσσερα τμήματα.

Κατά την ιεραρχική ένωση περιοχών (bottom up [5]), αρχικά απαιτείται μια καλή διαμέριση της εικόνας, στην οποία κάθε περιοχή είναι ομοιογενής. Έπειτα επιχειρείται η μίξη των ομοιότερων γειτονικών περιοχών. Το κυριότερο πρόβλημα με την τεχνική αυτή είναι η παραγωγή μιας καλής αρχικής διαμέρισης της εικόνας, γιατί τα λάθη που προκύπτουν από το στάδιο αυτό δεν είναι δυνατό να διορθωθούν.

Τέλος, η τμηματοποίηση μιας εικόνας είναι εφικτή με συσσωρεύσεις γύρω από κάποιο ή κάποια σημεία (region growing [9]). Επιλέγονται δηλαδή κάποια αρχικά σημεία τα οποία είναι αντιπροσωπευτικά των σημαντικών περιοχών της εικόνας (Σχήμα 8). Στη συνέχεια, τα σημεία αυτά ομαδοποιούνται (επαναληπτικά) με τα γειτονικά τους με βάση κάποια κριτήρια ομοιότητας, έως ότου όλα τα σημεία της εικόνας ενταχθούν σε κάποια περιοχή. Και στην μέθοδο αυτή, το πρόβλημα είναι η επιλογή των αντιπροσωπευτικών σημείων που συνήθως γίνεται διαλογικά.



Σχήμα 8: (α) Τομογραφική εικόνα της καρδιάς, (β) επιλογή των αντιπροσωπευτικών περιοχών (region) από τον χρήστη

1.2.5 Τεχνικές βασισμένες σε γράφους

Μια πολύ σημαντική τεχνική για την λύση του προβλήματος της τμηματοποίησης μιας εικόνας αποτελεί η αντιμετώπισή του ως πρόβλημα διάσπασης ή σύνδεσης γράφου [31,11, 12]].

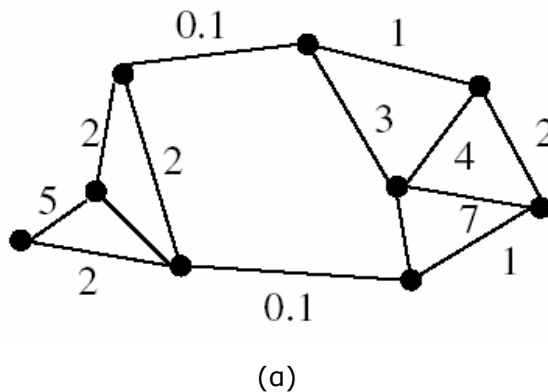
Γράφος G , είναι ένα σύνολο το οποίο αποτελείται από κορυφές V (vertices) και ακμές E (edges) δηλαδή $G = \{V, E\}$. Κάθε ακμή ενώνει ένα ζεύγος κορυφών και ισχύει $E \subset V \times V$. Δύο κορυφές ενός γράφου ενώνονται όταν υπάρχει μια σειρά ακμών με αφετηρία την μια κορυφή του γράφου και τέλος την άλλη. Σε κάθε ακμή ενός ζεύγους κορυφών v_i, v_j αντιστοιχεί μια αριθμητική τιμή $w(i,j)$ που ονομάζεται βάρος του γράφου (weighted graph) και υποδηλώνει την ομοιότητα των κορυφών i, j . Όταν ο γράφος είναι κατευθυνόμενος, (Σχήμα 9(α)) τότε το βάρος στις κορυφές A, B είναι $w(A,B)$ και η κορυφή A είναι η αφετηρία και η κορυφή B το τέλος της ακμής. Ενώ όταν ο γράφος είναι μη κατευθυνόμενος (Σχήμα 9(β)) τότε $w(A,B) = w(B,A)$.

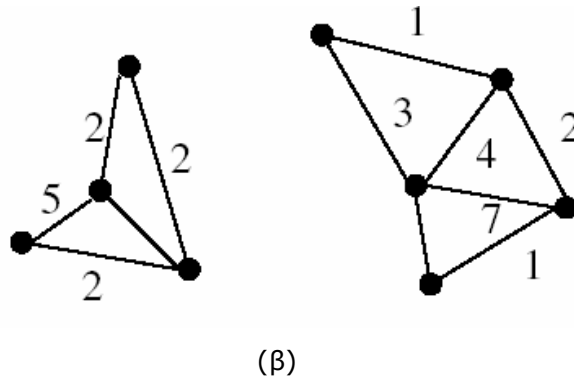


Σχήμα 9: (α) Κατευθυνόμενος γράφος, (β) μη κατευθυνόμενος γράφος.

Στο πρόβλημα τμηματοποίησης εικόνας με γράφο τα εικονοστοιχεία της εικόνας αποτελούν τις κορυφές του γράφου και κάθε ακμή έχει ένα βάρος το οποίο όσο μεγαλύτερο είναι, τόσο ομοιότερα είναι τα δύο εικονοστοιχεία που συνδέει. Κάθε κόμβος του γράφου μπορεί να αποτελείται από πολλά εικονοστοιχεία τα οποία αποτελούν μια περιοχή της εικόνας. Τα τμήματα αυτά ονομάζονται συνεκτικά (connected components). Έστω $G_i = \{V_i \cup E_i\}$ ένα συνεκτικό τμήμα του γράφου $G = \{V, E\}$. Το τμήμα αυτό θεωρείται υπογράφος όπου όλες του οι κορυφές ενώνονται, αλλά δεν υπάρχει ακμή στο E που να ενώνει κάποιο στοιχείο του V_i με στοιχείο του V_j με $i \neq j$. Έτσι αν στον αρχικό γράφο υπάρχουν n συνεκτικά τμήματα τότε $G' = \{V_1 \cup V_2 \dots V_n, E_1 \cup E_2 \dots E_n\}$, [1].

Σε μια τέτοια αντιμετώπιση του προβλήματος της τμηματοποίησης μιας εικόνας ο στόχος είναι η διάσπαση του γράφου σε τόσα συνεκτικά τμήματα όσες είναι και τα φυσικά αντικείμενα της εικόνας. Εάν ο γράφος διασπαστεί στις ακμές όπου η τιμή του βάρους είναι μικρή, τότε τα συνεκτικά τμήματα που δημιουργούνται αποτελούνται από κόμβους με μεγάλη ομοιότητα, Σχήμα 10.



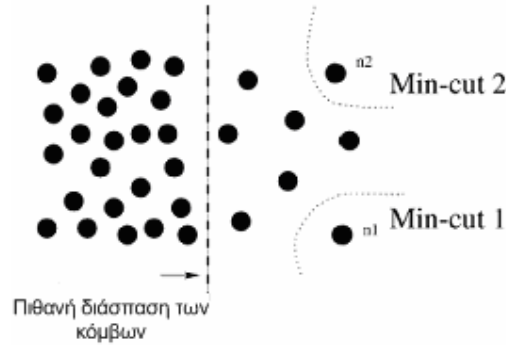


Σχήμα 10: (α) Μη κατευθυνόμενος γράφος με βάρη που χαρακτηρίζουν την ομοιότητα ανάμεσα στους κόμβους, (β) Δημιουργία συνεκτικών τμημάτων από διάσπαση ακμών του γράφου με μικρό βάρος.

Χαρακτηριστικό παράδειγμα αυτής της τεχνικής αποτελεί η μέθοδος [13] που έχει προταθεί από τους Jianbo Shi και Jitendra Malik που βασίζεται σε ένα γραφοθεωρητικό κριτήριο- *normalized cut*. Αν $G=(V,E)$ ο γράφος, μπορεί να διασπαστεί σε δύο ανεξάρτητα σύνολα A,B όπου $A \cup B = V, A \cap B = \emptyset$ αν διαγραφεί η ακμή που ενώνει δύο ακμές. Ο βαθμός ανομοιογένειας των συνόλων A,B που προκύπτουν δίνεται από το άθροισμα των βαρών w μεταξύ δύο κορυφών που έχουν διαγραφεί από τον γράφο.

$$cut(A,B) = \sum_{u \in A, v \in B} w(u,v) \quad (4)$$

Έτσι για τον καλύτερο διαχωρισμό του γράφου σε δύο μέρη (bi-partition) απαιτείται η ελαχιστοποίηση του παραπάνω κριτηρίου. Υπάρχουν όμως περιπτώσεις όπου η ελαχιστοποίηση κριτηρίου δεν συντελεί σε μια «καλή» τμηματοποίηση γιατί είναι δυνατό να απομονωθούν στο ελάχιστο κάποιοι κόμβοι του γράφου. Πιο αναλυτικά, η ελαχιστοποίηση του κριτηρίου επιτυγχάνεται με την διάσπαση μικρών συνόλων από απομονωμένους κόμβους του γράφου. Εξ' ορισμού η τιμή του $cut(A,B)$ μεταβάλλεται ανάλογα με τον αριθμό των ακμών ανάμεσα στα δύο μέρη όπου πραγματοποιήθηκε η διάσπαση. Στο Σχήμα 11 παρουσιάζεται μια τέτοια περίπτωση. Στο παράδειγμα αυτό εάν υποθέσουμε ότι το βάρος των ακμών είναι αντιστρόφως ανάλογο με την απόσταση των δύο κόμβων, τότε με την διάσπαση του n_1 ή n_2 το κριτήριο έχει πολύ μικρή τιμή. Ενώ εάν η διάσπαση πραγματοποιούνταν σε κάποιο άλλο πιθανό σημείο όπως φαίνεται στο σχήμα, το αποτέλεσμα είναι πιο αντιπροσωπευτικό χωρίς να ελαχιστοποιείται το κριτήριο.



Σχήμα 11: Περίπτωση όπου η ελαχιστοποίηση του κριτηρίου $cut(A,B)$ επιτυγχάνεται με μια όχι τόσο καλή διάσπαση του γράφου, ενώ σε μια άλλη πιθανή διάσπαση το αποτέλεσμα είναι πιο αντιπροσωπευτικό χωρίς να ελαχιστοποιείται το κριτήριο.

Το πρόβλημα αυτό λύνεται με την κανονικοποίηση του $cut(A,B)$, οπότε τότε το κριτήριο ονομάζεται *normalized cut* και δίνεται από τον τύπο

$$Ncut(A, B) = \frac{cut(A, B)}{asso(A, V)} + \frac{cut(A, B)}{asso(B, V)} \quad (5)$$

όπου $asso(A, V) = \sum_{u \in A, t \in V} w(u, t)$ είναι το άθροισμα όλων των βαρών από τις κορυφές του συνόλου A προς όλες τις κορυφές του γράφου, όμοια και το $asso(A, V)$.

Μια ακόμα προσέγγιση στο πρόβλημα της τμηματοποίησης μιας εικόνας με χρήση γράφων έχει προταθεί [14] στην οποία γίνεται χρήση πολλών κριτηρίων ομοιογένειας όπως χρώμα, φωτεινότητα, υφή, σχήμα σύμφωνα με τα οποία δημιουργούνται σχετικοί πίνακες (affinity matrices). Ο συνδυασμός των πινάκων αυτών με χρήση PCA (*Principal Components Analysis*) δίνει την τελική επιλογή του χαρακτηριστικού για την σύνδεση δυο κορυφών του γράφου, και επομένως και το βάρος στην αντίστοιχη ακμή.

Πιο αναλυτικά ορίζεται ένας γράφος $G = (R, E)$ όπου R είναι οι κορυφές του οι οποίες συμβολίζουν τα εικονοστοιχεία της εικόνας και E οι ακμές που συμβολίζουν την ομοιότητα των κορυφών που συνδέουν. Αυτό το μέτρο ομοιότητας είναι το βάρος $w_k(i, j)$ στις ακμές i,j ανάλογα με το κριτήριο κ. Ακόμα, $w_k(i, j) = w_k(j, i)$ και $w_k(i, j) = 0$ αν i,j θεωρούνται ανόμοια ως προς το κριτήριο κ. Για την επιλογή του κριτηρίου σύμφωνα με το οποίο θα δοθεί ένα βάρος στην ακμή του γράφου για δύο συγκεκριμένες κορυφές του, εφαρμόζεται το PCA. Με τη μέθοδο αυτή επιλέγεται το κριτήριο το οποίο δίνει το πιο έντονο χαρακτηριστικό

(δηλαδή το ιδιοδιάνυσμα με την μεγαλύτερη ιδιοτιμή) και έτσι προσεγγίζονται τα δεδομένα με ένα γραμμικό υπόχωρο χρησιμοποιώντας το κριτήριο του μέσου τετραγωνικού σφάλματος [12].

Κάθε κορυφή του γράφου (δηλαδή κάθε τμήμα της εικόνας) χαρακτηρίζεται από τα εξής κριτήρια: υφή, το οποίο αποτελείται από επτά συνιστώσες ομοιογένεια (Hom), αντίθεση (Cont), εντροπία (Ent), συσχέτιση (Cor), τοπική ομοιογένεια (HoLo), κατεύθυνση (Dir), σταθερότητα (Unif), φωτεινότητα (από το χρωματικό μοντέλο Lab), και χρώμα το οποίο αποτελείται από τέσσερις συνιστώσες το ab (από το Lab), το R, το G και το B από το RGB και τέλος ένα χαρακτηριστικό που δίνει το περίγραμμα από την παράγωγο της έντασης IG. Τα διάφορα χρωματικά μοντέλα που χρησιμοποιούνται για την αναπαράσταση του χρώματος περιγράφονται στην ενότητα 2.1. Έτσι το βάρος μιας ακμής εξαρτάται από την απόσταση $D_k(i, j)$ των δύο κορυφών i, j ανάλογα με το κριτήριο k και δίνεται από τον τύπο

$$w_k(i, j) = \exp\left[\frac{-(D_k(i, j))^2}{2\sigma_k^2}\right] \quad \forall k \in \left\{ \begin{array}{l} L, ab, R, G, B, IG, Hom, Cont, \\ Ent, Cor, HoLo, Dir, Unif \end{array} \right\} \quad (6)$$

Η τιμή της $D_k(i, j)$ αλλά και η τιμή της απόκλισης σ_k^2 διαφέρει ανάλογα με το k . Το πεδίο τιμών της $D_k(i, j)$ είναι το $[0, 1]$ και επίσης

$$D_k(i, j) = \begin{cases} 0 & i, j \text{ όμοια} \\ 1 & i, j \text{ διαφορετικά} \end{cases} \quad (7)$$

Πιο αναλυτικά, αν το k αντιπροσωπεύει την φωτεινότητα L από το Lab τότε,

$$D_L(i, j) = \frac{\sqrt{(L(i) - L(j))^2}}{100} \quad (8)$$

Αν το k αντιπροσωπεύει το χρώμα ab του χρωματικού μοντέλου Lab, τότε

$$D_{ab}(i, j) = \frac{\sqrt{(a(i) - a(j))^2 + (b(i) - b(j))^2}}{240} \quad (9)$$

Αν αντιπροσωπεύει τις συνιστώσες του RGB τότε,

$$D_{Col}(i, j) = \frac{\sqrt{(Col(i) - Col(j))^2}}{256} \quad (10)$$

Αν το k αντιπροσωπεύει την παράγωγο της έντασης των δύο κορυφών του γράφου τότε,

$$D_{IG}(i, j) = \frac{\sqrt{(IG(i) - IG(j))^2}}{256} \quad (11)$$

Και τέλος αν η απόσταση των δύο κορυφών υπολογίζεται σύμφωνα με το χαρακτηριστικό της υφής τότε,

$$D_{IG}(i, j) = \sqrt{(X(i) - X(j))^2} \quad \forall X \in \{Hom, Cont, Ent, Cor, HoLo, Dir, Unif\} \quad (12)$$

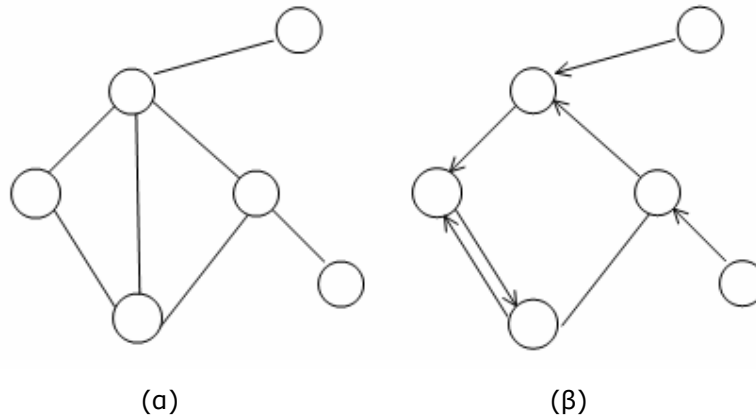
Η αναγωγή του προβλήματος της τμηματοποίησης της εικόνας σε ένα πρόβλημα γράφων έχει ενθαρρυντικά αποτελέσματα όχι μόνο στα σχετικά άρθρα που σύντομα περιλήφθηκαν [14, 15] αλλά και σε άλλες μελέτες [13, 15, 32, 33]. Συμπερασματικά, το πλεονέκτημα στις μεθόδους αυτές είναι ότι παράλληλα με το κριτήριο ομοιότητας που χρησιμοποιείται για την διάσπαση ενός γράφου (το οποίο καθορίζεται κάθε φορά ανάλογα με το χαρακτηριστικό βάσει του οποίου γίνεται η τμηματοποίηση), χρησιμοποιείται και η θέση τους στην εικόνα, οι αποστάσεις μεταξύ τους, και η γειτνίαση όλων των περιοχών.

1.2.6 Υβριδικές τεχνικές

Οι υβριδικές τεχνικές είναι οι πιο δημοφιλείς γιατί συνδυάζουν τις θετικές ιδιότητες των παραπάνω μεθόδων και παράλληλα περιορίζουν τα λάθη που προκύπτουν από την κάθε μέθοδο ξεχωριστά. Πιο αναλυτικά, συνδυάζονται οι τεχνικές ακμών με τις τεχνικές περιοχών έτσι ώστε τα αποτελέσματα της τμηματοποίησης να είναι ακριβέστερα [15, 16, 17].

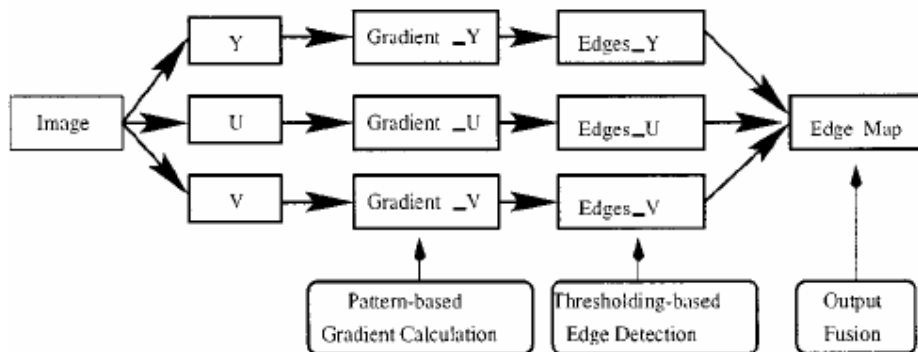
Ένας υβριδικός αλγόριθμος για τμηματοποίηση δυσδιάστατων και τρισδιάστατων εικόνων που έχει προταθεί [18], συνδυάζει τεχνικές που βασίζονται στις ακμές και στις περιοχές μέσω του μορφολογικού αλγόριθμου των διαιρετικών γραμμών. Αρχικά γίνεται μια επεξεργασία της εικόνας ώστε να ελαττωθεί ο θόρυβος, ενώ διατηρούνται οι ακμές. Έπειτα υπολογίζεται η παράγωγος και ανιχνεύονται οι διαιρετικές γραμμές του μέτρου της παραγώγου της εικόνας. Τέλος, γίνεται μια ιεραρχική ένωση περιοχών (bottom-up) για την παραγωγή ικανοποιητικών διαμερίσεων. Η παραδοσιακή μέθοδος υλοποίησης της ιεραρχικής διαδικασίας μίξης περιοχών είναι αυτή που χρησιμοποιεί τον γράφο γειτνίασης περιοχών, όλες οι ακμές του οποίου αποθηκεύονται σε μια ουρά προτεραιότητας (σωρός). Σε κάθε βήμα της διαδικασίας μίξης προσδιορίζεται η ακμή ελαχίστου κόστους και οι αντίστοιχες περιοχές ενώνονται. Ο συγκεκριμένος αλγόριθμος, είναι γρηγορότερος γιατί χρησιμοποιεί έναν επιπλέον γράφο (Γράφος Ομοιότερου Γείτονα) [19] μέσω του οποίου το μέγεθος και ο χρόνος επεξεργασίας της ουράς προτεραιότητας ελαττώνονται δραστικά (Σχήμα 12). Η τελική έξοδος του

αλγόριθμοι είναι μια διαμέριση της εικόνας η οποία μέσω του γράφου γειτνίασης περιοχών, προσφέρει πληροφορία που μπορεί να χρησιμοποιηθεί από υψηλότερου επιπέδου διαδικασίες.



Σχήμα 12: Γράφος Γειτνίασης και ενδεχόμενος Γράφος Ομοιότερου Γείτονα [19]

Επιπρόσθετα, έχουν προταθεί αλγόριθμοι οι οποίοι συνδυάζουν την εξαγωγή ακμών με βάση το χρώμα, με μεθόδους εξάπλωσης σπόρων SRG (seeded region growing) [19]. Αρχικά οι ακμές με βάση το χρώμα εξασφαλίζονται με μια ιστροπική ανίχνευση ακμών (isotropic edge detection) και με μια γρήγορη ανάλυση ιστογράμματος με χρήση κατωφλιού (entropic thresholding technique) (Σχήμα 13). Με τον τρόπο αυτό έχουν εξαχθεί οι πιο σημαντικές γεωμετρικές δομές (περιοχές) της εικόνας. Τα κέντρα αυτών των γειτονικών περιοχών αποτελούν τους αρχικούς σπόρους που χρησιμοποιούνται στο SRG.



Σχήμα 13: Ανίχνευση ακμών με βάση το χρώμα κατά τους Jianping Fan, David. K. Y. Yau [19]

Στη συνέχεια, ενώνονται στην ίδια περιοχή τα εικονοστοιχεία που βρίσκονται σε γειτονικές ακμές που μπορούν να ενωθούν με χρήση ελάχιστου επικαλύπτοντος δέντρου (minimum spanning tree). Τα κέντρα $(X_{i,j}^c, Y_{i,j}^c)$ δύο τέτοιων περιοχών E_i, E_j δίνονται από τους τύπους :

$$X_{i,j}^c = \frac{\sum_{(x,y) \in \{E_i, E_j\}} x}{\sum_{(x,y) \in \{E_i, E_j\}} \delta(x,y)} \quad Y_{i,j}^c = \frac{\sum_{(x,y) \in \{E_i, E_j\}} y}{\sum_{(x,y) \in \{E_i, E_j\}} \delta(x,y)} \quad (13)$$

όπου

$$\delta(x,y) = \begin{cases} 1, & (x,y) \in \{E_i, E_j\} \\ 0, & \text{αλλιώς} \end{cases} \quad (14)$$

Τέλος, έχουν προταθεί χρήσιμες μέθοδοι για τμηματοποίηση εικόνας με τη χρήση παραμετρικών κατανομών (Parametrical Distributional Clustering). Οι μέθοδοι αυτές έχουν το πλεονέκτημα της προσέγγισης των τοπικών χαρακτηριστικών μιας εικόνας με κάποια κατανομή και έτσι πραγματοποιείται μια καλή στατιστική ανάλυση των ιδιοτήτων της εικόνας. Έτσι αποφεύγονται τα λάθη που προκαλούνται από μετρήσεις οι οποίες είναι ευαίσθητες στον θόρυβο. Πιο συγκεκριμένα, όπως έχει ήδη προταθεί [20] για την τμηματοποίηση εικόνας αρχικά θεωρούνται οι αντικείμενα όπου, $i=1, \dots, n$. Ο σκοπός είναι τα αντικείμενα αυτά να ομαδοποιηθούν σε k τμήματα.

Για κάθε αντικείμενο o_i ορίζεται ένα σύνολο από n_i παρατηρήσεις $X_i = \{x_{i1}, x_{i2}, \dots, x_{in_i}\}, x_{ij} \in R^d$. Ορίζεται επίσης η μεταβλητή M_{in} με $n=1, 2, \dots, k$ ως εξής, $M_{in} = 1$ αν το o_i αντικείμενο συμπεριλαμβάνεται στο τμήμα n , και $M_{in} = 0$ αλλιώς. Οι τιμές αυτές εισάγονται στον πίνακα $M \in M = \{0,1\}^{n \times k}$. Στο σημείο αυτό γίνεται η παραδοχή για κάθε τμήμα ενός αντικειμένου ότι ακολουθεί Gaussian κατανομή. Επομένως οι X_i παρατηρήσεις λαμβάνονται από ένα μικτό μοντέλο Gaussian κατανομών ανάλογα με το τμήμα n του συγκεκριμένου αντικειμένου. Δεδομένου του τμήματος n του αντικειμένου στο οποίο ανήκει η πιθανότητα της παρατήρησης x δίνεται από τον τύπο

$$p(x|v) = \sum_{\alpha=1}^l p_{\alpha|v} g(x|\mu_{\alpha}, \sigma_{\alpha}) \quad (15)$$

όπου

$g_{\alpha}(x) = g(x|\mu_{\alpha}, \sigma_{\alpha})$ Gaussian κατανομή με μέση τιμή μ_{α} και διασπορά σ_{α}

Έστω Θ το σύνολο που αποτελείται από τις ελεύθερες παραμέτρους αυτού του μοντέλου, $\Theta = \{p_v, p_{\alpha|v}, \mu_\alpha \mid \alpha = 1, \dots, l; v = 1, \dots, \kappa\}$. Για τον προσδιορισμό των παραμέτρων αυτών απαιτείται η μεγιστοποίηση της πιθανοφάνειας που δίδεται από τον τύπο

$$p(X, M \mid \Theta) = \prod_{i \leq n} \prod_{v \leq k} [p_v \cdot p(X_i \mid v, \Theta)]^{M_{iv}} \quad (16)$$

Με την μέθοδο αυτή μετατρέπεται το πρόβλημα μιας καλής τμηματοποίησης, σε πρόβλημα εύρεσης μέγιστης πιθανοφάνειας με παραμετρικές μεθόδους για τις οποίες θα γίνει αναφορά στην ενότητα 2.2.1. Μειονέκτημα τέτοιων μεθόδου αποτελεί η παραδοχή ως προς την κατανομή των αντικειμένων της εικόνας όπως ήδη αναφέρθηκε, καθώς και ο καθορισμός του πλήθους των τελικών αντικειμένων της εικόνας μετά την τμηματοποίησή της.

1.3 Η προτεινόμενη μέθοδος τμηματοποίησης

Στην εργασία αυτή παρουσιάζεται μια μέθοδος τμηματοποίησης εικόνων η οποία κατατάσσεται στις τεχνικές τμηματοποίησης με γράφους όπως αυτές αναλύθηκαν στην ενότητα 1.2.5. Καινοτομία του αλγορίθμου αποτελεί το γεγονός ότι η απόφαση για την ένωση δύο περιοχών της εικόνας λαμβάνεται από μια μη παραμετρική στατιστική δοκιμασία (ενότητα 2.2.2). Έτσι η τμηματοποίηση της εικόνας γίνεται χωρίς υποθέσεις για την κατανομή που ακολουθεί το χαρακτηριστικό ως προς το οποίο γίνεται η τμηματοποίηση. Επιπρόσθετα, δεν απαιτείται εκ των προτέρων γνώση του πλήθους των περιοχών στις οποίες πρέπει να διαμεριστεί η εικόνα. Στην εν λόγω εργασία γίνεται εφαρμογή της μη παραμετρικής στατιστικής δοκιμασίας Mann Whitney (ενότητα 2.2.2.1) για τμηματοποίηση εικόνων ως προς το χρώμα.

