

546
AY



ΤΕΧΝΟΛΟΓΙΚΟ ΕΚΠΑΙΔΕΥΤΙΚΟ ΙΔΡΥΜΑ ΠΕΙΡΑΙΑ

ΣΧΟΛΗ ΤΕΧΝΟΛΟΓΙΚΩΝ ΕΦΑΡΜΟΓΩΝ

Τμήμα Αυτοματισμού

**Σχεδίαση και κατασκευή συστήματος
μηχανικής όρασης**

Πτυχιακή εργασία

ΤΟΥ

Μέρμηγκα Αλέξανδρου

A.M. 30000

Υπεύθυνοι Καθηγητές:

Αλαφοδήμος Κωνσταντίνος, Νικολάου Γρηγόριος

Αθήνα, Νοέμβριος 2012



Ευχαριστίες

Με την ολοκλήρωση αυτής της πτυχιακής θα ήθελα να ευχαριστήσω θερμά τον καθηγητή του ΤΕΙ Πειραιά, στο Τμήμα Αυτοματισμού και επιβλέποντα της παρούσας πτυχιακής εργασίας κ. Γρηγόρη Νικολάου για την υπομονή και εμπιστοσύνη του καθώς και την καθοδήγηση που μου παρείχε.

Τέλος, πρέπει να ευχαριστήσω ιδιαίτερα την οικογένεια μου για την αμέριστη συμπαράσταση και υποστήριξη που μου προσέφερε καθώς επίσης και την Αναστασία που με βοήθησε τόσο ψυχολογικά όσο και έμπρακτα στη διεκπεραίωση αυτής της εργασίας.

Σκοπός της εργασίας

Σκοπός της παρούσας πτυχιακής εργασίας είναι η σχεδίαση και η κατασκευή ενός συστήματος μηχανικής όρασης. Συγκεκριμένα για τον σκοπό αυτό αναπτύχθηκε ένας ολοκληρωμένος αλγόριθμος αναγνώρισης και κατάτμησης της πινακίδας του οχήματος. Ο αλγόριθμος αυτός υλοποιήθηκε με της πλατφόρμα MATLAB R2011a.

Περίληψη Κεφαλαίων

Κεφάλαιο 1^ο : Στο πρώτο κεφάλαιο της πτυχιακής εργασίας παρουσιάζονται ορισμένα εισαγωγικά στοιχεία πάνω στην χρήση της ψηφιακής επεξεργασίας εικόνας. Πληροφορίες όσον αφορά την πρακτική εφαρμογή της σε ποικίλους τομείς όπως στην ιατρική, στην εγκληματολογία, στην αρχαιολογία, στην αστρονομία κ.α..

Κεφάλαιο 2^ο : Σε αυτό το κεφάλαιο περιγράφονται και αναλύονται τα χρωματικά μοντέλα. Τα χρωματικά μοντέλα αναλύουν/συνθέτουν την ψηφιακή εικόνα σε αρκετά μικρά ομοιόμορφα στοιχεία, τα εικονοστοιχεία (pixels).

Κεφάλαιο 3^ο : Σε αυτό το κεφάλαιο γίνεται αναφορά στην ψηφιακή απεικόνιση των εικόνων, ενώ διακρίνονται και οι βασικοί τύποι των εικόνων που χρησιμοποιούνται στο MatLab.

Κεφάλαιο 4^ο : Σε αυτό το κεφάλαιο παρουσιάζονται οι βασικές έννοιες της θεωρίας των συνόλων ώστε να γίνει πιο κατανοητή η μορφολογική επεξεργασία εικόνας. Πιο ειδικά αναλύονται οι έννοιες Dilation, Erosion, Skeletonization, Thinning κ.α..

Κεφάλαιο 5^ο : Το πέμπτο κεφάλαιο αφιερώνεται, ίσως στο πιο σημαντικό και δύσκολο κομμάτι της ψηφιακής επεξεργασίας εικόνας, στην κατάτμηση της ψηφιακής εικόνας. Αναλύονται οι μέθοδοι κατάτμησης με την χρήση Κατωφλίωσης και Ανίχνευσης ακμών.

Κεφάλαιο 6^ο : Σε αυτό το κεφάλαιο παρουσιάζεται το σύστημα αυτόματης αναγνώρισης πινακίδων. Πιο συγκεκριμένα γίνεται επισκόπηση του συστήματος και αναλύονται τα επιμέρους βήματα της υλοποίησης του αλγόριθμου. Ο εντοπισμός της πινακίδας και η κατάτμηση της.

Κεφάλαιο 7^ο : Στο τελευταίο κεφάλαιο παρουσιάζονται κάποια χαρακτηριστικά παραδείγματα από τα αποτελέσματα της χρήσης του αλγορίθμου, ώστε να τονιστούν τα σημεία που ο αλγόριθμος δίνει ιδιαίτερη έμφαση, τις δυσκολίες και τους τρόπους που αντιμετωπίστηκαν.

Λέξεις Κλειδιά:

Ψηφιακή επεξεργασία εικόνας, μηχανική όραση, αυτόματη αναγνώριση χαρακτήρων, κατάτμηση εικόνας, αυτόματη αναγνώριση πινακίδων, ANPR, Automatic Number Plate Recognition, LPR, License Plate Recognition, MatLab.

Πίνακας Περιεχομένων

Κεφάλαιο 1	8
Εισαγωγή	8
1.1 Τι είναι η μηχανική όραση	8
1.2 Τι είναι η επεξεργασία εικόνας	9
1.3 Εφαρμογές της ψηφιακής επεξεργασίας εικόνας	11
1.3.1 Απεικόνιση Ακτινών γ	12
1.3.2 Απεικόνιση Ακτινών X	13
1.3.3 Απεικόνιση στην υπεριώδη ζώνη.....	14
1.3.4 Απεικόνιση στις ορατές και υπέρυθρες ζώνες.....	16
Κεφάλαιο 2	18
Χρωματικά Μοντέλα	18
2.1 Εισαγωγικά	18
2.2 Μοντέλο RGB	19
2.3 Μοντέλο CMY(K)	22
2.4 Μοντέλο HIS	24
Κεφάλαιο 3	28
Ψηφιακή απεικόνιση εικόνων	28
3.1 Εικόνες και ψηφιακές εικόνες	28
3.2 Τύποι ψηφιακών εικόνων	30
3.3 Απόκτηση εικόνας και δειγματοληψία	32
Κεφάλαιο 4	35
Μορφολογική Επεξεργασία Εικόνας	35
4.1 Βασικές έννοιες από την θεωρία συνόλων	35
4.2 Λογικές πράξεις εικόνων	37
4.3 Συστολή και Διαστολή	38
4.3.1 Διαστολή (<i>Dilation</i>).....	39
4.3.2 Συστολή (<i>Erosion</i>).....	40

4.4 Άνοιγμα και Κλείσιμο	42
4.4.1 Άνοιγμα (<i>Opening</i>).....	42
4.4.2 Κλείσιμο (<i>Closing</i>).....	43
4.5 Βασικοί Μορφολογικοί Αλγόριθμοι.....	44
4.5.1 Σκελετοποίηση (<i>Skeletonization</i>)	44
4.5.2 Λέπτυνση (<i>Thinning</i>).....	45
4.5.3 Συμπλήρωση Οπών (<i>Filling Holes</i>).....	46
4.5.4 Συνδεδεμένα Στοιχεία (<i>Connected Components</i>)	47
Ανάλυση Συνδεδεμένων Στοιχείων (<i>Connected Components Analysis</i>)	47
Ονοματισμός Στοιχείων (<i>Components Labeling</i>).....	48
Κεφάλαιο 5.....	49
Κατάτμηση Ψηφιακής Εικόνας	49
5.1 Εισαγωγικά	49
5.2 Κατωφλίωση.....	49
5.2.1 Καθολική Κατωφλίωση (<i>Global Thresholding</i>).....	49
5.2.2 Τοπική Κατωφλίωση (<i>Local Thresholding</i>)	50
5.3 Ανίχνευση ακμών	51
5.3.1 Ανάλυση Ακμών.....	52
5.3.2 Τελεστές Roberts.....	53
5.3.3 Τελεστές Sobel.....	55
5.3.4 Τελεστές Prewitt.....	56
5.4 Εξαγωγή Περιγράμματος	58
Κεφάλαιο 6.....	60
Παρουσίαση Συστήματος Αυτόματης Αναγνώρισης Πινακίδων.....	60
6.1 Εισαγωγικά	60
6.2 Επισκόπηση Συστήματος.....	61
6.3 Διαδικασία Επεξεργασίας.....	63
6.4 Εντοπισμός Πινακίδας	64
6.5 Κατάτμηση Πινακίδας.....	71
Κεφάλαιο 7.....	79
Συμπεράσματα	79

7.1 Πειραματικά αποτελέσματα	79
Βιβλιογραφία	84

Κεφάλαιο 1

Εισαγωγή

1.1 Τι είναι η μηχανική όραση

Οι άνθρωποι είμαστε κυρίως οπτικά όντα. Στηριζόμαστε σε σημαντικό βαθμό στην όρασή μας για να αντιληφθούμε τον κόσμο γύρω μας. Δεν κοιτάμε τα πράγματα μόνο για να τα αναγνωρίσουμε ή να τα ταξινομήσουμε, αλλά και για να εξετάσουμε τις διαφορές τους ή να αποκτήσουμε μια γενική προσέγγιση μίας σκηνής με μια γρήγορη ματιά. Για το λόγο αυτό οι άνθρωποι ανέκαθεν χρησιμοποίησαν οπτικές αναπαραστάσεις, δηλαδή εικόνες, για να εκφράσουν τα συναισθήματά τους ή για να εμπλουτίσουν τις διηγήσεις τους.

Με το πέρασμα των χρόνων οι άνθρωποι έχουν εξελίξει πολύ τις οπτικές τους δεξιότητες. Μπορούν να αναγνωρίσουν ένα πρόσωπο μέσα σε κλάσματα δευτερολέπτων, να ξεχωρίζουν τα χρώματα, να επεξεργαστούν πολύ γρήγορα μεγάλο ποσοστό οπτικών πληροφοριών.

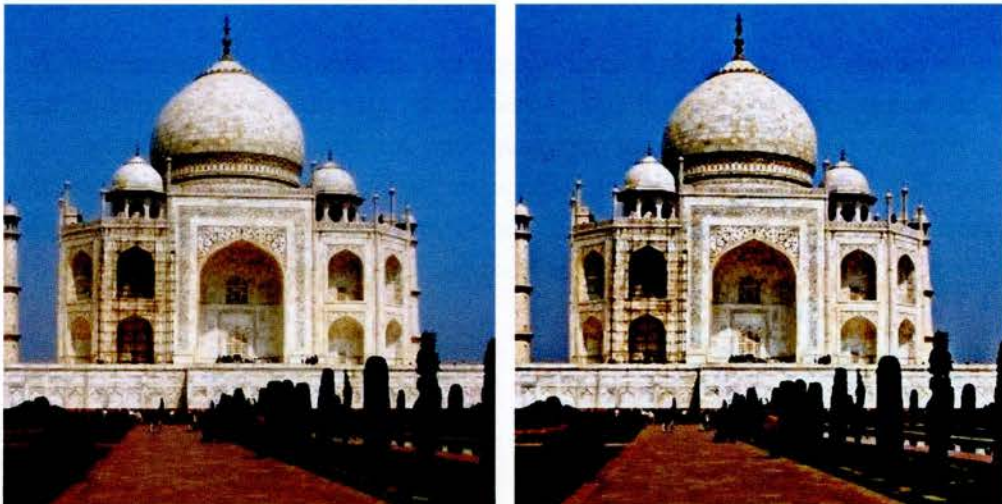
Ωστόσο, ο κόσμος μας βρίσκεται σε συνεχή κίνηση. Κοιτάζοντας κάτι για αρκετή ώρα αυτό θα αλλάξει με κάποιο τρόπο. Ακόμα και μία συμπαγή δομή όπως για παράδειγμα ένα κτίριο ή ένα βουνό, ακόμα και αυτό θα αλλάξει την εμφάνισή του ανάλογα την ώρα της ημέρας (ημέρα ή νύχτα) που το παρατηρούμε, από την ποσότητα του ηλιακού φωτός (φωτεινό ή συννεφιασμένο) που πέφτει πάνω του ή από διάφορες σκιές που υπάγονται σε αυτό.

Η εικόνα σήμερα αποτελεί μία από τις σημαντικότερες πηγές δεδομένων. Ένας τεράστιος όγκος ψηφιακών εικόνων λαμβάνεται καθημερινά σε όλα τα μέρη του κόσμο από ειδικά μηχανήματα με σκοπό να χρησιμοποιηθεί από εξειδικευμένους επιστήμονες. Για παράδειγμα τέτοιες εικόνες λαμβάνονται στον τομέα της ιατρικής, της μετεωρολογίας, της βιολογίας και σε άλλες επιστήμες. Όλες αυτές οι εικόνες περιέχουν πολύτιμα στοιχεία, τα οποία θα πρέπει πρώτα να ανιχνευθούν, να μετρηθούν και στη συνέχεια να καταχωρηθούν σε έναν υπολογιστή όπου θα γίνει η επεξεργασία και η ταξινόμηση τους. Σκοπός της μηχανικής όρασης είναι η άντληση όλων αυτών των στοιχείων με τη χρήση εξειδικευμένων αλγόριθμων και συστημάτων.

1.2 Τι είναι η επεξεργασία εικόνας

Η επεξεργασία εικόνας έχει ως σκοπό την αλλαγή του χαρακτήρα μιας ψηφιακής εικόνας προκειμένου να βελτιώσει τις εικαστικές τις πληροφορίες που περιέχει, ώστε να ερμηνεύονται καλύτερα και πιο κατάλληλα από ένα σύστημα επεξεργασίας εικόνας. Στην παρούσα πτυχιακή εργασία θα ασχοληθούμε με την ψηφιακή επεξεργασία εικόνας, η οποία περιλαμβάνει την χρήση ενός υπολογιστή για να αλλάξουμε την φύση μιας ψηφιακής εικόνας. Στη συνέχεια της ενότητας δίνονται ορισμένες από τις βασικές τεχνικές που χρησιμοποιεί η ψηφιακή επεξεργασία εικόνας.

Ενίσχυση άκμων: Με την ενίσχυση των ακμών μιας ψηφιακής εικόνας επιτυγχάνουμε η παραγόμενη εικόνα να φαίνεται πιο έντονη. Ένα παράδειγμα φαίνεται στην Εικόνα (1), στο οποίο η δεύτερη εικόνα είναι σαφώς πιο ευκρινής. Η ευκρίνεια των ακμών αποτελεί ζωτικό στοιχείο της εκτύπωσης, προκειμένου μια εικόνα να εμφανίζεται όσο καλύτερα γίνεται, στην εκτυπωμένη σελίδα.



α	β
---	---

Εικόνα 1 Ενίσχυση των ακμών μιας εικόνας. (α) Αρχική εικόνα. (β) Η εικόνα μετά την ενίσχυση ακμών.

Αφαίρεση θορύβου: Όλες οι συσκευές καταγραφής, τόσο αναλογικές όσο και ψηφιακές, έχουν χαρακτηριστικά που τις καθιστούν επιρρεπείς στο θόρυβο. Ο θόρυβος είναι ένα πολύ συχνό φαινόμενο στην μετάδοση δεδομένων, το οποίο μπορεί να οφείλεται σε τυχαία σφάλματα ή στα διάφορα ηλεκτρονικά μέρη της συσκευής καταγραφής ή ακόμα και στον αλγόριθμο επεξεργασίας που η συσκευή χρησιμοποιεί. Ο θόρυβος μπορεί να λάβει διαφορετικές μορφές. Έτσι είναι κατανοητό πως κάθε είδος θορύβου απαιτεί και διαφορετική μέθοδο για την αφαίρεση του. Στην Εικόνα 2(α) διακρίνεται θόρυβος με την ονομασία salt and pepper (αλάτι και πιπέρι). Το χαρακτηριστικό αυτού του θορύβου είναι τα λευκά και μαύρα εικονοστοιχεία της εικόνας που είναι πολύ διαφορετικά στο χρώμα ή την ένταση από τα γύρω τους.



α β

Εικόνα 2 Αφαίρεση του θορύβου από μία εικόνα. (α) Αρχική εικόνα. (β) Η εικόνα μετά την αφαίρεση του θορύβου.

Αφαίρεση θολώματος: Το θόλωμα σε μια εικόνα προκαλείται από την κίνηση είτε του φωτογραφιζόμενου αντικείμενου είτε της συσκευής λήψης. Ένα παράδειγμα δίνεται στην Εικόνα (3). Το θόλωμα που προκαλείται από την κίνηση μπορεί να συμβεί όταν η ταχύτητα του κλείστρου της φωτογραφικής μηχανής είναι πολύ μεγάλη σε σχέση με την ταχύτητα του αντικείμενου. Σε φωτογραφίες αντικειμένων που κινούνται γρήγορα : όπως για παράδειγμα αθλητές ή οχήματα το πρόβλημα του θολώματος είναι σημαντικό.



α β

Εικόνα 3 Αφαίρεση του θολώματος από μία εικόνα. (α) Αρχική εικόνα. (β) Η εικόνα μετά την αφαίρεση του θολώματος.

Ανίχνευση ακμών: Η ανίχνευση των ακμών μιας εικόνας μπορεί να είναι απαραίτητη για την μέτρηση των αντικειμένων σε μία εικόνα, ένα τέτοιο παράδειγμα διακρίνεται στην Εικόνα (4) . Εντοπίζοντας τις ακμές μιας εικόνας μπορούμε να μετρήσουμε την περίμετρο ή των εμβαδό των αντικειμένων που περιέχει. Μπορούμε να χρησιμοποιήσουμε και αλγόριθμους ανίχνευσης ακμών ως πρώτο βήμα στην διερεύνηση των ακμών όπως είδαμε παραπάνω. Από τα αποτελέσματα των ακμών, θα δούμε ότι μπορεί να είναι απαραίτητο να ενισχύσουμε ελαφρώς την πρώτη εικόνα ώστε να κάνεις τις ακμές πιο ξεκάθαρες.



α β

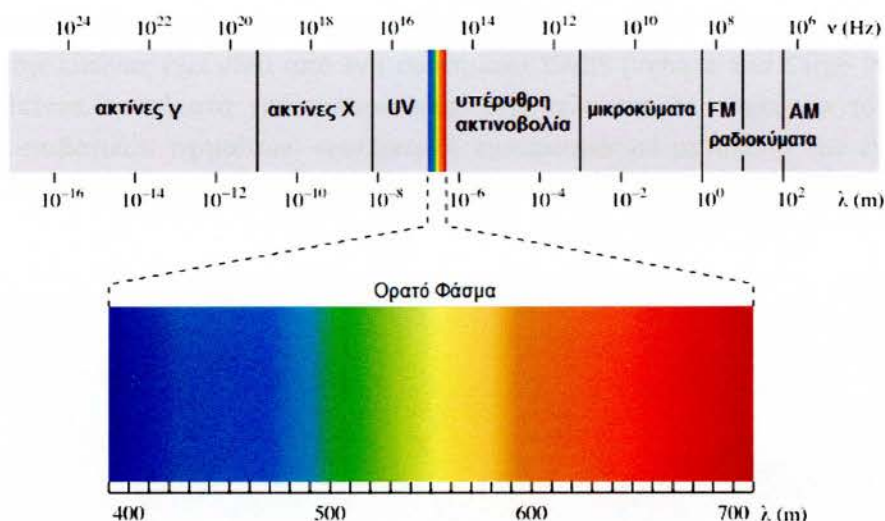
Εικόνα 4 Ανίχνευση ακμών σε μία εικόνα. (α) Αρχική εικόνα. (β) Εφαρμογή ανίχνευσης ακμών.

1.3 Εφαρμογές της ψηφιακής επεξεργασίας εικόνας

Η χρήση εφαρμογών που για την επίλυση προβλημάτων χρησιμοποιούν τον τομέα της ψηφιακής επεξεργασίας εικόνας συνεχώς αυξάνεται. Σήμερα, δεν υπάρχει σχεδόν κανένας τομέας της καθημερινής ζωής μας που να μην επηρεάζεται σε κάποιο βαθμό από την ψηφιακή επεξεργασία. Ποικίλα συστήματα επεξεργασίας εικόνας έχουν διεισδύσει για τα καλά στη ζωή μας, τα οποία χρησιμοποιούνται μέσα στο γραφείο μας, σε καταστήματα, σε εργοστάσια ή ακόμα και σε ιατρικά κέντρα. Ακόμα τέτοια συστήματα έχουν τη δυνατότητα να αναλύουν δορυφορικές φωτογραφίες και να υποβοηθούν στην κατασκευή χαρτών από αεροφωτογραφίες, στην πρόγνωση του καιρού, στην αναγνώριση και παρατήρηση αστικών, γεωργικών ή θαλάσσιων περιοχών. Παρακάτω θα παρουσιαστούν αρκετοί τομείς εφαρμογών, καθένας από τους οποίους χρησιμοποιεί τεχνικές επεξεργασίας εικόνας.

Οι τομείς εφαρμογών της ψηφιακής επεξεργασίας εικόνας είναι τόσοι πολύ, ώστε είναι απαραίτητο να οργανωθούν με κάποιο τρόπο, στην προσπάθεια να γίνει αντιληπτό το εύρος χρήσης τους. Ένας από τους πιο απλούς τρόπους να κατανοήσουμε το πλήθος των εφαρμογών της ψηφιακής επεξεργασίας εικόνας είναι να τις κατηγοριοποιήσουμε ανάλογα με την πηγή τους (για παράδειγμα ακτίνες Χ, ακτίνες γ και ούτω καθεξής).

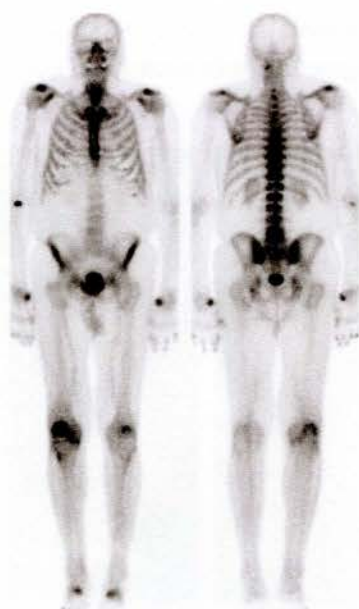
Η κύρια πηγή ενέργειας για τις εικόνες που χρησιμοποιούνται σήμερα είναι το ηλεκτρομαγνητικό φάσμα. Τα ηλεκτρομαγνητικά κύματα μπορούμε να τα αντιληφθούμε σαν σωματίδια που διαδίδονται όπως τα ημιτονοειδή κύματα τα οποία ταξιδεύουν σε ένα κυματοειδή σχηματισμό που κινείται με την ταχύτητα του φωτός. Κάθε ένα από αυτά τα σωματίδια περιέχει ένα ορισμένο ποσό ενέργειας, το οποίο ονομάζεται φωτόνιο. Έτσι αν κατηγοριοποιήσουμε τις φασματικές ομάδες σύμφωνα με την ενέργεια ανά φωτόνιο παίρνουμε το φάσμα που φαίνεται στην Εικόνα (5), και ονομάζεται ηλεκτρομαγνητικό φάσμα. Το ηλεκτρομαγνητικό φάσμα εκτείνεται θεωρητικά από σχεδόν μηδενικές συχνότητες έως το άπειρο και χωρίζεται σε επιμέρους ζώνες. Αυτές είναι τα ραδιοκύματα, τα μικροκύματα, η υπέρυθη ακτινοβολία, η ορατή ακτινοβολία (φως), η υπεριώδης ακτινοβολία, οι ακτίνες Χ και οι ακτίνες γ.



Εικόνα 5 Το ηλεκτρομαγνητικό φάσμα

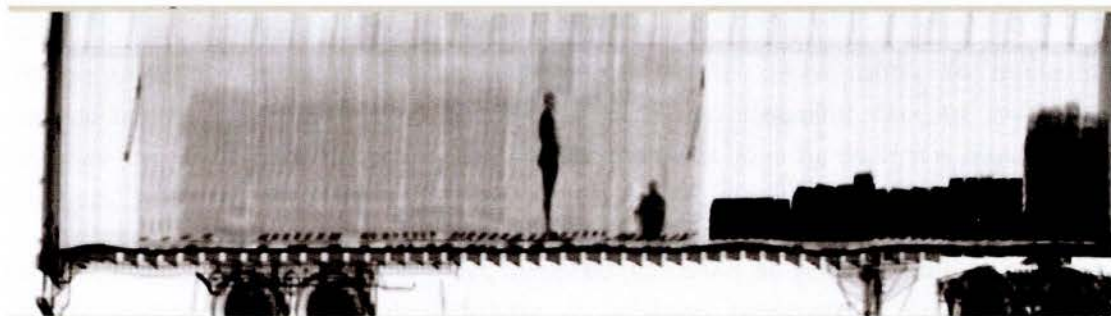
1.3.1 Απεικόνιση Ακτινών γ

Σημαντικές εφαρμογές της ψηφιακής επεξεργασίας εικόνας βασίζονται στις ακτίνες γ, συμπεριλαμβανομένης της πυρηνικής ιατρικής ή ακόμα και την επίβλεψη τελωνείων, στρατιωτικών βάσεων και άλλων κρίσιμων εγκαταστάσεων. Στην πυρηνική ιατρική ο σκοπός είναι να διοχετευτεί στον ασθενή ένα ραδιενεργό ισότοπο που απορροφάται σε μεγαλύτερες ποσότητες απ' ορισμένους ιστούς, απ' ότι από άλλους, καθιστώντας δυνατή την εξέταση συγκεκριμένων οργάνων. Για την ανίχνευση της εκπεμπόμενης ακτινοβολίας χρησιμοποιούνται ανιχνευτές ακτινών γ. Στην εικόνα Εικόνα (6) παρουσιάζεται η πλήρης σάρωση των οστών (σπινθηρογράφημα), ενός ασθενούς που ελήφθη με την χρήση ακτινών γ. Οι εικόνες αυτού του είδους χρησιμοποιούνται για να εντοπιστούν σημεία οστικής παθολογίας όπως λοιμώξεις ή όγκοι.



Εικόνα 6 Σπινθηρογράφημα - Μετάσταση Καρκίνου σε Σπόνδυλο

Στην παρακάτω εικόνα φαίνεται η ακτινογραφία ενός οχήματος που έγινε με χρήση ακτινών γ. Η λήψη της εικόνας έχει γίνει από ένα συστήματα VACIS (Vehicle and Cargo Inspection Systems). Τέτοια συστήματα χρησιμοποιούνται από τελωνειακές αρχές για τον έλεγχο φορτηγών, επιβατικών οχημάτων, κοντέινερ ή εμπορευμάτων με σκοπό τον εντοπισμό λαθραίων φορτίων ή επικίνδυνων υλικών.

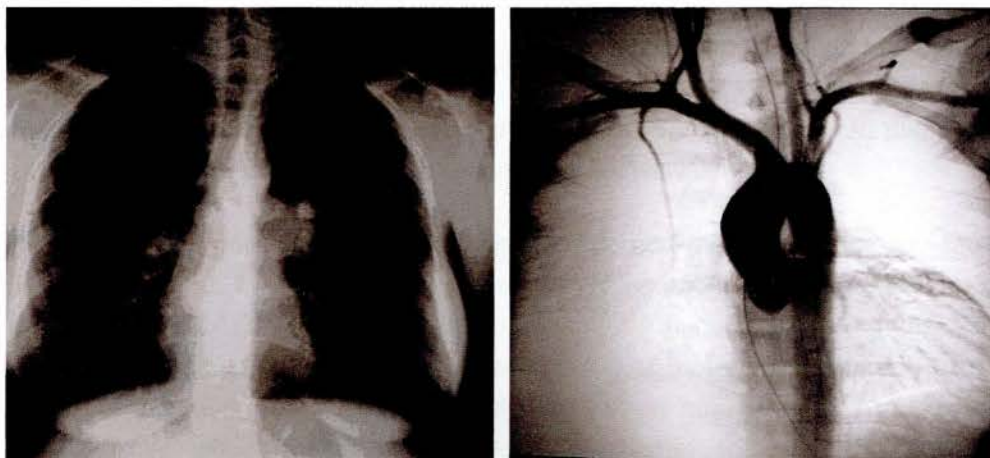


Εικόνα 7 Επιθεώρηση οχήματος με σύστημα VACIS

1.3.2 Απεικόνιση Ακτινών Χ

Οι ακτίνες Χ είναι από τις παλαιότερες πηγές ακτινοβολίας ηλεκτρομαγνητικής ενέργειας που χρησιμοποιούνται για την λήψη εικόπων. Οι πιο γνωστή χρήση των ακτινών Χ είναι στην ιατρική διάγνωση, αλλά επίσης χρησιμοποιούνται ευρέως στην βιομηχανία και σε άλλους τομείς όπως η αστρονομία.

Η Εικόνα 8(α) δείχνει μια ακτινογραφία θώρακος που παράγεται με την τοποθέτηση του ασθενούς μεταξύ μιας πηγής ακτινών Χ και ενός φιλμ ευαίσθητου στην ενέργεια των ακτινών Χ. Η αρχή της λήψης ακτινογραφιών βασίζεται στο γεγονός ότι οι ακτίνες Χ διαπερνούν σε ποικίλο βαθμό τα μαλακά μέρια του σώματος και τον αέρα, ενώ αντίθετα η πορεία τους διακόπτεται στα συμπαγή μέρια. Η εικόνα που λαμβάνεται στο φιλμ στη συνέχεια αιχμαλωτίζεται από ένα φωτοευαίσθητο σύστημα ψηφιοποίησης, για περαιτέρω επεξεργασία.



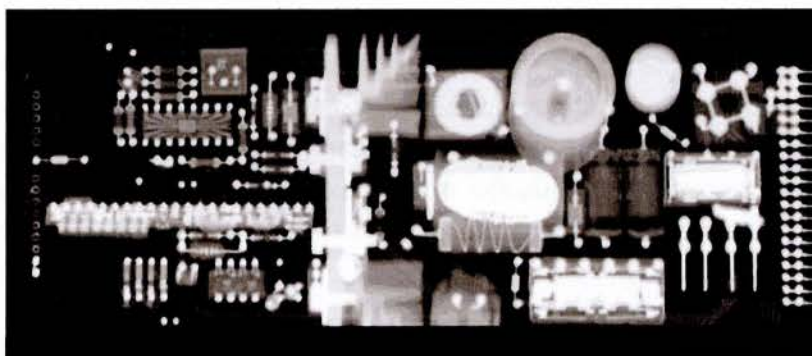
α β

Εικόνα 8 (α) Ακτινογραφία θώρακος. (β) Αγγειογραφία αορτής

Η Εικόνα 8(β) παρουσιάζει μια αγγειογραφία αορτής, που έχει ληφθεί με τη μέθοδο της ψηφιακής αφαιρετικής αγγειογραφίας. Η αγγειογραφία είναι μια ακτινολογική μέθοδος με την οποία απεικονίζονται τα αιμοφόρα αγγεία του ανθρωπίνου σώματος αφού εγχυθεί στο εσωτερικό τους ένα σκιαγραφικό μέσο. Στην ακτινογραφία που λαμβάνεται απεικονίζονται πολύ καλά τα αγγεία που έχουν σκιαγραφηθεί επειδή το σκιαγραφικό μέσο απορροφά την προσπίπτουσα ακτινοβολία και έτσι στο φιλμ αποτυπώνεται η σκιά του αγγείου, αφού αυτό περιέχει το σκιαγραφικό. Με τη μέθοδο αυτή μπορούν να σκιαγραφηθούν όλα τα αγγεία.

Σήμερα υπάρχουν πολύ εξελιγμένα μηχανήματα για το σκοπό αυτό, τα οποία με ηλεκτρονική επεξεργασία αφαιρούν από την τελική ακτινογραφία όλες τις σκιές των οργάνων που δεν ενδιαφέρουν τον εξεταστή και αφήνουν μόνο τις σκιές των αγγείων του ασθενούς.

Τεχνικές παρόμοιες με αυτές που μόλις αναφέρθηκαν αλλά σε γενικές γραμμές αφορούν υψηλής ενέργειας ακτίνες Χ εφαρμόζονται σε διάφορες βιομηχανικές διεργασίες. Στην εικόνα 9 φαίνεται μια εικόνα ακτίνων Χ ενός ηλεκτρονικού κυκλώματος. Τέτοιες εικόνες αντιπροσωπεύουν κυριολεκτικά εκατοντάδες εφαρμογές των ακτίνων Χ που χρησιμοποιούνται για την εξέταση κυκλωμάτων όπως για εξαρτήματα που λείπουν ή για κομμένους διαδρόμους πάνω στις πλακέτες.



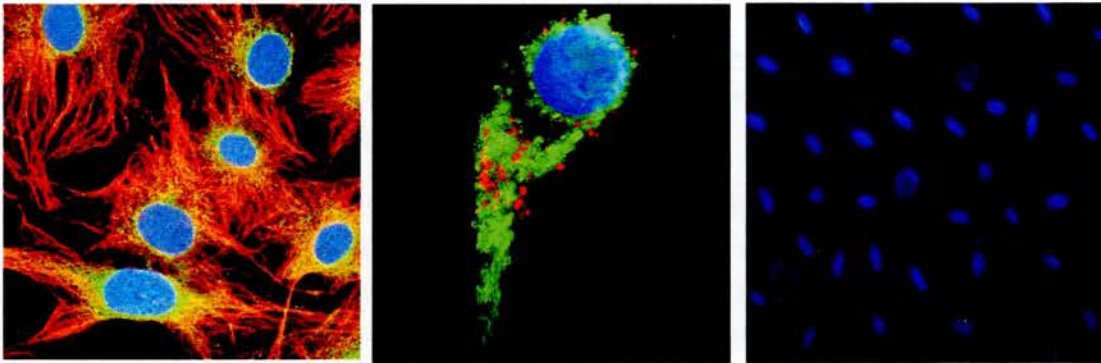
Εικόνα 9 Ακτινογραφία ηλεκτρονικού κυκλώματος

1.3.3 Απεικόνιση στην υπεριώδη ζώνη

Η χρήση της υπεριώδους ακτινοβολίας βρίσκει πρακτική εφαρμογή σε ποικίλους τομείς όπως στην ιατρική, στην εγκληματολογία, στην αρχαιολογία αλλά και στην αστρονομία. Παρακάτω θα παρουσιαστούν απεικονίσεις, στην υπεριώδη ζώνη, με χρήση μικροσκοπίου αλλά και ψηφιακής κάμερας για τη λήψη υπεριωδών εικόνων.

