

ΑΤΕΙ Κρήτης Παράρτημα Χανίων τμ. Ηλεκτρονικής

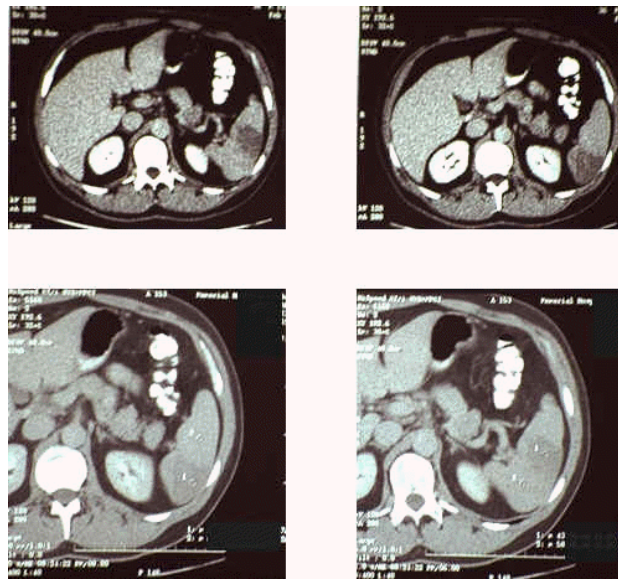
# Προηγμένες εφαρμογές των μαθηματικών στην ψηφιακή επεξεργασία σήματος με χρήση της Matlab

**Iterative Shadowgraphic Method (ISM)**

Πτυχιακή εργασία του Παναγιώτη Αργυρέα A.M.:2996

Εισηγητής Dr. Δημήτρης Πλιάκης

**Τίτλος: Προηγμένες εφαρμογές των μαθηματικών στην ψηφιακή επεξεργασία σήματος με χρήση MatLab®**



## Σκοπός της εργασίας

Σε αυτήν την εργασία επιλύω την διακριτή εξίσωση Poisson με διάφορους μαθηματικούς αλγόριθμους επιλύσεις γραμμικών συστημάτων επίσης κάνω ειδικότερη αναφορά στην επεξεργασία εικόνας και σήματος με σκοπό την βελτίωση του τελικού αποτελέσματος. Αυτά γίνονται με χρήση της **MatLab® 2009<sup>a</sup>** και του πακέτου Image Processing Toolbox. Με σκοπό να μελετήσουμε την χρήση της σαν φίλτρο βελτιώσεις εικόνων μέσω τις σκιαγραφικής επαναληπτικής μεθόδου.

## Περίληψη Κεφαλαίων

Κεφάλαιο 1<sup>ο</sup>: Ορισμένα εισαγωγικά στοιχεία πάνω στην χρήση της ψηφιακής επεξεργασίας σήματος και ιστορικά στοιχεία όσον αφορά την χρήση της στην ιατρική και στο διάστημα.

Κεφάλαιο 2<sup>ο</sup>: Σε αυτό το κεφάλαιο εξετάζουμε το τι είναι εικόνα στον ψηφιακό κόσμο, πως την αντιλαμβάνονται οι υπολογιστές σαν μαθηματικές απεικονίσεις πινάκων και το πώς αποθηκεύονται και πόσο χώρο απαιτούν για αυτό. Επίσης κάνω αναφορά στους διάφορους τύπους ψηφιακής εικόνας που υπάρχουν.

Κεφάλαιο 3<sup>ο</sup>: Σε αυτό το κεφάλαιο αναφέρομαι στα συστήματα συντεταγμένων στο Spatial Domain και την διαδικασία της συνέλιξης.

Κεφάλαιο 4<sup>ο</sup>: Σε αυτό το κεφάλαιο ασχολούμαι με την διαδικασία της τμηματοποίησης της εικόνας σε Grayscale εικόνες, μετά ασχολούμαι με την μαθηματική μορφολογική επεξεργασία εικόνας πιο ειδικά με το Dilation και το Erosion κ.α..

Κεφάλαιο 5<sup>ο</sup>: Σε αυτό το κεφάλαιο μελετάμε την διακριτή εξίσωση Poisson για 2 διαστάσεις και μερικούς τρόπους επίλυσης μέσω **MatLab®** και το τελικό αποτέλεσμα της επεξεργασίας.

Κεφάλαιο 6<sup>ο</sup>: Σε αυτό το κεφάλαιο μελετάμε την ISM Iterative Shadowgraphic Method και την μαθηματική της απόδειξη.

Ευχαριστώ την οικογένεια μου για την συμπαράσταση και την υπομονή τους , την Χαρά Παππά για την αγάπη και υπομονή της και τέλος τον ξάδελφο και συνάδελφο Γιάννη Κουντουράκη.

Ευχαριστώ πολύ.

## Περιεχόμενα

ΚΕΦΑΛΑΙΟ 1	8
1.1 Εισαγωγή	8
1.1.1 Εισαγωγή στην ψηφιακή επεξεργασία σήματος	8
1.1.2 Εφαρμογή της επεξεργασίας εικόνας στην ιατρική.	9
1.1.3 Εφαρμογή της επεξεργασίας εικόνας στο διάστημα.	11
1.2 Εισαγωγή στις σκιαγραφικές μεθόδους	12
1.2.1 Εισαγωγικά	12
1.2.2 Εναλλακτικές μέθοδοι ανάκτησης φάσης	13
1.2.3 Σκιαγραφικές μέθοδοι	13
ΚΕΦΑΛΑΙΟ 2	15
2.1 Τι είναι ψηφιακή εικόνα	15
2.2 Πλήθος bits που απαιτούνται για την αποθήκευση μιας εικόνας	16
2.3 Ευκρίνεια εικόνας	17
2.4 Χαρακτηριστικά εικόνων	19
2.4.1 Οι Εικόνες στο MatLab® και στο Image Processing Toolbox	19
2.4.2 Κατηγορίες δεδομένων εικόνας	18
2.5 Τύποι εικόνων	22
2.5.1 Εισαγωγικά	22
2.5.2 Indexed Images	22
2.5.3 Grayscale Images	23
2.5.4 Δυαδική εικόνα (Binary Image)	23
2.5.5 RGB Εικόνες	24

ΚΕΦΑΛΑΙΟ 3	27
3.1 Συστήματα Συντεταγμένων	27
3.1.1 Σύστημα Συντεταγμένων των Pixel	27
3.1.2 Σύστημα Συντεταγμένων Επιπέδου	28
3.2 Spatial Domain	29
3.2.1 Εισαγωγή	29
3.2.2 Δυσδιάστατη συνέλιξη (2D Convolution)	30
3.2.3 Spatial Filtering	
3.3 Ανάλυση εικόνας στο πεδίο των συχνοτήτων (Frequency Domain)	35
3.3.1 Εισαγωγή	35
3.3.2 Διακριτός μετασχηματισμός Fourier (2-DFT)	36
ΚΕΦΑΛΑΙΟ 4	40
4.1 Τμηματοποίηση εικόνας (Image Segmentation)	41
4.2 Μορφολογική επεξεργασία εικόνας (Morphological Processing)	46
4.2.1 Dilation	47
4.2.2 Περιγραφή λειτουργίας Dilation	48
4.2.3 Εφαρμογή του dilation	49
4.3 Erosion	53
4.3.1 Περιγραφή λειτουργίας Erosion	53
4.3.2 Εφαρμογή του erosion	55
4.4 Opening	59
4.5 Closing	60
ΚΕΦΑΛΑΙΟ 5	
5.1 Διακριτή εξίσωση Poisson	63
5.1.1 Δισδιάστατη	63

5.1.2 Λύση της εξίσωσης στις δύο διαστάσεις	63
5.2 Τρισδιάστατη	65
5.2.1 Λύση της εξίσωσης στις τρεις διαστάσεις	65
<b>ΚΕΦΑΛΑΙΟ 6</b>	
6.1 Επαναληπτική Σκιαγραφική Μέθοδος	67
6.1.1 Εισαγωγικά	67
6.1.2 Μαθηματικά Πρότυπα	67
6.2 Σύγκλιση της Αναπτυξιακής Φάσης	69
6.2.1 Λύση της εξίσωσης στις τρεις διαστάσεις	69
6.2.2 Το αποτέλεσμα του θορύβου στο ISM	73
6.3 Πλεονεκτήματα και περιορισμοί	74
<b>ΠΑΡΑΡΤΗΜΑ Α</b>	74
A.1 Γραμμικές εξισώσεις	74
A.1.1 Εισαγωγή στις γραμμικές εξισώσεις	74
A.1.2 Η λύση στην Matlab	75
A.2 Η μέθοδος Jacobi στην Matlab	78
A.2.1 Εισαγωγή	78
A.2.2 Εφαρμογή στην Matlab	78
<b>ΒΙΒΛΙΟΓΡΑΦΙΑ</b>	79

## ΚΕΦΑΛΑΙΟ 1

### 1.1 Εισαγωγή

#### 1.1.1 Εισαγωγή στην ψηφιακή επεξεργασία σήματος

Η ραγδαία εξέλιξη των υπολογιστών κυρίως μετά το 1975, επέτρεψαν την ανάπτυξη ενός νέου κλάδου που περιγράφεται γενικά ως ψηφιακή επεξεργασία εικόνας (ΨΕΕ). Η ψηφιακή επεξεργασία εικόνας αποτελεί πλέον ολόκληρη επιστήμη και έχει ευρύτερες εφαρμογές όπως για παράδειγμα, την αυτοματοποίηση γραφείου, τη ρομποτική και την όραση μηχανής (computer vision). Με τη λέξη εικόνα δεν νοείται απλά η απεικόνιση μιας σκηνής αλλά είναι ένα μέσο, ένας τρόπος, με τον οποίο μπορούμε να αποτυπώσουμε πληροφορίες διαφόρων ειδών. Έτσι, έγγραφα, ιατρικά δεδομένα (υπερηχογραφήματα, μαγνητικές τομογραφίες κ.λ.π.), διαστημικά δεδομένα κ.α. μπορούν να ψηφιοποιηθούν και να επεξεργασθούν ως εικόνες. Γενικά μπορούμε να πούμε ότι η ΨΕΕ αναπτύχθηκε για να αντιμετωπίσει τα ακόλουθα κύρια προβλήματα:

- Τη ψηφιοποίηση (digitization), κωδικοποίηση εικόνων με στόχο την εκτύπωση, αποθήκευση και μετάδοση τους.
- Τη βελτιστοποίηση (enhancement) και την αποκατάσταση (restoration) των εικόνων με στόχο την καλύτερη απεικόνιση και κατανόηση τους.
- Την τμηματοποίηση (segmentation) και την περιγραφή εικόνων.
- Την ανάλυση και την κατανόηση των εικόνων.

Μια διαφορετική κατηγοριοποίηση των από την παραπάνω ανάλυση μπορούμε να συμπεράνουμε ότι τα θέματα που αντιμετωπίζει η ΨΕΕ εικόνας αφορούν αφενός την αξιοποίηση των μέσων και αφετέρου την κατανόηση του περιεχομένου των εικόνων με απώτερο στόχο την προσέγγιση της ανθρώπινης όρασης. Με την έννοια αυτή η ΨΕΕ ταυτίζεται με θέματα ρομποτικής όρασης (robot vision), αναγνώρισης προτύπων (pattern recognition) και τεχνητής νοημοσύνης (artificial intelligence). Οι επιστημονικοί αυτοί κλάδοι δεν είναι οι μόνοι συγγενείς κλάδοι. Γενικά μπορούμε να αναφέρουμε ότι η ΨΕΕ συγγενεύει άμεσα με τους κλάδους:

- Ψηφιακή Επεξεργασία Σημάτων.
- Ρομποτική Όραση και Όραση Μηχανής.
- Τεχνητή Νοημοσύνη.



- Αναγνώριση Προτύπων.
- Νευρωνικά Δίκτυα (Neural Networks).
- Ασαφή Λογική (Fuzzy Logic).
- Κωδικοποίηση.
- Γραφικά Υπολογιστών (Computer Graphics).

### 1.1.2 Εφαρμογή της επεξεργασίας εικόνας στην ιατρική.

Το 1895, ο Wilhelm Conrad Röntgen ανακάλυψε ότι οι ακτίνες X θα μπορούσαν να περάσουν μέσα από στερεά υλικά . Η ιατρική ξεσηκώθηκε από τη δυνατότητα του να μπορεί κάποιος να δει μέσα στο ζωντανό ανθρώπινο σώμα. Ιατρικά συστήματα ακτινών X διαδόθηκαν σε όλο τον κόσμο μόνο σε μερικά έτη. Παρά την προφανή τους επιτυχία, η απεικόνιση των ιατρικών ακτινών X περιορίστηκε από τέσσερα προβλήματα μέχρι να κάνει την εμφάνισή τους το DSP (Digital Signal Processing) και άλλες σχετικές τεχνικές κατά τη διάρκεια της δεκαετίας του '70.

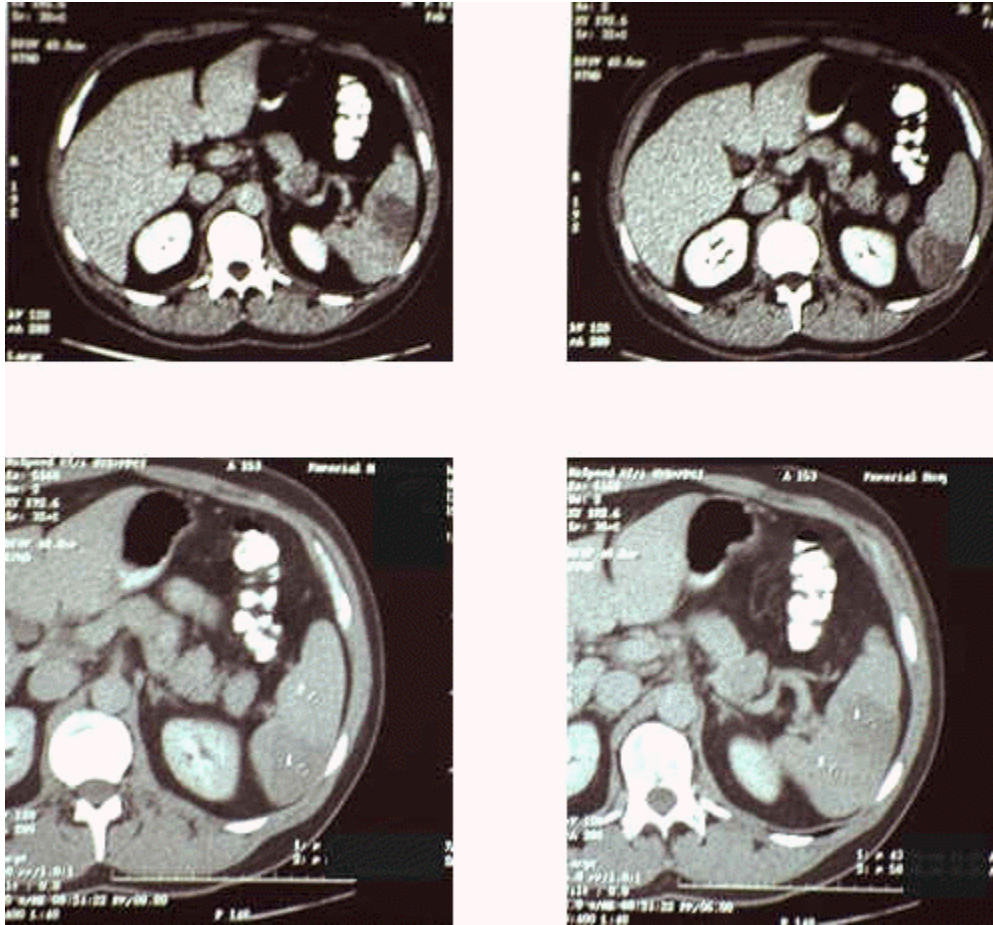
Ένα από τα προβλήματα που παρουσιάστηκαν στη ιατρική χρήση των ακτινών X, είναι η επικάλυψη που εμφανίζουν μεταξύ τους κάποιες δομές του ανθρωπίνου σώματος, με αποτέλεσμα να μην γίνονται άμεσα ορατές οι περιοχές που παρουσιάζουν ενδιαφέρον. Για παράδειγμα, κάποια τμήματα της καρδιάς μπορεί να μην είναι ορατά πίσω από νεύρα. Επίσης, με την χρήση αυτής της μεθόδου δεν είναι πάντα δυνατό να ξεχωρίσουν μεταξύ τους παρόμοιοι ιστοί. Παραδείγματος χάριν, μπορεί να είναι εμφανής ο διαχωρισμός του κόκαλου από τον μαλακό ιστό, αλλά δεν διακρίνεται η διαφορά ενός όγκου από ένα ζωτικό όργανο όπως είναι το συκώτι. Ακόμη οι εικόνες ακτινών X παρουσιάζουν την ανατομία του ανθρωπίνου σώματος, δηλαδή τη δομή του και όχι τη φυσιολογία, δηλαδή τη λειτουργία του.

Το πρόβλημα των επικαλυμμένων δομών του σώματος λύθηκε το 1971 με την εισαγωγή του πρώτου σαρωτή τομογραφίας με τη χρήση ηλεκτρονικού υπολογιστή γνωστό και ως CAT scanner (Computed axial Tomography). Η τομογραφία με τη χρήση υπολογιστή είναι ένα κλασικό παράδειγμα της ψηφιακής επεξεργασίας σήματος. Οι ακτίνες X από πολλές κατευθύνσεις περνούν μέσα από το τμήμα του σώματος του ασθενή που εξετάζεται. Αντί της απλής μορφής των εικόνων που προκύπτουν από τις ακτίνες X που έχουν ανιχνευθεί, τα σήματα μετατρέπονται σε ψηφιακά δεδομένα και αποθηκεύονται σε έναν υπολογιστή. Οι πληροφορίες αυτές χρησιμοποιούνται έπειτα για να υπολογιστούν οι εικόνες που εμφανίζονται σαν τομές μέσα στο σώμα.

Αυτές οι εικόνες παρουσιάζουν πολύ μεγαλύτερη λεπτομέρεια συγκριτικά με τις συμβατικές τεχνικές, επιτρέποντας σημαντικά την καλύτερη διάγνωση και θεραπεία. Ο αντίκτυπος της τομογραφίας ήταν σχεδόν τόσο μεγάλος όσο και η αρχική εισαγωγή των της απεικόνισης εικόνων ακτίνων X. Μέσα μόνο σε μερικά έτη, κάθε μεγάλο νοσοκομείο στον κόσμο είχε πρόσβαση σε τομογράφο σάρωσης. [1]



Εικ.1 Αξονικός τομογράφος

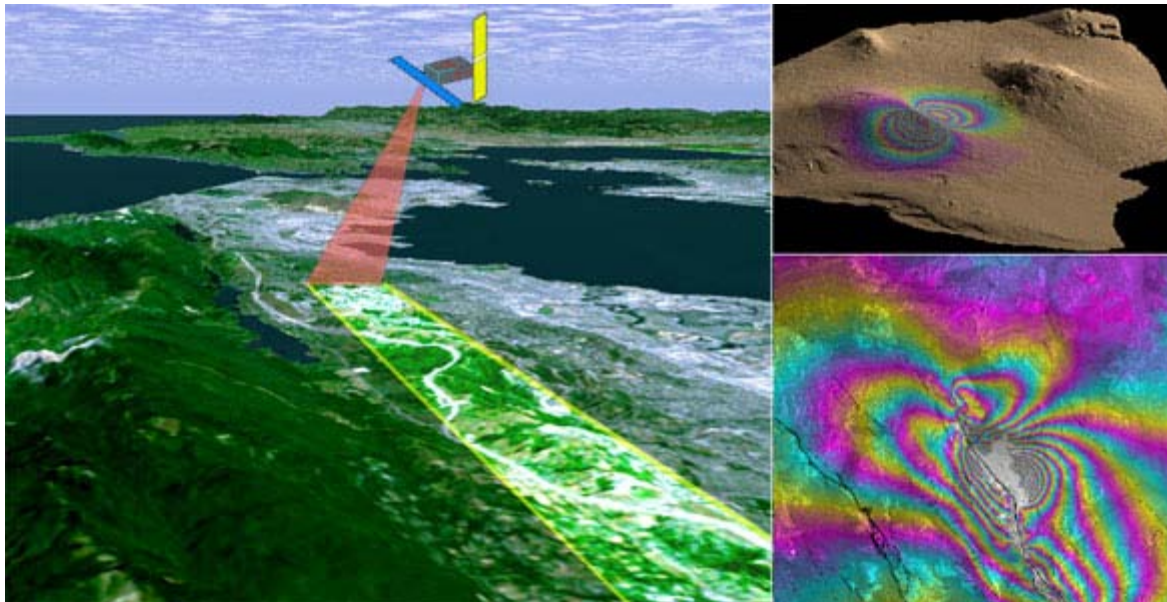


Εικ.2 Αποτελέσματα σάρωσης αξονικού τομογράφου

### 1.1.3 Εφαρμογή της επεξεργασίας εικόνας στο διάστημα.

Μερικές φορές, πρέπει να προκύψουν όσο το δυνατόν περισσότερα αποτελέσματα από μια κακή εικόνα. Αυτό το φαινόμενο εμφανίζεται συχνά οι εικόνες λαμβάνονται από μη επανδρωμένους δορυφόρους και από διαστημικά οχήματα εξερεύνησης. Η ψηφιακή επεξεργασία εικόνας μπορεί να βελτιώσει την ποιότητα των εικόνων που λαμβάνονται κάτω από εξαιρετικά δυσμενείς συνθήκες, με διάφορους τρόπους: ρύθμιση αντίθεσης και φωτεινότητας, ανίχνευση ακρών, μείωση θορύβου, ρύθμιση εστίασης, αύξηση ευκρίνειας κλπ. Εικόνες που παρουσιάζουν χωρική διαστρέβλωση, όπως συναντιέται σε μια επίπεδη εικόνα που απεικονίζει έναν σφαιρικό πλανήτη, μπορεί επίσης να στρεβλωθεί για μια διορθωμένη παρουσίαση. Πολλές μεμονωμένες εικόνες μπορούν επίσης να συνδυαστούν σε μια ενιαία βάση δεδομένων, επιτρέποντας στις πληροφορίες να επιδειχθούν με μοναδικούς τρόπους. Για παράδειγμα, μια τηλεοπτική ακολουθία που εξομοιώνει μια εναέρια πτήση πάνω από την επιφάνεια ενός απομακρυσμένου πλανήτη.

Ο μεγάλος όγκος πληροφοριών στις εικόνες είναι ένα πρόβλημα για τα συστήματα που πωλούνται σε ποσότητα στο ευρύ κοινό. Τα εμπορικά συστήματα πρέπει να είναι φτηνά, και αυτό δεν συνεργάζεται καλά με τις μεγάλες μνήμες και τους υψηλούς ρυθμούς μετάδοσης δεδομένων. Η απάντηση δίνεται από την ψηφιακή επεξεργασία με την συμπίεση εικόνας. Όπως ακριβώς με τα σήματα φωνής, οι εικόνες περιέχουν ένα τεράστιο ποσό περιττών πληροφοριών, και μπορούν να απεικονιστούν με τη χρήση κατάλληλων αλγόριθμων οι οποίοι μειώνουν των αριθμό των bits. Η τηλεόραση και άλλες κινούμενες εικόνες είναι ιδιαίτερα κατάλληλες για τη συμπίεση, δεδομένου ότι το μεγαλύτερο μέρος της εικόνας παραμένει ίδιο από το καρέ σε καρέ. Εμπορικά προϊόντα απεικόνισης που εκμεταλλεύονται αυτήν την τεχνολογία περιλαμβάνουν: βίντεο τηλέφωνα, προγράμματα υπολογιστών που απεικονίζουν κινούμενες εικόνες και ψηφιακή τηλεόραση. [1]



Εικ.3 D.S.P. από δορυφόρο

## 1.2 Εισαγωγή στις σκιαγραφικές μεθόδους

### 1.2.1 Εισαγωγικά

Τις τελευταίες τρεις δεκαετίες η εξέλιξη της Ψ.Ε.Ε. και η ραγδαία εξέλιξη της υπολογιστικής ικανότητας των προσωπικών υπολογιστών έχουν ανοίξει των δρόμων στον σχεδιασμό πολυπλοκότερων μαθηματικών οπτικών φίλτρων, επιτρέποντας να βελτιώσουμε την εικόνα με τρόπους που παλαιότερα ήταν εφικτή μόνο μέσω οπτικών τρόπων.

Όπως ανέφερα σε προηγούμενη παράγραφο, η Ψ.Ε.Ε. χρησιμοποιείται ευρέως σε εφαρμογές επεξεργασίας εικόνων που λαμβάνονται από δορυφόρους, μια ακόμη εντυπωσιακή

χρήση της Ψ.Ε.Ε. σε διαστημικές εφαρμογές είναι σε εικόνες που έχουν ληφθεί μέσω του διαστημικού τηλεσκοπίου Hubble και βελτιώθηκαν από την raw μορφή τους με τον αλγόριθμο από-συνέλιξης CLEAN. Πιο συγκεκριμένα, η βελτίωση των υπολογιστών (της υπολογιστικής τους δύναμη) βελτίωσε τη μέτρηση της γωνίας φάσης του διαδιδόμενου κύματος. Έτσι η ελεύθερη διάδοση κυματικών επιφανειών μπορεί να υπολογιστεί μέσα σε κλάσματα δευτερόλεπτου μέσω της επεξεργασίας ολογραμμάτων (ψηφιακή εγγραφή ενός εμπλεκόμενου πεδίου). Τα ολογράμματα σήμερα είναι περισσότερο διαδεδομένα και ειδικότερα σε αντικείμενα με λεπτή επιφάνεια φάσης όπως στην μικροσκοπία. Ένα άλλο πεδίο είναι τα μεγαλύτερα αντικείμενα, ας πούμε στην περίπτωση του αξονικού τομογράφου ο σκοπός μας επιτυγχάνεται μέσω αλγεβρικών μεθόδων ανακατασκευής του τελικού αντικειμένου και συγκεκριμένα μετρήσεις και φωτογραφίες του αντικείμενου από διάφορες γωνίες.

### 1.2.2 Εναλλακτικές μέθοδοι ανάκτησης φάσης

Οι προηγούμενες διαδικασίες απαιτούν μια ακτίνα αναφοράς. Με ένα σένσορα ο οποίος χρησιμοποιεί το διάμηκες εκτόπισμα των σημείων εστίασης μιας σειράς φακών για να πάρει δείγματα της κλήσης των επιφανειών διάδοσης των κυμάτων. Τώρα από αυτά τα δείγματα μπορούμε με επαναληπτικές “χαλαρές” μαθηματικές μεθόδους, όπου χαλαρές σημαίνει ότι το λάθος στους υπολογισμούς αντιμετωπίζεται σαν ένα όριο το οποίο ελαχιστοποιείται ή χαλαρώνει μέσα σε ένα όριο ανοχής το οποίο έχουμε ορίσει κατάλληλα, να ανακατασκευάσουμε το αρχικό αντικείμενο. Αυτό έγινε εφικτό μέσω της κατασκευής σειράς μικρό-φάκων που προσαρμόστηκαν σε CCD κάμερες υψηλής ανάλυσης και αυτός ο σένσορας ονομάζεται Shack-Hartmann. Τελευταία είναι εφικτό μέσω αυτού του σένσορα να ανακτήσουμε μοναδιαίες επιφάνειες φάσης.

### 1.2.3 Σκιαγραφικές μέθοδοι

Στις σκιαγραφικές μεθόδους οι διαφορές στην κατανομή της έντασης στα δύο πεδία διάδοσης είναι συνδεδεμένη με τις ανακλάσεις των φακών και τα διαμήκη συστατικά των αντικειμένων. Η ανάκτηση των προφίλ φάσης γίνεται μέσω της λύσης μερικών ελλειπτικών διαφορικών σε πλέγμα.