Πιο αναλυτικά, στην εργασία αυτή κατασκευάζεται ένας μη κατευθυνόμενος γράφος $G = \{V, E\}$ όπου κάθε ακμή συνδέει δύο κορυφές με ένα βάρος το οποίο χαρακτηρίζει την ομοιότητα των δύο κορυφών. Ο γράφος αρχικά αποτελείται από πολλές κορυφές οι οποίες ανάλογα με την ομοιότητα τους ενώνονται επαναληπτικά ανά δύο. Σκοπός της ένωσης των κορυφών του γράφου είναι να δημιουργηθούν τόσο υπογράφοι όσα είναι και τα φυσικά αντικείμενα που απεικονίζονται στην εικόνα [12]. Με τον τρόπο αυτό εξάγονται αντικείμενα καθένα από τα οποία παρουσιάζουν ομοιομορφία ως προς κάποιο χαρακτηριστικό. Εδώ το χαρακτηριστικό αυτό είναι το χρώμα.

Έστω R η εικόνα ως σύνολο σημείων. Η τμηματοποίηση της σε k περιοχές R_1, R_2, \dots, R_k πρέπει να είναι τέτοια ώστε η ένωσή τους να αποτελεί ολόκληρη την εικόνα R (σχέση (17)), αλλά και κάθε σημείο της εικόνας να ανήκει σε μια περιοχή (σχέση (28)).

$$\bigcup_{k=1}^K R_k = R \quad (17)$$

$$R_i \cap R_j = \emptyset, \quad i \neq j \quad (18)$$

Ο αλγόριθμος της εν λόγω εργασίας αποτελείται από πέντε βασικές διαδικασίες όπως αναφέρονται συνοπτικά στο Σχήμα 14. Τα τρία πρώτα αλγοριθμικά βήματα αφορούν στην προεπεξεργασία της εικόνας έτσι ώστε η εικόνα να μετατραπεί σε γράφο όπου κάθε κορυφή αποτελεί μια μικρή περιοχή της εικόνας.

- Βήμα 1:** Ανιστροπική Εξομάλυνση (Anisotropic Diffusion)
 - Βήμα 2:** Υπερτμηματοποίηση (over-segmentation)
 - Βήμα 3:** Εφαρμογή κατωφλιού (size thresholding)
 - Βήμα 4:** Δημιουργία γράφου
 - Βήμα 5:** Ένωση ακμών του γράφου
 - Βήμα 6:** Δημιουργία τελικής εικόνας από τον γράφο

Σχήμα 14: Περιγραφή του αλγορίθμου

Πιο αναλυτικά, αρχικά εφαρμόζεται η διαδικασία της Ανιστροπικής Εξομάλυνσης έτσι ώστε να εξομαλυνθεί η εικόνα από το θόρυβο, όπως περιγράφεται στην ενότητα 3.1.1. ενώ παράλληλα διατηρούνται οι ασυνέχειες φωτεινότητας [21]. Έπειτα πραγματοποιείται μια αρχική διαμέριση της εικόνας σε πολλές μικρές περιοχές (ενότητα 3.1.2 Υπερτμηματοποίηση) όπου κάθε τέτοια περιοχή έχει ομοιόμορφο χρώμα χωρίς μεγάλες αντιθέσεις και για κάθε μια περιοχή ικανοποιούνται οι σχέσεις (20) και (21). Στη συνέχεια, ορίζεται ένα κατώφλι το οποίο προσδιορίζει το ελάχιστο πλήθος των εικονοστοιχείων μιας περιοχής. Έτσι, εάν μια περιοχή αποτελείται από πολύ μικρό αριθμό εικονοστοιχείων τότε αυτή εμπερικλείεται σε μια άλλη γειτονική της περιοχή. Η επιλογή της περιοχής αυτής περιγράφεται στην ενότητα 3.1.3.

Οι περιοχές που έχουν δημιουργηθεί τελικά από την διαδικασία της υπερτμηματοποίησης και της εφαρμογής του κατωφλιού αποτελούν τους κόμβους ενός γράφου. Έτσι δημιουργείται ο γράφος όπου ο κάθε κόμβος χαρακτηρίζεται από ένα συγκεκριμένο χρώμα, και κάθε ακμή που ενώνει ένα ζεύγος κόμβων έχει ένα βάρος το οποίο εκφράζει το βαθμό ομοιότητας των δύο περιοχών. Η ομοιότητα δύο περιοχών καθορίζεται ανάλογα με το χρώμα κάθε περιοχής και ο υπολογισμός του βάρους γίνεται με την εφαρμογή μιας Μη Παραμετρικής Στατιστικής δοκιμασίας [22]. Οι ιδιότητες των δοκιμασιών αυτών αναλύονται στην ενότητα 2.2.2. Η συγκεκριμένη δοκιμασία που εφαρμόζεται στην εν λόγω εργασία ονομάζεται «Mann-Whitney» και περιγράφεται εκτενώς στην ενότητα 2.2.2.1, ενώ η εφαρμογή της για την επεξεργασία μιας εικόνας περιγράφεται στην ενότητα 3.2.3. Κάθε φορά που ενώνονται δύο κόμβοι και δημιουργείται μια νέα περιοχή μεταβάλλονται και τα χαρακτηριστικά κάθε κόμβου, το χρώμα, το πλήθος των εικονοστοιχείων που αποτελούν κόμβο και άλλα χαρακτηριστικά που περιγράφονται στην 3.2.5. Για το λόγο αυτό επαναπροσδιορίζονται τα χαρακτηριστικά αυτά και εφαρμόζεται πάλι ο αλγόριθμος «Mann-Whitney» για τα νέα δεδομένα. Ο αλγόριθμος λαμβάνει τέλος όταν η επαναληπτική εφαρμογή της δοκιμασίας αυτής καταλήξει σε περιοχές που δεν επιδέχονται περαιτέρω συνένωση.

2 Βασικές έννοιες

Στο κεφάλαιο αυτό γίνεται μια σύντομη περιγραφή εννοιών οι οποίες κρίνονται απαραίτητες για την ανάλυση και επεξήγηση του αλγόριθμου που παρουσιάζεται στην εργασία αυτή. Έτσι, στην ενότητα 2.1 αναλύεται η έννοια του χρώματος και περιγράφονται οι τρόποι με τους οποίους γίνεται η αναπαράστασή του στην επεξεργασία εικόνων. Ακόμα περιγράφονται συγκεκριμένα χρωματικά μοντέλα που χρησιμοποιούνται στην εργασία αυτή. Στις ενότητες 2.2.1 και 2.2.2 επεξηγούνται όροι της στατιστικής, όπως *Παραμετρική μέθοδος* και *Μη Παραμετρική μέθοδος*. Στην εν λόγω εργασία υλοποιήθηκε μια μη παραμετρική στατιστική δοκιμασία για την σύγκριση δύο περιοχών της εικόνας, η οποία ονομάζεται «Mann Whitney» ως κριτήριο ομοιότητας δύο περιοχών. Στην ενότητα 2.2.2.1 περιγράφεται αναλυτικά η συγκεκριμένη δοκιμασία. Οι παραμετρικές μέθοδοι αναφέρονται στην 2.2.1 έτσι ώστε να γίνει σύγκριση ανάμεσα στις δύο μεθόδους, και να αιτιολογηθεί η επιλογή των μη παραμετρικών μεθόδων στην συγκεκριμένη εφαρμογή.

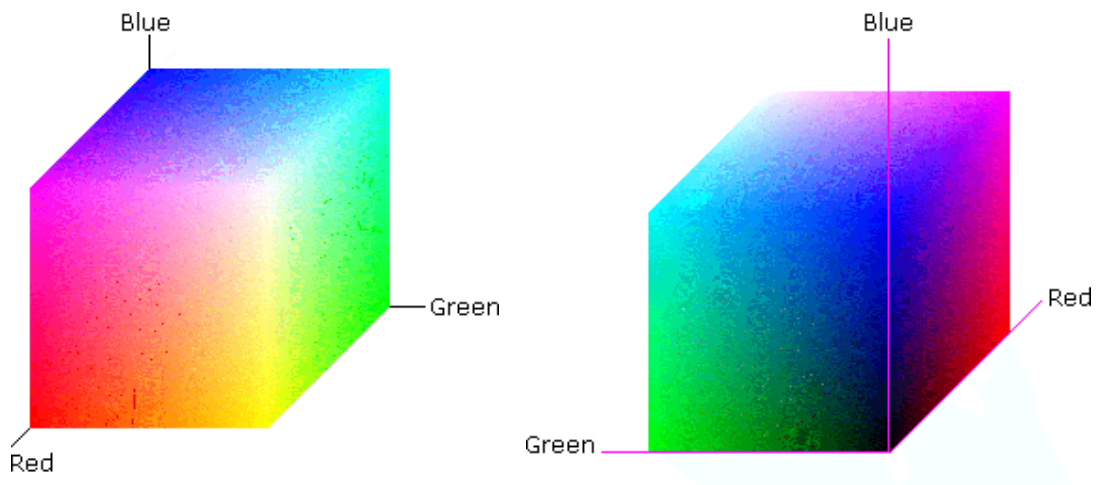
2.1 Αναπαράσταση του χρώματος στην επεξεργασία εικόνων

Για την ανάλυση και επεξεργασία μιας έγχρωμης εικόνας επομένως ο τρόπος με τον οποίο θα αναπαρασταθεί η πληροφορία του χρώματος για να ληφθεί αυτή η πληροφορία, είναι καθοριστική. Έτσι χρησιμοποιούνται διάφορα χρωματικά μοντέλα τα οποία αναπαριστούν την πληροφορία αυτή σε δύο, τρεις, ή τέσσερις διαστάσεις. Δηλαδή σε κάθε εικονοστοιχείο αντιστοιχούν δύο, τρεις, ή τέσσερις συνιστώσες χρώματος αντίστοιχα, οι οποίες είναι δυνατό να συνδυαστούν με διάφορους τρόπους.

Στην εν λόγω εργασία η επιλογή του χρωματικού μοντέλου έχει ιδιαίτερη σημασία για την τμηματοποίηση η οποία βασίζεται στην πληροφορία κάθε εικονοστοιχείου ως προς το χρώμα. Τα χρωματικά μοντέλα που χρησιμοποιούνται στην συγκεκριμένη εργασία είναι τα εξής RGB, HSV και YUV.

2.1.1 RGB

Στο RGB τα τρία βασικά χρώματα κόκκινο, πράσινο, μπλε, αναπαριστώνται στο τρισδιάστατο καρτεσιανό σύστημα συντεταγμένων, όπως φαίνεται στο Σχήμα 15 από διαφορετική οπτική γωνία. Στην αρχή των συντεταγμένων είναι το μαύρο (απουσία χρώματος), δηλαδή εκεί οι συνιστώσες του κόκκινου, του πράσινου και του μπλε είναι μηδέν, ενώ το άσπρο βρίσκεται στην απέναντι γωνία του κύβου εκεί όπου όλες οι συνιστώσες έχουν την μέγιστή τους τιμή. Όπως και τα περισσότερα χρωματικά μοντέλα, έτσι και το RGB είναι κανονικοποιημένο. Όλες οι συνιστώσες παίρνουν τιμές από το μηδέν μέχρι το 1, άρα το μαύρο χρώμα έχει συντεταγμένες (0.0, 0.0, 0.0) και το άσπρο (1.0, 1.0, 1.0).

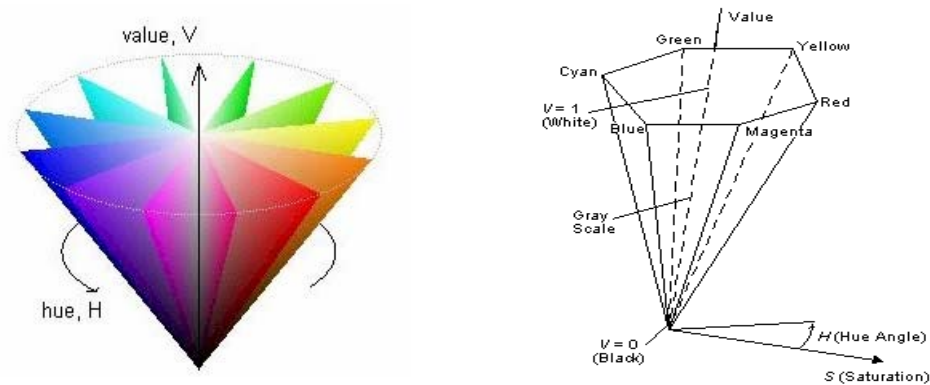


Σχήμα 15: Αναπαράσταση RGB

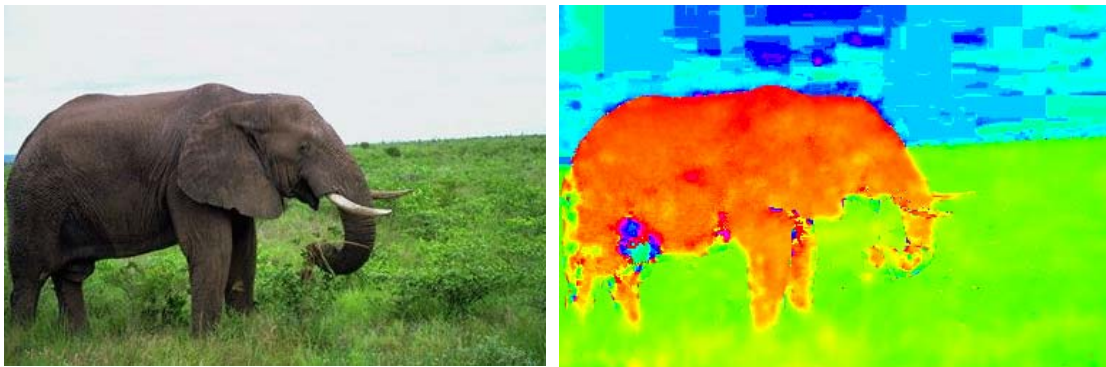
2.1.2 HSV

Το HSV (Hue, Saturation, Value) χρησιμοποιείται κυρίως από καλλιτέχνες και παράλληλα θεωρείται πιο αντιληπτικό για τον άνθρωπο και καταλληλότερο για αναπαράσταση και επεξεργασία εικόνων. Η συνιστώσα Hue είναι το χρώμα, Saturation είναι το επίπεδο του γκρι χρώματος και τέλος Value η φωτεινότητα.

Από το Σχήμα 17 μπορεί να παρατηρήσει κανείς ότι το χρώμα αποδίδεται ικανοποιητικά από τη μια συνιστώσα, όμως στα σημεία όπου το saturation είναι πολύ μικρό δεν γίνεται σωστή αναπαράσταση του πραγματικού χρώματος. Και αυτό το χρωματικό μοντέλο είναι κανονικοποιημένο και οι τιμές των συνιστωσών φαίνονται στο Σχήμα 16.



Σχήμα 16: Αναπαράσταση HSV

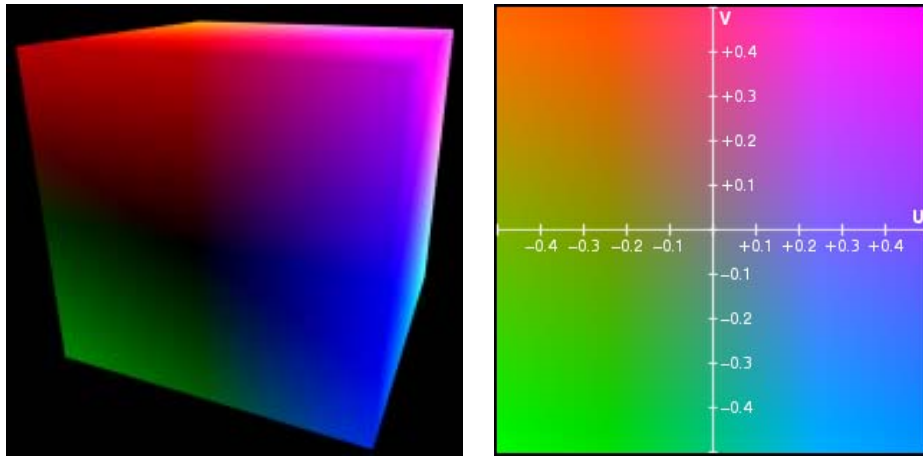


Σχήμα 17: Η συνιστώσα του χρώματος (hue) στο HSV

2.1.3 YUV

Τέλος, στην εργασία αυτή χρησιμοποιείται και το χρωματικό μοντέλο YUV. Η συνιστώσα Y αναπαριστά την φωτεινότητα (luminance) , ενώ η U και V το χρώμα (chroma). Το Y παίρνει τιμές στο διάστημα [0,1] ενώ τα U, V στο [-0.5,

0.5]. Το πλεονέκτημα του YUV είναι ότι η χρωματική πληροφορία είναι ανεξάρτητη από τη φωτεινότητα. Έτσι, αγνοώντας τις δύο συνιστώσες U,V το Y δίνει μια εικόνα στις αποχρώσεις του γκρι. Στο Σχήμα 18 φαίνεται η τρισδιάστατη αναπαράσταση του YUV από την πλευρά όπου $Y=0$. Το απόλυτο μαύρο είναι στα σημεία όπου $Y=U=V=0$.



Σχήμα 18: Αριστερά: τρισδιάστατη αναπαράσταση YUV, δεξιά: αναπαράσταση των U,V για $Y=0,5$

2.2 Στατιστικές δοκιμασίες

Μια «Στατιστική δοκιμασία» χρησιμοποιείται για την σύγκριση παρατηρούμενης στατιστικής παραμέτρου (συχνότητα, μέση τιμή, διακύμανση κ.τ.λ) προς την αντίστοιχη θεωρητική παράμετρο. Στις δοκιμασίες αυτές διατυπώνεται μια υπόθεση H_0 και ορίζεται η στάθμη σημαντικότητας η οποία δίνει μια τιμή κρίσεως που καθορίζει την περιοχή απόρριψης της H_0 και μια περιοχή αποδοχής της H_0 . Οι δοκιμασίες αυτές πραγματοποιούνται είτε με χρήση κάποιων στατιστικών παραμέτρων του δείγματος και ονομάζονται «Παραμετρικές δοκιμασίες», είτε χωρίς καμία γνώση παραμέτρων του δείγματος και για το λόγο αυτό ονομάζονται « Μη Παραμετρικές δοκιμασίες». Οι δοκιμασίες για τις οποίες γίνεται περαιτέρω ανάλυση στην εργασία αυτή είναι οι δοκιμασίες ομοιογένειας για μεγάλα δείγματα.

2.2.1 Παραμετρικές δοκιμασίες

Οι παραμετρικές δοκιμασίες έχουν εφαρμογές όχι μόνο σε στατιστικές αναλύσεις αλλά και σε οποιοδήποτε πεδίο κρίνεται απαραίτητη η λήψη μιας απόφασης δεδομένης μιας υπόθεσης [23]. Με τις δοκιμασίες αυτές γίνεται σύγκριση μιας παρατηρούμενης στατιστικής παραμέτρου είτε συχνότητα, είτε μέση τιμή, είτε διακύμανση, προς την αντίστοιχη θεωρητική παράμετρο.

Η πιο χαρακτηριστική παραμετρική δοκιμασία είναι η δοκιμασία ομοιογένειας με τη χρήση μέσων τιμών. Έστω δύο δείγματα Δ_1 και Δ_2 με μεγέθη η_1, η_2 αντίστοιχα ενώ οι μέσες τιμές τους ως προς ένα χαρακτηριστικό είναι m_1, m_2 . Το ερώτημα είναι εάν με κάποια στάθμη σημαντικότητας τα δύο αυτά δείγματα προέρχονται από τον ίδιο πληθυσμό. Έστω ότι το μεν δείγμα Δ_1 εξήχθη από ένα πληθυσμό Π_1 του οποίου το χαρακτηριστικό έχει μέση τιμή μ_1 και διακύμανση σ_1^2 , το δε δεύτερο δείγμα Δ_2 εξήχθη από ένα πληθυσμό Π_2 του οποίου το χαρακτηριστικό έχει μέση τιμή μ_2 και διακύμανση σ_2^2 . Κατά την υπόθεση H_0 , τα δύο δείγματα θεωρούνται ομογενή δηλαδή ισχύει $\mu_1 = \mu_2$. Εάν δίνονται οι τιμές των διακυμάνσεων σ_1^2, σ_2^2 υπολογίζεται ο λόγος

$$z = \frac{m_1 - m_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \quad (19)$$

Εάν σ_1^2, σ_2^2 άγνωστοι, τότε υπολογίζεται ο λόγος

$$z = \frac{m_1 - m_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad (20)$$

Όπου s_1^2, s_2^2 δίνονται από τις σχέσεις (21),(22) και x_{1i} ($i=1,2,\dots,n_1$), x_{2i} ($i=1,2,\dots,n_2$) είναι η τιμή του χαρακτηριστικού ως προς το οποίο γίνεται η σύγκριση των δειγμάτων των Δ_1, Δ_2 .

$$s_1^2 = \frac{\sum_{i=1}^{n_1} (x_{1i} - m_1)^2}{n_1 - 1} \quad (21)$$

$$s_2^2 = \frac{\sum_{i=1}^{n_2} (x_{2i} - m_2)^2}{n_2 - 1} \quad (22)$$

Στο σημείο αυτό γίνεται χρήση της εξής ιδιότητας [26], εάν από οποιοδήποτε στατιστικό σύνολο μέσης τιμής μ και διακύμανσης σ^2 , εξαχθεί δείγμα μεγέθους n , του οποίου η μέση τιμή είναι m , τότε η μεταβλητή u δίνεται από την ακόλουθη σχέση,

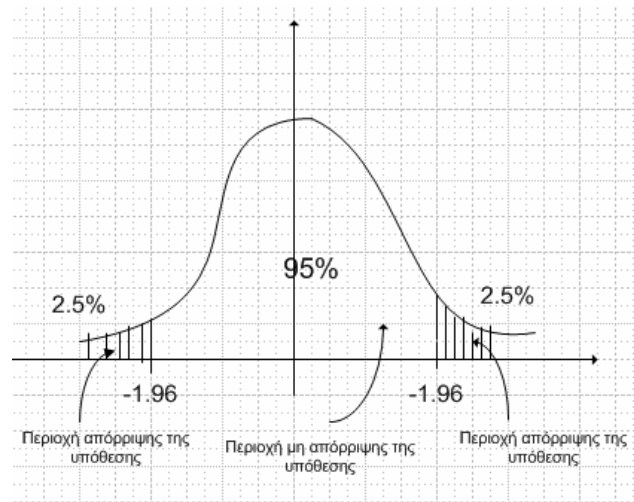
$$u = \frac{m - \mu}{\sqrt{\frac{\sigma^2}{n}}} \quad (23)$$

ακολουθεί κατά προσέγγιση κανονική κατανομή $N(0,1)$. [26]. Η προσέγγιση αυτή έχει καλύτερα αποτελέσματα όσο το μέγεθος του δείγματος γίνεται μεγαλύτερο.

Στην περίπτωση λοιπόν που σ_1^2, σ_2^2 γνωστά, και όπως αναφέρθηκε παραπάνω η m_1 ακολουθεί κανονική κατανομή μέσης τιμής μ_1 και διακύμανσης $\frac{\sigma_1^2}{n_1}$ και η m_2 όμοια κανονική κατανομή με μέση τιμή μ_2 και διακύμανση $\frac{\sigma_2^2}{n_2}$, τότε και η $m_1 - m_2$ ακολουθεί κανονική κατανομή με μέση τιμή $\mu_1 - \mu_2$ και επειδή τα δείγματα είναι ανεξάρτητα η διακύμανση είναι $\sigma_d^2 = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$ άρα και η μεταβλητή της σχέσης (19) ακολουθεί κανονική κατανομή $N(0,1)$.

Στην περίπτωση όπου σ_1^2, σ_2^2 δεν είναι γνωστά, τότε αυτές εκτιμούνται όπως αποδεικνύεται [24] με τις αντίστοιχες τιμές (21) (22). Και σε αυτή τη περίπτωση ισχύει ότι η $m_1 - m_2$ ακολουθεί κανονική κατανομή με μέση τιμή $\mu_1 - \mu_2$ και η διακύμανση είναι $s_d^2 = \frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}$ άρα και η μεταβλητή της σχέσης (20) ακολουθεί κανονική κατανομή $N(0,1)$.

Έτσι, σύμφωνα με την δοκιμασία αυτή εάν $|z| < 1.96$, η διαφορά $|m_1 - m_2|$ δεν είναι σημαντική στην στάθμη 5%, και επομένως η υπόθεση H_0 γίνεται δεκτή, διαφορετικά η διαφορά θεωρείται υπολογίσιμη και απορρίπτεται η υπόθεση H_0 . Η στάθμη σημαντικότητας 5% δηλώνει ότι $P(|z| < 1.96) = 0.95$, Σχήμα 19.



Σχήμα 19: Περιοχές απορρίψεως και μη απορρίψεως της H_0 για στάθμη σημαντικότητας 5%

Γενικά, ακόμα και για την σύγκριση περισσότερων από δύο παρατηρούμενων μέσων τιμών γίνεται η υπόθεση για το προς μελέτη χαρακτηριστικό των πληθυσμών ότι ακολουθεί κανονική κατανομή και επιπλέον ότι οι διακυμάνσεις του χαρακτηριστικού των πληθυσμών είναι ίσες. Αυτή η παραδοχή είναι περιοριστική για τις διάφορες εφαρμογές των δοκιμασιών αυτών. Ακόμα, για να είναι δυνατή η υπόθεση της κανονικότητας της κατανομής του πληθυσμού πρέπει το δείγμα να είναι επαρκές. Διαφορετικά, αν οι προς διάθεση παρατηρήσεις είναι περιορισμένες δεν μπορεί να γίνει καμία υπόθεση για την κατανομή του πληθυσμού γεγονός που επίσης περιορίζει την εφαρμογή των παραμετρικών δοκιμασιών.

2.2.2 Μη παραμετρικές δοκιμασίες

Σε αντίθεση με τις παραμετρικές μεθόδους, μια στατιστική δοκιμασία ονομάζεται μη παραμετρική όταν η εφαρμογή της δεν εξαρτάται από υποθέσεις σχετικές με την κατανομή του χαρακτηριστικού των πληθυσμών από τα οποία ελήφθησαν τα δείγματα [23]. Το ζητούμενο σε αυτή την περίπτωση όπως και στις Παραμετρικές δοκιμασίες είναι εάν μια υπόθεση H_0 , η οποία αναφέρεται σε ολόκληρο τον πληθυσμό, βρίσκεται σε αντίφαση με τα αποτελέσματα της παρατήρησης στο δείγμα, και αν είναι επιβεβλημένο να προτιμηθεί αντί της H_0 μια άλλη υπόθεση H_1 . Η διαφορά μεταξύ των παραμετρικών και των μη παραμετρικών δοκιμασιών είναι ότι στις μη παραμετρικές μεθόδους συγκρίνονται οι τάξεις του χαρακτηριστικού των δειγμάτων και όχι οι μέσες τιμές ή άλλες στατιστικές παράμετροι. Η έννοια της τάξης αναλύεται παρακάτω. Τέλος, με χρήση των δοκιμασιών αυτών λαμβάνονται έγκυρα αποτελέσματα ακόμα και αν υπάρχει μικρό στατιστικό δείγμα προς μελέτη.