Με την πάροδο του χρόνου εμφανίστηκαν πολλοί προηγμένοι «συγγενείς» του κλασικού οπτικού μικροσκοπίου. Ένας εξ αυτών είναι το οπτικό μικροσκόπιο φθορισμού, το οποίο χρησιμοποιεί υπεριώδη ακτινοβολία. Είναι γνωστό πως κάποιες χημικές ουσίες απορροφούν την υπεριώδη ακτινοβολία και την επανεκπέμπουν στιγμιαία ως ορατό φως. Το φαινόμενο αυτό ονομάζεται φθορισμός. Αν στο δείγμα που θέλουμε να μελετήσουμε υπάρχουν τέτοιες φθορίζουσες ουσίες, τότε μελετώντας το φως που εκπέμπεται, παίρνουμε χρήσιμες πληροφορίες για το δείγμα. Υπάρχουν πολλές τέτοιες χημικές ουσίες στα κύτταρα που έχουν ως χαρακτηριστικό τους τον αυτοφθορισμό. Ωστόσο αν δεν

υπάρχουν μπορούν κατά βούληση να χρησιμοποιηθούν φθορίζουσες ουσίες, όπως η φλουορεσκεΐνη, για να αποκαλύψουν τα σημεία που θέλουμε να μελετήσουμε σε διάφορα κύτταρα.

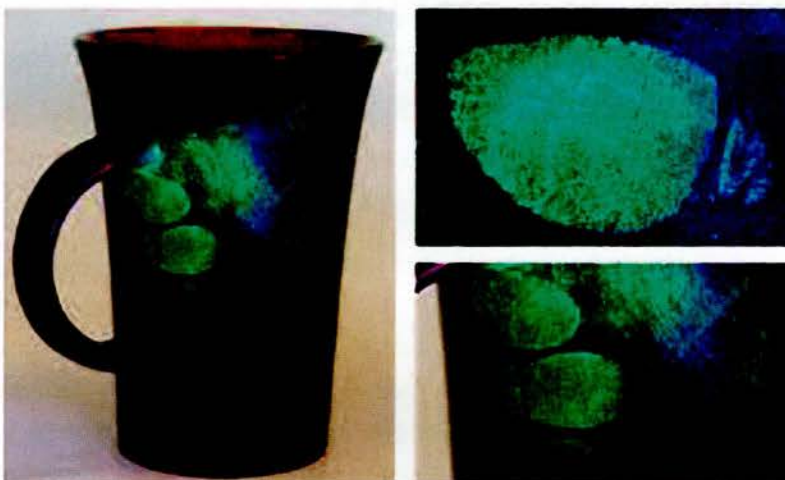


α β γ

Εικόνα 10 Στιγμιότυπα από μικροσκόπιο φθορισμού. (α) Ομάδα ινοβλαστων κύτταρων. (β) Ανθρώπινο κύτταρο. (γ) Επιφάνεια φύλλου ελιάς

Οι Εικόνες 10(α) και (β) δείχνουν τα αποτελέσματα της τυπικής ικανότητας του μικροσκοπίου φθορισμού. Η εικόνα 10(α) δείχνει μια εικόνα από μικροσκόπιο φθορισμού μιας ομάδας ινοβλαστών, κύτταρων απαραίτητων για την ανανέωση των ιστών. Οι πυρήνες των κυττάρων αποτυπώνονται με μπλε χρώμα ενώ οι νηματοειδείς πρωτεΐνες που σχηματίζουν τον κυτταροσκελετό με πορτοκαλί. Η Εικόνα 10(β) δείχνει το αποτύπωμα ενός ανθρώπινου κύτταρου. Με μπλε χρώμα φαίνεται ο πυρήνας του και με πράσινο το κυτταρόπλασμα. Στην Εικόνα 10(γ) παρατηρούμε την επιφάνεια ενός αποτριχωμένου φύλλου ελιάς όπως φαίνεται στο μικροσκόπιο φθορισμού. Ο έντονος μπλε φθορισμός οφείλεται στην ύπαρξη κυρίως φερουλικού οξέος στους επιεφυμενιδικούς κηρούς των καταφρακτικών.

Επίσης το υπεριώδες φως έχει ιδιότητες που το καθιστούν ένα πολύ ισχυρό ερευνητικό εργαλείο για την εγκληματολογία, καθώς έχει την ικανότητα να κάνει πολλές ουσίες να φθορίζουν. Το υπεριώδες φως επιτρέπει σε ερευνητές και εγκληματολόγους να εξετάσουν τις ενδείξεις και να ανακτήσουν στοιχεία που θα μπορούσαν να μην έχουν εντοπιστεί στο παρελθόν.



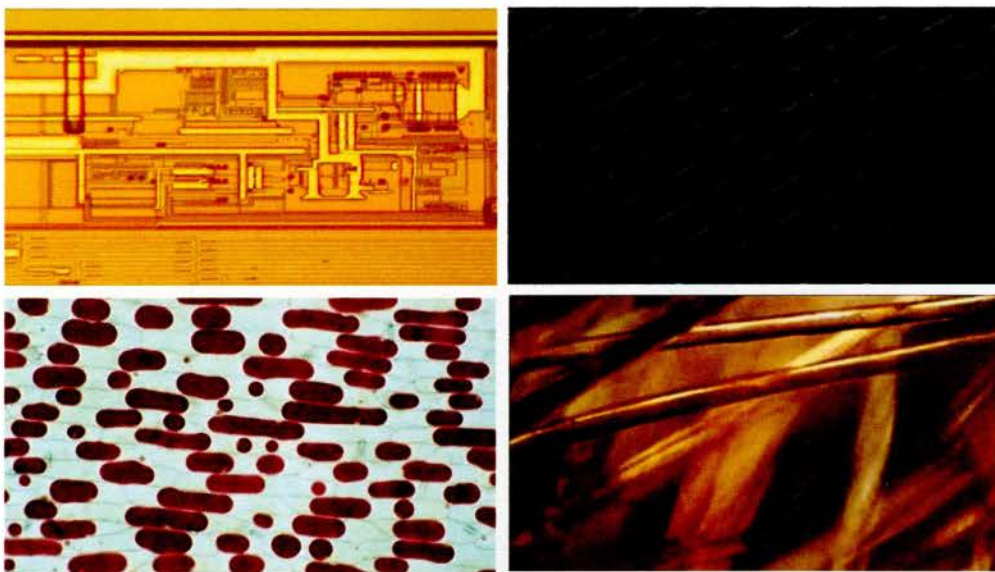
Εικόνα 11 Απεικόνιση δακτυλικού αποτυπώματος με υπεριώδες φως

1.3.4 Απεικόνιση στις ορατές και υπέρυθρες ζώνες

Το φως που βλέπουμε με τα μάτια μας είναι πραγματικά ένα πολύ μικρό τμήμα του ηλεκτρομαγνητικού φάσματος, το οποίο καλύπτει την ορατή ζώνη. Λαμβάνοντας υπόψη ότι με την ορατή ζώνη του ηλεκτρομαγνητικού φάσματος είμαστε πιο εξοικειωμένοι αφού τη χρησιμοποιούμε σ' όλες τις καθημερινές μας δραστηριότητες δεν αποτελεί έκπληξη ότι η χρήση εφαρμογών ψηφιακής επεξεργασίας εικόνας, στην εν λόγω ζώνη, είναι κατά πολύ μεγαλύτερη από όλες τις άλλες ζώνες μαζί.

Οι εικόνες που έχουν ληφθεί στην υπέρυθρη ζώνη συχνά χρησιμοποιούνται σε συνδυασμό με εικόνες της ορατής ζώνης. Για το λόγο αυτό οι δυο αυτές ζώνες του ηλεκτρομαγνητικού φάσματος θα παρουσιαστούν από κοινού στη παρούσα ενότητα. Παρακάτω θα γίνει αναφορά σε τομείς που γίνεται χρήση των εικόνων των δυο αυτών ζωνών όπως στην οπτική μικροσκοπία, στην αστρονομία, στην τηλεπισκόπηση, στην βιομηχανία και στην επιβολή του νόμου.

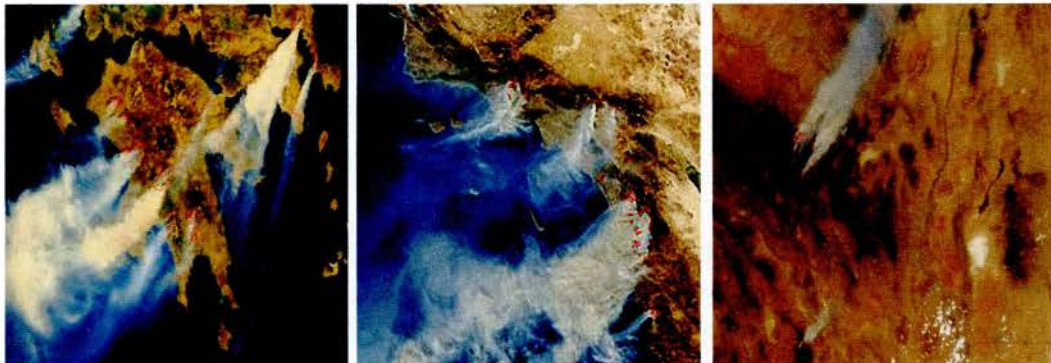
Παρακάτω φαίνονται αρκετά παραδείγματα εικόνων που έχουν ληφθεί με ένα οπτικό μικροσκόπιο. Τα παραδείγματα προέρχονται από ιατρικές έρευνες και μικροεπιθεωρήσεις για το χαρακτηρισμό υλικών στη βιομηχανία. Ακόμα και με ένα απλό μικροσκόπιο το εύρος των εφαρμογών που μπορούμε να καλύψουμε είναι υπερβολικά μεγάλο για να μπορέσουμε να αναφερθούμε με λεπτομέρειες εδώ.



Εικόνα 12 Στιγμιότυπα από οπτικό μικροσκόπιο. (α) Επιφάνεια επεξεργαστή Pentium. (β) Επιφάνεια CD. (γ) Κύτταρα από αίμα βατράχου. (δ) Ξανθιές τρίχες από τετράχρονο αγόρι.

Ένας άλλος σημαντικός τομέας της ψηφιακής επεξεργασίας εικόνας είναι η τηλεπισκόπηση. Τέτοιες εικόνες λαμβάνονται από το διάστημα με χρήση δορυφόρων, όπως ο LANDSAT της NASA. Η πρωταρχική λειτουργία του LANDSAT είναι να αποκτήσει και να μεταδώσει εικόνες στην γη από το διάστημα για σκοπούς παρακολούθησης των περιβαλλοντικών συνθηκών στον πλανήτη. Οι εικόνες αρχειοθετούνται σε σταθμούς ανά τον κόσμο και αποτελούν μια μοναδική πηγή πληροφοριών για την παγκόσμια έρευνα και τις εφαρμογές

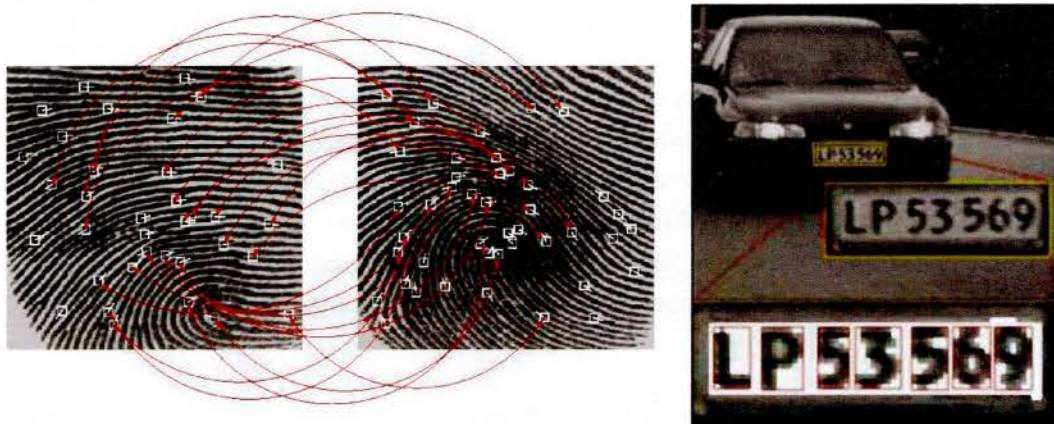
στη γεωργία, στη χαρτογραφία, τη γεωλογία, τη δασοκομία, τη χωροταξία, την επιτήρηση, την εκπαίδευση αλλά και σε άλλους τομείς.



Εικόνα 13 Εντοπισμός πυρκαγιών από δορυφόρο

Στις παραπάνω εικόνες διακρίνεται περιοχές, σε διάφορα μέρη του κόσμου, στις οποίες έχουν εκδηλωθεί εστίες πυρκαγιών. Οι εστίες αυτές έχουν εντοπιστεί με την αποστολή δεδομένων από το δορυφόρο σε επίγειους σταθμούς, όπου με τη χρήση ανεπτυγμένων αλγορίθμων επεξεργασίας εικόνων είναι εφικτός ο εντοπισμός των πυρκαγιών. Οι εικόνες που αποστέλλονται στους επίγειους σταθμούς λαμβάνονται τόσο στο ορατό όσο και στο υπέρυθρο μέρος του ηλεκτρομαγνητικού φάσματος. Οι εστίες των πυρκαγιών είναι εύκολα ορατές και στις τρεις εικόνες, αφού έχουν χρωματιστεί με έντονο κόκκινο χρώμα.

Τα παραδείγματα της ψηφιακής επεξεργασίας εικόνας στο οπτικό φάσμα πραγματικά είναι αμέτρητα. Η εικόνα 14(α) δείχνει το δακτυλικό αποτύπωμα ενός αντίχειρα. Εικόνες από τα δακτυλικά αποτυπώματα συνήθως επεξεργάζονται από υπολογιστή για να βελτιωθεί η ποιότητα τους, με σκοπό να βρεθούν ομοιότητες με αντίστοιχα δείγματα σε βάσεις δεδομένων που πιθανόν να ταιριάζουν.



Εικόνα 14 Παραδείγματα της ψηφιακής επεξεργασίας εικόνας στο οπτικό φάσμα. (α) Ταίριασμα δακτυλικών αποτυπωμάτων. (β) Αυτόματη αναγνώριση αριθμού κυκλοφορίας.

Τέλος η εικόνα του οχήματος που διακρίνεται παραπάνω αποτελεί παράδειγμα της αυτόματης ανάγνωσης αριθμών κυκλοφορίας. Οι πινακίδες κυκλοφορίας και άλλες εφαρμογές για την αναγνώριση χαρακτήρων χρησιμοποιούνται ευρέως για την παρακολούθηση της κυκλοφορίας και για επιτήρηση. Η εφαρμογή αυτή της ψηφιακής επεξεργασίας εικόνας είναι που θα μας απασχολήσει σε επόμενα κεφάλαια.

Κεφάλαιο 2

Χρωματικά Μοντέλα

2.1 Εισαγωγικά

Ο σκοπός των χρωματικών μοντέλων (ονομάζονται επίσης χρωματικοί χώροι ή συστήματα χρωμάτων) είναι να διευκολύνουν τον προσδιορισμό των χρωμάτων σε ορισμένα πρότυπα, σύμφωνα με ένα γενικό και αποδεκτό τρόπο. Στην ουσία ένα τέτοιο σύστημα είναι μια προδιαγραφή ενός τρισδιάστατου συστήματος συντεταγμένων, όπου κάθε χρώμα αντιπροσωπεύεται από ένα μόνο σημείο.

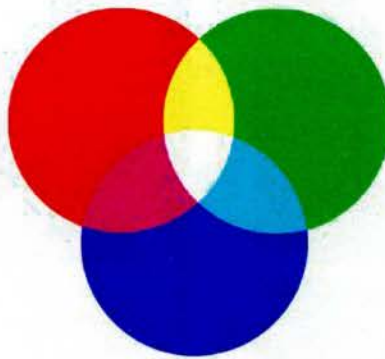
Για την περιγραφή των χρωμάτων χρησιμοποιούμε τα χρωματικά μοντέλα. Κάθε ένα από αυτά χρησιμοποιεί κάποια βασικά χρώματα και μία μέθοδο για την περιγραφή των χρωματικών αντιλήψεων που μπορεί να έχουμε. Τα χρωματικά μοντέλα αναλύουν/συνθέτουν την εικόνα σε αρκετά μικρά ομοιόμορφα στοιχεία, τα εικονοστοιχεία (pixels), με βάση κάποιες μεθόδους. Συνήθως αποτελούνται από ένα χρώμα ή τον συνδυασμό κάποιων λίγων χρωμάτων, των βασικών του κάθε μοντέλου. Έτσι προσεγγίζουν αρκετά την εικόνα που βλέπουμε, με διαφορετική ακρίβεια το καθένα. Τα χρώματα αυτά και οι μέθοδοι διαφέρουν από μοντέλο σε μοντέλο.

Τα περισσότερα μοντέλα χρωμάτων που χρησιμοποιούνται σήμερα είτε σχετίζονται με το υλικό (hardware), όπως στις έγχρωμες οθόνες και τους εκτυπωτές, είτε με εφαρμογές (software) όπου ο σωστός χειρισμός του χρώματος είναι ο τελικός στόχος, όπως στη δημιουργία έγχρωμων γραφικών ή ακόμα και κινούμενων σχεδίων. Όσον αφορά την ψηφιακή επεξεργασία εικόνας, το μοντέλο που είναι περισσότερο προσανατολισμένο προς το υλικό και χρησιμοποιείται πιο συχνά στην πράξη είναι το μοντέλο RGB, το οποίο χρησιμοποιείται από τις έγχρωμες οθόνες και μια ευρεία γκάμα έγχρωμων βιντεοκαμερών. Τα μοντέλα CMY και CMYK χρησιμοποιούνται στις έγχρωμες εκτυπώσεις. Ενώ το μοντέλο HIS, το οποίο ανταποκρίνεται σε μεγάλο βαθμό με τον τρόπο που οι άνθρωποι περιγράφουν και να ερμηνεύουν το χρώμα. Το μοντέλο HIS έχει επίσης το πλεονέκτημα ότι διαχωρίζει τις έγχρωμες πληροφορίες μιας εικόνας από τις γκρι, καθιστώντας το κατάλληλο για πολλές τεχνικές που χρησιμοποιούνται στις εικόνες με αποχρώσεις του γκρι και αναπτύσσονται στην παρούσα πτυχιακή εργασία.

Σήμερα υπάρχουν ποικίλα μοντέλα χρωμάτων που χρησιμοποιούνται εξαιτίας του γεγονότος ότι η επιστήμη που ασχολείται με τα χρώματα αποτελεί από μόνη της ένα τεράστιο επιστημονικό πεδίο που περιλαμβάνει πολλούς τομείς εφαρμογών. Τα μοντέλα που περιγράφονται παρακάτω αποτελούν τα κορυφαία μοντέλα που χρησιμοποιούνται στην επεξεργασία εικόνας.

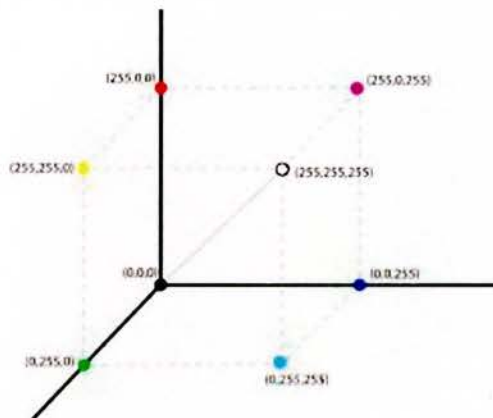
2.2 Μοντέλο RGB

Το πρότυπο RGB είναι ένα μοντέλο χρωμάτων στο οποίο τα χρώματα κόκκινο, πράσινο και μπλε προστίθενται μαζί, με διάφορους τρόπους, ώστε να αναπαράγουν ένα όσο πιο ευρύ φάσμα χρωμάτων γίνεται. Ειδικά συνδυάζοντάς τα ανά δύο παίρνουμε τα τρία δευτερεύοντα χρώματα, κίτρινο (Y) (Κόκκινο και Πράσινο), κυανό (C) (Πράσινο και Μπλε) και Ιώδες (M) (Μπλε και Κόκκινο) όπως φαίνεται και στην παρακάτω εικόνα. Έτσι κάθε χρώμα καθορίζεται από 3 συντεταγμένες που προσδιορίζουν τη συμμετοχή καθενός από τα βασικά χρώματα. Ίσες αναλογίες αντιστοιχούν σε τόνους του γκρι, μεταξύ λευκού και μαύρου. Το όνομα του μοντέλου προέρχεται από τα αρχικά των τριών βασικών χρωμάτων: κόκκινο (**R** – Red), πράσινο (**G** – Green) και μπλε (**B** - Blue).



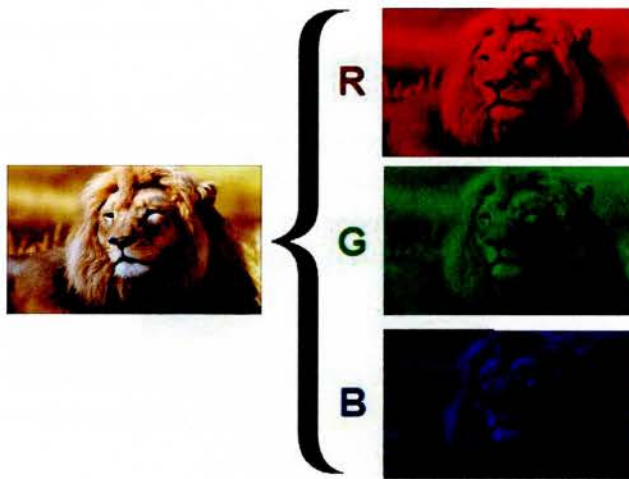
Εικόνα 15 Χρωματικό μοντέλο RGB

Το χρωματικό μοντέλο RGB μπορεί να αναπαρασταθεί με τη χρήση ενός κύβου χρωμάτων σε ένα τρισδιάστατο σύστημα συντεταγμένων. Στην αρχή των αξόνων είναι η κορυφή του κύβου που αντιστοιχεί στο μαύρο χρώμα. Στις κορυφές του κύβου που βρίσκονται πάνω στους άξονες του συστήματος συντεταγμένων βρίσκονται τα βασικά χρώματα (κόκκινο, πράσινο, μπλε). Τα δευτερογενή χρώματα βρίσκονται στις τρεις κορυφές του κύβου που βρίσκονται απέναντι από τα αντίστοιχα βασικά χρώματα, ενώ στην κορυφή διαγώνια απέναντι από το μαύρο βρίσκεται το λευκό. Στη διαγώνιο μεταξύ μαύρου και λευκού βρίσκονται όλες οι αποχρώσεις του γκρι. Η διαγώνιος αυτή ονομάζεται και γραμμή έντασης.



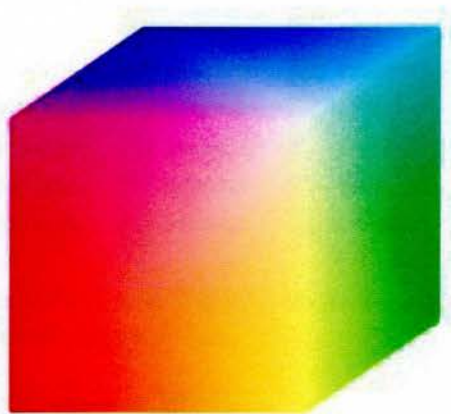
Εικόνα 16 Κύβος RGB

Κάθε χρώμα στο σύστημα αυτό προσδιορίζεται από ένα σημείο στον κύβο με τρεις συντεταγμένες, την τιμή για το κόκκινο (R), το πράσινο (G) και το μπλε (B) χρώμα. Έτσι οι εικόνες που αναπαριστώνται στο μοντέλο χρωμάτων RGB αποτελούνται ουσιαστικά από τρία ξεχωριστά κανάλια ή τρεις ξεχωριστές εικόνες, μια για κάθε βασικό χρώμα. Με αυτό τον τρόπο όταν τροφοδοτείται μια οθόνη με μια εικόνα RGB, αυτό που βλέπει ο παρατηρητής δεν είναι τίποτα άλλο από τον συνδυασμό των τριών καναλιών, τα οποία παράγουν μια έγχρωμη σύνθετη εικόνα. Θα πρέπει να τονιστεί πως το μοντέλο RGB εξαρτάται από τη συσκευή απεικόνισης, είναι δηλαδή "device dependent". Ίδιες συντεταγμένες θα δώσουν διαφορετικά χρώματα σε διαφορετικές συσκευές ή με διαφορετικές ρυθμίσεις της ίδιας συσκευής, πράγμα που δεν συμβαίνει με άλλα μοντέλα.



Εικόνα 17 Οι 3 συνιστώσες RGB

Ο αριθμός των bits που χρησιμοποιείται για την αναπαράσταση κάθε pixel στο μοντέλο RGB ονομάζεται βάθος χρώματος. Έτσι είναι κατανοητό πως μια RGB εικόνα αποτελείται από μια κόκκινη, μια πράσινη και μια μπλε εικόνα των 8-bits. Υπό αυτές τις συνθήκες κάθε pixel χρώματος μιας πλήρους έγχρωμης εικόνας RGB έχει βάθος χρώματος 24-bits. Ο συνολικός αριθμός των διαφορετικών αποχρώσεων μιας 24-bits εικόνας είναι $(28)^3 = 16.777.216$, τα οποία είναι αρκετά χρώματα για κάθε εικόνα.

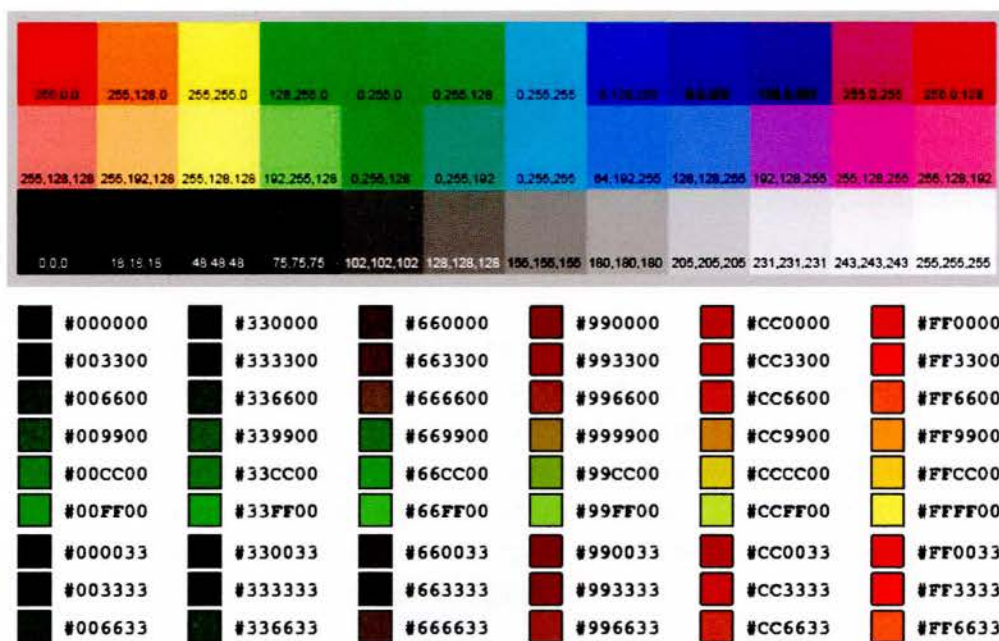


Εικόνα 18 Κύβος RGB 24bits

Ένα χρώμα στο μοντέλο χρώματος RGB μπορεί να περιγραφεί με το να προσδιορίσουμε την ποσότητα του κόκκινου, του πράσινο και του μπλε χρώματος που περιλαμβάνει. Έτσι η απόχρωση του επιλεγόμενου χρώματος καθορίζεται από την τριπλή τιμή (R, G, B), όπου κάθε συστατικό στοιχείο μπορεί να κυμαίνεται από μηδέν έως ένα συγκεκριμένο ανώτατο όριο τιμής. Εάν όλα τα συστατικά είναι στο μηδέν το αποτέλεσμα είναι μαύρο, ενώ αν όλα είναι στο μέγιστο, το αποτέλεσμα είναι το πιο απόλυτο λευκό. Τα χρώματα μπορούν να περιγραφούν ποσοτικά με διάφορους τρόπους.

- Οι τιμές των χρωμάτων μπορούν να τοποθετηθούν στη κλίμακα 0 (ελάχιστο) ως 1 (μέγιστο). Έτσι το κόκκινο είναι (1.0, 0.0, 0.0), ενώ το μπλε (0.0, 0.0, 1.0).
- Άλλος τρόπος είναι να γραφούν με ποσοστά, από 0% (ελάχιστο) ως 100% (μέγιστο). Τότε το κόκκινο είναι 100%, 0%, 0%.
- Στην πληροφορική, οι τιμές αποθηκεύονται ως ακέραιοι αριθμοί από 0 έως 255, δηλαδή το εύρος ενός byte (1 byte περιέχει 8 bits) όπου μπορεί να προσφέρει 256 διαφορετικές τιμές. Αυτές μπορούν να αναπαρασταθούν είτε ως δεκαδικό, είτε ως δεκαεξαδικό αριθμοί. Όμως με τη χρήση σύγχρονου ψηφιακού εξοπλισμού το εύρος ανεβαίνει από 0 ως 65536, αφού για την αναπαραγωγή ενός χρώματος χρησιμοποιούνται 16 bits.

Ακολουθούν δυο χρωματικοί πίνακες για να γίνει πιο κατανοητός ο τρόπος περιγραφής των χρωμάτων στο χρωματικό μοντέλο RGB.

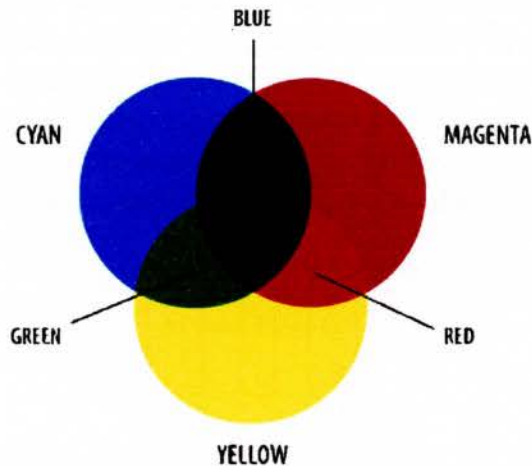


Εικόνα 19 Χρωματικοί κώδικες μοντέλου RGB. (α) Αναπαράσταση στο δεκαδικό σύστημα. (β) Αναπαράσταση στο δεκαεξαδικό σύστημα.

2.3 Μοντέλο CMY(K)

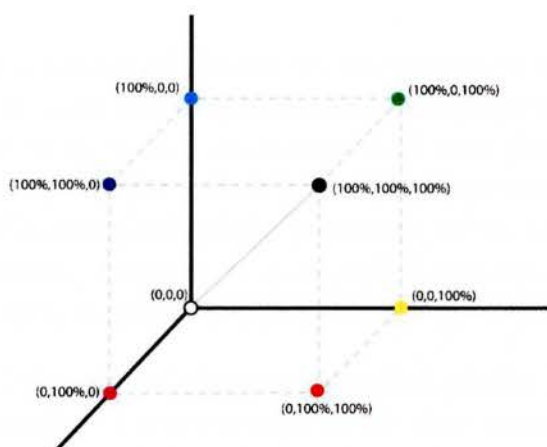
Στο χρωματικό μοντέλο RGB που εξετάστηκε πιο πάνω, τα βασικά χρώματα λειτουργούν *προσθετικά*. Ξεκινώντας δηλαδή με βάση το μαύρο και προσθέτοντάς τα σε διάφορες αναλογίες παίρνουμε τα υπόλοιπα χρώματα ως και το λευκό. Γι' αυτό τα βασικά χρώματα χρησιμοποιούνται για τη σύνθεση των άλλων χρωμάτων όταν η βάση είναι το μαύρο, και η προσθήκη χρώματος αυξάνει τη φωτεινότητα, όπως για παράδειγμα στις οθόνες καθοδικού σωλήνα.

Για τις εκτυπώσεις όμως, ξεκινάμε με τη φωτεινότητα στο μέγιστο, το λευκό, και κάθε προσθήκη χρώματος τη μειώνει. Επιδιώκουμε να εμφανίσουμε σε ένα σημείο της σελίδας ένα χρώμα, τοποθετώντας μελάνια που θα απορροφούν όλα τα άλλα χρώματα του ορατού φάσματος και θα αντανακλούν μόνο αυτό. Η προσθήκη χρώματος λειτουργεί δηλαδή *αφαιρετικά*. Τα πιο πρόσφορα χρώματα για να χρησιμοποιηθούν σαν βασικά στο μοντέλο αυτό είναι τα δευτερεύοντα χρώματα του RGB, δηλαδή το κυανό (**C** – Cyan), το ιώδες (**M** – Magenta) και το κίτρινο (**Y** – Yellow).



Εικόνα 20 Χρωματικό μοντέλο CMY(K)

Στο πιο πάνω διάγραμμα φαίνεται πως ανά δυο τα βασικά χρώματα του μοντέλου CMY μας δίνουν (λιγότερο φωτεινά) ως δευτερεύοντα τα βασικά χρώματα του R,G,B. Με τα παραπάνω αντιλαμβανόμαστε πως ο χρωματικός χώρος CMY έχει ίδια μορφή με τον RGB. Είναι ένας κύβος όπου οι συντεταγμένες κάθε σημείου-χρώματος μας λένε τι αναλογία C, M και Y απαιτείται για να πάρουμε στο χαρτί το επιθυμητό χρώμα. Όπως και με το μοντέλο χρωμάτων RGB κάθε χρώμα εκπροσωπείται από τρεις τιμές C, M και Y με τιμές της τάξης του $[0,1]$ ή στην 8bit εκδοχή του, της τάξης του $[0,255]$.



Εικόνα 21 Κύβος CMY(K)

Οι περισσότερες συσκευές που ψεκάζουν χρωστικές ουσίες σε χαρτί, όπως οι έγχρωμοι εκτυπωτές και τα φωτοτυπικά μηχανήματα, απαιτούν η εισαγωγή δεδομένων να γίνεται με βάση το χρωματικό πρότυπο CMY ή να εκτελείται μετατροπή από RGB σε CMY σε εσωτερικό επίπεδο. Η μετατροπή γίνεται με τη χρήση της παρακάτω εξίσωσης, η οποία μας δείχνει ότι όλες οι χρωματικές τιμές είναι μεταξύ του διαστήματος [0, 1]. Η εξίσωση (6,2 - 1) αποδεικνύει ότι το φως που αντανακλάται από μια επιφάνεια επικαλυμμένη με καθαρό κυανό δεν περιέχει καθόλου κόκκινο (δηλαδή, $C = 1 - R$ στην εξίσωση).

$$\begin{bmatrix} C \\ M \\ Y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

Ομοίως, το καθαρό ματζέντα δεν αντανακλά το πράσινο και το καθαρό κίτρινο δεν αντανακλά το μπλε. Η εξίσωση (6,2 - 1) αποκαλύπτει επίσης ότι οι RGB τιμές μπορούν να βρεθούν εύκολα από ένα σύνολο τιμών CMY αφαιρώντας τις επιμέρους τιμές CMY από 1, δηλαδή το απόλυτο λευκό.