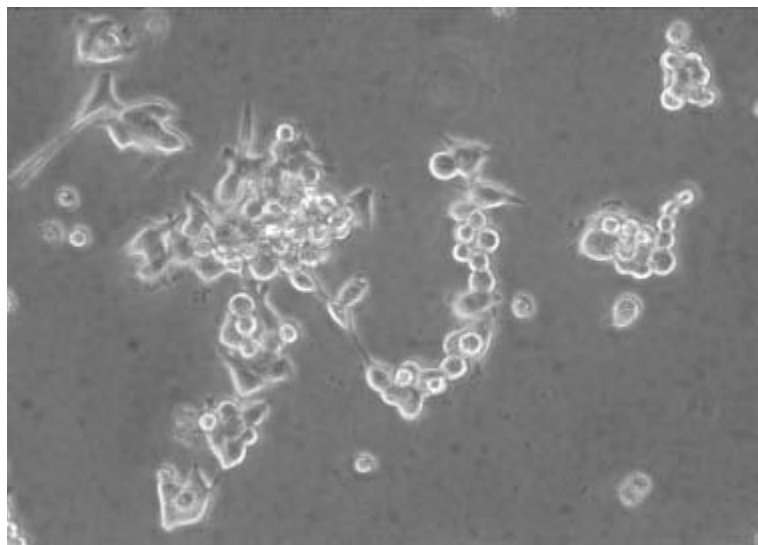
Ένα από τα θετικά των μεθόδων που χρησιμοποιούν την αυτό-αναφορά είναι στην δυνατότητα να δίνουμε προσδιορισμό στις διαταραχές των επιφανειών όπου το αρχικό φως είναι χαμηλής συνοχής. Μερικές τυπικές εφαρμογές είναι στα ψηφιακά μικροσκόπια, σε εικόνες από το διάστημα με το φως των αστεριών και σε διάφορες άλλες εφαρμογές φασματικής φωτογραφίας.

Μερική περιορισμοί αυτών των μεθόδων είναι η δυνατότητα ανακατασκευής μόνο μη-περιστρεφόμενων αντικειμένων. Η οποία δυνατότητα δίνεται σε άλλους επαναληπτικούς αλγορίθμους οι οποίοι βρίσκουν εφαρμογή σε πολυποίκιλα πεδία όπως την ικανοποίηση της

διάχυσης ηλεκτρονίων και στην παραγωγή ολογραμμάτων μέσω υπολογιστή. Βέβαια αυτοί οι αλγόριθμοι έχουν πολύ αργό ρυθμό σύγκλισης και σε μεγάλης ανάλυσης εικόνες έχουν το πρόβλημα της αποτελμάτωσης.[29]



Εικόνα από το Hubble



Εικόνα φασματικής μικροσκοπίας

## ΚΕΦΑΛΑΙΟ 2

### 2.1 Τι είναι ψηφιακή εικόνα

Η μετάβαση από τον αναλογικό κόσμο στον ψηφιακό συνεπάγεται τη μετατροπή αναλογικών σημάτων σε ψηφιακά. Έτσι μια πραγματική εικόνα μεταφέρεται στον ψηφιακό κόσμο με τη μορφή διακεκριμένου σήματος που έχει τη μορφή ψηφιακών πινάκων. Μια ψηφιακή εικόνα μπορεί να είναι δυαδική (binary images), μονοχρωματική αποχρώσεων του γκρι (gray-level ή gray-scale images) ή έγχρωμη (color image). Μια ψηφιακή εικόνα αποχρώσεων του γκρι διαστάσεων  $N \times M$  παριστάνεται από έναν δυσδιάστατο πίνακα ακέραιων αριθμών  $I(i,j)$ ,  $i=0,\dots,N-1$  και  $j=0,\dots,M-1$ , όπου  $0 \leq I(i,j) \leq G-1$ . Το  $G$  ισούται συνήθως με μια δύναμη του 2, δηλαδή  $G = 2^m$  με συνηθέστερη τιμή το  $m = 8$  που αντιστοιχεί σε 256 αποχρώσεις του γκρι. Σημειώνεται ότι προς διευκόλυνση της ψηφιακής επεξεργασίας των εικόνων όχι μόνον το  $G$  αλλά και οι διαστάσεις είναι επιθυμητό να είναι δυνάμεις του δύο.

$$I(i,j) = \begin{bmatrix} I(0,0) & \dots & I(0,M-1) \\ \vdots & \ddots & \vdots \\ I(N-1,0) & \dots & I(N-1,M-1) \end{bmatrix}$$

Η τιμή  $I(i,j)$  είναι ανάλογη της φωτεινότητας της εικόνας στο εικονοστοιχείο (pixel)  $(i,j)$  και συνεπώς ο πίνακας  $I(i,j)$  είναι ουσιαστικά μια διακεκριμένη συνάρτηση που εκφράζει την ένταση της φωτεινότητας της εικόνας (light intensity function) σε κάθε εικονοστοιχείο. Η λέξη pixel (ή pel ορισμένες φορές) προέρχεται και αντικαθιστά ουσιαστικά τη φράση «picture element».

Η απλούστερη μορφή μιας εικόνας είναι η δυαδική μορφή. Μια δυαδική εικόνα έχει μόνον δύο στάθμες φωτεινότητας που συνήθως είναι το μαύρο και το άσπρο. Το μαύρο αντιστοιχεί στη τιμή 0 και το άσπρο στην τιμή 1 (ή σε αντιστοίχιση με τις gray-level εικόνες στην τιμή 255). Μια δυαδική εικόνα καταλαμβάνει μικρότερη μνήμη και οι επεξεργασίες της απαιτεί μικρότερο υπολογιστικό κόστος. Σε δυαδική μορφή μπορούν να απεικονισθούν σημαντικές πληροφορίες όπως είναι το εμβάδόν και η θέση αντικειμένων, η μορφή αντικειμένων κ.ά. Σημειώνεται ότι πάρα πολλές και σημαντικές εφαρμογές της ΨΕΕ όπως είναι η οπτική αναγνώριση χαρακτήρων (OCR: Optical Character Recognition), η αναγνώριση υπογραφής (signature recognition) και η αναγνώριση αποτυπωμάτων (fingerprint recognition) γίνονται συνήθως με τη χρήση δυαδικών εικόνων.

Οι έγχρωμες ψηφιακές εικόνες αποτελούν το μέσο για την απεικόνιση του πραγματικού κόσμου. Μια έγχρωμη ψηφιακή εικόνα αποτελείται από τρεις gray-level εικόνες. Δηλαδή, το χρώμα κάθε εικονοστοιχείου έχει τρεις συνιστώσες που αντιστοιχούν στις γκρι αποχρώσεις των αντίστοιχων εικονοστοιχείων των τριών γκρι εικόνων. Από μαθηματικής άποψης, μια ψηφιακή έγχρωμη εικόνα διαστάσεων  $N \times M$  μπορεί να συμβολιστεί ως  $I_c(i,j)$ ,  $I = 0, \dots, N-1$  και  $j = 0, \dots, M-1$ , όπου  $0 \leq I_c(i,j) \leq G-1$  για κάθε  $c = 1, 2, 3$ . Έτσι, το χρώμα κάθε εικονοστοιχείου  $(i, j)$  προκύπτει από το συνδυασμό τριών χρωματικών συνιστωσών:

$$\text{Color}(i,j) = (I1(i,j), I2(i,j), I3(i,j))$$

Για παράδειγμα, στο χρωματικό σύστημα RGB, το κάθε χρώμα συντίθεται από τα χρώματα κόκκινο (Red), πράσινο (Green) και μπλε (Blue). Παρακάτω θα περιγραφούν αναλυτικότερα οι παραπάνω τύποι εικόνων. [2][3]

## 2.2 Πλήθος bits που απαιτούνται για την αποθήκευση μιας εικόνας

Μια εικόνα διαστάσεων  $N \times M$  και πλήθους αποχρώσεων  $G=2^m$  απαιτεί  $b = N \cdot M \cdot m$  bits για να αποθηκευτεί. Το  $m$  ονομάζεται βάθος bit (bit depth) ή βάθος χρώματος (color depth) και εκφράζει πόση χρωματική πληροφορία για εμφάνιση ή εκτύπωση έχει η εικόνα. Μεγαλύτερο βάθος bit σημαίνει περισσότερες διαθέσιμες αποχρώσεις και αποχρώσεις που αποδίδονται με μεγαλύτερη ακρίβεια στις ψηφιακές εικόνες. Παρακάτω ακολουθεί ένας πίνακας με ενδεικτικά παραδείγματα για το απαιτούμενο πλήθος bits για την αποθήκευση μιας εικόνας. [4]

**Πίν.1 Απαιτούμενος αριθμός bits μιας εικόνας**

α/α	Τύπος εικόνας	N	M	m	bits	Bytes
1	Δυαδική	100	100	1	10.000	1.250
2	Αποχρώσεων του γκρι	100	100	8	80.000	10.000
3	Έγχρωμη	100	100	24	240.000	30.000
4	Δυαδική	256	256	1	65.536	8.192
5	Αποχρώσεων του γκρι	256	256	8	524.288	65.536
6	Έγχρωμη	256	256	24	1.572.864	196.608
7	Δυαδική	512	512	1	262.144	32.768
8	Αποχρώσεων του γκρι	512	512	8	2.097.152	256.144
9	Έγχρωμη	512	512	24	6.291.456	786.432



### 2.3 Ευκρίνεια εικόνας

Η ευκρίνεια μιας εικόνας καθορίζει το πόσο καλά μπορούμε να βλέπουμε τις λεπτομέρειες μιας εικόνας. Ισούται με το πλήθος των εικονοστοιχείων ανά μονάδα επιφάνειας και συνήθως μετριέται σε pixels/in<sup>2</sup> ή διαφορετικά σε dpi (dots per inches). Είναι φανερό ότι η ευκρίνεια εξαρτάται τόσο από τις διαστάσεις όσο και από το πλήθος των αποχρώσεων της κάθε εικόνας. Αν για παράδειγμα κρατήσουμε σταθερό το  $m$  και μεταβάλλουμε (μειώνουμε) τις διαστάσεις μιας εικόνας θα εμφανίσει το φαινόμενο του σκακιού (checkerboard). Δηλαδή, η εικόνα και κατατεμαχιστεί σε τετράγωνα ομοιόμορφες χρωματικά περιοχές .

Εικ.4



Ολόκληρη η εικόνα

6,17% του αρχικού αριθμού των pixel

Εικ.5



2.78% του αρχικού αριθμού των pixel

1,54% του αρχικού αριθμού των pixel

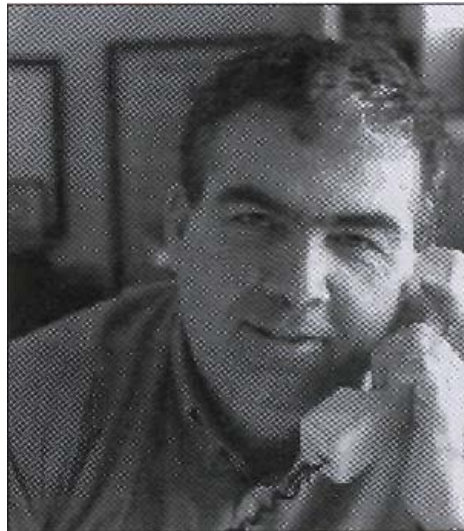
Αν από την άλλη μεριά κρατήσουμε σταθερές τις διαστάσεις και μειώσουμε το  $m$  τότε θα εμφανιστούν πάλι οπτικά ομοιόμορφες χρωματικά περιοχές, που όμως η ευκρίνεια τους θα καθορίζεται από το  $m$  και τη συγγένεια τοπικά των αποχρώσεων. [5][6]



$m=8$



$m=6$



$m=4$



m=8

m=6

Εικόνες από 6 έως 10: Επίδραση της μεταβολής του  $m$  σε μια εικόνα.

## 2.4 Χαρακτηριστικά εικόνων

### 2.4.1 Οι Εικόνες στο MatLab® και στο Image Processing Toolbox

Η βασική δομή δεδομένων στο MatLab® είναι ο πίνακας, μια καθορισμένη διαμόρφωση πραγματικών ή πολύπλοκων στοιχείων. Αυτό το αντικείμενο ταιριάζει απολύτως φυσικά στην αναπαράσταση εικόνων, πραγματικών καθορισμένων τιμών χρώματος ή intensity data.

Το MatLab® αποθηκεύει τις περισσότερες εικόνες σαν δυδιάστατους πίνακες, που κάθε στοιχείο του πίνακα αντιστοιχεί σε ένα συγκεκριμένο εικονοστοιχείο της εικόνας. Τα στοιχεία του κάθε πίνακα αντικατοπτρίζουν συνήθως την φωτεινότητα του κάθε pixel. Για παράδειγμα, μια εικόνα που αποτελείται από 200 σειρές και 300 στήλες διαφορετικών κουκίδων χρώματος θα αποθηκευτούν σε ένα πίνακα διαστάσεων 200 x 300. Κάποιες εικόνες, όπως οι RGB, απαιτούν τρισδιάστατους πίνακες, όπου το πρώτο επίπεδο στην τρίτη διάσταση αντιπροσωπεύει τις αποχρώσεις του κόκκινου εικονοστοιχείου το δεύτερο επίπεδο του πράσινου και το τρίτο του μπλε. [7]

### 2.4.2 Κατηγορίες δεδομένων εικόνας

Αν και οι συντεταγμένες κάθε pixel μιας εικόνας είναι ακέραιοι αριθμοί, οι τιμές της φωτεινότητας που αντιπροσωπεύει το καθένα από αυτά δεν περιορίζονται μόνο στο σύνολο των ακεραίων. Συνήθως στο MatLab® τα περισσότερα δεδομένα των εικόνων αποθηκεύονται σε πίνακες κλάσης double. Τα δεδομένα σε αυτούς τους πίνακες αποθηκεύονται ως 64 bit αριθμοί (8 bytes/pixel), κινητής υποδιαστολής (floating point).

Για επεξεργασία εικόνας, ωστόσο αυτή η αναπαράσταση των δεδομένων δεν είναι πάντα η ιδανική. Ο αριθμός των εικονοστοιχείων στην εικόνα μπορεί να είναι πολύ μεγάλος, για παράδειγμα μια 1000\*1000 εικόνα έχει 1.000.000 εικονοστοιχεία.

Αφού κάθε εικονοστοιχείο αντιπροσωπεύεται από τουλάχιστον ένα στοιχείο του πίνακα αυτή η εικόνα απαιτεί περίπου 8 megabytes μνήμης. Για να μειώσουμε τις απαιτήσεις της μνήμης, το MatLab® υποστηρίζει την αποθήκευση των δεδομένων σε πίνακες των 8-bit ή των 16-bit μη προσημασμένων ακεραίων, τάξης unit8 και unit16. Αυτοί οι πίνακες απαιτούν το ένα όγδοο ή το ένα το τέταρτο της μνήμης που απαιτούν οι double πίνακες.[8]

Ακολουθεί ένας συγκεντρωτικός πίνακας των τύπων των δεδομένων εικόνας.

Πιν.2

Κλάσεις δεδομένων εικόνας	Περιγραφή
<b>double</b>	Αυξημένη ακρίβεια, κινητής υποδιαστολής αριθμοί στο προσεγγιστικό εύρος τιμών [-10308, 10308], (8 Bytes/pixel)
<b>uint8</b>	Μη προσημασμένος 8-bit ακέραιος στο πεδίο τιμών [0,255], (1 Byte/pixel)
<b>uint16</b>	Μη προσημασμένος 16-bit ακέραιος στο πεδίο τιμών [0,65535], (2 Bytes/pixel)
<b>uint32</b>	Μη προσημασμένος 32-bit ακέραιος στο πεδίο τιμών [0,4294967295], (4 Bytes/pixel)
<b>int8</b>	Προσημασμένος 8-bit ακέραιος στο πεδίο τιμών [-128,127], (1 Byte/pixel)
<b>int16</b>	Προσημασμένος 16-bit ακέραιος στο πεδίο τιμών [-32768,32767], (2 Bytes/pixel)
<b>Int32</b>	Προσημασμένος 32-bit ακέραιος στο πεδίο τιμών [-2147483648,2147483647], (2 Bytes/pixel)
<b>Single</b>	Απλή ακρίβεια, κινητής υποδιαστολής αριθμοί στο προσεγγιστικό εύρος τιμών [-1038, 1038], (4 Bytes/pixel)
<b>Char</b>	Χαρακτήρες (2 Bytes/pixel)
<b>logical</b>	Οι τιμές είναι 0 και 1, (1 Byte/pixel)

## 2.5 Τύποι εικόνων

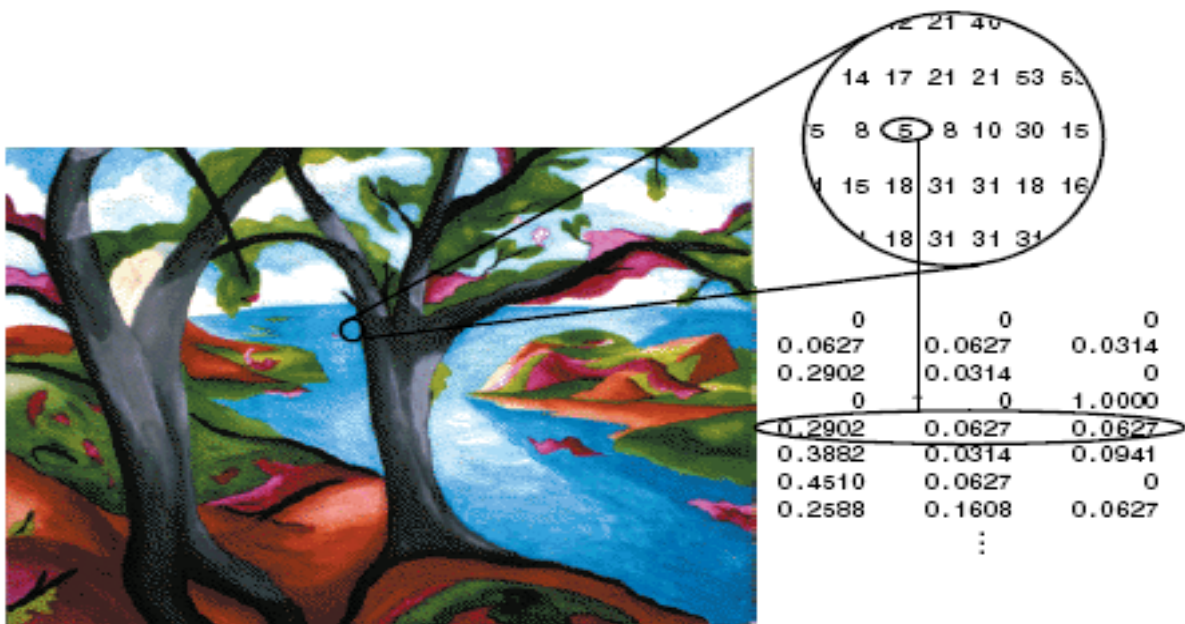
### 2.5.1 Εισαγωγικά

Το Image Processing Toolbox υποστηρίζει τέσσερις βασικούς τύπου εικόνων:

- Indexed Images
- Intensity Images
- Binary Images
- RGB Images

### 2.5.2 Indexed Images

Μια indexed εικόνα αποτελείται από έναν πίνακα δεδομένων  $X$  και ένα πίνακα χρωμάτων,  $map$ . Ο πίνακας δεδομένων μπορεί να είναι της τάξεως `uint8`, `uint16` ή `double`. Ο πίνακας χρωμάτων μπορεί να είναι ένας  $m \times 3$  πίνακας τάξεως `double` που περιέχει κινητής υποδιαστολής τιμές στο διάστημα  $[0 \ 1]$ . Κάθε σειρά του πίνακα χρωμάτων, καθορίζει την κόκκινη, την πράσινη, και τη μπλε συνισταμένη ενός μόνο χρώματος. Το χρώμα κάθε εικονοστοιχείου της εικόνας καθορίζεται χρησιμοποιώντας της αντιπροσωπευτική τιμή του  $X$  σαν ευρετήριο στον πίνακα χρωμάτων. Η τιμή 1 δείχνει στην πρώτη σειρά του πίνακα, η τιμή 2 στην δεύτερη σειρά κ.ο.κ..

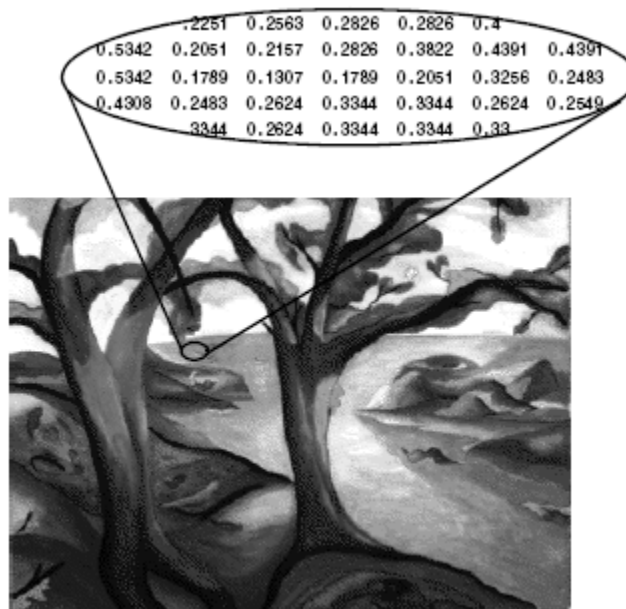


Εικ.11 Indexed εικόνα –Πίνακας χρωμάτων

### 2.5.3 Grayscale Images

Μια grayscale εικόνα είναι ένας πίνακας δεδομένων,  $I$ , που οι τιμές του αντιπροσωπεύουν την ένταση του κάθε pixel μέσα στο διάστημα τιμών από 0 έως 1. Το MatLab® αποθηκεύει μια grayscale εικόνα σαν ένα μονοδιάστατο πίνακα με κάθε στοιχείο του πίνακα να αντιπροσωπεύει ένα εικονοστοιχείο της εικόνας. Ο πίνακας μπορεί να είναι της τάξεως `double`, `unit8` ή `unit16`. Τα στοιχεία στον πίνακα δεδομένων της εικόνας αντιπροσωπεύουν διάφορες εντάσεις ή αλλιώς αποχρώσεις του γκρι, όπου η ένταση 0 συνήθως αντιπροσωπεύει το μαύρο, και η ένταση 1 ή 255 ή 65535 το άσπρο.

Το σχήμα παρακάτω αναπαριστά μια grayscale εικόνα τάξεως `double` στο διάστημα τιμών  $[0,1]$ :

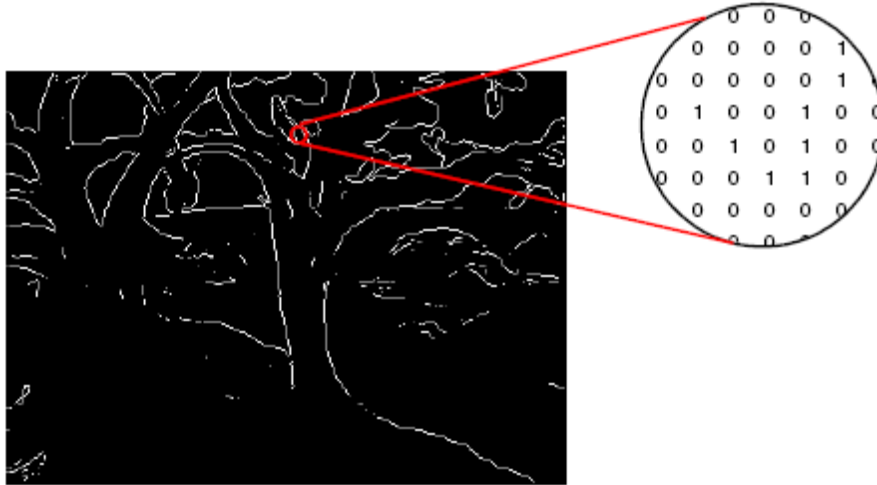


Εικ.12 Grayscale εικόνα

### 2.5.4 Δυαδική εικόνα (Binary Image)

Σε μια binary εικόνα κάθε εικονοστοιχείο παίρνει μια από τις μόνο δυο διακριτές τιμές, 0 και 1. Ουσιαδώς οι δυο αυτές τιμές ανταποκρίνονται στο on και στο off αντίστοιχα. Μια binary εικόνα αποθηκεύεται σαν λογικός πίνακας (logical data class) από μηδενικά και άσσους.

Το σχήμα παρακάτω αντιπροσωπεύει μια binary εικόνα:



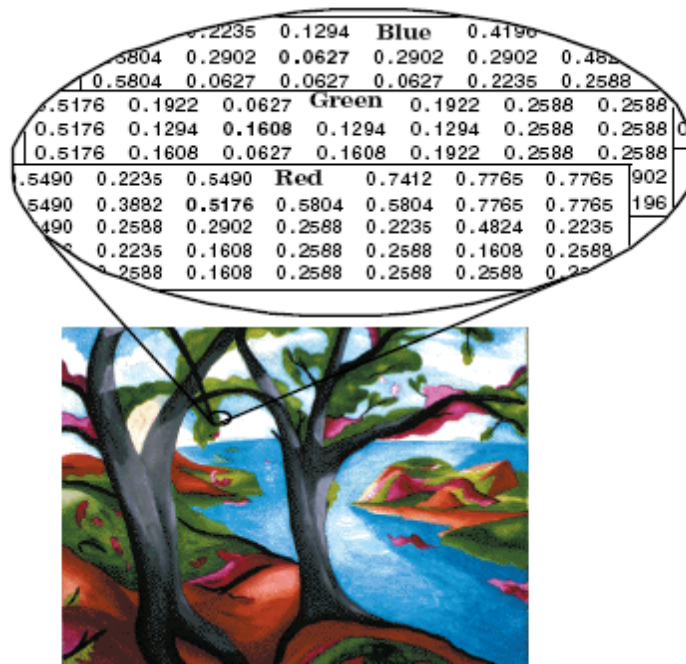
Εικ.13 Binary εικόνα

### 2.5.5 RGB Εικόνες

Μια RGB εικόνα μερικές φορές αναφέρεται σαν εικόνα πραγματικών χρωμάτων και αποθηκεύεται στο MatLab® σαν ένας πίνακας δεδομένων  $m \times n \times 3$  τα οποία καθορίζουν την κόκκινη, την πράσινη και την μπλε συνιστώσα για κάθε ξεχωριστό εικονοστοιχείο. Το χρώμα κάθε εικονοστοιχείου καθορίζεται από τον συνδυασμό του κόκκινου, του πράσινου και του μπλε που αποθηκεύονται σε κάθε χρωματικό επίπεδο στη θέση του εικονοστοιχείου. Οι RGB εικόνες αποθηκεύονται σαν 24-bit εικόνες, όπου καθεμία από τις χρωματικές συνιστώσες είναι 8-bit. Αυτό αποδίδει περίπου 16 εκατομμύρια χρώματα. Η ακρίβεια με την οποία μια κανονική εικόνα αναπαρίσταται μας οδήγησε στον όρο «εικόνα πραγματικών χρωμάτων».