Στο σημείο αυτό θα αναλυθεί η έννοια της τάξης του χαρακτηριστικού με βάση το οποίο γίνεται η σύγκριση των δύο δειγμάτων. Ας υποθεθεί ότι λαμβάνεται ένα δείγμα A από 9 ανθρώπους από τους οποίους καταγράφεται το ύψος τους. Έστω ότι η καταμέτρηση έδωσε τα εξής αποτελέσματα: 159, 172, 164, 164, 168, 172, 172, 176, 157. Γίνεται ταξινόμηση των δεδομένων αυτών από το μικρότερο ύψος στο μεγαλύτερο, και ο πίνακας έχει ως εξής

Τιμή της μεταβλητής	157	159	164	164	168	172	172	172	176
----------------------------	-----	-----	-----	-----	-----	-----	-----	-----	-----

Έτσι, μετά την ταξινόμηση την τάξη 1 θα έχει η τιμή 157 της μεταβλητής, την τάξη 2 θα έχει η τιμή 159 της μεταβλητής. Η επόμενη όμως τιμή της μεταβλητής, η 164, λαμβάνεται από δύο άτομα. Αυτές οι δύο ίσες τιμές της μεταβλητής πρέπει να καταλάβουν την 3^η και 4^η τάξη. Έτσι σε αυτές τις ίσες τιμές δίνεται η τάξη $3.5 = \frac{3+4}{2}$. Εν συνεχεία, την 5^η τάξη θα καταλάβει η τιμή 168 της μεταβλητής. Η επόμενη όμως τιμή της μεταβλητής, η 172, λαμβάνεται από τρία άτομα. Οι τρεις αυτές ίσες τιμές της μεταβλητής πρέπει να καταλάβουν την 6^η, 7^η, και 8^η τάξη. Έτσι σε αυτές τις τιμές της μεταβλητής δίνεται η τάξη $7 = \frac{6+7+8}{3}$. Και τέλος την τάξη 9 θα έχει η τιμή 176 της μεταβλητής. Ο παρακάτω πίνακας δείχνει τις τελικές τιμές της τάξης των διάφορων τιμών της μεταβλητής.

Τιμή της μεταβλητής	157	159	164	164	168	172	172	172	176
Τάξη της μεταβλητής	1	2	3.5	3.5	5	7	7	7	9

Γενικά, υπάρχουν διάφορες δοκιμασίες μη παραμετρικής στατιστικής οι οποίες συγκρίνουν πληθυσμούς ως προς συγκεκριμένο χαρακτηριστικό. Για παράδειγμα η δοκιμασία Wilcoxon συγκρίνει δύο συζευγμένα δείγματα [24], η δοκιμασία Kruskal Wallis συγκρίνει κ ανεξάρτητα δείγματα [25]. Οι δοκιμασίες αυτές βρίσκουν εφαρμογές σε στατιστικές αναλύσεις αλλά και σε πειράματα στον κλάδο της Βιολογίας και της Ιατρικής (π.χ σε ελέγχους αποτελεσματικότητας φαρμάκων, εμβόλιων).

Στην συγκεκριμένη εργασία χρησιμοποιείται η δοκιμασία Mann-Whitney με την οποία γίνεται η σύγκριση δύο ανεξάρτητων δειγμάτων και το οποίο αναλύεται εκτενώς στην ενότητα 2.2.2.1. Η δοκιμασία αυτή κρίνεται η καταλληλότερη για δύο λόγους. Πρώτον, επειδή δεν είναι γνωστή η κατανομή που ακολουθεί το χαρακτηριστικό ως προς το οποίο γίνεται η τμηματοποίηση (χρώμα, υφή, κίνηση), επομένως μια μη παραμετρική δοκιμασία η οποία συγκρίνει δύο πληθυσμούς χωρίς να απαιτείται γνώση της κατανομής του χαρακτηριστικού κρίνεται κατάλληλη για τέτοιες συγκρίσεις. Και δεύτερον, η δοκιμασία Mann-Whitney συγκρίνει ακριβώς δύο ανεξάρτητους πληθυσμούς, όπου στην εν λόγω εργασία οι πληθυσμοί που θέτονται προς σύγκριση είναι δύο κόμβοι του γράφου οι οποίοι εξ'ορισμού είναι ανεξάρτητοι.

2.2.2.1 Η δοκιμασία Mann Whitney

Η δοκιμασία Mann Whitney είναι μια δοκιμασία ομογένειας δύο ανεξάρτητων δειγμάτων. Επειδή όμως ανήκει στις μη παραμετρικές μεθόδους, για την σύγκριση των δειγμάτων δεν απαιτείται χρήση στατιστικών παραμέτρων αλλά η τάξη των τιμών της μεταβλητής στα δείγματα, όπως αναπτύχθηκε στην ενότητα 2.2.2.

Κατά την δοκιμασία των Mann Whitney [26] δίδονται δύο ανεξάρτητα δείγματα A, B μεγεθών n_A , n_B αντίστοιχα. Τα δείγματα αυτά συγκρίνονται ως προς ένα χαρακτηριστικό. Δεν γίνεται όμως καμία υπόθεση για την κατανομή του χαρακτηριστικού στους δύο πληθυσμούς, αρκεί το χαρακτηριστικό να είναι

«διατάξιμο». Έστω H_0 η υπόθεση τα δύο δείγματα να είναι ομογενή. Έστω Γ το δείγμα όπου $\Gamma = A \cup B$, τότε το πλήθος των στοιχείων του Γ είναι $n_A + n_B$ επειδή $A \cap B = \emptyset$. Ακόμα αν $\text{rank}(i)$ είναι η τάξη του i -οστού στοιχείου του δείγματος A και k το πλήθος των εμφανίσεών του στο δείγμα τότε ως sum_A ορίζεται το παρακάτω άθροισμα. Ανάλογες είναι και οι σχέσεις για τα A και για το δείγμα B .

$$\text{sum}_A = \sum_{i=1}^{n_A} k \cdot \text{rank}(i) \quad (24.1)$$

$$\text{sum}_B = \sum_{i=1}^{n_B} k \cdot \text{rank}(i) \quad (24.2)$$

Για να ελεγχθεί η υπόθεση H_0 σχηματίζονται οι μεταβλητές u_A και u_B , οι οποίες χαρακτηρίζουν την διάταξη των τιμών του χαρακτηριστικού ως προς το οποίο γίνεται η σύγκριση των δύο συνόλων A, B στο σύνολο $\Gamma = A \cup B$.

$$u_A = n_A \cdot n_B + \frac{n_A \cdot (n_A + 1)}{2} - \text{sum}_A \quad (25.1)$$

$$u_B = n_A \cdot n_B + \frac{n_B \cdot (n_B + 1)}{2} - \text{sum}_B \quad (25.1)$$

Μπορεί να παρατηρήσει κανείς ότι οι μεταβλητές u_A και u_B δίνουν την μέγιστη τιμή που μπορεί να πάρει η sum_A ή η sum_B αντίστοιχα. Έστω οι δύο πληθυσμοί A, B όπου οι τιμές του A του χαρακτηριστικού ως προς το οποίο συγκρίνονται είναι μεγαλύτερες από αυτές του B . Ακόμα υποτίθεται ότι το πλήθος εμφάνισης κάθε τιμής του χαρακτηριστικού είναι 1. Τότε εφαρμόζοντας τις εξισώσεις (24.1) και (24.2) στο παράδειγμα αυτό μπορεί κανείς να υπολογίσει το εξής άθροισμα

$$\text{sum}_A = (n_B + 1) + (n_B + 2) + \dots + (n_B + n_A)$$

$$\text{sum}_A = n_A \cdot n_B + (1 + 2 + \dots + n_A)$$

$$\text{sum}_A = n_A \cdot n_B + \frac{n_A \cdot (n_A + 1)}{2}$$

Επομένως αν τα δύο δείγματα A και B ακολουθούν ίδιες κατανομές με την ίδια παράμετρο θέσης τότε τα A και B θεωρούμε ότι προέρχονται από τον ίδιο πληθυσμό (Order Statistics). Εάν λοιπόν u_A και u_B ακολουθούν με ικανοποιητική προσέγγιση κανονική κατανομή, της οποίας η μέση τιμή m_u , και η διακύμανση σ_u^2 όπως φαίνεται στις σχέσεις (26), (27) αποδεικνύεται [23] ότι η υπόθεση H_0 είναι ορθή, εάν δηλαδή τα δείγματα είναι ομογενή.

$$m_u = \frac{n_A \cdot n_B}{2} \quad (26)$$

$$\sigma_u^2 = \frac{n_A \cdot n_B \cdot (n_A + n_B + 1)}{12} \quad (27)$$

Ισοδύναμα, εάν η μεταβλητή w που δίνεται από την εξίσωση (28) ακολουθεί την κανονική κατανομή $N(0,1)$ τότε τα δύο δείγματα θεωρούνται ομογενή. Ο έλεγχος είναι δυνατό να γίνει για οποιαδήποτε μεταβλητή u_A ή u_B .

$$w = \frac{u_A - m_u}{\sigma_u} = \frac{u_A - \frac{n_A n_B}{2}}{\sqrt{\frac{n_A n_B (n_A + n_B + 1)}{12}}} \quad (28.1)$$

$$w = \frac{u_B - m_u}{\sigma_u} = \frac{u_B - \frac{n_A n_B}{2}}{\sqrt{\frac{n_A n_B (n_A + n_B + 1)}{12}}} \quad (28.2)$$

Έτσι από τους πίνακες της κανονικής κατανομής εάν $|w| < 1.96$ η υπόθεση H_0 γίνεται δεκτή στην στάθμη 5%, δηλαδή η υπόθεση ότι τα δύο δείγματα είναι ομογενή, δηλαδή με πιθανότητα 0.95 η H_0 γίνεται δεκτή $P(|w| < 1.96) = 0.95$.

3 Ανάλυση μεθόδου

Στο προηγούμενο κεφάλαιο αναφέρθηκαν γενικά θέματα τμηματοποίησης εικόνας και βασικές έννοιες οι οποίες είναι απαραίτητες για την ανάλυση του εν λόγω αλγορίθμου, ενώ στο συγκεκριμένο κεφάλαιο γίνεται εκτενής ανάλυση της μεθόδου. Στην ενότητα 3.1, περιγράφονται οι μέθοδοι που χρησιμοποιήθηκαν ώστε να αναχθεί το πρόβλημα της τμηματοποίησης της εικόνας σε πρόβλημα διάσπασης ενός γράφου. Στη συνέχεια (ενότητα 3.1.3), περιγράφεται η κατασκευή του γράφου γειτνίασης περιοχών στον οποίο οι κόμβοι παριστάνουν περιοχές της εικόνας, και οι ακμές τις σχέσεις μεταξύ τους. Το κριτήριο ομοιότητας ως προς το χρώμα που χρησιμοποιείται ως βάρος τις ακμές του γράφου δίνεται από τη μη παραμετρική δοκιμασία Mann Whitney (ενότητα 3.2.3). Στο γράφο αυτό εφαρμόζεται μια επαναληπτική διαδικασία ένωσης δύο περιοχών και λαμβάνουν χώρα οι ανάλογες πράξεις ανανέωσης του γράφου οι οποίες περιγράφονται στην ενότητα 3.2.5.

Τα αποτελέσματα του αλγορίθμου παρουσιάζονται στην ενότητα 4 για ένα μεγάλο πλήθος έγχρωμων εικόνων και συγκρίνονται τα αποτελέσματα για τα διάφορα χρωματικά μοντέλα που έχουν χρησιμοποιηθεί σε κάθε εικόνα. Τέλος, παρουσιάζονται τα αποτελέσματα για διάφορες τιμές της παραμέτρου D η οποία έχει ορίζεται στην ενότητα 3.2.2 και καθορίζει το πλήθος του δείγματος που λαμβάνεται από την κάθε περιοχή της εικόνας και σχολιάζεται η αξιοπιστία του αλγορίθμου για τις διάφορες τιμές της παραμέτρου D αλλά και των διάφορων χρωματικών μοντέλων.

3.1 Προεπεξεργασία εικόνας

Όπως ήδη αναφέρθηκε μετά την επεξεργασία της εικόνας δημιουργείται ο γράφος ομοιότητας γειτονικών περιοχών της εικόνας. Ο γράφος έχει τόσες κορυφές όσες είναι και οι περιοχές του γράφου. Θα μπορούσε κανείς να θεωρήσει αρχικά ότι κάθε εικονοστοιχείο αποτελεί και κορυφή του γράφου. Το υπολογιστικό κόστος όμως στην περίπτωση αυτή είναι μεγάλο. Για το λόγο αυτό κρίνεται απαραίτητη η υπερτμηματοποίηση της εικόνας σε πολλές ομοιογενείς περιοχές. Η διεργασία αυτή περιγράφεται αναλυτικά στην παράγραφο 3.1.2. Με την εφαρμογή της Ανισοτροπικής Διάχυσης στην εικόνα, η ομαδοποίηση των εικονοστοιχείων γίνεται πιο εύκολα αφού εξομαλύνεται η εικόνα, ενώ παράλληλα διατηρούνται οι

ασυνέχειες φωτεινότητας στα σημεία όπου υπάρχουν έντονες αντιθέσεις φωτεινότητας.

3.1.1 Ανισοτροπική Εξομάλυνση (Anisotropic diffusion)

Στο αρχικό αυτό στάδιο κρίνεται απαραίτητη η εφαρμογή κάποιου αλγόριθμου για την εξομάλυνση (smoothing) της εικόνας από θόρυβο. Από τις διάφορες μεθόδους που έχουν προταθεί κατά καιρούς για την υλοποίηση της εργασίας αυτής επιλέχθηκε η Ανισοτροπική Εξομάλυνση όπως προτάθηκε από τους Michael J. Black, Guillermo Sapiro, David H. Marimont και David Heeger το 1998 [22] η οποία βέβαια βασίζεται σε μια προσέγγιση των Perona – Malik αλλά με την βοήθεια στατιστικών εργαλείων (*robust statistics*).

Πιο αναλυτικά, οι Perona – Malik πρότειναν Ανισοτροπική Εξομάλυνση [26] όπου η βασική εξίσωση είναι η εξής,

$$I_s^{t+1} = I_s^t + \frac{\lambda}{|n_s|} \sum_{p \in n_s} g(\nabla I_{s,p}) \nabla I_{s,p} \quad (29)$$

όπου I_s^t είναι η εικόνα μετά από t επαναλήψεις, s το τρέχον εικονοστοιχείο της εικόνας, $|n_s|$ είναι το πλήθος των γειτόνων του τρέχοντος εικονοστοιχείου, λ σταθερά ($\lambda \in \mathbb{R}^+$) η οποία καθορίζει τον βαθμό της εξομάλυνσης (diffusion rate) και η συνάρτηση g καθορίζει τις ακμές της εικόνας (edge stopping function)

Με τον τρόπο αυτό γίνεται εξομάλυνση (smoothing) στις τιμές φωτεινότητας της αρχικής εικόνας διατηρώντας όμως τις αντιθέσεις φωτεινότητας με αποτέλεσμα να διατηρούνται οι ακμές της εικόνας. Το ίδιο αποτέλεσμα επιτυγχάνεται μη τη χρήση στατιστικών μεθόδων (*robust statistics* [27]). Πιο συγκεκριμένα, αρκεί να ελαχιστοποιηθεί η νόρμα του σφάλματος (error norm) $\rho(\cdot)$, η οποία αποτελεί την συνάρτηση όπου προσδιορίζει τις ακμές της εικόνας (edge stopping functions).

$$\min_I \sum_{s \in I} \sum_{p \in n_s} \rho(I_p - I_s, \sigma) \quad (30)$$

Για την ελαχιστοποίηση της ποσότητας αυτής αρκεί να ελαχιστοποιείται η διαφορά στις τιμές φωτεινότητας μεταξύ κάθε εικονοστοιχείου I_s , και τον γείτονά του I_p .

Η λύση αυτού του προβλήματος δίνεται από την εξίσωση (31) όπου για συνάρτηση g χρησιμοποιείται μια συνάρτηση ψ όπως φαίνεται στην παρακάτω σχέση, και σ είναι μια σταθερά που εξαρτάται από την επιλογή της συνάρτησης g .

$$I_s^{t+1} = I_s^t + \frac{\lambda}{|n_s|} \sum_{p \in n_s} \psi(I_p - I_s^t, \sigma) \quad (31)$$

Για κάθε εικόνα I υπολογίζεται η μεταβλητή σ δίνεται από τον ακόλουθο τύπο

$$\sigma = 1.4826 \cdot \text{median}(\|\nabla I - \text{median}(\|\nabla I\|)\|) \quad (32)$$

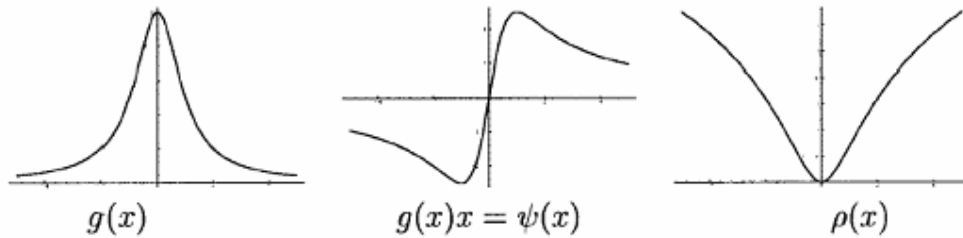
Για την συνάρτηση ψ που εμφανίζεται στη σχέση (31) σύμφωνα με την σχετική εργασία [22] είναι δυνατό να χρησιμοποιηθούν διάφορες συναρτήσεις όπως η Lorentzian (Σχήμα 20), Tukey (Σχήμα 21) και Huber (Σχήμα 22) με τις αντίστοιχες νόρμες σφάλματος.

Στην συγκεκριμένη υλοποίηση επιλέγεται η συνάρτηση Tukey όπου ο τύπος της δίδεται από τη σχέση (33) και η νόρμα σφάλματος από την σχέση (34):

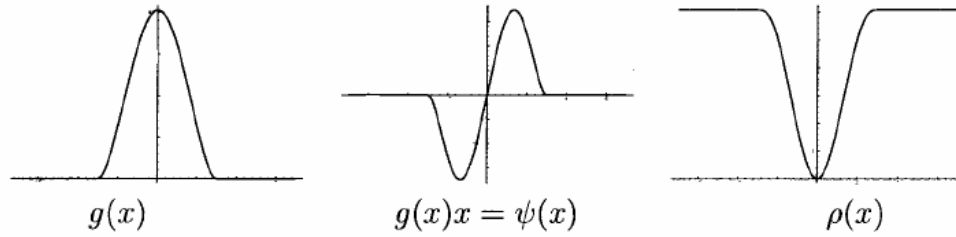
$$\psi(x, \sigma_o) = \begin{cases} x \left[1 - (x/\sigma_o)^2 \right]^2, & |x| \leq \sigma_o \\ 0, & |x| > \sigma_o \end{cases} \quad (33)$$

$$\rho(x, \sigma_o) = \begin{cases} \frac{x^2}{\sigma_o^3} - \frac{x^4}{\sigma_o^4} + \frac{x^6}{3\sigma_o^6}, & |x| \leq \sigma_o \\ \frac{1}{3}, & \text{αλλιώς} \end{cases} \quad (34)$$

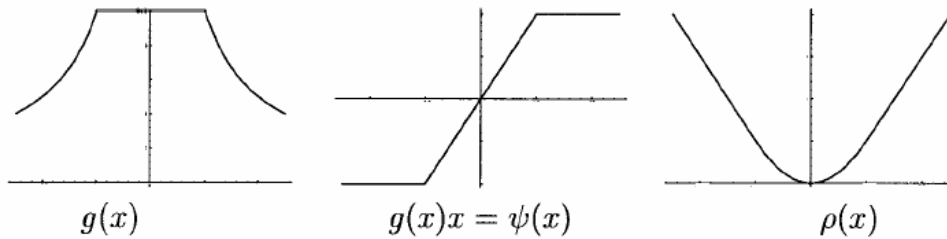
Για να έχει εφαρμογή η μέθοδος σε όλες τις εικόνες ανεξάρτητα από την επιλογή της συνάρτησης ψ , επιλέγεται κατάλληλο σ ώστε όλες οι συναρτήσεις ψ που αναφέρθηκαν να μεταβληθούν τόσο, ώστε να οριστεί ένα κοινό κατώφλι σ_o [22], που καθορίζει τα όρια (outliers) κάθε περιοχής μιας εικόνας [28] όπως φαίνεται στο Σχήμα 23.



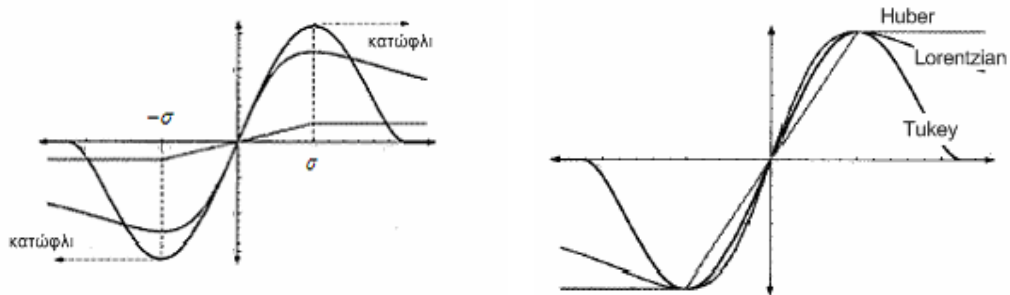
Σχήμα 20: Συνάρτηση Lorentzian



Σχήμα 21: Συνάρτηση Tukey



Σχήμα 22: Συνάρτηση Huber

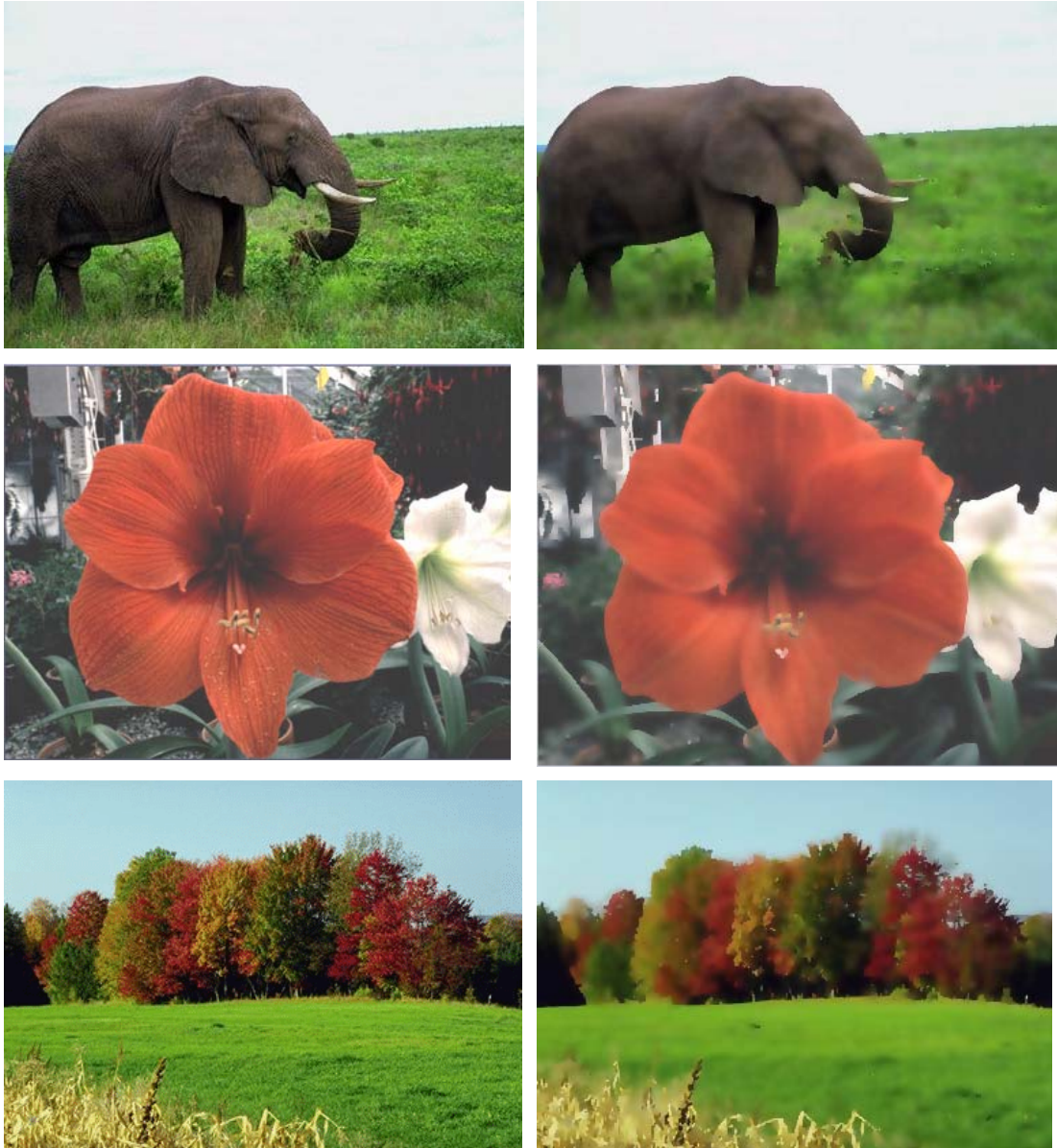


Σχήμα 23: Οι τιμές του σ για τον καθορισμό των ακμών μιας εικόνας και η αντίστοιχη μετατροπή των συναρτήσεων Huber, Lorentzian, και Tukey

Για την συνάρτηση Tukey αυτή η τιμή του σ_0 είναι $\sigma_0 = \sqrt{5} \cdot \sigma$, και η παράμετρος λ της σχέσης (31) δίδεται από τον ακόλουθο τύπο

$$\lambda = \frac{1}{\psi(\sigma, \sigma_0)} \quad (35)$$

Στο Σχήμα 24 φαίνονται τα αποτελέσματα εικόνων μετά την εφαρμογή της Ανισοτροπικής Εξομάλυνσης στις αντίστοιχες εικόνες.



Σχήμα 24: Στα αριστερά η αρχική εικόνα και δεξιά η εικόνα μετά την εφαρμογή της Ανιστροπικής Εξομάλυνσης

3.1.2 Υπερτμηματοποίηση (Over-segmentation)

Στο σημείο αυτό, εφόσον έχει προηγηθεί η εξομάλυνση της εικόνας από το θόρυβο με την εφαρμογή της Ανισοτροπικής Εξομάλυνσης (ενότητα 3.1.1) πραγματοποιείται μια αρχική διαμέριση της εικόνας σε πολλές περιοχές. Έστω R ολόκληρη η εικόνα ως σύνολο σημείων. Η τμηματοποίηση σε k περιοχές R_1, R_2, \dots, R_k πρέπει να είναι τέτοια ώστε η ένωσή τους να αποτελεί ολόκληρη την εικόνα R , αλλά και κάθε σημείο της εικόνας να ανήκει σε μια περιοχή δηλαδή,

$$\bigcup_{k=1}^K R_k = R \quad (36)$$

$$R_i \cap R_j = \emptyset, \quad i \neq j \quad (37)$$

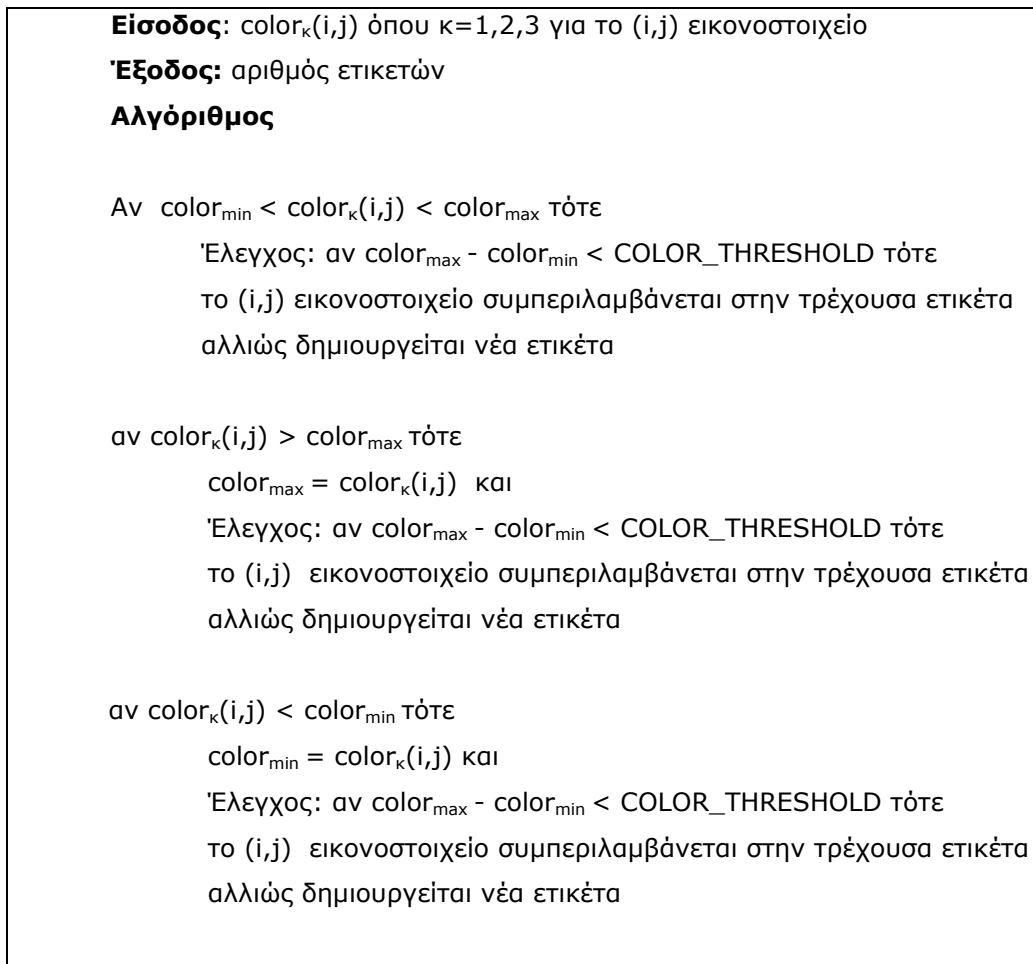
Το πλήθος k των αρχικών περιοχών δεν είναι καθορισμένο από το χρήστη αλλά προκύπτει ανάλογα με την ανομοιομορφία της εικόνας ως προς το χρώμα.