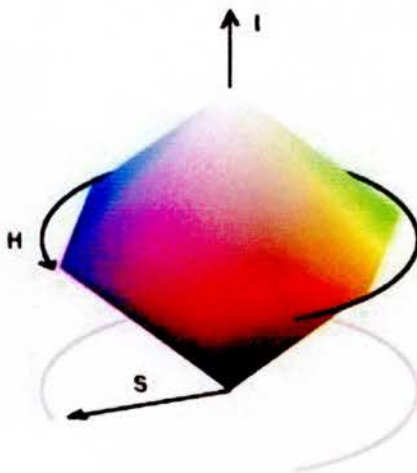
$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} C \\ M \\ Y \end{bmatrix}$$

Τα μελάνια όμως από τη φύση τους δεν μπορούν να αποδώσουν συγκεκριμένα μήκη κύματος – χρώματα (όπως τα pixels μίας οθόνης) αλλά μία ευρεία περιοχή του χρωματικού φάσματος. Το αποτέλεσμα είναι συνδυασμός των τριών βασικών χρωμάτων να δίνει ένα καφετί χρώμα αντί για το μαύρο. Για το λόγο αυτό προστέθηκε στο μοντέλο CMY και το μαύρο μελάνι με αποτέλεσμα να προκύψει το χρωματικό μοντέλο CMYK (Cyan – Magenta – Yellow – Black).

2.4 Μοντέλο HIS

Στα δυο προηγούμενα κεφάλαια είδαμε πως δημιουργούνται τα χρώματα στα χρωματικά μοντέλα RGB και CMY, αλλά και πως μεταφερόμαστε εύκολα από το ένα σύστημα στο άλλο. Όπως προαναφέρθηκε, τα χρωματικά συστήματα αυτά είναι ιδανικά για την παραγωγή έγχρωμης εικόνας (όπως στην λήψη εικόνας από μια έγχρωμη κάμερα ή για την εμφάνιση μιας εικόνας σε μια οθόνη ή σε χαρτί), αλλά η χρήση τους για την περιγραφή των χρωμάτων είναι πολύ πιο περιορισμένη. Επιπλέον, το σύστημα RGB ταιριάζει απόλυτα με το γεγονός ότι το ανθρώπινο μάτι είναι ιδιαίτερα οξυδερκή στο κόκκινο, το πράσινο και το μπλε.

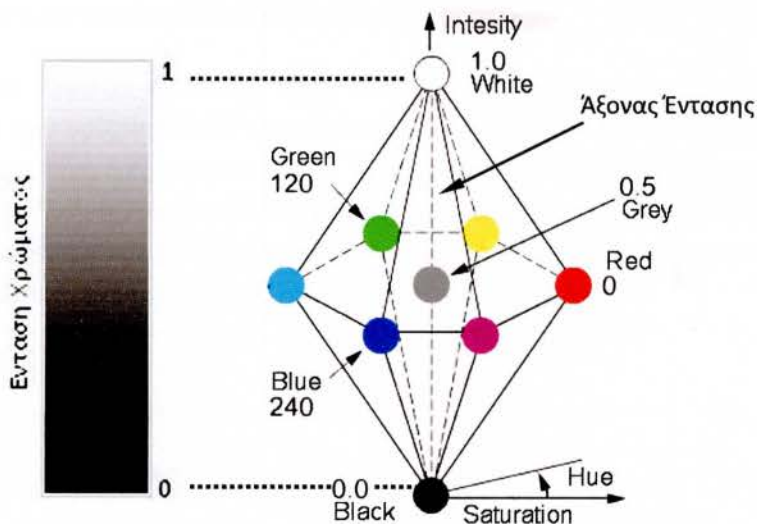
Δυστυχώς, το RGB, το CMY και άλλα παρόμοια χρωματικά μοντέλα δεν είναι κατάλληλα για την περιγραφή των χρωμάτων από την άποψη ότι δεν είναι πρακτικά για την ανθρώπινη ερμηνεία. Για παράδειγμα, ένας άνθρωπος δεν αναφέρεται στο χρώμα ενός αυτοκινήτου, δίνοντας το ποσοστό κάθε ενός από τα βασικά χρώματα που το συνθέτουν. Επιπλέον, κανένας μας δεν σκέπτεται ότι μια έγχρωμη εικόνα αποτελείται από τρεις βασικές εικόνες που συνδυάζονται για να σχηματίσουν μια πλήρως έγχρωμη εικόνα. Όταν οι άνθρωποι δουν ένα έγχρωμο αντικείμενο μπορούν να το περιγράψουν με βάση την απόχρωση, τον κορεσμό (καθαρότητα) και την ένταση (φωτεινότητα) του. Η απόχρωση είναι χρωματικό χαρακτηριστικό που περιγράφει ένα καθαρό χρώμα (καθαρό κίτρινο, πορτοκαλί ή κόκκινο), ενώ ο κορεσμός δίνει μια μέτρηση του βαθμού στον οποίο ένα καθαρό χρώμα αραιώνεται με λευκό φως. Η φωτεινότητα, που είναι ένας από τους παράγοντες κλειδί στην περιγραφή της αίσθησης του χρώματος, αποτελεί μια υποκειμενική περιγραφή η οποία πρακτικά είναι αδύνατο να μετρηθεί.



Εικόνα 22 Χρωματικό μοντέλο HIS

Γνωρίζουμε ότι η ένταση (επίπεδο του γκρι) είναι μια από της πιο χρήσιμες περιγραφές των μονοχρωματικών εικόνων. Αυτή η ποσότητα είναι σίγουρα μετρήσιμη και εύκολα ερμηνεύσιμη. Το χρωματικό μοντέλο που ακολουθεί, αποκαλείται HSI (ή HSL) (Απόχρωση (Hue), Κορεσμός (Saturation), Ένταση(Intensity)) και έχει την ιδιότητα να διαχωρίζει την ένταση από της υπόλοιπες πληροφορίες (απόχρωση και κορεσμό) μιας εικόνας. Ο χώρος χρώματος HSI είναι ένα πολύ ελκυστικό μοντέλο χρωμάτων για ανάπτυξη αλγορίθμων επεξεργασίας εικόνας με βάση τις περιγραφές του χρώματος που είναι έμφυτες και ενστικτώδεις για τους ανθρώπους.

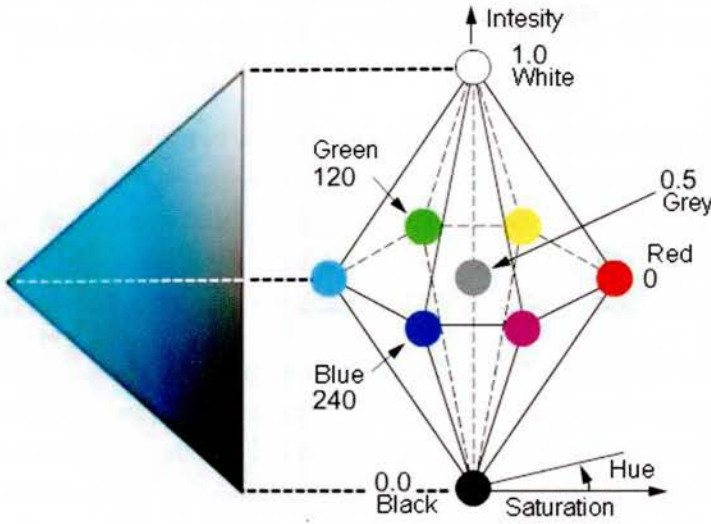
Όπως παρουσιάστηκε στην ενότητα 2.1, μια έγχρωμη εικόνα RGB μπορεί να θεωρηθεί ότι αποτελείται από τρεις μονοχρωματικές εικόνες (που αντιπροσωπεύουν το κόκκινο, το πράσινο και το μπλε) με διαφορετική ένταση η καθεμία. Έτσι είναι φανερό πως μπορούμε να αντλήσουμε την τιμή της έντασης οποιοδήποτε χρώματος. Για να γίνει αυτό θα πρέπει απλά να θεωρήσουμε τον άξονα έντασης που διακρίνεται με διακεκομμένες γραμμές στην Εικόνα 23. Στην κορυφή του άξονα βρίσκεται το απόλυτο λευκό ενώ στη βάση το απόλυτο μαύρο. Τα ενδιάμεσα σημεία περιέχουν τις αποχρώσεις του γκριζου. Η τομή του επιπέδου (ύψος) που ανήκει το χρώμα, την ένταση του οποίου αναζητούμε, με τον άξονα έντασης θα μας δώσει ένα σημείο μεταξύ του διαστήματος [0,1]. Το 0 χαρακτηρίζει τη βάση του άξονα, ενώ για την κορυφή έχουμε 1. Επίσης σημειώνουμε ότι ο κορεσμός (καθαρότητα) του χρώματος αυξάνεται ανάλογα της απόστασης από τον άξονα έντασης. Στην πραγματικότητα, ο κορεσμός των σημείων στον άξονα της έντασης είναι μηδέν, όπως αποδεικνύεται από το γεγονός ότι όλα τα σημεία κατά μήκος αυτού του άξονα είναι γκρι.



Εικόνα 23 Διαβαθμίσεις Άξονα Έντασης

Τέλος για να γίνει κατανοητό με πιο τρόπο και η απόχρωση μπορεί να καθοριστεί από ένα δεδομένο σημείο RGB, ας θεωρήσουμε τρία σημεία, το μαύρο και το λευκό, δηλαδή την κορυφή και τη βάση του άξονα έντασης. Για τρίτο σημείο επιλέγουμε το κυανό που βρίσκεται στις 180°. Αυτομάτως δημιουργείται ένα νοητό τρίγωνο με κορυφές τα τρία σημεία που επιλέξαμε, ενώ μια από τις πλευρές του είναι ο άξονας έντασης. Αφού δύο τις κορυφές είναι το μαύρο και άσπρο και το τρίτο είναι ένα τυχαίο σημείο χρώματος, όλα τα σημεία του τριγώνου θα έχουν την ίδια απόχρωση, μιας και το μαύρο και το άσπρο στοιχείο δεν μπορούν να αλλάξουν την απόχρωση (φυσικά, η ένταση και ο κορεσμός των σημείων

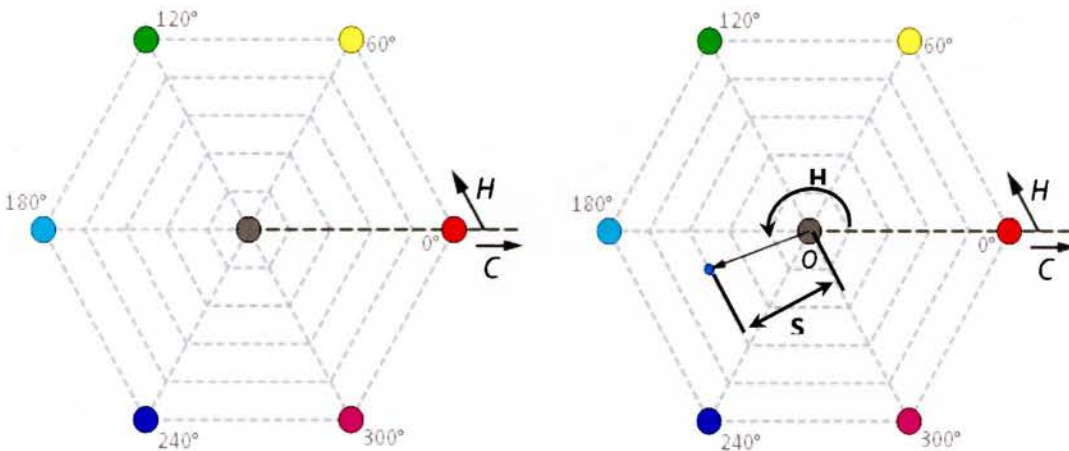
σε αυτό το τρίγωνο θα μπορούσαν να ήταν διαφορετικές). Περιστρέφοντας το νοητό τρίγωνο γύρω από τον κατακόρυφο άξονα της έντασης, δηλαδή απλά αλλάζοντας το χρώμα που επιλέξαμε άρα και τις μοίρες, το νοητό τρίγωνο θα αποκτήσει άλλη απόχρωση.



Εικόνα 24 Διαβαθμίσεις Κορεσμού (S)

Σύμφωνα με τα παραπάνω, φθάνουμε στο συμπέρασμα ότι η απόχρωση, ο κορεσμός και η τιμή της έντασης που απαιτούνται για να σχηματίσουν ένα HSI χώρο χρώματος μπορούν να ληφθούν από τον κύβο χρωμάτων RGB. Δηλαδή μπορούμε να μετατρέψουμε οποιοδήποτε σημείο RGB σε αντίστοιχο σημείο στο HSI χρωματικό μοντέλο.

Δεδομένου ότι τα επίπεδα κινούνται προς τα επάνω και προς τα κάτω στον άξονα της έντασης, τα όρια που ορίζονται από την τομή του κάθε επιπέδου με τις πλευρές του κύβου έχουν είτε ένα τριγωνικό ή εξαγωνικό σχήμα. Αυτό μπορεί να απεικονιστεί πολύ πιο εύκολα παρατηρώντας την Εικόνα 25(α). Σε αυτή βλέπουμε ότι τα βασικά χρώματα απέχουν μεταξύ τους 120° . Τα δευτερεύοντα χρώματα απέχουν 60° από τα βασικά, πράγμα που σημαίνει ότι η γωνία μεταξύ των δευτερευόντων χρωμάτων είναι επίσης 120° .

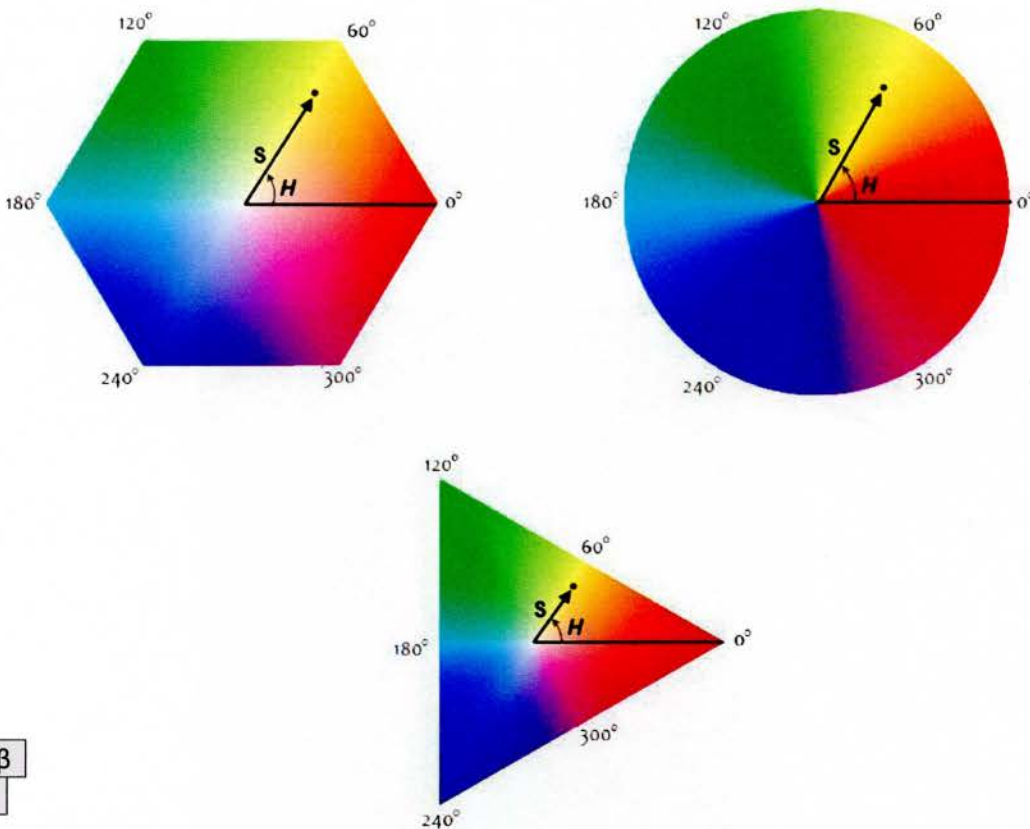


Εικόνα 25 Αναπαράσταση χρωματικού μοντέλου HIS. (α) Κάτοψη χρωματικού μοντέλου HIS. (β) Κάτοψη με επιλογή ενός τυχαίου σημείου

Στην Εικόνα 25(β) φαίνεται το ίδιο εξαγωνικό σχήμα και ένα τυχαίο σημείο χρώματος (φαίνεται σαν τελεία). Η απόχρωση του τυχαίου σημείου καθορίζεται από μια γωνία που σχηματίζεται μεταξύ του κόκκινου χρώματος και του τυχαίου σημείου. Συνήθως (αλλά όχι πάντα) ως αρχή των αξόνων ορίζεται το κόκκινο χρώμα που αντιστοιχεί στις 0° , άρα έχει και απόχρωση 0. Θα πρέπει να σημειωθεί πως η απόχρωση αυξάνεται από τα αριστερά προς τα δεξιά.

Ο κορεσμός (η απόσταση από τον κατακόρυφο άξονα) είναι το μήκος του ευθύγραμμου τμήματος από την αρχή των αξόνων μέχρι το τυχαίο σημείο. Ως αρχή των αξόνων καθορίζεται το σημείο του επιπέδου του χρώματος που εφάπτεται με τον κάθετο άξονα της έντασης. Τα σημαντικά στοιχεία του HSI χρωματικού χώρου είναι ο κάθετος άξονας της έντασης, το μήκος του ευθύγραμμου τμήματος και η γωνία που κάνει με το κόκκινο χρώμα.

Ως εκ τούτου, δεν είναι περίεργο να δούμε το χρωματικό πρότυπο HSI να αναπαρίσταται, εκτός από εξαγώνο όπου παρουσιάστηκε στην παρούσα εργασία, με τρίγωνο ή ακόμα και με κύκλο όπως φαίνεται στην Εικόνα 26(β) και (γ). Το σχήμα που επιλέγεται πραγματικά δεν έχει σημασία, δεδομένου ότι κάθε ένα από αυτά τα σχήματα μπορεί να είναι κυρτό σε ένα από τα άλλα δύο με γεωμετρικό μετασχηματισμό.



α	β
γ	

Εικόνα 26 Αναπαράσταση χρωματικού μοντέλου HIS ως (α) εξαγώνο, (β) κύκλο, (γ) τρίγωνο

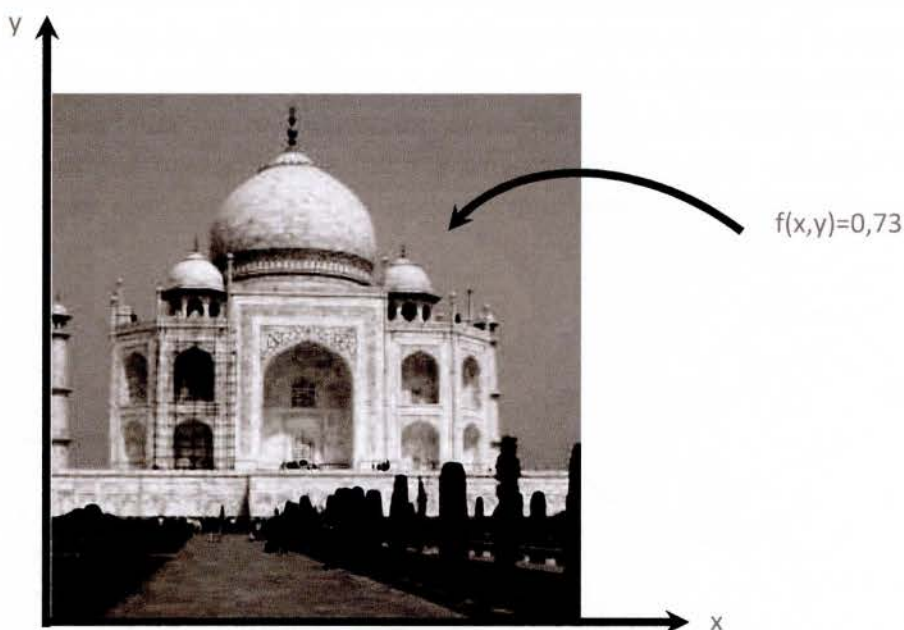
Κεφάλαιο 3

Ψηφιακή απεικόνιση εικόνων

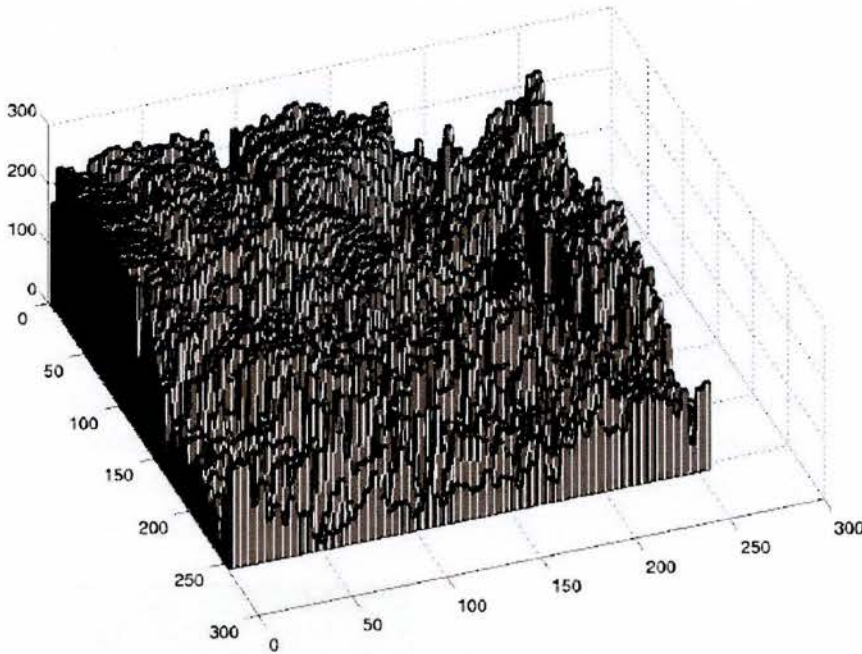
3.1 Εικόνες και ψηφιακές εικόνες

Μια ψηφιακή εικόνα αποτελείται από ένα αριθμό γραμμών και στηλών, το πλήθος των οποίων μας δίνει το συνολικό αριθμό των εικονοστοιχείων (pixels) που αποτελούν την ψηφιακή εικόνα. Τα εικονοστοιχεία μοιάζουν σαν μικρά πλακίδια, καθένα από τα οποία περιέχει μια τιμή. Αν υποθέσουμε ότι έχουμε μια ασπρόμαυρη εικόνα, δηλαδή μια εικόνα με αποχρώσεις του γκρι, τότε τα εικονοστοιχεία θα παίρνουν τιμές μεταξύ του διαστήματος $0 - 255$. Η τιμή αυτή δείχνει πόσο φωτεινό είναι το συγκεκριμένο εικονοστοιχείο.

Κάθε εικόνα είναι ουσιαστικά μια δυσδιάστατη συνάρτηση $f(x,y)$, όπου x και y είναι οι συντεταγμένες και f η τιμή φωτεινότητας σε οποιοδήποτε σημείου, όπως φαίνεται στην εικόνα 27. Μπορούμε να υποθέσουμε ότι σε μια τέτοια εικόνα οι τιμές φωτεινότητας της μπορεί να είναι οποιοσδήποτε πραγματικός αριθμός της τάξης 0 (μαύρο) σε 255 (λευκό). Το εύρος του x και του y θα εξαρτώνται από το πλήθος των γραμμών και των στηλών της εικόνας. Μια τέτοια συνάρτηση μπορεί να απεικονιστεί και με τον τρόπο που παρουσιάζεται στην Εικόνα (28). Η έννοια της εικόνας ως συνάρτηση, είναι ζωτικής σημασίας για την ανάπτυξη και εφαρμογή τεχνικών επεξεργασίας εικόνας.



Εικόνα 27 Η εικόνα ως συνάρτηση



Εικόνα 28 Η Εικόνα 27 σχεδιασμένη σαν συνάρτηση δύο μεταβλητών

Μια ψηφιακή εικόνα διαφέρει από μια φωτογραφία στο ότι οι τιμές των x και y , καθώς και του $f(x,y)$ είναι όλες διακριτές. Μια ψηφιακή εικόνα, όπως είδαμε παραπάνω, μπορεί να θεωρηθεί σαν μια σειρά από σημεία δειγματοληψίας, καθένα από τα οποία έχει μια ιδιαίτερη κβαντοποιημένη τιμή φωτεινότητας. Τα εικονοστοιχεία που περιβάλλουν ένα δεδομένο εικονοστοιχείο αποτελούν την “γειτονιά” του, όπου κάθε γειτονιά χαρακτηρίζεται από το σχήμα της. Μπορούμε να μιλήσουμε για παράδειγμα για γειτονικά pixels σε διάταξη 3x3 ή 3x5. Εκτός από πολύ ειδικές καταστάσεις, οι γειτονιές έχουν μονό αριθμό γραμμών και στηλών. Αυτό εξασφαλίζει ότι το τρέχον pixel είναι στο κέντρο της γειτονιάς. Ένα παράδειγμα δίνεται στην Εικόνα (29). Εάν μια γειτονιά έχει άρτιο αριθμό γραμμών ή στηλών (ή και τα 2), τότε είναι αναγκαίο να διευκρινιστεί ποιο pixel στην γειτονιά είναι το τρέχον pixel.

48	219	168	145	244	188	120	58
49	218	87	94	133	35	17	148
174	151	75	178	225	3	253	194
77	127	87	139	44	228	149	135
138	229	136	113	250	51	109	167
37	211	186	177	67	75	131	44
175	164	78	154	64	167	85	97
96	209	213	213	223	65	112	234

Τρέχον pixel

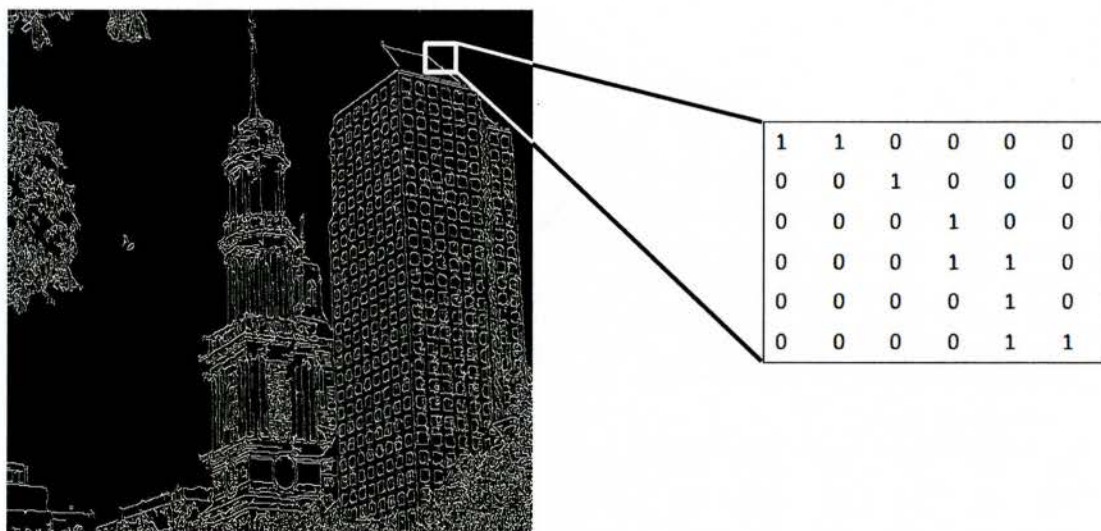
3 x 5 περιοχή

Εικόνα 29 Pixels σε μια περιοχή 3x5

3.2 Τύποι ψηφιακών εικόνων

Στην παρούσα ενότητα θα παρουσιαστούν οι 4 βασικοί τύποι εικόνων :

Δυαδικές (Binary): Σε αυτή την κατηγορία εικόνων κάθε ρixel μπορεί να παίρνει δύο μόνο πιθανές τιμές, 0 ή 1. Συνήθως τα δύο χρώματα που χρησιμοποιούνται για την αναπαράσταση μιας δυαδικής εικόνας είναι το μαύρο και το άσπρο, αν και οποιοδήποτε ζευγάρι χρωμάτων μπορεί να χρησιμοποιηθεί. Δεδομένου ότι υπάρχουν μόνο 2 πιθανές τιμές για κάθε ρixel χρειαζόμαστε μόνο ένα bit ανά ρixel. Τέτοιες εικόνες μπορούν κατά συνέπεια να είναι πολύ αποδοτικές από άποψη αποθήκευσης και ταχύτητας επεξεργασίας. Οι δυαδικές εικόνες είναι κατάλληλες για αρχιτεκτονικά σχέδια, εφαρμογές οπτικής αναγνώρισης χαρακτήρων ή δακτυλικών αποτυπωμάτων. Ένα παράδειγμα μιας δυαδικής εικόνας φαίνεται παρακάτω. Στην εικόνα, διακρίνονται μόνο δύο χρώματα : λευκό για της ακμές και μαύρο για το φόντο.



Εικόνα 30 Παράδειγμα δυαδικής εικόνας

Αποχρώσεις του γκρι (Grayscale): Οι εικόνες αυτού του είδους είναι γνωστές και σαν ασπρόμαυρες εικόνες. Κάθε ρixel μιας τέτοιας εικόνας προβάλλεται σαν μια απόχρωση του γκρι, συνήθως από 0 (απόλυτο μαύρο) ως 255 (φωτεινό λευκό). Σε αυτό το εύρος κάθε ρixel μπορεί να αναπαρασταθεί από 8 bits ή από ένα byte. Τέτοιες εικόνες χρησιμοποιούνται στην ιατρική και σε άλλες επιστήμες, αφού οι 256 αποχρώσεις του γκρι είναι επαρκείς για την αναγνώριση των πιο φυσικών αντικειμένων. Ένα παράδειγμα μιας εικόνας αποχρώσεων του γκρι διακρίνεται στην Εικόνα 31(α).

Έγχρωμες (RGB): Σε αυτήν την περίπτωση κάθε ρixel έχει ένα συγκεκριμένο χρώμα. Η περιγραφή του κάθε χρώματος καθορίζεται από την ποσότητα του κόκκινου, του πράσινου και του μπλε χρώματος που περιέχει. Αν κάθε ένα από αυτά τα χρώματα έχει εύρος 0-255, αυτό δίνει ένα σύνολο $255^2 = 16.777.215$ διαφορετικών πιθανών χρωμάτων. Έτσι έχουμε ένα αρκετά μεγάλο πλήθος χρωμάτων για κάθε εικόνα. Μια τέτοια εικόνα θεωρείται ως αποτέλεσμα μιας στοίβας τριών πινάκων, που αντιπροσωπεύουν τις τιμές του κόκκινου,

του πράσινου και του μπλε για κάθε pixel. Αυτό σημαίνει ότι για κάθε pixel αντιστοιχούν 3 τιμές. Στην Εικόνα 31(β) παρουσιάζεται ένα παράδειγμα μιας έγχρωμης εικόνας.



230	223	232	234	235	232	145
237	236	236	234	233	234	154
255	255	255	251	230	236	162
99	90	67	35	98	237	132
222	152	255	125	129	246	131
154	199	255	156	189	241	142
216	132	162	165	179	235	127



49	55	56	57	52	53
58	60	60	58	55	57
58	58	54	53	55	56
83	78	72	69	68	69
88	91	91	84	83	82
69	76	83	78	76	75
61	69	73	78	76	76

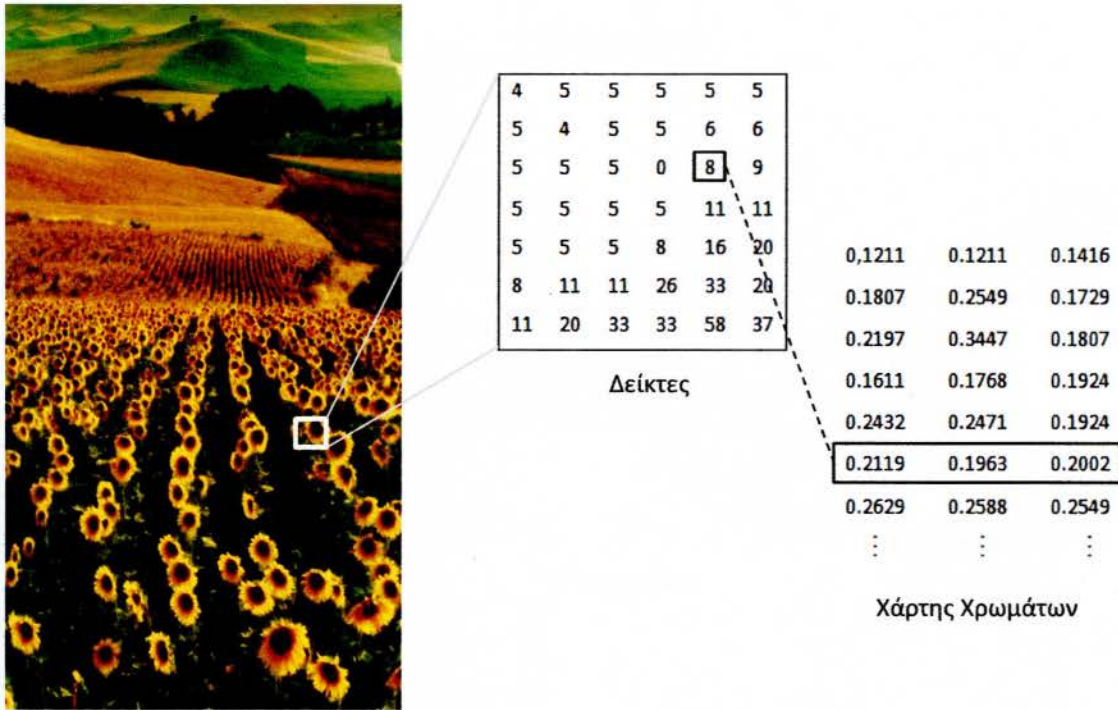
64	76	82	79	78	78
93	93	91	91	86	86
88	82	88	90	88	89
126	119	113	108	111	110
137	136	132	126	126	120
105	108	114	114	118	113
96	103	112	108	111	107

66	80	77	80	87	77
81	93	96	99	86	85
83	83	91	94	92	88
135	128	126	112	107	105
141	129	129	112	117	115
95	99	109	108	112	109
84	93	107	101	105	102

α
β

Εικόνα 31 (α) Εικόνα με αποχρώσεις του γκρι. (β) Έγχρωμη εικόνα

Δείκτης θέσης (Indexed): Οι περισσότερες έγχρωμες εικόνες έχουν μόνο ένα μικρό υποσύνολο από τα 16.777.215 διαφορετικά πιθανά χρώματα. Κάθε εικόνα για ευκολία στον χειρισμό, κατά την επεξεργασία, και την αποθήκευση της διαθέτει έναν σχετικό χρωματικό χάρτη, ο οποίος έχει μια λίστα από όλα τα χρώματα που περιέχονται σε αυτήν την εικόνα. Κάθε pixel έχει μια τιμή ή οποία δίνει το χρώμα της (όπως και για μια εικόνα RGB), αλλά και έναν δείκτη για το που είναι αυτό το χρώμα στο χάρτη. Με αυτόν τον τρόπο, αν μια εικόνα έχει 256 χρώματα ή λιγότερα τότε ο δείκτης τιμών θα απαιτήσει ένα byte κάθε φορά για αποθήκευση. Η Εικόνα 32 παρουσιάζει ένα παράδειγμα, όπου διακρίνονται οι δείκτες και η αντιστοιχία τους στον χρωματικό χάρτη.



