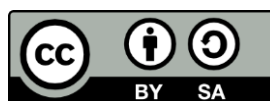


**ΤΕΧΝΟΛΟΓΙΚΟ ΕΚΠΑΙΔΕΥΤΙΚΟ ΙΔΡΥΜΑ
ΚΕΝΤΡΙΚΗΣ ΜΑΚΕΔΟΝΙΑΣ
ΣΧΟΛΗ ΤΕΧΝΟΛΟΓΙΚΩΝ ΕΦΑΡΜΟΓΩΝ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ**

ΨΗΦΙΑΚΗ ΕΠΕΞΕΡΓΑΣΙΑ ΕΙΚΟΝΑΣ

Χαράλαμπος Π. Στρουθόπουλος
Καθηγητής



Ιούλιος 2015

ΠΕΡΙΕΧΟΜΕΝΑ

1. Κεφάλαιο 1 – Εισαγωγή και ορισμοί
 - 1.1 Είδη ψηφιακών εικόνων.
 - 1.2 Βασικές έννοιες
 - 1.3 Ιστόγραμμα
2. Κεφάλαιο 2 – Επεξεργασίες σε δυαδικές εικόνες
 - 2.1. Κέντρο βάρους αντικειμένου
 - 2.2. Κωδικοποίηση δυαδικής εικόνας κατά μήκος διαδρομής.
 - 2.3. Συνδεδεμένα στοιχεία.
 - 2.4. Κωδικοποίηση αλυσίδας
 - 2.5. Ο μετασχηματισμός του Hough.
3. Κεφάλαιο 3 – Επεξεργασίες σε εικόνες αποχρώσεων του γκρι
 - 3.1. Βελτίωση εικόνων και αποκατάσταση εικόνων
 - 3.1.1. Εξομάλυνση
 - 3.1.2. Εξισορρόπηση ιστογράμματος
 - 3.2. Κατωφλίωση
 - 3.2.1 Κατωφλίωση με βάση τη διασπορά
 - 3.2.2 Κατωφλίωση με βάση την εντροπία
 - 3.2.3 Πολυκατωφλίωση με χρήση νευρωνικού δικτύου
 - 3.3. Κωδικοποίηση εικόνων αποχρώσεων του γκρι
 - 3.3.1 Κωδικοποίηση Huffman
 - 3.3.2 Κωδικοποίηση LZW
 - 3.3.3 Κωδικοποίηση 2Δ-διακριτού μετασχηματισμό συνημιτόνου
 - 3.4. Ανίχνευση ακμών
 - 3.4.1 Ανίχνευση ακμών με την χρήση πρώτων παραγώγων
 - 3.4.2 Ανίχνευση ακμών με την χρήση του τελεστή Laplace
 - 3.5 Μεταβολή μεγέθους – Παρεμβολή τιμών
4. Κεφάλαιο 4 - Μορφολογία ψηφιακών δυαδικών εικόνων
 - 4.1: Μεταφορά (Translation)
 - 4.2: Ανάκλαση (Reflection)
 - 4.3: Συμπλήρωμα (Complement)
 - 4.4: Τομή (Intersection)
 - 4.5: Ένωση (Union)
 - 4.6: Διαφορά (Difference)
 - 4.7: Διαστολή (Dilation)

- 4.8: Διάβρωση (Erosion)
- 4.9: Άνοιγμα (Opening)
- 4.10: Κλείσιμο (Closing).
- 4.11: Εφαρμογές – Παραδείγματα

ΚΕΦΑΛΑΙΟ 1

Εισαγωγή και ορισμοί

Η ψηφιακή επεξεργασία εικόνας (ΨΕΕ) αποτελεί έναν ευρύ επιστημονικό κλάδο που αναπτύχθηκε με την ραγδαία εξέλιξη των υπολογιστών. Ο όρος εικόνα χρησιμοποιείται ευρύτερα από την απλή απεικόνιση ενός σκηνικού και περιλαμβάνει την αποτύπωση κάθε είδους πληροφοριών. Τα υπερηχογραφήματα, οι μαγνητικές τομογραφίες, οι δορυφορικές φωτογραφίες κ.α. μπορούν να επεξεργαστούν ως ψηφιακές εικόνες.

Οι στόχοι της ΨΕΕ είναι οι εξής:

- 1) Η ψηφιοποίηση και κωδικοποίηση εικόνων με σκοπό την αποθήκευση, μετάδοση και εκτύπωσή τους.
- 2) Η βελτίωση και η αποκατάσταση των εικόνων με σκοπό την καλύτερη απεικόνισή τους.
- 3) Η ανάλυση και κατανόηση των εικόνων

Η ΨΕΕ συνεργάζεται με τους παρακάτω επιστημονικούς κλάδους:

- 1) Ψηφιακή Επεξεργασία Σημάτων (ΨΕΣ)
- 2) Ρομποτική όραση
- 3) Τεχνητή Νοημοσύνη
- 4) Αναγνώριση Προτύπων
- 5) Νευρωνικά Δίκτυα
- 6) Ασαφής Λογική
- 7) Κωδικοποίηση
- 8) Γραφικά Η/Υ

Η ψηφιακή εικόνα είναι ένα πεπερασμένο σύνολο περιοχών όπου κάθε περιοχή είναι χρωματισμένη με χρώμα που προέρχεται από ένα πεπερασμένο σύνολο χρωμάτων. Στις περισσότερες των περιπτώσεων, μια ψηφιακή εικόνα είναι ένα ορθογώνιο, διαιρεμένο με γραμμές και στήλες σε ορθογώνιες περιοχές που κάθε μία έχει συγκεκριμένο χρώμα. Μια τέτοια περιοχή ονομάζεται στοιχείο της εικόνας ή εικονοστοιχείο. Στην αγγλική λέγεται *pixel* ή *pel*, όρος ο οποίος προέρχεται από τη σύντμηση των λέξεων *picture element*. Από τα παραπάνω καταλαβαίνουμε πως μια ψηφιακή εικόνα είναι ένα διδιάστατο ψηφιακό σήμα.