Στο συγκεκριμένο σημείο του αλγορίθμου, δημιουργείται ένας πίνακας M_{ixj} όπου ixj είναι το μέγεθος της εικόνας δηλαδή το πλήθος των εικονοστοιχείων της. Σε κάθε θέση ixj του πίνακα αποθηκεύεται η κάθε ετικέτα του αντίστοιχου εικονοστοιχείου. Η ετικέτα υπολογίζεται με βάση το χρώμα του εικονοστοιχείου.

Για κάθε εικονοστοιχείο που δεν έχει κάποια ετικέτα ελέγχεται η τιμή κάθε συνιστώσας ως προς την τρέχουσα μέγιστη και την ελάχιστη τιμή. Πιο αναλυτικά, έστω (i,j) το τρέχον εικονοστοιχείο και $color_k(i,j)$ όπου $k=1,2,3$ η τιμή του χρώματος της πρώτης, δεύτερης και τρίτης συνιστώσας του χρωματικού μοντέλου αντίστοιχα για το τρέχον εικονοστοιχείο της εικόνας. Η μέγιστη τιμή είναι $color_{max}$ και ελάχιστη $color_{min}$ των εικονοστοιχείων που έχουν ελεγχθεί. Ακόμα ορίζεται ένα κατώφλι $COLOR_THRESHOLD$ για την διαφορά της τιμής του χρώματος, κοινό προς κάθε συνιστώσα. Για όσα εικονοστοιχεία ισχύει $color_{min} < color_k(i,j) < color_{max}$ και $color_{max} - color_{min} < COLOR_THRESHOLD$ αντιστοιχίζεται μια κοινή ετικέτα. Αλλιώς δημιουργείται μια νέα ετικέτα και ανανεώνονται οι τιμές των $color_{min}$ και $color_{max}$. Ο σχετικός αναδρομικός αλγόριθμος περιγράφεται στο Σχήμα 25.

Στο στάδιο αυτό του αλγορίθμου για την αναπαράσταση του χρώματος χρησιμοποιείται το μοντέλο RGB (ενότητα 2.1.1). Όπως αναφέρθηκε, σε αυτό το στάδιο του αλγορίθμου χρησιμοποιούνται οι διαφορές στις τιμές κάθε συνιστώσας (Σχήμα 25) ανάμεσα στα εικονοστοιχεία χωρίς να απαιτείται μεγάλη ακρίβεια για την αναπαράσταση κάθε απόχρωσης. Ακόμα και αν δύο εικονοστοιχεία έχουν παρόμοιο χρώμα αλλά διαφορετική ετικέτα λόγω του περιορισμού $color_{min} < color_k(i,j) < color_{max}$, στην συνέχεια είναι δυνατό να αποκτήσουν την ίδια ετικέτα κατά την σύνδεση των κόμβων του γράφου. Έτσι δεν αλλοιώνεται το

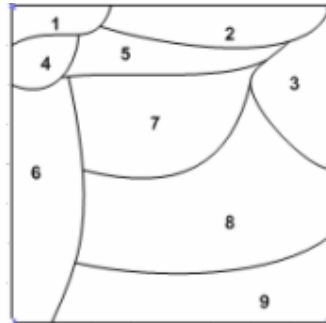
αποτέλεσμα της τμηματοποίησης λόγω του περιορισμού στην αρχική διαμέριση της εικόνας.



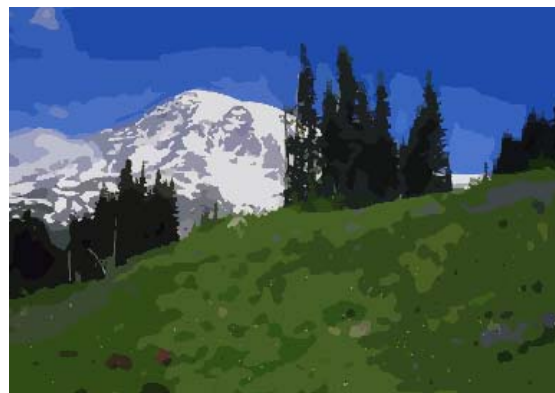
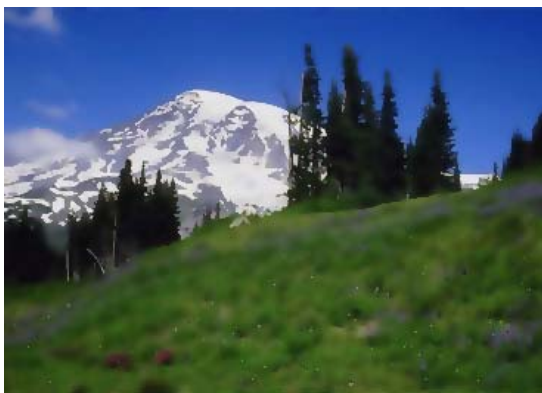
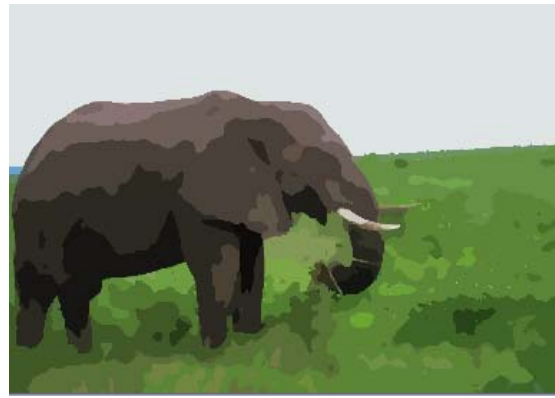
Σχήμα 25: Αλγόριθμος επιλογής ετικέτας

Με την διαδικασία αυτή κάθε εικονοστοιχείο έχει αποκτήσει μια ετικέτα. Αν $M_{i \times j}$ είναι ο πίνακας με τον οποίο αναπαριστάται η εικόνα, τότε σε κάθε θέση του πίνακα αποθηκεύεται η ετικέτα που χαρακτηρίζει το (i,j) εικονοστοιχείο της εικόνας. Με τον τρόπο αυτό μετατρέπεται η εικόνα σε πολλές περιοχές-υπερτμηματοποίηση, όπως φαίνεται στο Σχήμα 26 για έναν πίνακα ετικετών μιας υποτιθέμενης εικόνας μεγέθους 7×7 . Το αποτέλεσμα της υπερτμηματοποίησης για συγκεκριμένες εικόνες με χρωματικό μοντέλο RGB, HSV και YUV φαίνονται στο Σχήμα 27, Σχήμα 28, Σχήμα 29 αντίστοιχα.

1	1	1	2	2	2	3
4	4	5	5	5	3	3
6	6	7	7	7	3	3
6	6	7	7	8	8	8
6	6	8	8	8	8	8
6	6	9	9	9	9	9
6	6	6	9	9	9	9



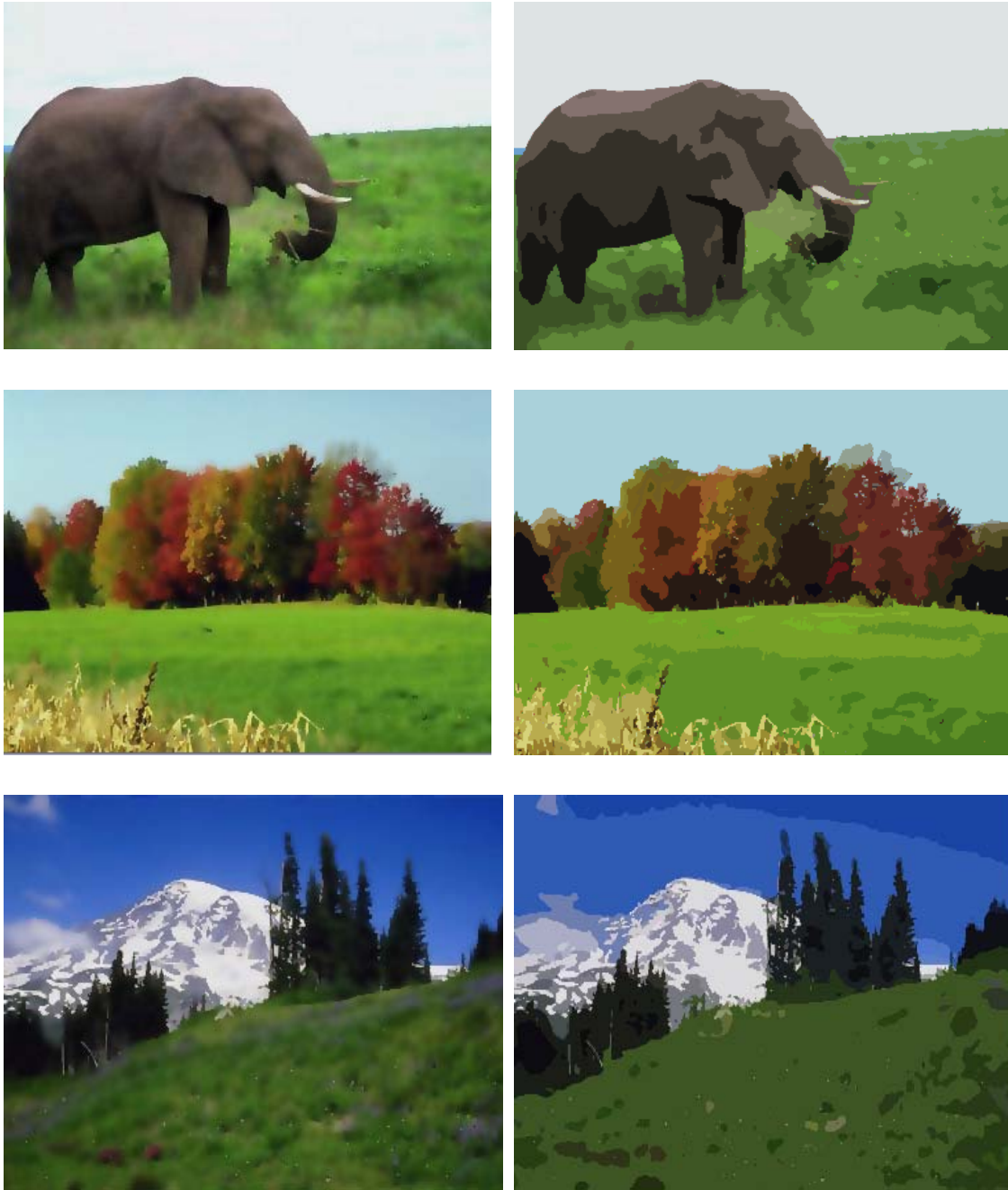
Σχήμα 26: Μετατροπή του πίνακα ετικετών σε περιοχές



Σχήμα 27: Στα αριστερά η εικόνα μετά την εφαρμογή της Ανισοτροπικής Εξομάλυνσης και δεξιά η εικόνα μετά την εφαρμογή του αλγόριθμου επιλογής ετικέτας σε RGB (υπερτμηματοποίηση).



Σχήμα 28: Στα αριστερά η εικόνα μετά την εφαρμογή της Ανιστροπικής Εξομάλυνσης και δεξιά η εικόνα μετά την εφαρμογή του αλγόριθμου επιλογής ετικέτας σε HSV (υπερτμηματοποίηση)



Σχήμα 29: Στα αριστερά η εικόνα μετά την εφαρμογή της Ανιστροπικής Εξομάλυνσης και δεξιά η εικόνα μετά την εφαρμογή του αλγόριθμου επιλογής ετικέτας σε YUV (υπερτμηματοποίηση)

3.1.3 Εφαρμογή κατώφλιου (size thresholding)

Έπειτα από την εφαρμογή της υπερτμηματοποίησης, η εικόνα αποτελείται από πολλές περιοχές με κοινά χαρακτηριστικά ως προς το χρώμα. Παρόλο που το αποτέλεσμα είναι αρκετά αντιπροσωπευτικό για μια αρχική διαμέριση της εικόνας, παρατηρούνται στο Σχήμα 27 μερικές περιοχές που δεν προσδίδουν ιδιαίτερη χρωματική πληροφορία και δεν αντιπροσωπεύουν κάποια ξεχωριστή περιοχή, λόγω του μικρού αριθμού των εικονοστοιχείων που την απαρτίζουν.

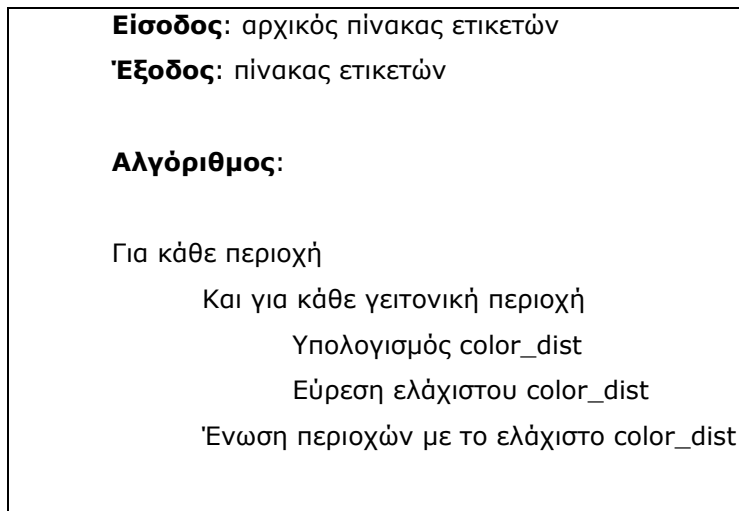
Για λόγους οικονομίας υπολογιστικού χρόνου στο συγκεκριμένο σημείο του αλγορίθμου ορίζεται ένα κατώφλι `SIZE_THRESHOLD` ίσο με 200 το οποίο προσδιορίζει το ελάχιστο πλήθος των εικονοστοιχείων μιας περιοχής. Εάν μια περιοχή αποτελείται από πιο μικρό αριθμό εικονοστοιχείων τότε εμπερικλείεται σε γειτονική της περιοχή ως προς το χρώμα και αποτελούν πλέον μια ενιαία περιοχή με κοινή ετικέτα. Η πιο κοντινή περιοχή (`color_dist`) ως προς το χρώμα επιλέγεται έπειτα από σύγκριση της μέσης τιμής του χρώματος της συγκεκριμένης περιοχής ($mean_color_k$ όπου $k=1,2,3$ για τις τρεις συνιστώσες του χρωματικού μοντέλου) με τη μέση τιμή των γειτονικών της περιοχών ($mean_neighbor_color_k$).

Πιο αναλυτικά, υπολογίζεται η διαφορά της μέσης τιμής της χρωματικής πληροφορίας μιας συνιστώσας της μικρής περιοχής με τη μέση τιμή της χρωματικής πληροφορίας της ίδιας συνιστώσας των γειτονικών περιοχών και συμβολίζεται ως $dist_k$, $k=1,2,3$. Η διαφορά αυτή είναι πολύ μικρή, γεγονός που δικαιολογεί την ένταξη του σε μια μεγαλύτερη περιοχή χωρίς καν να εφαρμοστούν πιο πολύπλοκοι αλγόριθμοι που απαιτούν περισσότερο υπολογιστικό χρόνο. Η διαφορά αυτή υπολογίζεται για κάθε συνιστώσα k του χρωματικού μοντέλου. Έτσι υπάρχει μια ομοιομορφία ως προς το πλήθος των εικονοστοιχείων που αντιστοιχίζονται σε κάθε ετικέτα. Για να υπολογιστεί επομένως η πιο κοντινή περιοχή (`color_dist`) αρκεί να βρεθεί το μικρότερο τιμή της Ευκλείδειας απόστασης των $dist_k$.

$$dist_k = (mean_color_k - mean_neighbor_color_k)^2 \quad \text{όπου} \quad k = 1,2,3 \quad (41)$$

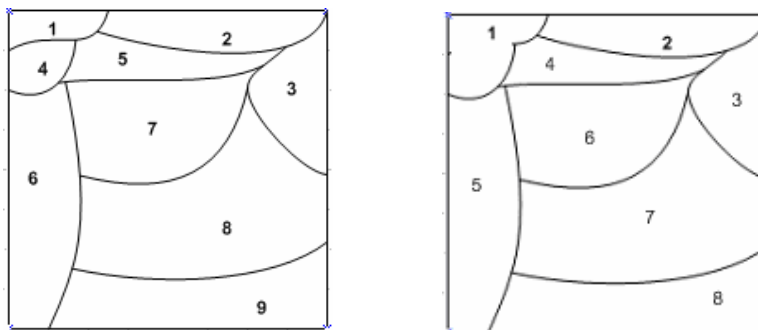
$$color_dist = \left(\sum_{k=0}^2 dist_k \right)^{\frac{1}{2}} \quad (42)$$

Ο αλγόριθμος για τον πίνακα ετικετών φαίνεται στο Σχήμα 30 .

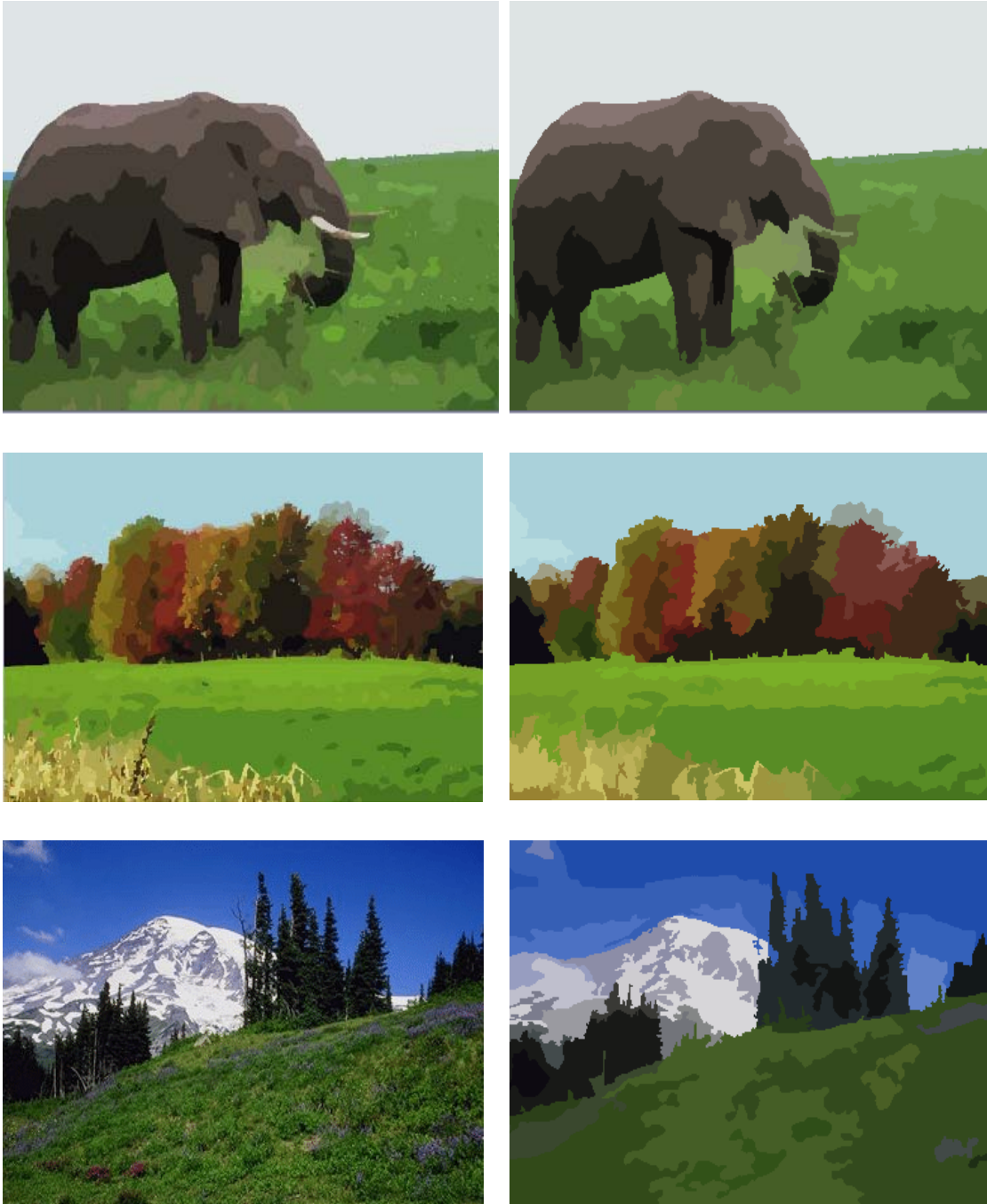


Σχήμα 30: Αλγόριθμος δημιουργίας πίνακα ετικετών

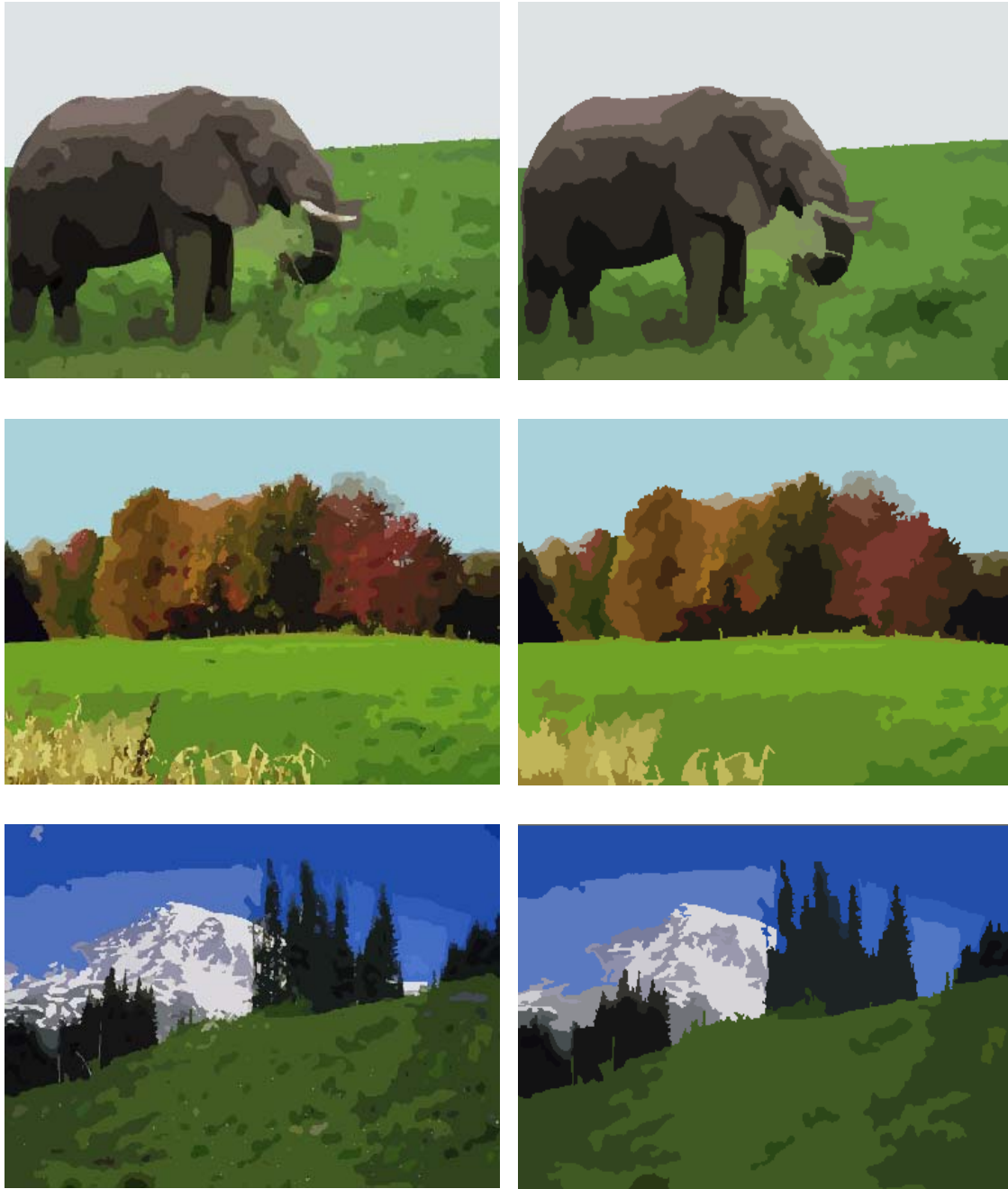
Για το παράδειγμα στο Σχήμα 31 η περιοχή με την ετικέτα 4 έχει πολύ μικρό αριθμό εικονοστοιχείων. Έπειτα υπολογίζεται το color_dist για τις γειτονικές περιοχές 1, 2, 5,6. Έστω ότι την μικρότερη απόσταση ως προς το χρώμα είναι η απόσταση του 1 με το 4, και σχηματικά η ένωση φαίνεται στο Σχήμα 31. Για συγκεκριμένες εικόνες στο Σχήμα 32 φαίνεται η εφαρμογή του κατωφλιού σε μια εικόνα μετά την υπερτμηματοποίηση για τα χρωματικά μοντέλα RGB, HSV και YUV.