Εικόνα 32 Παράδειγμα έγχρωμης εικόνας με δείκτες χρωμάτων

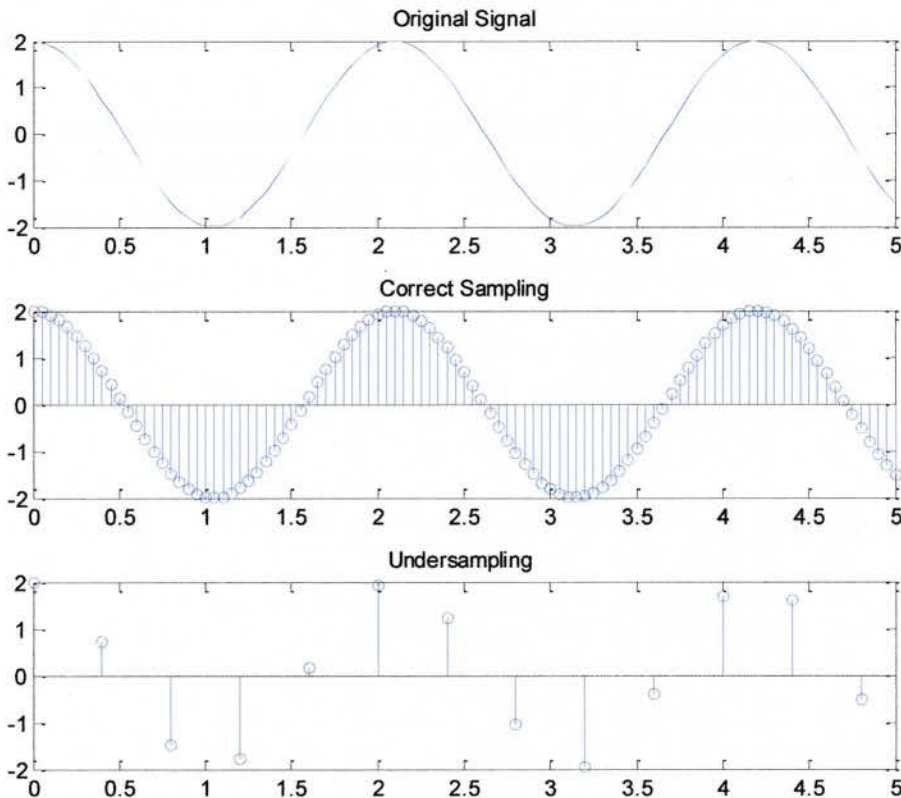
3.3 Απόκτηση εικόνας και δειγματοληψία

Τα σήματα που αντιμετωπίζουμε στην φύση είναι σήματα συνεχούς χρόνου. Η επεξεργασία των σημάτων συνεχούς χρόνου είναι δύσκολη και τις περισσότερες φορές αρκετά δαπανηρή. Αν με κάποιο τρόπο, μέσω μίας αξιόπιστης μεθόδου, καταφέρναμε να μετατρέψουμε τα σήματα συνεχούς χρόνου σε σήματα διακριτού χρόνου, τότε θα μπορούσαν να χρησιμοποιηθούν οι ηλεκτρονικοί υπολογιστές. Στις περισσότερες των περιπτώσεων, η επεξεργασία σημάτων διακριτού χρόνου είναι πολύ πιο εύκολη και προτιμητέα από την επεξεργασία σημάτων συνεχούς χρόνου. Με τον όρο δειγματοληψία αναφερόμαστε στην τεχνική του να «βλέπουμε» ένα συνεχές σήμα, κρατώντας τις τιμές του μερικές επιλεγμένες χρονικές στιγμές, ενώ το αγνοούμε τις υπόλοιπες.

Για παράδειγμα ας υποθέσουμε ότι παίρνουμε την συνάρτηση

$$y = 2\sigma\upsilon\upsilon(3x)$$

Στο πρώτο γράφημα διακρίνεται το αρχικό, αναλογικό, σήμα. Αν τώρα υποθέσουμε ότι έχουμε πάρει δείγματα του αρχικού σήματος σε 100 σημεία, όπως φαίνεται στο δεύτερο γράφημα, μπορούμε σαφώς να ανακατασκευάσουμε το αρχικό σήμα και όλες τις ιδιότητες του. Στο τρίτο γράφημα διακρίνονται λιγότερα από 10 δείγματα του αρχικού σήματος. Αυτό δείχνει ένα παράδειγμα «υποδειγματοληψίας» όπου ο αριθμός των σημείων δεν αρκούν για να ανακατασκευάσουν το αρχικό σήμα.

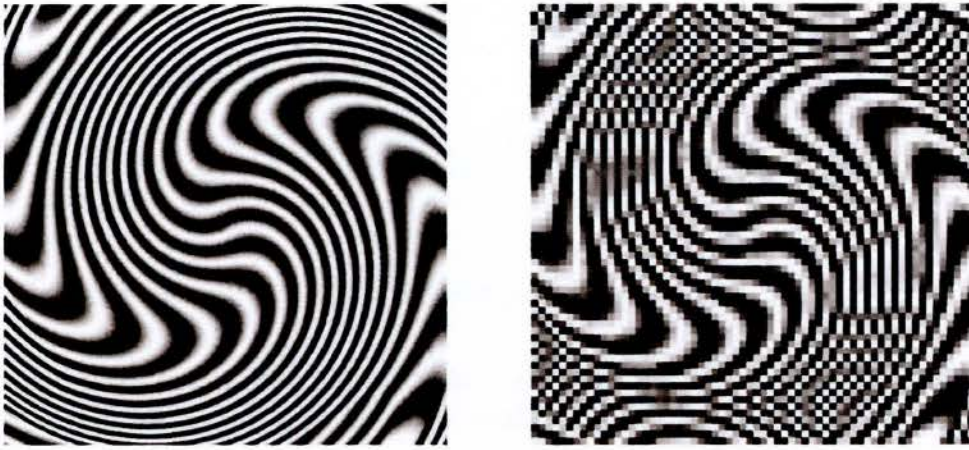


Εικόνα 33 Παράδειγμα δειγματοληψίας σήματος

Προκειμένου να διασφαλίσουμε ότι θα έχουμε αρκετά σημεία δειγματοληψίας, απαιτείται ότι η συχνότητα δειγματοληψίας (fs) δεν είναι μεγαλύτερη από το μισό της συχνότητας του σήματος μας. Αυτό είναι γνωστό και ως το θεώρημα Nyquist. Το θεώρημα Nyquist ή θεώρημα της δειγματοληψίας λέει πρακτικά ότι μια συνεχής συνάρτηση μπορεί να ανακατασκευαστεί από τα δείγματα της υπό τον όρο ότι η συχνότητα δειγματοληψίας είναι τουλάχιστον δύο φορές από τη μέγιστη συχνότητα της συνάρτησης.

$$fs \geq 2 f_{max}$$

Η δειγματοληψία μιας εικόνας απαιτεί την χρήση του κριτηρίου Nyquist, αφού θεωρούμε μια εικόνα ως συνεχή συνάρτηση δύο μεταβλητών και επιθυμούμε να πάρουμε δείγματα από αυτήν ώστε να παράγουμε μια ψηφιακή εικόνα. Ένα παράδειγμα φαίνεται στην Εικόνα



α β

Εικόνα 34 Παράδειγμα δειγματοληψίας, (α) Σωστή δειγματοληψία, (β) Υποδειγματοληψία

34(α) όπου μια εικόνα παρουσιάζεται με σωστή χρήση του κριτηρίου Nyquist και άρα σωστή δειγματοληψία. Αντίθετα στην Εικόνα 34(β) οι οδοντωτές άκρες στην εικόνα αποτελούν παράδειγμα υποδειγματοληψίας.

Κεφάλαιο 4

Μορφολογική Επεξεργασία Εικόνας

4.1 Βασικές έννοιες από την θεωρία συνόλων

Η έννοια του συνόλου δεν είναι εύκολο να προσδιοριστεί με απόλυτη μαθηματική ακρίβεια. Λέμε απλά ότι σύνολο είναι μία συλλογή αντικειμένων, όπου κάθε αντικείμενο που περιέχεται σ' ένα σύνολο ονομάζεται στοιχείο του συνόλου.

Τα στοιχεία των συνόλων με τα οποία ασχολείται η ψηφιακή επεξεργασία εικόνας είναι οι συντεταγμένες των pixels που αντιπροσωπεύουν αντικείμενα ή άλλα χαρακτηριστικά που παρουσιάζουν ενδιαφέρον σε μια εικόνα.

Έστω ότι A είναι ένα σύνολο στο δειγματικό χώρο Ω . Αν $\alpha = \{\alpha_1, \alpha_2\}$ είναι ένα στοιχείο του συνόλου A , τότε λέμε πως το α ανήκει στο σύνολο A και το συμβολίζουμε με :

$$\alpha \in A.$$

Παρομοίως, αν το α δεν είναι στοιχείο του A , τότε λέμε πως το α δεν ανήκει στο σύνολο A και το συμβολίζουμε με :

$$\alpha \notin A.$$

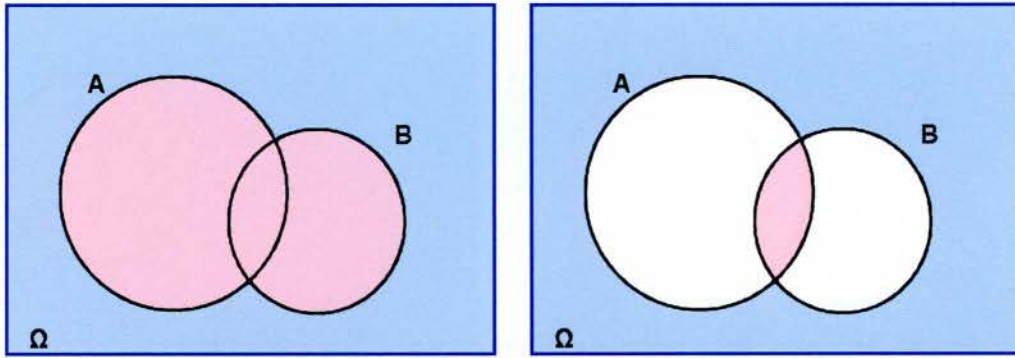
Το σύνολο που δεν έχει κανένα στοιχείο ονομάζεται μηδενικό ή κενό σύνολο και συμβολίζεται με το σύμβολο \emptyset . Ένα σύνολο καθορίζεται από το περιεχόμενο που βρίσκεται μεταξύ δύο αγκυλών $\{\}$. Πολλοί συγγραφείς χρησιμοποιούν δύο άδειες αγκύλες για να παρουσιάσουν ένα κενό σύνολο.

Η ένωση δύο συνόλων A και B είναι το σύνολο των στοιχείων που ανήκουν είτε στο A είτε στο B , είτε και στα δύο μαζί και συμβολίζεται ως εξής :

$$A \cup B.$$

Παρόμοια, η τομή των δύο συνόλων A και B είναι το σύνολο των στοιχείων που ανήκουν και στο A και στο B μαζί και συμβολίζεται ως εξής :

$$A \cap B.$$



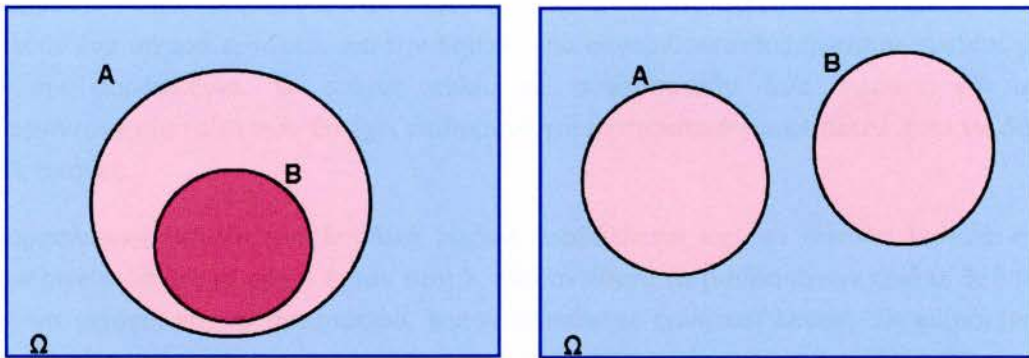
Εικόνα 35 (α) Ένωση δύο συνόλων. (β) Τομή δύο συνόλων.

Αν κάθε στοιχείο ενός συνόλου A είναι επίσης στοιχείο ενός άλλου συνόλου B , τότε λέμε ότι το σύνολο A είναι υποσύνολο του B και συμβολίζεται ως εξής :

$$A \subseteq B.$$

Δύο σύνολα A και B ονομάζονται ασυμβίβαστα όταν δεν έχουν κανένα κοινό στοιχείο. Σε αυτήν την περίπτωση συμβολίζονται :

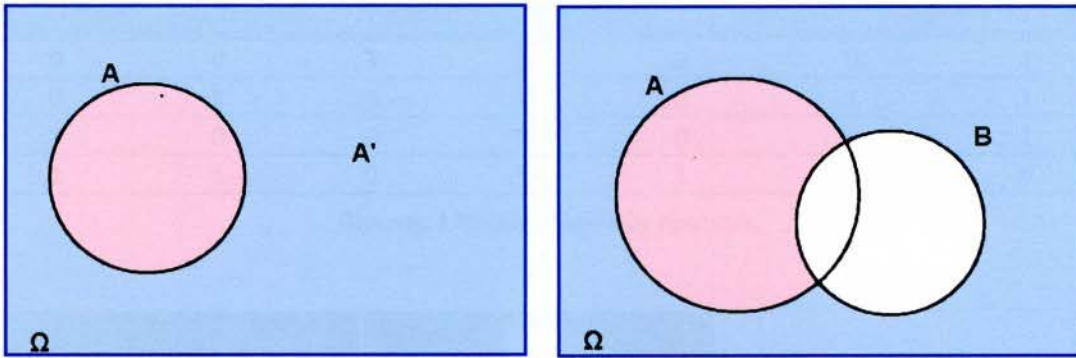
$$A \cap B = \emptyset.$$



Εικόνα 36 (α) Το B αποτελεί γνήσιο υποσύνολο του A . (β) Ασυμβίβαστα σύνολα.

Κάθε σύνολο έχει το συμπληρωματικό του. Έτσι το σύνολο A έχει ως συμπληρωματικό του το σύνολο A' , το οποίο περιέχει όσα στοιχεία δεν περιέχει το σύνολο A . Όπως φαίνεται και στην παρακάτω εικόνα για το συμπλήρωμα ενός συνόλου ισχύουν τα εξής :

$$A \cup A' = \Omega \quad \text{και} \quad A \cap A' = \emptyset.$$



Εικόνα 37 (α) Το A' αποτελεί συμπληρωματικό του A . (β) Η διαφορά συνόλων $A-B$.

Τέλος η διαφορά μεταξύ των συνόλων A και B ορίζεται το πλήθος των στοιχείων που περιλαμβάνονται στο σύνολο A και όχι στο B και συμβολίζεται ως εξής :

$$A - B.$$

4.2 Λογικές πράξεις εικόνων

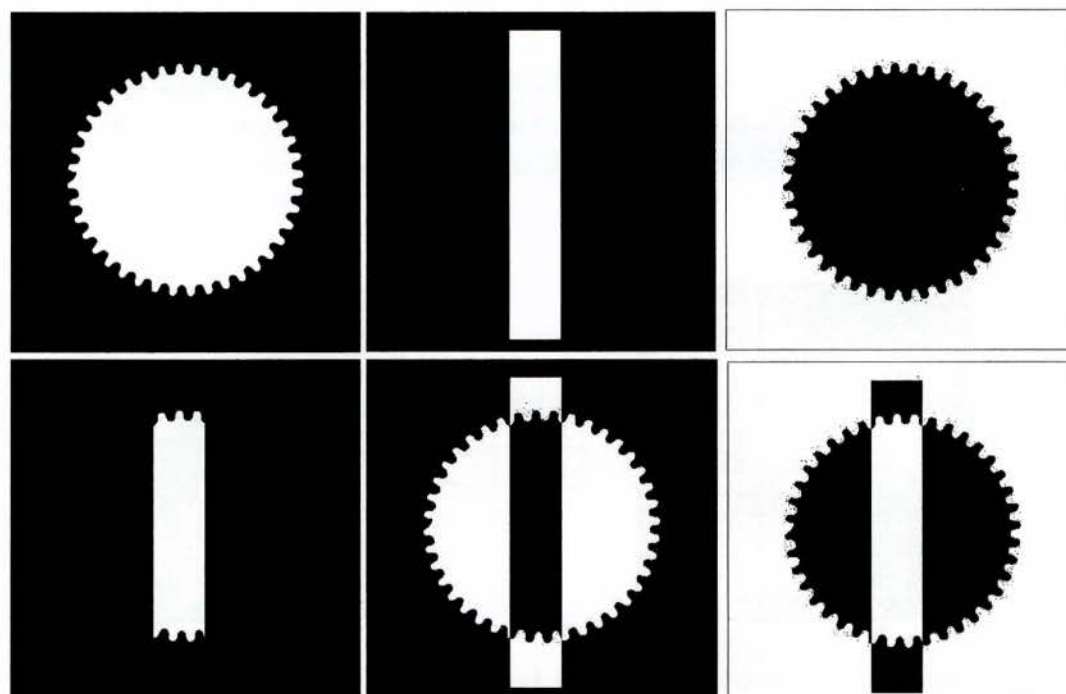
Η πλειοψηφία των εφαρμογών που βασίζονται στην μορφολογική επεξεργασία εικόνας περιλαμβάνουν δυαδικές εικόνες. Οι λογικές πράξεις, αν και απλές από την φύση τους, αποτελούν ένα ισχυρό εργαλείο για την δημιουργία αλγορίθμων επεξεργασίας εικόνας με βάση την μορφολογία. Ο στόχος είναι να συνδυαστούν δύο εικόνες για να δημιουργήσουν μία τρίτη που θα έχει επιθυμητά χαρακτηριστικά γνωρίσματα από τις δύο αρχικές εικόνες.

Στη μορφολογική θεωρία μια δυαδική εικόνα προβάλλεται ως ένα σύνολο λευκών και μαύρων pixels. Τα λευκά pixels έχουν τιμή 1, ενώ αντίθετα τα μαύρα έχουν τιμή 0. Βέβαια μπορεί να ισχύσει και το αντίστροφο, χρησιμοποιώντας αρνητική λογική. Οι κυριότερες λογικές πράξεις που χρησιμοποιούνται στην επεξεργασία εικόνας είναι η AND, η OR και η NOT. Οι ιδιότητες τους συνοψίζονται στον Πίνακα 1. Οι πράξεις αυτές ουσιαστικά μπορούν να συνδυαστούν και να σχηματίσουν οποιαδήποτε άλλη λογική πράξη.

Οι λογικές πράξεις εκτελούν τις βασικές λειτουργίες από την άλγεβρα Boole και πραγματοποιούνται σε ένα προς ένα pixel τα pixel μεταξύ δύο οι περισσότερων εικόνων. Αυτό δεν ισχύει στην περίπτωση της NOT, η οποία εφαρμόζεται αποκλειστικά στα pixels μιας ενιαίας εικόνας. Για παράδειγμα η συνάρτηση AND, δύο δυαδικών μεταβλητών, δίνει 1 μόνο όταν και οι δύο μεταβλητές είναι 1. Έτσι η παραγόμενη εικόνα έχει 1 στις θέσεις όπου και οι δύο εικόνες που συνδυάζονται έχουν 1. Σε οποιαδήποτε άλλη περίπτωση το νέο pixel θα έχει τιμή 0.

A	B	NOT A	A OR B	A AND B	A XOR B	A NAND B
0	0	1	0	0	0	1
0	1	1	1	0	1	1
1	0	0	1	0	1	1
1	1	0	1	1	0	0

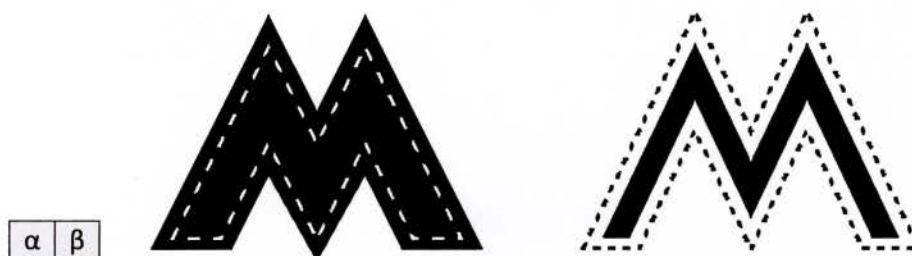
Πίνακας 1 Ιδιότητες λογικών πράξεων.



Εικόνα 38 Παραδείγματα λογικών πράξεων με εικόνες. (α) Αρχική εικόνα A. (β) Αρχική εικόνα B. (γ) Λογική πράξη NOT στην A. (δ) Λογική πράξη AND των A και B. (ε) Λογική πράξη XOR των A και B. (ζ) Λογική πράξη XNOR των A και B.

4.3 Συστολή και Διαστολή

Οι συναρτήσεις Συστολή (Erosion) και Διαστολή (Dilation) είναι οι δύο θεμελιώδεις συναρτήσεις της μορφολογική επεξεργασία εικόνας, με την έννοια ότι μπορούν να σχηματίσουν όλες τις άλλες συναρτήσεις που θα παρουσιαστούν και θα αναδειχθούν στο παρόν κεφάλαιο.

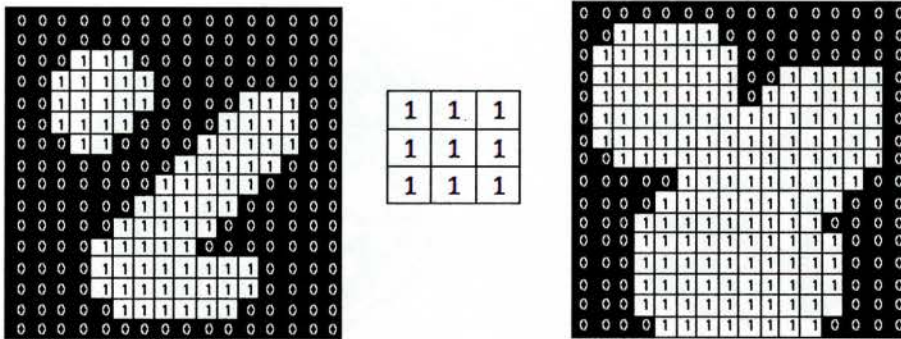


Εικόνα 39 (α) Διαστολή του γράμματος M. (β) Συστολή του γράμματος M.

4.3.1 Διαστολή (Dilation)

Η Διαστολή είναι μια λειτουργία που “μεγαλώνει” ή “παχαίνει” τα αντικείμενα σε μια δυαδική εικόνα. Η βασική επίδραση της είναι να αυξάνει σταδιακά τα όρια των pixels, των περιοχών που βρίσκονται στο προσκήνιο (π.χ. λευκά pixels με τιμή έντασης 1). Έτσι οι περιοχές των pixels που βρίσκονται προσκήνιο θα αυξηθούν σε μέγεθος, ενώ τα κενά τμήματα εντός των περιοχών αυτών θα γίνουν μικρότερα.

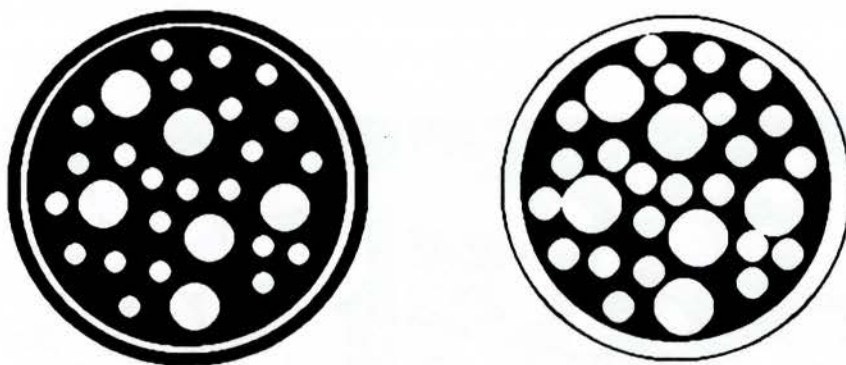
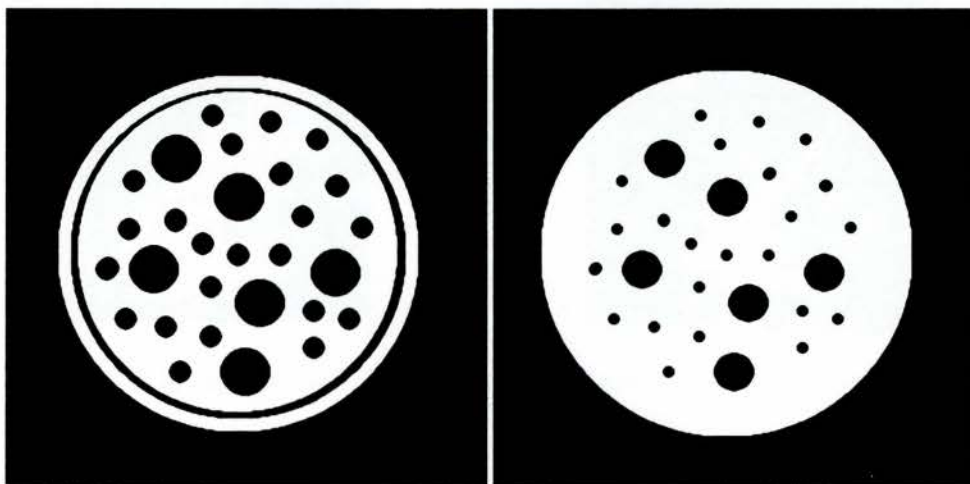
Ο τελεστής Διαστολής δέχεται σαν είσοδο του δύο δεδομένα. Το πρώτο είναι η εικόνα που πρόκειται να διασταλεί. Ο τρόπος και η έκταση της «πάχυνσης» ελέγχεται από ένα σύνολο σημείων (συνήθως μικρό), γνωστό ως πυρήνας (kernel), που αποτελεί το δεύτερο δεδομένο εισόδου. Αυτός ο πυρήνας είναι ουσιαστικά που καθορίζει το ποσοστό της διαστολής στις περιοχές της εικόνας εισόδου.



α β γ

Εικόνα 40 (α) Αρχική εικόνα) Πυρήνας 3x3. (γ) Εφαρμογή Διαστολής στην αρχική εικόνα.

Σαν παράδειγμα Διαστολής ας θεωρήσουμε την Εικόνα 40(α) ως μια δυαδική εικόνα, ενώ ως πυρήνα ένα τετράγωνο 3 × 3, όπως στην Εικόνα 40(β). Το αποτέλεσμα από την εφαρμογή του πυρήνα φαίνεται στην Εικόνα 40(γ), στην οποία παρατηρούμε τη διαστολή των pixels με τιμή έντασης 1.



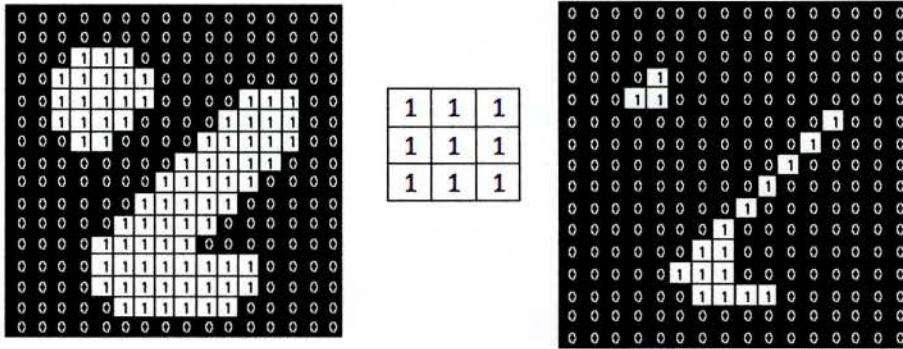
α	β
γ	δ

Εικόνα 41 (α) Αρχική εικόνα. (β) Εφαρμογή Διαστολής στην αρχική εικόνα (α). (γ) Αρχική εικόνα. (δ) Εφαρμογή Διαστολής στην αρχική εικόνα (γ).

Παραπάνω διακρίνονται δύο παραδείγματα Διαστολής, όπου από τις αρχικές εικόνες (α) και (γ) έχουν προέλθει οι εικόνες (β) και (δ). Στις παραγόμενες εικόνες παρατηρούμε την αύξηση της έκτασης των λευκών περιοχών. Παράλληλα στην Εικόνα 41(β) είναι προφανή η μείωση τις διαμέτρου των περιοχών με τιμή έντασης 0 (μαύρων περιοχών).

4.3.2 Συστολή (*Erosion*)

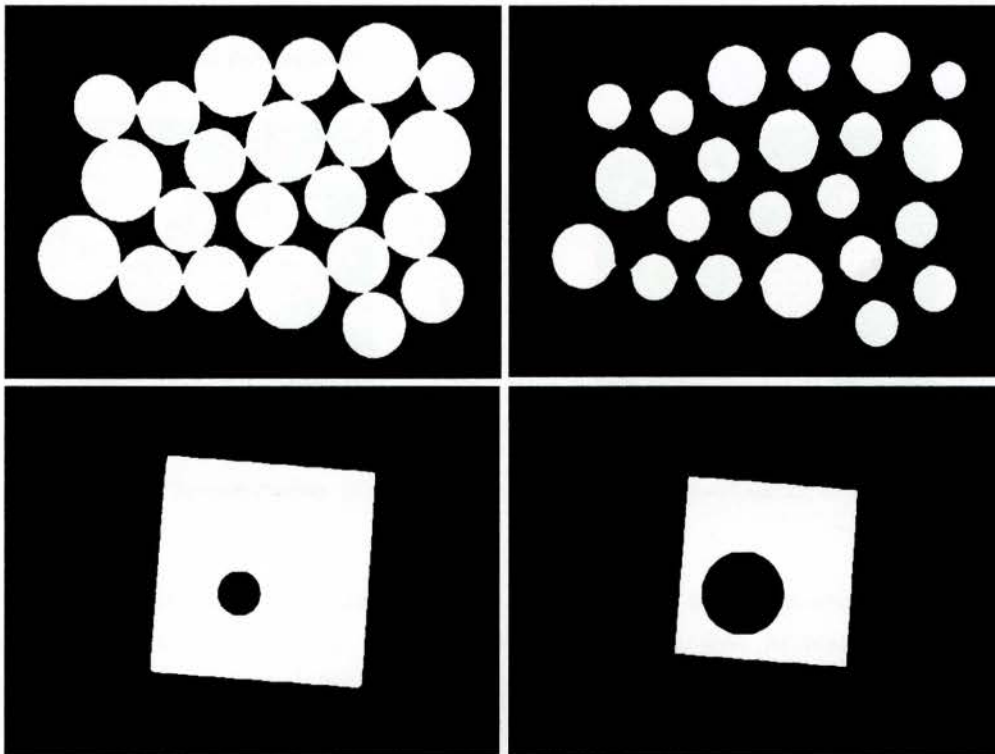
Η Συστολή “συρρικνώνει” ή “αραιώνει” αντικείμενα σε μια δυαδική εικόνα. Όπως και στην Διαστολή, ο τρόπος και η έκταση της συρρίκνωσης ελέγχεται από τον πυρήνα, ενώ αντίθετα από τη Διαστολή τα όρια των pixels, των περιοχών που έχουν τιμή έντασης 1 συρρικνώνονται σε μέγεθος και τα κενά τμήματα στις περιοχές αυτές γίνονται μεγαλύτερα. Η Εικόνα 31 απεικονίζει μια διαδικασία Συστολής με ένα πυρήνα διαστάσεων 3x3.



α β γ

Εικόνα 42 (α) Αρχική εικόνα. (β) Πυρήνας 3x3. (γ) Εφαρμογή Συστολής στην αρχική εικόνα.

Παρακάτω διακρίνονται δύο παραδείγματα Συστολής, όπου από τις αρχικές εικόνες (α) και (γ) έχουν προέλθει οι εικόνες (β) και (δ). Στις παραγόμενες εικόνες παρατηρούμε την μείωση της έκτασης των λευκών περιοχών. Παράλληλα στην Εικόνα 32(δ) είναι προφανή η αύξηση τις διαμέτρου των περιοχών με τιμή έντασης 0 (μαύρων περιοχών).



α β
γ δ

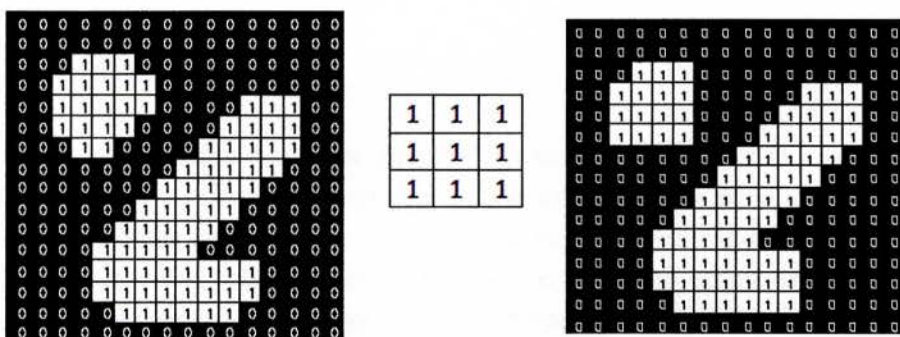
Εικόνα 43 (α) Αρχική εικόνα. (β) Εφαρμογή Συστολής στην αρχική εικόνα (α). (γ) Αρχική εικόνα. (δ) Εφαρμογή Συστολής στην αρχική εικόνα (γ).

4.4 Άνοιγμα και Κλείσιμο

Στην παραπάνω ενότητα παρουσιάστηκε η Διαστολή, η οποία επεκτείνει τα όρια των αντικειμένων μιας δυαδικής εικόνας, καθώς και η Συστολή η οποία τα συρρικνώνει. Στην παρούσα ενότητα θα παρουσιαστούν δύο άλλες μορφολογικές λειτουργίες : το Άνοιγμα (Opening) και το Κλείσιμο (Closing).