Ένας πίνακας RGB στο MatLab μπορεί να είναι της τάξεως double, unit8, ή unit16. Η παρακάτω εικόνα αναπαριστά μια RGB εικόνα τάξεως double:





Εικ. 14 RGB Εικόνα και χρωματικές συνιστώσες (Red, Green, Blue)

Σε ένα RGB πίνακα τάξης double κάθε στοιχείο χρώματος αντιπροσωπεύει μια τιμή μεταξύ 0 και 1. Ένα εικονοστοιχείο με την τιμή κάθε συνιστώσας να είναι μηδέν (0,0,0), είναι μαύρο και ένα εικονοστοιχείο με τιμές (1,1,1) είναι άσπρο. Οι συνιστώσες των τριών χρωμάτων για κάθε εικονοστοιχείο αποθηκεύονται σε μια τρίτη διάσταση του πίνακα δεδομένων. Για παράδειγμα το κόκκινο, το πράσινο και το μπλε συστατικό του εικονοστοιχείου που βρίσκεται στη θέση (10,5) αποθηκεύεται στον RGB (10,5,1), RGB (10,5,2), RGB (10,5,3) πίνακα αντίστοιχα.



Εικ. 15 RGB Εικόνα



Εικ.16 Red Συνιστώσα



**Εικ.17 Green Εικόνα**

**Εικ.18 Blue Συνιστώσα**

Παρατηρούμε ότι για κάθε ξεχωριστό επίπεδο χρώματος περιέχει μια περιοχή άσπρου. Το άσπρο ανταποκρίνεται στις υψηλότερες τιμές κάθε ξεχωριστού χρώματος. Για παράδειγμα στο κόκκινο επίπεδο της εικόνας το άσπρο αντιπροσωπεύει την υψηλότερη συγκέντρωση τιμών καθαρού κόκκινου. Όταν το κόκκινο αναμιγνύεται με πράσινο ή μπλε εμφανίζονται γκρίζα εικονοστοιχεία. Οι μαύρες περιοχές στις γκρίζες εικόνες δείχνουν εικονοστοιχεία που δεν περιέχουν τιμές του κόκκινου.[9]

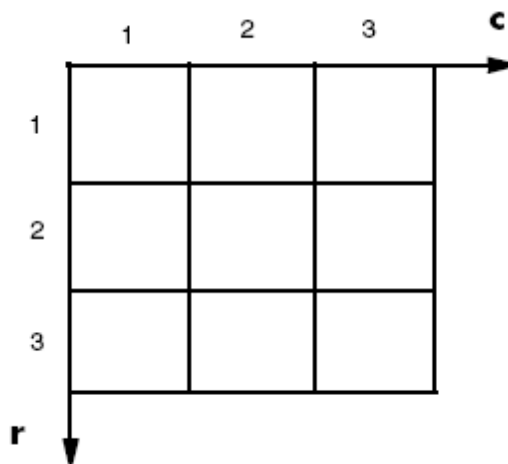
## ΚΕΦΑΛΑΙΟ 3

### 3.1 Συστήματα Συντεταγμένων

Μια θέση σε μια εικόνα μπορεί να εκφραστεί με διάφορα συστήματα συντεταγμένων, ανάλογα με το περιβάλλον. Τα κυριότερα συστήματα που εφαρμόζονται στην επεξεργασία εικόνας είναι το σύστημα συντεταγμένων των pixel (**Pixel Coordinates**) και το σύστημα συντεταγμένων επιπέδου (**Spatial Coordinates**).

#### **3.1.1 Σύστημα Συντεταγμένων των Pixel**

Γενικά, η πιο εύκολη μέθοδος για τον προσδιορισμό ενός σημείου σε μια εικόνα, είναι η χρήση του συστήματος συντεταγμένων των pixels. Σε αυτό το σύστημα συντεταγμένων, η εικόνα θεωρείται σαν ένα πλέγμα διακριτών στοιχείων, διατεταγμένα από το πάνω έως το κάτω μέρος της εικόνας και από αριστερά έως δεξιά όπως φαίνεται και στην παρακάτω εικόνα.



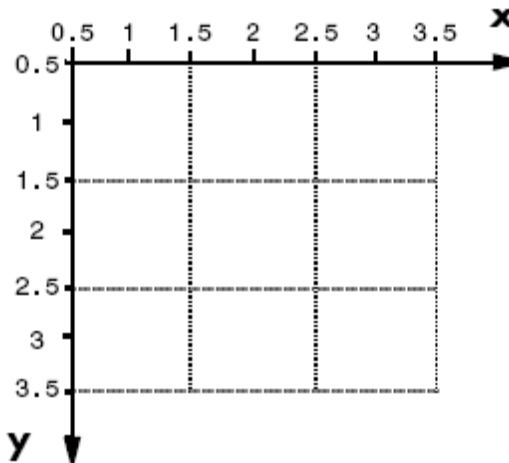
Σχ.1 Σύστημα Συντεταγμένων Επιπέδου

Για τις συντεταγμένες των pixel, η πρώτη συνιστώσα  $r$  (γραμμές) αυξάνει προς τα κάτω ενώ η δεύτερη συνιστώσα  $c$  (στήλες) αυξάνει προς τα δεξιά. Οι συντεταγμένες των pixels έχουν ακέραιες τιμές και εύρος μεταξύ ένα και του μήκους των στηλών ή των γραμμών. Μεταξύ των συντεταγμένων των pixel και των συντεταγμένων των στοιχείων ενός πίνακα, υπάρχει ένα προς ένα αντιστοιχία.

Για παράδειγμα, τα στοιχεία του pixel στην πέμπτη γραμμή, δεύτερη στήλη αποθηκεύονται στο στοιχείο (5,2) του πίνακα. [10]

### 3.1.2 Σύστημα Συντεταγμένων Επιπέδου

Στο σύστημα συντεταγμένων των pixels, ένα pixel θεωρείται σαν μια διακριτή μονάδα, το οποίο προσδιορίζεται από ένα μοναδικό ζευγάρι συντεταγμένων. Με βάση τα παραπάνω ένα σημείο με συντεταγμένες (5.3,2.2), δεν υφίσταται. Ωστόσο, είναι πολύ χρήσιμο να θεωρηθεί ένα pixel σαν ένα μικρό τετράγωνο. Με βάση αυτή την υπόθεση, ένα σημείο όπως είναι το (5.3,2.2) είναι διαφορετικό από το σημείο (5,2). Στο σύστημα συντεταγμένων επιπέδου, τα σημεία μιας εικόνας είναι θέσεις πάνω σε ένα επίπεδο, και περιγράφονται σε σχέση με τα  $x$  και  $y$  (όχι  $r$  και  $c$  όπως στο σύστημα συντεταγμένων των pixels). Η ακόλουθη φωτογραφία παρουσιάζει το σύστημα συντεταγμένων επιπέδου που χρησιμοποιείται στις εικόνες. Θα πρέπει να σημειωθεί ότι η συνιστώσα  $y$  αυξάνει προς τα κάτω.



Σχ.2 Σύστημα συντεταγμένων επιπέδου

Το σύστημα συντεταγμένων επιπέδου παρουσιάζει πολλά κοινά στοιχεία με το σύστημα συντεταγμένων των pixels. Για παράδειγμα, στο σύστημα συντεταγμένων επιπέδου το κεντρικό σημείο κάθε pixel έχει τις ίδιες συντεταγμένες στο σύστημα συντεταγμένων των pixels, για το ίδιο pixel. Ωστόσο, υπάρχουν κάποιες σημαντικές διαφορές. Στις συντεταγμένες των pixels, η πάνω αριστερή γωνία της εικόνας είναι η (1,1), ενώ στις συντεταγμένες επιπέδου το ίδιο σημείο είναι το (0.5,0.5). Η διαφορά αυτή οφείλεται στο ότι, το σύστημα συντεταγμένων των pixels είναι διακριτό, ενώ το σύστημα συντεταγμένων επιπέδου είναι συνεχές. Επίσης η πάνω αριστερή γωνία είναι πάντα το σημείο (1,1) στις συντεταγμένες των pixels, αλλά στο σύστημα συντεταγμένων επιπέδου μπορεί να προσδιοριστεί μια μη προεπιλεγμένη αρχή συντεταγμένων. Επίσης, η διάταξη της οριζόντιας και κατακόρυφης συνιστώσας είναι ανάστροφη στο σύστημα συμβόλων των δυο συστημάτων. Όπως αναφέρθηκε νωρίτερα, οι συντεταγμένες των pixels εκφράζονται σαν  $(r,c)$ , ενώ οι συντεταγμένες των σημείων του επιπέδου εκφράζονται σαν  $(x,y)$ . [11]

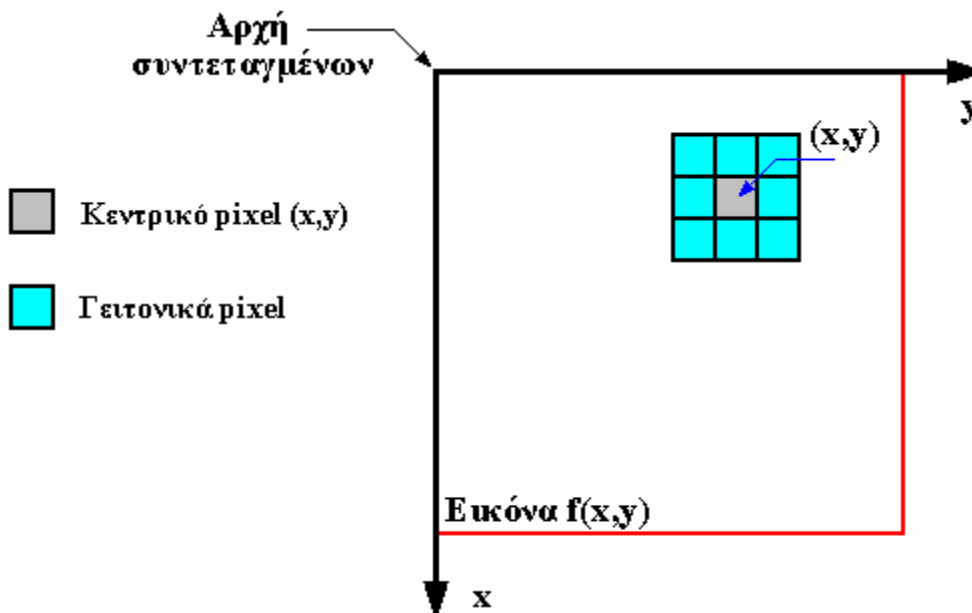
## 3.2 Spatial Domain

### 3.2.1 Εισαγωγή

Ο όρος **spatial domain** αναφέρεται στο σύνολο των pixel που συνθέτουν την εικόνα. Οι τεχνικές επεξεργασίας της εικόνας που ανήκουν στην κατηγορία του spatial domain επενεργούν άμεσα στα pixel της εικόνας. Οι spatial domain επεξεργασίες δηλώνονται από την ακόλουθη σχέση

$$g(x,y) = T[f(x,y)]$$

επεξεργασία και  $T$  είναι τελεστής που επενεργεί στην εικόνα  $f$ . Ο τελεστής ορίζεται από ένα συγκεκριμένο σύνολο γειτονικών σημείων ενός σημείου  $(x,y)$ . Η κύρια μέθοδος καθορισμού των γειτονικών σημείων του επιπέδου της εικόνας, του σημείου  $(x,y)$ , είναι η χρήση μιας τετράγωνης ή ορθογώνιας περιοχής κεντραρισμένη στο σημείο  $(x,y)$ , όπως φαίνεται και στην παρακάτω εικόνα.



Σχ. 3 Γειτονική περιοχή μεγέθους 3x3 του σημείου  $(x,y)$

Το κέντρο της περιοχής μετακινείται από pixel σε pixel, ξεκινώντας από την πάνω αριστερή γωνία της εικόνας, και καθώς κινείται, περικλείει διαφορετικά γειτονικά σημεία κάθε φορά. Ο τελεστής  $T$  εφαρμόζεται σε κάθε σημείο  $(x,y)$  για να προκύψει η έξοδος  $g$ , σε αυτή την περιοχή. Μόνο τα γειτονικά pixels χρησιμοποιούνται για τον υπολογισμό της νέας τιμής  $g$  του

σημείου  $(x,y)$ . Η διαδικασία στηρίζεται σε μια βασική εφαρμογή της ψηφιακής επεξεργασίας εικόνας την δυσδιάστατη συνέλιξη.[11]

### 3.2.2 Δυσδιάστατη συνέλιξη (2D Convolution)

Η δυσδιάστατη συνέλιξη είναι μια από τις θεμελιώδεις εφαρμογές για την ανάλυση της εικόνας. Η εφαρμογή αυτή προσδίδει μια νέα τιμή σε ένα δεδομένο pixel, η οποία προκύπτει από μια στατιστική ανάλυση των τιμών των γειτονικών pixel, μιας  $k \times k$  γειτονικής περιοχής του κεντρικού pixel. Οι τιμές των γειτονικών pixel τοποθετούνται μια τετραγωνική μήτρα, καλούμενη συχνά και ως μάσκα φιλτραρίσματος (**Filter Mask**) ή πυρήνας συνέλιξης (**Kernel**), και η ίδια η μάσκα εφαρμόζεται σε κάθε pixel μιας φωτογραφίας. Επιλέγοντας διαφορετικά μεγέθη μάσκας, διαφορετικές διαδικασίες μπορούν να εκτελεστούν. Οι μάσκες αυτές χρησιμοποιούνται σαν φίλτρα χαμηλών συχνοτήτων για την εξομάλυνση των απότομων μεταβολών στις τιμές των pixel και σαν φίλτρα υψηλών συχνοτήτων τονίζοντας τα χαρακτηριστικά της εικόνας.

### 3.2.3 Spatial Filtering

Η ανάλυση της εικόνας επικαλείται την διακριτή μορφή της δυσδιάστατης συνέλιξης. Αυτό καθορίζεται από την ακόλουθη σχέση μεταξύ των στοιχείων  $g(x, y)$  της αρχικής εικόνας, και των στοιχείων  $i h(\xi, n)$  του πυρήνα συνέλιξης **H**, και όπου τα στοιχεία της εικόνας που προκύπτει από την παραπάνω διαδικασία, όπου  $x,y,\xi,n$  ακέραιοι.

$$g_0(x, y) = \sum_{\xi=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} g_i(\xi, n)h(x - \xi, y - n)$$

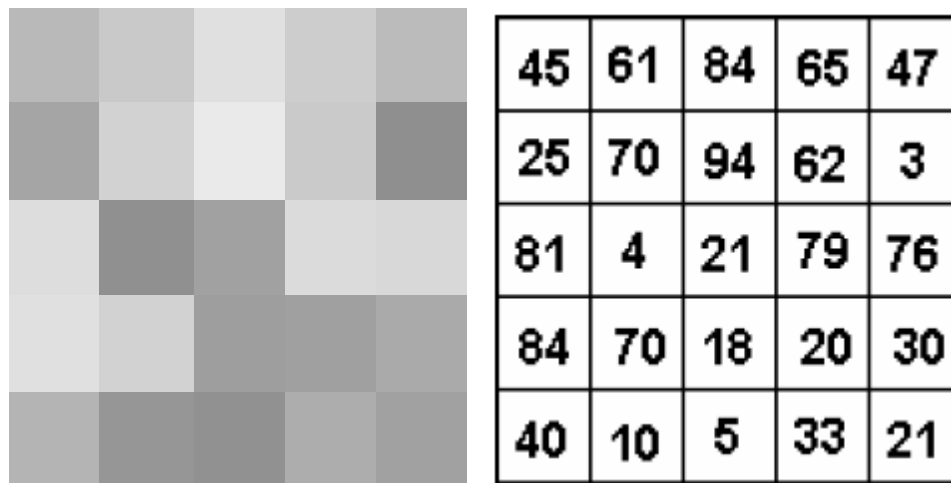
Οι συντελεστές του πυρήνα συνέλιξης **H** παριστάνουν μια ιδιαίτερη προσέγγιση της αναλυτικής μορφής της απόκρισης που περιγράφει το επιθυμητό φίλτρο. Οι πυρήνες στην πλειοψηφία τους είναι τετραφωνικοί πίνακες,  $k_x = k_y \equiv k$ , όπου  $k$  είναι μονός αριθμός και πολύ μικρότερος από την γραμμική διάσταση  $N$  της εικόνας. Η γενική έκφραση μπορεί να αποδοθεί απλούστερα από την ακόλουθη σχέση.

$$g_0(x, y) = \sum_{\xi=-(k-1)/2}^{(k-1)/2} \sum_{n=-(k-1)/2}^{(k-1)/2} g_i(\xi, n)h(x - \xi, y - n)$$

Στην πιο απλή μορφή μιας δυσδιάστατης λειτουργίας συνέλιξης, σε μια ψηφιακή εικόνα, χρησιμοποιείται ένας πυρήνας συνέλιξης. Οι πυρήνες συνήθως σχηματίζονται από μονό αριθμό γραμμών και στηλών υπό μορφή τετραγώνου. Η πιο συνήθεις μορφή είναι μεγέθους  $3 \times 3$ , αλλά

χρησιμοποιούνται και μεγαλύτερων διαστάσεων convolution kernels, ανάλογα με την εφαρμογή και το επιθυμητό αποτέλεσμα της επεξεργασίας. Η λειτουργία συνέλιξης εκτελείται ξεχωριστά σε κάθε pixel της αρχικής εικόνας και περιλαμβάνει τρεις διαδοχικές διαδικασίες. Η λειτουργία ξεκινάει με την επίστρωση του πυρήνα συνέλιξης στην αρχική εικόνα με τέτοιο τρόπο που το κεντρικό pixel του πυρήνα να αντιστοιχεί στη θέση κάθε pixel της αρχικής εικόνας. Το pixel αυτό ονομάζεται target pixel. Στη συνέχεια η ακέραιη τιμή κάθε pixel της αρχικής εικόνας, πολλαπλασιάζεται με την αντίστοιχη τιμή του επικαλυμμένου πυρήνα. Στο τρίτο βήμα, υπολογίζεται το άθροισμα των τιμών που προέκυψαν από το προηγούμενο βήμα. Τέλος, η τιμή του target pixel αντικαθίσταται από το αποτέλεσμα του αθροίσματος που καθορίστηκε στο τρίτο στάδιο.

Ακολουθεί ένα παράδειγμα που περιγράφει αναλυτικά τα στάδια της λειτουργίας της συνέλιξης. Έστω ένα κομμάτι μιας grayscale εικόνας, το οποίο έχει μέγεθος 25 pixel. Στην σχήμα 5 φαίνονται οι τιμές κάθε pixel της αρχικής εικόνας.



Σχ.4&5 Αρχική εικόνα και τιμές των pixel της αρχικής εικόνας

Έστω ο ακόλουθος πυρήνας συνέλιξης μεγέθους 3 x 3, που χρησιμοποιείται για τον τονισμό των χαρακτηριστικών της εικόνας.

-1	0	-1
0	5	0
-1	0	-1

Πυρήνας της συνέλιξης

Στο πρώτο στάδιο της λειτουργίας, ο πίνακας του πυρήνα επικαλύπτει την αρχική εικόνα έτσι ώστε το κεντρικό pixel του πυρήνα να αντιστοιχεί κάθε φορά, σε μια διαφορετική θέση ενός pixel κάθε φορά, όπως φαίνεται και στην παρακάτω εικόνα.

-1	0	-1	65	47
45	61	84	62	3
0	5	0	79	76
25	70	94	20	30
-1	0	-1	81	4
81	4	21	79	76
84	70	18	20	30
40	10	5	33	21

Σχ.6 Επικάλυψη της εικόνας από τον πυρήνα της συνέλιξης(1<sup>ο</sup> στάδιο)

Στη συνέχεια η τιμή κάθε pixel της αρχικής εικόνας, πολλαπλασιάζεται με την αντίστοιχη τιμή του επικαλυμμένου πυρήνα, και υπολογίζεται το άθροισμα όλων των πολλαπλασιασμών.

$$-1 * 45 + 0 * 61 - 1 * 84 + 0 * 25 + 5 * 70 + 0 * 94 - 1 * 81 + 0 * 4 - 1 * 21 = 119$$

Το αποτέλεσμα του αθροίσματος είναι η νέα τιμή του pixel της εικόνας.. Για να εκτελεστεί η συνέλιξη σε μια ολόκληρη εικόνα, πρέπει τα παραπάνω στάδια της λειτουργίας να

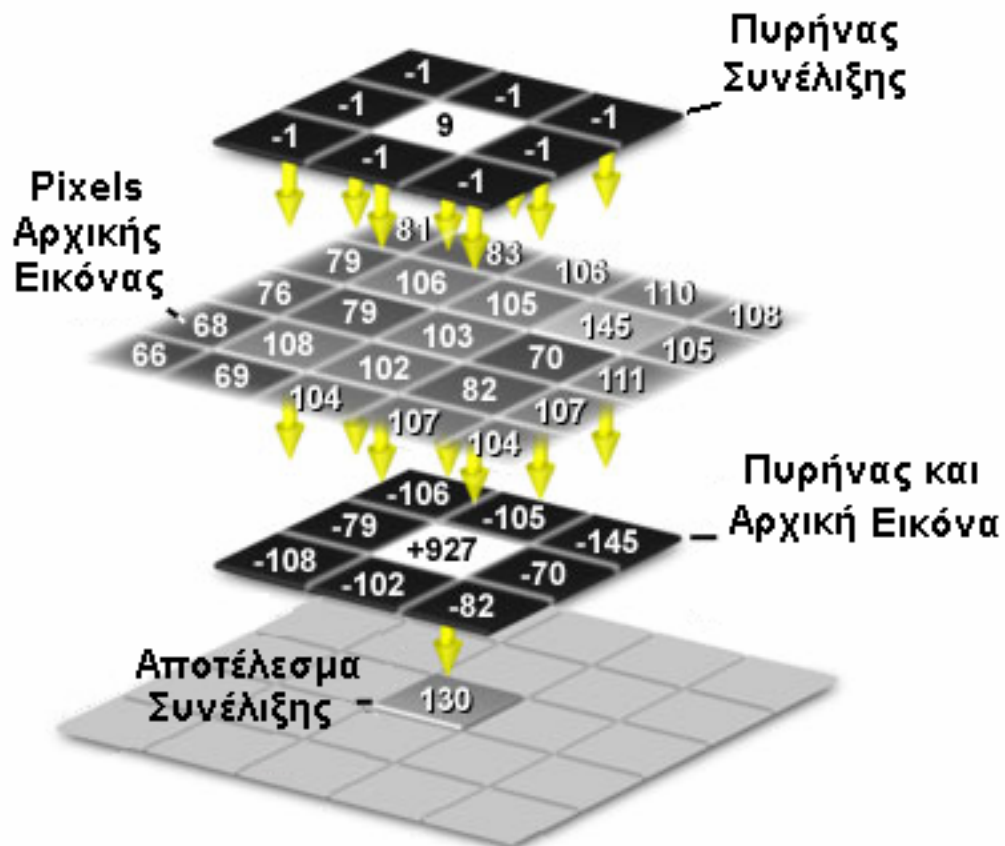


επαναληφθούν για κάθε pixel της αρχικής εικόνας. Το αποτέλεσμα της συνέλιξης για την αρχική εικόνα του παραδείγματος φαίνεται στην εικόνα που ακολουθεί

	119	261	82	
	-201	-117	250	
	203	-36	-23	

Σχ.7 Αποτέλεσμα συνέλιξης της αρχικής εικόνας

Συγκεντρωτικά τα βήματα της διαδικασίας της συνέλιξης φαίνονται στο ακόλουθο σχήμα.



Σχ.8 Διαδικασία συνέλιξης

Γενικά, οι αριθμητικές τιμές που χρησιμοποιούνται στους πυρήνες συνέλιξης είναι ακέραιοι αριθμοί με έναν διαιρέτη που μπορεί να ποικίλει ανάλογα με την επιθυμητή λειτουργία. Επίσης, επειδή πολλές διαδικασίες συνέλιξης οδηγούν σε αρνητικές τιμές, συχνά εφαρμόζονται αντισταθμισμένες τιμές για την επαναφορά μιας θετικής τιμής. Οι πίνακες των πυρήνων για 8 bit grayscale εικόνες συχνά περιορίζονται με διαιρέτες και αντισταθμίσεις που επιλέγονται έτσι ώστε όλες οι τιμές που προκύπτουν από τη διαδικασία της συνέλιξης να βρίσκονται εντός του διαστήματος τιμών 0-255 (Double Class). Η εφαρμογή της συνέλιξης σε μια εικόνα αλλοιώνει αναπόφευκτα τα pixels που βρίσκονται κοντά στα όρια της εικόνας. Ο αριθμός των pixel που αλλοιώνονται, αλλά και κατά πόσο είναι εμφανής η αλλοίωση αυτή, εξαρτάται από το μέγεθος του πυρήνα συνέλιξης. Όσο μεγαλύτερος είναι ο πίνακας του πυρήνα τόσο μεγαλύτερος είναι και ο αριθμός των pixel που αλλοιώνονται στα όρια της εικόνας, με αποτέλεσμα να γίνεται και περισσότερο εμφανής στην φιλτραρισμένη εικόνα. Το φαινόμενο αυτό είναι γνωστό ως Border Effect και συμβαίνει όταν το κεντρικό pixel του πυρήνα επικαλύπτει κάποιο από τα pixels που βρίσκονται στα άκρα της εικόνας. Σε αυτή την περίπτωση ένα τμήμα του πυρήνα βρίσκεται εκτός των ορίων της εικόνας, γεγονός που προκαλεί εσφαλμένα αποτελέσματα στους υπολογισμούς.

-1	?	0	?	-1	?				
0	?	5	45	0	61	84	65	47	
-1	?	0	25	-1	70	94	62	3	
		81	4	21	79	76			
		84	70	18	20	30			
		40	10	5	33	21			

Σχ.9 Border Effect

?	?	?	?	?
?				?
?				?
?				?
?	?	?	?	?

Σχ.10 Αποτέλεσμα συνέλιξης

Οι πυρήνες συνέλιξης είναι χρήσιμοι για ένα ευρύ φάσμα λειτουργιών της ψηφιακής επεξεργασίας εικόνας, συμπεριλαμβανομένης της ομαλοποίησης των μεταβολών στην ένταση των pixel στις εικόνες που περιέχουν θόρυβο (μέσος όρος), της αύξησης της έντασης των ακμών, χρησιμοποιώντας Laplacian, φίλτρα κλίσης κλπ. Επιπλέον, η τοπική αντίθεση μπορεί να ρυθμιστεί μέσω της χρήσης του μέγιστου, του ελάχιστου, ή των median φίλτρων, και οι εικόνες μπορούν να μετασχηματιστούν από το πεδίο του χώρου στο πεδίο των συχνοτήτων (μετασχηματισμός Fourier) με τους πυρήνες συνέλιξης.[12][13]

Η διαδικασία των γειτονικών pixel αποτελείται όπως αναφέρθηκε και στην προηγούμενη παράγραφο αποτελείται από τα ακόλουθα βήματα:

- 1) Καθορισμός κεντρικού σημείου (x, y).

- 2) Εκτέλεση της λειτουργίας, η οποία περιλαμβάνει μόνο τα pixel μιας προκαθορισμένης “γειτονιάς” του κεντρικού σημείου.
- 3) Το αποτέλεσμα της λειτουργίας είναι η απόκριση της διαδικασίας στο κεντρικό σημείο.
- 4) Η διαδικασία επαναλαμβάνεται για κάθε σημείο της εικόνας.

Η διαδικασία αυτή είναι γνωστή στην βιβλιογραφία με τον όρο **spatial filtering**. Εάν οι υπολογισμοί που εκτελούνται στα γειτονικά pixel είναι γραμμικοί (**Linear**), τότε η διαδικασία ονομάζεται **linear spatial filtering**, διαφορετικά είναι μη γραμμικοί και ονομάζεται **nonlinear spatial filtering**.