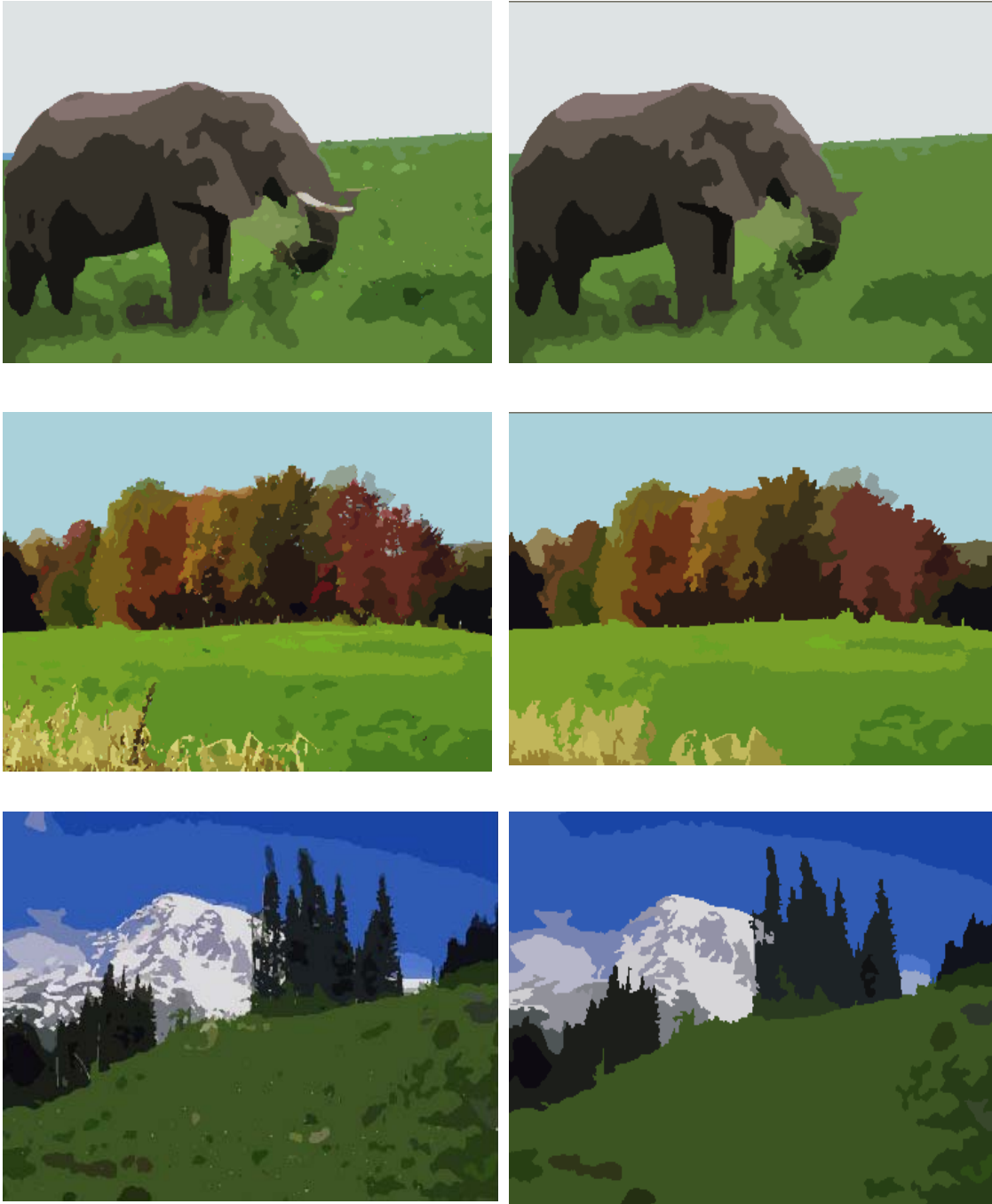
Σχήμα 31: Ο πίνακας ετικετών μετά την εφαρμογή του κατωφλιού (size thresholding).



Σχήμα 32: Στα αριστερά η εικόνα μετά την υπερτμηματοποίηση σε RGB και δεξιά η εικόνα μετά την εφαρμογή του κατοφλιού.



Σχήμα 33: Στα αριστερά η εικόνα μετά την υπερμηματοποίηση σε HSV και δεξιά η εικόνα μετά την εφαρμογή του κατοφλιού.



Σχήμα 34: Στα αριστερά η εικόνα μετά την υπερτμηματοποίηση σε YUV και δεξιά η εικόνα μετά την εφαρμογή του κατοφλιού.

3.2 Επεξεργασία δεδομένων της εικόνας με τη δομή γράφου

Όπως αναλύθηκε στην ενότητα 3.1.3 η αρχική εικόνα έχει μετατραπεί σε μια εικόνα από πολλές περιοχές και σε κάθε μια από τις οποίες αντιστοιχεί μια ετικέτα. Έχοντας σαν είσοδο την εικόνα αυτή στόχος του αλγορίθμου στο στάδιο αυτό, είναι η δημιουργία ενός γράφου με τόσους κόμβους όσους είναι και τα τμήματα της εικόνας. Κάθε κόμβος μπορεί να συνδεθεί μόνο με κάποιο γειτονικό του κόμβο. Η έννοια της γειτνίασης αναλύεται στην ενότητα 3.2.1. Στην ενότητα 3.2.2 ορίζεται η παράμετρος η οποία καθορίζει το πλήθος των εικονοστοιχείων κάθε περιοχής που θα συμμετάσχει στον υπολογισμό του βάρους των ακμών του γράφου. Οι κόμβοι του γράφου ενώνονται με ένα βάρος το οποίο προέρχεται από τη δοκιμασία Mann-Whitney όπως αναλύεται στην ενότητα 3.2.3 για όλες τις συνιστώσες του χρωματικού μοντέλου που επιλέγεται στον αλγόριθμο. Στη συνέχεια επιλέγεται το ζεύγος κόμβων (ενότητα 3.2.4) που θα ενωθεί πρώτο και θα έχει πλέον μια ετικέτα και έπειτα γίνεται η ανανέωση όλων των δεδομένων του γράφου (ενότητα 3.2.5). Ο αλγόριθμος είναι επαναληπτικός, έτσι για τον νέο γράφο που δημιουργήθηκε μετά από την ένωση αυτή εφαρμόζεται πάλι η δοκιμασία Mann-Whitney και υπολογίζονται τα βάρη για τις νέες περιοχές. Ο αλγόριθμος σταματάει όταν δεν υπάρχει κάποιο βάρος ανάμεσα σε δύο κόμβους κάτω από ένα συγκεκριμένο κατώφλι, ορισμένο από την δοκιμασία Mann-Whitney. Με τον τρόπο αυτό παράγεται η τελική εικόνα τμηματοποίησης.

3.2.1 Κατασκευή γράφου γειτονικών περιοχών

Ο αρχικός γράφος $G = \{V, E\}$ αποτελείται από τόσους κόμβους όσες είναι και οι περιοχές στις οποίες έχει χωριστεί η εικόνα μετά εφαρμογή κατωφλιού (size thresholding), όπως περιγράφηκε αναλυτικά στην 3.1.3. Κάθε κόμβος μπορεί να ενωθεί με έναν άλλο κόμβο μόνο αν οι περιοχές που αντιπροσωπεύουν θεωρούνται γειτονικές.

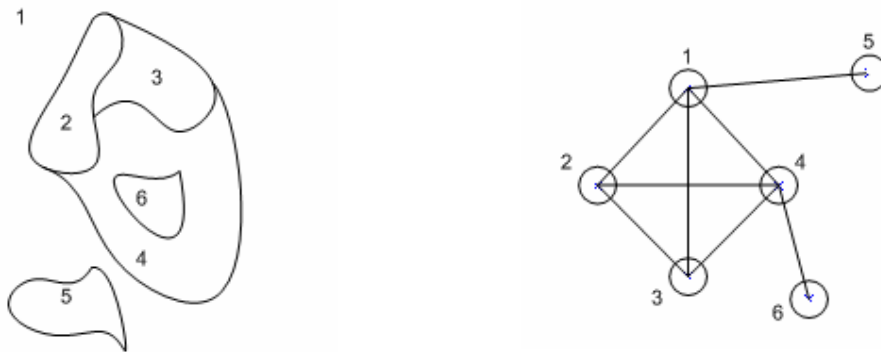
Δύο περιοχές R_i, R_j είναι γειτονικές όταν υπάρχει στοιχείο της R_i του οποίου η γειτονιά περιέχει τουλάχιστον ένα στοιχείο της R_j . Η μαθηματική διατύπωση του γράφου γειτνίασης περιοχών $G = \{V, E\}$ της διαμέρισης Δ_k της εικόνας είναι,

$$V = \{R_1, R_2, \dots, R_K\} \quad (43)$$

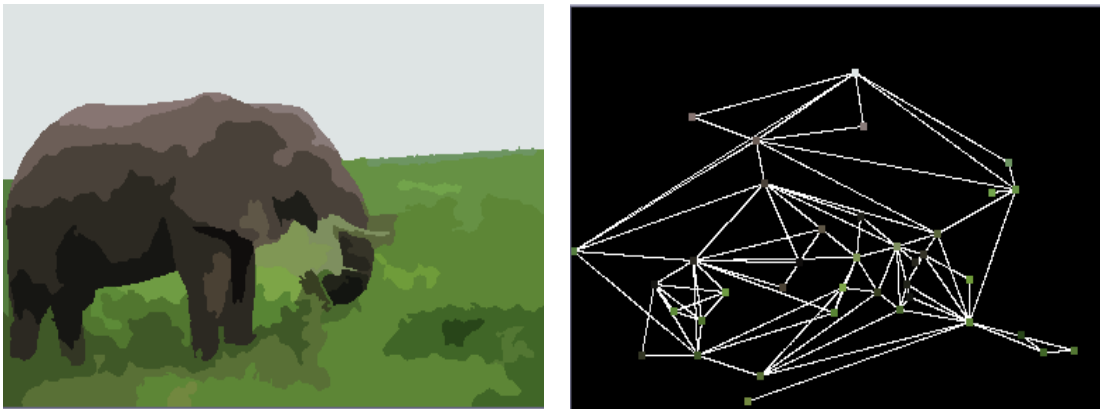
$$R_i = \{x_{i,1}, x_{i,2}, \dots, x_{i,N_i}\}, i \in \{1, 2, \dots, K\} \quad (44)$$

$$E = \{\{R_i, R_j\}, 1 \leq i, j \leq K, i \neq j, R_i, R_j \text{ γειτονικές}\} \quad (45)$$

Ένα παράδειγμα ενός πιθανού γράφου γειτνίασης δίνεται στο Σχήμα 35, ενώ στο Σχήμα 36 παρουσιάζεται ο αρχικός γράφος όπου φαίνονται όλες οι πιθανές συνδέσεις των γειτονικών περιοχών της εικόνας μετά την εφαρμογή κατωφλιού (size thresholding).



Σχήμα 35: Εικόνα ετικετών και ο αντίστοιχος γράφος γειτνίασης



Σχήμα 36: Αριστερά η εικόνα μετά την εφαρμογή κατωφλιού σε RGB και δεξιά ο γράφος με όλες τις πιθανές συνδέσεις των περιοχών της αριστερής εικόνας.

Στο Σχήμα 36 φαίνεται ο γράφος των γειτονικών περιοχών της εικόνας μετά την εφαρμογή κατωφλιού. Ο γράφος αποτελείται από τόσους κόμβους όσες είναι και οι περιοχές της εικόνας μετά την εφαρμογή του κατωφλιού. Κάθε κόμβος συνδέεται μόνο με τους γειτονικούς κόμβους με κάποιο βάρος που αναλύεται στην παρακάτω παράγραφο 3.2.3.

3.2.2 Καθορισμός μεγέθους δείγματος σε κάθε περιοχή

Στο κεφάλαιο αυτό ορίζεται μια παράμετρος D η οποία καθορίζει τον τρόπο με τον οποίο επιλέγονται τα εικονοστοιχεία κάθε περιοχής ώστε να χρησιμοποιηθούν ως δεδομένα στην δοκιμασία Mann Whitney η οποία εφαρμόζεται ανάμεσα σε δύο περιοχές για τον υπολογισμό του βάρους της ακμής τους στον γράφο. Η παράμετρος αυτή μεταβάλλεται έτσι ώστε κατά τον υπολογισμό του βάρους να λαμβάνονται υπόψη τα εικονοστοιχεία κάθε περιοχής ανά ένα, ανά δύο, ανά τέσσερα, κ.ο.κ. δηλαδή για περιορισμένο δείγμα.

Η συγκεκριμένη εφαρμογή της μη παραμετρικής στατιστικής δοκιμασίας ομοιογένειας Mann Whitney έχει εφαρμογή σε οποιαδήποτε δείγματα, αρκεί αυτά να θεωρούνται «μεγάλα». Στην στατιστική αποδεικνύεται ότι ένα «μεγάλο» δείγμα αποτελείται από τουλάχιστον 30 στοιχεία. Έτσι, δεδομένου ότι η μικρότερη περιοχή στην εν λόγω εργασία ορίστηκε να είναι μεγέθους μεγαλύτερου από 200 εικονοστοιχεία (παράγραφος 3.1.3), η μέγιστη τιμή που μπορεί να έχει η παράμετρος D είναι 8.

Πιο αναλυτικά, έστω δύο περιοχές της εικόνας A και B για τις οποίες εφαρμόζεται η δοκιμασία Mann Whitney και n_A, n_B το πλήθος των εικονοστοιχείων των περιοχών A και B αντίστοιχα. Με την παράμετρο διαλογής D μεταβάλλεται το μέγεθος κάθε περιοχής σύμφωνα με τους τύπους (46) (47).

$$n_A^D = \frac{n_A}{D} \quad D = 2, \dots, 8 \quad (46)$$

$$n_B^D = \frac{n_B}{D} \quad D = 2, \dots, 8 \quad (47)$$

Εάν $\Gamma = A \cup B$ είναι η ένωση των περιοχών A και B , το σύνολο Γ θα αποτελείται από n_Γ^D εικονοστοιχεία.

$$n_\Gamma^D = \frac{n_A + n_B}{D} \quad D = 2, \dots, 8 \quad (48)$$

Επειδή $n_\Gamma^D < n_\Gamma$ ο υπολογιστικός χρόνος του αλγορίθμου μειώνεται σημαντικά, όπως αναλύεται και στο κεφάλαιο 4 όπου παρουσιάζονται αναλυτικά τα αποτελέσματα για διάφορες τιμές του D .

3.2.3 Υπολογισμός βαρών του γράφου με την δοκιμασία Mann-Whitney

Η δοκιμασία Mann Whitney περιγράφηκε εκτενώς στην ενότητα 2.2.2.1 ως μια διαδικασία μη παραμετρικής στατιστικής. Στην ενότητα αυτή περιγράφεται η εφαρμογή της διαδικασίας αυτής για δεδομένα που αφορούν χρωματική πληροφορία μιας εικόνας. Στο σημείο αυτό επιλέγεται ένα από τα χρωματικά μοντέλα (RGB, HSV ή YUV) που αναλύθηκαν στην ενότητα 2.1.

Έτσι η μεταβλητή που θα υποστεί ταξινόμηση για να υλοποιηθεί ο αλγόριθμος της δοκιμασίας είναι η τιμή κάθε συνιστώσας του χρωματικού μοντέλου κάθε εικονοστοιχείου. Οι δύο πληθυσμοί από τους οποίους λαμβάνονται τα δείγματα είναι οι περιοχές για τις οποίες γίνεται ο έλεγχος για την πιθανή ένωσή τους, ώστε να αποτελέσουν ένα ενιαίο τμήμα της εικόνας.

Έστω A και B οι δύο περιοχές για τις οποίες γίνεται ο έλεγχος για την ομοιότητά τους ως προς το χρώμα. Το ερώτημα είναι αν αυτές οι περιοχές ενωθούν σε μια ενιαία περιοχή, κατά πόσο το αποτέλεσμα είναι ομογενές, κατά πόσο δηλαδή μπορεί κανείς να θεωρήσει τις δύο αυτές περιοχές ως ένα τμήμα της εικόνας (υπόθεση H_0). Έστω Γ το σύνολο όπου $\Gamma = A \cup B$ και n_A και n_B το πλήθος των δειγμάτων που λαμβάνεται από τις περιοχές A και B αντίστοιχα. Δηλαδή το Γ αποτελείται από όλες τις τιμές μιας συνιστώσας του χρωματικού μοντέλου για όλα τα εικονοστοιχεία που αποτελούν τις δύο περιοχές A, B. Τότε το πλήθος των στοιχείων του Γ είναι $n_A + n_B$ αφού εξ'ορισμού ισχύει $A \cap B = \emptyset$.

Επειδή η δοκιμασία Mann Whitney είναι μια δοκιμασία Μη Παραμετρικής Στατιστικής, η σύγκριση των δύο δειγμάτων γίνεται με βάση την τάξη κάθε εικονοστοιχείου όπως ορίστηκε στην παράγραφο 2.2.2. Έστω $\text{rank}(i)$ η τάξη του i -οστού εικονοστοιχείου του δείγματος Γ . Τα δεδομένα στη συγκεκριμένη εφαρμογή είναι η τιμή του χρώματος κάθε εικονοστοιχείου για κάθε συνιστώσα του χρωματικού μοντέλου. Όπως αναλύθηκε και στην παράγραφο 2.2.2, $n_A + n_B$ τιμές ταξινομούνται από τη μικρότερη στη μεγαλύτερη τιμή. Η μικρότερη τιμή της μεταβλητής είναι τάξης 1, η επόμενη είναι τάξης 2 κ.ο.κ. Σε περίπτωση που η i -οστή τιμή της μεταβλητής είναι ίδια με την τιμή $i+1$ της ταξινόμησης, τότε και οι δύο τιμές είναι τάξης $\frac{i+(i+1)}{2}$.

Για να ελεγχθεί η υπόθεση H_0 σχηματίζεται η μεταβλητή u .

$$u = n_A \cdot n_B + \frac{n_A \cdot (n_A + 1)}{2} - \text{sum}_A \quad (49)$$

Όπου η μεταβλητή sum_A δίνεται από τον σχέση (50) για το i εικονοστοιχείο του συνόλου A τιμή του οποίου παρατηρείται k φορές στο Γ .

$$sum = \sum_{i=0}^{n_A} k \cdot rank(i) \quad (50)$$

Αποδεικνύεται όπως αναλύθηκε στην ενότητα 2.2.2.1 ότι εάν η υπόθεση H_0 είναι ορθή, εάν δηλαδή τα δείγματα είναι ομογενή, τότε η μεταβλητή u ακολουθεί με ικανοποιητική προσέγγιση κανονική κατανομή, της οποίας η μέση τιμή m_u , και η διακύμανση σ_u^2 δίνεται από τους τύπους αντίστοιχα

$$m_u = \frac{n_A \cdot n_B}{2} \quad (51)$$

$$\sigma_u^2 = \frac{n_A \cdot n_B \cdot (n_A + n_B + 1)}{12} \quad (52)$$

Συνεπώς η μεταβλητή w ακολουθεί την κανονική κατανομή $N(0,1)$ όπου η w δίνεται από τον τύπο,

$$w = \frac{u - m_u}{\sigma_u} = \frac{u - \frac{n_A n_B}{2}}{\sqrt{\frac{n_A n_B (n_A + n_B + 1)}{12}}} \quad (53)$$

Σύμφωνα λοιπόν με την δοκιμασία Mann Whitney εάν $|w| < T$ όπου $T=1.96$ γίνεται δεκτή η υπόθεση H_0 δηλαδή η υπόθεση ότι τα δύο δείγματα είναι ομογενή, άρα μπορούν να αποτελέσουν μια ενιαία περιοχή της εικόνας.

Στην επόμενη ενότητα περιγράφεται ο αλγόριθμος με τον οποίο επιλέγεται το ζεύγος των κόμβων που θα ενωθούν, αλλά και ο τρόπος με τον οποίο συνδυάζονται οι τρεις συνιστώσες του χρώματος για τα βάρη των ακμών.

3.2.4 Επαναληπτικός αλγόριθμος ένωσης κόμβων του γράφου

Στην παρούσα εργασία η τμηματοποίηση της εικόνας γίνεται με βάση το χρώμα. Στο κεφάλαιο 2.1 αναλύθηκαν τα χρωματικά μοντέλα τα οποία χρησιμοποιούνται. Ακριβώς επειδή το χρώμα αποδίδεται με τρεις συνιστώσες, τα βάρη του γράφου πρέπει να εκφράζουν ομοιότητα μεταξύ δύο κόμβων ως προς όλες τις συνιστώσες.

Τα βάρη επομένως για κάθε ζεύγος γειτονικών κόμβων υπολογίζονται με τη δοκιμασία Mann-Whitney η οποία αναλύθηκε στην ενότητα 2.2.2.1, για κάθε συνιστώσα του χρωματικού μοντέλου. Έτσι για κάθε ακμή υπολογίζονται τρία βάρη w_1, w_2, w_3 όπως περιγράφηκε στην ενότητα 3.2.3, που αντιπροσωπεύουν την ομοιότητα των περιοχών ως προς την αντίστοιχη συνιστώσα του χρωματικού μοντέλου. Μόνο αν οι δύο περιοχές είναι συγκρίσιμες ως προς όλες τις συνιστώσες του χρώματος υπάρχει πιθανότητα να ενωθούν. Πιο αναλυτικά, οι πιθανές τιμές που μπορεί να πάρει το βάρος του γράφου ανάμεσα σε δύο κόμβους είναι,

$$-T < w_i \leq T \quad i = 1, 2, 3 \quad (54)$$

Στο σημείο αυτό πρέπει να τονιστεί ότι ο εν λόγω αλγόριθμος βασίζεται σε θεωρήματα της μη παραμετρικής στατιστικής, καθώς και ότι όλα τα κατώφλια που χρησιμοποιήθηκαν στην δοκιμασία Mann Whitney βασίζονται σε πίνακες βασικών κατανομών. Έχοντας εφαρμόσει τον αλγόριθμο σε κάθε συνιστώσα του χαρακτηριστικού ως προς το οποίο γίνεται η τμηματοποίηση, έχουν βρεθεί όλα τα ζεύγη των περιοχών που σύμφωνα με την δοκιμασία Mann Whitney μπορούν να αποτελέσουν ενιαία περιοχή. Εδώ πρέπει να ληφθεί μια απόφαση για το ζεύγος που θα ενωθεί πρώτα. Είναι λογικό πως η σειρά με την οποία ενώνονται οι περιοχές παίζει σημαντικό ρόλο στο τελικό αποτέλεσμα της τμηματοποίησης, αφού μετά από κάθε ένωση μεταβάλλονται τα δεδομένα κάθε περιοχής. Έτσι, γίνεται η παραδοχή ότι το ζεύγος των κόμβων με το μικρότερο άθροισμα των w_i , δηλαδή το ζεύγος με την μεγαλύτερη ομοιότητα ως προς όλες τις συνιστώσες, πρέπει να ενωθεί πρώτα.

Επομένως το τελικό βάρος ανάμεσα σε δύο γειτονικούς κόμβους και για τις τρεις συνιστώσες ενός χρωματικού μοντέλου δίδεται από τον τύπο (55) ενώ αν δεν ικανοποιείται έστω και μια συνθήκη από την (54) τότε $w = 0$.

$$w = |w_1| + |w_2| + |w_3| \quad (55)$$

Για να ενωθούν επομένως δύο περιοχές πρέπει να θεωρούνται ομογενείς ως προς κάθε μια συνιστώσα του χρωματικού μοντέλου μετά την εφαρμογή της δοκιμασίας Mann-Whitney. Εάν έστω και για μια συνιστώσα δεν ικανοποιείται το κριτήριο της δοκιμασίας, τότε οι περιοχές θεωρούνται ανομοιογενείς. Έτσι, τα βάρη στις ακμές του γράφου υπολογίζονται από τον τύπο (55) για δύο περιοχές που θεωρούνται ομογενείς διαφορετικά το βάρος είναι μηδέν.

Εάν για παράδειγμα έχει επιλεγεί το χρωματικό μοντέλο HSV τότε υπολογίζονται τρία βάρη ομοιότητας, το βάρος $w(h)$ που χαρακτηρίζει την ομοιότητα των δύο περιοχών ως προς τη πρώτη συνιστώσα Hue, το βάρος $w(s)$ ως προς το Saturation, και τέλος το $w(v)$ ως προς το Value. Και για το συγκεκριμένο

παράδειγμα του χρωματικού μοντέλου HSV ο τύπος για το βάρος δίνεται στην σχέση (56).

$$w = |w(h)| + |w(s)| + |w(v)| \quad (56)$$

Όλα τα βάρη του γράφου αποθηκεύονται σε έναν πίνακα που ονομάζεται πίνακας ομοιότητας. Όσο μικρότερο είναι το βάρος, τόσο μεγαλύτερη είναι η ομοιότητα των κόμβων. Έτσι, όταν τα βάρη διαταχθούν από το μικρότερο στο μεγαλύτερο στον πίνακα ομοιότητας, το πρώτο στοιχείο του πίνακα είναι το $\min_{w(i,j)}$ δείχνει ποιες περιοχές παρουσιάζουν μεγαλύτερη ομοιότητα και επομένως το ζεύγος των κόμβων (i,j) ενώνονται και αποτελούν μια ενιαία περιοχή με μια καινούρια ετικέτα.

Όπως αναφέρθηκε στην ενότητα 3.2.1 κάθε κόμβος μπορεί να ενωθεί μόνο με τους γείτονες του. Έτσι, ο αλγόριθμος για τον υπολογισμό των βαρών ανάμεσα σε κάθε κόμβο και των γειτονικών του κόμβων αντίστοιχα, περιγράφεται στο Σχήμα 37.

Είσοδος: αρχικός γράφος γειτνίασης

Έξοδος: τελικός γράφος με βάρη

Αλγόριθμος:

Για κάθε κόμβο

Και για κάθε γειτονικό κόμβο

Εφαρμογή Mann Whitney τεστ για την πρώτη συνιστώσα (υπολογισμός w_1).

Εφαρμογή Mann Whitney τεστ για την δεύτερη συνιστώσα (υπολογισμός w_2).

Εφαρμογή Mann Whitney τεστ για την τρίτη συνιστώσα (υπολογισμός w_3).

Αν έστω και ένα από τα w_1, w_2, w_3 είναι μεγαλύτερα από το κατώφλι, τότε οι δύο κόμβοι δεν ενώνονται.

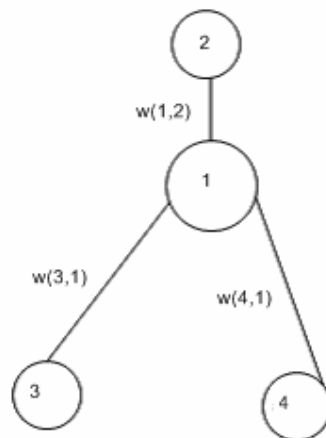
Αλλιώς, υπολογισμός του w και αποθήκευσή του στον πίνακα ομοιότητας.

Διάταξη του πίνακα ομοιότητας από το μικρότερο στο μεγαλύτερο.

Ένωση περιοχών (i,j) που τους αντιστοιχεί το βάρος $\min_{w(i,j)}$.

Σχήμα 37: Αλγόριθμος υπολογισμού βαρών

Ο αλγόριθμος διακόπτεται όταν πλέον δεν υπάρχει κανένα ζεύγος κόμβων που να μπορεί να αποτελέσει μια κοινή περιοχή της εικόνας. Πιο αναλυτικά, έστω ο γράφος στο Σχήμα 38, όπου $w(1,2)$, $w(3,1)$ και $w(4,1)$ είναι τα βάρη των αντίστοιχων κόμβων. Τα βάρη αυτά έχουν υπολογιστεί με τον τρόπο που περιγράφηκε στην ενότητα 3.2.3 με την επαναληπτική διαδικασία που περιγράφηκε παραπάνω. Η δοκιμασία αυτή δίνει ένα κατώφλι $T'=3T$, το οποίο όταν υπερβαίνεται οι δύο περιοχές δεν είναι δυνατό να ενωθούν και έτσι διαγράφεται η συγκεκριμένη ακμή. Στον πίνακα ομοιότητας αποθηκεύονται όλα τα $w(i,j)$ για τα οποία ισχύει $0 < w(i,j) < T'$. Όταν όμως δεν υπάρχουν κόμβοι i,j τέτοιοι ώστε $w(i,j) < T$ τότε ο πίνακας ομοιότητας μηδενίζεται. Στην περίπτωση αυτή δεν υπάρχει κανένα ζεύγος κόμβων το οποίο σύμφωνα με το κριτήριο Mann Whitney να μπορεί να ενωθεί και να αποτελέσει κοινή περιοχή στην εικόνα. Στο σημείο αυτό η τμηματοποίηση έχει ολοκληρωθεί και τα τμήματα της εικόνας είναι όσοι είναι και οι κόμβοι του γράφου στο βήμα αυτό.