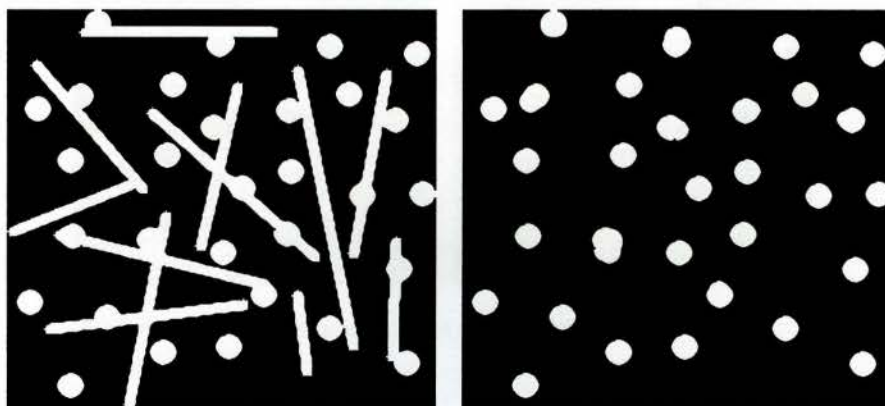
4.4.1 Άνοιγμα (Opening)

Η βασική επίδραση της λειτουργίας του Ανοίγματος μοιάζει σε μεγάλο βαθμό με τη Συστολή, δεδομένου ότι τείνει να αφαιρέσει κάποια από τα ρικελ που βρίσκονται στα άκρα των αντικειμένων μιας δυαδικής εικόνας. Ωστόσο αλλοιώνει το σχήμα των αντικειμένων σε μικρότερο βαθμό από την Συστολή. Όπως και με άλλους μορφολογικούς τελεστές, η ακριβής λειτουργία του τελεστή Ανοίγματος καθορίζεται από ένα πυρήνα. Σκοπός του τελεστή είναι να διατηρήσει τις περιοχές (με τιμή έντασης 1) της δυαδικής εικόνας που έχουν παρόμοιο σχήμα με τον πυρήνα ή του μοιάζουν απόλυτα, εξαλείφοντας παράλληλα όλες τις άλλες περιοχές της εικόνας. Παρακάτω φαίνεται η επίδραση του τελεστή Ανοίγματος.



Εικόνα 44 (α) Αρχική εικόνα. (β) Πυρήνας 3x3. (γ) Εφαρμογή Ανοίγματος στην αρχική εικόνα.

Με απλά λόγια σαν Άνοιγμα ορίζεται η Συστολή που ακολουθείται από μια Διαστολή χρησιμοποιώντας τον ίδιο πυρήνα και για τις δύο λειτουργίες. Συνεπώς ο τελεστής Ανοίγματος απαιτεί, δύο εισόδους: μια εικόνα και ένα πυρήνα.



Εικόνα 45 (α) Αρχική εικόνα. (β) Εφαρμογή Ανοίγματος στην αρχική εικόνα.

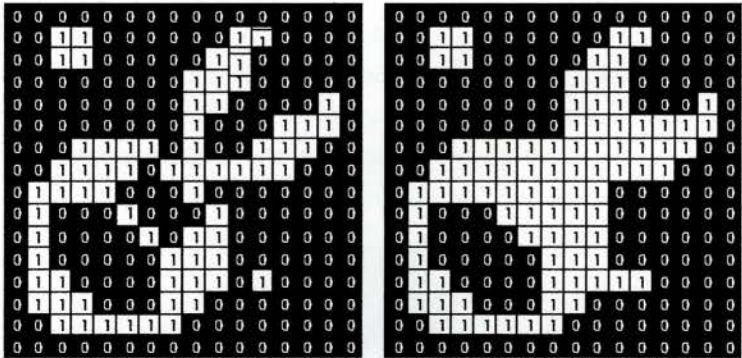
Στην Εικόνα 45 εμφανίζεται ένα παράδειγμα που φαίνονται έντονα η χρησιμότητα του τελεστή Ανοίγματος. Πιο συγκεκριμένα στην Εικόνα 45(α) διακρίνεται ένα μείγμα ευθειών γραμμών και κάποιων μικρών κύκλων. Αν υποθέσουμε πως θέλουμε να μετρήσουμε τους κύκλους θα πρέπει να απομακρύνουμε τις γραμμές. Αυτό επιτυγχάνεται χρησιμοποιώντας ως πυρήνα ένα κύκλο με διάμετρο 11 (Εικόνα 46), ώστε να μοιάζει με τις περιοχές που θέλουμε να διατηρήσουμε.

0	0	0	0	1	1	1	0	0	0	0
0	0	1	1	1	1	1	1	1	0	0
0	1	1	1	1	1	1	1	1	1	0
0	1	1	1	1	1	1	1	1	1	0
1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1
0	1	1	1	1	1	1	1	1	1	0
0	0	1	1	1	1	1	1	1	0	0
0	0	0	1	1	1	1	0	0	0	0

Εικόνα 46 Πυρήνας 11x11

4.4.2 Κλείσιμο (Closing)

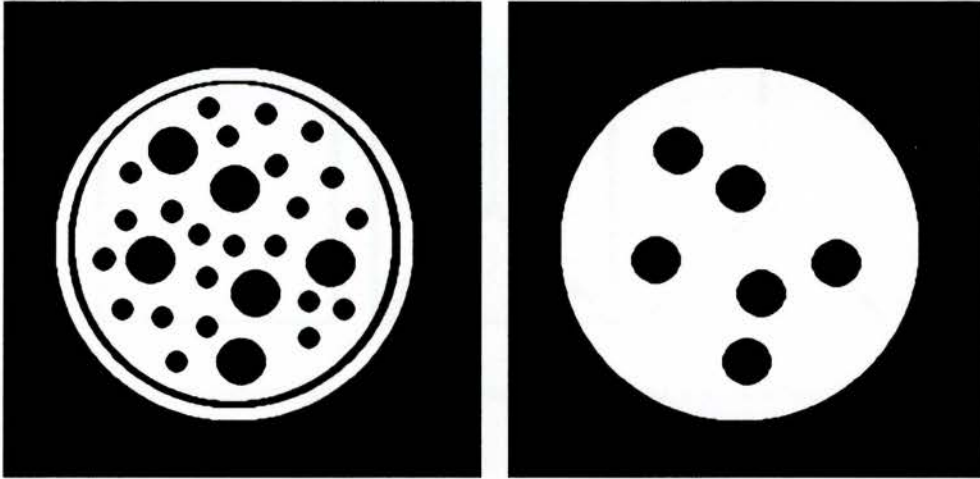
Ο τελεστής του Κλεισίματος έχει την ικανότητα να εξομαλύνει το περίγραμμα των αντικειμένων μιας δυαδικής εικόνας, να εξαλείφει τα μικρά κενά στο εσωτερικό των αντικειμένων, καθώς και να γεμίζει τα κενά στο περίγραμμα. Μοιάζει πάρα πολύ με τον τελεστή της Διαστολής, αφού τείνει να αυξάνει τα όρια των αντικειμένων μιας δυαδικής εικόνας, αλλοιώνοντας όμως το σχήμα τους λιγότερο. Για την εφαρμογή του τελεστή είναι και εδώ απαραίτητη η χρήση ενός πυρήνα, ο οποίος θα διατηρήσει τις περιοχές (με τιμή έντασης 0) της δυαδικής εικόνας που έχουν παρόμοιο σχήμα με τον πυρήνα ή του μοιάζουν απόλυτα, εξαλείφοντας παράλληλα όλες τις άλλες περιοχές της εικόνας. Παρακάτω φαίνεται η επίδραση του τελεστή Κλεισίματος.



α β

Εικόνα 47 (α) Αρχική εικόνα. (β) Εφαρμογή Κλεισίματος στην αρχική εικόνα.

Ο τελεστής Κλεισίματος μπορεί να παραχθεί εφαρμόζοντας αρχικά τον τελεστή Διαστολής και στη συνέχεια τον τελεστή Συστολής, αρκεί να γίνει χρήση του ίδιου πυρήνα και στις δύο περιπτώσεις. Στην Εικόνα 48 παρουσιάζεται ένα παράδειγμα χρήσης του τελεστή Κλεισίματος. Στην Εικόνα 48(α) διακρίνεται ένας κύκλος που περιέχει μικρότερα και μεγαλύτερα κενά. Αν είναι επιθυμητή η εξάλειψη των μικρών κενών και η διατήρηση των μεγάλων, θα πρέπει να χρησιμοποιήσουμε ως πυρήνα ένα δίσκο με διάμετρο μεγαλύτερη από αυτή των μικρών κενών και ταυτόχρονα μικρότερη των μεγάλων.



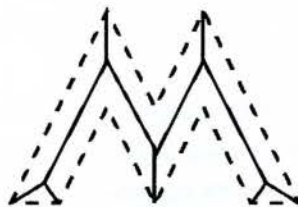
α β

Εικόνα 48 (α) Αρχική εικόνα. (β) Εφαρμογή Κλεισίματος στην αρχική εικόνα.

4.5 Βασικοί Μορφολογικοί Αλγόριθμοι

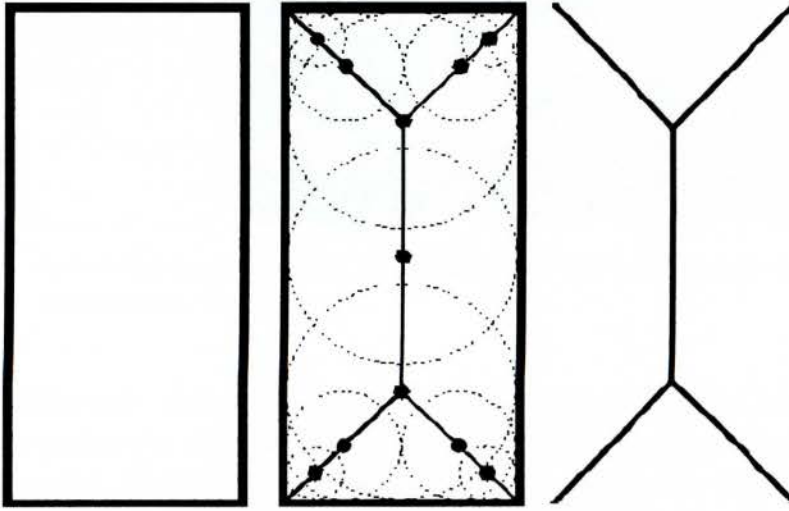
4.5.1 Σκελετοποίηση (Skeletonization)

Η Σκελετοποίηση είναι μια διαδικασία στη μορφολογική επεξεργασία εικόνας κατά την οποία προκαλείται μείωση της περιμέτρου ενός αντικειμένου, σε μια δυαδική εικόνα. Το νέο αντικείμενο που θα προκύψει αποτελεί ουσιαστικά μια λεπτή έκδοση του αρχικού αντικειμένου, τα όρια της οποίας ισαπέχουν από την αρχική περίμετρο. Ο σκελετός που θα προκύψει τονίζει συνήθως γεωμετρικές και τοπολογικές ιδιότητες του αρχικού σχήματος, όπως η συνδεσιμότητα, η τοπολογία, το μήκος και το πλάτος του. Σε συνδυασμό με την ισαπέχουσα απόσταση από τα όρια του αρχικού αντικειμένου, ο σκελετός μπορεί επίσης να χρησιμεύσει ως αναπαράσταση του αρχικού αντικειμένου, αφού περιέχει όλες τις πληροφορίες που είναι αναγκαίες για την ανακατασκευή του.



Εικόνα 49 Σκελετοποίηση του γράμματος M

Ένας άλλος τρόπος για να γίνει πιο κατανοητή η διαδικασία της Σκελετοποίησης είναι να θεωρήσουμε κύκλους που εφάπτονται στα όρια του αρχικού αντικειμένου, τα κέντρα των οποίων αποτελούν το παραγόμενο σχήμα. Το σχήμα αυτό είναι ο σκελετός.



α β γ

Εικόνα 50 (α) Αρχική εικόνα. (β) Εφαρμογή σκελετισμού. (γ) Σκελετοποιημένη εικόνα.

Στις περισσότερες εφαρμογές της ψηφιακής επεξεργασίας εικόνας θα πρέπει να γίνει μετατροπή μιας έγχρωμης ή ασπρόμαυρης εικόνας σε δυαδική και στη συνέχεια να εφαρμοστεί Σκελετοποίηση. Αν τώρα εμείς πρέπει να βρούμε το σχήμα ενός αντικειμένου, υπάρχουν περιπτώσεις που ο σκελετός μπορεί να γίνει αρκετά πολύπλοκος. Παρακάτω διακρίνεται ένα τέτοιο παράδειγμα σκελετισμού μιας εικόνας.

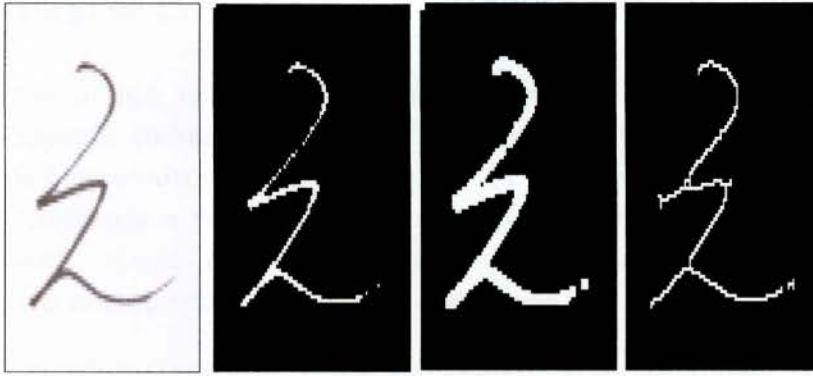


α β γ

Εικόνα 51 (α) Αρχική εικόνα. (β) Δυαδική εικόνα. (γ) Σκελετοποιημένη εικόνα.

4.5.2 Λέπτυνση (*Thinning*)

Μια άλλη μέθοδος για να μειώσουμε την περίμετρο ενός αντικειμένου είναι με τη χρήση της Λέπτυνσης. Η Λέπτυνση είναι και αυτή μια μορφολογική διαδικασία που μπορεί να εφαρμοστεί σε μία δυαδική εικόνα. Είναι μια απλή και γρήγορη τεχνική η οποία δεν είναι αρκετά αξιόπιστη, καθώς προσπαθεί να μειώσει το πάχος όλων των γραμμών αφήνοντας μία λεπτή γραμμή στο πάχος ενός pixel.



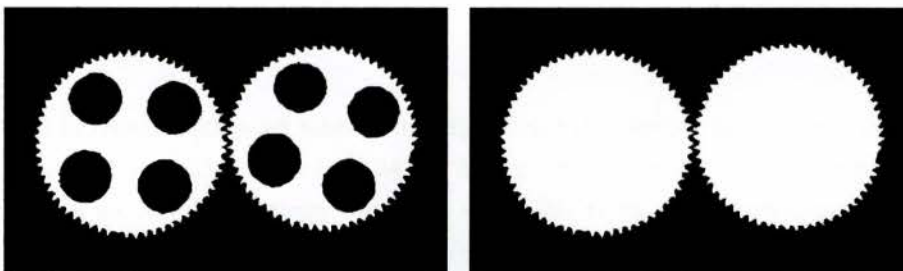
α β γ δ

Εικόνα 52 (α) Αρχική εικόνα. (β) Δυαδική εικόνα. (γ) Εφαρμογή Διαστολής. (δ) Εφαρμογή Λέπτυνσης.

Παραπάνω διακρίνουμε ένα παράδειγμα εφαρμογής Λέπτυνσης. Όπως συμβαίνει με τις περισσότερες εφαρμογές στην ψηφιακή επεξεργασία εικόνας, έτσι και εδώ η αρχική ασπρόμαυρη εικόνα Εικόνα 52(α) μετατρέπεται σε δυαδική Εικόνα 52(β). Παρατηρούμε όμως ότι υπάρχουν σημεία ασυνέχειας. Για να τα εξαλείψουμε χρησιμοποιούμε τη διαδικασία της Διαστολής στην δυαδική εικόνα. Στην εικόνα που προκύπτει Εικόνα 52(γ) εφαρμόζεται η διαδικασία της Λέπτυνσης. Έτσι έχουμε μια νέα δυαδική εικόνα πολύ πιο ξεκάθαρη από αυτή που είχαμε αρχικά.

4.5.3 Συμπλήρωση Οπών (Filling Holes)

Μια ακόμα χαρακτηριστική λειτουργία της μορφολογικής επεξεργασίας εικόνας είναι και η Συμπλήρωση Οπών (Filling Holes). Σαν οπή μπορεί να οριστεί μια περιοχή του φόντου (pixels με τιμή έντασης 0), η οποία περιβάλλεται από μία μεγαλύτερη περιοχή που βρίσκεται στο προσκήνιο (pixels με τιμή έντασης 1). Σκοπός του τελεστή Συμπλήρωσης Οπών είναι να καλύψει αυτές τις οπές.



α β

Εικόνα 53 (α) Αρχική εικόνα. (β) Εφαρμογή Συμπλήρωσης Οπών στην αρχική εικόνα.

Στην Εικόνα 53(α) παρουσιάζεται η δυαδική απεικόνιση δύο οδοντωτών τροχών. Με τη χρήση του τελεστή Συμπλήρωσης Οπών οι τέσσερις οπές που υπάρχουν σε κάθε οδοντωτό τροχό θα συμπληρωθούν από pixels με τιμή έντασης 1. Έτσι πλέον ο κάθε οδοντωτός τροχός θα παρουσιάζεται σαν ένα ενιαίο σχήμα χωρίς οπές.

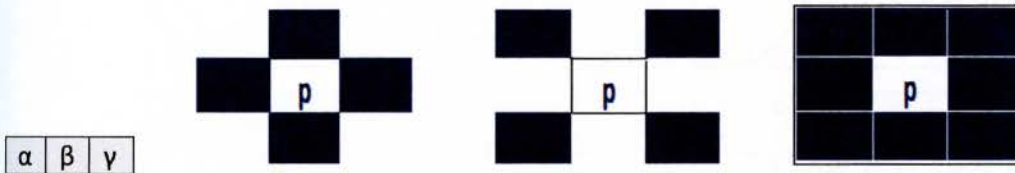
4.5.4 Συνδεδεμένα Στοιχεία (*Connected Components*)

Η συνδεσιμότητα μεταξύ των pixels αποτελεί μία από τις θεμελιώδεις έννοιες στην ψηφιακή επεξεργασία εικόνας και απλοποιεί πολλούς από τους ορισμούς της. Για να θεωρηθούν δύο ή περισσότερα pixels συνδεδεμένα θα πρέπει να προσδιορίζεται αν είναι γειτονικά και παράλληλα η τιμή έντασης τους να ικανοποιεί ένα συγκεκριμένο κριτήριο ομοιότητας. Αυτό οδηγεί στην έννοια των Συνδεδεμένων Στοιχείων (*Connected Components*), που αναφέρονται επίσης και σαν αντικείμενα.

Σε πολλές εφαρμογές επεξεργασίας εικόνας δεν αρκεί απλά η εύρεση τους αλλά κρίνεται απαραίτητος και ο προσδιορισμός τους. Έτσι σε μία ψηφιακή εικόνα με πολλά αντικείμενα θα πρέπει να προσδιορίσουμε τα Συνδεδεμένα Στοιχεία για να βρούμε τις θέσεις των αντικειμένων. Ο προσδιορισμός αυτός ορίζεται ως Ονοματισμός Στοιχείων (*Components Labeling*).

Ανάλυση Συνδεδεμένων Στοιχείων (*Connected Components Analysis*)

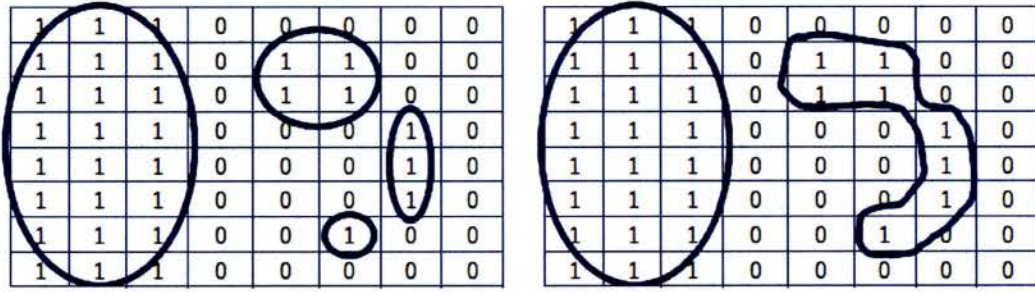
Ο αλγόριθμος των Συνδεδεμένων Στοιχείων έχει την ικανότητα να βρίσκει όλα τα συνδεδεμένα στοιχεία μιας ψηφιακής εικόνας, με βάση την συνδεσιμότητα τους. Για να γίνει αυτό πιο κατανοητό ας θεωρήσουμε ένα pixel p , μιας ψηφιακής εικόνας, στις συντεταγμένες (x,y) . Το pixel αυτό έχει δύο οριζόντια και δύο κάθετα γειτονικά pixels των οποίων οι συντεταγμένες είναι $(x + 1,y)$, $(x-1,y)$, $(x,y+1)$ και $(x,y-1)$. Αυτό το σύνολο των 4-γειτονικών pixels του p , συμβολίζεται με $N_4(p)$ και παρουσιάζεται στην Εικόνα 54(α).



Εικόνα 54 (α) Pixel p και 4-γείτονες. (β) Pixel p και οι 4 διαγώνιοι γείτονες. (γ) Pixel p και 8-γείτονες.

Τα τέσσερα διαγώνια γειτονικά pixels του p έχουν συντεταγμένες $(x + 1, y + 1)$, $(x + 1, y - 1)$, $(x - 1, y - 1)$ και $(x - 1, y + 1)$. Η Εικόνα 54(β) δείχνει τα γειτονικά αυτά pixels, τα οποία συμβολίζονται με $N_D(p)$. Η ένωση του $N_4(p)$ και $N_D(p)$ είναι τα 8-γειτονικά pixel του p και συμβολίζονται με $N_8(p)$.

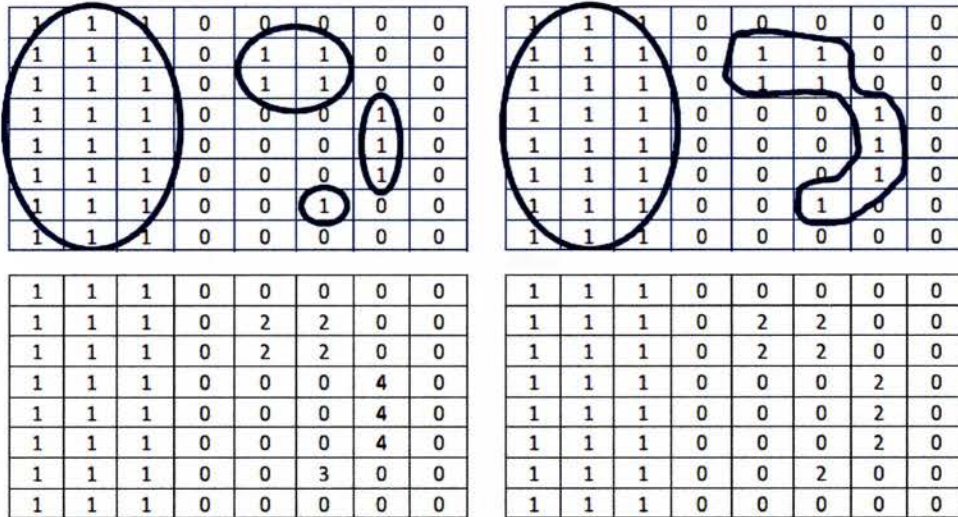
Ο αλγόριθμος των Συνδεδεμένων Στοιχείων για να λειτουργήσει θα πρέπει να οριστεί η μορφή γειτνίασης μεταξύ των pixels, η πιο κοινή είναι η 4 και 8 γειτνίαση. Η Εικόνα 44 δείχνει την επίδραση που μπορεί να έχει η γειτνίαση στον καθορισμό του αριθμού των συνδεδεμένων στοιχείων σε μια ψηφιακή εικόνα. Η Εικόνα 55(α) δείχνει μια μικρή δυαδική εικόνα με 4 συνδεδεμένα στοιχεία και 4-γειτνίαση, ενώ η Εικόνα 55(β) δείχνει ότι η επιλογή της 8-γειτνίασης μειώνει τον αριθμό των συνδεδεμένων στοιχείων σε 2.



Εικόνα 55 (α) 4 συνδεδεμένα στοιχεία με 4-γειτνίαση. (β) 2 συνδεδεμένα στοιχεία με 8-γειτνίαση.

Ονοματισμός Στοιχείων (Components Labeling)

Αφού έχουμε εντοπίσει τα συνδεδεμένα στοιχεία σε μια ψηφιακή εικόνα το επόμενο βήμα είναι μπορέσουμε να προσδιορίσουμε με μεγαλύτερη ακρίβεια τη θέση τους. Ο προσδιορισμός αυτός γίνεται με τον αλγόριθμο Ονοματισμού Στοιχείων, ο οποίος σαρώνει όλη την ψηφιακή εικόνα και ομαδοποιεί τα pixels σε στοιχεία, σύμφωνα με την συνδεσιμότητά τους. Έτσι όλα τα pixels ενός συνδεδεμένου στοιχείου έχουν κοινή τιμή έντασης και κατά κάποιον τρόπο συνδέονται μεταξύ τους. Μόλις όλα τα στοιχεία έχουν εντοπιστεί, σε κάθε pixel δίνεται μία τιμή ανάλογα με το στοιχείο που ανήκει.



Εικόνα 56 (α) 4 συνδεδεμένα στοιχεία με 4-γειτνίαση. (β) 2 συνδεδεμένα στοιχεία με 8-γειτνίαση. (γ) Ονοματισμός στοιχείων με 4-γειτνίαση. (δ) Ονοματισμός στοιχείων με 8-γειτνίαση

Εφαρμόζοντας των αλγόριθμο Ονοματισμού Στοιχείων για τις εικόνες 56(α) και 56(β) προκύπτουν οι νέες ψηφιακές εικόνες. Τα pixels των νέων εικόνων χαρακτηρίζονται από ένα μοναδικό ακέραιο αριθμό ανάλογα με το συνδεδεμένο στοιχείο στο οποίο ανήκουν. Ο ακέραιος αριθμός αυτός κυμαίνεται από 1 μέχρι N, όπου N είναι το πλήθος των συνδεδεμένων στοιχείων της νέας εικόνας. Στην Εικόνα 56(α) ισχύει $N=4$, ενώ στην Εικόνα 56(β) $N=2$. Επομένως τα pixels με ένδειξη 1 ανήκουν στην πρώτο συνδεδεμένο στοιχείο ενώ τα pixels με ένδειξη 2 ανήκουν στο δεύτερο και ούτω καθεξής. Τα pixel του φόντου διατηρούν την τιμή 0.

Κεφάλαιο 5

Κατάτμηση Ψηφιακής Εικόνας

5.1 Εισαγωγικά

Στην μηχανική όραση, κατάτμηση ορίζεται η διαδικασία του διαχωρισμού μιας ψηφιακής εικόνας σε επιμέρους τμήματα (segments). Σκοπός της κατάτμησης είναι η απλοποίηση και η αναπαράσταση μιας ψηφιακής εικόνας με τέτοιο τρόπο ώστε να είναι ευκολότερη και πιο ουσιαστική η ανάλυση της. Η κατάτμηση ψηφιακής εικόνας χρησιμοποιείται συνήθως για τον εντοπισμό αντικειμένων ή ορίων.

Το αποτέλεσμα της κατάτμησης μιας ψηφιακής εικόνας είναι ένα σύνολο από σχήματα ή αντικείμενα που καλύπτουν ολόκληρη την εικόνα. Όλα τα pixels σε μια περιοχή, που διαχωρίζεται κατά την διαδικασία της κατάτμησης, έχουν κάποια κοινά χαρακτηριστικά, όπως το χρώμα, η ένταση ή η υφή.

Η κατάτμηση μιας ψηφιακής εικόνας αποτελεί ίσως το πιο σημαντικό και δυσκολότερο κομμάτι στην επεξεργασία εικόνας και είναι καθοριστικό για την επιτυχία ή μη της τελικής ανάλυσης. Για το λόγο αυτό, ο αριθμός των μεθόδων κατάτμησης εικόνας που έχουν αναπτυχθεί είναι σχεδόν ο ίδιος με τον αριθμό των προβλημάτων κατάτμησης που υπάρχουν.

5.2 Κατωφλίωση

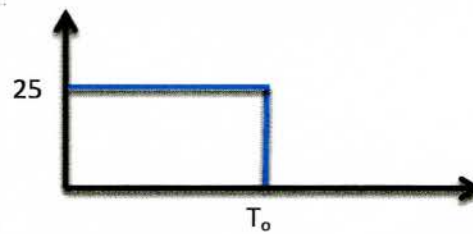
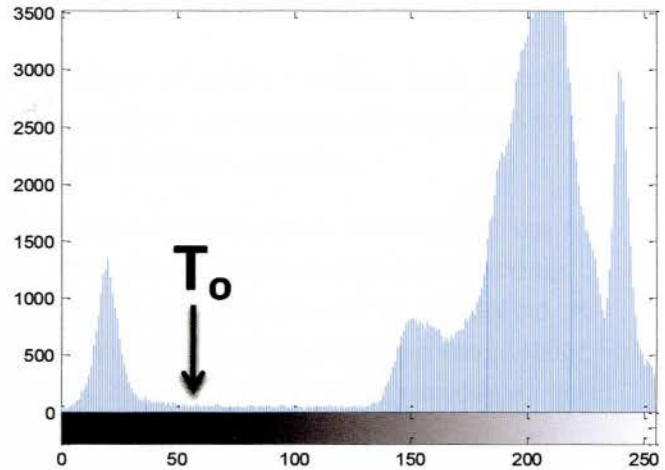
Κατωφλίωση (*Thresholding*) είναι η απλούστερη μέθοδος που χρησιμοποιείται για την κατάτμηση των ψηφιακών εικόνων. Σε πολλές εφαρμογές της μηχανικής όρασης είναι απαραίτητος ο διαχωρισμός των περιοχών μιας ψηφιακής εικόνας, που αντιστοιχούν σε αντικείμενα που μας ενδιαφέρουν, από τις περιοχές της εικόνας που αντιστοιχούν στο παρασκήνιο (background). Η κατωφλίωση συχνά προσφέρει έναν εύκολο και βολικό τρόπο, ώστε να πραγματοποιηθεί αυτός ο διαχωρισμός με βάση τις διαφορετικές τιμές έντασης ή τα χρώματα μεταξύ των περιοχών του προσκήνιου (foreground) και του παρασκήνιου της εικόνας. Με την εφαρμογή της κατωφλίωσης σε μια εικόνα αποχρώσεων του γκρι παράγεται μια δυαδική εικόνα.

Οι κυριότερες μέθοδοι κατωφλίωσης παρουσιάζονται στις επόμενες ενότητες.

5.2.1 Καθολική Κατωφλίωση (Global Thresholding)

Όταν το διάστημα των τιμών έντασης ενός αντικειμένου είναι διαφορετικών αποχρώσεων του γκρι από την υπόλοιπη εικόνα, τότε μπορούμε να χρησιμοποιήσουμε τον απλούστερο τύπο κατωφλίωσης, την Καθολική Κατωφλίωση (*Global Thresholding*). Στην Εικόνα 57(α) παρατηρούμε πως το χρώμα των χαρακτήρων είναι πιο έντονο από το χρώμα του

υπόβαθρου. Αυτό έχει σαν αποτέλεσμα το ιστογράμμα της εικόνας, που δίνεται στην Εικόνα 57(β) να αποτελείται από δύο λοβούς, διαχωρισμένους μεταξύ τους. Αν εμείς θέλουμε να διατηρήσουμε τους μαύρους χαρακτήρες της Εικόνας 57(α), τότε αρκεί να θεωρήσουμε σαν τιμή κατωφλίσωσης $T_0=60$.



α	β
γ	δ

Εικόνα 57 (α) Αρχική Εικόνα. (β) Ιστογράμμα αρχικής εικόνας. (γ) Εικόνα με Καθολική Κατωφλίωση. (δ) Βαθυπερατό φίλτρο.

Με την τιμή αυτή διαχωρίζονται ξεκάθαρα οι δύο λοβοί του ιστογράμματος. Κάνοντας χρήση του βαθυπερατού φίλτρου της Εικόνας 46(δ) όσα pixels της αρχικής εικόνας έχουν τιμή έντασης μικρότερη ή ίση με $T_0=60$, στην Εικόνα 46(γ), γίνονται μαύρα ενώ όσα έχουν μεγαλύτερη γίνονται λευκά.

Στο παραπάνω παράδειγμα χρησιμοποιήθηκε μία και μοναδική τιμή κατωφλίου, σύμφωνα με την οποία έγινε και η σύγκριση των τιμών έντασης όλων των pixels της αρχικής εικόνας. Το μεγάλο μειονέκτημα της Καθολικής Κατωφλίσωσης είναι πως λαμβάνει αποκλειστικά την ένταση, σαν μέγεθος για να προσδιορίσει τη σχέση μεταξύ των pixels της ψηφιακής εικόνας. Έτσι δεν υπάρχει καμία βεβαιότητα ότι τα προσδιοριζόμενα pixels είναι συνεχόμενα.

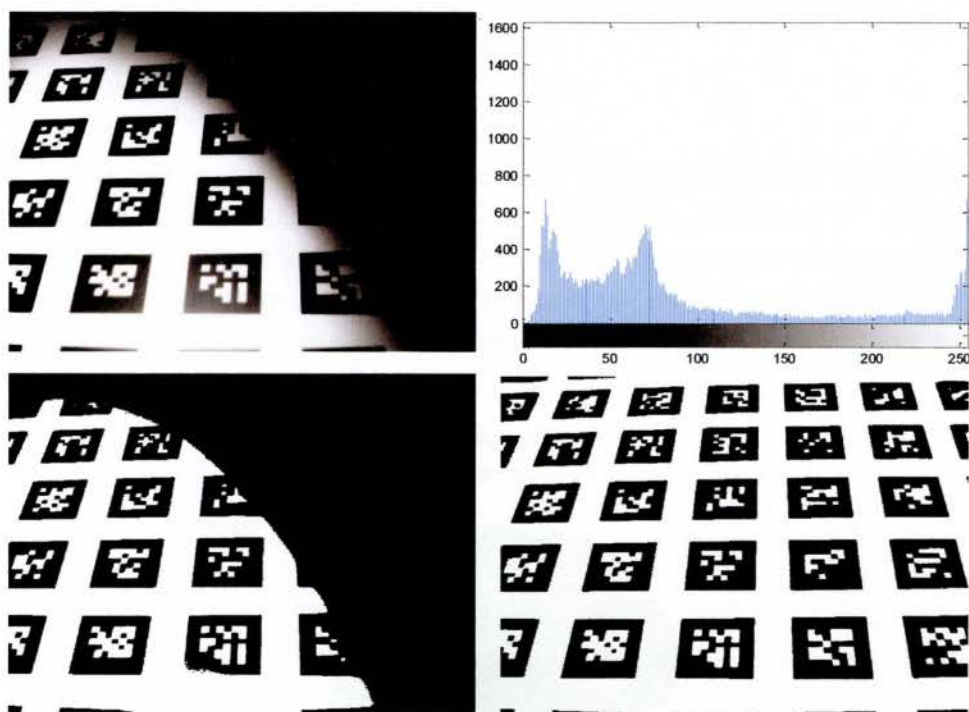
5.2.2 Τοπική Κατωφλίωση (Local Thresholding)

Πολλές φορές δεν είναι δυνατόν να γίνει χρήση της Καθολικής Κατωφλίσωσης με σκοπό να απομονωθεί πλήρως ένα αντικείμενο. Η Καθολική Κατωφλίωση χρησιμοποιείται για την κατάτμηση μιας εικόνας θέτοντας όλα τα pixels των οποίων η τιμή έντασης είναι πάνω από την τιμή κατωφλίου ως προσκήνιο και όλα τα υπόλοιπα pixels, τα οποία έχουν τιμή κάτω από την τιμή κατωφλίου, ως υπόβαθρο.