Αν κάθε χρώμα κωδικοποιηθεί με έναν αριθμό τότε η ορθογώνια ψηφιακή εικόνα περιγράφεται από έναν πίνακα αριθμών $J \times K$, όπου J το πλήθος των γραμμών και K το πλήθος των στηλών της ψηφιακής εικόνας. Η τιμή $I(j,k)$ με $k=0,1,2,\dots,K-1$ και $j=0,1,2,\dots,J-1$ είναι ο κωδικός του χρώματος της ψηφιακής εικόνας.

Η μετατροπή μιας εικόνας σε ψηφιακή μορφή ουσιαστικά είναι η μετατροπή ενός δισδιάστατου αναλογικού σήματος σε ψηφιακό και απαιτεί τις διαδικασίες της δειγματοληψίας και του κβαντισμού.

1.1 Είδη ψηφιακών εικόνων

Υπάρχουν τρία είδη ψηφιακών εικόνων που χαρακτηρίζονται από το πλήθος των χρωμάτων που περιέχουν:

- 1) *Δυαδικές εικόνες (binary images)*: Κάθε εικονοστοιχείο των εικόνων μπορεί να χρωματιστεί με ένα από δύο χρώματα (συνήθως άσπρο ή μαύρο). Για κάθε εικονοστοιχείο απαιτείται ένα bit πληροφορίας, π.χ. με τιμή μηδέν (0) για το μαύρο και ένα (1) για λευκό. Οι εικόνες των εγγράφων που αποτελούνται μόνο από το χρώμα του χαρτιού και της μελάνης αναπαρίστανται σε δυαδική ψηφιακή μορφή.
- 2) *Εικόνες αποχρώσεων του γκρι (gray level images)*: Κάθε εικονοστοιχείο των εικόνων μπορεί να χρωματιστεί με μία από τις αποχρώσεις του γκρι οι οποίες ξεκινούν από το μαύρο και καταλήγουν στο λευκό. Από αυτές τις αποχρώσεις συνήθως λαμβάνονται 256 αντιπροσωπευτικές που κωδικοποιούνται με τιμές $0,1,2,\dots,255$. Η απόχρωση κάθε εικονοστοιχείου προφανώς απαιτεί πληροφορία ενός byte.
- 3) *Έγχρωμες εικόνες (color images)* στις οποίες κάθε εικονοστοιχείο χρωματίζεται με χρώματα που προέρχονται από την ανάμειξη των αποχρώσεων του κόκκινου, πράσινου και μπλε (*RGB*). Για κάθε ένα από τα τρία αυτά χρώματα λαμβάνονται 256 αποχρώσεις δηλαδή πληροφορία του ενός byte. Συνεπώς κάθε εικονοστοιχείο της έγχρωμης εικόνας, απαιτεί 3 bytes.

1.2 Βασικές έννοιες

Το σύνολο των χρωμάτων που μπορούν να χρησιμοποιηθούν για τον χρωματισμό των εικονοστοιχείων της εικόνας λέγεται χρωματική παλέτα. Εάν C είναι το πλήθος των χρωμάτων, τότε για την κωδικοποίησή τους απαιτούνται B bits και ισχύουν οι σχέσεις

$$C=2^B \Leftrightarrow B=\log_2 C$$

Το B ονομάζεται *βάθος bit (bit depth)* της ψηφιακής εικόνας. Εάν η εικόνα έχει K στήλες και J γραμμές τότε για την απεικόνισή της απαιτούνται $J \times K \times B$ bits. Ο Πιν.1.2.1 παρουσιάζει ενδεικτικές τιμές των παραπάνω μεγεθών.

Είδος εικόνας	J	K	B	Bits	bytes
Δυαδική	100	100	2	20000	2500
Αποχρώσεων του γκρι	100	100	8	80000	10000
Έγχρωμη RGB	100	100	24	240000	30000
Δυαδική	512	512	2	524288	65536
Αποχρώσεων του γκρι	512	512	8	2097152	262144
Έγχρωμη RGB	512	512	24	6291456	786432

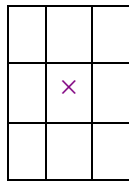
Πίνακας 1.2.1.