Οι γραμμικές διαδικασίες αποτελούνται από τους πολλαπλασιασμούς καθενός από τα γειτονικά pixel με τους αντίστοιχους συντελεστές της μάσκας και από το άθροισμα των αποτελεσμάτων για τη λήψη της απόκρισης, κάθε σημείου  $(x, y)$ . Όπως και το γραμμικό φίλτράρισμα, έτσι και το nonlinear spatial filtering βασίζεται στις λειτουργίες μιας περιοχής  $m \times n$  γειτονικών pixels, μεταφέροντας το κεντρικό σημείο σε κάθε pixel της εικόνας. Ωστόσο, αν και το linear spatial filtering είναι βασισμένο στον υπολογισμό του αθροίσματος των αποτελεσμάτων, που αποτελεί μια γραμμική διαδικασία, το nonlinear spatial filtering είναι βασισμένο σε μη γραμμικές λειτουργίες των γειτονικών pixels. Για παράδειγμα, μια μη γραμμική λειτουργία μπορεί να είναι η αντικατάσταση της τιμής του κεντρικού σημείου με την μέγιστη τιμή από τα γειτονικά pixel.[14][15]

### **3.3 Ανάλυση εικόνας στο πεδίο των συχνοτήτων (Frequency Domain)**

#### **3.3.1 Εισαγωγή**

Σε πολλές εφαρμογές επεξεργασίας εικόνας, το βασικό ενδιαφέρον ανάλυσης δεν συγκεντρώνεται στα χαρακτηριστικά των αντικειμένων ούτε και στην διάταξή τους, αλλά στη βασική δομή τους και στο βαθμό ακρίβειας που μπορεί να επεξεργαστεί μια σειρά πολλαπλών πανομοιότυπων αντικειμένων. Αυτά τα χαρακτηριστικά είναι πιο εύκολο να αναλυθούν στο πεδίο των συχνοτήτων και όχι στο πεδίο του χώρου  $(x,y)$  της εικόνας. Στο πεδίο των συχνοτήτων η ανάλυση γίνεται με την βοήθεια του μετασχηματισμού Fourier. Ο μετασχηματισμός αυτός προσφέρει μεγάλη ευελιξία στη σχεδίαση και στην υλοποίηση φίλτρων, για τον τονισμό των χαρακτηριστικών, την αποκατάσταση μιας εικόνας και σε ένα πλήθος άλλων πρακτικών εφαρμογών.

### 3.3.2 Διακριτός μετασχηματισμός Fourier (2-D Discrete Fourier Transform)

Ο δυσδιάστατος διακριτός μετασχηματισμός Fourier (DTF), μιας εικόνας  $f(x,y)$  διαστάσεων  $M \times N$ , για  $x = 0,1,2,\dots,M-1$  και  $y = 0,1,2,\dots,N-1$ , δίνεται από την ακόλουθη εξίσωση.[16]

$$F(u, v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi(\frac{ux}{M} + \frac{vy}{N})}$$

για  $u = 0,1,2,\dots,M-1$  και  $v = 0,1,2,\dots,N-1$ . Το πεδίο των συχνοτήτων είναι απλά ένα σύστημα συντεταγμένων, όπου η τιμή κάθε ζεύγους δίνεται από τη μεταβλητή  $F(u,v)$ , με τα  $u$  και  $v$  να είναι οι μεταβλητές συχνότητας. Το πεδίο συχνοτήτων είναι ανάλογο του spatial domain, όπου η τιμή κάθε ζεύγους δίνεται από τη μεταβλητή  $f(x,y)$ , με τα  $x$  και  $y$  να είναι οι spatial μεταβλητές. Ο Μετασχηματισμός Fourier αποτελεί έναν μαθηματικό τρόπο απεικόνισης μιας εικόνας, ο οποίος ουσιαστικά αντιπροσωπεύει το σύνολο των μέτρων και των φάσεων, των διαφόρων οπτικών συχνοτήτων, της εικόνας.

Ο αντίστροφος μετασχηματισμός Fourier δίνεται από την ακόλουθη εξίσωση.

$$f(x, y) = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) e^{j2\pi(\frac{ux}{M} + \frac{vy}{N})}$$

για  $x = 0,1,2,\dots,M-1$  και  $y = 0,1,2,\dots,N-1$ . Έτσι γνωρίζοντας την τιμή  $F(u,v)$ , μπορούμε να υπολογίσουμε την τιμή  $f(x,y)$  με τη χρήση του αντίστροφου μετασχηματισμού Fourier. Επειδή οι δείκτες των πινάκων στο MATLAB ξεκινούν από το 1 και όχι από το 0, οι τιμές  $F(1,1)$  και  $f(1,1)$  αντιστοιχούν στις μαθηματικές ποσότητες  $F(0,0)$  και  $f(0,0)$  του μετασχηματισμού καθώς και στον αντίστροφό του. Η τιμή του μετασχηματισμού στην αρχή των συντεταγμένων του πεδίου των συχνοτήτων  $F(0,0)$  ονομάζεται dc συνιστώσα του μετασχηματισμού Fourier. Αυτή η ορολογία προέρχεται από την ηλεκτροτεχνία, όπου το “dc” δηλώνει το συνεχές ρεύμα (ρεύμα μηδενικής συχνότητας). Εάν τώρα ορίσω στην DTF  $u=0$  και  $v=0$  τότε θα έχω.

$$\begin{aligned} F(0,0) &= \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi(\frac{0x}{M} + \frac{0y}{N})} \Rightarrow \\ F(0,0) &= \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \Rightarrow \\ F(0,0) &= MN * f(x, y) \end{aligned}$$

Από αυτόν τον τύπο προκύπτει ότι η τιμή  $F(0,0)$  είναι ίση με το άθροισμα όλων των τιμών της συνάρτησης  $f(x,y)$ . [17]

Η κύρια μέθοδος ανάλυσης του μετασχηματισμού Fourier είναι το φάσμα του, για παράδειγμα η ένταση της μεταβλητής  $F(u,v)$ , το οποίο απεικονίζεται σαν εικόνα. Θεωρώντας ότι οι μεταβλητές  $R(u,v)$  και  $I(u,v)$  αποτελούν την πραγματική και φανταστική συνιστώσα της μεταβλητής  $F(u,v)$  το συνολικό μέτρο των συχνοτήτων ορίζεται από την ακόλουθη σχέση:

$$|F(u, v)| = [R^2(u, v) + I^2(u, v)]^{\frac{1}{2}} \quad (1)$$

Η συνολική φάση των συχνοτήτων ορίζεται από την εξίσωση:

$$\varphi(u, v) = \tan^{-1} \left[ \frac{R^2(u,v)}{I^2(u,v)} \right]^{\frac{1}{2}} \quad (2)$$

Από τις εξισώσεις 1,2 προκύπτει η πολική μορφή της συνάρτησης  $F(u,v)$ :

$$|F(u, v)| = |F(u, v)| e^{-j\varphi(u,v)}$$

### **3.4 Ιστόγραμμα εικόνας**

Το ιστόγραμμα μιας εικόνας αποχρώσεων του γκρι περιέχει σημαντικές πληροφορίες για την εικόνα και για το λόγο αυτό είναι ένα από τα σημαντικότερα εργαλεία στην επεξεργασία ψηφιακών εικόνων. Μπορεί να χρησιμοποιηθεί για τη βελτιστοποίηση της εικόνας, την τροποποίηση των χαρακτηριστικών της, την μετατροπή της σε εικόνα με λιγότερες αποχρώσεις, την εξαγωγή χαρακτηριστικών της εικόνας κ.α.

Η βελτιστοποίηση μιας εικόνας που προκύπτει από την τροποποίηση του ιστογράμματος της είναι επιθυμητή σε πολλές εφαρμογές. Σε μια σημαντική ομάδα προβλημάτων μια εικόνα ενισχύεται τροποποιώντας την αντίθεση της ή/και τη δυναμική περιοχή των φωτεινοτήτων της. Για παράδειγμα μια τυπική εικόνα ακόμα και μη υποβιβασμένη, θα φαίνεται καλύτερα αν οι ακμές της τονιστούν περισσότερο. Επίσης, αν για κάποιο λόγο μια εικόνα με μεγάλη δυναμική περιοχή καταγραφτεί σε ένα μέσο με μικρή δυναμική περιοχή, η αντίθεση της και κατά συνέπεια οι λεπτομέρειες της ελαττώνονται ιδιαίτερα στις πολύ φωτεινές και πολύ σκοτεινές περιοχές. Αυξάνοντας την τοπική αντίθεση και μειώνοντας τη δυναμική περιοχή μπορούμε να ενισχύσουμε την ποιότητα αυτής της εικόνας. Σε μια άλλη κατηγορία προβλημάτων, μια υποβιβασμένη εικόνα μπορεί να βελτιωθεί μειώνοντας τον υποβιβασμό της. Παραδείγματα υποβιβασμού είναι το θόλωμα, ο τυχαίος θόρυβος βάθους και ο θόρυβος στίγματος.

Το ιστόγραμμα μιας εικόνας είναι εύκολο να προσδιορισθεί. Εκφράζει τη κατανομή των αποχρώσεων του γκρι στην εικόνα και στις περισσότερες περιπτώσεις καθορίζει την εικόνα. Ένα ιστόγραμμα είναι ένα γράφημα που στον οριζόντιο άξονα έχει τις φωτεινότητες από 0-255 και στον κατακόρυφο άξονα το πλήθος των εικονοστοιχείων που έχουν κάθε φωτεινότητα. Ανάλογα με την εφαρμογή, ο κατακόρυφος άξονας μπορεί να κανονικοποιηθεί με βάση τη μέγιστη τιμή του ιστογράμματος.

Στην περίπτωση που η μέγιστη τιμή του ιστογράμματος αντιστοιχηθεί με το ένα, τότε το κανονικοποιημένο ιστόγραμμα που προκύπτει μας δίνει ουσιαστικά την κατανομή πυκνότητας-πιθανότητας των επιπέδων του γκρι στην εικόνα. Σύμφωνα με τα παραπάνω, το ιστόγραμμα  $h(k)$ , μιας ψηφιακής εικόνας  $I$ , διαστάσεων  $N \times M$ , με αποχρώσεις του γκρι στην περιοχή  $[0, L - 1]$  εικόνας μπορεί να προσδιοριστεί με τη βοήθεια της σχέσης:

$$h(k) = \begin{cases} h(k) + 1, & \text{αν } I(i, k) = k \\ h(k), & \text{διαφορετικά} \end{cases}$$

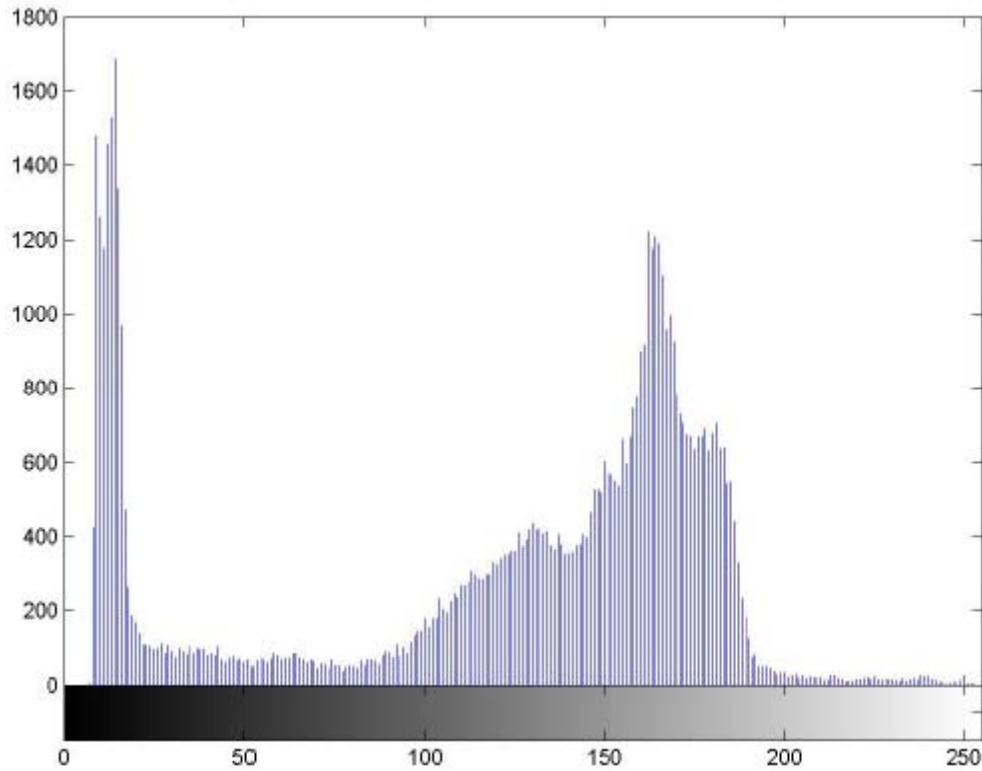
όπου  $(i, j) \in \{0, 1, \dots, L-1\}$ . Από τα παραπάνω καταλήγω ότι:

$$\sum_{k=0}^{L-1} h(k) = N * M$$

Παρακάτω παρουσιάζεται ένα παράδειγμα ιστογράμματος μιας τυπικής grayscale εικόνας.



**Grayscale εικόνα**



**Και το ιστόγραμμα της**

Οι τεχνικές τροποποίησης ιστογράμματος μετασχηματίζουν την αρχική ζώνη φωτεινότητας με τη βοήθεια μιας γραμμικής ή μη γραμμικής συνάρτησης μετασχηματισμού  $T(h(k))$ . Με αυτόν τον τρόπο, οι φωτεινότητες της εικόνας απεικονίζονται σε μια νέα ζώνη και προκύπτει έτσι ένα νέο ιστόγραμμα από τη σχέση:

$$h'(k) = T(h(k)), k = 0, 1, 2, \dots, L - 1$$

έτσι ώστε  $h'(k) \in \{0, \dots, L-1\}$ . Η διαδικασία αυτή συναντάται σε πολλά προγράμματα επεξεργασίας εικόνας ως μια τεχνική τροποποίηση ιστογράμματος μέσω καμπύλων.[19]

## ΚΕΦΑΛΑΙΟ 4

### 4.1 Τμηματοποίηση εικόνας (Image Segmentation)

Η εύρεση κατωφλίων (Thresholds) είναι μία από τις ισχυρότερες τεχνικές για την τμηματοποίηση εικόνων που απεικονίζονται με αποχρώσεις του γκρι. Η εφαρμογή των τεχνικών κατωφλίου σε μια ψηφιακή εικόνα βασίζεται στην υπόθεση ότι τα εικονοστοιχεία του αντικειμένου (προσκήνιο - foreground) μπορούν να διαχωριστούν από τα εικονοστοιχεία του φόντου (παρασκήνιο - background) με βάση τις τιμές των αποχρώσεων τους. Για τον προσδιορισμό των δύο κύριων αποχρώσεων του γκρι με τη χρήση ενός κατωφλίου (bi-level ή global thresholding) η εικόνα μετατρέπεται σε εικόνα δύο κύριων αποχρώσεων με τη χρήση μιας μοναδικής τιμής κατωφλίου για όλα τα εικονοστοιχεία της. Η τεχνική αυτή χρησιμοποιείται ευρέως για τη μετατροπή μιας εικόνας σε δυαδική. Η προσέγγιση αυτή δίνει ικανοποιητικά αποτελέσματα σε περιπτώσεις όπως των εικόνων κειμένου, όπου είναι γνωστό εκ των προτέρων ότι η εικόνα περιέχει μόνο δύο κύριες αποχρώσεις του γκρι (π.χ. άσπρο, μαύρο). Συνεπώς οι μέθοδοι ενός συνολικού κατωφλίου πρέπει να εφαρμόζονται σε εικόνες όπου είναι ξεκάθαρη η διαφορά μεταξύ των αποχρώσεων του προσκήνιου με το παρασκήνιο. Έχουν προταθεί αρκετές τεχνικές για την αυτόματη επιλογή ενός καθολικού κατωφλίου.

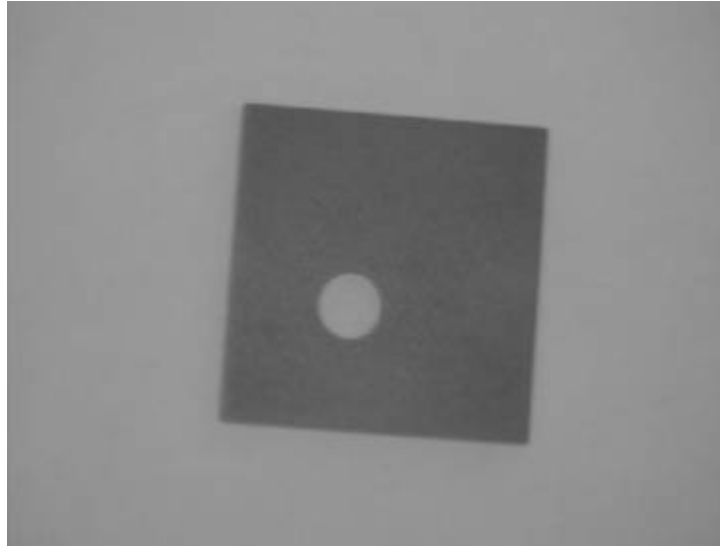
Για την τμηματοποίηση περίπλοκων εικόνων είναι αναγκαία η χρήση τεχνικών επιλογής κατωφλίων για πολλά επίπεδα (multilevel thresholding). Στη γενική περίπτωση αυτό δεν είναι απλό πρόβλημα. Υπάρχουν αρκετές δυσκολίες για την επιλογή κατωφλίων σε εικόνες με πολλά αντικείμενα και οφείλονται κυρίως στις κατανομές των αποχρώσεων, στα μικρά και στα αλληλεπικαλυπτόμενα αντικείμενα. Ο προσδιορισμός του βέλτιστου αριθμού των τελικών αποχρώσεων είναι ένα σημαντικό πρόβλημα. Αν και υπάρχουν αλγόριθμοι προσδιορισμού του βέλτιστου αριθμού των αποχρώσεων στις περισσότερες των τεχνικών ο αριθμός αυτός θεωρείται εκ των προτέρων γνωστός.

Αρκετές τεχνικές έχουν προταθεί για την αντιμετώπιση των δυσκολιών που προαναφέρθηκαν. Οι περισσότερες από αυτές βασίζονται στην ταύτιση των ακμών (edge matching) και την ταξινόμηση. Οι μέθοδοι αυτές είναι εφαρμόσιμες σε εικόνες με καλές ακμές και δεν βασίζονται στο ιστόγραμμα των αποχρώσεων. Σε μεγάλο αριθμό περιπτώσεων τα αντικείμενα που περιέχονται σε μια εικόνα με πολλά επίπεδα του γκρι, έχουν κοντινή μεταξύ τους φωτεινότητα. Για παράδειγμα, στην εικόνα 21 φαίνεται ότι το αντικείμενο έχει φωτεινότητα κοντά στο μαύρο ενώ το φόντο θα έχει φωτεινότητα υψηλή. Στο σχήμα 6 φαίνεται το ιστόγραμμα του εικόνα 21. Όπως παρατηρούμε, στο ιστόγραμμα αυτό μπορούμε να διακρίνουμε δύο "λόφους" (hills) και μια κοιλάδα (valley) μεταξύ τους. Είναι φανερό ότι ο αριστερός λόφος αντιπροσωπεύει τα εικονοστοιχεία που πιθανότατα ανήκουν στο

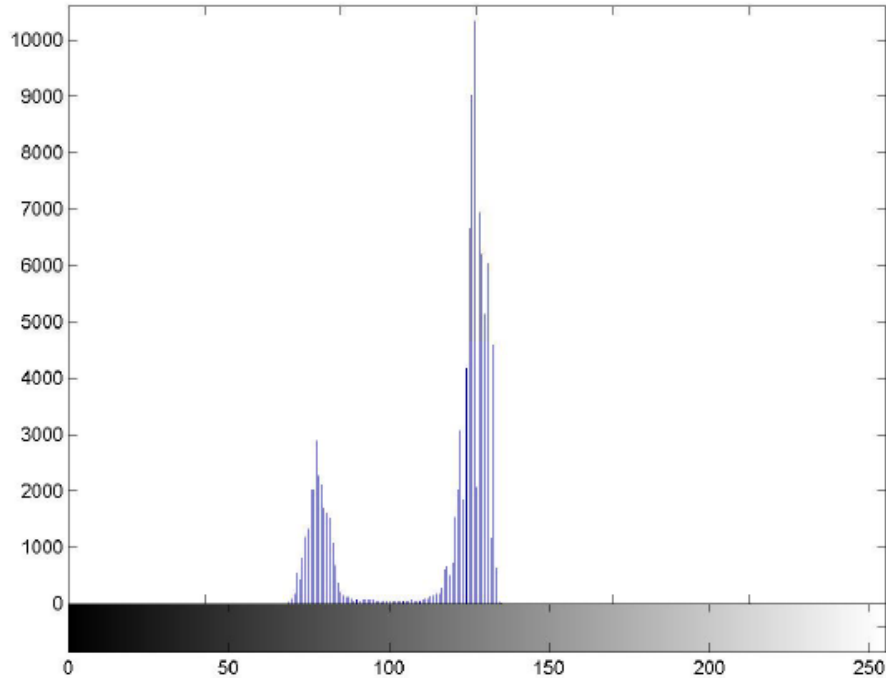


αντικείμενο, ενώ αντίθετα ο δεξιός λόφος αντιστοιχεί σε εικονοστοιχεία που ανήκουν πιθανότατα στο φόντο.

Αυτό που δεν μπορούμε με σαφήνεια να πούμε είναι που ανήκουν τα εικονοστοιχεία της κοιλάδας. Για τα εικονοστοιχεία αυτά δεχόμαστε ότι μπορούν να ανήκουν και στους χαρακτήρες αλλά και στο φόντο. Όμως, είναι σαφές ότι όσο πιο κοντά στον αριστερό λόφο βρίσκεται ένα εικονοστοιχείο τόσο μεγαλύτερη είναι η πιθανότητα να αποτελεί εικονοστοιχείο αντικειμένου και το αντίστροφο.



**Εικ.21 Μια εικόνα Grayscale**

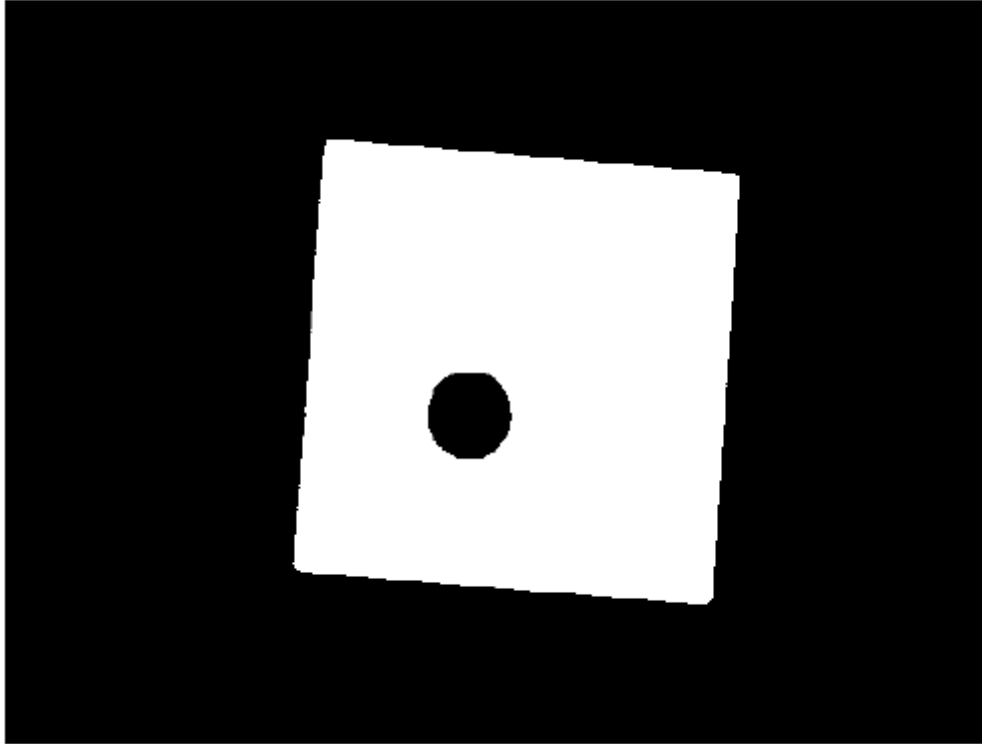


Σχ.6 Το ιστόγραμμα της εικ.21

Επειδή οι εικόνες διαφέρουν σε χαρακτηριστικά η μια από την άλλη δεν είναι δυνατόν με χρήση απλού thresholding να γίνει ο διαχωρισμός background από foreground. Μπορεί να γίνει εύκολα αντιληπτό εάν μια εικόνα ανήκει σε αυτήν την κατηγορία κοιτώντας το ιστόγραμμα έντασης της. Στην περίπτωση όπου τα pixels μιας εικόνας είναι εύκολο να διαχωριστούν από τα pixels του background με βάση την τιμή της έντασης των εικονοστοιχείων τότε στο ιστόγραμμα της εικόνας θα είναι ευδιάκριτη μια κορυφή που θα αντιστοιχεί στα pixels του foreground. Έτσι μια τιμή για το threshold μπορεί εύκολα να καθοριστεί η οποία θα απομονώνει αυτά τα pixels. Αν προσδιοριστεί με μια κατάλληλη μέθοδο μια τιμή κατωφλίου  $T$  στην περιοχή της κοιλάδας, τότε μπορούμε να διαχωρίσουμε την εικόνα και να την μετατρέψουμε σε δυαδική σύμφωνα με την ακόλουθη σχέση:

$$bw(i, j) = \begin{cases} 0, & \text{αν } I(i, j) > T \\ 1, & \text{αν } I(i, j) \leq T \end{cases}$$

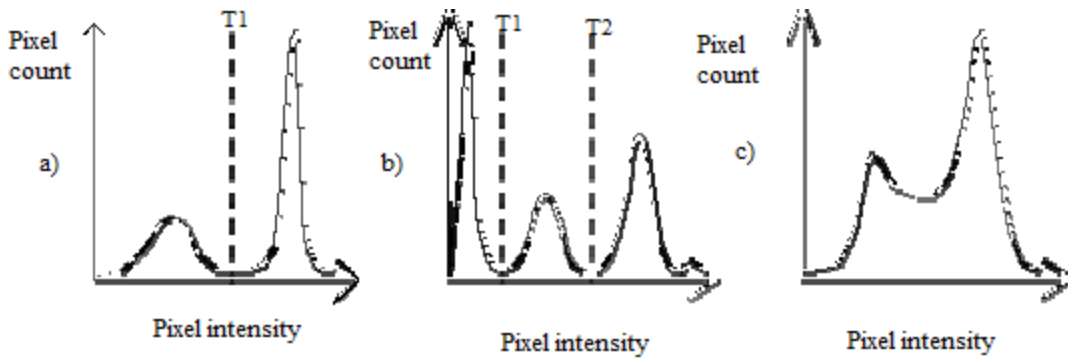
όπου  $(i, j)$  οι συντεταγμένες του εικονοστοιχείου  $I(i, j)$ . Έτσι, στην τελική δυαδική εικόνα το αντικείμενο θα έχει εικονοστοιχεία με τιμή 1 ενώ τα εικονοστοιχεία του φόντου θα έχουν τιμή 0. Στο σχήμα 7 παρουσιάζεται το αποτέλεσμα της εφαρμογής της σχέσης στην εικόνα 21 για τιμή κατωφλίου  $T = 120$ .



Σχ.7 Δυαδική μορφή της Grayscale εικόνας.