Σχήμα 38: Μορφή γράφου μετά από την ένωση δύο κόμβων

3.2.5 Ανανέωση της δομής του γράφου μετά την ένωση δύο περιοχών

Στην 3.2.4 αναλύθηκε ο τρόπος με τον οποίο επιλέγονται οι δύο κόμβοι του γράφου που θα ενωθούν. Κάθε φορά που πραγματοποιείται μια τέτοια ένωση

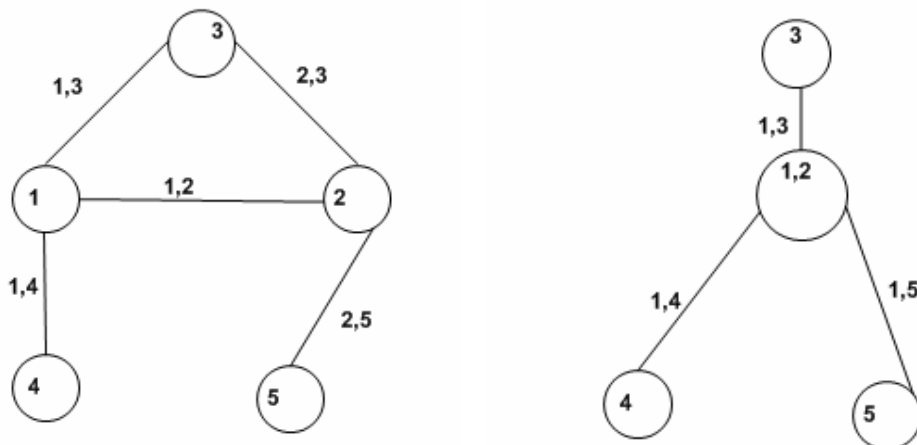
όμως μεταβάλλονται τα δεδομένα και τα χαρακτηριστικά κάθε περιοχής και πρέπει τα στοιχεία αυτά να ανανεώνονται.

Κατά την ένωση δύο περιοχών κάποιες ακμές διαγράφονται όπως για παράδειγμα οι ακμές (1,2) , (2,3) και (2,5) του Σχήμα 39 αριστερά, ενώ κάποιες άλλες εμφανίζονται όπως η (1,5) στο Σχήμα 39 δεξιά μετά την ένωση δύο κόμβων.

Πιο γενικά το σύνολο των γειτόνων του ζεύγους των περιοχών (i,j) διαμερίζεται σε τρία σύνολα,

- A_i : σύνολο γειτόνων της R_i που δεν είναι γείτονες της R_j .
- A_j : σύνολο γειτόνων της R_j που δεν είναι γείτονες της R_i .
- $A_{i,j}$: σύνολο κοινών γειτόνων των R_i, R_j .

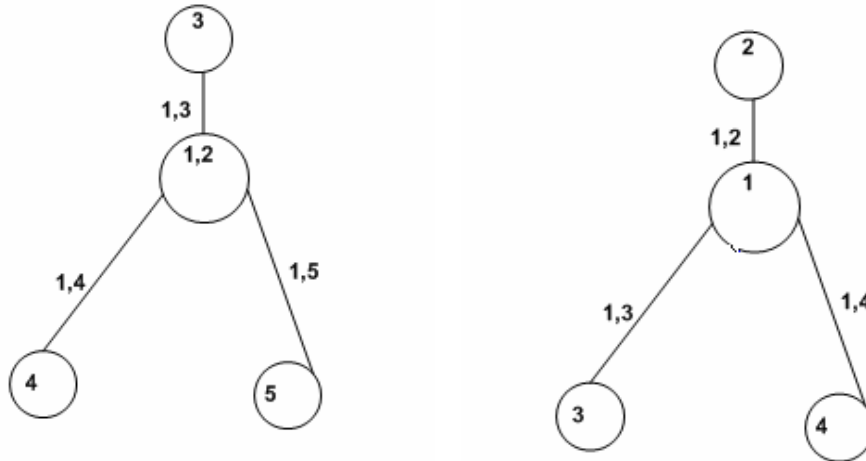
Κατά την ένωση των περιοχών R_i, R_j με $i < j$, οι ακμές του συνόλου A_i εξακολουθούν να υφίστανται αλλά η ομοιότητα μεταξύ των άκρων τους μεταβάλλεται. Οι ακμές του συνόλου A_j διαγράφονται αλλά εισάγονται ισάριθμες με άκρο το i στη θέση του j . Από τις ακμές του συνόλου $A_{i,j}$ διαγράφονται οι ακμές των οποίων το ένα άκρο είναι ίσο με το j .



Σχήμα 39: Παράδειγμα ένωσης περιοχών

Βέβαια για λόγους ομοιομορφίας αλλάζει και η αρίθμηση των ετικετών όπως φαίνεται στο Σχήμα 40. Είναι φανερό ότι μετά από μια ένωση και αλλαγή των ετικετών αλλάζουν και οι γείτονες των κόμβων. Έτσι, πριν αλλάξουν οι ετικέτες αναδιαμορφώνονται και οι γείτονες. Είναι λογικό πως αν ενωθούν δύο περιοχές για παράδειγμα οι 1 και 2 όπως φαίνεται στο Σχήμα 39 τότε οι γείτονες της

καινούριας περιοχής θα έχει όλους τους γείτονες των 1 και 2 μαζί. Βέβαια, στον προσδιορισμό των γειτόνων δεν πρέπει να υπολογιστούν δύο φορές οι κοινοί γείτονες των 1 και 2 όπως είναι η περιοχή 3.



Σχήμα 40: Διαμόρφωση ετικετών μετά από μια ένωση περιοχών

Ακόμα αν n_k το πλήθος των εικονοστοιχείων της περιοχής k και n_l το πλήθος των εικονοστοιχείων της περιοχής l τότε μετά την ένωση των k και l η νέα περιοχή θα έχει n_{k+l} πλήθος εικονοστοιχείων, δηλαδή

$$n_{k+l} = n_k + n_l \quad (57)$$

Τέλος μετά την ένωση δύο κόμβων μεταβάλλεται και το αντιπροσωπευτικό χρώμα της περιοχής. Όπως έχει ήδη αναφερθεί για κάθε περιοχή είναι αποθηκευμένη η μέση τιμή του χρώματος για κάθε χρωματική συνιστώσα $mean_color(i)$, και έχει οριστεί στην ενότητα 3.1.2. Έτσι η νέα μέση τιμή του χρώματος της καινούριας περιοχής δίνεται από τη σχέση,

$$mean_color_{k+l}(i) = \frac{mean_color_k(i) \cdot n_k + mean_color_l(i) \cdot n_l}{n_{k+l}} \quad (58)$$

Με τον τρόπο αυτό έχουν υπολογιστεί όλα τα νέα δεδομένα μιας καινούριας περιοχής και ο αλγόριθμος εφαρμόζεται ξανά από την αρχή για τον καινούριο αυτό

γράφο. Συνοψίζοντας, ο επαναληπτικός αλγόριθμος της εν λόγω εργασίας φαίνεται στο Σχήμα 41.

Είσοδος: αρχικός εικόνα

Έξοδος: εικόνα μετά από τμηματοποίηση ως προς το χρώμα

Αλγόριθμος:

Εφαρμογή Ανισοτροπικής Εξομάλυνσης.

Υπερτμηματοποίηση εικόνας.

Εφαρμογή κατωφλιού.

Κατασκευή αρχικού γράφου όπου όλοι οι κόμβοι ενώνονται με τους γειτονικούς κόμβους .

Επιλογή χρωματικού μοντέλου (RGB, HSV, YUV).

Για κάθε κόμβο του γράφου

Υπολογισμός βαρών $w(i,j)$ στις ακμές του γράφου σύμφωνα με τη δοκιμασία Mann Whitney.

Αν δεν υπάρχουν i,j τέτοια ώστε $w(i,j) < T'$

Τότε έξοδος

αλλιώς

Υπολογισμός του πίνακα ομοιότητας.

Εύρεση $\min_{w(i,j)}$ και το ζεύγος των (i,j) στο οποίο αντιστοιχεί αυτή τη τιμή.

Ένωση των κόμβων i,j .

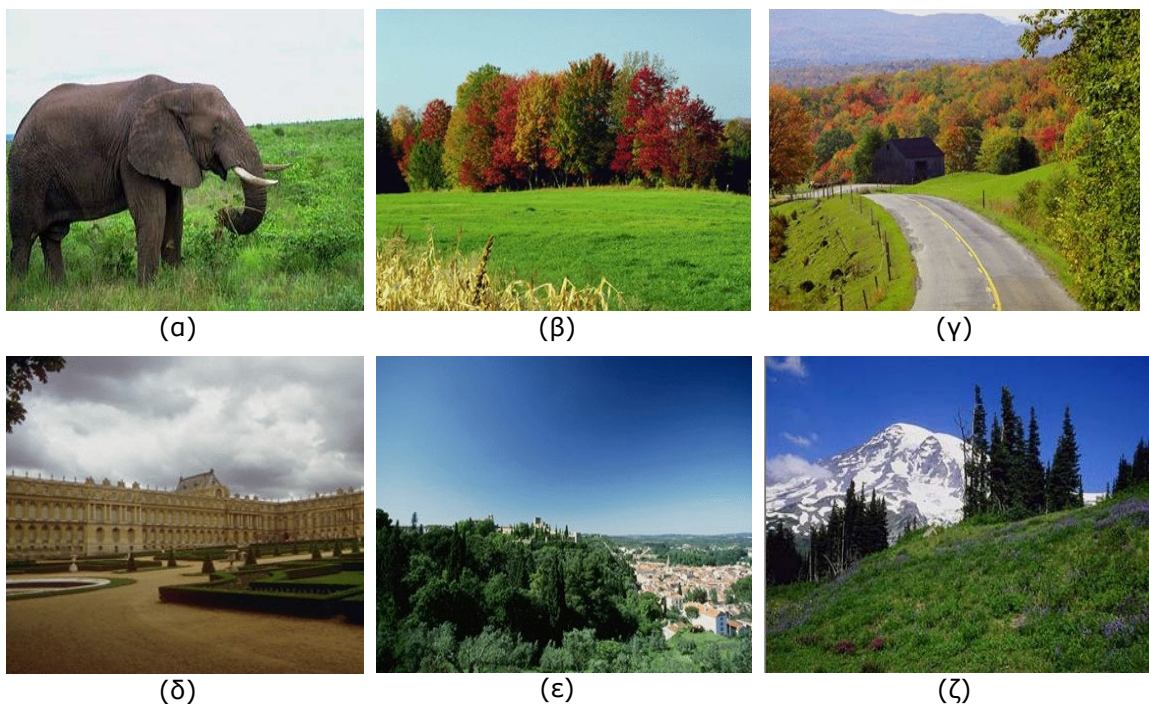
Ανανέωση του γράφου μετά την ένωση των i,j .

Σχήμα 41: Ο επαναληπτικός γραφοθεωρητικός αλγόριθμος τμηματοποίησης εικόνας

4 Αποτελέσματα

Στο κεφάλαιο αυτό παρουσιάζονται τα αποτελέσματα του αλγορίθμου για τα χρωματικά μοντέλα RGB, HSV και YUV τα οποία αναλύθηκαν εκτενώς στην ενότητα 2.1.1, 2.1.2 και 2.1.3 αντίστοιχα. Ακόμα γίνονται πειράματα για διάφορες τιμές της παραμέτρου D (ενότητα 3.2.2) η οποία καθορίζει τον τρόπο με τον οποίο επιλέγονται τα εικονοστοιχεία κάθε περιοχής ώστε να χρησιμοποιηθούν ως δεδομένα στην δοκιμασία Mann Whitney,

Ένα δείγμα από τις εικόνες στις οποίες εφαρμόστηκε ο αλγόριθμος της εργασίας αυτής φαίνονται στο Σχήμα 42.



Σχήμα 42: Ένα δείγμα από τις εικόνες στις οποίες εφαρμόστηκε ο αλγόριθμος.

Οι τιμές των παραμέτρων όπου χρησιμοποιήθηκαν για τα αποτελέσματα αυτά φαίνονται στον παρακάτω πίνακα. Η παράμετρος SIZE_THRESHOLD χρησιμοποιείται για τον καθορισμό του ελάχιστου πλήθους εικονοστοιχείων που πρέπει να έχει μια περιοχή μετά την υπερμνηματοποίηση για να ληφθεί υπόψη στον αλγόριθμο (παράγραφος 3.1.2), η παράμετρος COLOR_THRESHOLD εκφράζει την διαφορά της μέγιστης από την ελάχιστη τιμή του χρώματος κάθε περιοχής (παράγραφος 3.1.2) και η παράμετρος D που καθορίζει το πλήθος των εικονοστοιχείων του υποσυνόλου στο οποίο εφαρμόζεται η δοκιμασία Mann Whitney (παράγραφος 3.2.2).

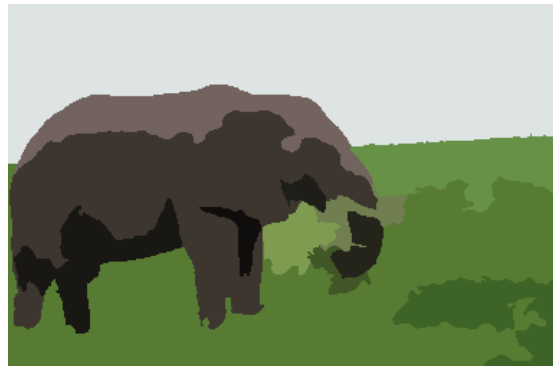
SIZE_THRESHOLD	200
COLOR_THRESHOLD	40
D	2,...,8

4.1 Σύγκριση αποτελεσμάτων ως προς το χρωματικό μοντέλο που χρησιμοποιήθηκε

Αρχικά, χρησιμοποιήθηκε το χρωματικό μοντέλο RGB για την αναπαράσταση του χρώματος και η τιμή της παραμέτρου διαλογής $D=1$. Έτσι το μέγεθος του δείγματος που λαμβάνεται για την εφαρμογή της δοκιμασίας Mann Whitney ανάμεσα σε δύο περιοχές, ισούται με το άθροισμα των εικονοστοιχείων των δύο περιοχών. Τα αποτελέσματα φαίνονται στο Σχήμα 43.



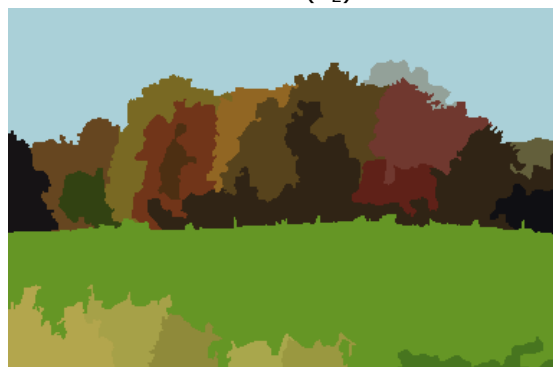
(α_1)



(α_2)



(β_1)



(β_2)



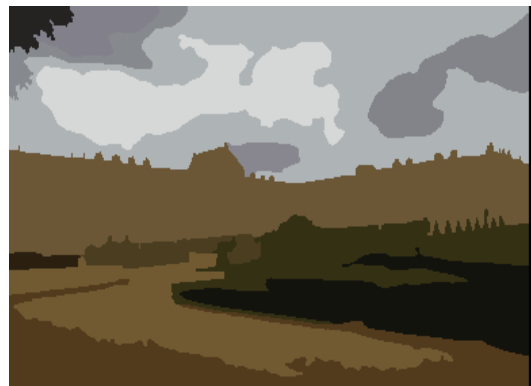
(γ_1)



(γ_2)



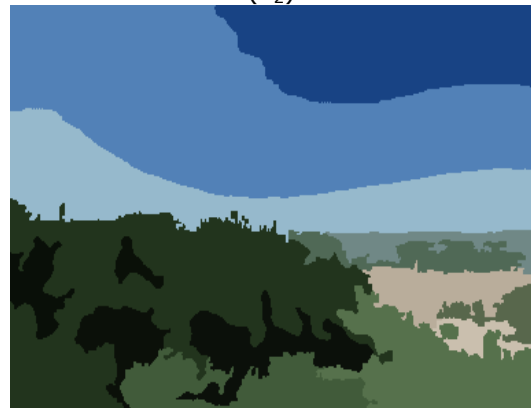
(δ_1)



(δ_2)



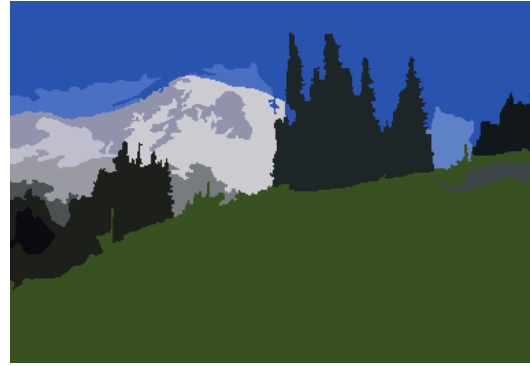
(ϵ_1)



(ϵ_2)



(ζ₁)



(ζ₂)

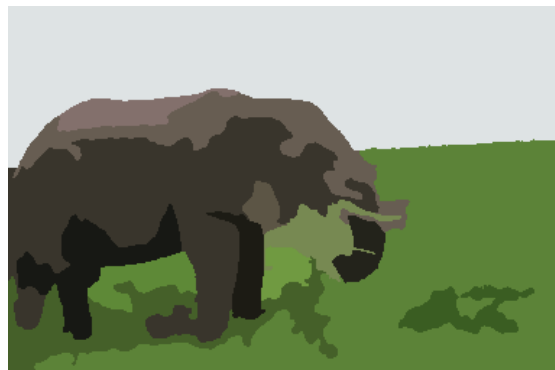
Σχήμα 43: Τμηματοποίηση έγχρωμων εικόνων σε RGB όπου η στατιστική δοκιμασία Mann Whitney έγινε με πλήρες δείγμα.

Τα αποτελέσματα της εφαρμογής του αλγορίθμου για το χρωματικό μοντέλο RGB είναι ικανοποιητικά. Πιο συγκεκριμένα για την εικόνα (β₁) οι σημαντικές περιοχές αποδίδονται σωστά μετά την τμηματοποίηση (β₂) όπως και στην (ζ₁). Στην περίπτωση όμως της (α₁) ο αλγόριθμος δεν έδωσε και τόσο καλά αποτελέσματα αφού το κύριο αντικείμενο της εικόνας (ελέφαντας) θα έπρεπε να αποτελεί ξεχωριστό αντικείμενο από το φόντο (το γρασίδι). Το χρωματικό μοντέλο RGB δεν βοηθάει στο να διαφοροποιηθούν ικανοποιητικά οι αποχρώσεις του καφέ με το πράσινο σε ορισμένα σημεία όπως φαίνεται στην (α₂).

Στη συνέχεια επαναλαμβάνεται ο αλγόριθμος για το χρωματικό μοντέλο HSV, όπου το δείγμα είναι επαρκές, δηλαδή $D=1$, Σχήμα 44.



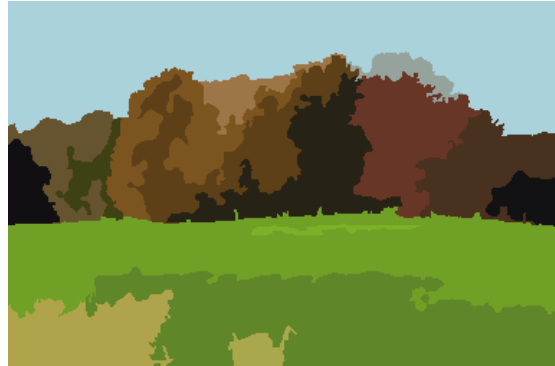
(α₁)



(α₁)



(β_1)



(β_2)



(γ_1)



(γ_2)



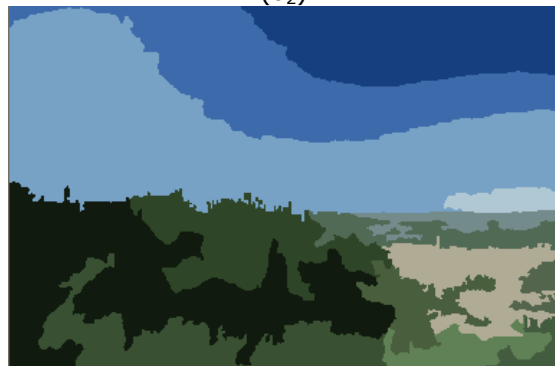
(δ_1)



(δ_2)



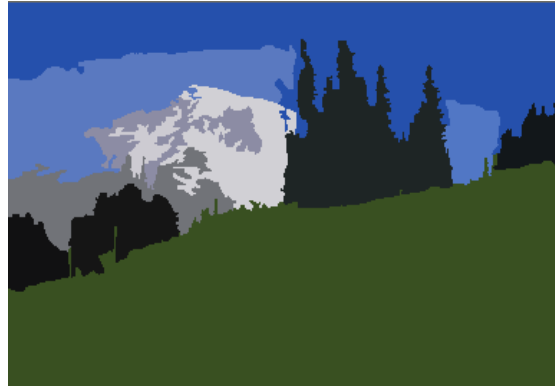
(ϵ_1)



(ϵ_2)



(ζ₁)



(ζ₂)

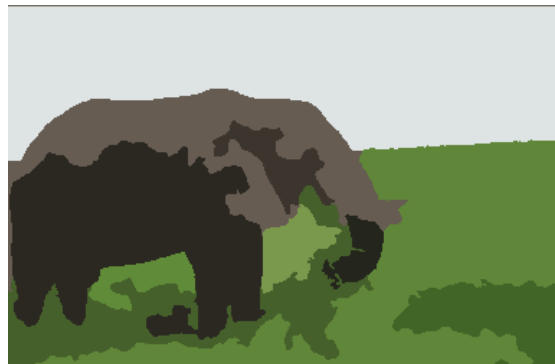
Σχήμα 44: Τμηματοποίηση έγχρωμων εικόνων σε HSV όπου η στατιστική δοκιμασία Mann Whitney έγινε με πλήρες δείγμα.

Στο Σχήμα 44 τα αποτελέσματα είναι πιο αντιπροσωπευτικά σε σχέση με αυτά στο Σχήμα 43, ιδιαίτερα στο παράδειγμα (α₁). Ακόμα στα παραδείγματα (ε₁) και (ζ₁) όπου τα χρώματα δεν είναι εύκολα διαχωρίσιμα κυρίως λόγω της σκίασης ανάμεσα στα αντικείμενα, τα αποτελέσματα (ε₂) και (ζ₂) είναι αρκετά ικανοποιητικά.

Στη συνέχεια επαναλαμβάνεται ο αλγόριθμος για το χρωματικό μοντέλο YUV, όπου το δείγμα είναι επαρκές, δηλαδή D=1, Σχήμα 45.



(α₁)



(α₂)



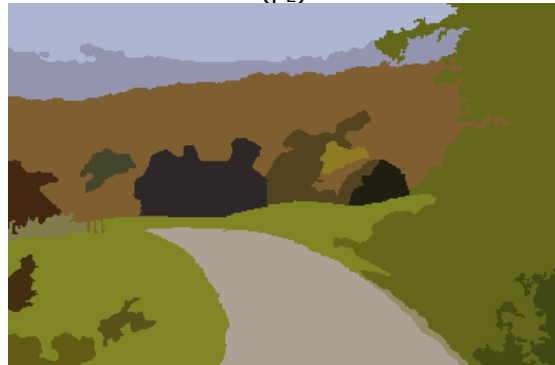
(β_1)



(β_2)



(γ_1)



(γ_2)



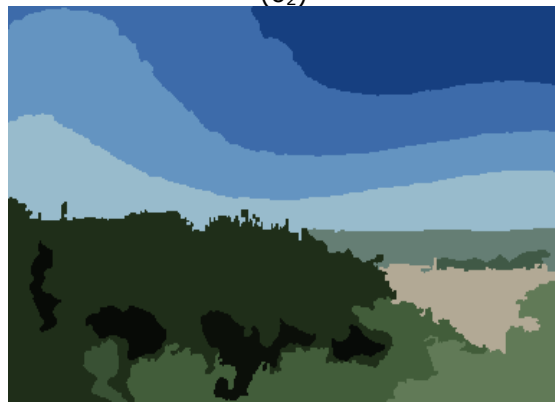
(δ_1)



(δ_2)



(ϵ_1)



(ϵ_2)



(ζ₁)



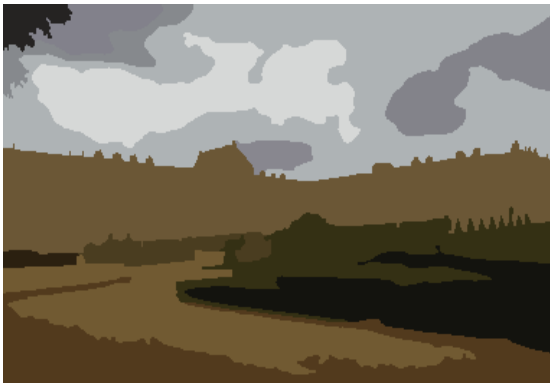
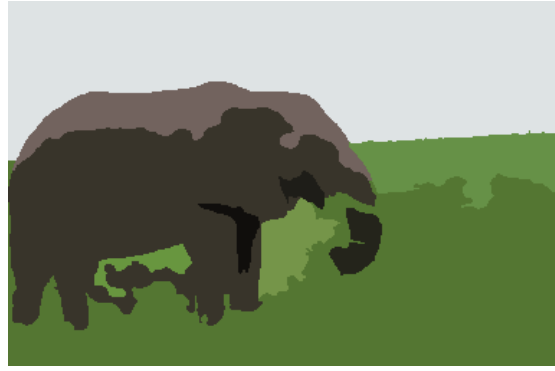
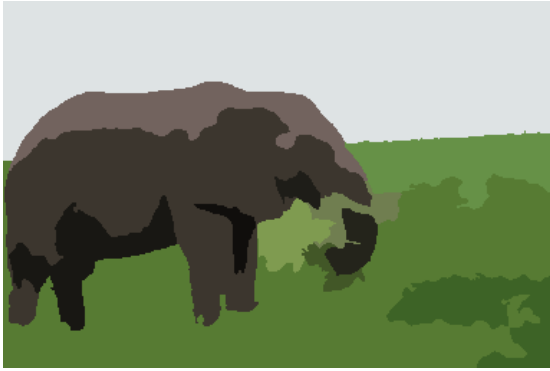
(ζ₂)

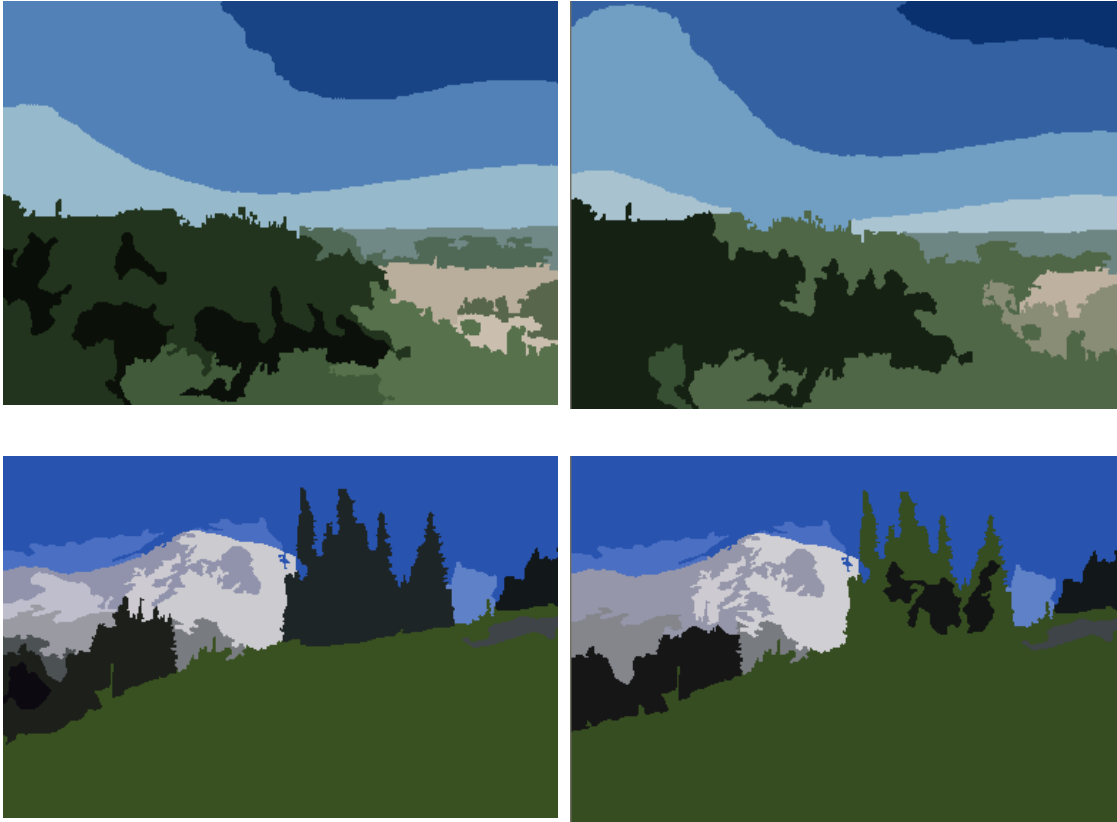
Σχήμα 45: Τμηματοποίηση έγχρωμων εικόνων σε YUV όπου η στατιστική δοκιμασία Mann Whitney έγινε με πλήρες δείγμα.