Σε περιπτώσεις με σκιές φωτισμού ή αυξομειώσεις της έντασης του φωτός η Καθολική Κατωφλίωση αντιμετωπίζει προβλήματα καθώς θεωρεί σταθερή την φωτεινότητα της ψηφιακής εικόνας. Την λύση σε αυτό το πρόβλημα δίνει μια πιο εξελιγμένη μέθοδος κατωφλίωσης, η Τοπική Κατωφλίωση (*Local Thresholding*) ή αλλιώς Προσαρμοσμένη Κατωφλίωση (*Adaptive Thresholding*), η οποία ανταπεξέρχεται με πιο αποτελεσματικό τρόπο στις μεταβαλλόμενες συνθήκες φωτισμού.

Η Τοπική Κατωφλίωση χωρίζει την ψηφιακή εικόνα σε επιμέρους εικόνες, όπου εντοπίζει μια τιμή κατωφλίωσης σε κάθε μία από αυτές ξεχωριστά. Η διαδικασία εντοπισμού γίνεται με βάση το εύρος των τιμών έντασης των pixels κάθε επιμέρους εικόνας.

Στο παρακάτω παράδειγμα (Εικόνα 58) διακρίνεται η αρχική εικόνα Εικόνα 58(α) και το αντίστοιχο ιστόγραμμα της Εικόνα 58(β), στο οποίο δε διακρίνονται ξεχωριστές κορυφές από τις οποίες θα μπορούσαμε να ξεχωρίσουμε τα αντικείμενα από το υπόβαθρο. Η εφαρμογή της Καθολικής Κατωφλίωσης αποτυγχάνει εντελώς Εικόνα 58(γ), για τους λόγους που προαναφέρθηκαν. Αντίθετα η εφαρμογή της Τοπικής Κατωφλίωσης καταφέρνει και μετατρέπει την αρχική εικόνα σε δυαδική αποτελεσματικά.



Εικόνα 58 (α) Αρχική εικόνα. (β) Ιστόγραμμα αρχικής εικόνας. (γ) Καθολική Κατωφλίωση. (δ) Τοπική Κατωφλίωση.

α	β
γ	δ

5.3 Ανίχνευση ακμών

Η ανίχνευση ακμών αποτελεί μια από τις θεμελιώδεις τεχνικές που χρησιμοποιούνται στη ψηφιακή επεξεργασία και την ανάλυση της εικόνας. Δεν υπάρχει ακόμα ακριβής και ευρέως αποδεκτός μαθηματικός ορισμός της ακμής. Παρόλα αυτά σαν ακμή είθισται να ορίζουμε την απότομη αλλαγή στις τιμές φωτεινότητας, δηλαδή το νοητό σύνορο που δημιουργείται μεταξύ δυο περιοχών με διαφορετικό χρωματισμό, φωτισμό, διαφορετικό

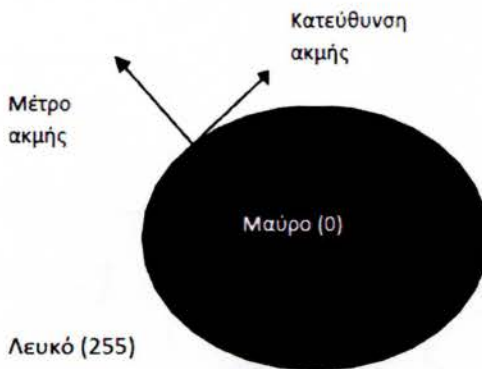
βάθος και υφή. Έτσι είναι κατανοητό πως οι ακμές μπορούν να βρεθούν από τον υπολογισμό της πρώτης παραγωγού όλων των σημείων της εικόνας και τον εντοπισμό των τοπικών μέγιστων

Στη διαδικασία ανίχνευση ακμών μιας ψηφιακής εικόνας μπορεί να παρουσιαστούν αρκετές δυσκολίες. Οι ακμές μπορεί να οφείλονται σε μικρότερες ή μεγαλύτερες μεταβολές στην φωτεινότητα των pixels της εικόνας. Ακόμα και η εμφάνιση θορύβου σε μια εικόνα μπορεί να προκαλέσει αλλοιώσεις στα όρια των αντικειμένων, οδηγώντας σε εσφαλμένη ανίχνευση ακμών. Επίσης σε πολλές περιπτώσεις η ύπαρξη σκιών ή και οι διαφορετικές τιμές φωτεινότητας μπορούν να ανιχνευτούν σαν ψευδοακμές ενώ δεν αντιστοιχούν σε πραγματικές ακμές.

Στην ιδανική περίπτωση, το αποτέλεσμα της εφαρμογής ανίχνευσης ακμών σε μια ψηφιακή εικόνα θα οδηγήσει σε ένα σύνολο συνδεδεμένων λεπτών γραμμών που δείχνουν τα όρια των αντικειμένων. Έτσι εφαρμόζοντας έναν αλγόριθμο ανίχνευσης ακμών σε μια εικόνα μειώνεται σημαντικά ο όγκος των δεδομένων προς, διατηρώντας παράλληλα τις σημαντικές διαρθρωτικές ιδιότητες της εικόνας. Αν το βήμα της ανίχνευσης ακμών είναι επιτυχή, η μετέπειτα επεξεργασία της αρχικής εικόνας μπορεί να απλοποιηθεί σημαντικά.

5.3.1 Ανάλυση Ακμών

Υπάρχει μεγάλη ποικιλία αλγορίθμων που αφορούν στην επίλυση αυτού του προβλήματος, όμως όλοι βασίζονται στην έννοια της κλίσης της συνάρτησης φωτεινότητας $I(k,j)$ στη θέση (k,j) ενός εικονοστοιχείου της εικόνας. Μια ακμή είναι ιδιότητα του κάθε εικονοστοιχείου ξεχωριστά και υπολογίζεται από την συμπεριφορά της συνάρτησης της εικόνας σε μια περιοχή γειτονικών εικονοστοιχείων. Πρόκειται για διανυσματική μεταβλητή με μέτρο και κατεύθυνση.



Εικόνα 59 Κατεύθυνση και μέτρο ακμής

Το μέτρο της ακμής παρουσιάζει πόσο απότομη είναι η μεταβολή της συνάρτησης φωτεινότητας (ισχυρή, αδύναμη ακμή) και η κατεύθυνση μας δίνει τον προσανατολισμό της ακμής στην εικόνα.

Το μέτρο της ακμής υπολογίζεται από τη σχέση:

$$|\bar{\nabla}f(x, y)| = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2}$$

Ενώ για τον υπολογισμό της κατεύθυνσης έχουμε:

$$\varphi = \arg\left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right)$$

Για τον εντοπισμό των ακμών σε μία ψηφιακή εικόνα έχουν αναπτυχθεί αρκετές μέθοδοι, οι οποίες υπολογίζουν διακριτές προσεγγίσεις των μερικών παραγώγων κατά κατεύθυνση για την υπό επεξεργασία εικόνα. Αυτό γίνεται με την συνέλιξη της εικόνας και ενός μικρού μητρώου που στόχο έχει να ενισχύσει την ένταση των ακμών. Ορισμένες από τους αυτές τις μεθόδους θα αναλυθούν στις παρακάτω ενότητες.

5.3.2 Τελεστές Roberts

Οι τελεστές (μητρώα) Roberts ήταν οι πρώτοι που προτάθηκαν για τον υπολογισμό της πρώτης παραγώγου της συνάρτησης φωτεινότητας μιας ψηφιακής εικόνας. Το κάθε μητρώο αποτελείται από ένα πίνακα 2x2, όπου ο κάθε πίνακας προέρχεται περιστρέφοντας τον άλλον κατά 90°.

+1	0
0	-1

G_x

0	+1
-1	0

G_y

Τα μητρώα αυτά έχουν σχεδιαστεί για να λειτουργούν το καθένα ξεχωριστά σε κάθε κλίση, κάθετα (G_y) και οριζόντια (G_x). Πιο συγκεκριμένα καθένα από τα δύο μητρώα μπορεί να εφαρμοστεί στην αρχική εικόνα, για να παράγει ξεχωριστά αποτελέσματα για κάθε κατεύθυνση. Αυτά μπορούν στη συνέχεια να συνδυαστούν μαζί και να προκύψει το

απόλυτο μέτρο της συνολικής κλίσης σε κάθε σημείο της αρχικής εικόνας. Το απόλυτο μέτρο δίνεται από τη σχέση:

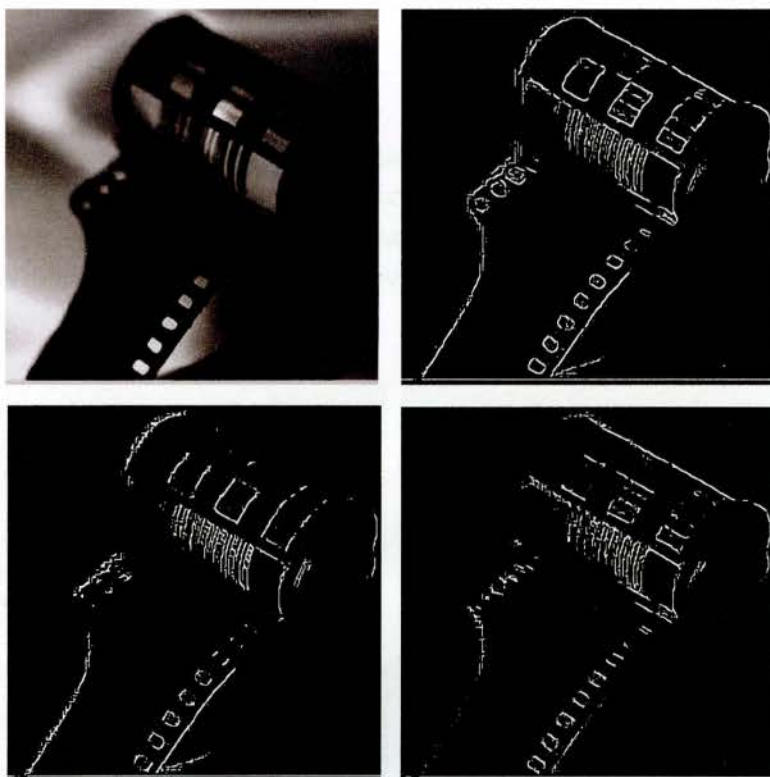
$$|G| = \sqrt{G_x^2 + G_y^2}$$

Ωστόσο το απόλυτο μέτρο υπολογίζεται, για λόγους ταχύτητας, κατά προσέγγιση από τη σχέση :

$$|G| = |G_x^2| + |G_y^2|$$

Η γωνία της κατεύθυνσης των ακμών υπολογίζεται από τη σχέση:

$$\theta = \arctan\left(\frac{G_y}{G_x}\right) - \frac{3\pi}{4}$$



α	β
γ	δ

Εικόνα 60 (α) Αρχική εικόνα. (β) Ανίχνευση ακμών με τη χρήση και των δύο τελεστών (G_x, G_y). (γ) Κάθετη ανίχνευση ακμών (G_y). (δ) Οριζόντια ανίχνευση ακμών (G_x).

Η ανίχνευση των ακμών σε μία ψηφιακή εικόνα με τη χρήση των τελεστών Roberts χαρακτηρίζεται από το πλεονέκτημα της ταχύτητας, καθώς τα δύο μητρώα είναι μικρών

διαστάσεων. Ωστόσο στην περίπτωση που οι ακμές δεν είναι αρκετά απότομες η ανίχνευση τους παρουσιάζει δυσκολίες.

5.3.3 Τελεστές Sobel

Οι τελεστές Sobel προσεγγίζουν την μερική παράγωγο πρώτης τάξης κατά κατεύθυνση για μια ψηφιακή εικόνα. Υπάρχουν 8 διαφορετικά μητρώα διαστάσεων 3x3, καθένα από τα οποία προέρχεται περιστρέφοντας το άλλο κατά 45°, με τα οποία μπορούμε να υπολογίσουμε την μερική παράγωγο. Στην περίπτωση που μας ενδιαφέρει να εντοπίσουμε τις ακμές μας αρκούν δύο μητρώα, τα οποία διαφέρουν μεταξύ τους κατά 90°.

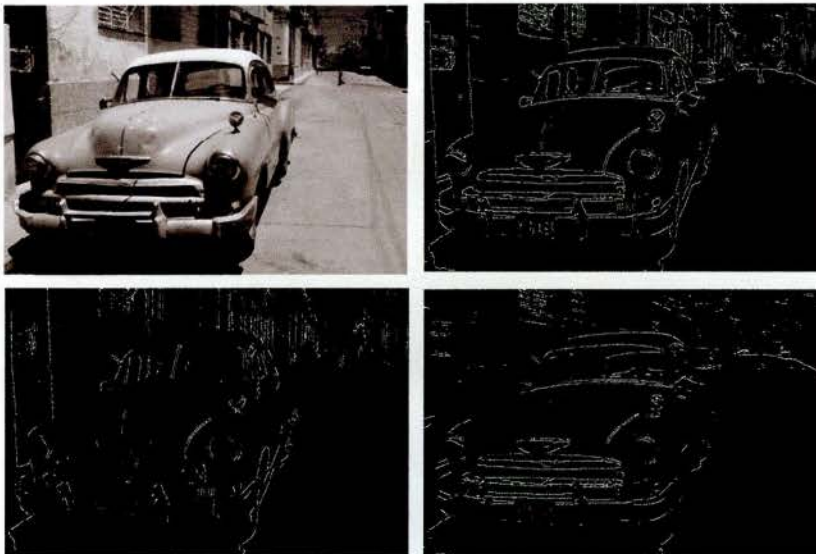
-1	0	+1
-2	0	+2
-1	0	+1

G_x

+1	+2	+1
0	0	0
-1	-2	-1

G_y

Η διαδικασία εντοπισμού των ακμών παραμένει ίδια με αυτή των τελεστών Roberts. Τα αποτελέσματα ωστόσο διαφέρουν καθώς οι τελεστές Sobel είναι μεγαλύτερων διαστάσεων, επομένως πιο αργό στον υπολογισμό τους. Παρόλα αυτά οι τελεστές Sobel μπορούν να ανταποκριθούν καλύτερα στην ανίχνευση ακμών που οι τιμές έντασης τους δεν αλλάζουν απότομα.



α	β
γ	δ

Εικόνα 61 (α) Αρχική εικόνα. (β) Ανίχνευση ακμών με τη χρήση και των δύο τελεστών (G_y, G_x). (γ) Κάθετη ανίχνευση ακμών (G_y). (δ) Οριζόντια ανίχνευση ακμών (G_x).

5.3.4 Τελεστές Prewitt

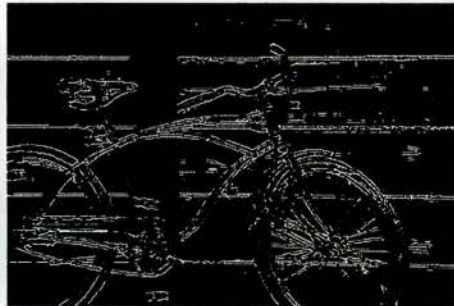
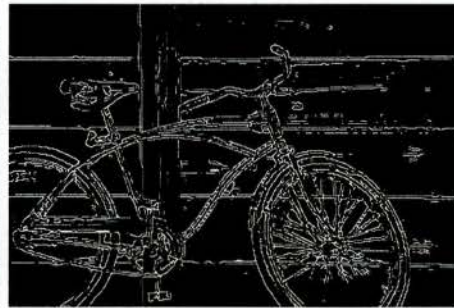
Οι τελεστές Prewitt προσεγγίζουν την μερική παράγωγο πρώτης τάξης κατά κατεύθυνση μιας ψηφιακής εικόνας όπως ακριβώς και οι τελεστές Sobel. Παρόμοια και σε αυτή την περίπτωση υπάρχουν 8 διαφορετικά μητρώα διαστάσεων 3×3 . Η μοναδική διαφορά με τους τελεστές Sobel είναι οι τιμές που περιλαμβάνουν τα μητρώα. Σε κάθε περίπτωση το άθροισμα κάθε μητρώου είναι 0. Παρακάτω στην Εικόνα 62(α) εφαρμόζονται οι τελεστές Prewitt, για την ανίχνευση ακμών.

+1	+1	+1
+1	-2	+1
-1	-1	-1

G_x

+1	+1	-1
+1	-2	-1
+1	+1	-1

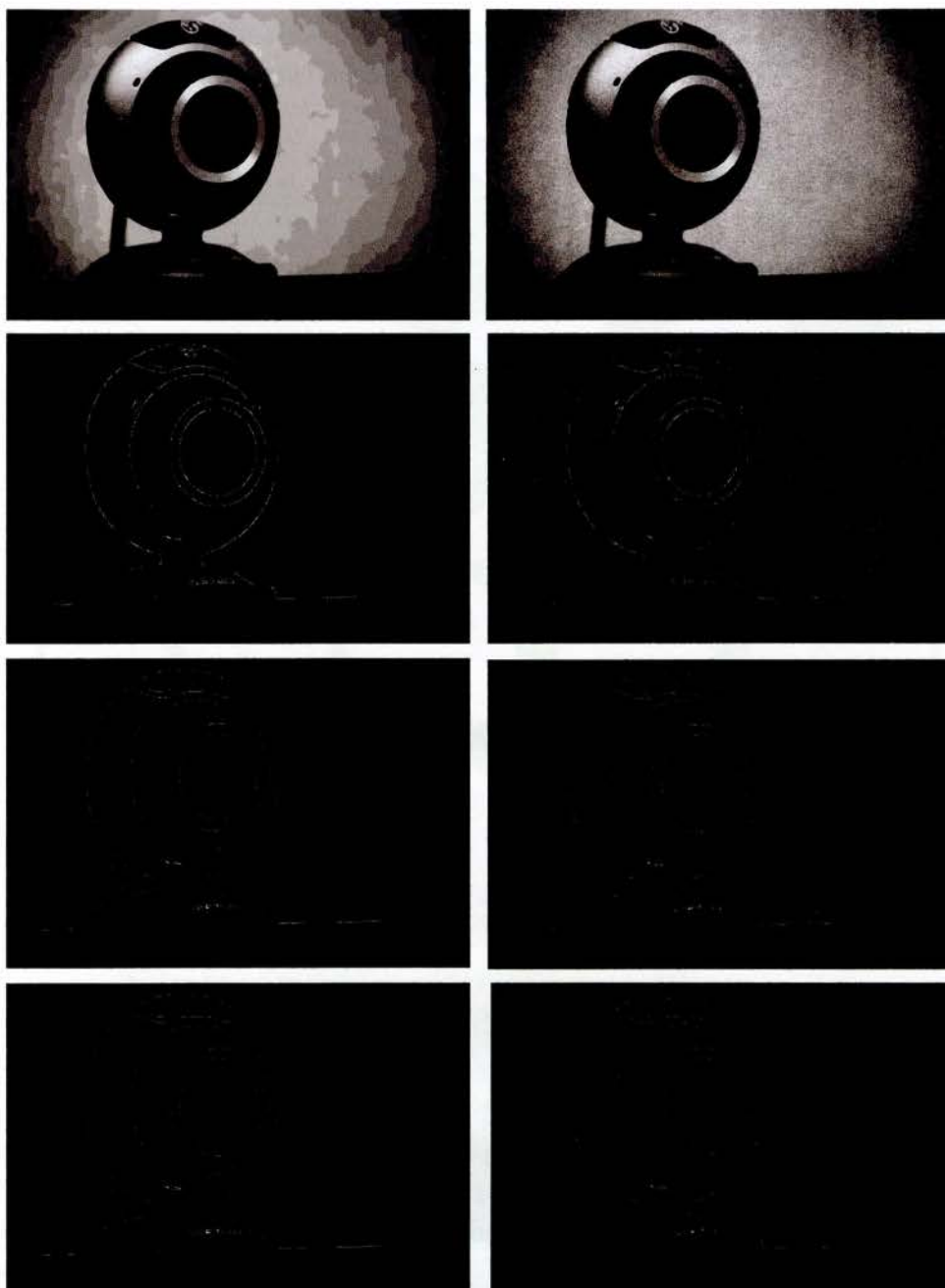
G_y



α	β
γ	δ

Εικόνα 62 (α) Αρχική εικόνα. (β) Ανίχνευση ακμών με τη χρήση και των δύο τελεστών (G_y, G_x). (γ) Κάθετη ανίχνευση ακμών (G_y). (δ) Οριζόντια ανίχνευση ακμών (G_x).

Στην Εικόνα 63(α) γίνεται ανίχνευση ακμών χρησιμοποιώντας και τους τρεις τελεστές που παρουσιάστηκαν παραπάνω. Παρατηρούμε πως όλοι οι τελεστές καταφέρνουν να εντοπίσουν σε ικανοποιητικό βαθμό το σχήμα του αντικειμένου. Αντίθετα στην Εικόνα 63(ε), η οποία έχει προσβληθεί από θόρυβο σημαντικής ισχύος, ο τελεστής Roberts (Εικόνα 63(στ)) αποδεικνύεται επιρρεπής στον θόρυβο και δεν καταφέρει να αποδώσει τις ακμές του αντικειμένου σωστά. Οι τελεστές Sobel και Prewitt παραμένουν αποτελεσματικοί στην αναγνώριση του αντικειμένου.

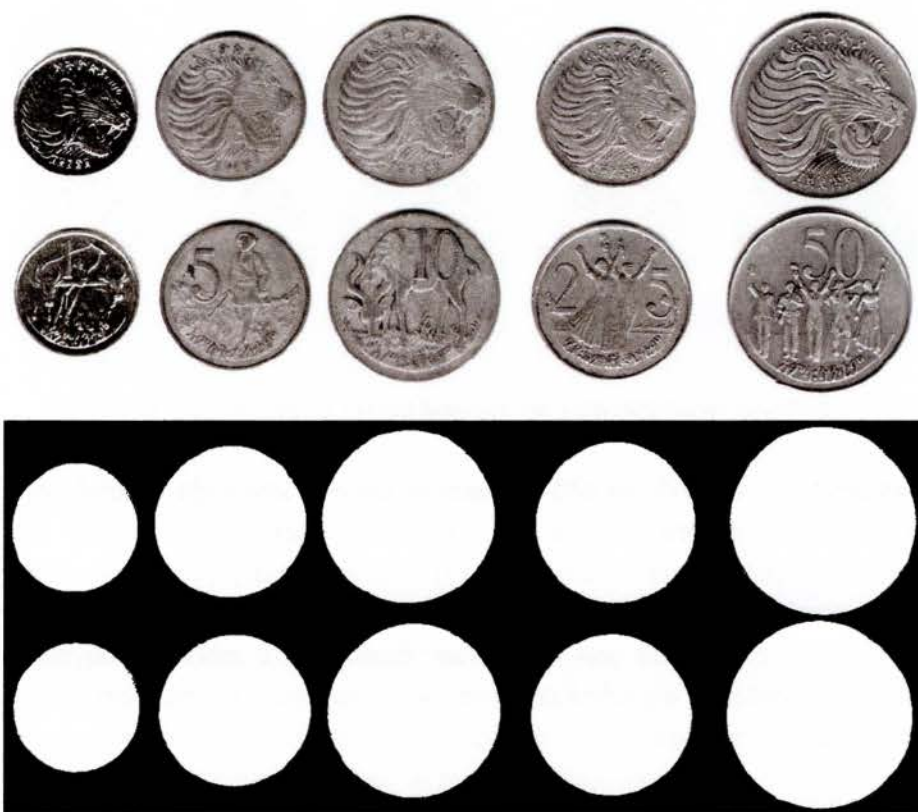


α	ε
β	στ
γ	ζ
δ	η

Εικόνα 63 (α) Αρχική εικόνα. (β) Ανίχνευση ακμών κατά Roberts. (γ) Ανίχνευση ακμών κατά Sobel. (δ) Ανίχνευση ακμών κατά Prewitt. (ε) Αρχική εικόνα με θόρυβο. (στ) Ανίχνευση ακμών κατά Roberts. (ζ) Ανίχνευση ακμών κατά Sobel. (η) Ανίχνευση ακμών κατά Prewitt.

5.4 Εξαγωγή Περιγράμματος

Σε πολλές εφαρμογές της ψηφιακής επεξεργασίας εικόνας μετά από την διαδικασία της ανίχνευσης ακμών σε μια εικόνα είναι απαραίτητο να εντοπιστούν τα περιγράμματα των αντικειμένων που περιλαμβάνει. Εντοπίζοντας το περίγραμμα ενός αντικείμενου μπορούμε να βγάλουμε διάφορα συμπεράσματα, που αφορούν το σχήμα του ή την περίμετρο του. Η εξαγωγή περιγράμματος (*boundary extraction*) αποτελεί σημαντικό θέμα στην ψηφιακή επεξεργασία εικόνας, με αποτέλεσμα να έχουν αναπτυχθεί διάφορες προσεγγίσεις για αυτό το σκοπό. Η βασική ιδέα είναι να καταλήξουμε σε μια επαναληπτική καμπύλη ώστε να προσεγγίσουμε τα όρια του αντικειμένου.



Εικόνα 64 (α) Αρχική εικόνα. (β) Εφαρμογή κατωφλίωσης

Για την διαδικασία εξαγωγής του περιγράμματος σε μια ψηφιακή εικόνα θα πρέπει να έχει προηγηθεί κατωφλίωση στην αρχική εικόνα. Ας πάρουμε για παράδειγμα την Εικόνα 64(α) που απεικονίζει 10 νομίσματα διαφορετικού μεγέθους. Αν θέλουμε να εντοπίσουμε και να εξαγάγουμε το περίγραμμα των νομισμάτων θα πρέπει αρχικά να εφαρμόσουμε κατωφλίωση Εικόνα 64(β), ώστε να διαχωρίσουμε τα pixels που βρίσκονται στο υπόβαθρο από τα υπόλοιπα.



Εικόνα 65 (α) Εντοπισμός περιγράμματος. (β) Εξαγωγή περιγράμματος

Κατά την διαδικασία εξαγωγής του περιγράμματος όλα τα pixels της εικόνας σαρώνονται δύο φορές. Κατά την πρώτη σάρωση τα pixels που ανήκουν στο εσωτερικό των κερμάτων σημειώνονται με μία τιμή διάφορη από 0 και 255. Στην Εικόνα 65(α) για να γίνει πιο αντιληπτός ο εντοπισμός του περιγράμματος έχουν χρησιμοποιηθεί χρώματα. Στην δεύτερη σάρωση πραγματοποιείται ο εντοπισμός των pixels που βρίσκονται στο εσωτερικό των κερμάτων από την τιμή που προηγουμένως τους έχει δοθεί και ταυτόχρονα εξισώνονται με τα pixels του υπόβαθρου, δηλαδή παίρνουν τιμή 0. Επομένως τα μόνα pixels που παραμένουν με τιμή 1 στην εικόνα είναι τα pixels του περιγράμματος.

Κεφάλαιο 6

Παρουσίαση Συστήματος Αυτόματης Αναγνώρισης Πινακίδων

6.1 Εισαγωγικά

Η αυτόματη αναγνώριση των αριθμών κυκλοφορίας αποτελεί ένα αναγκαίο βήμα για τη μετάβαση του πολιτισμού μας στα συστήματα ευφυών μεταφορών. Τα συστήματα ευφυών μεταφορών (Intelligent Transport Systems-ITS) είναι ένας συνδυασμός τεχνολογιών πληροφορικής και επικοινωνιών, εφαρμοσμένων στον τομέα των μεταφορών με στόχο την αποδοτικότερη, ασφαλέστερη και οικονομικότερη, κυκλοφορία των ατόμων ή των εμπορευμάτων, κάνοντας χρήση νέων τεχνολογιών. Σήμερα τα οχήματα παίζουν πολύ μεγάλο ρόλο στις μετακινήσεις των ανθρώπων και το εμπόριο. Επίσης, η χρήση των οχημάτων συνεχώς αυξάνεται, λόγω της αύξησης του πληθυσμού και των ανθρωπίνων αναγκών, τα τελευταία χρόνια. Ως εκ τούτου, ο έλεγχος των οχημάτων είναι ένα μεγάλο πρόβλημα, πολύ δύσκολο στην επίλυση του. Τα συστήματα αυτόματης αναγνώρισης αριθμών κυκλοφορίας χρησιμοποιούνται με σκοπό τον αποτελεσματικό έλεγχο.

Η αυτόματη αναγνώριση πινακίδας αποτελεί ουσιαστικά μια μορφή αυτόματης αναγνώρισης του οχήματος. Είναι μια εφαρμογή επεξεργασίας εικόνας που χρησιμοποιείται για τον προσδιορισμό των οχημάτων μόνο από τις πινακίδες τους. Τα συστήματα αυτόματης αναγνώρισης πινακίδων πραγματικού χρόνου διαδραματίζουν σημαντικό στην αυτόματη παρακολούθηση των κανόνων κυκλοφορίας και της επιβολής του νόμου σε δημόσιους δρόμους. Δεδομένου ότι κάθε όχημα φέρει μία μοναδική πινακίδα, είναι δυνατή η αναγνώριση του χωρίς να χρειάζεται να φέρει εξωτερικές κάρτες, ετικέτες ή πομπούς που να είναι αναγνωρίσιμοι. Λόγω αυτού έχουν αναπτυχθεί ποικίλες τεχνικές αναγνώρισης, ενώ συστήματα αναγνώρισης αριθμών κυκλοφορίας χρησιμοποιούνται σε οδούς κυκλοφορίας οχημάτων, σε χώρους ασφάλειας, όπως χώροι στάθμευσης, στον έλεγχο των συνόρων ή ακόμα και στον εντοπισμό κλεμμένων οχημάτων.

Στους χώρους στάθμευσης οι αριθμοί κυκλοφορίας των οχημάτων χρησιμοποιούνται για τον υπολογισμό της διάρκειας παραμονής του οχήματος εντός του χώρου. Όταν ένα όχημα διέρχεται από μια είσοδο του χώρου στάθμευσης ο αριθμός κυκλοφορίας του αναγνωρίζεται και αποθηκεύεται σε μια βάση δεδομένων. Έτσι όταν το όχημα θέλει να εξέλθει του χώρου στάθμευσης ο αριθμός κυκλοφορίας του αναγνωρίζεται ξανά και συγκρίνεται με τους αριθμούς κυκλοφορίας που είναι αποθηκευμένοι στη βάση δεδομένων. Η διαφορά μεταξύ της ώρας εισόδου και εξόδου του οχήματος χρησιμοποιείται για τον υπολογισμό των τελών στάθμευσης του οχήματος. Η αυτόματη αναγνώριση του αριθμού κυκλοφορίας μπορεί να χρησιμοποιηθεί και για τον έλεγχο πρόσβασης. Για παράδειγμα η τεχνολογία αυτή χρησιμοποιείται από πολλές εταιρείες ώστε να επιτρέπουν την πρόσβαση μόνο σε εξουσιοδοτημένο προσωπικό.

Σε πολλά κράτη συστήματα αναγνώρισης αριθμών κυκλοφορίας είναι εγκατεστημένα ανιχνεύοντας και παρακολουθώντας τα σύνορα. Κάθε όχημα μπορεί να καταχωρηθεί σε μια κεντρική βάση δεδομένων και να συγκριθεί με μια μαύρη λίστα για τον εντοπισμό κλεμμένων οχημάτων. Στον έλεγχο κίνησης των οδών κυκλοφορίας τα οχήματα μπορεί να κατευθυνθούν μέσω εναλλακτικών διαδρομών στον προορισμό τους, διευκολύνοντας τον έλεγχο συμφόρησης σε οδούς με έντονη κυκλοφορία κατά τις ώρες αιχμής.

Ο σχεδιασμός ενός συστήματος ANPR είναι πεδίο έρευνας της τεχνητής νοημοσύνης, της τεχνητής όρασης και των νευρωνικών δικτύων. Σκοπός της παρούσας πτυχιακής είναι να σχεδιάσει και να κατασκευάσει ένα σύστημα αυτόματης αναγνώρισης αριθμών κυκλοφορίας, μελετώντας της αλγοριθμικές και μαθηματικές αρχές των συστημάτων μηχανικής όρασης.

6.2 Επισκόπηση Συστήματος

Είναι σημαντικό να φτιαχτεί ένας αλγόριθμος, ο οποίος θα επιστρέφει το αποτέλεσμα της επεξεργασμένης εικόνας όσο το δυνατό γρηγορότερα και έγκυρα, ώστε να μην επιτρέπει χαμηλά επίπεδα αναγνώρισης. Στην επιστήμη των υπολογιστών, η πολυπλοκότητα ενός αλγορίθμου συχνά υπολογίζεται από το πλήθος των πολλαπλασιασμών που πραγματοποιούνται ώσπου να φτάσουμε στο επιθυμητό αποτέλεσμα. Για να διατηρηθεί ο αριθμός των πράξεων σε όσον το δυνατό χαμηλότερο επίπεδο θα πρέπει να είμαστε βέβαιοι ότι μόνο οι απαραίτητοι υπολογισμοί πραγματοποιούνται.

Υπάρχει όμως και ένας άλλος αποτελεσματικός τρόπος για τη μείωση της πολυπλοκότητας: μειώνοντας το μέγεθος και αλλάζοντας τον τύπο της προς ανάλυση εικόνας. Με τη μείωση του μεγέθους της εικόνας το πλήθος των απαιτούμενων υπολογισμών θα μειωθεί σημαντικά. Παράλληλα η εικόνα που ο αλγόριθμος θα επεξεργάζεται θα είναι ασπρόμαυρη (grayscale). Στην περίπτωση που έχουμε έγχρωμη (RGB) εικόνα αυτή θα μετατρέπεται σε ασπρόμαυρη. Έχουμε ήδη αναφέρει πιο πάνω ότι οι έγχρωμες εικόνες αποτελούνται ουσιαστικά από 3 κανάλια, δηλαδή 3 πίνακες, αυτό έχει σαν αποτέλεσμα οποιαδήποτε επεξεργασία γίνεται στην εικόνα να γίνεται σε τουλάχιστον ένα κανάλι, κάτι που οπωσδήποτε αυξάνει την πολυπλοκότητα των πράξεων. Στην Εικόνα 24 φαίνεται ένα παράδειγμα μιας τέτοιας μετατροπής. Η αρχική έγχρωμη εικόνα μετατρέπεται σε ασπρόμαυρη, ώστε να επεξεργάζεται με πιο γρήγορο τρόπο από τον αλγόριθμο.



Εικόνα 66 Μετατροπή έγχρωμης εικόνας σε ασπρόμαυρη

Θα πρέπει εδώ να τονιστεί πως για την κατασκευή του αλγόριθμου χρησιμοποιήσαμε δείγματα ασπρόμαυρων εικόνων που είχαν ληφθεί από αληθινές συνθήκες σε διάφορους χώρους στάθμευσης. Πιο συγκεκριμένα είχαμε στη διάθεση μας συλλογές εικόνων από 6 χώρους στάθμευσης. Ο συνολικός αριθμός των εικόνων που χρησιμοποιήθηκαν είναι περίπου 400. Όλες οι εικόνες δείχνουν τη μπροστινή όψη του οχήματος, λιγότερο ή περισσότερο κοντά, με ή χωρίς κλίση. Σε καμία περίπτωση δεν μπήκε κανένα απολύτως κριτήριο στις εικόνες που χρησιμοποιήθηκαν. Μέσα στο δείγμα έχουν γίνει δεκτές εικόνες με πολύ χαμηλό φωτισμό, ακόμα και κάτω από 50 Lux. Εικόνες με πινακίδες παλαιού τύπου, στις οποίες οι αλφαριθμητικοί χαρακτήρες δεν έχουν, θεωρητικά, το επαρκές μέγεθος για να αναγνωριστούν σωστά. Ακόμα και εικόνες που έχουν ληφθεί σε νυχτερινό περιβάλλον και τα φώτα των οχημάτων “τυφλώνουν” το σύστημα λήψης.