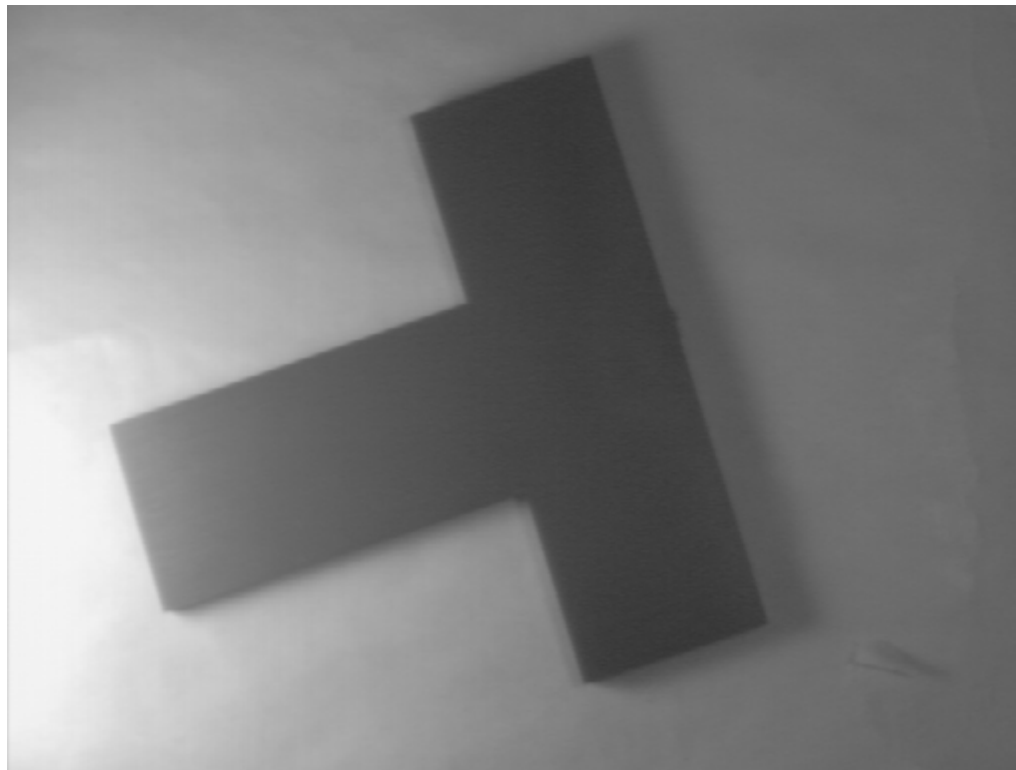
Η διαδικασία που αναπτύχθηκε παραπάνω ανήκει στις μεθόδους του απόλυτου κατωφλίου. Είναι όμως πιθανό να έχουμε εικόνες που περιέχουν αντικείμενα με διαφορετικές μεταξύ τους φωτεινότητες ή εικόνες με ανομοιογενή φωτισμό. Σε αυτές τις περιπτώσεις είναι σχεδόν αδύνατον η εφαρμογή ενός απόλυτου thresholding να αποφέρει αποδεκτά αποτελέσματα. Σε τέτοιες περιπτώσεις απαιτείται η εφαρμογή πιο πολύπλοκων μεθόδων thresholding.

Το σχήμα 8 δείχνει μερικά χαρακτηριστικά ιστογράμματα μαζί με τις επιλογές thresholding.

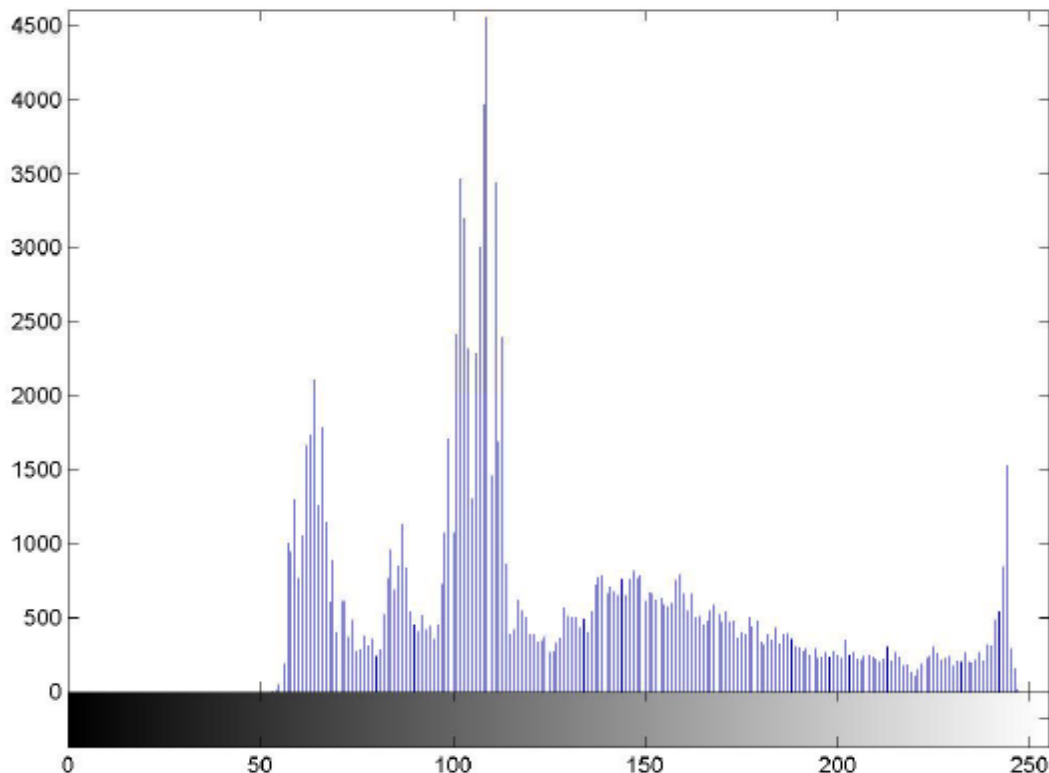


Σχ.8 Μορφές ιστογράμματος Grayscale εικόνας

Στην πρώτη περίπτωση του σχήματος 8 απεικονίζεται η πιο απλή μορφή του ιστογράμματος όπου η κατανομή των pixels του background και foreground εμφανίζονται σαν δυο ευδιάκριτες κορυφές. Έτσι ο διαχωρισμός τους μπορεί να γίνει εύκολα με την χρήση ενός απόλυτου threshold  $T_1$ . Στην δεύτερη περίπτωση το ιστόγραμμα είναι πιο πολύπλοκο. Υποθέτοντας ότι τα pixels του foreground βρίσκονται στην μεσαία κορυφή, μεταξύ των δυο κοιλάδων, επιλέγονται δυο τιμές threshold  $T_1$  και  $T_2$  από τα όρια αυτής της κορυφής.. Στην Τρίτη περίπτωση οι δύο κορυφές μπλέκονται μεταξύ τους και είναι σχεδόν αδύνατον να γίνει ο διαχωρισμός με την χρήση ενός και μόνο threshold. Παρακάτω εικόνα 22 παρουσιάζεται μια εικόνα με ανομοιογενή φωτισμό και στο σχήμα 9 το ιστόγραμμα της εικόνας.

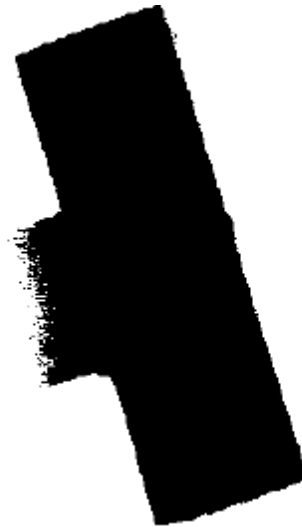


Εικ.22 Grayscale εικόνα με ανομοιογενή φωτισμό.

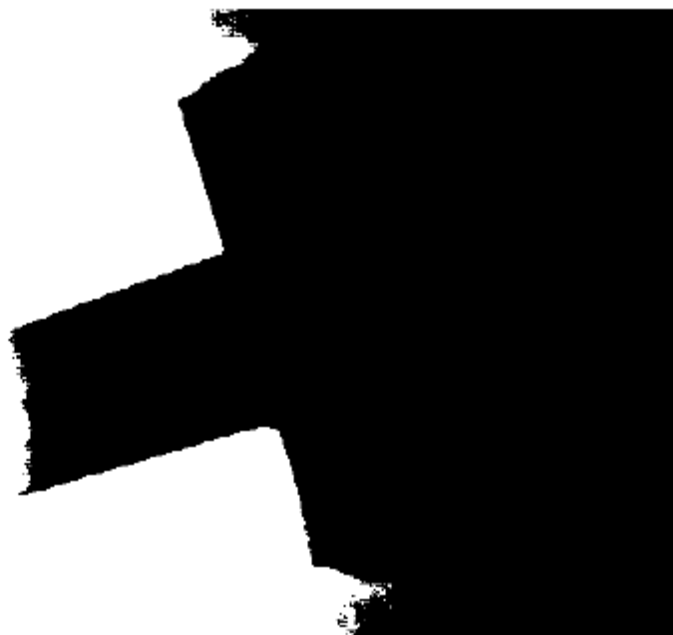


Σχ.9 Το ιστόγραμμα της εικ.22

Λόγω του ανομοιογενή φωτισμού της εικόνας οι δυο κορυφές συμπίπτουν και ο διαχωρισμός τους με καθολικό threshold δεν δίνει ικανοποιητικά αποτελέσματα. Στις ακόλουθες εικόνες παρουσιάζεται το αποτέλεσμα της χρήσης απλού thresholding με τιμές 80 και 120 αντιστοίχως. Είναι εμφανές ότι σε τέτοιες περιπτώσεις απαιτείται η χρήση προσαρμοστικού thresholding.[19][20][21]



Εικ.23 Δυαδική εικόνα της εικόνας 22 με κατώφλι  $T=80$ .



Εικ.24 Δυαδική εικόνα της εικόνας 22 με κατώφλι  $T=120$ .

## 4.2 Μορφολογική επεξεργασία εικόνας (Morphological Processing)

### 4.2.1 Dilation

Η διαδικασία του dilation είναι μια από τις δυο βασικές διαδικασίες στον τομέα μαθηματική μορφολογικής επεξεργασίας, η άλλη είναι το erosion. Μπορεί να εφαρμοστεί είτε σε δυαδικές εικόνες (binary images), είτε σε grayscale εικόνες. Όταν εφαρμοστεί σε μια δυαδική

εικόνα θα έχει σαν αποτέλεσμα να μικρύνει το μέγεθος των αντικείμενων που βρίσκονται στο foreground.

#### 4.2.2 Περιγραφή λειτουργίας Dilation

Η διαδικασία του dilation για να λειτουργήσει χρειάζεται δυο πίνακες δεδομένων. Ο πρώτος είναι η εικόνα που πρόκειται να γίνει dilate και ο δεύτερος είναι (συνήθως αρκετά μικρότερος) ονομάζεται στοιχείο δόμησης (structural element) συμφωνά με τον οποίο θα γίνει και το dilation της εικόνας. Ο μαθηματικός ορισμός του dilation για δυαδικές εικόνες είναι: Ας επιθέσουμε ότι  $X$  είναι ένα σύνολο ευκλειδείων συντεταγμένων που αντιστοιχούν στην αρχική δυαδική εικόνα και  $K$  είναι το σύνολο των συντεταγμένων του στοιχείου δόμησης. Το  $Kx$  είναι το  $K$  με το pixel-origin του στο σημείο  $x$ . Κατόπιν το dilation της εικόνας  $X$  από στοιχείο δόμησης  $K$ , είναι απλά το σύνολο όλων των σημείων  $X$  έτσι ώστε το  $Kx$  να είναι ένα υποσύνολο του  $X$ . Ο μαθηματικός ορισμός για την διαδικασία του dilation σε grayscale εικόνες είναι ίδιος εκτός από τον τρόπο με τον οποίο προκύπτει το σύνολο συντεταγμένων που σχετίζονται με την αρχική εικόνα. Σαν παράδειγμα dilation σε binary εικόνα, υποθέστε ότι το στοιχείο δόμησης έχει μέγεθος  $3 \times 3$  και σχήμα τετραγώνου, με την pixel-origin στο κέντρο του όπως φαίνεται Σχήμα 10. Σημειώστε ότι στο επόμενο σχήμα και στα υπόλοιπα του κεφαλαίου όλα τα pixels των αντικειμένων στο foreground θα απεικονίζονται με τιμή 1 και τα pixels του background με 0.

1	1	1
1	1	1
1	1	1

Set of co-ordinate points -  $\{(-1,-1),(0,1),(1,-1),(-1,0),(0,0),(1,0),(-1,1),(0,1),(1,1)\}$

Σχ.10 Structural Elements  $3 \times 3$  σχήματος τετραγώνου

Για να υπολογίσουμε το αποτέλεσμα του dilation στην αρχική binary εικόνα με το παραπάνω SE, θα πρέπει να εξετάσουμε κάθε ένα από τα pixels του background. Για κάθε background pixel (από εδώ και στο εξής input pixel) τοποθετείται το SE πάνω από την αρχική εικόνα έτσι ώστε το pixel-origin του SE να συμπίπτει με τις συντεταγμένες του input pixel. Εάν έστω και για ένα pixel του SE, το αντίστοιχο pixel της εικόνας είναι foreground pixel, τότε το input pixel θέτεται στην τιμή του foreground pixel. Εάν όλα τα pixels που αντιστοιχούν στο SE ανήκουν στο background τότε το input pixel αφήνεται ως έχει .

Παραδείγματος χάριν το  $3 \times 3$  τετραγωνικό SE που παρουσιάστηκε παραπάνω όταν εφαρμοστεί σε μια binary εικόνα θα έχει σαν αποτέλεσμα να τεθούν σε τιμές των foreground pixels όλα τα background pixels τα οποία γειτονεύουν με foreground pixels. Τέτοια pixels συνήθως βρίσκονται στις άκρες των φωτεινών περιοχών, έτσι πρακτικά το αποτέλεσμα αυτής

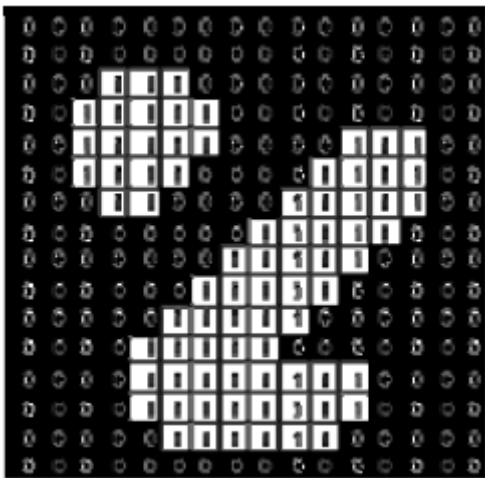
της διαδικασίας είναι να μεγαλώνουν οι περιοχές του foreground και να μικραίνουν οι μαύρες περιοχές μέσα σε αυτές.

Όταν εφαρμόσουμε dilation σε μια grayscale εικόνα τότε το pixel-origin του SE όταν μεταφέρεται πάνω στην εικόνα παίρνει την μέγιστη τιμή από τα pixel τα οποία καλύπτονται από το SE.

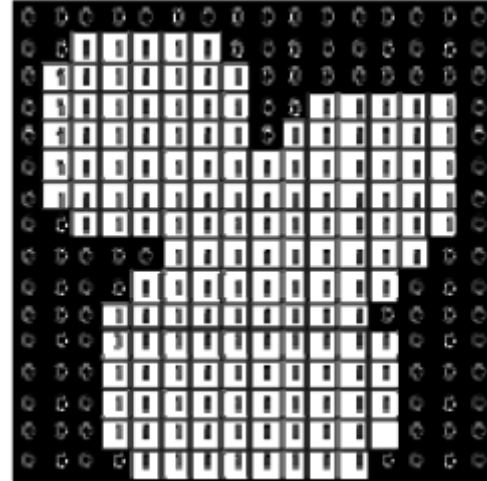
Η διαδικασία του dilation είναι αντίθετη του erosion, δηλαδή dilation των εικονοστοιχείων του background ισοδυναμεί με dilation των εικονοστοιχείων του foreground.

### 4.2.3 Εφαρμογή του dilation

Οι περισσότερες εφαρμογές αυτής της διαδικασίας προϋποθέτουν η εικόνα που τίθεται προς επεξεργασία να έχει δυαδική μορφή (binary image). Μια τέτοια εικόνα μπορεί να προέλθει από μια grayscale εικόνα στην οποία έχει γίνει threshold. Το στοιχείο δόμησης που θα χρησιμοποιήσουμε στο παράδειγμα μας είναι ένα τετράγωνος πίνακας 3x3, σαν αυτόν που περιγράψαμε παραπάνω. Στην εικόνα 26 παρουσιάζονται τα αποτελέσματα που είχε η διαδικασία του dilation στην εικόνα 25 με την χρήση του προαναφερθέντος SE.



Εικ.25 Αρχική δυαδική εικόνα.

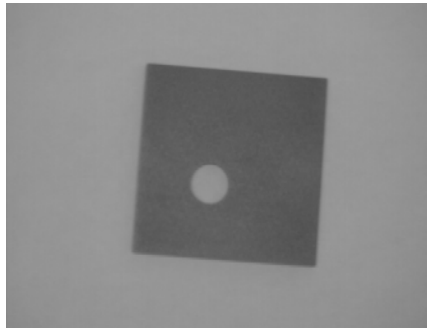


Εικ.26 Παράδειγμα Dilation με την χρήση με την χρήση ενός τετραγώνου SE 3x3.

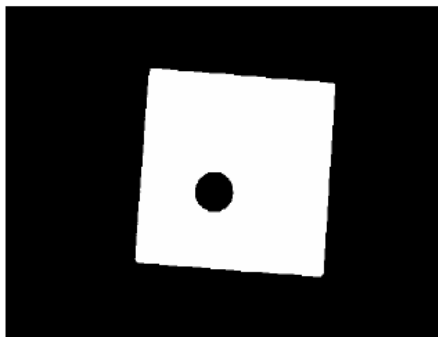
Το 3x3 τετράγωνο SE είναι πιθανώς το πιο κοινό στοιχείο δόμησης που χρησιμοποιείται στις εφαρμογές του Dilation. Ένα μεγαλύτερο SE θα προκαλούσε πιο ακραία αποτελέσματα εάν και τα ίδια αποτελέσματα θα μπορούσαν να επιτευχθούν με την εφαρμογή επαναλαμβανόμενων διαδικασιών dilation με την χρήση ενός μικρότερου SE αλλά παρομοίου σχήματος. Στις



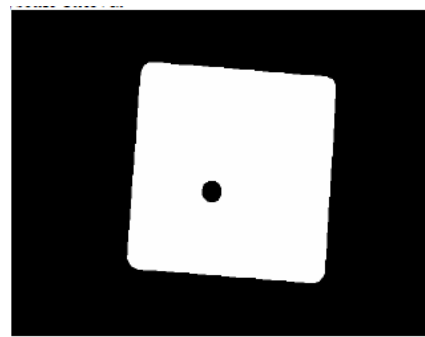
περιπτώσεις όπου απαιτείται η χρήση ενός μεγαλύτερου SE συνήθως επιλέγεται ένας SE σχήματος δίσκου.



Εικ.27 Grayscale εικόνα



Εικ.28 Threshold εικόνα.



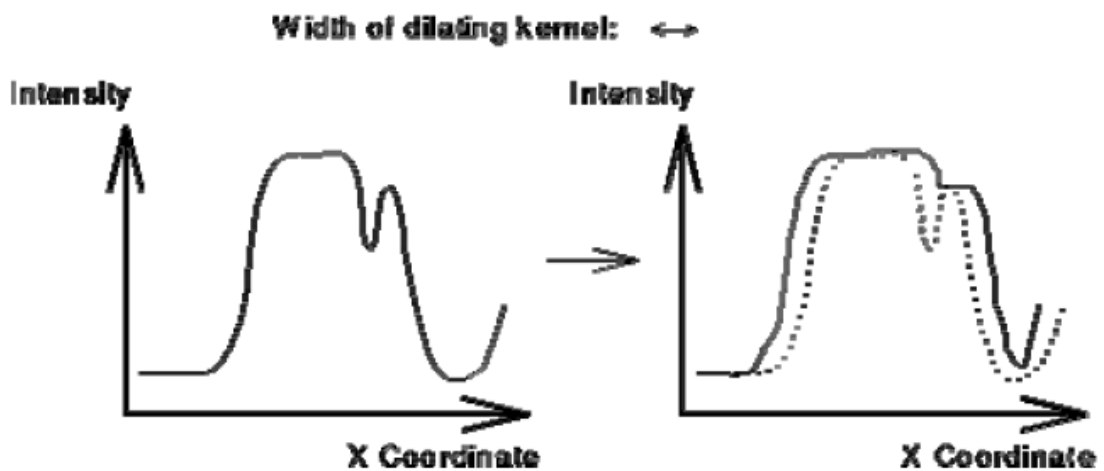
Εικ.29 Dilated εικόνα με SE δίσκου διαμέτρου 11 pixel.

Στην εικόνα 28 παρουσιάζεται το αποτέλεσμα που έχει το thresholding στην αρχική εικόνα 27.

Η εικόνα 29 μας δείχνει την επίδραση που έχει η εφαρμογή διπλού dilation στην thresholded εικόνα με την χρήση ενός SE σχήματος δίσκου και ακτίνας 11 εικονοστοιχείων. Παρατηρείται ότι οι άκρες του τετραγώνου έχουν στρογγυλέψει. Γενικά όταν εφαρμοστεί dilation σε συνδυασμό με ένα SE σχήματος δίσκου οι κυρτές άκρες των αντικειμένων θα στρογγυλέψουν ενώ οι κοίλες άκρες θα διατηρηθούν όπως έχουν.

Η διαδικασία του dilate μπορεί να γίνει κατευθυντική με τη χρησιμοποίηση ενός λιγότερο συμμετρικού στοιχείου δόμησης, όπως ένα SE το οποίο θα έχει 10 pixels μήκος και 1 pixel ύψους. Ομοίως, ένα 3x3 τετράγωνο SE με pixel-origin το μεσαίο στην κορυφαία σειρά θα «διέστειλε» περισσότερο την κάτω περιοχή του αντικειμένου παρά την άνω.

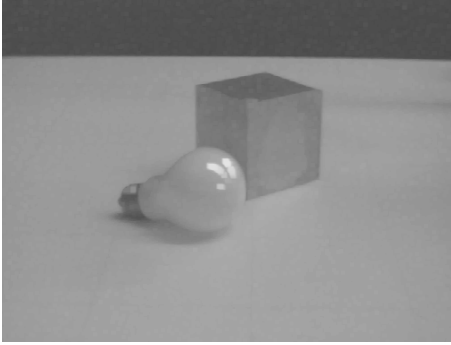
Η εφαρμογή του dilation σε μια grayscale εικόνα με την χρήση ενός επιπέδου SE σχήματος δίσκου, γενικά θα αυξήσει την φωτεινότητα στην εικόνα. Φωτεινές περιοχές που περιβάλλονται από σκοτεινές θα μεγεθυνθούν ενώ οι σκοτεινές περιοχές που περιβάλλονται φωτεινές θα συρρικνωθούν. Τα μικρά σκοτεινά σημεία στις εικόνες θα εξαφανιστούν καθώς θα πάρουν την τιμή της έντασης που έχουν τα γειτονικά σημεία και τα μικρά φωτεινά σημεία θα μεγαλώσουν. Η επίδραση που έχει το dilation είναι πιο εμφανής σε περιοχές όπου η ένταση των τιμών αλλάζει απότομα, ενώ οι περιοχές όπου η τιμή της έντασης είναι ομοιόμορφη δεν επηρεάζονται σημαντικά (εκτός από τις άκρες αυτών των περιοχών). Στο σχήμα 11 βλέπουμε ένα κάθετο οριζόντιο τμήμα μιας grayscale εικόνας και το αποτέλεσμα που έχει το dilation αυτής με την χρήση ενός επιπέδου SE σχήματος δίσκου.



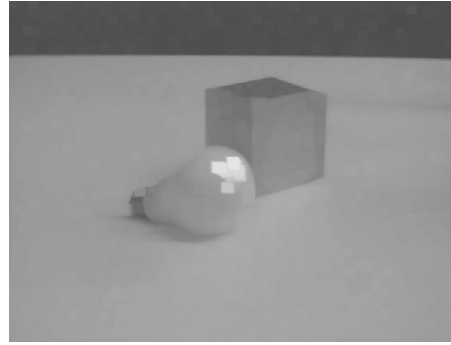
Σχ.11 Grayscale dilation που χρησιμοποιεί ένα SE σχήματος δίσκου. Τα διαγράμματα απεικονίζουν ένα κάθετο τμήμα της Grayscale εικόνας.



Εικ.30 Αρχική εικόνα



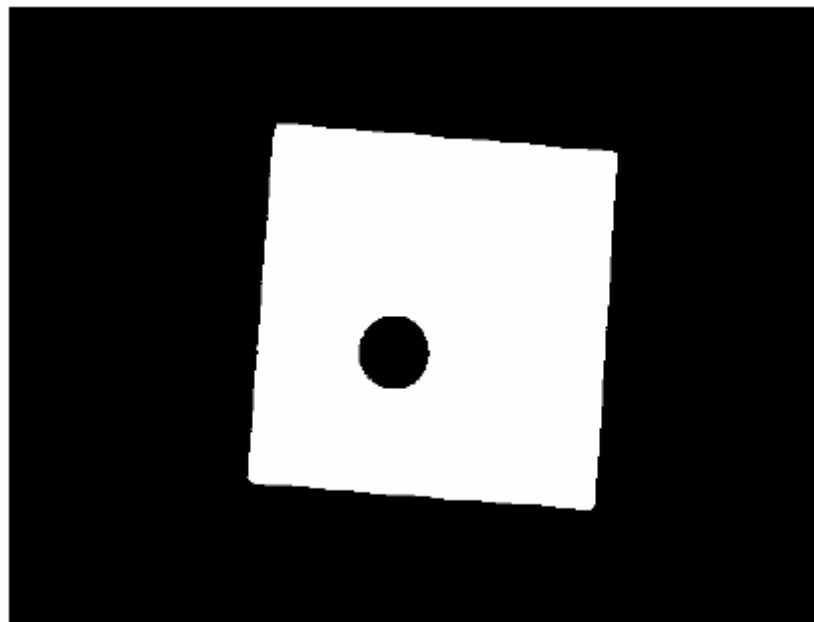
**Εικ.31** Η εικόνα μετά από 2 επαναλήψεις του dilation.



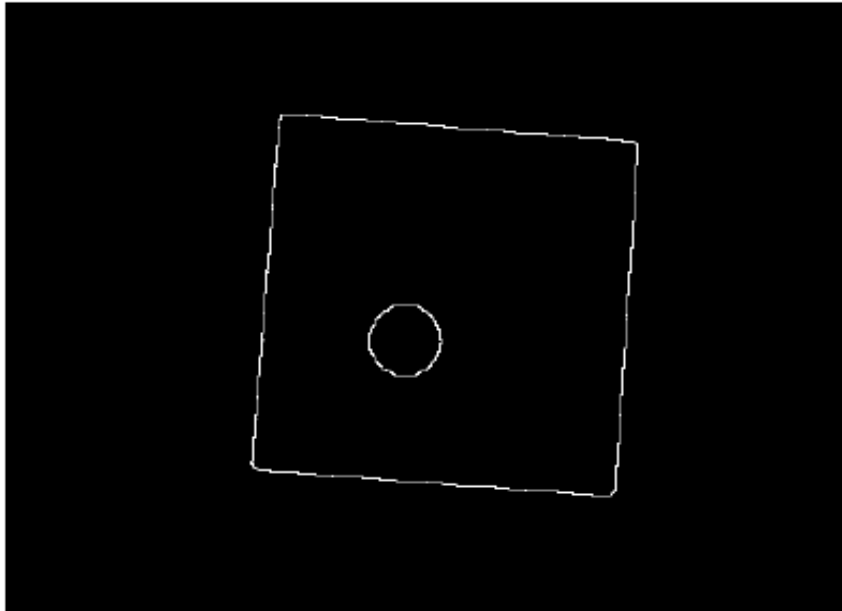
**Εικ.35** Η εικόνα μετά από 5 επαναλήψεις του dilation.