Ευκρίνεια της εικόνας είναι το πλήθος των εικονοστοιχείων ανά μονάδα επιφάνειας και καθορίζει πόσο λεπτομερής είναι η ψηφιακή αναπαράσταση της εικόνας. Η ευκρίνεια E μιας εικόνας διαστάσεων $J \times K$ και εμβαδού A δίνεται από την σχέση

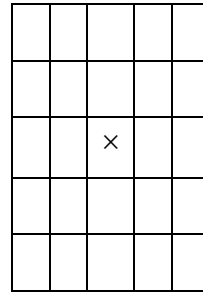
$$E = \frac{J \times K}{A} \quad (1.2.1)$$

και μετρείται σε πλήθος εικονοστοιχείων ανά μονάδα επιφάνειας π.χ. pixels/mm² ή dpi (dots per inch : κουκίδες ανά ίντσα). Η ευκρίνεια εξαρτάται τόσο από το πλήθος των εικονοστοιχείων όσο και από τις φυσικές διαστάσεις της εικόνας.

Μια ομάδα γειτονικών εικονοστοιχείων λέγεται γειτονιά. Σε μια γειτονιά $S_{M \times N}$ με M γραμμές και N στήλες μιας εικόνας διαστάσεων $J \times K$ υπάρχει ένα κεντρικό εικονοστοιχείο (j_c, k_c) όταν M, N είναι περιττοί αριθμοί. Η θέση των εικονοστοιχείων της S αναφέρονται συχνά, σχετικά με την θέση του κεντρικού εικονοστοιχείου της. Η πιο συνήθης γειτονιά είναι τριών (3) γραμμών και τριών (3) στηλών και λέγεται γειτονιά 3×3 . Στο Σχ.1.2.1 φαίνονται 3×3 και 5×5 γειτονιές.



Γειτονιά 3×3



Γειτονιά 5×5

Σχήμα 1.2.1: Γειτονιές 3×3 και 5×5 με τα κεντρικά τους εικονοστοιχεία.

Σε κάθε εικονοστοιχείο μιας γειτονιάς με M γραμμές και N στήλες όπου M και N περιττοί ακέραιοι, μπορούμε να αντιστοιχίσουμε έναν συντελεστή $w(m,n)$, $m=-(M-1)/2, \dots, (M-1)/2$, $n=-(N-1)/2, \dots, (N-1)/2$ στην γραμμή

$$r = (M-1)/2 + m + 1, r = 1 \dots M,$$

και στην στήλη

$$c = (N-1)/2 + n + 1, c = 1 \dots N,$$

Ονομάζουμε μάσκα W τον πίνακα S με στοιχεία τις τιμές τις μάσκας σύμφωνα με την σχέση

$$s_{rc} = w(m,n).$$

Για παράδειγμα για $N=M=3$ ο πίνακας S είναι

$$\begin{vmatrix} w(-1,-1) & w(-1,0) & w(-1,1) \\ w(0,-1) & w(0,0) & w(0,1) \\ w(1,-1) & w(1,0) & w(1,1) \end{vmatrix}$$

Οι συντελεστές της μάσκας και οι τιμές φωτεινότητας των εικονοστοιχείων μπορούν να εμπλακούν σε χρήσιμους υπολογισμούς για την επεξεργασία της εικόνας. Ο συνηθέστερος υπολογισμός δίνεται από την σχέση

$$A(j,k) = \sum_m \sum_n w(m,n) \cdot I(j+m, k+n)$$

όπου (j,k) εικονοστοιχείο της εικόνας I .

Στο παρακάτω σχήμα δίνονται μάσκες 3×3 και η εφαρμογή τους σε μία 3×3 περιοχή εικονοστοιχείων.

$$1/9 \begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline \end{array} \quad \begin{array}{|c|c|c|} \hline 240 & 250 & 240 \\ \hline 240 & 100 & 150 \\ \hline 240 & 250 & 250 \\ \hline \end{array}$$

$$\frac{240 + 250 + 240 + 240 + 100 + 150 + 240 + 250 + 250}{9} \approx 191$$

$$1/17.64 \begin{array}{|c|c|c|} \hline 1.41 & 1 & 1.41 \\ \hline 1 & 8 & 1 \\ \hline 1.41 & 1 & 1.41 \\ \hline \end{array} \quad \begin{array}{|c|c|c|} \hline 240 & 250 & 240 \\ \hline 240 & 100 & 150 \\ \hline 240 & 250 & 250 \\ \hline \end{array}$$

$$\frac{1.41 \cdot 240 + 250 + 1.41 \cdot 240 + 240 + 8 \cdot 100 + 150 + 1.41 \cdot 240 + 250 + 1.41 \cdot 250}{17.64} \approx 138$$

Σχήμα 1.2.2.

Αν (j_c, k_c) είναι το κεντρικό εικονοστοιχείο της γειτονιάς S της εικόνας $I_{J \times K}$ η τιμή A μπορεί να αποδοθεί ως τιμή φωτεινότητας του εικονοστοιχείου (j_c, k_c) μιας νέας εικόνας $I'_{J \times K}$. Αν αυτό εφαρμοσθεί για όλες τις γειτονιές της εικόνας I , τότε λέμε ότι η νέα εικόνα I' προέκυψε από το φιλτράρισμα της I με την μάσκα W . Αν περιγράψουμε μια τέτοια πράξη με όρους της ψηφιακής επεξεργασίας σήματος τότε η εφαρμογή της μάσκας W σε όλα τα εικονοστοιχεία της εικόνας ισοδυναμεί με την έξοδο ενός γραμμικού και ανεξάρτητου από την μετατόπιση (LTI) συστήματος με απόκριση κρουστικής διέγερσης της μορφής

$$h(j, k) = \sum_m \sum_n w(m, n) \cdot \delta(j + m, k + n)$$

h=	w(1,1)	w(1,0)	w(1,-1)
	w(0,1)	w(0,0)	w(0,-1)
	w(-1,1)	w(-1,0)	w(-1,-1)