Εικόνα 67 Χαρακτηριστικά δείγματα ασπρόμαυρων εικόνων

Όλα τα παραπάνω δείγματα περιέχουν πινακίδες από ελληνικά οχήματα. Αυτό σημαίνει πως ο αριθμός κυκλοφορίας αποτελείται από 3 γράμματα, του Λατινικού αλφάβητου, και 4 ψηφία. Μέσα όμως στα δείγματα χρησιμοποιήσαμε και μία συλλογή με περίπου 120 έγχρωμες εικόνες, με πινακίδες από τη Σλοβακία, όπου χρησιμοποιείται άλλη διάταξη. Οι εικόνες αυτές έχουν τραβηχτεί από ψηφιακή φωτογραφική μηχανή και είναι καλύτερης ανάλυσης, ενώ έχουν τραβηχτεί κάτω από συνθήκες ηλιοφάνειας, χωρίς τη χρήση τεχνητού φωτισμού. Τα οχήματα απεικονίζονται είτε από το μπροστινό μέρος είτε από πίσω.

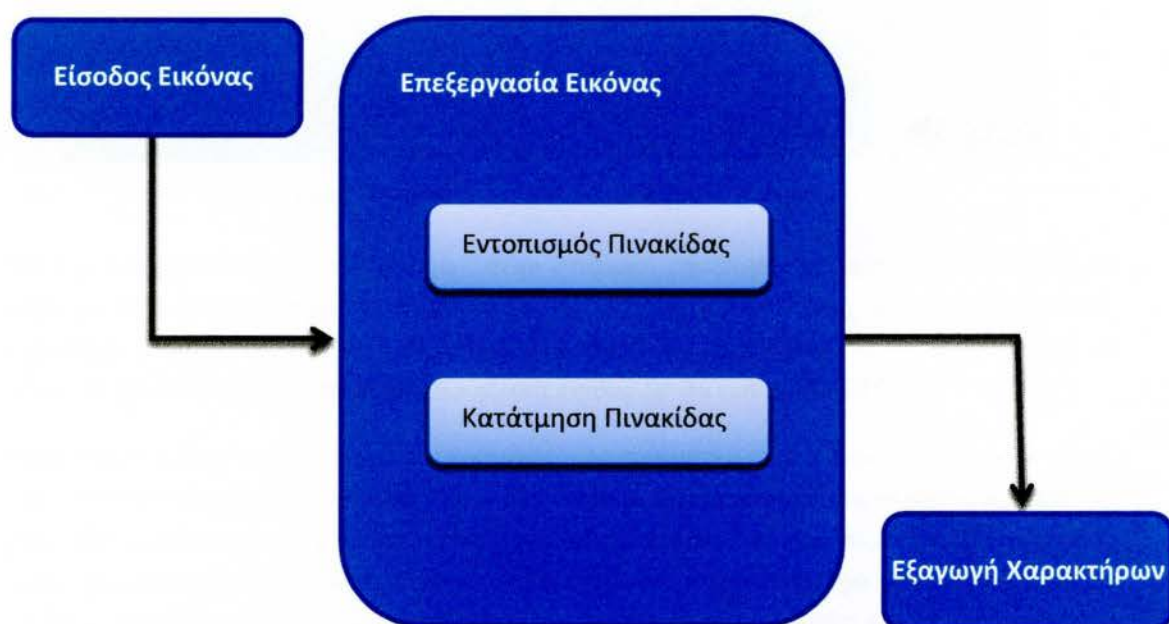


Εικόνα 68 Χαρακτηριστικά δείγματα έγχρωμων εικόνων

6.3 Διαδικασία Επεξεργασίας

Η επεξεργασία της αναγνώρισης υποδιαιρείται σε τρία στάδια, καθένα από τα οποία περιέχει συναφείς δραστηριότητες. Όλη η διαδικασία της αναγνώρισης περιλαμβάνεται στο παρακάτω διάγραμμα. Σαν πρώτο βήμα ορίζεται να εντοπιστούν περιοχές στην εικόνα που πιθανόν να είναι η πινακίδα. Αρκετές από τις περιοχές που θα εντοπιστούν θα αποκλειστούν εξ' αρχής από την πιθανότητα να είναι η πινακίδα καθώς δε θα πληρούν τα κριτήρια του μεγέθους. Δε θα είναι δηλαδή ορθογώνιες ή θα έχουν πολύ μικρό ή μεγάλο σχήμα. Η περιοχή που θα μοιάζει να πληρεί τις προϋποθέσεις που έχουν οριστεί, και ουσιαστικά θα είναι η πινακίδα του οχήματος, θα αποκόπτεται από την υπόλοιπη εικόνα και θα αποστέλλεται για περαιτέρω επεξεργασία.

Όταν η περιοχή που είναι πιθανό να περιέχει την πινακίδα εντοπιστεί, θα πρέπει να καθοριστεί καλύτερα η θέση της πινακίδας. Αφού εντοπίσουμε με μεγαλύτερη ακρίβεια τη θέση της πινακίδας τότε είμαστε σε θέση να διατηρήσουμε μόνο τα αλφαριθμητικά σύμβολα διώχνοντας καθετί περιττό. Στη συνέχεια και αφού έχουμε στην εικόνα μόνο αλφαριθμητικά σύμβολα, τότε μπορούμε να τα περιστρέψουμε προς την κατεύθυνση που χρειάζεται ώστε να είναι σε ευθεία. Αυτό το βήμα θα μας βοηθήσει στην κατάτμηση της πινακίδας και στον διαχωρισμό του κάθε αλφαριθμητικού συμβόλου από τα υπόλοιπα.



Εικόνα 69 Διάγραμμα Συστήματος Αναγνώρισης Πινακίδας

6.4 Εντοπισμός Πινακίδας

Ο συνήθης τρόπος εντοπισμού της πινακίδας είναι η εύρεση, στην εικόνα, περιοχών με έντονο λευκό χρώμα. Η ιδέα πίσω από αυτόν τον τρόπο επίλυσης του προβλήματος είναι ότι η πινακίδα έχει πολύ λευκό χρώμα και παράλληλα διαθέτει και μία αντανακλαστική επιφάνεια, η οποία βοηθάει στην ανίχνευση της.



Εικόνα 70 Χαρακτηριστικά παραδείγματα πινακίδων

Με την φράση “έντονο λευκό χρώμα”, στην ψηφιακή επεξεργασία εικόνας, θεωρούνται τα pixels με τιμή φωτεινότητας που πλησιάζει το 255. Στην περίπτωση που έχουμε εικόνες οχημάτων που έχουν ληφθεί με κακό φωτισμό και άρα έχουν χαμηλή αντίθεση, η μεγαλύτερη τιμή φωτεινότητας σε μία τέτοια εικόνα δε θα ξεπερνά το 160.

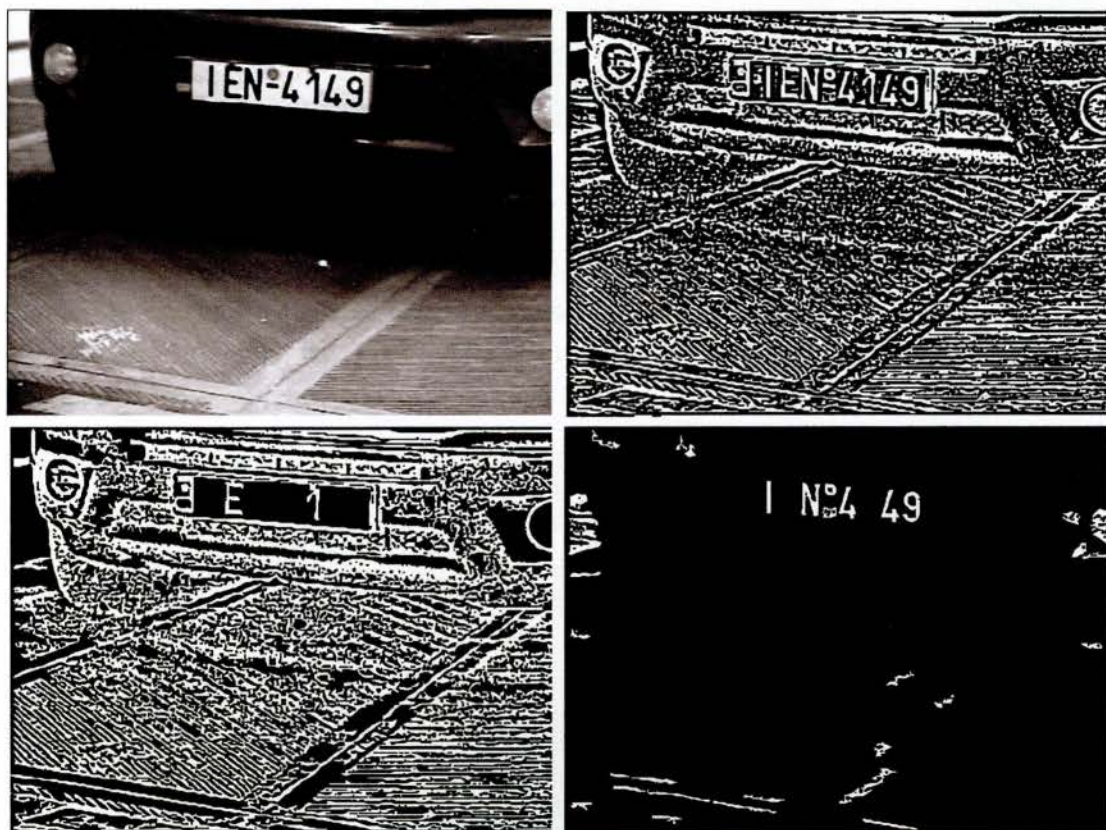
Ένας άλλος λόγος που υστερεί η παραπάνω υλοποίηση είναι ότι παλαιότερες πινακίδες, στην Ελλάδα τουλάχιστον πριν το 2004, δεν διέθεταν αντανακλαστική επιφάνεια, ώστε να είναι πιο ευδιάκριτες όταν πέφτει πάνω τους φως. Στην παρακάτω εικόνα διακρίνονται δύο διαφορετικού τύπου πινακίδες. Ο τύπος (α) χρησιμοποιούταν στη χώρα μας από το 1983 ως το 2004. Αυτές οι πινακίδες δεν διαθέτουν αντανακλαστική μεμβράνη και δεν είναι εύκολα εντοπίσιμες με τον παραπάνω τρόπο. Ο δεύτερος τύπος πινακίδας που παρουσιάζεται, ο πιο καινούργιος, χρησιμοποιείται από το 2004 έως και σήμερα. Τέτοιες πινακίδες είναι πιο εύκολα εντοπίσιμες, αφού διαθέτουν την απαραίτητη μεμβράνη.



Εικόνα 71 (α) Πινακίδα παλαιού τύπου. (β) Σύγχρονος τύπος πινακίδας.

Ένα άλλο μειονέκτημα των πινακίδων παλαιού τύπου, το οποίο δε θα μας απασχολήσει στο παρόν κεφάλαιο, είναι ο τρόπος που παρουσιάζεται το ψηφίο 1. Η πάνω διαγώνια γραμμή είναι αρκετά μικρή και δύσκολα μπορεί να αναγνωριστεί. Ένας άνθρωπος θα μπορούσε να καταλάβει πως πρόκειται για το ψηφίο 1 επειδή βρίσκεται στη δεξιά μεριά της πινακίδας, εκεί που βρίσκονται οι αριθμοί. Αν όμως η πινακίδα που πρόκειται να επεξεργαστούμε δε διαθέτει το πρότυπο: 3 γράμματα αριστερά, 4 αριθμοί δεξιά, τότε είναι δύσκολη η αναγνώριση της. Αυτό έχει σαν αποτέλεσμα πολλές φορές το ψηφίο 1 να αναγνωρίζεται σαν το γράμμα I.

Ο αλγόριθμος που θα παρουσιαστεί στην πτυχιακή εργασία εντοπίζει την πινακίδα με τον εξής τρόπο. Αρχικά η εικόνα των αποχρώσεων του γκρι (grayscale) μετατρέπεται σε δυαδική (binary). Στη συνέχεια από τη δυαδική εικόνα απομακρύνουμε όσα κομμάτια είναι κάτω από 230 pixels (Εικόνα 72(γ)). Αυτομάτως θα απομακρυνθούν και κάποια από τα αλφαριθμητικά σύμβολα τις πινακίδας, αν όχι όλα. Στο παράδειγμα που ακολουθεί δεν απομακρύνονται όλα τα σύμβολα. Παραμένουν το γράμμα E και το ψηφίο 1.



α	β
γ	δ

Εικόνα 72 (α) Αρχική εικόνα. (β) Δυαδική εικόνα. (γ) Αφαίρεση των σχημάτων με μέγεθος μικρότερο από 230 pixels. (δ) Εμφάνιση των σχημάτων με μέγεθος μικρότερο από 230 pixels.

Ο λόγος που παραμένουν είναι οι βίδες που συγκρατούν την πινακίδα. Οι βίδες είναι τοποθετημένες ακριβώς πάνω από τα δύο σύμβολα. Αυτό έχει σαν αποτέλεσμα να μην θεωρούνται σαν ανεξάρτητα κομμάτια με εμβαδό μικρότερο από 230 pixels, αλλά μέρος ενός μεγαλύτερου κομματιού. Για να φτάσουμε στην Εικόνα 72 (δ) θα πρέπει να αφαιρέσουμε τις δύο προηγούμενες εικόνες, την (β) και την (γ). Το αποτέλεσμα της

αφαίρεσης αυτής είναι η Εικόνα 30 (δ), η οποία περιέχει όλα τα κομμάτια της δυαδικής εικόνας με κάτω από 230 pixels μέγεθος.

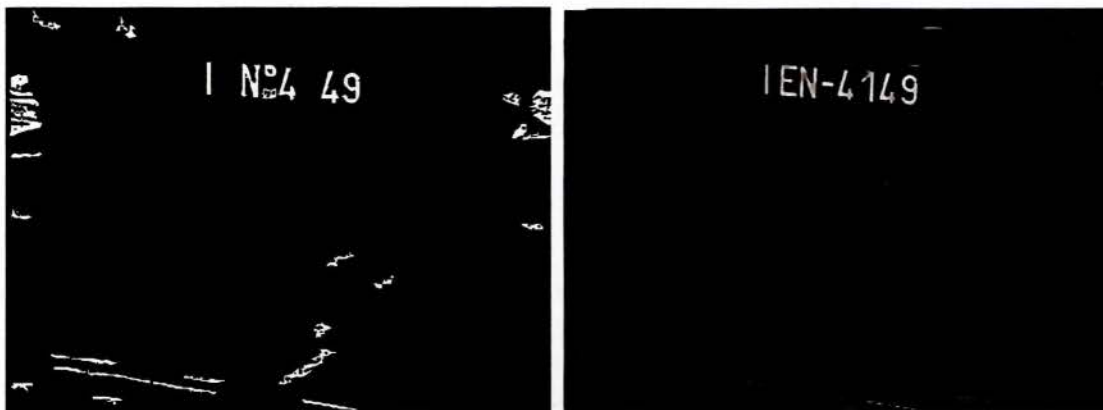
Στο επόμενο βήμα εφαρμόζουμε στην αρχική εικόνα φίλτρο bottom hat. Το φίλτρο αυτό έχει την ικανότητα να τονίζει τα σκούρα τμήματα μιας εικόνας που βρίσκονται πάνω σε πιο λευκό φόντο. Το αποτέλεσμα του φίλτρου φαίνεται στην Εικόνα 73 (β). Παρατηρούμε πόσο πολύ έχουν τονισθεί τα αλφαριθμητικά σύμβολα της πινακίδας σε σχέση με την υπόλοιπη εικόνα.



α β

Εικόνα 73 (α) Αρχική εικόνα. (β) Εφαρμογή φίλτρου bottom hat.

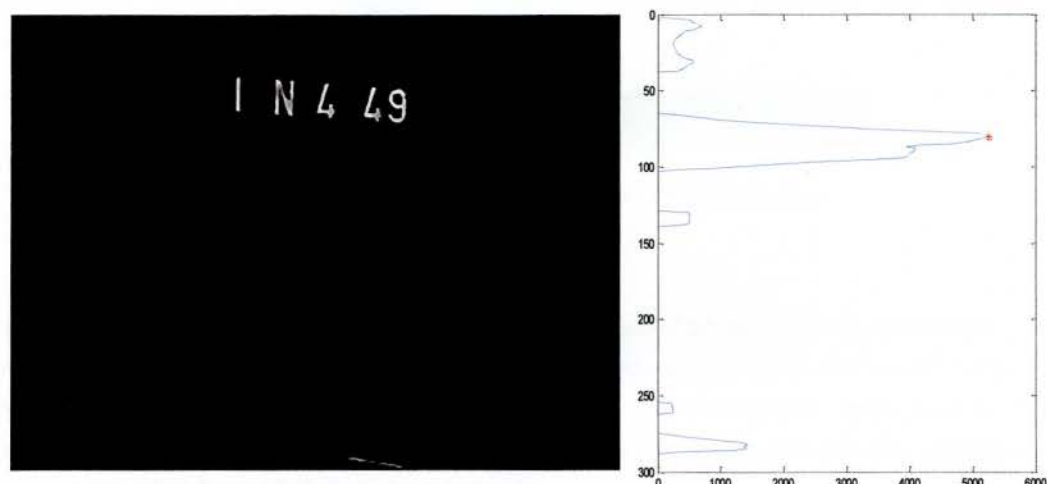
Έχοντας ολοκληρώσει τα δύο παραπάνω βήματα έχουμε καταλήξει στις παρακάτω δύο εικόνες. Σε μια δυαδική εικόνα που περιέχει όλα τα κομμάτια της αρχικής εικόνας με κάτω από 230 pixels μέγεθος (Εικόνα 72 (δ)) και σε μία εικόνα που έχει προκύψει ύστερα από την εφαρμογή φίλτρου bottom hat (Εικόνα 73 (β)).



α β

Εικόνα 74 (α) Σχημάτα με μέγεθος μικρότερο από 230 pixels.
(β) Εφαρμογή φίλτρου bottom hat.

Η δυαδική εικόνα αποτελείται από 0 και 1. Τα μαύρα pixels αντιστοιχούν στο 0 και τα λευκά στο 1. Πολλαπλασιάζοντας αυτές τις δύο εικόνες θα πάρουμε σαν αποτέλεσμα την Εικόνα 75(α). Ο πολλαπλασιασμός των δύο εικόνων γίνεται κατά σημείο. Έτσι τα σημεία που η δυαδική εικόνα έχει τιμή 1 θα παραμείνουν εμφανή, ενώ τα υπόλοιπα θα μηδενιστούν.



α β

Εικόνα 75 (α) Τελική εικόνα. (β) Απεικόνιση κάθετης προβολής.

Μετά και το παραπάνω στάδιο είμαστε σε θέση να βρούμε την κάθετη προβολή (vertical projection) της τελικής εικόνας. Η κάθετη προβολή είναι το γράφημα που διακρίνεται παραπάνω (Εικόνα 75 (β)) και δίνει το άθροισμα καθεμιάς γραμμής της εικόνας. Στο σημείο που το γράφημα παρουσιάζει τη μέγιστη τιμή του βρίσκεται και η πινακίδα. Σε αντίθεση με την κάθετη προβολή υπάρχει και η οριζόντια προβολή (horizontal projection), η οποία μας δίνει το άθροισμα καθεμιάς στήλης της εικόνας.

Πιο συγκεκριμένα ας θεωρήσουμε την διακριτή συνάρτηση $f(x,y)$. Τότε η κάθετη προβολή p_y , της συνάρτησης f είναι το άθροισμα όλων των pixels για καθεμιά από τις y_{th} γραμμές της εικόνας. Με τον ίδιο τρόπο οριζόντια προβολή p_x , της συνάρτησης f είναι το άθροισμα όλων των pixels για καθεμιά από τις x_{th} στήλες της εικόνας.

Μαθηματικά η κάθετη και η οριζόντια προβολή ορίζονται από τις παρακάτω σχέσεις:

$$p_x(x) = \sum_{j=0}^{h-1} f(x, j) \qquad p_y(y) = \sum_{i=0}^{w-1} f(i, y)$$

όπου w και h είναι οι διαστάσεις της εικόνας.

Με την εφαρμογή και της κάθετης προβολής η νέα εικόνα που θα χρησιμοποιήσουμε για να εντοπίσουμε την πινακίδα θα είναι η οριζόντια λωρίδα που ακολουθεί.



Εικόνα 76 Οριζόντιο εντοπισμός πινακίδας.

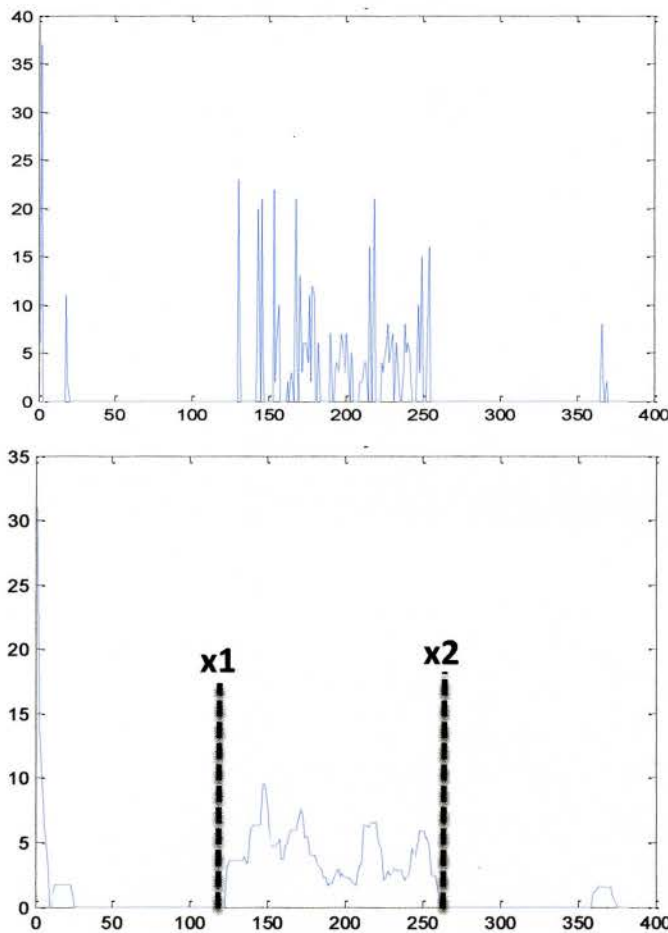
Για να βρούμε τις δύο πλευρές της πινακίδας και να προχωρήσουμε στον πλήρη εντοπισμό της θα χρησιμοποιήσουμε την οριζόντια προβολή που αναφέραμε παραπάνω. Πριν την χρησιμοποίηση της οριζόντιας προβολής θα πρέπει να εφαρμόσουμε κάποιο φίλτρο ανίχνευσης ακμών. Πιο συγκεκριμένα χρησιμοποιούμε ανίχνευση ακμών κατά κάθετη

κατεύθυνση, με τη χρήση τελεστών Sobel. Το αποτέλεσμα φαίνεται στην εικόνα που ακολουθεί.



Εικόνα 77 Ανίχνευση ακμών κατά Sobel με χρήση μήτρας κάθετης κλίσης.

Παρατηρούμε από την Εικόνα 77 ότι η θέση της πινακίδας διακρίνεται ξεκάθαρα. Για να καταφέρουμε να εντοπίσουμε τα ακριβή όρια της και να την απομονώσουμε πλήρως θα υπολογίσουμε την οριζόντια προβολή, η οποία παρουσιάζεται στην Εικόνα 78(α). Εξομαλύνοντας το γράφημα της οριζόντιας προβολής θα προκύψει ένα νέο γράφημα, στο οποίο τα σημεία x_1 και x_2 θα αποτελούν τα όρια της πινακίδας.



α
β

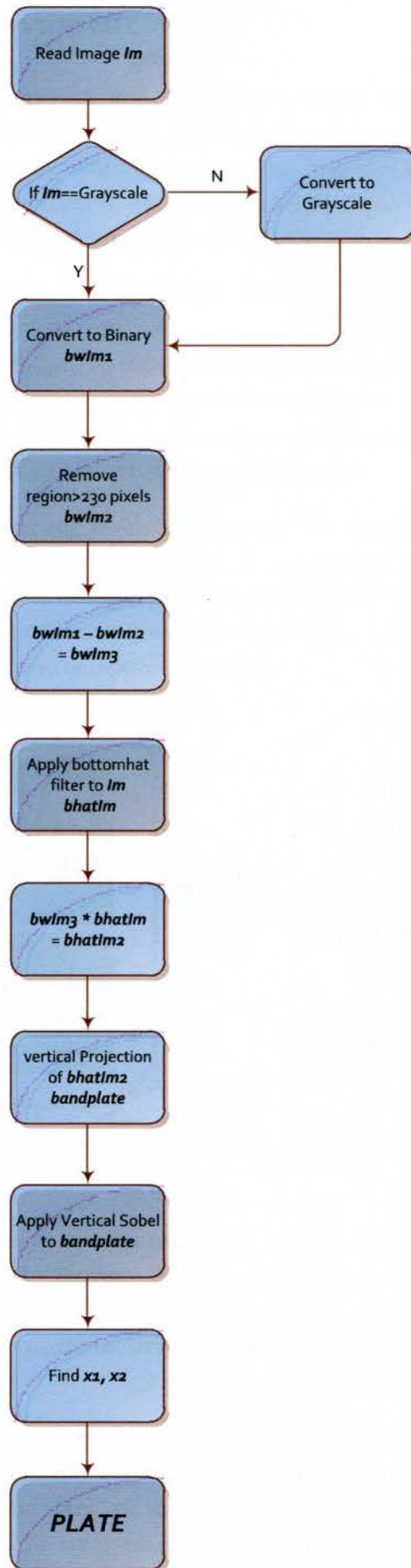
Εικόνα 78 (α) Οριζόντια προβολή Εικόνας 77. (β) Εξομαλύνση οριζόντιας προβολής.

Υπολογίζοντας τις τιμές x_1 και x_2 θα έχουμε βρει επακριβώς και τα πλάγια όρια της πινακίδας.



Εικόνα 79 Εντοπισμός πινακίδας.

Το πρώτο στάδιο, ο εντοπισμός της πινακίδας ολοκληρώθηκε. Τα βήματα που ακολουθήσαμε φαίνονται στο παρακάτω διάγραμμα ροής.

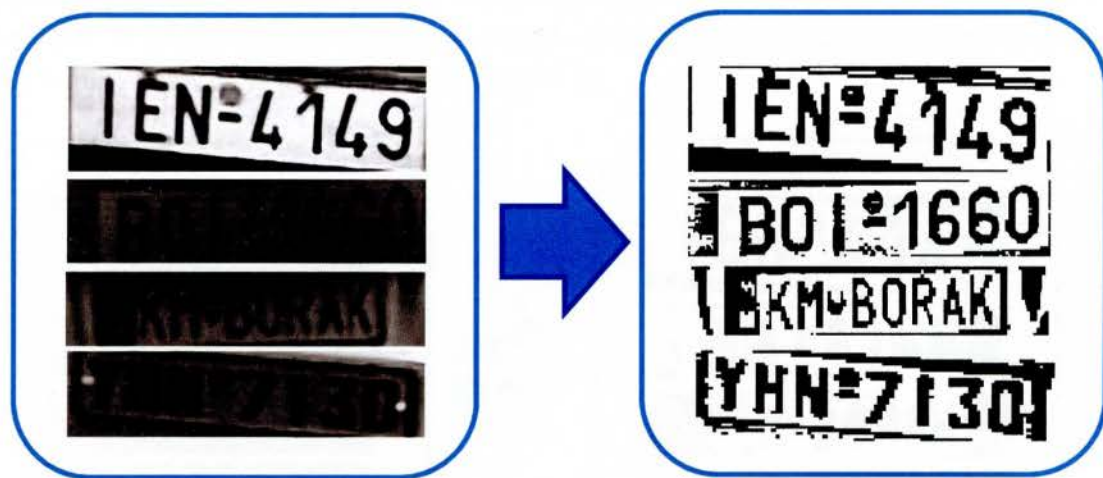


6.5 Κατάτμηση Πινακίδας

Στην ενότητα αυτή θα παρουσιαστεί το δεύτερο τμήμα του συστήματος αυτόματης αναγνώρισης πινακίδων, η κατάτμηση της πινακίδας του οχήματος. Παραπάνω έγινε αναφορά στην διαδικασία εντοπισμού της πινακίδας, η οποία αποτελεί το πρώτο στάδιο του αλγόριθμου που θα αναπτυχθεί. Αφού έχουμε εξασφαλίσει τον επιτυχή εντοπισμό της πινακίδας μπορούμε να προχωρήσουμε στην κατάτμηση της.

Υπάρχουν πολλοί παράγοντες που καταστούν την κατάτμηση της πινακίδας μια αρκετά δύσκολη διαδικασία, όπως ο θόρυβος της εικόνας, το πλαίσιο της πινακίδας, οι βίδες που τη συγκρατούν πάνω στο όχημα, το αναγνωριστικό σήμα της χώρας κυκλοφορίας, η κλίση που μπορεί να έχει γίνει η λήψη ή ακόμα και διακυμάνσεις στο φωτισμό. Για την εξάλειψη των προβλημάτων που προκαλούν οι παραπάνω παράγοντες έχει δημιουργηθεί ένας αρκετά ισχυρός αλγόριθμος με σημαντικά ποσοστά επιτυχίας.

Το πρώτο βήμα κατάτμησης της πινακίδας αποτελεί η μετατροπή της ψηφιακής εικόνας που περιέχει την πινακίδα σε δυαδική. Η μετατροπή γίνεται χρησιμοποιώντας προσαρμοστική κατωφλίωση (*Adaptive Thresholding*), έτσι ώστε να διαχωριστούν με μεγαλύτερη αποτελεσματικότητα τα pixels του υπόβαθρου από τα υπόλοιπα ακόμα αν η εικόνα δεν χαρακτηρίζεται από ομοιόμορφο φωτισμό.

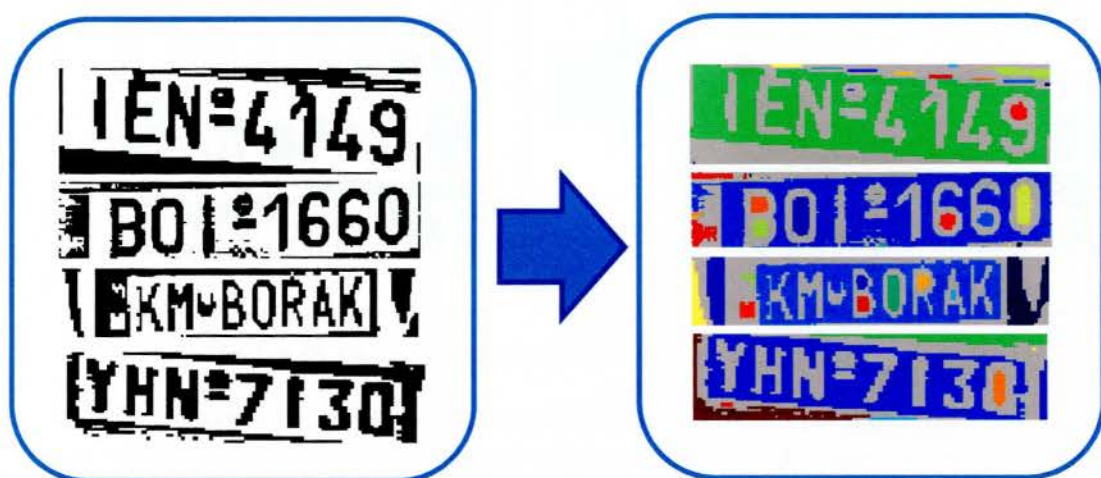


Εικόνα 80 Μετατροπή πινακίδας σε δυαδική.

Έχοντας πραγματοποιήσει το παραπάνω βήμα θα έχουμε μια δυαδική εικόνα που αποτελείται από 0 και 1. Ωστόσο η κατάτμηση των χαρακτήρων δεν μπορεί να πραγματοποιηθεί εύκολα, καθώς η δυαδική εικόνα περιλαμβάνει, πέρα από τα αλφαριθμητικά σύμβολα, και αντικείμενα τα οποία δυσκολεύουν την κατάτμηση και θα πρέπει να απομακρυνθούν.

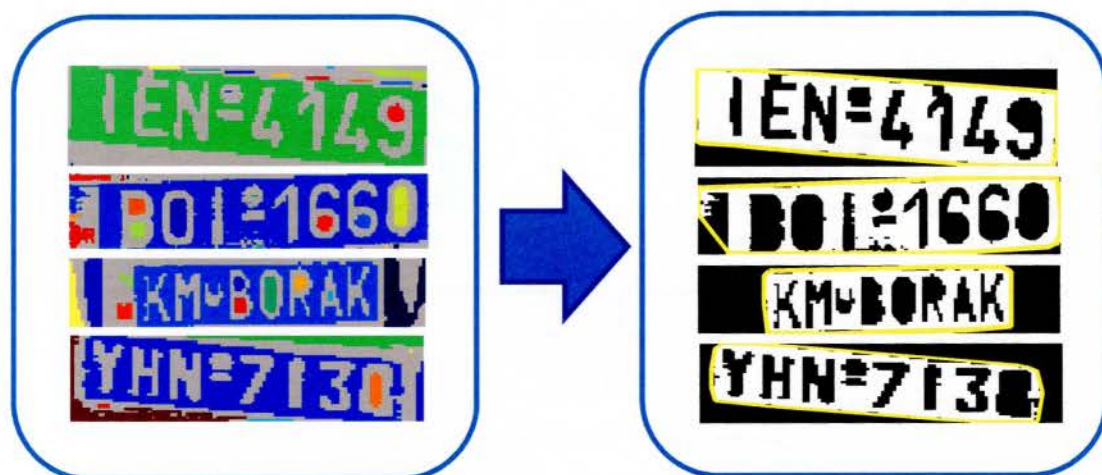
Για την απομάκρυνση των περιττών αντικειμένων αρχικά θα πρέπει να απομονώσουμε αποκλειστικά την πινακίδα διατηρώντας μόνο αυτή. Παρατηρούμε στην (Εικόνα 80) ότι το μεγαλύτερο αντικείμενο σε όλες τις δυαδικές εικόνες είναι η πινακίδα. Για να την απομονώσουμε θα πρέπει να κάνουμε χρήση της μεθόδου των Συνδεδεμένων Στοιχείων

(*Connected Components*), η οποία παρουσιάστηκε στην ενότητα 4.5.4.. Με τη μέθοδο αυτή μπορούμε να ξεχωρίσουμε και να ονοματίσουμε το κάθε αντικείμενο ξεχωριστά.