Η εικόνα 31 παρουσιάζει τα αποτελέσματα της διαδικασίας του dilation σε μια grayscale εικόνα (εικ.30). Σε αυτήν την εφαρμογή χρησιμοποιήθηκε ένα επίπεδο SE 3x3 και η διαδικασία του dilation επαναλήφθηκε δυο φορές. Τα χαρακτηριστικά στην επιφάνεια της λαμπάς έχουν μεγεθυνθεί και έχουν πάρει ένα στρογγυλό σχήμα. Η σκούρα περιοχή μέσα στον κύβο έχει μικρύνει αφού οι περιβάλλουσες περιοχές είναι πιο φωτεινές. Τελικά, οι περιοχές έχουν αποκτήσει μια ομοιομορφία στο εσωτερικό τους με την εξαφάνιση των σκοτεινών σημείων.

Η διαδικασία του dilation μπορεί επίσης να χρησιμοποιηθεί και για εφαρμογές ανίχνευσης ακρών αντικειμένων. Αφού, εφαρμόσουμε την διαδικασία του dilation στην φωτογραφία που επεξεργαζόμαστε υστέρτα αφαιρούμε από την dilated εικόνα την αρχική. Αυτό που απομένει είναι μονό τα αποτελέσματα που είχε το dilation στην αρχική εικόνα και κατ' επέκταση οι άκρες του αντικειμένου.[22][23]



**Εικ.33** Dilated εικόνα της εικ.28.



Εικ.34 Αποτέλεσμα αφαίρεσης αρχικής εικόνας από την dilated.

## **4.3 Erosion**

Η διαδικασία του erosion είναι μια από τις δυο βασικές διαδικασίες στον τομέα της μαθηματικής μορφολογικής επεξεργασίας, η άλλη είναι το dilation (διαστολή). Μπορεί να εφαρμοστεί είτε σε δυαδικές εικόνες (binary images), είτε σε grayscale εικόνες. Όταν εφαρμοστεί σε μια δυαδική εικόνα θα έχει σαν αποτέλεσμα να μικρύνει το μέγεθος των αντικείμενων που βρίσκονται στο foreground.

### **4.3.1 Περιγραφή λειτουργίας Erosion**

Η διαδικασία του erosion για να λειτουργήσει χρειάζεται δυο πίνακες δεδομένων. Ο πρώτος είναι η εικόνα που πρόκειται να γίνει erode και ο δεύτερος είναι (συνήθως αρκετά μικρότερος) ονομάζεται στοιχείο δόμησης ( structural element) συμφωνά με τον οποίο θα γίνει και το erosion της εικόνας.

Ο μαθηματικός ορισμός του erosion για δυαδικές εικόνες είναι: Ας υποθέσουμε ότι  $X$  είναι ένα σύνολο ευκλειδίων συντεταγμένων που αντιστοιχούν στην αρχική δυαδική εικόνα, και  $K$  είναι το σύνολο των συντεταγμένων του στοιχείου δόμησης. Το  $K_x$  είναι το  $K$  με το pixel origin του στο σημείο  $x$ . Κατόπιν το erosion της εικόνας  $X$  από το στοιχείο δόμησης  $K$  είναι απλά το σύνολο όλων των σημείων  $X$  έτσι ώστε το  $K_x$  να είναι ένα υποσύνολο  $X$ .

Ο μαθηματικός ορισμός για την διαδικασία του erosion σε grayscale εικόνες είναι ίδιος εκτός από τον τρόπο με τον οποίο προκύπτει το σύνολο συντεταγμένων που σχετίζονται με την αρχική εικόνα.

Σαν παράδειγμα erosion σε binary εικόνα, υποθέστε ότι το στοιχείο δόμησης είναι ένα 3x3 τετράγωνο, με το pixel-origin στο κέντρο του όπως φαίνεται στο παρακάτω σχήμα (σχ.12).

1	1	1
1	1	1
1	1	1

Set of co-ordinate points -  $\{(-1,-1),(0,1),(1,-1),(-1,0),(0,0),(1,0),(-1,1),(0,1),(1,1)\}$

Σχ.12 Ένα 3x3 τετράγωνο Structural Element

Για να υπολογίσουμε το αποτέλεσμα του erosion στην αρχική binary εικόνα με το παραπάνω SE, θα πρέπει να εξετάσουμε κάθε ένα από τα pixels του foreground. Για κάθε foreground pixel τοποθετείται το SE πάνω από την αρχική εικόνα έτσι ώστε το pixel-origin του SE να συμπίπτει με τις συντεταγμένες του input pixel. Εάν για κάθε pixel του SE, το αντίστοιχο pixel της εικόνας είναι foreground pixel, τότε το input pixel αφήνεται όπως έχει.. Εάν όμως κάποιο από τα pixels που αντιστοιχούν στο SE ανήκουν στο background τότε το input pixel θέτεται και αυτό σαν background.

Παραδείγματος χάριν το 3x3 τετράγωνο SE που παρουσιάστηκε παραπάνω όταν εφαρμοστεί σε μια binary εικόνα θα έχει σαν αποτέλεσμα να αφαιρεθούν οποιαδήποτε εικονοστοιχεία, τα οποία ανήκουν στο foreground αλλά δεν περιβάλλονται από αλλά εικονοστοιχεία του foreground. Τέτοια εικονοστοιχεία βρίσκονται στις άκρες των λευκών περιοχών του foreground, έτσι αυτή η διαδικασία έχει σαν αποτέλεσμα την σμίκρυνση του μεγέθους των αντικειμένων ( οι «τρύπες» μέσα στα αντικείμενα μεγαλώνουν).

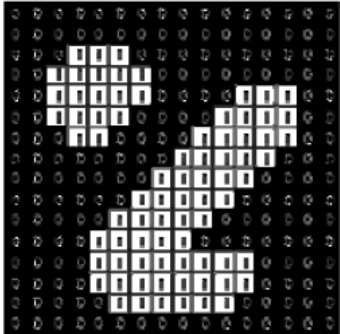
Όταν εφαρμόσουμε erode σε μια grayscale εικόνα τότε το pixel-origin του SE όταν μεταφέρεται πάνω στην εικόνα παίρνει την ελάχιστη τιμή από τα pixel τα οποία καλύπτονται από το SE.

Η διαδικασία του erode είναι αντίθετη του dilate, δηλαδή erosion των εικονοστοιχείων του foreground ισοδυναμεί με dilation των εικονοστοιχείων του background.

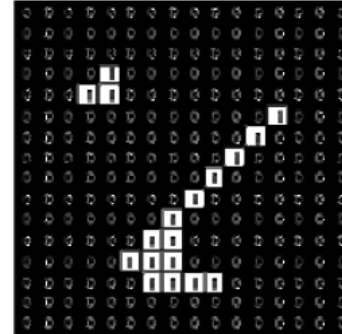
#### 4.3.2 Εφαρμογή του erosion

Οι περισσότερες εφαρμογές αυτής της διαδικασίας προϋποθέτουν η εικόνα που τίθεται προς επεξεργασία να έχει δυαδική μορφή (binary image). Μια τέτοια εικόνα μπορεί να προέλθει από μια grayscale εικόνα που έχει γίνει thresholded χρησιμοποιήση εικόνας thresholding.

Το 3 x 3 τετράγωνο SE είναι ένα από το πιο κοινά χρησιμοποιούμενα στοιχειά δόμησης στην διαδικασία του erosion, αλλά μπορούν να χρησιμοποιηθούν και αλλά ανάλογα με την εφαρμογή. Ένα μεγαλύτερο SE θα έχει σαν αποτέλεσμα ένα ακραίο αποτέλεσμα erosion, αν και το ίδιο αποτέλεσμα μπορεί να επιτευχθεί χρησιμοποιώντας ένα μικρότερο SE, ίδιου σχήματος, πολλές φορές. Στις εφαρμογές όπου χρησιμοποιείται μεγάλο SE είναι σύνηθες να επιλέγεται ένα, κατά προσέγγιση, κυκλικού σχήματος SE εν αντίθεση με ένα τετραγώνου σχήματος.

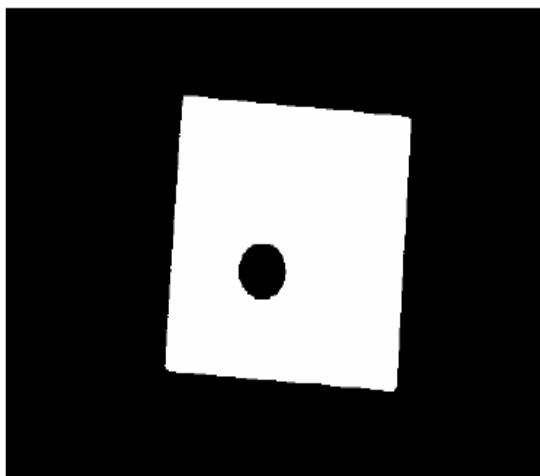


Εικ.35 Αρχική Binary εικόνα

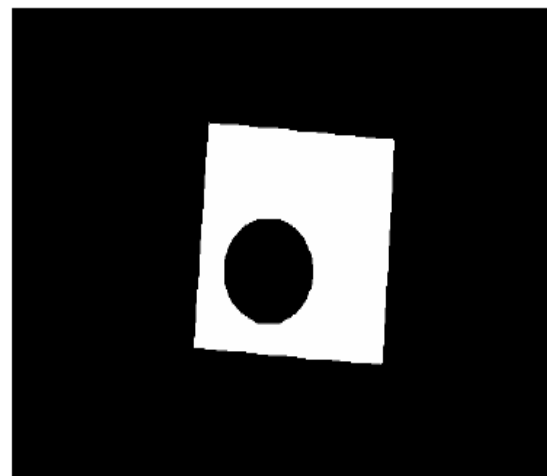


Εικ.36 Παράδειγμα Erosion με την χρήση ενός τετραγώνου SE 3x3.

Στις εικόνες 37 και 38 φαίνεται το αποτέλεσμα του erosion με την χρήση ενός structural element με κυκλικό σχήμα και διάμετρο 11 pixels. Είναι εμφανές ότι ενώ το μέγεθος του αντικειμένου έχει μικρύνει και η τρυπά στο εσωτερικό του έχει μεγαλώσει, το σχήμα του αντικειμένου διατηρείται. Γενικά η χρήση ενός σφαιρικού SE μπορεί να αλλοιώσει λίγο τις άκρες των αντικειμένων στρογγυλεύοντας τις γωνιές, αλλά θα διατηρηθεί το σχήμα του αντικειμένου.



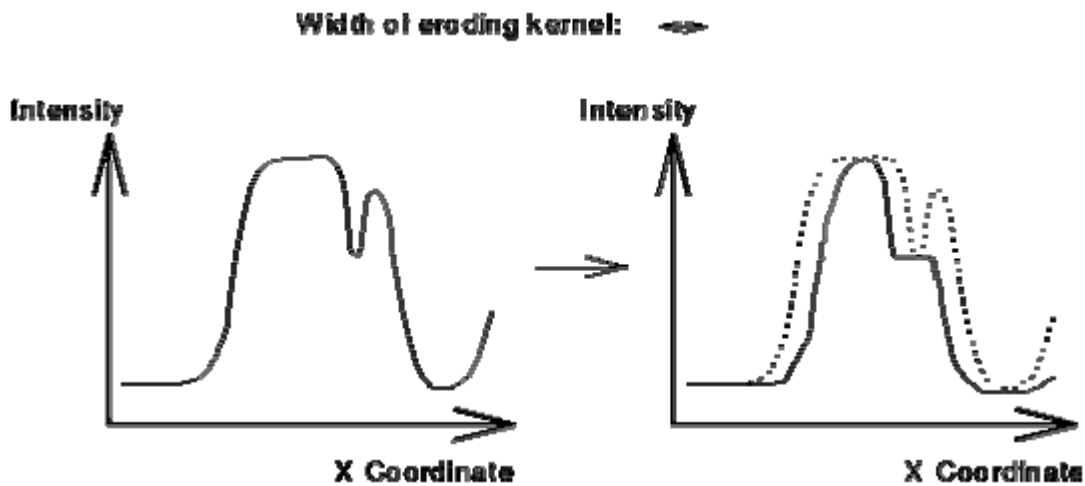
Εικ.37 Αρχική εικόνα



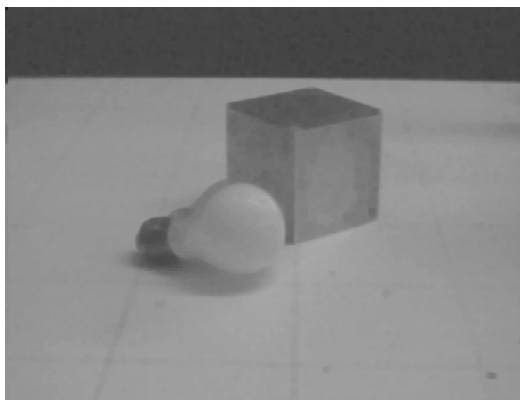
Εικ.38 Eroded εικόνα με SE δίσκο διαμέτρου 11 pixel.

Το dilation μπορεί να γίνει και κατευθυντικό με τη χρησιμοποίηση μιας λιγότερο συμμετρικής δομής SE. π.χ. ένα στοιχείο δόμησης που είναι έχει 10 εικονοστοιχεία μήκος και 1 εικονοστοιχείο ύψος θα προκαλέσει erosion μονό στην οριζόντια κατεύθυνση. Ομοίως, 3X3 τετραγωνικό στοιχείο δόμησης με pixel-origin το μεσαίο στην κορυφαία σειρά αντί για το pixel στο κέντρο, εάν χρησιμοποιηθεί για την διαδικασία του erode θα έχει καλύτερα αποτελέσματα στην κάτω περιοχή του αντικειμένου παρά στην άνω.

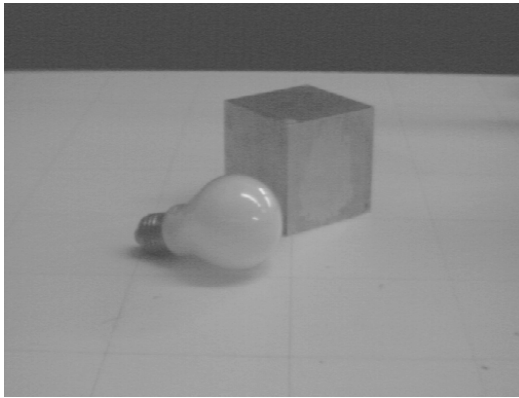
Όταν εφαρμόσουμε grayscale erosion με ένα επίπεδο δίσκο SE θα “σκουράνει” την εικόνα. Φωτεινές περιοχές που περιβάλλονται από σκοτεινές περιοχές θα συρρικνωθούν. Τα μικρά φωτεινά σημεία στις εικόνες θα εξαφανιστούν καθώς θα υποστούν erosion και εξομοιωθούν με τις σκούρες γύρω περιοχές, ενώ οι μικρές σκούρες περιοχές θα μεγαλώσουν σε μέγεθος. Η επίδραση που έχει το erosion είναι μεγαλύτερη σε περιοχές όπου η ένταση των τιμών αλλάζει απότομα, ενώ οι περιοχές όπου η τιμή της έντασης είναι ομοιόμορφη δεν επηρεάζονται σημαντικά. Στις εικόνες 37 και 38 βλέπουμε ένα κάθετο οριζόντιο τμήμα μιας grayscale εικόνας και το αποτέλεσμα που έχει το erosion αυτής με την χρήση ενός επιπέδου SE σχήματος δίσκου. Παρατηρούμε ότι με αυτό το SE αποκόπτονται οι αιχμές και μεγαλώνουν οι κοιλάδες.



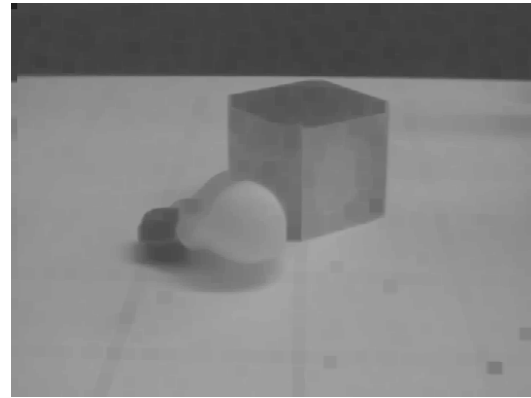
Σχ.13 Grayscale erosion που χρησιμοποιεί ένας SE σχήματος δίσκου. Τα διαγράμματα απεικονίζουν ένα κάθετο τμήμα της Grayscale εικόνας.



**Εικ.39 Αρχική εικόνα**



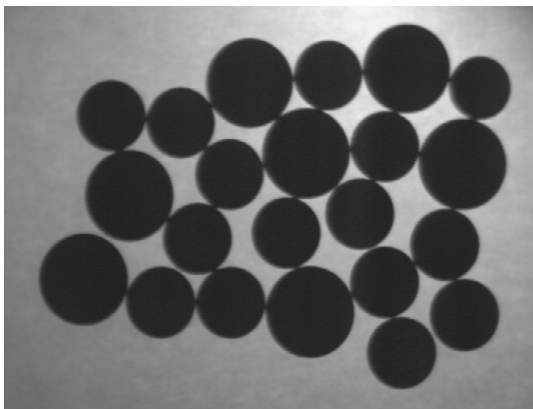
**Εικ.40 Διπλό erosion με SE 3x3 τετράγωνο**



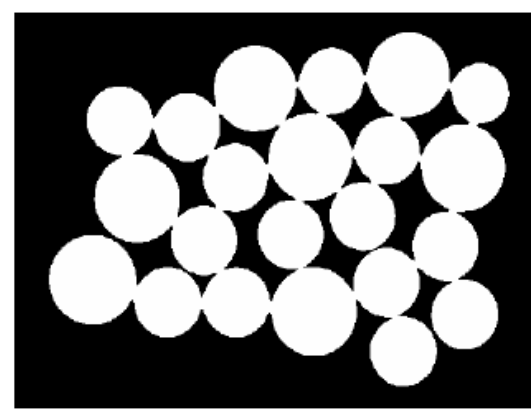
**Εικ.41 ErosionX5 με το ίδιο SE 3x3 τετράγωνο.**

Η εικόνα 40 έχει προκύψει από διπλή εφαρμογή erosion στην αρχική εικόνα (εικ.40) με την χρήση ενός τετράγωνου SE 3x3. Παρατηρούμε ότι τα έντονα χαρακτηριστικά έχουν εξαφανιστεί ενώ οι πιο οι περιοχές έχουν μεγαλύτερη ομοιομορφία λόγω της εξάλειψης των πιο φωτεινών σημείων και τέλος το σώμα του κύβου έχει αυξηθεί σε μέγεθος δεδομένου ότι είναι σκοτεινότερο από το περιβάλλον του. Στην επόμενη εικόνα (εικ.41) φαίνεται το αποτέλεσμα αν στην αρχική εικόνα εφαρμόσουν erosion πέντε φορές με το ίδιο SE.

Υπάρχουν πολλές εφαρμογές του erosion. Μια από αυτές είναι ο χωρισμός κάποιων αντικειμένων σε μια δυαδική εικόνα έτσι ώστε να είναι δυνατή η καταμέτρηση τους μέσω κάποιου αλγορίθμου.

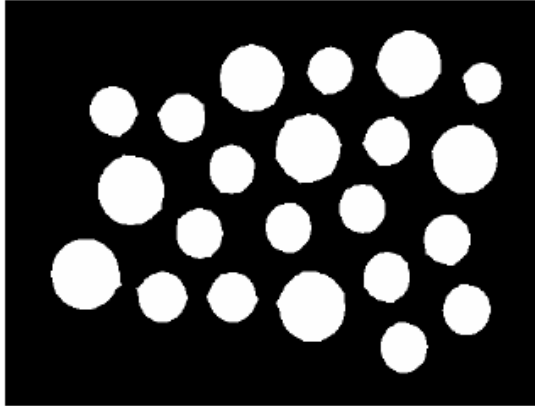


**Εικ.42 Αρχική εικόνα**

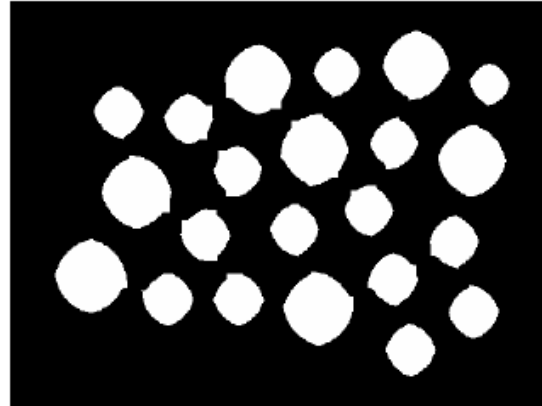


**Εικ.43 Thresholded Εικόνα**





Εικ.44 Erosion με δίσκο διαμέτρου 11 pixels



Εικ.45 Erosion με SE τετράγωνο 9x9

Στην εικόνα 42 βλέπουμε κάποια νομίσματα (σκούρες περιοχές) σε φωτεινό background. Εφαρμόζουμε global thresholding και παίρνουμε το αποτέλεσμα που παρουσιάζεται στο σχήμα 43. Η μέτρηση των νομισμάτων δεν μπορεί να γίνει ακόμα γιατί όπως είναι εμφανές οι πλευρές των νομισμάτων εφάπτονται κάνοντας έτσι δύσκολο για το αλγόριθμο καταμέτρησης να ξεχωρίσει το ένα νόμισμα από το άλλο. Για να απλοποιήσουμε την διαδικασία εφαρμόζουμε erosion με ένα SE σχήματος δίσκου και διαμέτρου 11 εικονοστοιχεία. Το αποτέλεσμα παρουσιάζεται στο σχήμα 44. Τώρα πια τα νομίσματα έχουν χωριστεί το ένα από το άλλο και το σχήμα τους έχει διατηρηθεί σε μεγάλο βαθμό.

Στο σχήμα 45 έχουμε και πάλι εφαρμόσει erosion στην αρχική φωτογραφία αλλά αυτήν την φορά έχουμε χρησιμοποιήσει ένα τετράγωνο SE 9x9 αντί για δίσκο. Και τα δυο SE έχουν περίπου το ίδιο μέγεθος. Τα νομίσματα έχουν διαχωριστεί όπως και στην προηγούμενη περίπτωση αλλά το σχήμα των νομισμάτων έχει αλλοιωθεί σημαντικά.[24][25]

#### **4.4 Opening**

Το opening (άνοιγμα) και closing (κλείσιμο) είναι δυο πολύ σημαντικοί χειριστές της μαθηματικής μορφολογίας. Και οι δυο προέρχονται από τις θεμελιώδεις διαδικασίες του dilation και του erosion. Όπως και οι χειριστές από τους οποίους προέρχονται έχουν εφαρμογές σε binary εικόνες αλλά και σε grayscale. Η βασική επίδραση που έχει το άνοιγμα μιας φωτογραφίας είναι παρόμοιο με αυτό του erosion δηλαδή αφαιρούνται μερικά από τα foreground (φωτεινά) pixels από τα άκρα των περιοχών των foreground pixels. Όμως έχει λιγότερο διαβρωτικό χαρακτήρα στα στοιχεία της εικόνας. Η εφαρμογή του ανοίγματος μιας εικόνας γίνεται με την χρήση ενός SE, το οποίο και προσδιορίζει την ακριβής επίδραση που θα έχει η διαδικασία. Όταν εφαρμόσουμε την διαδικασία του ανοίγματος σε μια εικόνα τότε οι περιοχές των foreground pixels που έχουν παρόμοιο σχήμα με το SE ή μπορεί να χωρέσει ολόκληρο το SE μέσα τους, διατηρούνται και όλες οι υπόλοιπες περιοχές αφαιρούνται.

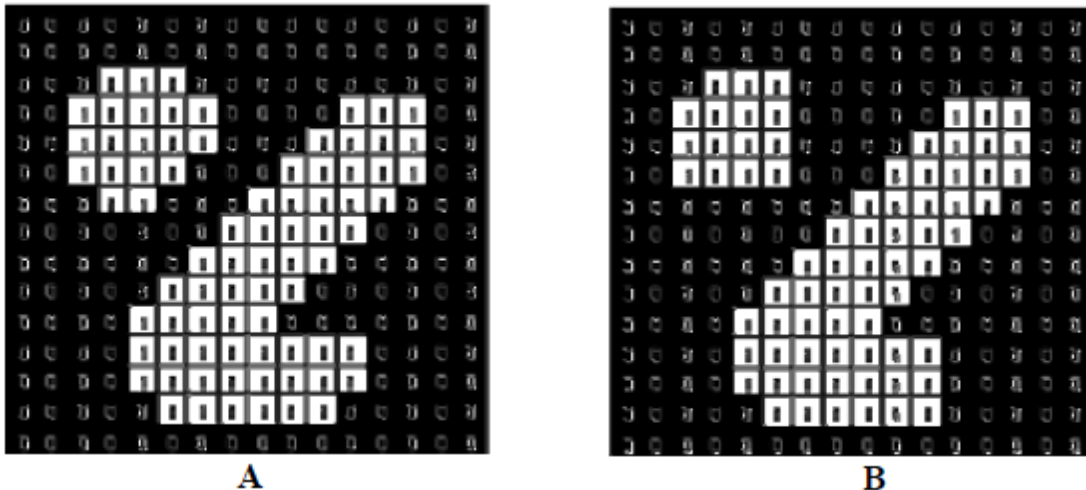
## Πως λειτουργεί

Η διαδικασία του ανοίγματος μιας φωτογραφίας περιγράφεται ως το erosion της εικόνας ακολουθούμενο από το dilation αυτής χρησιμοποιώντας και στις δυο περιπτώσεις το ίδιο SE. Για να εφαρμοστεί την διαδικασία του opening θα πρέπει να χρησιμοποιηθούν δυο πίνακες δεδομένων: την εικόνα που θα ανοιχθεί και το στοιχείο δόμησης.. Το άνοιγμα των foreground pixels μιας εικόνας είναι η αντίστροφη διαδικασία του κλεισίματος (closing) των background pixel της ίδιας εικόνας με το ίδιο στοιχείο δόμησης.

## Εφαρμογή του Opening

Αν και με απλό dilation μια εικόνας μπορούν να αφαιρεθούν οι ανεπιθύμητες μικρές περιοχές των foreground όπως π.χ. ο 'salt' θόρυβος αρκετά αποτελεσματικά, μειονεκτεί στο ότι επηρεάζει όλα τα foreground pixels αδιακρίτως. Το opening παρακάμπτει αυτό το πρόβλημα εφαρμόζοντας πρώτα erosion και μετά dilation στην εικόνα.

Τα αποτελέσματα του opening μπορούν εύκολα να οπτικοποιηθούν. Φανταστείτε το στοιχείο δόμησης να γλιστράει στα όρια από κάθε περιοχής foreground pixels από την εσωτερική πλευρά χωρίς να αλλάζει ο προσανατολισμός του. Όλα τα pixel του foreground που καλύπτονται από το στοιχείο δόμησης ενώ αυτό βρίσκεται ολόκληρο μέσα στην περιοχή του foreground, θα διατηρηθούν. Όλα τα υπόλοιπα pixels που δεν μπορούν να καλυφθούν από το SE χωρίς αυτό να πέφτει και στην περιοχή του background θα υποστούν την διαδικασία του erosion. Όταν τελειώσει η διαδικασία του opening τότε η περιοχή του foreground θα έχει διαμορφωθεί έτσι ώστε το στοιχείο δόμησης θα μπορεί να τοποθετηθεί οπουδήποτε πάνω της και να μην «πέφτει» πάνω σε εικονοστοιχείο του background. Έτσι περαιτέρω άνοιγμα της εικόνας δεν θα επέφερε καμιά αλλαγή. Στις εικόνες 46 και 47 φαίνεται η επίδραση του ανοίγματος με ένα τετράγωνο 3x3 στοιχείου δόμησης.