και θα δίνεται από τη σχέση $I' = I ** h$

1.3. Ιστόγραμμα

Αν η ψηφιακή εικόνα αποτελείται από αποχρώσεις του γκρι μπορούμε να δημιουργήσουμε την κατανομή του πλήθους των εικονοστοιχείων που έχουν την ίδια τιμή

απόχρωσης για κάθε απόχρωση. Η κατανομή αυτή λέγεται ιστόγραμμα των αποχρώσεων της εικόνας και δίνεται μαθηματικά από την σχέση

$$h(g) = \sum_{\forall I(j,k)=g} 1 \quad (1.3.1)$$

όπου $g=0,1,2..G-1$, G το πλήθος των αποχρώσεων, $j=0, \dots, J-1$, $k=0, \dots, K-1$, J το πλήθος των γραμμών, K το πλήθος των στηλών της εικόνας, $I(j,k)$ η τιμή της φωτεινότητας στο σημείο (j,k) και $h(g)$ το πλήθος των εικονοστοιχείων με απόχρωση g . Το ιστόγραμμα δηλαδή μας πληροφορεί πόσα εικονοστοιχεία υπάρχουν με συγκεκριμένη τιμή απόχρωσης. Η συνάρτηση

$$H(g) = \frac{h(g)}{\sum_{g=0}^{G-1} h(g)} \quad (1.3.2)$$

λέγεται κανονικοποιημένο ιστόγραμμα και δίνει την πιθανότητα ένα τυχαίο εικονοστοιχείο της εικόνας να έχει απόχρωση g .

ΚΕΦΑΛΑΙΟ 2

ΕΠΕΞΕΡΓΑΣΙΕΣ ΣΕ ΔΥΑΔΙΚΕΣ ΕΙΚΟΝΕΣ

2.1 Κέντρο βάρους αντικειμένου.

Μία ψηφιακή δυαδική εικόνα με J γραμμές και K στήλες παριστάνεται με έναν πίνακα $I_{J \times K}$ και κάθε στοιχείο I_{jk} , $j=0, \dots, J-1$, $k=0, \dots, K-1$ παίρνει τιμή μηδέν (0) ή ένα (1). Αν η εικόνα αναπαριστά ένα αντικείμενο τότε το πλήθος των εικονοστοιχείων του αντικειμένου που έχουν τιμή '1' δίνεται από την σχέση

$$N = \sum_{j=0}^{J-1} \sum_{k=0}^{K-1} I_{jk}$$

Το κέντρο βάρους του αντικειμένου βρίσκεται στην θέση (\bar{j}, \bar{k}) της εικόνας σύμφωνα με τις σχέσεις

$$\bar{k} = \frac{\sum_{j=0}^{J-1} \sum_{k=0}^{K-1} k \cdot I_{jk}}{N},$$

$$\bar{j} = \frac{\sum_{j=0}^{J-1} \sum_{k=0}^{K-1} j \cdot I_{jk}}{N}$$

2.2 Κωδικοποίηση δυαδικής εικόνας κατά μήκος διαδρομής

Μία δυαδική εικόνα μπορεί να κωδικοποιηθεί με κατάλληλο αλγόριθμο ώστε να ελαττωθεί η απαιτούμενη ποσότητα πληροφορίας για την αποθήκευση ή την μετάδοσή της και να διευκολυνθεί η αναγνώριση του περιεχομένου της. Ένας τέτοιος αλγόριθμος είναι η *κωδικοποίηση κατά μήκος διαδρομής* (*RLE: Run Length Encoding*). Σύμφωνα με αυτόν ονομάζουμε *συστοιχία* μία ομάδα από διαδοχικά εικονοστοιχεία με την ίδια τιμή (όλα '1' ή όλα '0') και το πλήθος των εικονοστοιχείων ως *μήκος της συστοιχίας*. Διατρέχουμε κάθε σειρά της εικόνας και γράφουμε την θέση του πρώτου εικονοστοιχείου και το μήκος κάθε

συστοιχίας από '1' ή εναλλακτικά την θέση του πρώτου και του τελευταίου εικονοστοιχείου κάθε συστοιχίας. Μία άλλη προσέγγιση είναι να γράψουμε το μήκος των διαδοχικών συστοιχειών από '1' και '0' δεχόμενοι ότι η πρώτη συστοιχία αποτελείται πάντα από '1' έστω και μηδενικού μήκους . Στο Σχ.2.2.1 που ακολουθεί φαίνεται η κωδικοποίηση των τριών πρώτων γραμμών μιας δυαδικής εικόνας με τους τρόπους που αναφέρθηκαν.