Με τη χρήση του χρωματικού μοντέλου YUV τα αποτελέσματα (Σχήμα 45) στο δείγμα αυτών των έγχρωμων εικόνων είναι ενθαρρυντικά τόσο σε απλές εικόνες όπως είναι οι (α₁) και (β₁) αλλά και σε πιο πολύπλοκες εικόνες όπως οι (γ₁) και (δ₁) τα αποτελέσματα των εικόνων μετά την τμηματοποίηση (γ₂) και (δ₂) αποδίδουν τις περιοχές της αρχικής εικόνας με ακρίβεια. Σε ακόμα πιο πολύπλοκες εικόνες βέβαια, όπως είναι οι (ε₁) και (ζ₁) όπου υπάρχουν αρκετές περιοχές με σκιές και τέτοια υφή που δεν βοηθούν στην τμηματοποίηση τους ως προς το χρώμα.

4.2 Σύγκριση αποτελεσμάτων για διάφορα μεγέθη του δείγματος

Στην ενότητα 2.2.2 αναφέρθηκε μια ιδιότητα των μη παραμετρικών δοκιμασιών ότι η δοκιμασία διατηρεί την εγκυρότητα της ακόμα και σε μικρό δείγμα. Για το λόγο αυτό δοκιμάστηκε ο εν λόγω αλγόριθμος για μειωμένο δείγμα χρωματικής πληροφορίας. Εδώ παρουσιάζονται τα αποτελέσματα για D=2 δηλαδή το μέγεθος του δείγματος κάθε περιοχής A είναι $n_A^2 = \frac{n_A}{2}$ και τα εικονοστοιχεία έχουν επιλεγεί ανά δύο. Στο Σχήμα 46 συγκρίνεται η τμηματοποίηση σε RGB και D=1, με την αντίστοιχη τμηματοποίηση με D=2.



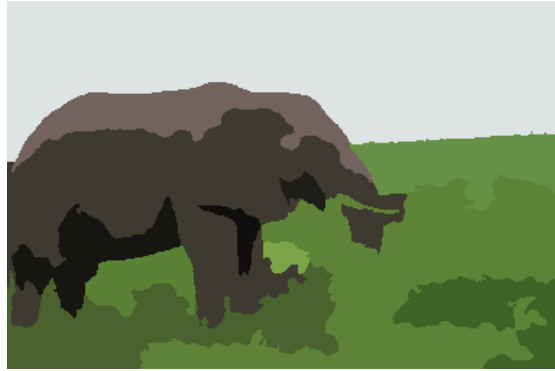
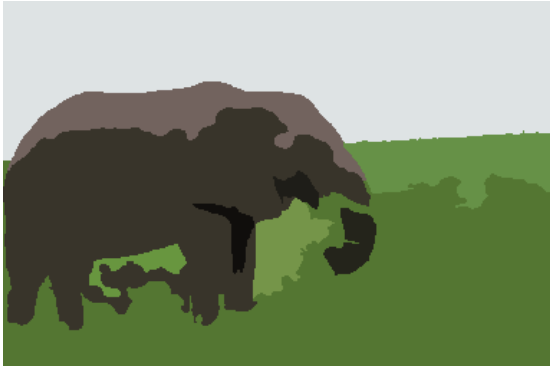


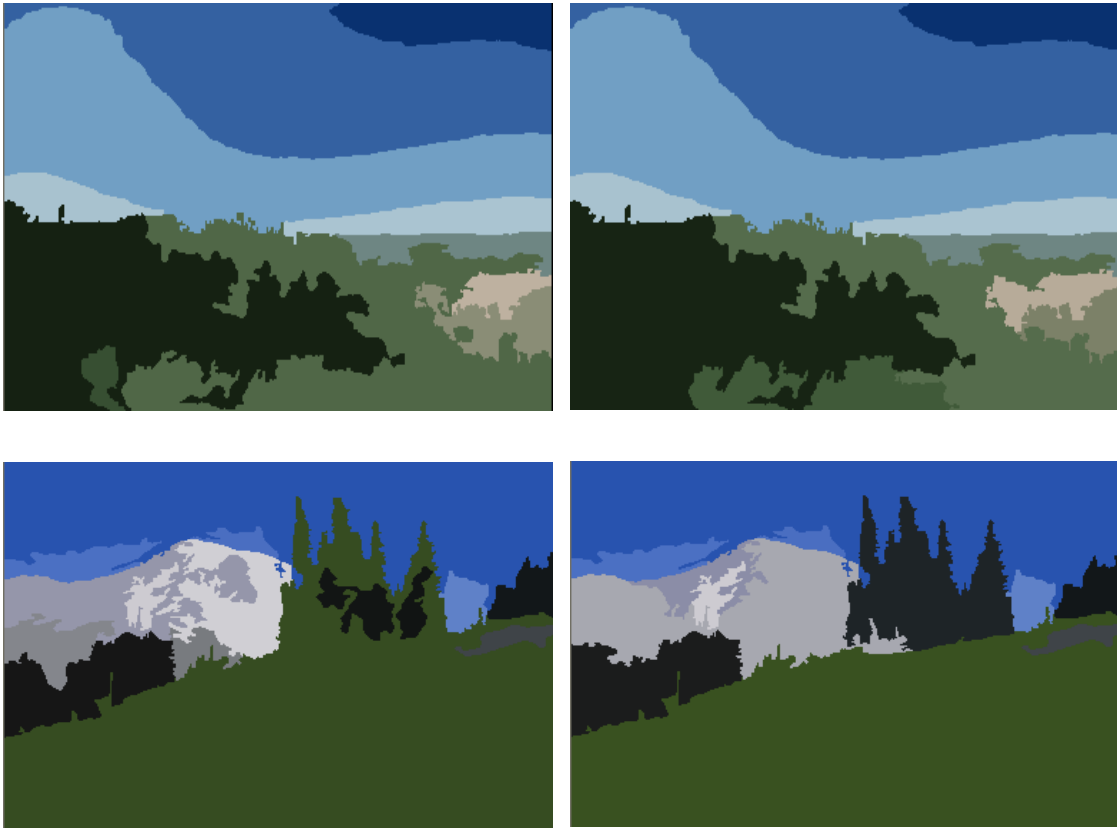
Σχήμα 46: Τμηματοποίηση έγχρωμων εικόνων σε RGB και D=1 ενώ δεξιά με D=2.

Τα αποτελέσματα αποδεικνύουν ότι ο αλγόριθμος είναι αξιόπιστος ακόμα και για μικρότερο δείγμα που λαμβάνεται από κάθε περιοχή. Αυτό είναι αναμενόμενο αφού όπως αναφέρθηκε στην ενότητα 2.2.2 αυτό αποτελεί μια ιδιότητα των μη παραμετρικών διαδικασιών.

Αξίζει να σημειωθεί ότι με την αλλαγή της παραμέτρου D αλλάζει σημαντικά ο υπολογιστικός χρόνος που απαιτείται για να ολοκληρωθεί η τμηματοποίηση της εικόνας. Το φαινόμενο αυτό είναι λογικό επειδή ο αλγόριθμος επεξεργάζεται κάθε στοιχείο των περιοχών που συγκρίνονται. Έτσι όταν μειώνεται το πλήθος των στοιχείων που λαμβάνουν μέρος στον αλγόριθμο, μειώνεται ανάλογα και ο υπολογιστικός χρόνος. Στο συγκεκριμένο παράδειγμα, όπου D=2, ο χρόνος σε όλα τα παραδείγματα μειώθηκε στο μισό ενώ τα αποτελέσματα της τμηματοποίησης δεν αλλοιώθηκαν από τα αντίστοιχα αποτελέσματα με πλήρες δείγμα.

Ακόμα παρουσιάζονται τα αποτελέσματα για D=8 δηλαδή το μέγεθος του δείγματος κάθε περιοχής A είναι $n_A^2 = \frac{n_A}{8}$. Στο Σχήμα 47 συγκρίνεται η τμηματοποίηση σε RGB και D=2, με την αντίστοιχη τμηματοποίηση με D=8.

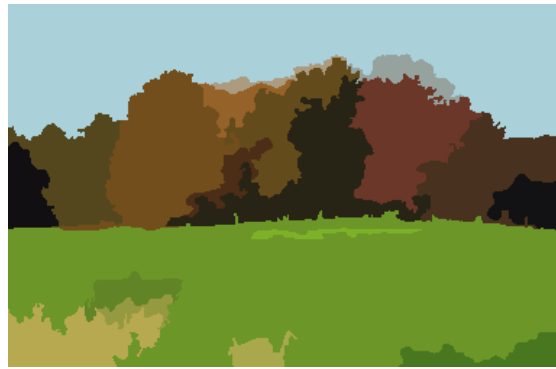
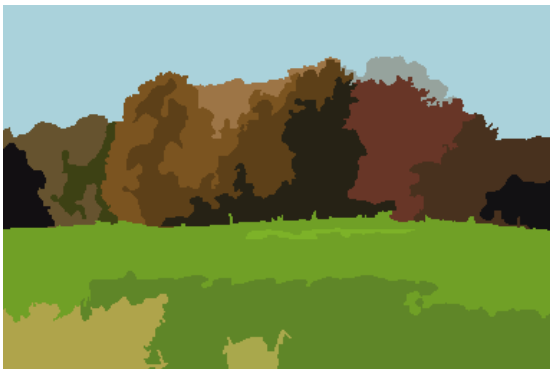
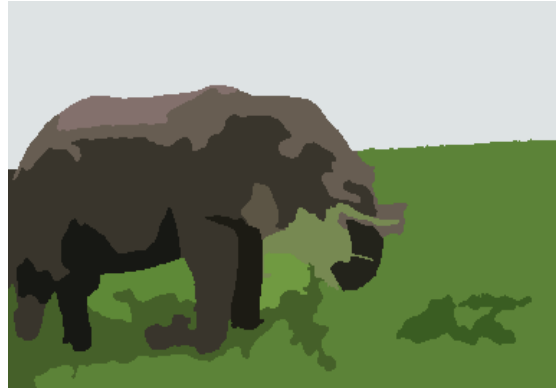
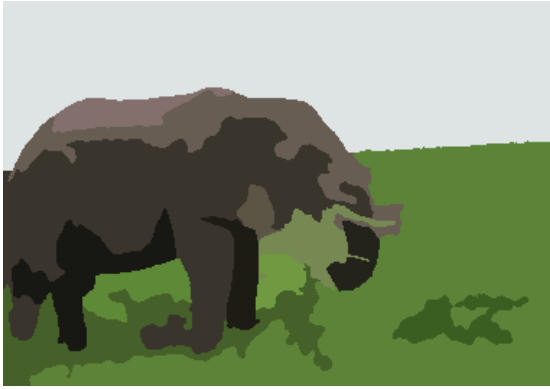


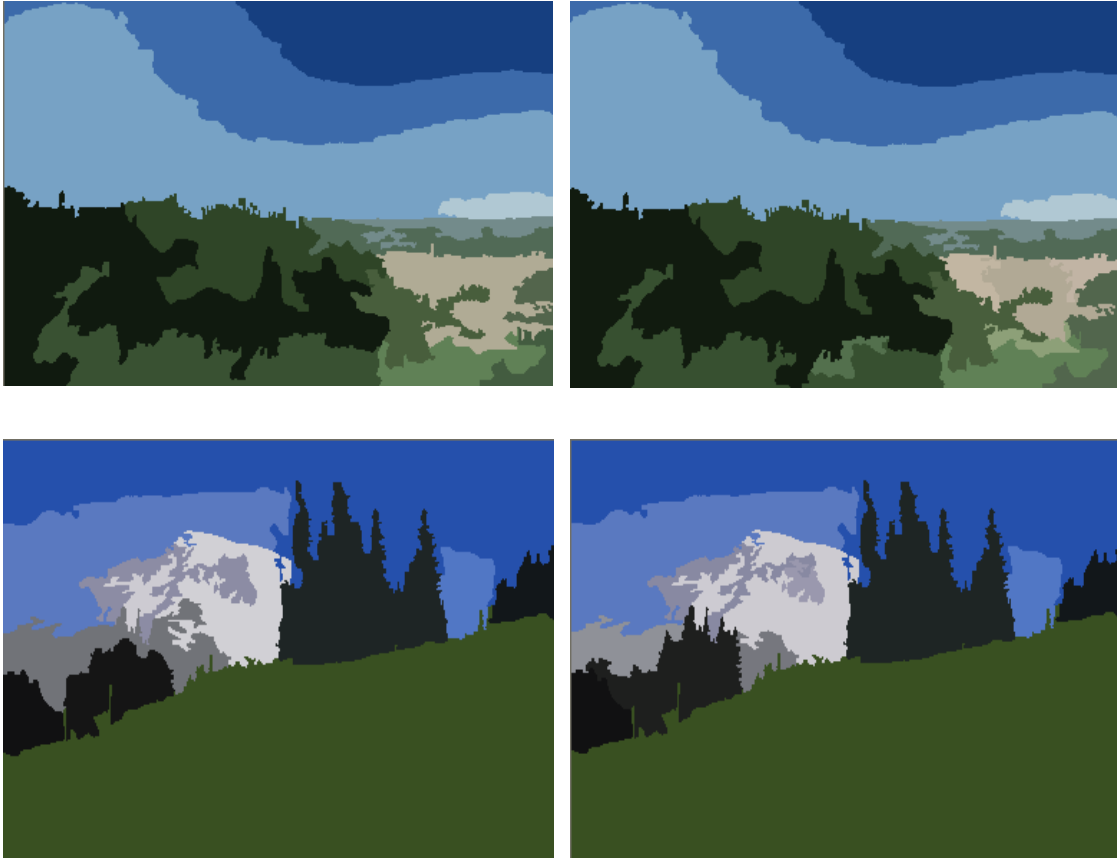


Σχήμα 47: Τμηματοποίηση έγχρωμων εικόνων σε RGB και $D=2$ ενώ δεξιά $D=8$.

Στο συγκεκριμένο παράδειγμα, όπου $D=2$, ο χρόνος σε όλα τα παραδείγματα μειώθηκε στο μισό ενώ τα αποτελέσματα της τμηματοποίησης δεν αλλοιώθηκαν σε σχέση με τα αντίστοιχα στα οποία χρησιμοποιήθηκε δείγμα υπολογισμένο με $D=2$.

Στη συνέχεια επαναλαμβάνεται ο αλγόριθμος για το χρωματικό μοντέλο HSV και $D=2$. Στο Σχήμα 48 συγκρίνονται τα αποτελέσματα της τμηματοποίησης σε HSV και $D=1$, με την αντίστοιχη τμηματοποίηση με $D=2$.

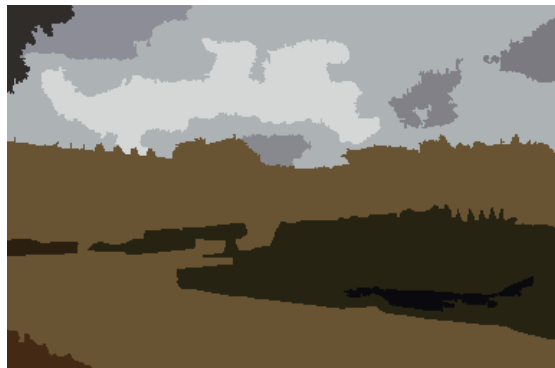
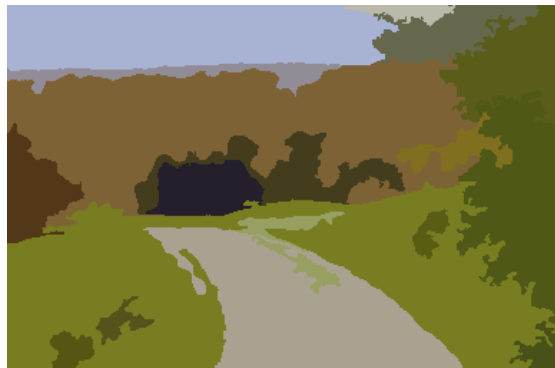
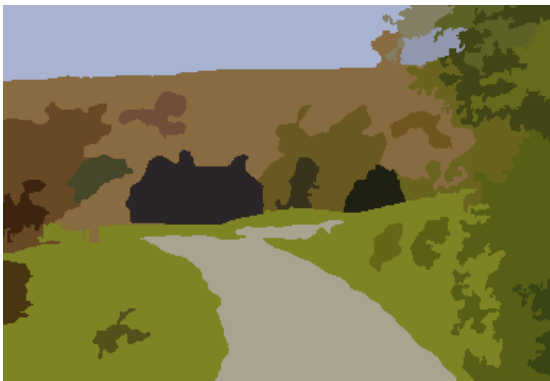
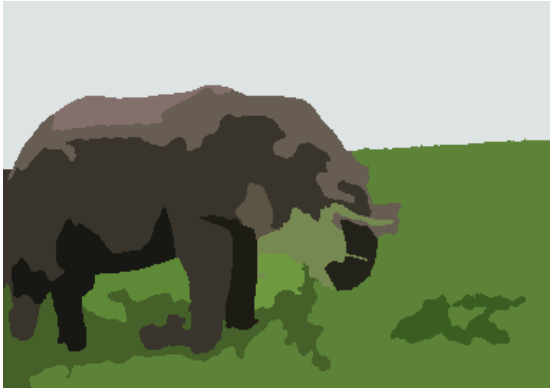


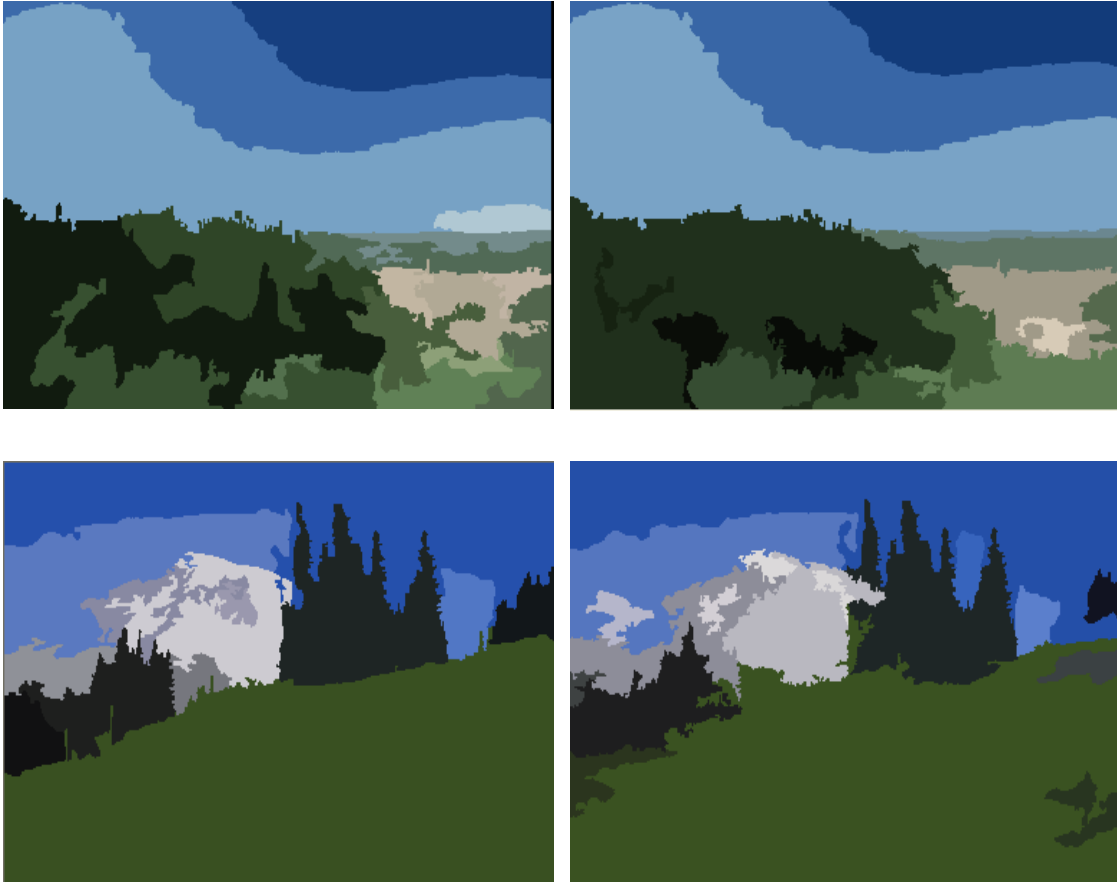


Σχήμα 48: Τμηματοποίηση έγχρωμων εικόνων σε HSV με $D=1$ και $D=2$.

Στο Σχήμα 48 επίσης παρατηρείται ότι για $D=2$, ο χρόνος σε όλα τα παραδείγματα μειώθηκε, ενώ τα αποτελέσματα της τμηματοποίησης με τον συγκεκριμένο αλγόριθμο δεν αλλοιώθηκαν σε σχέση με τα αντίστοιχα αποτελέσματα πλήρους δείγματος.

Στη συνέχεια επαναλαμβάνεται ο αλγόριθμος για το χρωματικό μοντέλο HSV και $D=2$. Στο Σχήμα 49 συγκρίνονται τα αποτελέσματα της τμηματοποίησης σε HSV και $D=2$, με την αντίστοιχη τμηματοποίηση με $D=8$.

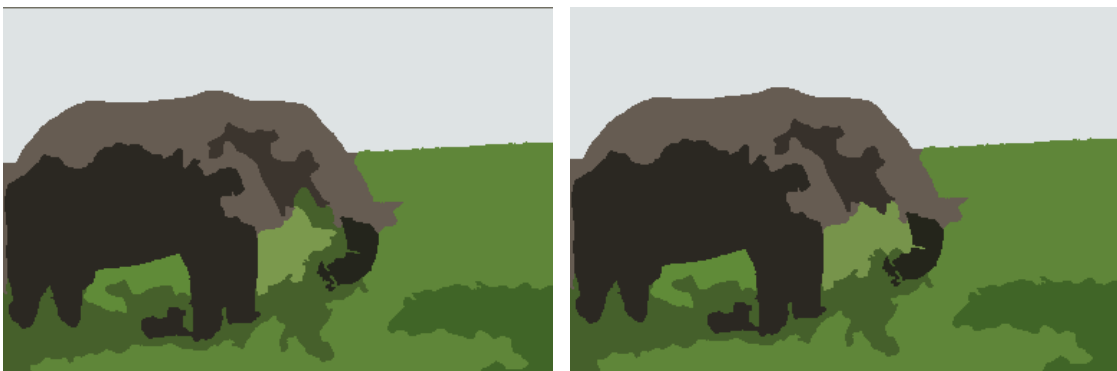


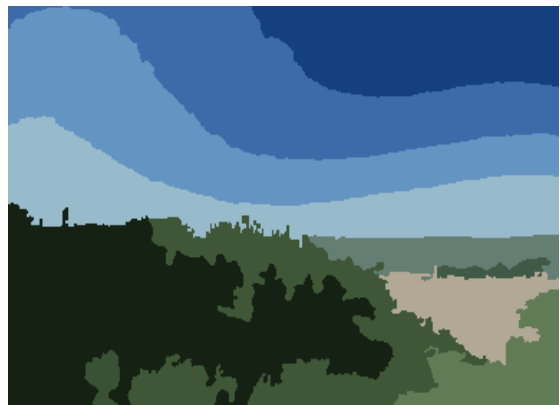
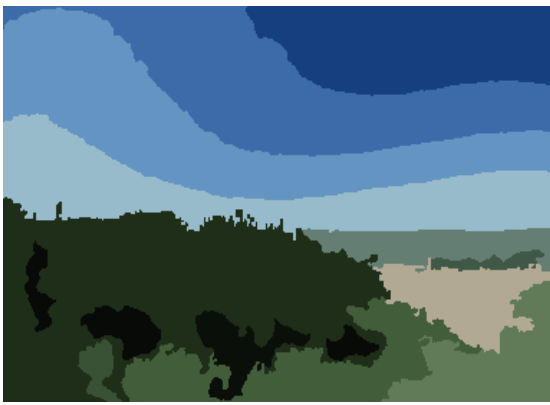
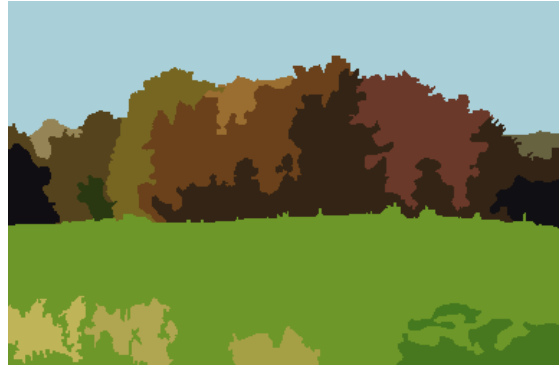
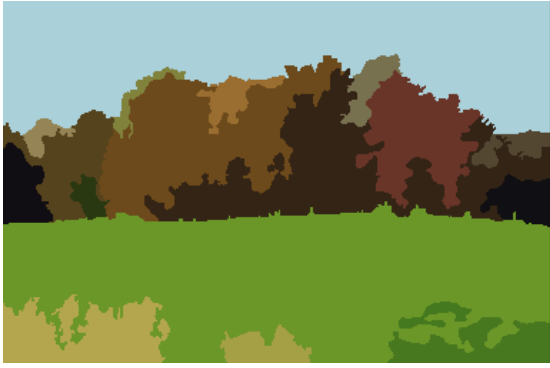


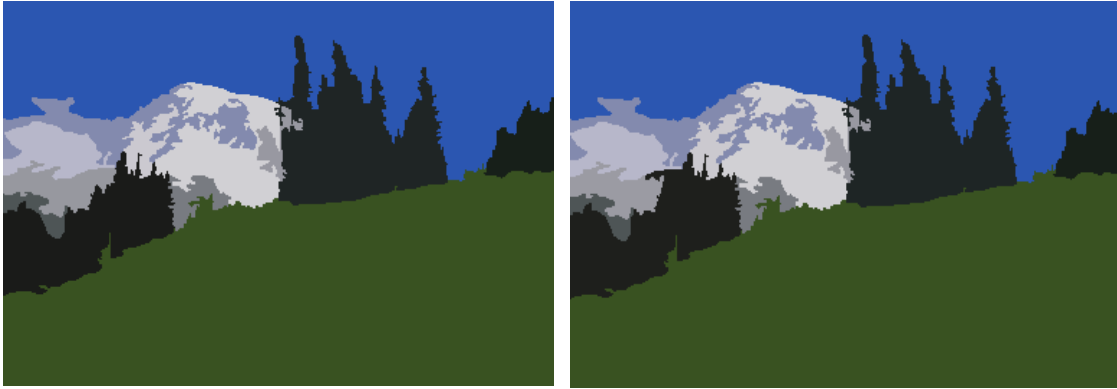
Σχήμα 49: Τμηματοποίηση έγχρωμων εικόνων σε HSV με $D=2$ και $D=8$.