Εικ.46&47 Επίδραση του ανοίγματος χρησιμοποιώντας μια 3x3 τετραγωνική δόμηση στοιχείου.

Όπως και στην περίπτωση του dilation και του erosion ένα από τα πιο συνηθισμένα SE είναι ο τετράγωνος πίνακας 3X3. Η επίδραση που παρουσιάζεται στο παραπάνω σχήμα είναι σχετικά ήπια εάν ληφθεί υπ' όψιν ότι το στοιχείο δόμησης είναι αρκετά συμπαγές και χωράει εύκολα μέσα στα όρια της foreground περιοχής ακόμη και πριν από την διαδικασία του ανοίγματος. Για να αυξηθεί η επίδραση μπορούν να εφαρμοστούν πολλαπλές διαδικασίες του erosion ακολουθούμενες από τον ίδιο αριθμό διαδικασιών του dilation. Πρακτικά αυτή η διαδικασία είναι σαν να ανοιγόταν η εικόνα με ένα μεγαλύτερο τετράγωνο στοιχείο δόμησης.[26][27][28]

#### 4.5 Closing

Το κλείσιμο (closing) είναι ένας πολύ σημαντικός χειριστής στην μαθηματική μορφολογία. Όπως και ο αντίστροφος χειριστής του το άνοιγμα προέρχονται από τους θεμελιώδεις χειριστές του erosion και dilation. Εφαρμογές του κλεισίματος μπορούν να βρεθούν και για τις δυαδικές αλλά και για τις grayscale εικόνες. Το κλείσιμο μιας εικόνας παρουσιάζει κάποιες ομοιότητες με το dilate αυτής στο και στις δύο περιπτώσεις εκτείνονται τα όρια του φωτεινού αντικειμένου ( foreground), μονό που στην περίπτωση του closing το σχήμα του αντικειμένου διατηρείται καλύτερα. Όπως και οι υπόλοιποι μορφολογικοί χειριστές η ακριβής επίδραση που θα έχουν καθορίζεται από το στοιχείο δόμησης που θα επιλέγει. Η γενική επίδραση που έχει η διαδικασία είναι να διατηρούνται οι περιοχές του background που έχουν παρόμοιο σχήμα με το στοιχείο δόμησης ή περιοχές που μπορεί να χωρέσει πλήρως μέσα τους το στοιχείο δόμησης, ενώ όλες οι υπόλοιπες περιοχές του background αφαιρούνται.

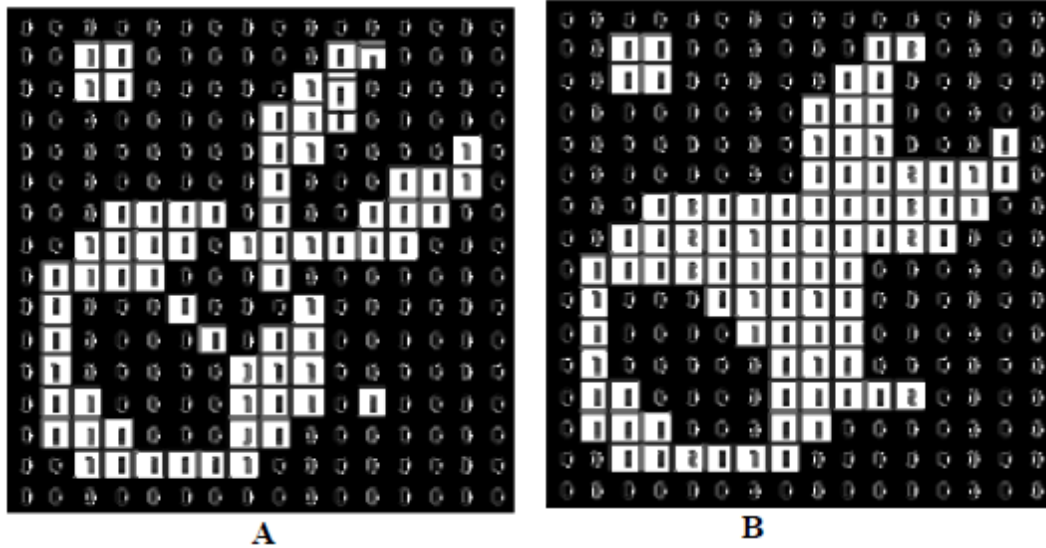
## Πώς λειτουργεί

Η διαδικασία του κλεισίματος είναι η ακριβώς ανάποδη από την διαδικασία του ανοίγματος. Είναι δηλαδή το dilation μιας εικόνας ακολουθούμενο από το erosion αυτής χρησιμοποιώντας το ίδιο SE. Έτσι για να εφαρμοστεί η διαδικασία του closing θα πρέπει να χρησιμοποιηθούν δυο πίνακες δεδομένων: η εικόνα που θα κλειστεί και το στοιχείο δόμησης. Το κλείσιμο των foreground pixels μιας εικόνας είναι η αντίστροφη διαδικασία του ανοίγματος (opening) των background pixels της ίδιας εικόνας με το ίδιο στοιχείο δόμησης.

## Εφαρμογή του Closing

Μια από τις εφαρμογές του closing είναι η αφαίρεση του pepper θορύβου, οι “τρύπες” χρώματος που υπάρχουν στο background. Η δυσκολία του να γίνει κάτι τέτοιο με την διαδικασία του dilation έγκειται στο ότι θα αλλοιωθούν όλες οι περιοχές των pixels αδιακρίτως. Εφαρμόζοντας την διαδικασία του erosion στην εικόνα αμέσως μετά από το dilation και χρησιμοποιώντας το ίδιο στοιχείο δόμησης, δηλαδή κάνοντας closing στην εικόνα μειώνουμε αυτό το πρόβλημα.

Τα αποτελέσματα του closing μπορούν εύκολα να οπτικοποιηθούν. Φανταστείτε το στοιχείο δόμησης να γλιστράει γύρω από κάθε περιοχή foreground pixels χωρίς να αλλάζει ο προσανατολισμός του. Για κάθε pixel του background που βρίσκεται στα σύνορα αυτών των περιοχών γίνεται έλεγχος εάν το στοιχείο δόμησης μπορεί να περιέχει αυτό το pixel χωρίς να περιέχει παράλληλα και κάποιο κομμάτι της foreground περιοχής. Εάν αυτό ισχύει τότε το pixel παραμένει background pixel διαφορετικά γίνεται foreground pixel. Όταν τελειώσει η διαδικασία του κλεισίματος τότε η περιοχή του background θα έχει διαμορφωθεί έτσι ώστε το στοιχείο δόμησης θα μπορεί να τοποθετηθεί οπουδήποτε πάνω της και να μην «πέφτει» πάνω σε εικονοστοιχείο του foreground. Έτσι περαιτέρω κλείσιμο της εικόνας δεν θα επέφερε καμιά αλλαγή. Στην εικόνα απεικονίζεται η επίδραση που έχει το κλείσιμο στην εικ.48 με την χρήση ενός τετράγωνου 3x3 στοιχείου δόμησης.



Εικ.48&49 Επίδραση του κλεισίματος χρησιμοποιώντας μια 3x3 τετραγωνική δόμηση στοιχείου

Όπως και για το dilation και το erosion το τετράγωνο 3x3 στοιχείο δόμησης είναι από τα πιο κοινώς χρησιμοποιούμενα. Για επιτευχθεί το κλείσιμο μιας εικόνας με μεγαλύτερο SE δεν χρειάζεται να οριστεί κάποιο μεγαλύτερο SE αλλά τα ίδια αποτελέσματα μπορούν να επιτευχθούν με πολλαπλά dilation ακολουθούμενα από τον ίδιο αριθμό erosion.[27][28]

## ΚΕΦΑΛΑΙΟ 5

### 5.1 Διακριτή εξίσωση Poisson

#### 5.1.1 Δισδιάστατη

Θα λύσουμε την εξίσωση την διακριτή εξίσωση για 2 διαστάσεις με έναν μετασχηματισμό Laplace.

$$2D - \text{Laplacian}(u) = F(x, y) = \frac{d^2 u(x, y)}{dx^2} + \frac{d^2 u(x, y)}{dy^2}$$

Για  $(x, y)$  στο πεδίο Omega με μοναδιαίο τετράγωνο  $[0 < x, y < 1]$ . Για να κάνουμε την λύση πιο συγκεκριμένη πρέπει να ορίσουμε τις οριακές συνθήκες, δηλαδή να ορίσουμε την τιμή για  $u(x, y)$  στα όρια του Omega. Αυτό μπορεί να γίνει εύκολα με την απλή υπόθεση ότι:

$$u(x, y) = 0 \quad , \text{αν } (x, y) \in \text{Omega}$$

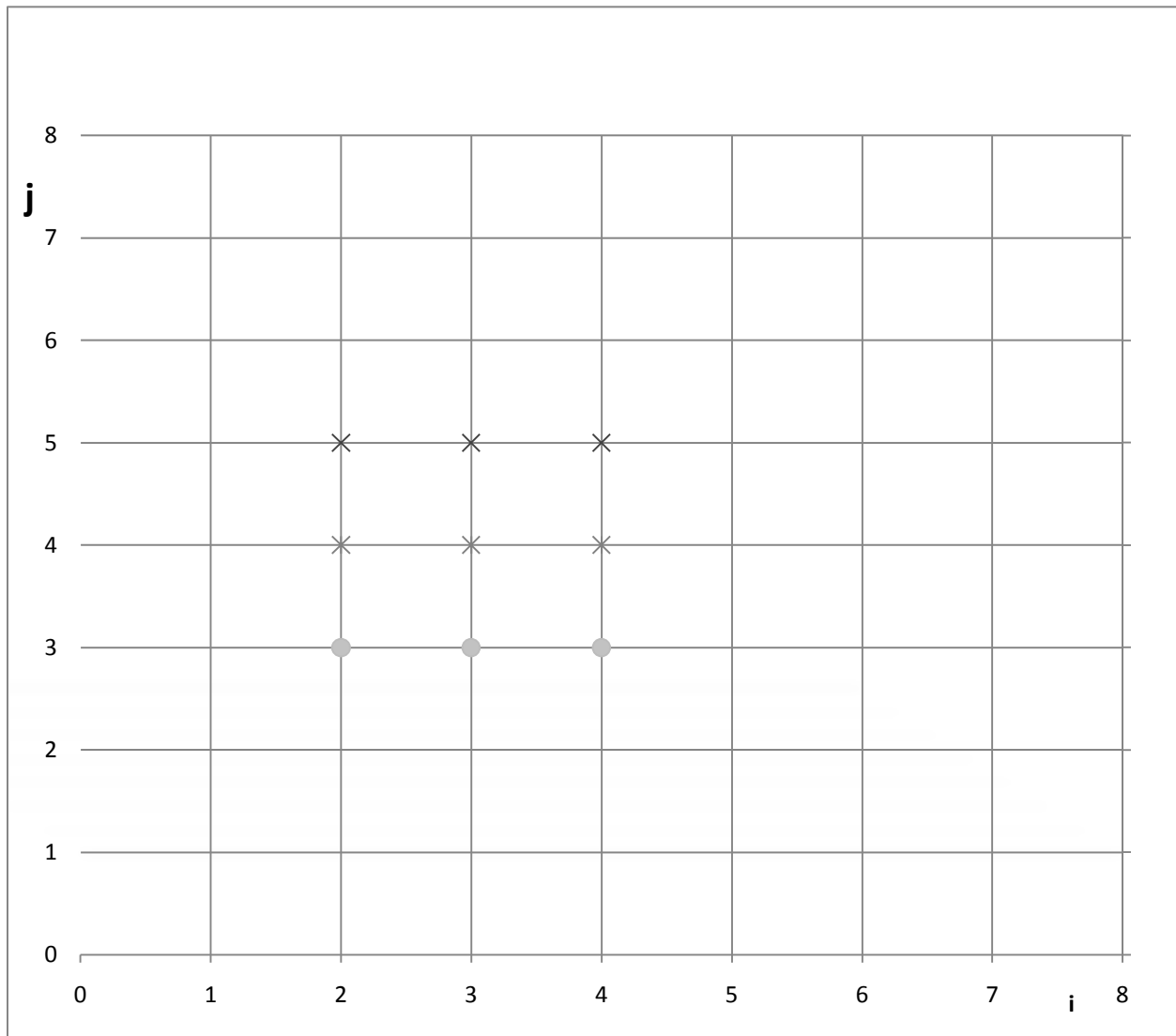
Αυτή η εξίσωση περιγράφει την θερμοκρασία "σταθερής-κατάστασης", ενός ομοιόμορφου τετραγώνου με όρια (σε θερμοκρασία)  $u=0$  και  $F(x, y)$  να είναι ίσο με της πηγες εξωτερικής στο  $(x, y)$ .

#### 5.1.2 Λύση της εξίσωσης στις δύο διαστάσεις

Θα διακρίνουμε την εξίσωση σε ορισμένο αριθμό διαφορικών. Χρησιμοποιώντας ένα πλαίσιο  $(n+1)$  με  $(n+1)$  μοναδιαίων τετράγωνων στο πεδίο Omega.

Όπου  $h = \frac{1}{n+1}$  οι διαστάσεις των τετραγώνων του πλαισίου και αφήνουμε το  $u(i, j)$  να είναι η λύση για

$$u(i, j) = \begin{cases} x = i * h, & \text{για } n = 7 \\ y = j * h, & \text{για } n = 7 \end{cases}$$



$$4 * u(i, j) - u(i - 1, j) - u(i + 1, j) - u(i, j - 1) - u(i, j + 1) = b(i, j)$$

Η άνωθεν γραμμική εξίσωση συσχετίζει την  $u(i, j)$  και τις τιμές των γειτονικών τετραγώνων (pixel στην περίπτωση μας).

Για  $\begin{cases} 1 \leq i \\ j \leq n \end{cases}$  τότε  $N = n^2$  (εξισώσεις αγνώστων).

Όταν  $(i,j)$  είναι κοντά στα όρια  $\begin{cases} i = 1 & \text{ή} & i = n \\ j = 1 & \text{ή} & j = n \end{cases}$  τότε ένα ή περισσότερα σημεία  $u(i\pm, j\pm)$  είναι πάνω στα όρια και για αυτό είναι ίσα με το 0.

$$b(i,j) = -f(i * h, j * h) * h^2$$

Η άνωθεν γραμμική εξίσωση είναι η κλιμακούμενη για το δεξί μέρος της εξίσωσης  $f(x,y)$  στο αντίστοιχο σημείο, στο πλαίσιο  $(i,j)$ . [30]

## 5.2 Τρισδιάστατη

*Στις τρεις διαστάσεις η εξίσωση είναι περίπου ίδια:*

$$3D - Laplacian(u) = F(x, y, z) = \frac{d^2u(x, y, z)}{dx^2} + \frac{d^2u(x, y, z)}{dy^2} + \frac{d^2u(x, y, z)}{dz^2}$$

### 5.2.1 Λύση της εξίσωσης στις τρεις διαστάσεις

Δεν θα μελετήσουμε τον τρόπο επίλυσης στις τρεις διαστάσεις, όμως θα προσθέσω ότι για την επίλυση της απαιτείται να δώσουμε τις οριακές συνθήκες για  $(x,y,z)$ .



Πιν.3 Αλγόριθμοι Επίλυσης (Με παράλληλη επεξεργασία δεδομένων)

Αλγόριθμος	Τύπος	Σειριακός Χρόνος	PRAM Parallel	Αποθηκευτικός	#Επεξεργαστών
Dense LU	D	$N^3$	N	$N^2$	$N^2$
Band LU	D	$N^2$	N	$N^{3/2}$	N
Inv(P)*bhat	D	$N^2$	logN	$N^2$	$N^2$
Jacobi	I	$N^2$	N	N	N
Sparce LU	D	$N^{3/2}$	$N^{1/2}$	$N*\log N$	N
CG	I	$N^{3/2}$	$N^{1/2}+\log N$	N	N
SOR	I	$N^{3/2}$	$N^{1/2}$	N	N
FFT	D	$N*\log N$	logN	N	N
MultiGrid	I	N	$(\log N)^2$	N	N
LowerBound	--	N	logN	N	--

## Κεφάλαιο 6

### 6.1 Επαναληπτική σκιαγραφική μέθοδος

#### 6.1.1 Εισαγωγικά

Η βασική ιδέα πίσω από τις ποσοτικές σκιαγραφικές μεθόδους αποτελεί ο υπολογισμός της επιφάνειας μέσω των μετρήσεων της έντασης ενός διασκορπισμένου πεδίου (την σκιά) ενός αντικείμενου, το οποίο απορρόφα, ασθενώς, την φάση από μια ακτίνα φωτός, οπίσθιας προβολής που μεταδίδεται διαμέσου του άξονα  $z$ . Δύο γραφήματα της έντασης της ακτινοβολίας που διαχέεται από το αντικείμενο καταγράφονται σε δύο επίπεδα μελέτης, κανονικοποιημένα ως προς τον άξονα  $z$  διάδοσης της ακτίνας του φωτός. Αυτά από μόνα τους είναι ικανά να πάρουν την διασπορά φάσης του φωτεινού πεδίου στο επίπεδο του αντικείμενου.

#### 6.1.2 Μαθηματικά πρότυπα

Ας υποθέσουμε ότι στο εγκάρσιο επίπεδο  $O$  περιλαμβάνεται και το αντικείμενο, αυτό το πεδίο πρέπει να χωρίζεται από το επίπεδο  $S$  (σκιά (Shadow)) από μια απόσταση  $\Delta z$ . Στην σκιαγραφική μέθοδο, ο δεσμός μεταξύ των αλλαγών στην ένταση και στο προφίλ της επιφάνειας παρέχεται από την γεωμετρική οπτική. Για λεπτά-αντικείμενα και για μικρό εύρος διασποράς, η μέθοδος καταλήγει στην λύση της ακόλουθης εξίσωσης Poisson:

$$\nabla^2 \varphi_0(x, y) = \frac{k}{\Delta z} \left[ 1 - \frac{I_S(x, y)}{I_O(x, y)} \right] \quad \text{εξ. 1}$$

Η εξίσωση αυτή θα αναφέρεται ως η σκιαγραφική εξίσωση

Το  $\varphi_0$  εδώ είναι το προσδοκώμενο φασικό πεδίο του επιπέδου  $O$ , το  $I_O(x, y)$  και το  $I_S(x, y)$  είναι διασταυρούμενα πεδία της έντασης όπως αυτή μετρείται στο  $O$  και  $S$  επίπεδο, και ο παράγοντας  $k$  είναι το κυματικό διάνυσμα διάδοσης. Εδώ πρέπει να παρατηρήσουμε ότι η εξίσωση προδιαθέτει ότι η διαφορά έντασης είναι ανάλογη με την καμπύλη της οπτικής φάσης του επιπέδου  $\varphi_0$ .

Η μέθοδος υπό συζήτηση αφαιρεί την διάθλαση από την απεικόνιση του γραφήματος φάσης αντικείμενου το οποίο έχει ληφθεί με την σκιαγραφική εξίσωση. Ο αλγόριθμος

υπολογίζει επαναληπτικά τα  $\psi_n$  του προφίλ φάσης που ξεκινάει από το  $\psi_0 = \varphi_0$ . Το εκτιμώμενο φασικό πεδίο, σε κάθε βήμα-επανάληψη, χρησιμοποιείται για να πλησιάσουμε το πεδίο στο επίπεδο του αντικειμένου, με τον τρόπο αυτό.

$$U_n = \sqrt{I_0} e^{i\psi_n} \text{ εξ. 2}$$

Το πεδίο αυτό μετά διαδίδεται στο επίπεδο της σκιάς (shadow plane) S και το αναμενόμενο πεδίο έντασης υπολογίζεται από το αποτέλεσμα της συνέλιξης

$$I_n = |F_{\Delta z} * U_n|^2 \text{ εξ. 3}$$

Όπου:

$$F_{\Delta z}(x) = \frac{k e^{ik\Delta z}}{i2\pi\Delta z} e^{ik/\Delta z(x^2+y^2)} \text{ εξ. 4}$$

Είναι ο πυρήνας Fresnel για την διάδοση στο ελεύθερο διάστημα.

Μια διόρθωση της τάξης  $\varphi_n$  στο προφίλ της φάσης  $\psi_n$  υπολογίζετε από την σκιαγραφική εξίσωση αν αντικαταστήσουμε το  $I_s$  με το  $I_n$  και το  $I_0$  με το  $I_s$  :

$$\nabla^2 \varphi_n = \frac{k}{\Delta z} \left(1 - \frac{I_n}{I_s}\right) \text{ εξ. 5}$$

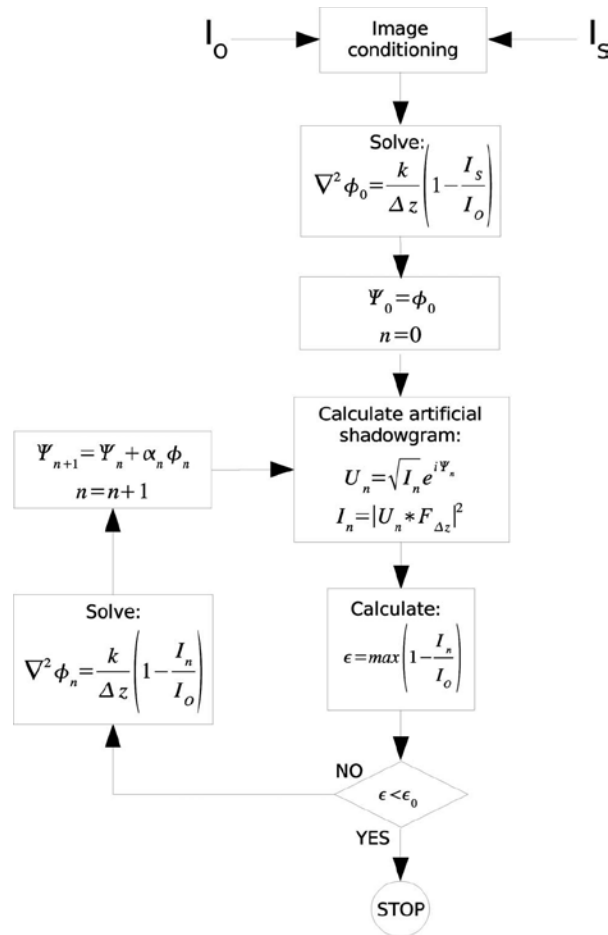
Τέλος η διόρθωση αυτή προστίθεται στο προφίλ φάσης που ήδη έχουμε υπολογίσει:

$$\psi_{n+1} = \psi_n + a_n \varphi_n \text{ εξ. 6}$$

Όπου  $a_n$  είναι μια προσεκτικά επιλεγμένη παράμετρος υπερχαλάρωσης (overrelaxation):

$$a_n = (-1)^n a \text{ εξ. 7}$$

Με  $a=1$  τότε έχουμε μια επαρκή παράμετρο που εξασφαλίζει την σύγκλιση της σειράς του προφίλ της φάσης  $\psi(x,y)$  του πεδίου στο επίπεδο του αντικειμένου O.



Block διάγραμμα της σκιαγραφικής μεθόδου

## 6.2 Σύγκλιση της φάσης ανάπτυξης

### 6.2.1 Απόδειξη του αλγορίθμου

Για να επιβεβαιώσουμε την σύγκλιση της εξίσωσης 6, πρέπει να αποδείξουμε ότι η σειρά συναρτήσεων  $\{\varphi_n\}$  συγκλίνουν ομοιόμορφα στο 0. Η αναλυτική απόδειξη είναι μεγάλη και πολύπλοκη και χρησιμοποιεί μερικές από τις πιο πρόσφατες ανακαλύψεις στο χώρο των μαθηματικών της φυσικής.

Η απόδειξη θα γίνει σε τρία μέρη. Στο πρώτο μέρος θα χρησιμοποιήσουμε προβλέψεις που έχουμε κυρίως πάρει από την εξίσωσή 3 ή από την μέθοδο σταθερής φάσης. Μέτα

εισαγάγουμε υποδιαιρέσεις των πεδίων της έντασης και της φάσης με τόση διάμετρο όση χρειάζεται έτσι ώστε η ένταση να παρουσιάζει μέτριες διακυμάνσεις.

Δεύτερον, προσπαθούμε να δείξουμε ότι το  $\{\varphi_n\}$  είναι μια σειρά *Cauchy*. Αυτό είναι και το πιο τεχνικό μέρος της απόδειξης, το οποίο χρησιμοποιεί τις *weighted estimates* και το σχέδιο επανάληψης Moser. Επειδή ο χώρος που καταλαμβάνουν οι ομαλές συναρτήσεις σε ένα δεσμευμένο πεδίο είναι σε σχήμα Banach, αυτό μας οδηγεί στο συμπέρασμα ότι η σειρά  $\{|\varphi_n|\}$  συγκλίνει στο 0.

Τρίτον και τελευταίο, εφαρμόζουμε το αρχικό αποτέλεσμα της σύγκλισης και το εναλλακτικό  $\alpha_n$  έτσι έχουμε την σύγκλιση στο αληθινό προφίλ της φάσης. Το  $\alpha_n$  στην περίπτωση μας υπαγορεύεται από το σχέδιο επανάληψης.

Θέτω  $\sigma = \Delta z/k$  και ορίζω  $\Omega \subset \mathbf{R}^2$  ως το πεδίο της εικόνας προς επεξεργασία. Από την εξίσωση 3 μπορούμε να πάρουμε τρεις χρήσιμες ανισότητες:

$$I_n \leq \frac{\text{vol}(\Omega)k}{(2\pi\Delta z)} \left[ \int_{\Omega} I_o(x, y, \Delta z) dx dy \right] \text{ av. 8}$$

$$|I_n - I_s| \leq C \frac{k}{\Delta z} I_o \int_{\Omega} [|\Psi_n - \Psi_m|^2 dx dy]^{1/2} \text{ av. 9}$$

$$1 - \sigma \nabla^2 \Psi_n + E^- \leq \frac{I_n}{I_o} \leq 1 - \sigma \nabla^2 \Psi_n + E^+ \text{ av. 10}$$

Όπου C είναι μία σταθερά που εξαρτάται από την περιοχή του αποτελέσματος της ολοκλήρωσης και τα  $E^\pm$  είναι δεσμευμένα από την σειρά  $O(\sigma^2)$ . Η παράγωγος της εξίσωση 8 είναι αμελητέα. Η δεύτερη ανισότητα προέρχεται από την ταυτότητα σύνθετων αριθμών  $|z|^2 - |w|^2 = z(\bar{z} - \bar{w}) + \bar{w}(z - w)$ , ακολουθούμενη από το τρίγωνο, τις ανισότητες Cauchy-Schwarz και το ημίτονο  $|\sin x| \leq |x|$ . Τέλος η εξ. 10 ακολουθεί την απευθείας προσέγγιση της μεθόδου στατικής φάσης του ολοκληρώματος της διάδοσης εξ. 3.