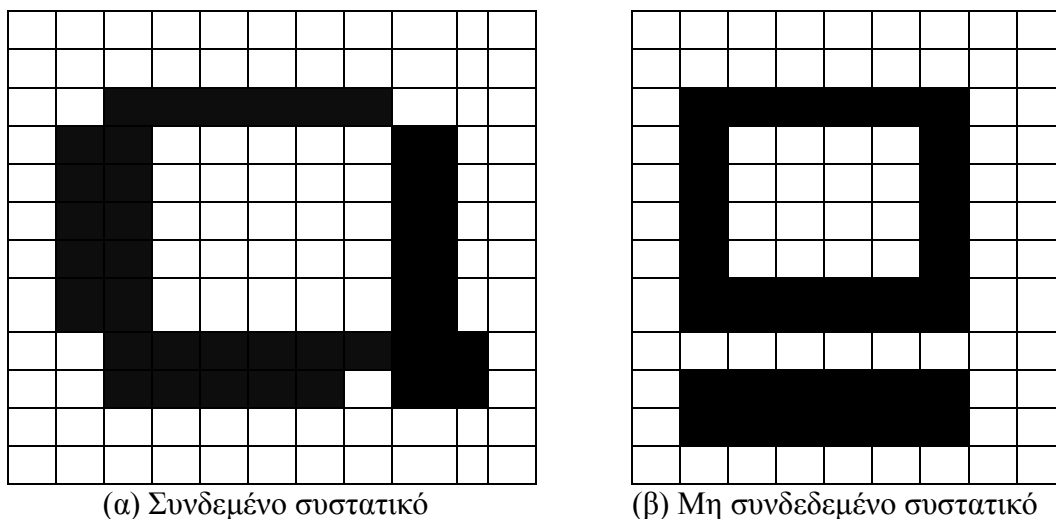
1	1	1	0	0	0	0	1	1	0	0
0	0	0	0	0	1	1	0	0	1	1
1	1	1	1	1	0	0	1	1	1	0

(0,3),(7,2)	(0,2),(8,9)	3,4,2,2
(5,2),(9,2)	(5,6),(9,10)	0,5,2,2,2
(0,5),(7,3)	(0,4),(7,9)	5,2,3,1

Σχήμα 2.2.1.

2.3 Συνδεδεμένα στοιχεία.

Θεωρούμε ένα σύνολο S εικονοστοιχείων μιας δυαδικής εικόνας με την ίδια τιμή (0 ή 1). Δύο εικονοστοιχεία ρ_1 και ρ_2 που ανήκουν στο S ονομάζονται *συνδεδεμένα* όταν υπάρχει διαδρομή από εικονοστοιχεία του S που οδηγεί από το ρ_1 στο ρ_2 . Ένα σύνολο εικονοστοιχείων λέγεται *συνδεδεμένο συστατικό (connected component)* όταν όλα τα εικονοστοιχεία του είναι μεταξύ τους συνδεδεμένα. Στο ακόλουθο Σχ.2.3.1 φαίνεται ένα συνδεδεμένο και ένα μη συνδεδεμένο συστατικό δυαδικής εικόνας.



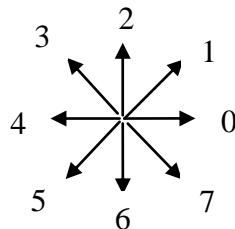
(α) Συνδεδεμένο συστατικό

(β) Μη συνδεδεμένο συστατικό

Σχήμα 2.3.1

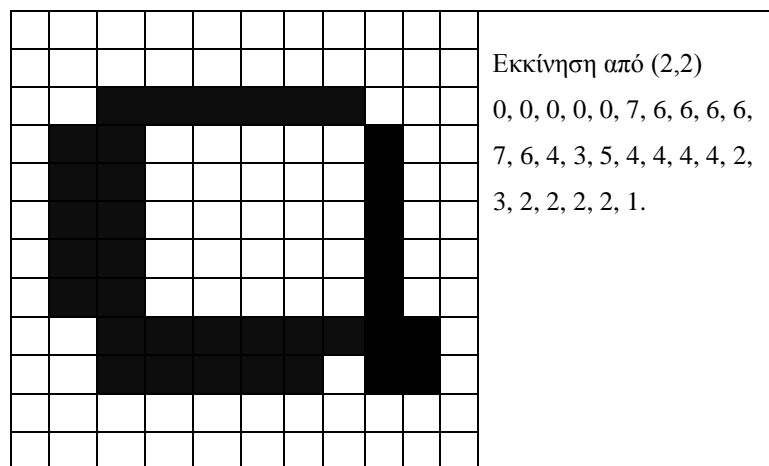
2.4 Κωδικοποίηση αλυσίδας.

Με την κωδικοποίηση αλυσίδας κωδικοποιούμε το περίγραμμα ενός αντικειμένου της εικόνας που είναι συνδεδεμένο. Για τον σκοπό αυτό καθορίζουμε και κωδικοποιούμε τις διευθύνσεις που ξεκινούν από το κεντρικό εικονοστοιχείο και καταλήγουν στα γειτονικά του σε μια 3×3 γειτονιά. όπως φαίνεται στο Σχ.2.4.1



Σχήμα 2.4.1.

Ακολουθως ξεκινώντας από οποιοδήποτε εικονοστοιχείο του εξωτερικού περιγράμματος του αντικειμένου (συνήθως το πάνω αριστερό εικονοστοιχείο) διατρέχουμε το περίγραμμα γράφοντας τον κωδικό της σχετικής διεύθυνσης κάθε εικονοστοιχείου με το επόμενο του. Στο Σχ.2.4.2 δείχνεται η εφαρμογή της κωδικοποίησης αλυσίδας.