Ομοίως και στο παράδειγμα αυτό επαληθεύεται ότι ο χρόνος μειώνεται ανάλογα με την μείωση του δείγματος.

Έπειτα, επαναλαμβάνεται ο αλγόριθμος για το χρωματικό μοντέλο YUV, για $D=2$. Στο Σχήμα 50 συγκρίνονται τα αποτελέσματα αυτά, με τα αντίστοιχα για $D=1$.





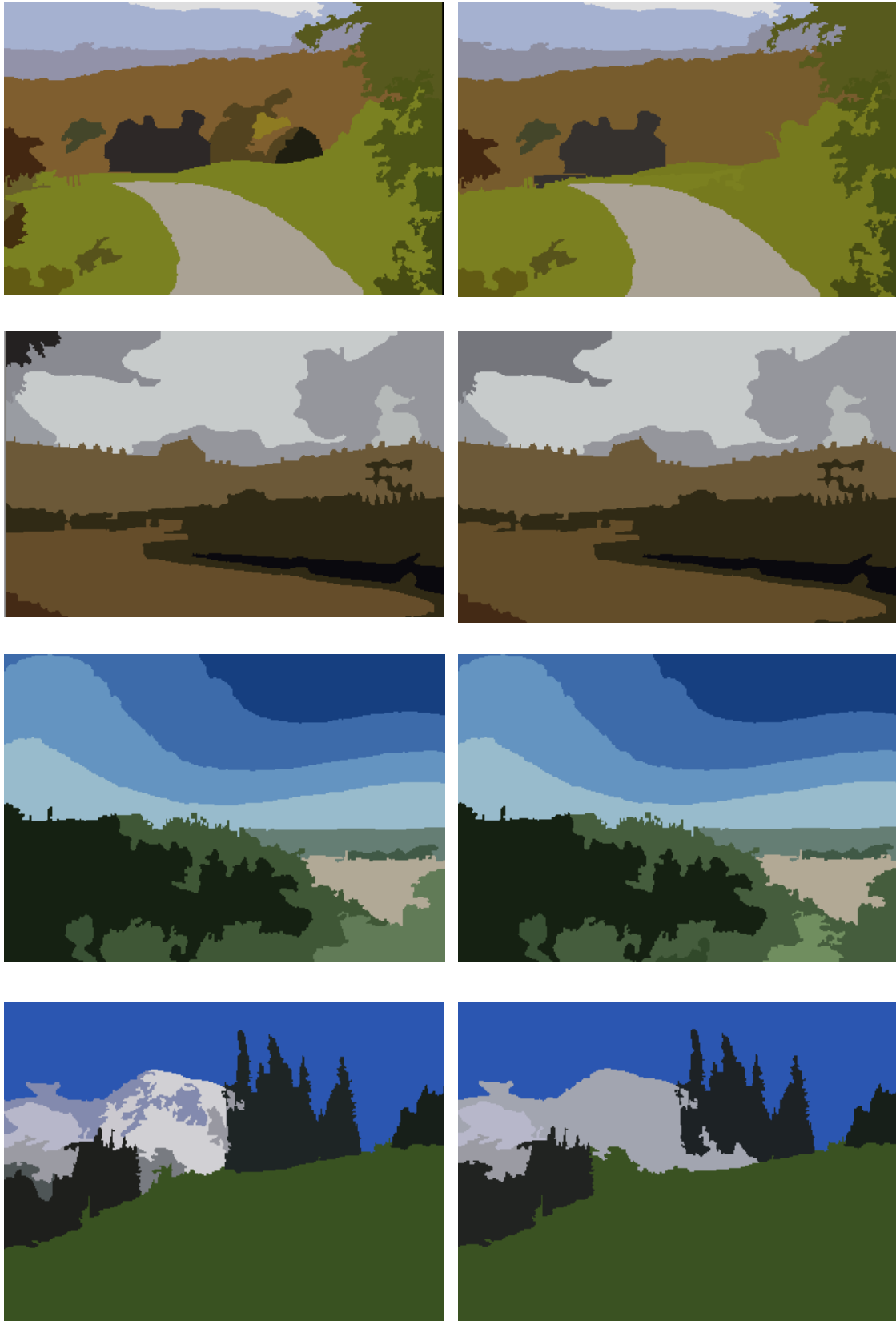


Σχήμα 50: Τμηματοποίηση έγχρωμων εικόνων σε YUV με $D=1$ ενώ δεξιά με $D=2$.

Και με τη χρήση του YUV χρωματικού μοντέλου ο αλγόριθμος διατηρεί την αξιοπιστία του καθώς μειώνεται το στατιστικό δείγμα.

Τέλος, επαναλαμβάνεται ο αλγόριθμος για το χρωματικό μοντέλο YUV, για $D=8$, και στο Σχήμα 51 συγκρίνονται τα αποτελέσματα αυτά, με τα αντίστοιχα για $D=2$.





Σχήμα 51: Τμηματοποίηση έγχρωμων εικόνων σε YUV με $D=2$ ενώ δεξιά με $D=8$.

Συμπερασματικά, το αποτέλεσμα της τμηματοποίησης δεν εξαρτάται άμεσα από το πλήθος του δείγματος που λαμβάνεται για τη σύγκριση δύο περιοχών, ενώ όπως είναι αναμενόμενο ο χρόνος επηρεάζεται από το μέγεθος του δείγματος. Έτσι λοιπόν όσο μειώνεται το δείγμα στο οποίο εφαρμόζεται η δοκιμασία Mann Whitney, τόσο μειώνεται και ο υπολογιστικός χρόνος του πειράματος. Πραγματικά μπορεί κανείς να παρατηρήσει ότι η ισχύς του κριτηρίου της δοκιμασίας Mann Whitney διατηρείται όσο το στατιστικό δείγμα είναι μεγάλο.

Εν κατακλείδι, ο συγκεκριμένος επαναληπτικός γραφοθεωρητικός αλγόριθμος εφαρμόστηκε σε πολλές εικόνες και για διάφορες τιμές του D . Έπειτα από πολλά πειράματα παρατηρήθηκε πως τα αποτελέσματα της τμηματοποίησης σε μεγάλο βαθμό αντιστοιχούν στα φυσικά αντικείμενα που απεικονίζονται στην αρχική εικόνα για $D \leq 8$.

Στα παρακάτω σχήματα (Σχήμα 52, Σχήμα 53, Σχήμα 54) παρουσιάζονται τα αποτελέσματα του αλγορίθμου για $D=10$ όπου επιβεβαιώνεται ότι το κριτήριο ομοιότητας δύο περιοχών δεν δίνει αξιόπιστα αποτελέσματα σε περιπτώσεις όπου το στατιστικό δείγμα είναι πολύ μικρό, όταν δηλαδή το πλήθος των εικονοστοιχείων μιας περιοχής είναι πολύ μικρό.



(α_1)



(α_2)



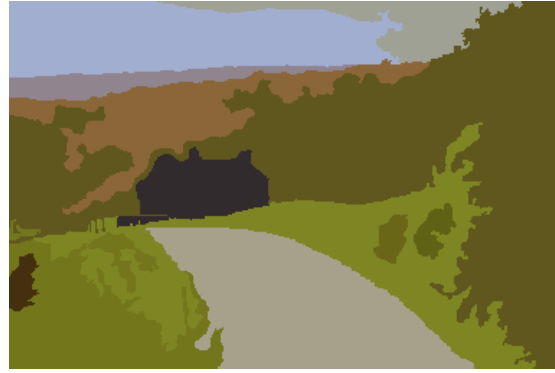
(β_1)



(β_2)



(γ_1)



(γ_2)



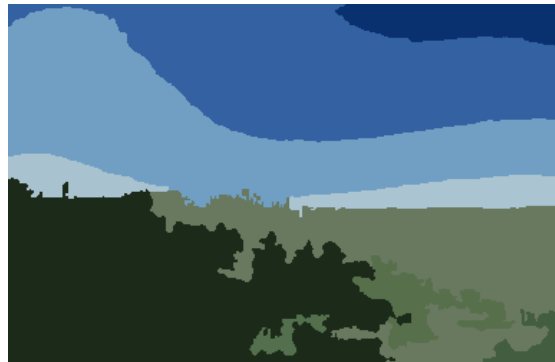
(δ_1)



(δ_2)



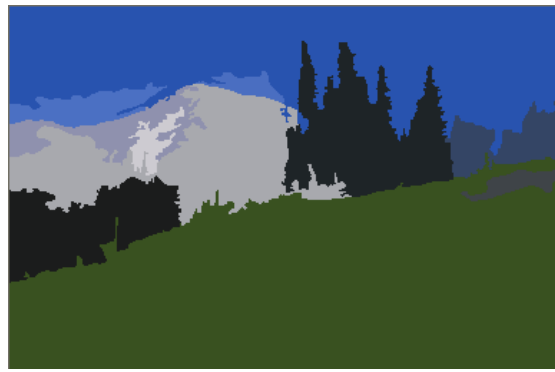
(ϵ_1)



(ϵ_2)



(ζ_1)



(ζ_2)

Σχήμα 52: Αριστερά, η αρχική εικόνα και δεξιά η τμηματοποίηση της εικόνας μετά την εφαρμογή του αλγορίθμου σε RGB με $D=10$



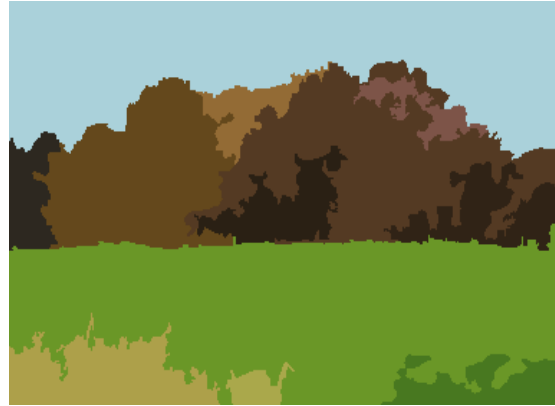
(α_1)



(α_2)



(β_1)



(β_2)



(γ_1)



(γ_2)



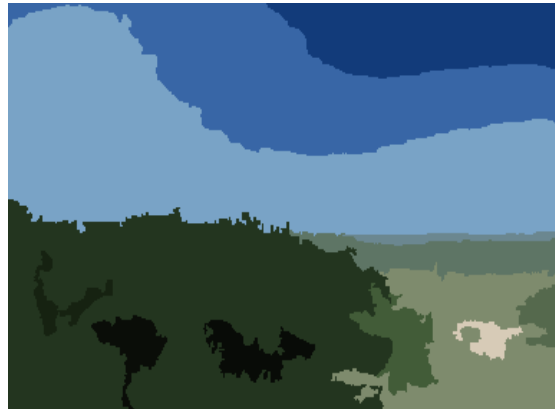
(δ_1)



(δ_2)



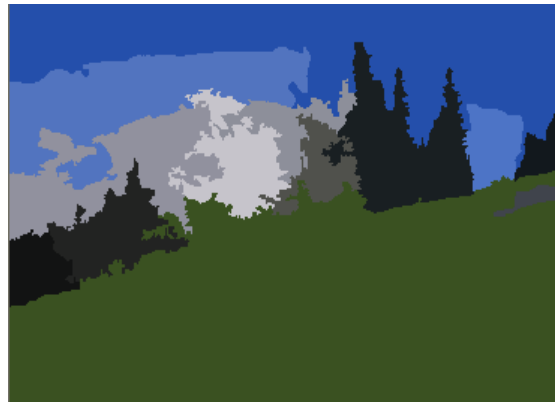
(ϵ_1)



(ϵ_2)



(ζ_1)

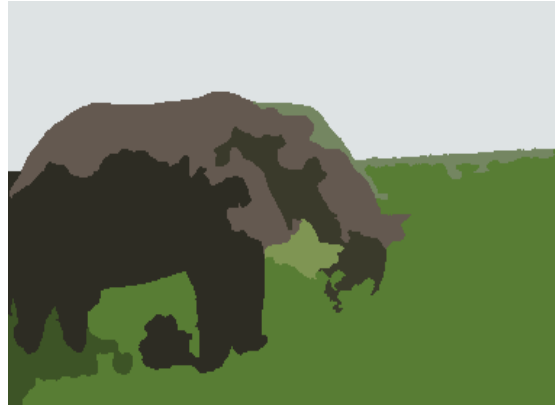


(ζ_2)

Σχήμα 53: Αριστερά, η αρχική εικόνα και δεξιά η τμηματοποίηση της εικόνας μετά την εφαρμογή του αλγορίθμου σε HSV με $D=10$.



(α_1)



(α_2)



(β_1)



(β_2)



(γ_1)



(γ_2)



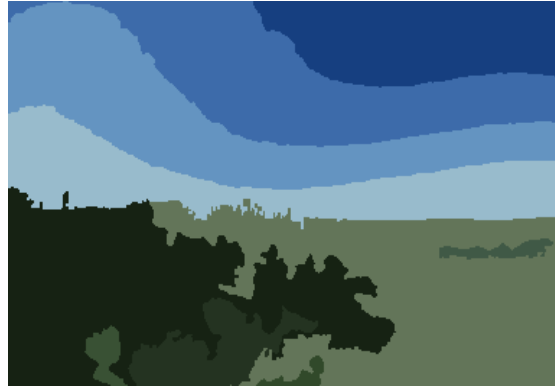
(δ_1)



(δ_2)



(ε₁)



(ε₂)



(ζ₁)



(ζ₂)

Σχήμα 54: Αριστερά, η αρχική εικόνα και δεξιά η τμηματοποίηση της εικόνας μετά την εφαρμογή του αλγορίθμου σε YUV με $D=10$.

5 Συμπεράσματα

Στην εργασία αυτή παρουσιάστηκε ένας αλγόριθμος για τμηματοποίηση έγχρωμων εικόνων με αποκλειστικό κριτήριο το χρώμα. Το πρόβλημα της τμηματοποίησης ανάγεται σε ένα πρόβλημα σύνδεσης κόμβων γράφου. Το βάρος των ακμών του γράφου συμβολίζει τον βαθμό ομοιότητας των αντίστοιχων κόμβων, άρα και την ομοιότητα δύο περιοχών της εικόνας ως προς το χρώμα. Καινοτομία του αλγορίθμου αποτελεί το γεγονός ότι η απόφαση για την ένωση δύο περιοχών της εικόνας λαμβάνεται από μια μη παραμετρική στατιστική δοκιμασία (ενότητα 2.2.2). Έτσι η τμηματοποίηση της εικόνας γίνεται χωρίς υποθέσεις για την κατανομή που ακολουθεί το χαρακτηριστικό ως προς το οποίο γίνεται η τμηματοποίηση. Επιπρόσθετα, δεν απαιτείται εκ των προτέρων γνώση του πλήθους των περιοχών στις οποίες πρέπει να διαμεριστεί η εικόνα. Στην εν λόγω εργασία γίνεται εφαρμογή της μη παραμετρικής στατιστικής δοκιμασίας Mann Whitney (ενότητα 2.2.2.1) για τμηματοποίηση εικόνων ως προς το χρώμα.

Τα αποτελέσματα της τμηματοποίησης του αλγορίθμου είναι ενθαρρυντικά, δεδομένου ότι δεν γίνεται καμία υπόθεση για την κατανομή του χρώματος, και επομένως δεν χρησιμοποιούνται στατιστικά μεγέθη όπως μέση τιμή, διασπορά, τα οποία απαιτούν την γνώση της κατανομής του χρώματος ή κάποια παραδοχή για την κατανομή. Στα παραδείγματα στο Σχήμα 42 φαίνεται πως σε περιπτώσεις όπως είναι οι εικόνες (α), (β), (ζ) όπου το χρώμα χαρακτηρίζει ικανοποιητικά τα αντικείμενα, ο αλγόριθμος έδωσε τις σημαντικότερες περιοχές ως προς το χρώμα σαν ξεχωριστά αντικείμενα. Ενώ στην εικόνα (δ) όπου η υφή έχει καθοριστικό ρόλο για τον χαρακτηρισμό των αντικειμένων, έχουν ομαδοποιηθεί περιοχές της εικόνας που έχουν μεν παρόμοιο χρώμα αλλά δεν αποτελούν κάποιο φυσικό αντικείμενο.

Ο αλγόριθμος εφαρμόστηκε για τα χρωματικά μοντέλα RGB, HSV, YUV για τις διάφορες τιμές της παραμέτρου που καθορίζει το μέγεθος των δειγμάτων. Μπορεί κανείς να διακρίνει πως τα αποτελέσματα της τμηματοποίησης για το χρωματικό μοντέλο YUV, Σχήμα 50 και Σχήμα 51, δεν επηρεάζονται από την μείωση των δειγμάτων όσο τα RGB, YUV. Έτσι, μειώνεται ο υπολογιστικός χρόνος του αλγορίθμου χωρίς να μεταβάλλονται σημαντικά το αποτέλεσμα της τμηματοποίησης.

Επιπρόσθετα, από τα αποτελέσματα της εφαρμογής του αλγορίθμου για διάφορα μεγέθη δείγματος επιβεβαιώνεται ότι οι μη παραμετρικές στατιστικές δοκιμασίες ομοιογένειας έχουν εφαρμογή σε οποιαδήποτε δείγματα, αρκεί αυτά να θεωρούνται «μεγάλα». Στην στατιστική αποδεικνύεται ότι ένα «μεγάλο» δείγμα

αποτελείται από τουλάχιστον 30 στοιχεία. Έτσι, δεδομένου ότι η μικρότερη περιοχή στην εν λόγω εργασία ορίστηκε να είναι μεγέθους μεγαλύτερου από 200 εικονοστοιχεία (παράγραφος 3.1.3), η μέγιστη τιμή που μπορεί να έχει η παράμετρος D είναι 8.

6 ΜΕΛΛΟΝΤΙΚΕΣ ΕΠΕΚΤΑΣΕΙΣ

Τα αποτελέσματα της τμηματοποίησης που παρουσιάστηκαν στο κεφάλαιο 4 είναι ενθαρρυντικά, δεδομένου ότι στην εργασία αυτή αποκλειστικό κριτήριο για την ομοιότητα των περιοχών της εικόνας είναι η χρωματική πληροφορία. Ειδικά στις περιπτώσεις όπου τα αντικείμενα της εικόνας χαρακτηρίζονται κυρίως από το χρώμα ο αλγόριθμος διαμερίζει την εικόνα με αρκετή ακρίβεια σε τόσες περιοχές όσα είναι και τα φυσικά αντικείμενα που απεικονίζονται σε αυτήν.

Μια επέκταση του αλγορίθμου αποτελεί ο συνδυασμός διάφορων κριτηρίων για την τμηματοποίηση έγχρωμων εικόνων. Στην εν λόγω εργασία ο αλγόριθμος εφαρμόστηκε αποκλειστικά για τη χρωματική πληροφορία κάθε εικονοστοιχείου. Ιδιαίτερο ενδιαφέρον έχει η εφαρμογή του αλγορίθμου χρησιμοποιώντας επιπρόσθετη πληροφορία, όπως για παράδειγμα την υφή (texture). Έτσι, ο αλγόριθμος θα μπορούσε να εφαρμοστεί στην περίπτωση όπου τα χαρακτηριστικά κάθε εικονοστοιχείου δίνονται από ένα διάνυσμα όχι μόνο με τρεις συνιστώσες (για κάθε συνιστώσα του χρωματικού μοντέλου) αλλά με μια συνιστώσα επιπλέον για την υφή. Στα παραδείγματα (δ) και (ε) στο Σχήμα 42 η πληροφορία της υφής είναι απαραίτητη για μια τμηματοποίηση σε πιο ευδιάκριτα φυσικά αντικείμενα. Στα αποτελέσματα που παρουσιάζονται στο Σχήμα 43, Σχήμα 44, Σχήμα 45 για τις αντίστοιχες εικόνες είναι φανερό πως ένα τέτοιο επιπλέον κριτήριο μπορεί να φανεί ιδιαίτερα χρήσιμο.

Επιπλέον μια πιο εκτεταμένη πειραματική επαλήθευση του αλγορίθμου για περισσότερες εικόνες αλλά και διάφορες τιμές των παραμέτρων που χρησιμοποιήθηκαν είναι δυνατό να βελτιώσουν σημαντικά τα αποτελέσματα της τμηματοποίησης. Ακόμα, η ποσοτική ανάλυση του υπολογιστικού χρόνου του αλγορίθμου για διάφορες τιμές των δειγμάτων κάθε κόμβου, μπορεί να συντελέσει στην εξαγωγή γενικευμένων συμπερασμάτων.

Τέλος, επεμβάσεις στον αλγόριθμο ως προς την προεπεξεργασία της εικόνας είναι δυνατό να βελτιώσουν το αποτέλεσμα της τμηματοποίησης. Η υπερτμηματοποίηση έχει καθοριστικό ρόλο στη δημιουργία των αρχικών κόμβων του γράφου. Όσο πιο ομογενείς είναι οι αρχικές περιοχές που δημιουργούνται στην υπερτμηματοποίηση, τόσο πιο αντιπροσωπευτικά είναι τα αποτελέσματα της τμηματοποίησης στα φυσικά αντικείμενα από τα οποία αποτελείται η εικόνα.

Εν κατακλείδι, θα μπορούσε κανείς να προσδιορίσει μια συγκεκριμένη εφαρμογή της τμηματοποίησης εικόνων, έτσι ώστε να χρησιμοποιηθούν τέτοιες τιμές παραμέτρων και κατάλληλη επεξεργασία εικόνων ώστε ο αλγόριθμος να αποδώσει τις περιοχές ενδιαφέροντος της συγκεκριμένης εφαρμογής.

7 Αναφορές

- [1] David A. Forsyth and Jean Ponce, "Computer Vision: A Modern Approach", Prentice Hall, 14 August, 2002.
- [2] Αριστειδης Αντωνίου, Μουρίκης Δημήτρης, Λέκτορας Πανεπιστημίου Αθηνών, "Ατλας υπερηχοτομογραφίας," 1^ο κεφάλαιο : Ορφανουδάκης Στέλλιος, Καθηγητής Πανεπιστημίων Yale και Κρήτης.
- [3] Y. W. Lim and S. U. Lee, "On the color image segmentation algorithm based on the thresholding and the fuzzy C-means technique," Pattern Recognition, vol.3,no.9, pp935-952, 1990.
- [4] "Digital Image Processing" W. K. Pratt, 2nd edition, J. Wiley and sons, 1991.
- [5] K. Mardia and T. Hainsworth. A Spatial Thresholding Method for Image Segmentation. IEEE Trans. On Pattern Analysis. And Mach. Intell., 10(6):919—927, November 1988.
- [6] D. Geiger and F. Girosi. Parallel and Deterministic Algorithms from MRF's: Surface Reconstruction. IEEE Trans. On Pattern Anal. And Mach. Intell, 13(5):401—412, May 1991.
- [7] D. Geiger and A.Yuille. "A Common Framework for Image Segmentation" Int. Journal of Comp. Vision, 6(3):227—243, 1991.
- [8] P. L. Palmer, H. Dabis, and J. Kittler, "A performance measure for boundary detection algorithms," Computer Visual Image Understanding, vol.63, pp.476-494, 1996.
- [9] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," in Proc. 1st ICCV, 1987, pp.259-267.
- [10] A. K. Jain, "Fundamentals of digital image processing," Prentice Hall, 1989.
- [11] Alan Gibbons, Algorithmic Graph Theory, Cambridge University Press, 1985.
- [12] "Efficient Graph-Based Image Segmentation," Pedro F. Felzenszwalb, Artificial Intelligence Lab, Massachusetts Institute of Technology Daniel P. Huttenlocher, Computer Science Department, Cornell University.

[13] Jianbo Shi and Jitendra Malik, Member, IEEE, "Normalized Cuts and Image Segmentation," IEEE Trans. Pattern Analysis on Machine Intelligence, vol. 22, no.8, August 2000.

[14] Z. Wu and R. Leahy, "An Optimal Graph Theoretic Approach to Data Clustering: Theory and Its Application to Image Segmentation," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 15, no. 11, pp. 1,101-1,113, Nov. 1993.

[15] T. Pavlidis and Y. T. Liow, "Integrating region growing and edge detection," IEEE Trans. Pattern Analysis Machine Intelligence, vol. 12, pp.225-233, 1990.

[16] J. Haddon and J. Boyce, "Image segmentation by unify region and boundary information," IEEE Trans. Pattern Analysis Machine Intelligence, vol. 12, pp.929-948, 1990.

[17] C. Chu and J. K. Aggarwal, "The integration of image segmentation maps using region and edge information," IEEE Trans. Pattern Analysis Machine Intelligence, vol. 15, pp.1241-1252, 1993.

[18] Χάρης Κωνσταντίνος , "Ένας υβριδικός αλγόριθμος για την τμηματοποίηση δισδιάστατων και τρισδιάστατων εικόνων," Πανεπιστήμιο Κρήτης, Σχολή Θετικών Επιστημών, Τμήμα Επιστήμης Υπολογιστών.

[19] Jianping Fan, David. K. Y. Yau, Member, IEEE, Ahmed. K. Elmagarmid, Senior Member, IEEE, and Walid G. Aref, Member, IEEE
"Automatic Image Segmentation by Integrating Color-Edge Extraction and Seeded Region Growing"

[20] Lothar Hermes, Thomas Zoller and Joachim M. Buhmann Rheinische Friedrich Wilhelms, Universitat Institut fur Informatik III, Romerstr. 164,
"Parametric Distributional Clustering for Image Segmentation"

[21] Michael J. Black, Member, IEEE, Guillermo Sapiro, Member, IEEE, David H. Marimont, Member, IEEE, and David Heeger," Robust Anisotropic Diffusion," IEEE Transactions on Image Processing, vol.7, no.3, March 1998.

[22] "Handbook of Parametric and Nonparametric Statistical Procedures," Third Edition by David Sheskin.

-
- [23] Ιωάννου Αν. Αναστασιάδου "Στοιχεία Ανώτερων Μαθηματικών και Στατιστικής II," σελίδες 191 – 247, Θεσσαλονίκη 1976.
- [24] Ιωάννου Αν. Αναστασιάδου, "Στοιχεία Ανώτερων Μαθηματικών και Στατιστικής II," σελίδες 175 – 182, Θεσσαλονίκη 1976.
- [25] Ιωάννου Αν. Αναστασιάδου «Στοιχεία Ανώτερων Μαθηματικών και Στατιστικής II» σελίδες 410 – 431 Θεσσαλονίκη 1976.
- [26] P. Perona and J. Malik, "Scale-space and edge detection using Anisotropic Diffusion," IEEE Trans. Pattern Analysis on Machine Intelligence., vol. 12, no. 7 pp. 629–639, July 1990.
- [27] P. J. Huber, "Robust Statistics," New York: Wiley, 1981.
- [28] P. J. Rousseauw and A. M. Leroy, "Robust Regression and Outlier Detection," New York: Wiley, 1987.
- [30] Graph Theory by Diestel. R Springer-Verlag, 1997
- [31] Charles Zahn, "Graph theoretical methods for detecting and describing Gestalt clusters," IEEE Transactions on Computation, vol. 20, pp. 68-86, 1971.
- [32] "Isoperimetric Graph Partitioning for Data Clustering and Image Segmentation" Leo Grady and Eric L. Schwartz, July 29, 2003