Για κάθε μια από τις επαναλήψεις n, υποθέτουμε ότι η λύση της  $\varphi_n$  εξίσωσης 5 και την πολλαπλασιάζουμε με την συνάρτηση βάρους  $w_n$  για να αποκτήσουμε ένα σετ συναρτήσεων  $\{\zeta_n\}$ :

$$\zeta_n = \omega_n \varphi_n \text{ εξ. 11}$$

Η συναρτήσεις βάρους είναι επιλεγμένες έτσι ώστε να απλοποιηθούν ως  $w_n = \sum_j P_n^j \chi_n^j$  όπου  $P_n^j$  είναι αναλυτικές συναρτήσεις και  $\chi_n^j$  είναι ομαλές συναρτήσεις αποκοπής καθορισμένες σε υπό-επίπεδα  $\Omega_n^j \subset \Omega$ .

Μετά εισάγουμε τις υπόλοιπες συναρτήσεις

$$K_{n,m} = \varphi_n - \zeta_m \quad n, m \in N \text{ εξ. 12}$$

Θα δείξουμε ότι  $\{K_{n,m}\}$  τείνει στο 0. Λόγω της τριγωνικής ανισότητας έχουμε:

$$|K_{n,m}| \geq |\varphi_n - \varphi_m| - |1 - \omega_m| |\varphi_m| \quad \text{εξ. 13}$$

Κατά την απόδειξη θα δείξουμε ότι  $\omega_n \rightarrow 1$ , που σημαίνει ότι  $\varphi_n$  είναι μια σειρά Cauchy. Λόγω του ότι αυτές οι συναρτήσεις σε ορισμένο διάστημα αποτελούν ένα Banach χώρο με την υπερκανονικότητα, τότε  $\{|\varphi_n|\}$  συγκλίνει. Θα χρειαστούμε και 2 ακόμα βοηθητικές συναρτήσεις:

$$\varepsilon_{n,m} = 1 - I_n/I_S - \omega_m \left(1 - \frac{I_m}{I_S}\right) \quad \text{εξ. 14 και}$$

$$\eta_m = c_m - \omega_m \quad \text{εξ. 15}$$

Πρέπει να εξασφαλίσουμε ότι η παράμετρος  $\varepsilon_{n,m} < \varepsilon$  για  $n, m > N(\varepsilon)$ . Αυτό επιτυγχάνεται με την κατάλληλη επιλογή συνάρτησης βαρύτητας. Ομοίως επιλέγουμε την σταθερά  $c_m$  έτσι ώστε  $\eta_m < \varepsilon$ . Αυτό είναι εφικτό λόγω των ανισοτήτων 9,10 και την ανισότητα του Poincare.

$$\begin{aligned} |I_n - I_m| &\leq \sum_j C \text{diam}(\Omega_N^j) \left[ \int_{\Omega_N^j} |\nabla(\Psi_n - \Psi_m)| \right]^{\frac{1}{2}} \\ &\leq C' \sum_j \frac{\text{diam}(\Omega_N^j)}{\sigma} \left( \int_{\Omega_N^j} \frac{|I_n - I_m|^2}{I_0^2} dx dy \right)^{\frac{1}{2}} \quad \text{εξ. 16} \end{aligned}$$

Όπου ο παράγοντας  $\text{diam}(\Omega_N^j)$  είναι η διάμετρος του πεδίου  $\Omega_N^j$ . Η τελευταία ανισότητα προέρχεται από την ολοκλήρωση σε μέρη και την ανισότητα Cauchy-Schwarz. Αυτό μας επιτρέπει να επιλέξουμε το  $\Omega_N^j$  έτσι ώστε σε κάθε βήμα  $n > N(\varepsilon)$  να είναι  $\frac{\text{diam}(\Omega_N^j)}{\sigma} < \varepsilon$ . Δηλαδή καθορίζουμε το μέγεθος της περιοχής έτσι ώστε να έχουμε μέτριες παραλλαγές στις τιμές της έντασης (σε  $\sigma$  μονάδες). Τώρα αφαιρώντας τις αν.10 για ένα οποιοδήποτε ζεύγος  $(n,m)$  και χρησιμοποιώντας τον κανόνα του Leibniz's την παρακάτω, προσεγγιστική, εξίσωση:

$$\nabla^2 K_{n,m} \sim \varepsilon_{n,m} + 2(\zeta_m H_m^0 + \nabla \zeta_m H_m^1 - \zeta_m |H_m^1|^2) \quad \text{εξ. 17}$$

Έχουμε όμως δείξει ότι:

$$H_n^0 = \frac{\nabla^2 \eta_n}{\eta_n}, H_n^1 = \frac{\nabla \eta_n}{\eta_n} \quad \text{εξ. 18}$$

Τις παραπάνω ανισότητες τις γράφουμε σε ασθενή μορφή, δηλαδή πολλαπλασιάζω και τα δυο μέρη με μια απαλή συνάρτηση αποκοπής  $\chi$  που στηρίζεται σε ένα από τα  $\Omega_N^j$  το οποίο ορίζω, απλά, ως  $\Omega$ . Έτσι έχουμε:

$$\int_{\Omega} \chi \nabla^2 K_n \sim \varepsilon + \int_{\Omega} \chi \zeta_m H_m^0 + x \nabla \zeta_m * H_m^1 + x \zeta_m |H_m^1|^2 \quad \text{εξ.19}$$

Το οποίο με ολοκλήρωση κατά μέρη μας δίνει, φυσικά με τα κατάλληλα  $\gamma_1$  και  $\gamma_2$ :

$$\int_{\Omega} \chi \nabla^2 K_n \leq \varepsilon + \gamma_1 \int_{\Omega} \chi \zeta_m (H_m^0 + |H_m^1|^2) + \gamma_2 \int_{\Omega} \zeta_m \nabla \chi * H_m^1 \quad \text{εξ.20}$$

Η σχέση 16 μας δείχνει ότι μπορούμε να χωρίσουμε την εικόνα μετά από κάθε επανάληψη έτσι ώστε  $I_n - I_m \rightarrow 0$ , περαιτέρω βελτίωση του  $\Omega$  οδηγεί στις επιλογές  $w_m \sim 1$  και  $c_m \sim 1$  οι οποίες είναι αναγκαίες για να κρατήσουμε το  $\eta_m$  και το  $\varepsilon_{n,m}$  μικρότερα από το  $\varepsilon$ . Το  $w_m \rightarrow 1$  επίσης σημαίνει ότι ο δεύτερος όρος του δευτέρου μέρους τις ανισότητας 13 εξαλείφεται γιατί το  $m \rightarrow \infty$ . Άρα μένει μόνο να δειχτεί ότι και το  $K_{n,m} \rightarrow 0$ . Αυτό επιτυγχάνεται από μια διαδικασία που λέγεται επαναληπτική μέθοδος Moser η οποία σε συνδυασμό με τις ανισότητες Hardy.

Η αρχή αυτών των μεθόδων είναι η παρακάτω ανισότητα ομαλών συναρτήσεων που ανήκει στο πεδίο  $\Omega$ .

$$\int_{\Omega} u^2 \leq C \int_{\Omega} |\nabla u|^2 \quad \text{εξ.21}$$

Όπου η C είναι μια σταθερά που εξαρτάται το τετράγωνο της διαμέτρου του πεδίου  $\Omega$ . Από την άνωθεν ανισότητα προκύπτει ότι για έναν οποιονδήποτε, θετικό, πραγματικό, αριθμό s έχουμε:

$$\left( \int_{\Omega} u^{2sl} \right)^{\frac{1}{2s}} \leq C_s [\text{vol}(\Omega)]^{2s+3/2s+2} \left( \int_{\Omega} u^{2l+2} |\nabla u|^2 \right)^{\frac{1}{2}} \quad \text{εξ.22}$$

Το δεξί μέρος τις 22 μετά από ολοκλήρωση κατά μέρη μας δίνει σε ένα ορισμένο ολοκλήρωμα τις μορφής:  $\int_{\Omega} \phi u^{2l-1} (-\nabla^2 u)$  εξ.23 που ισχύει για κάθε συνάρτηση u που ικανοποιεί την διαφορική εξίσωση 17 ή το αντίστοιχο ολοκλήρωμα τις σχέσης 20.

Μετά πάω στη σχέση 20 και βρίσκω το  $K_{n,m}^{2l-1} \chi$  μαζί με το  $\phi$  και χρησιμοποιώντας τους B3 και B4 βρίσκουμε την παρακάτω ανίσωση:

$$- \int_{\Omega} \chi K_{n,m}^{2l-1} \nabla^2 K_n \leq C_1 \int_{\Omega} \eta_m K_{n,m}^{2l-1} \left| \frac{\nabla(x\zeta_m)}{\sqrt{\chi\zeta_m}} \right|^2 \quad \text{εξ.24} \Rightarrow$$

$$\left( \int_{\Omega} \chi K_{n,m}^{2sl} \right)^{\frac{1}{2s}} \leq C_s [\text{vol}(\Omega)]^{\frac{2n+3}{2s+2}} \varepsilon \int_{\Omega} K_{n,m}^{2l-1} \left| \frac{\nabla(x\zeta_m)}{\sqrt{\chi\zeta_m}} \right|^2 \quad \text{εξ.25}$$

Η οποία με  $l=s^i$  επαναλήψεις δίνει  $\lim_{r \rightarrow \infty} \left[ \frac{1}{\text{vol}(\Omega)} \int_{\Omega} |f|^r \right]^{1/r} = \text{sup}_{\Omega} |f|$  εξ.26

Με  $\sup_{\Omega} |K_{m,n}| \rightarrow C\epsilon$  εξ.27

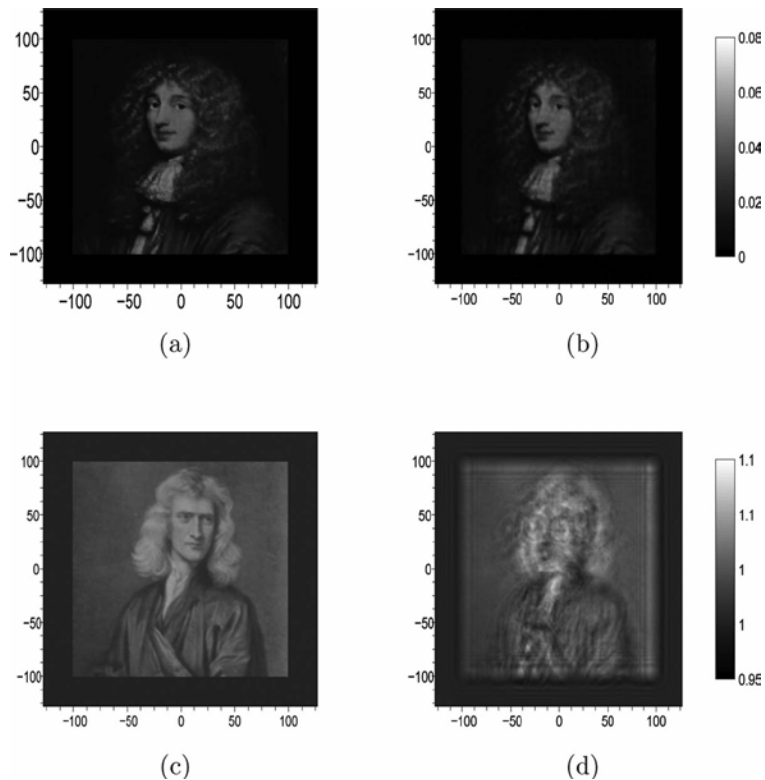
Η σχέση 27 μας δείχνει ότι το  $\{\varphi_n\}$  είναι μια σειρά Cauchy και μετατρέπεται σε μία ομαλή συνάρτηση. Αν η οριακή συνάρτηση δεν είναι 0, τότε από το δεύτερο μέρος της ανίσωσης 16 ( $n=m+1$ ) και λόγω της 5, δεν θα έχουμε  $I_n - I_m \rightarrow 0$ . Οπότε  $I_n \rightarrow I_S$  οπότε παίρνουμε:

$$\sup |\varphi_n| \rightarrow 0$$

Που είναι αρκετό για εξαγωγή μιας σειράς συναρτήσεων  $\{\varphi_n\}$  η οποία μειώνεται για θετικές τιμές και αυξάνει για αρνητικές τιμές και τελικά η σειρά συγκλίνει στο  $\sum_{n=0}^{\infty} (-1)^n \varphi_n$ .

### 6.2.2 Το αποτέλεσμα του θορύβου στον ISM

Παρά την αποτελεσματικότητα του αλγορίθμου σε εικόνες “καθαρές” χωρίς θόρυβο.



Για να έχει όμως και πρακτικές εφαρμογές πρέπει ο αλγόριθμος μας να είναι σταθερός απέναντι και στον θόρυβο, στις εφαρμογές ακόμη και μερικές ποσοστιαίες μονάδες λευκού ψευδοθορύβου είναι αρκετά για να καταστρέψουν την διαδικασία συλλογής



φάσης, το οποίο όμως δεν είναι κάτι που δεν περιμέναμε μιας και είναι ένα πρόβλημα που εμφανίζεται και σε άλλους αλγόριθμους ιδίου τύπου.

Η λύση σε μια τέτοια κατάσταση έρχεται από την λύση παρόμοιων προβλημάτων από-συνέλιξης. Έτσι σε κάθε επανάληψη κάνουμε μια κανονικοποίηση, μικρής κλίμακας, στην εξίσωσης 5 η οποία επιτυγχάνεται μέσω φιλτραρίσματος του μήκους κύματος.

Είναι εμφανές ότι για την πιστή χρησιμοποίηση του ISM αλγόριθμου θα πρέπει να γίνει μια σωστή εκτίμηση του είδους και του επιπέδου του θορύβου. Όσο για περιπτώσεις από-συνέλιξης, η μέγιστη φωτεινότητα εξαρτάται σχεδόν αποκλειστικά από τα επίπεδα σήματος προς θόρυβο (signal-to-noise levels).

### **6.3 Πλεονεκτήματα και περιορισμοί**

Ένα από τα μεγάλα πλεονεκτήματα του ISM είναι στην, αποδεδειγμένη, διόρθωση φάσης κατά την επανάληψη. Πράγμα το οποίο κάνει τον ISM πολύ εύρωστο απέναντι στα προβλήματα αποτελμάτωσης (σταματήματος). Αυτά του είδους προβλημάτων συναντώνται σε εικόνες χαμηλής φωτεινότητας. Επίσης ένα άλλο πλεονέκτημα είναι ότι σύγκλιση επιτυγχάνεται χωρίς να επηρεάζεται από το χωρισμό των επιπέδων  $O$  και  $S$ , σε αντίθεση με άλλες μεθόδους όπου θα η εικόνα θα πρέπει να έχει διακριτά επίπεδα. Ακόμα πρέπει να τονιστεί ότι ο ISM είναι σκιαγραφική μέθοδος και σαν τέτοια θα πρέπει να χρησιμοποιείται σε εφαρμογές εντοπισμού φάσης σε φασματικές περιοχές όπου ο έλεγχος της έντασης του φωτός και/ή interferometry είναι δύσκολες π.χ. EUV και soft to hard x-rays.

Τώρα όσο αναφορά τους περιορισμούς ένας πολύ βασικός είναι ότι λόγω της επαναληπτικής φύσης του ISM δεν μπορεί να χρησιμοποιηθεί σε εφαρμογές πραγματικού χρόνου (real-time) ειδικά σε πολύ μεγάλους πίνακες pixel, δηλαδή μεγάλες σε ανάλυση εικόνες. Για παράδειγμα μια τυπική εικόνα μεγέθους 256\*256 pixel σε έναν 64μπιτο, με CPU 3GHz χρειάζεται μερικά λεπτά. Ένα ακόμη σημαντικό μειονέκτημα, όχι του ISM μόνο αλλά των σκιαγραφικών μεθόδων γενικώς, είναι η αδυναμία παράγωγης μοναδιαίων μηκών κύματος. Πρέπει όμως να πούμε ότι και αυτό δεν είναι σημαντικό μειονέκτημα μιας και η ευαισθησία της μεθόδου μπορεί να φτάσει και τα επίπεδα του  $\lambda/100$  που την καθιστά ιδανική σε στόχους χαμηλής φωτεινότητας.[29]

## ΠΑΡΑΡΤΗΜΑ Α

### Α.1 Γραμμικές Εξισώσεις

#### Α.1.1 Εισαγωγή στις γραμμικές εξισώσεις

Ένα συχνό πρόβλημα στις επιστημονικές εφαρμογές είναι το πώς θα λύσουμε ένα σύστημα γραμμικών εξισώσεων.

Παρ.1

$$3x + 2y - z = 10$$

$$-x + 3y + 2z = 5$$

$$x - y - z = -1$$

Η Matlab είναι σχεδιασμένη να λύνει αυτού του είδους τα συστήματα εύκολα και απευθείας, όπως και θα δούμε.

Αν ο πίνακας A είναι:

$$A = \begin{bmatrix} 3 & 2 & -1 \\ -1 & 3 & 2 \\ 1 & -1 & -1 \end{bmatrix}$$

Και ο τα διανύσματα x και b είναι:

$$x = \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \quad b = \begin{bmatrix} 10 \\ 5 \\ -1 \end{bmatrix}$$

όπως βλέπετε είναι εύκολο να εκφράσουμε το παραπάνω σύστημα σε μορφή πινάκων:

$$\begin{bmatrix} 3 & 2 & -1 \\ -1 & 3 & 2 \\ 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 10 \\ 5 \\ -1 \end{bmatrix}$$

Η ακόμα πιο απλά σαν ένα μονοδιάστατο πίνακα:

$$\mathbf{Ax} = \mathbf{b}$$

Η λύση θα μπορούσε να γραφτεί και έτσι:

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$$

Όπου  $\mathbf{A}^{-1}$  είναι το ανεστραμμένος πίνακας  $\mathbf{A}$  (Δηλ. ο πίνακας που αν πολλαπλασιαστεί με  $\mathbf{A}$  δίνει το πίνακα  $\mathbf{I}$ ).

### A.1.2 Η λύση στην Matlab

Μπορούμε να λύσουμε στην **Matlab** το παραπάνω σύστημα απλώς εισάγοντας τις παρακάτω εντολές στο command line του command window:

```
A=[3 2 -1; -1 3 2; 1 -1 -1];
```

```
b=[10 5 -1];
```

```
x=A\b
```

Αυτό θα δώσει αποτέλεσμα:

```
x =
```

```
-2.0000
```

```
5.0000
```

```
-6.0000
```

Στην δίκη μας περίπτωση αυτό σημαίνει  $x=-2, y=5, z=-6$ .

Μπορούμε να ορίσουμε την πράξη  $\mathbf{A} \setminus \mathbf{b}$ , ως  $\mathbf{b}$  διαιρούμενο με το  $\mathbf{A}$ , ή ως τον ανάστροφο πίνακα  $\mathbf{A}$  πολλαπλασιαζόμενο με το  $\mathbf{b}$ , δηλαδή :

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$$

Θα μπορούσαμε επίσης να το γράψουμε στην **Matlab** σαν:

$$x = \text{inv}(A) * b$$

Επειδή η λειτουργία *inv* βρίσκει το ανάστροφο του πίνακα απευθείας. Όμως, το  $A \setminus b$  είναι πιο γρήγορο και αποτελεσματικό μιας και χρησιμοποιεί Gauss Elimination άρα και λιγότερες επαναλήψεις και χρόνο επεξεργασίας.

## A.2 Η Μέθοδος Jacobi στην Matlab

### A.2.1 Εισαγωγή

Η μέθοδος Jacobi είναι μια επαναληπτική μέθοδος επίλυσης γραμμικών συστημάτων της μορφής  $Ax=b$ . Όταν ο πίνακας  $A$  καλύπτει ορισμένα κριτήρια και ο αριθμός των επαναλήψεων επιλεγούν κατάλληλα μπορεί να είναι αρκετά ακριβής και πολύ πιο γρήγορες από τον Gaussian Elimination που αναφέραμε και πιο πάνω.

### A.2.2 Εφαρμογή στην Matlab

Η γενική επαναληπτική μέθοδος για να λύσουμε την  $Ax=b$  ορίζεται από τους όρους του

$$Sx^{new} = b + Tx^{old}$$

Όπου  $A=S-T$  και είναι και αρκετά εύκολο να λύσουμε συστήματα της μορφής  $Sx=b$ .

Η μέθοδος Jacobi εκμεταλλεύεται το γεγονός ότι τα διαγώνια συστήματα μπορούν να λυθούν με τον τρόπο της μίας διαίρεσης κατά άγνωστο. Δηλαδή σε  $O(n)$  Φλοπς (περάσματα). Για να εφαρμόσουμε την μέθοδο αυτή, γράφουμε  $A = L + D + U$  όπου  $D$  είναι ένας πίνακας  $n*n$  που περιέχει το άνω τριγωνικό μέρος του  $A$ . Ορίζω  $S$  την διαγώνιο της  $A$ . Τότε,  $T = S - A = -(L + U)$

**Παράδειγμα A : (5x5)**

$$A = \begin{bmatrix} 6 & 1 & 1 & 1 & 1 \\ 1 & 7 & 1 & 1 & 1 \\ 1 & 1 & 8 & 1 & 1 \\ 1 & 1 & 1 & 9 & 1 \\ 1 & 1 & 1 & 1 & 10 \end{bmatrix} \text{ και } B = \begin{bmatrix} -10 \\ -6 \\ 0 \\ 8 \\ 18 \end{bmatrix}$$

Η λύση θα για  $Ax=b$  είναι

$$x = \begin{bmatrix} -2 \\ -1 \\ 0 \\ 1 \\ 2 \end{bmatrix}$$

Ο κώδικας της **Matlab** είναι όπως παρατίθεται από κάτω και εισάγεται κατευθείαν στο command window :

**>> A=[6 1 1 1 1;1 7 1 1 1;1 1 8 1 1;1 1 1 9 1;1 1 1 1 10]**

```
A =  6   1   1   1   1
      1   7   1   1   1
      1   1   8   1   1
      1   1   1   9   1
      1   1   1   1  10
```

**>> b=[-10; -6; 0; 8; 18]**

```
b = -10
     -6
      0
      8
     18
```

```
>> S=diag(diag(A))
```

```
S =  6  0  0  0  0
      0  7  0  0  0
      0  0  8  0  0
      0  0  0  9  0
      0  0  0  0 10
```

```
>> T=S-A
```

```
T =  0 -1 -1 -1 -1
     -1  0 -1 -1 -1
     -1 -1  0 -1 -1
     -1 -1 -1  0 -1
     -1 -1 -1 -1  0
```

```
>> x=zeros(size(b))
```

```
x =  0
      0
      0
      0
      0
```

```
>> for k=1:10,x=S\b+T*x;
```

```
end
```

## ΒΙΒΛΙΟΓΡΑΦΙΑ

---

- [1] Steven W. Smith, “The Breadth and Depth of DSP”, The Scientist and Engineer’s Guide to Digital Signal Processing, Second Edition, σελ.9-10.
- [2] Νικόλαος Παπαμάρκος, «Τι είναι μια ψηφιακή εικόνα», Κεφάλαιο 1, Ψηφιακή Επεξεργασία & Ανάλυση Εικόνας, ISBN 960-387-352-7, Β.Γκιούρδας Εκδοτική, σελ.2-4.
- [3] Maria Petrou, Panagiota Bosdogianni, “What is an Image?”. Image Processing The Fundamentals, Willy Publications 1999, σελ. 1.
- [4] Νικόλαος Παπαμάρκος, «Πλήθος bits που απαιτούνται για την αποθήκευση μιας εικόνας», Κεφάλαιο 1, Ψηφιακή Επεξεργασία & Ανάλυση Εικόνας, Β.Γκιούρδας Εκδοτική,σελ.4-5.
- [5] Maria Petrou, Panagiota Bosdogianni, “What is meant by image resolution?”, Image Processing The Fundamentals, ISBN 0-471-99883-4, Willy Publications 1999, σελ. 2.
- [6] Νικόλαος Παπαμάρκος, «Ευκρίνεια Εικόνας», Κεφάλαιο 1, Ψηφιακή Επεξεργασία & Ανάλυση Εικόνας, Β.Γκιούρδας Εκδοτική,σελ.5-8.
- [7] Mathworks, Image Processing Toolbox, User’s guide Version 3, σελ. 2-4.
- [8] Rafael C. Gonzalez, Richard E. Woods and Steven L. Eddins, Digital Image Processing Using Matlab, “Data Classes”, Prentice Hall 2004, σελ. 23.
- [9] Mathworks, Image Processing Toolbox, User’s guide Version 3, σελ. 2-5 έως 2- 10.
- [10] Mathworks, Image Processing Toolbox, User’s guide Version 3, σελ. 2-28 έως 2- 31.
- [11] Rafael C. Gonzalez, Richard E. Woods and Steven L. Eddins, Digital Image Processing Using Matlab, “Intensity Transformation and Spatial Filtering ”, ISBN 0- 13-008519-7, Prentice Hall 2004, σελ. 65-66.
- [12] Michael Seu, Lawrence O’ Gorman, Michael J. Sammon, Practical Image Analysis, “Local Image Operations: Convolution”, ISBN 0-521-66065-3, Cambridge University 2004, σελ. 59-61.
- [13]<http://microscopy.fsu.edu/primer/java/digitalimaging/processing/kernelmaskoperation/index.html>
- [14] Rafael C. Gonzalez, Richard E. Woods and Steven L. Eddins, Digital Image Processing Using Matlab, “Linear Spatial Filtering “, Prentice Hall 2004, σελ.89.
- [15] Rafael C. Gonzalez, Richard E. Woods and Steven L. Eddins, Digital Image Processing Using Matlab, “Nonlinear Spatial Filtering “, Prentice Hall 2004, σελ. 96.

- [16] Rafael C. Gonzalez, Richard E. Woods and Steven L. Eddins, Digital Image Processing Using Matlab, “The 2-D Discrete Fourier Transformation ”, Prentice Hall 2004, σελ. 108-112.
- [17] Rafael C. Gonzalez, Richard E. Woods and Steven L. Eddins, Digital Image Processing Using Matlab, “Nonlinear Spatial Filtering ”, Prentice Hall 2004, σελ. 96.
- [18] Νικόλαος Παπαμάρκος, «Το ιστόγραμμα μιας εικόνας», Κεφάλαιο 1, Ψηφιακή Επεξεργασία & Ανάλυση Εικόνας, Β.Γκιούρδας Εκδοτική,σελ.121.
- [19] <http://www.cee.hw.ac.uk/hipr/html/threshld.html>
- [20] Νικόλαος Παπαμάρκος, «Κεφάλαιο 4 - Τμηματοποίηση εικόνας», Κεφάλαιο 1, Ψηφιακή Επεξεργασία & Ανάλυση Εικόνας, Β.Γκιούρδας Εκδοτική,σελ.145-209.
- [21] Rafael C. Gonzalez, Richard E. Woods and Steven L. Eddins, Digital Image Processing Using Matlab, “Thresholding ”, Prentice Hall 2004, σελ. 404-407.
- [22] <http://www.cee.hw.ac.uk/hipr/html/dilate.html>
- [23] Rafael C. Gonzalez, Richard E. Woods and Steven L. Eddins, Digital Image Processing Using Matlab, “Dilation ”, Prentice Hall 2004, σελ. 338.
- [24] <http://www.cee.hw.ac.uk/hipr/html/erode.html>
- [25] Rafael C. Gonzalez, Richard E. Woods and Steven L. Eddins, Digital Image Processing Using Matlab, “Thresholding ”, Prentice Hall 2004, σελ. 404-407.
- [26] <http://www.cee.hw.ac.uk/hipr/html/opening.html>
- [27] <http://www.cee.hw.ac.uk/hipr/html/closing.html>
- [28] Rafael C. Gonzalez, Richard E. Woods and Steven L. Eddins, Digital Image Processing Using Matlab, “Opening and Closing ”, Prentice Hall 2004, σελ. 347-350.
- [29] Dimitris Pliakis<sup>1</sup> and Stefano Minardi, Phase front retrieval by means of an iterative Shadowgraphic method.
- [30] <http://www.eecs.berkeley.edu/~demmel/cs267/lect>.