Σχήμα 2.4.2

2.5 Ο μετασχηματισμός του Hough

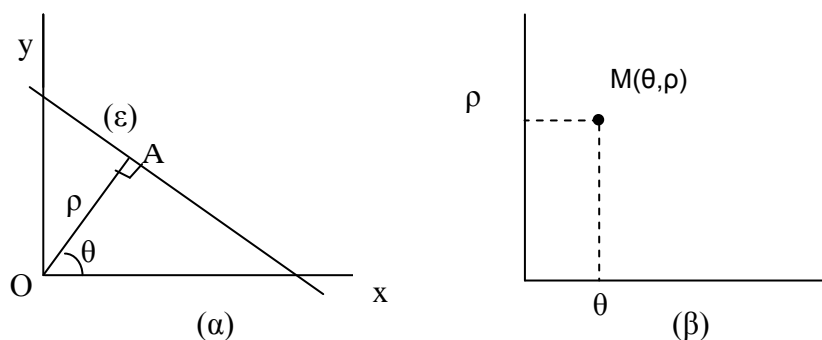
Συχνά οι ψηφιακές εικόνες περιέχουν γραμμές που ανήκουν σε σχήματα, τεχνικά σχέδια, γραφήματα, σειρές κειμένου ή άλλες αναπαραστάσεις. Σε πολλές εφαρμογές είναι επιθυμητή η εύρεση της θέσης και η αναγνώριση της μορφής των γραμμών (π.χ. ευθύγραμμα τμήματα, τόξα). Τέτοιες εφαρμογές είναι: η διανυσματική κωδικοποίηση τυπωμένων τεχνικών σχεδίων που ψηφιοποιήθηκαν από σαρωτές (scanners), η εύρεση περιοχών κειμένου σε έγγραφα, ο προσδιορισμός της υφής, βιομηχανικές εφαρμογές κατασκευών και ρομποτικής κ.α. Ο μετασχηματισμός του Hough (HT: Hough Transform) καταδεικνύει την ύπαρξη ευθειών σε μια εικόνα και αποτελεί την βάση πολλών τεχνικών για τον προσδιορισμό ευθυγράμμων τμημάτων, καμπυλών και μορφών που αναλύονται ανάλογα.

Ο HT προτάθηκε από τον Paul Hough το 1962 ως μέρος της κατασκευής μιας συσκευής ανίχνευσης της κίνησης σωματιδίων υψηλής ενέργειας και στόχευε στην αυτοματοποίηση και αντικατάσταση της οπτικής διαδικασίας που απαιτούσε πολλές ανθρωποώρες. Ο αρχικός αλγόριθμος εξελίχθηκε και η σημερινή διατύπωση του είναι η ακόλουθη.

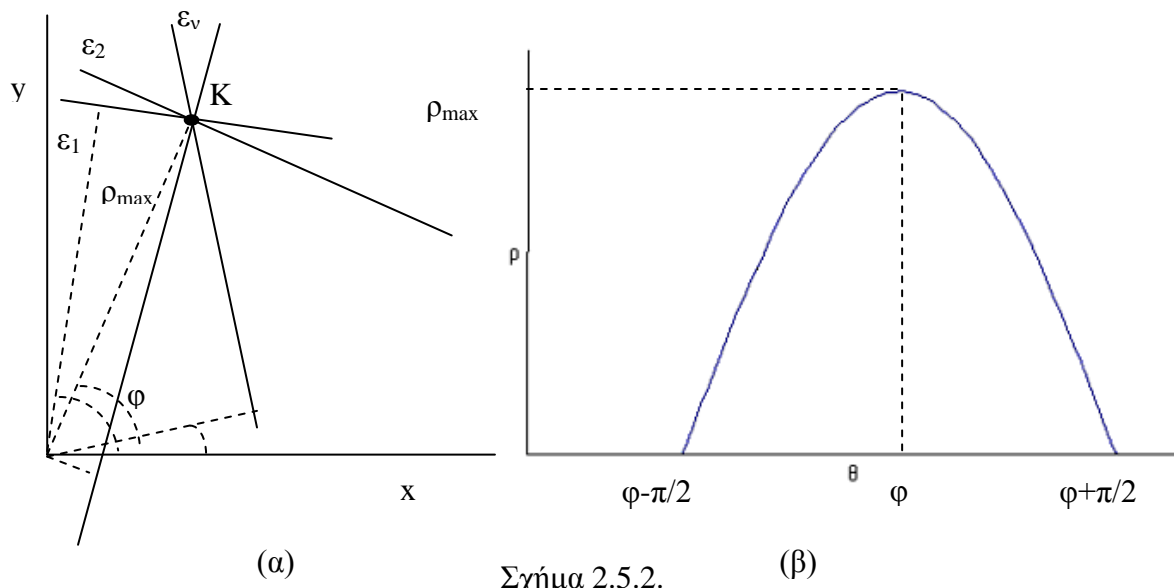
Κάθε ευθεία (ϵ) του καρτεσιανού επιπέδου Π_{xy} περιγράφεται από την πολική της εξίσωση

$$\rho = x \sin\theta + y \eta\mu\theta$$

όπου (x,y) σημείο της ευθείας και ρ, θ οι πολικές της παράμετροι (Σχ.2.5.1(α)).