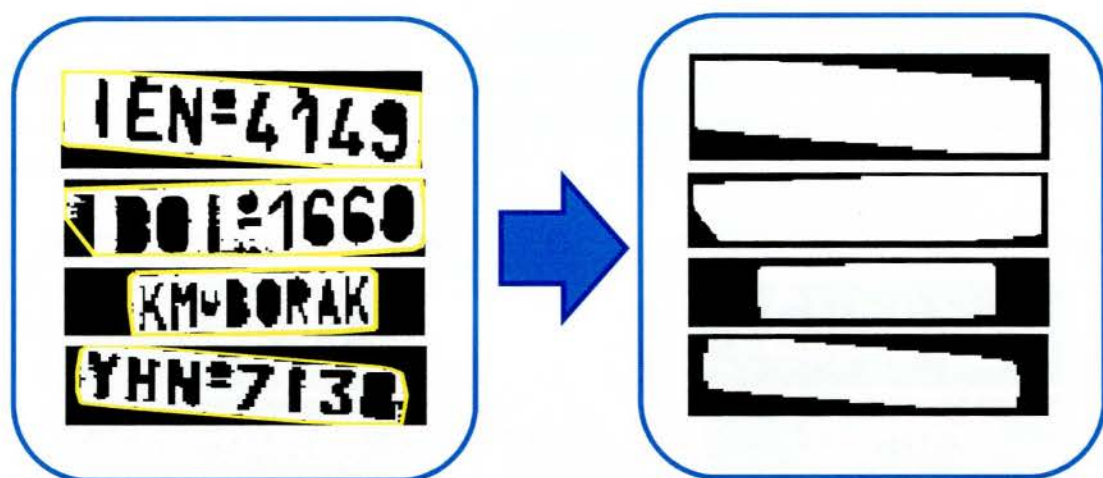
Εικόνα 81 Εύρεση των Συνδεδεμένων Στοιχείων από τη δυαδική εικόνα.

Για να γίνει πιο ευδιάκριτος ο διαχωρισμός των αντικειμένων της εικόνας, μετά τη χρήση της μεθόδου των συνδεδεμένων στοιχείων, κάθε αντικείμενο παρουσιάζεται με ένα ξεχωριστό χρώμα. Από την (Εικόνα 81) γίνεται αντιληπτό πιο καθαρά ότι η πινακίδα αποτελεί το μεγαλύτερο συνδεδεμένο στοιχείο. Αυτό που μένει είναι να απομονώσουμε την πινακίδα από τα περιττά αντικείμενα Εικόνα 82.



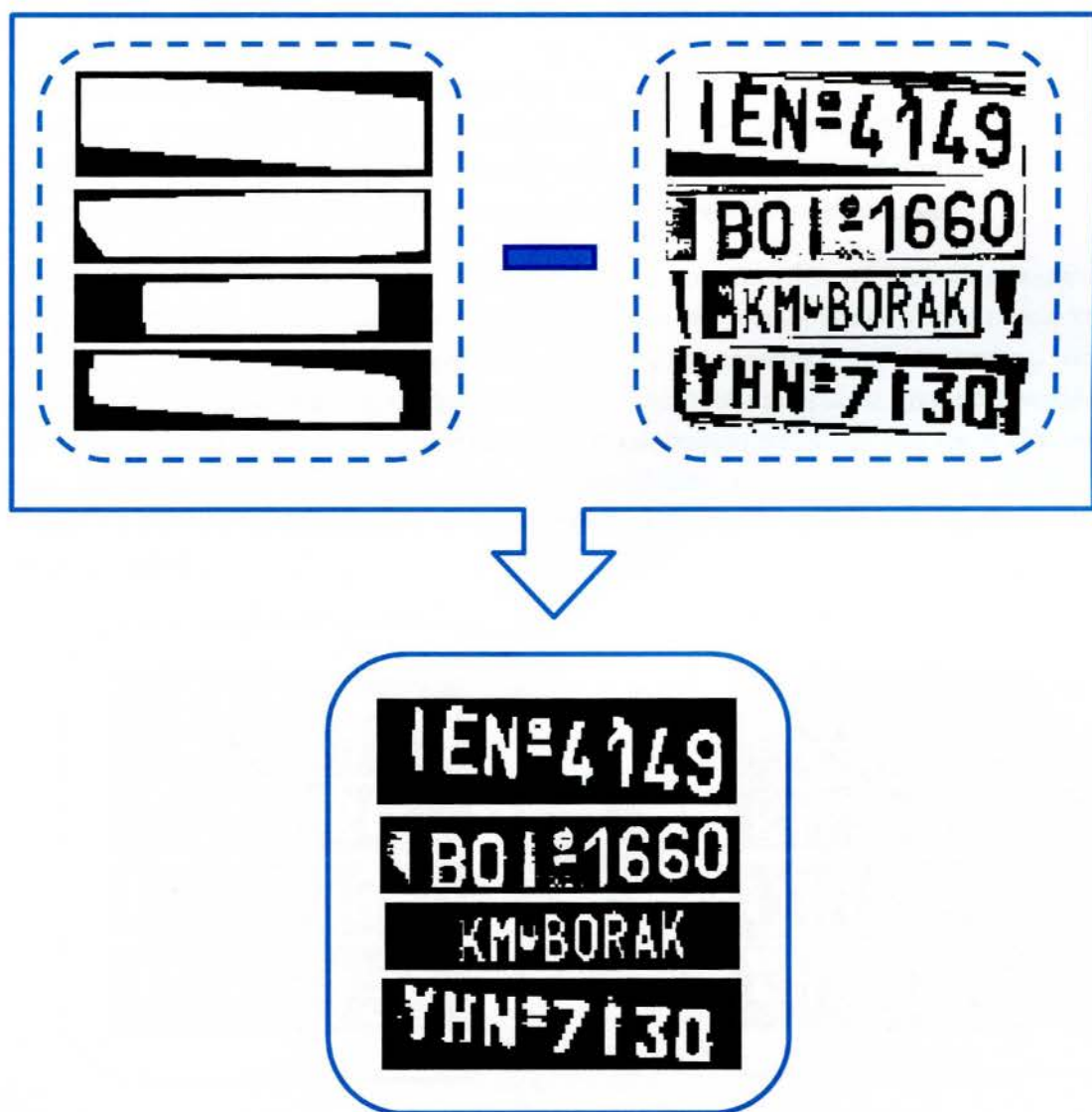
Εικόνα 82 Απομόνωση του μεγαλύτερου συνδεδεμένου στοιχείου και εύρεση του περιγράμματος του.

Στη συνέχεια θα πρέπει να βρεθεί το περίγραμμα της πινακίδας και να γίνει χρήση της μεθόδου Συμπλήρωσης Οπών (*Filling Holes*), η οποία παρουσιάστηκε στην ενότητα 4.5.3. Στην Εικόνα 82 εκτός από το μεγαλύτερο συνδεδεμένο στοιχείο παρουσιάζεται και ο εντοπισμός του περιγράμματος του.



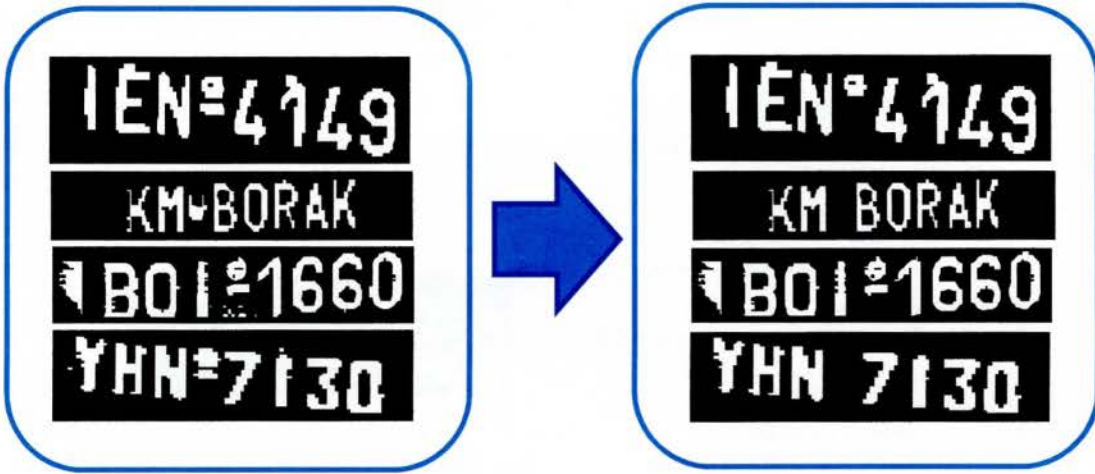
Εικόνα 83 Εφαρμογή Συμπλήρωσης Οπών στο περίγραμμα.

Με την εφαρμογή της μεθόδου Συμπλήρωσης Οπών όλο το εσωτερικό του περιγράμματος, που εμπεριέχει και τα αλφαριθμητικά σύμβολα, θα αποτελεί ένα ενιαίο αντικείμενο, το οποίο θα είναι και το μοναδικό που υπάρχει στην εικόνα (Εικόνα 83). Ουσιαστικά το αντικείμενο αυτό αποτελεί σχηματικά την ίδια την πινακίδα.



Εικόνα 84 Εμφάνιση αλφαριθμητικών συμβόλων πινακίδας.

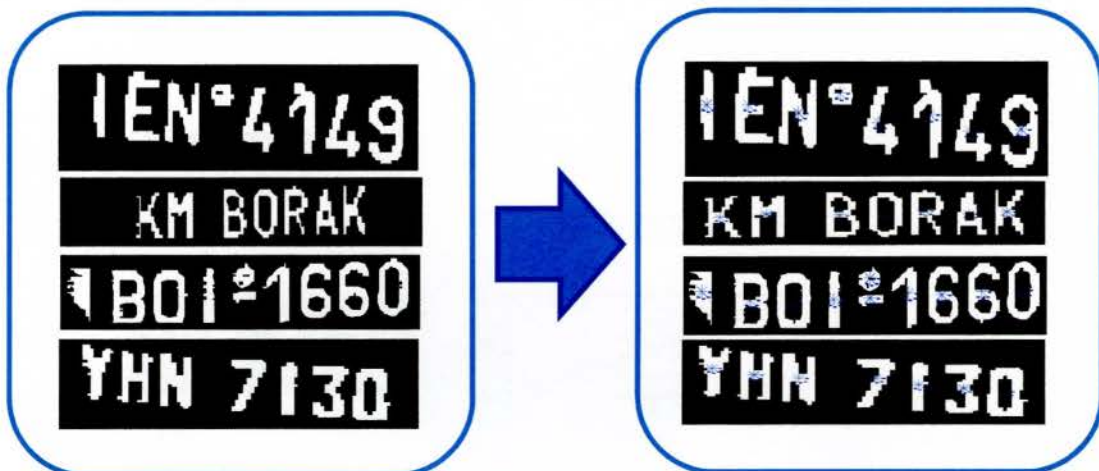
Πλέον απομένει να εμφανίσουμε τα αλφαριθμητικά σύμβολα της πινακίδας. Αυτό γίνεται αφαιρώντας την αρχική δυαδική εικόνα από το ενιαίο αντικείμενο που δημιουργήσαμε με τη συμπλήρωση σπών. Η παραγόμενη εικόνα θα περιέχει τα αλφαριθμητικά σύμβολα της πινακίδας (Εικόνα 84).



Εικόνα 85 Αφαίρεση μικρών αντικειμένων.

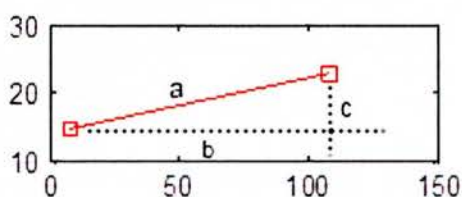
Ένα περαιτέρω βήμα μετά την εμφάνιση των αλφαριθμητικών χαρακτήρων είναι η αφαίρεση αντικειμένων με μέγεθος κάτω από 35 pixels (Εικόνα 85). Το βήμα αυτό χρειάζεται ιδιαίτερη προσοχή καθώς αν μία εικόνα έχει ληφθεί σε απόσταση μεγαλύτερη των 8 μέτρων, υπάρχει περίπτωση να απομακρύνουμε λανθασμένα τα σύμβολα.

Έχοντας πλέον εμφανίσει τα αλφαριθμητικά σύμβολα της πινακίδας και έχοντας αφαιρέσει σε μεγάλο βαθμό τα περιττά αντικείμενα θα πρέπει να ακολουθήσει η κατάρτιση της πινακίδας. Ένας από τους παράγοντες, ο οποίος αναφέρθηκε και παραπάνω, που επηρεάζουν τη διαδικασία της κατάρτισης είναι η κλίση που μπορεί να έχει μια πινακίδα. Αν μια εικόνα έχει ληφθεί με σημαντική κλίση η κατάρτιση της πινακίδας και η εξαγωγή των αλφαριθμητικών συμβόλων μπορεί να αποτελέσει μια αρκετά δύσκολη υπόθεση. Επομένως για να γίνει ευκολότερη η διαδικασία κατάρτισης θα πρέπει η υπάρχουσα κλίση να απαλειφθεί.



Εικόνα 86 Εντοπισμός κέντρου κάθε σχήματος της εικόνας.

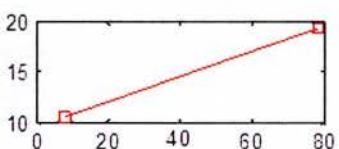
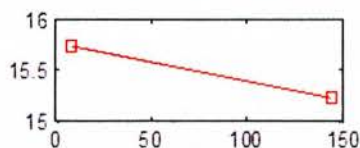
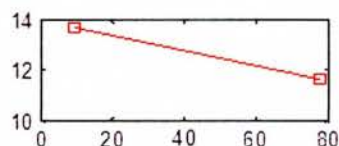
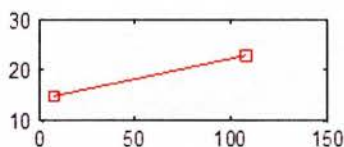
Αρχικά θα πρέπει να εντοπιστεί το κεντρικό ριχεί του κάθε σχήματος που υπάρχει στην εικόνα (Εικόνα 86). Για να εντοπιστεί αν υπάρχει μεγάλη κλίση σε μια εικόνα γίνεται χρήση των συντεταγμένων των κέντρων του πρώτου και του τελευταίου σχήματος που υπάρχουν στην εικόνα. Με αυτόν τον τρόπο εξετάζεται η ευθεία που ενώνει νοητά όλα τα σχήματα της εικόνας. Για παράδειγμα στην εικόνα με την πινακίδα IEN4149, ως πρώτο σχήμα θεωρείται το γράμμα I, ενώ ως τελευταίο το ψηφίο 9. Ενώνοντας τα κέντρα αυτών των δύο σχημάτων δημιουργείται το ευθύγραμμο τμήμα **a**, το οποίο αποτελεί την υποτείνουσα του νοητού ορθογώνιου τριγώνου που διακρίνεται στην παρακάτω εικόνα. Γνωρίζοντας με ακρίβεια τις συντεταγμένες των δύο κέντρων, μπορούμε εύκολα να υπολογίσουμε όλες τις πλευρές του ορθογώνιου τριγώνου.



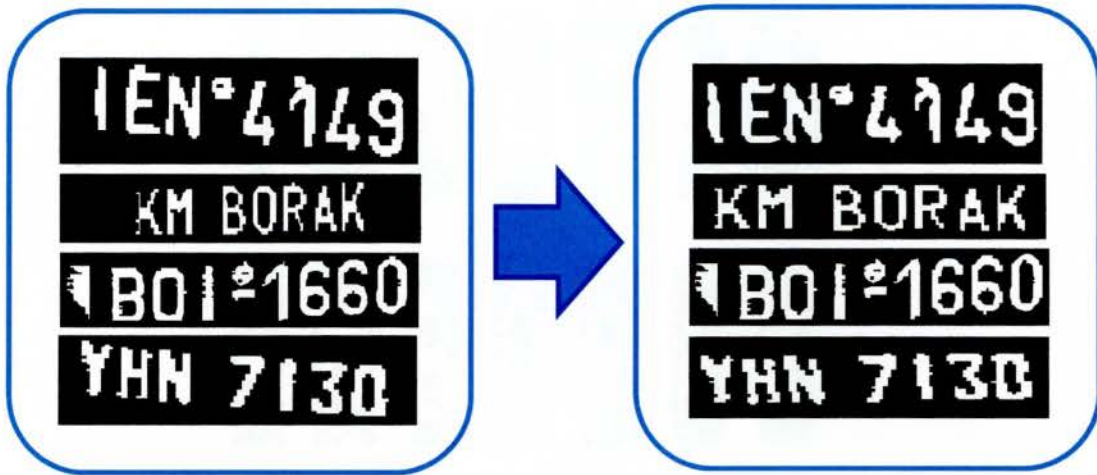
Εικόνα 87 Μεταφορά και ένωση των κέντρων του πρώτου και του τελευταίου σχήματος στο καρτεσιανό επίπεδο.

Για την περιστροφή της πινακίδας θα πρέπει να υπολογιστεί η γωνία που δημιουργείται μεταξύ των πλευρών **a** και **b**. Η γωνία αυτή θα μας υποδείξει πόσες μοίρες θα πρέπει να περιστραφεί η πινακίδα. Ο υπολογισμός της γωνίας προκύπτει υπολογίζοντας το παρακάτω τόξο ημιτόνου :

$$\varphi = \text{τοξημ} \left(\frac{c}{a} \right)$$

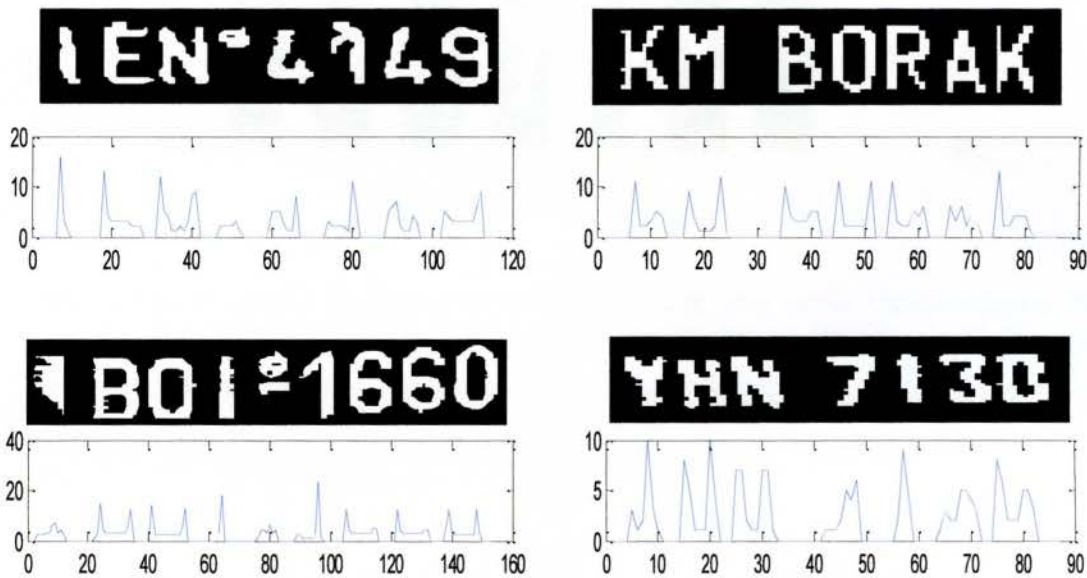


Εικόνα 88 Δημιουργία ευθύγραμμου τμήματος για τον υπολογισμό της κλίσης.



Εικόνα 89 Ευθυγράμμιση πινακίδας.

Έχοντας ολοκληρώσει και την ευθυγράμμιση της πινακίδας (Εικόνα 89), απομένει ο εντοπισμός και η εξαγωγή του κάθε χαρακτήρα. Στη διαδικασία αυτή θα γίνει χρήση της οριζόντιας προβολής, η οποία αναλύθηκε παραπάνω. Η οριζόντια προβολή θα παρουσιάσει τα όρια του κάθε σχήματος, βοηθώντας στον διαχωρισμό τους (Εικόνα 90).



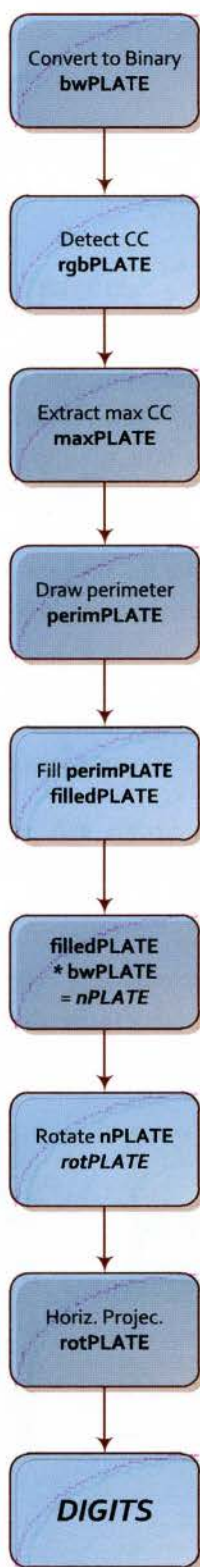
Εικόνα 90 Εύρεση οριζόντιας προβολής.

Μετά την ολοκλήρωση και της οριζόντιας προβολής είμαστε σε θέση να εντοπίσουμε και να ξεχωρίσουμε όλα τα σχήματα της εικόνας. Αυτό επιτυγχάνεται με την παρατήρηση της οριζόντιας προβολής. Τα σημεία που η οριζόντια προβολή έχει μηδενικές τιμές θα πρέπει να αποκοπούν, ώστε να μην είναι ορατά και να προβληθούν μόνο τα σημεία με τιμές μεγαλύτερες του 0. Τα σημεία με θετικές τιμές αποτελούν και τα σχήματα της εικόνας.



Εικόνα 91 Εξαγωγή χαρακτήρων.

Στην επόμενη σελίδα ακολουθεί το διάγραμμα ροής, στο οποίο παρουσιάζονται όλα τα στάδια για την ολοκλήρωση και του δεύτερου βήματος του αλγόριθμου, η κατάτμηση των χαρακτήρων της πινακίδας του οχήματος.



Κεφάλαιο 7

Συμπεράσματα

7.1 Πειραματικά αποτελέσματα

Στην ενότητα αυτή παρουσιάζονται κάποια χαρακτηριστικά παραδείγματα από τα αποτελέσματα της χρήσης του αλγορίθμου. Κάθε παράδειγμα συνοδεύεται και από κάποιες παρατηρήσεις, ώστε να τονιστούν τα σημεία που ο αλγόριθμος δίνει ιδιαίτερη έμφαση, ανάλογα με τη λήψη της ψηφιακής εικόνας.

Επίσης θα πρέπει να τονιστεί ότι στις εικόνες με μεταβαλλόμενο φωτισμό (Εικόνα 97 και Εικόνα 98) η αρχική εικόνα έχει βελτιωθεί με τροποποίηση του ιστογράμματος, από τον ίδιο τον αλγόριθμο.

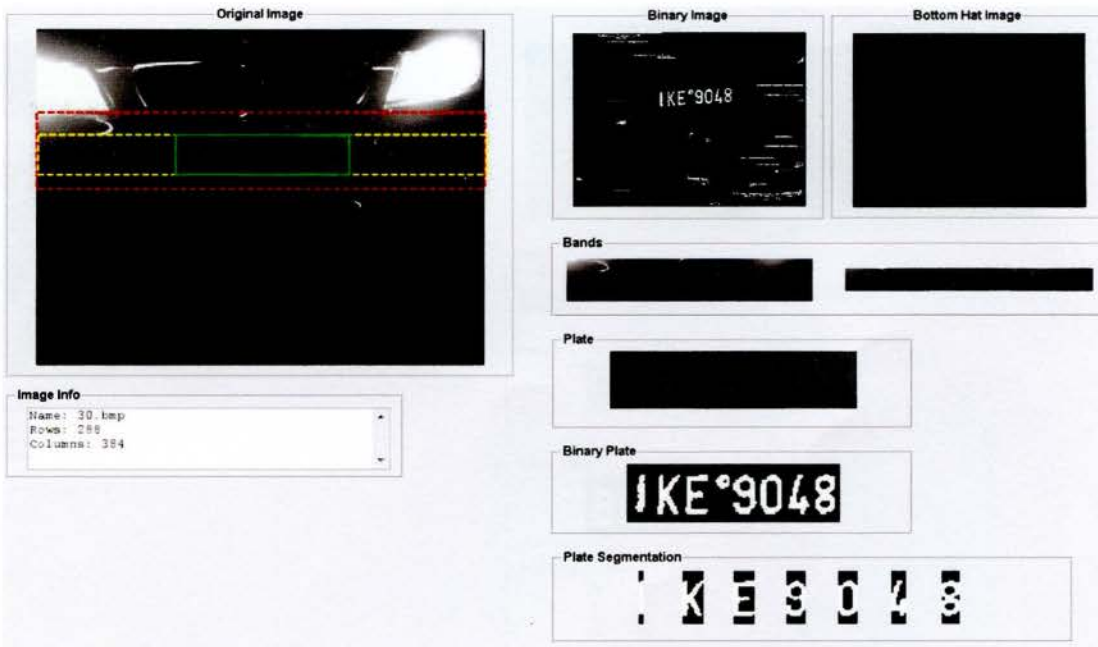
Στην Εικόνα 99 ο αλγόριθμος δεν καταφέρνει να διακρίνει την ύπαρξη των δύο τελευταίων χαρακτήρων, αφού η πινακίδα είναι φθαρμένη.

Ακόμα στην Εικόνα 100 ο αλγόριθμος δεν καταφέρνει να ξεχωρίσει τους δύο πρώτους χαρακτήρες. Η ύπαρξη της βίδας αναγνωρίζεται ως συνδεδετικό κομμάτι των δύο χαρακτήρων.

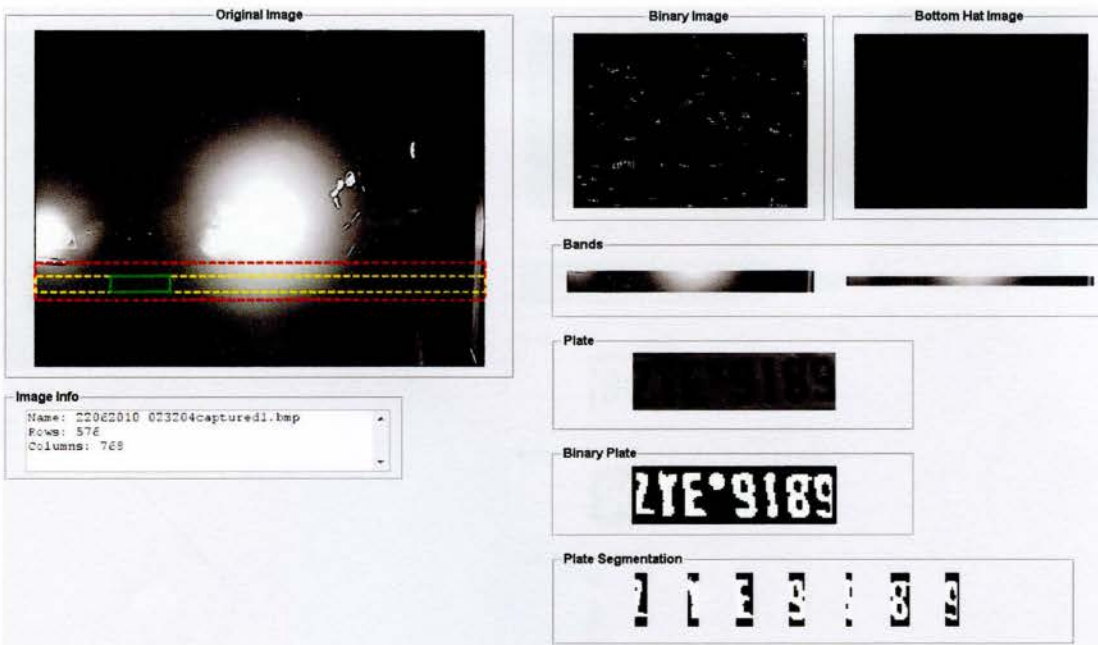
Τέλος στην (Εικόνα 101) και (Εικόνα 102) έχει γίνει μετατροπή του χρωματικού τους χάρτη από έγχρωμη σε αποχρώσεις του γκρι, ώστε να μειωθεί ο χρόνος επεξεργασίας.



Εικόνα 92 Παράδειγμα εφαρμογής του αλγορίθμου με κλίση πινακίδας.



Εικόνα 93 Παράδειγμα εφαρμογής του αλγόριθμου σε εικόνα με χαμηλή αντίθεση.



Εικόνα 94 Παράδειγμα εφαρμογής του αλγόριθμου σε εικόνα βραδυνης λήψης, χωρίς τη χρήση εξωτερικού φωτισμού.



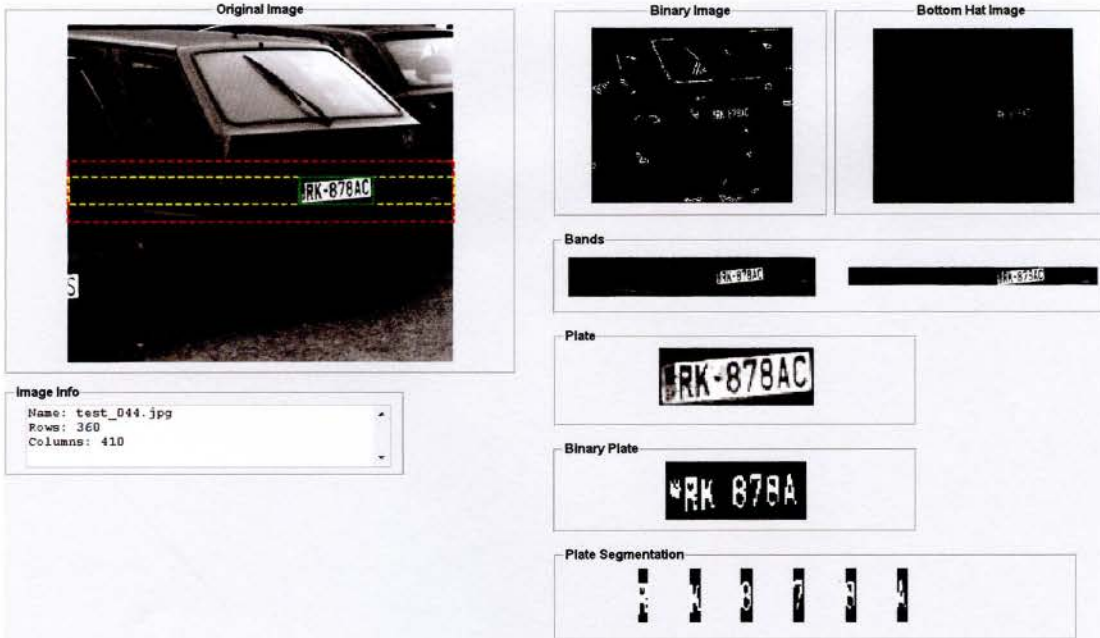
Εικόνα 95 Παράδειγμα εφαρμογής του αλγόριθμου σε φθαρμένη πινακίδα.



Εικόνα 96 Παράδειγμα εφαρμογής του αλγόριθμου σε πινακίδα με ύπαρξη βίδας ανάμεσα στους χαρακτήρες.



Εικόνα 97 Παράδειγμα εφαρμογής του αλγόριθμου σε όχημα λευκού χρώματος.



Εικόνα 98 Παράδειγμα εφαρμογής του αλγόριθμου σε μακρινή λήψη.

Βιβλιογραφία

ΕΛΛΗΝΟΓΛΩΣΣΗ

[1] Παπαμάρκος, Ν., (2005), *Ψηφιακή επεξεργασία εικόνας*, Θεσσαλονίκη: Εκδόσεις Ιδιωτική Έκδοση .

[2] Πήτας, Ι., (2005), *Ψηφιακή επεξεργασία & ανάλυση εικόνας*, Αθήνα: Εκδόσεις Γκιούρδας Β. .

[3] Σαγκριώτης, Ε. – Σέργιος, Θ., (2003), *Σήματα και Επεξεργασία Εικόνας – Ανάλυση Εικόνας και Αναγνώριση Προτύπων*, Πάτρα: Εκδόσεις Ελληνικό Ανοικτό Πανεπιστήμιο.

[4] Σύρκος, Γ. – Κούκος, Ι., (2002), *Εισαγωγή στη Σχεδίαση Συστημάτων Ελέγχου με το MATLAB*, Αθήνα: (Αυτοέκδοση).

[5] Έλληνας, Ι., (2005), *Εγχειρίδιο Μαθήματος Ψηφιακής Επεξεργασία Εικόνας & Βίντεο*

ΞΕΝΟΓΛΩΣΣΗ

[6] Gonzalez, R. – Woods, R., (2002), *Digital Image Processing*, New Jersey.: Prentice Hall, Second edition.

[7] Gonzalez, R. – Woods, R. – Eddins, S., (2004), *Digital Image Processing Using MATLAB*, New Jersey : Prentice Hall, First edition.

[8] McAndrew, A., (2004), *An Introduction to Digital Image Processing with MATLAB*, Victoria University of Technology

[9] Petrou, M. – Bosdogianni, P., (1999), *Image Processing The Fundamentals*, Willy Publications.

[10] Mathworks, (2011), *Image Processing Toolbox - User's guide*, Mathworks.

[11] Halarick, R. – Sharpio, L., (1993) *Computer and Robot Vision*, New York : Addison-Wesley.

[12] Seul, M. – Gorman, L. – Sammon, M., (2000), *Practical Algorithms for Image Analysis*, USA : Cambridge University Press

ΔΙΑΔΙΚΤΥΟ

[13] Image Morphology, http://en.wikipedia.org/wiki/Mathematical_morphology, (accessed August 2011)

[14] Segmentation image processing), <http://fourier.eng.hmc.edu/e161/lectures/morphology/node3.html>, (accessed August 2011).

[15] Automatic Plate Recognition, http://en.wikipedia.org/wiki/Automatic_number_plate_recognition, (accessed August 2011).

[16] License plate and text detection, <http://cmp.felk.cvut.cz/~zimmerk/lpd/index.html>, (accessed December 2011).

[17] License Plate Extraction, Processing, and Character Segmentation, <http://www.lebgeeks.com/forums/viewtopic.php?id=7786>, (accessed December 2011).

[18] Counting Teeth by Ned Gulley, <http://www.mathworks.com/matlabcentral/fileexchange/6002-counting-teeth/content/html/gearcount.html#4>, (accessed December 2011).

[19] Connected – Component Images in Matlab, <http://www.johnloomis.org/ece563/notes/BinaryImages/coins/coins.html>, (accessed December 2011).

[20] Disconnected component Labeling: part 1, <http://blogs.mathworks.com/steve/page/7/>, (accessed December 2011).