Σχ. 2.5.1



Σχήμα 2.5.2.

Το ρ είναι το μήκος του ευθύγραμμου τμήματος OA , $OA \perp (\varepsilon)$ και θ η γωνία που σχηματίζει το OA με τον άξονα Ox . Ισχύει ότι $\rho \geq 0$ και $-\pi \leq \theta < \pi$. Θεωρούμε το επίπεδο $\Pi_{\theta\rho}$ όπου στον κάθετο άξονα σημειώνουμε τις τιμές του ρ και στον οριζόντιο τις τιμές του θ . Οι πολικές παράμετροι (θ, ρ) της ευθείας (ε) καθορίζουν ένα σημείο $M(\theta, \rho)$ στο επίπεδο $\Pi_{\theta\rho}$ (Σχ.2.5.1.(β))

Με τον τρόπο αυτό η ευθεία (ε) του επιπέδου Π_{xy} αντιστοιχίζεται (μετασχηματίζεται) με την χρήση της πολικής της εξίσωσης σε ένα σημείο του επιπέδου $\Pi_{\theta\rho}$. Ισοδύναμα κάθε σημείο $M(\theta, \rho)$ του επιπέδου $\Pi_{\theta\rho}$ ορίζει μια ευθεία $\varepsilon \rightarrow \rho = x \cos \theta + y \sin \theta$ στο επίπεδο Π_{xy} . Από ένα σημείο $K(x_k, y_k)$ του επιπέδου Π_{xy} διέρχονται άπειρες ευθείες (ε_v) , $v=1, 2, \dots$ που ικανοποιούν την σχέση

$$\rho_v = x_k \sin \theta_v + y_k \eta \mu \theta_v$$

(Σχ.2.5.2(α)). Οι τιμές ρ_v και θ_v των μεταβλητών ρ και θ ανήκουν στην καμπύλη $\rho = x_k \sin \theta + y_k \eta \mu \theta$ του επιπέδου $\Pi_{\theta\rho}$. Η καμπύλη είναι ημιτονοειδής (Σχ.2.5.2(β)), όπως δείχνεται ακολούθως. Έστω ϕ η γωνία χOK . Ισχύει ότι

$$\varepsilon \phi \phi = \frac{y_k}{x_k} \Leftrightarrow x_k = y_k \sigma \phi \phi = y_k \varepsilon \phi \left(\frac{\pi}{2} - \phi \right) \Leftrightarrow$$

$$\rho = y_k \frac{\eta \mu \left(\frac{\pi}{2} - \phi \right)}{\sigma \nu \left(\frac{\pi}{2} - \phi \right)} \sigma \nu \theta + y_k \eta \mu \theta \Leftrightarrow$$

$$\rho = \frac{y_k}{\eta \mu \phi} \eta \mu \left(\theta + \frac{\pi}{2} - \phi \right)$$

Αν x_k και y_k είναι θετικά (1° τεταρτημόριο) και επειδή $\rho \geq 0$ έπεται ότι

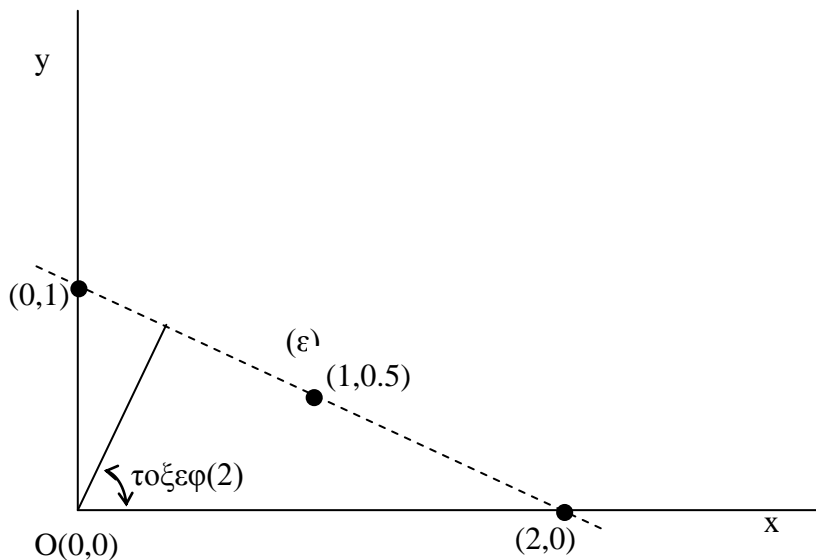
$$0 \leq \phi \leq \frac{\pi}{2} \quad (\alpha)$$

$$\phi - \frac{\pi}{2} \leq \theta \leq \phi + \frac{\pi}{2}, \quad (\beta)$$

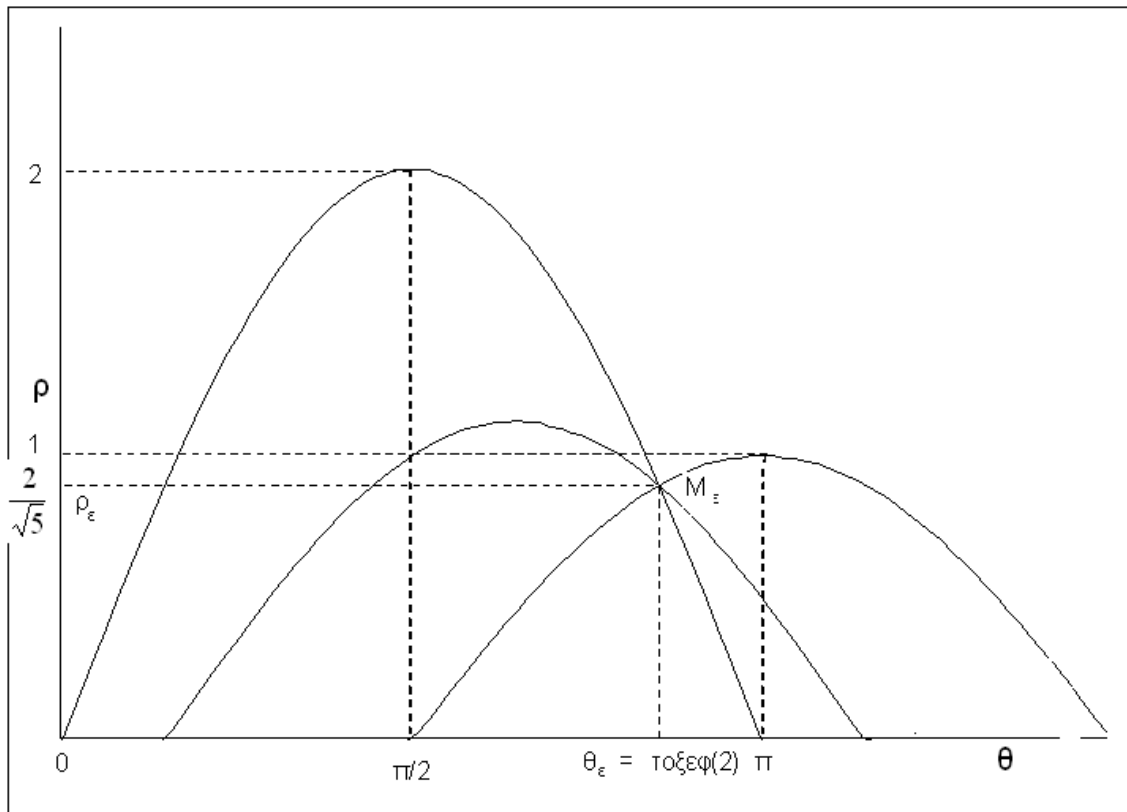
η μεγαλύτερη τιμή του ρ προκύπτει για $\theta = \phi$ και είναι

$$\rho_{\max} = \sqrt{x_k^2 + y_k^2} \quad (\gamma)$$

Ο αναλυτικός υπολογισμός όλων των δυνατών τιμών των ρ και θ των ευθειών που ορίζουν ανά δύο τα σημεία της εικόνας και η εύρεση κοινών τιμών είναι επίπονος και πρακτικά ασύμφορος. Για αυτό ακολουθούμε το ακόλουθο σκεπτικό. Έστω τρία τουλάχιστον συνευθειακά σημεία στο επίπεδο Π_{xy} με τις αντίστοιχες καμπύλες των ευθειών που διέρχονται από αυτά (Σχ.2.5.3(α)). Οι καμπύλες αυτές τέμνονται σε σημείο $M_\varepsilon(\theta_\varepsilon, \rho_\varepsilon)$ με θ_ε και ρ_ε τις πολικές παραμέτρους της ευθείας που διέρχεται από αυτά (Σχ.2.5.3(β)). Άρα αν κατασκευάσουμε τις καμπύλες του επιπέδου $\Pi_{\rho\theta}$ για όλα τα σημεία του επιπέδου Π_{xy} τα σημεία τομής τους έχουν συντεταγμένες τις πολικές παραμέτρους των ευθειών που ορίζουν τα σημεία αυτά. Προφανώς, όσες καμπύλες διέρχονται από ένα σημείο τομής στο $\Pi_{\rho\theta}$, τόσα σημεία του Π_{xy} ανήκουν στην ευθεία που ορίζεται από το σημείο τομής.



Σχ.2.5.2(α).



Σχ.2.5.2(β).

Προγραμματιστικά και υπολογιστικά η μέθοδος μπορεί να εφαρμοσθεί με τεχνικές όπως η ακόλουθη. Ορίζεται ένας δισδιάστατος πίνακας $H_M \times N$ που κβαντίζει το επίπεδο $\Pi_{\theta\rho}$ σε M γραμμές και N στήλες με μηδενικές αρχικές τιμές. Η ψηφιακή εικόνα με πίνακα $I_J \times K$ τοποθετείται στο 1^ο τεταρτημόριο. Οι συντεταγμένες ενός εικονοστοιχείου $P(k,j)$ με $I(k,j)=1$ θα είναι $x_P=k$, $y_P=J-j$. Επειδή η εικόνα βρίσκεται στο 1^ο τεταρτημόριο το διάστημα μεταβολής της γωνίας θ μπορεί να παίρνει τιμές από $-\pi/2$ έως π . Με βήμα

$$\Delta\theta = \frac{3\pi}{2N}$$

κάθε δείκτης $v=1, \dots, N$ αντιστοιχίζεται στο διάστημα $[\theta_v, \theta_v + \Delta\theta)$, $\theta_v = -\pi/2 + (v-1)\Delta\theta$.

Με βήμα

$$\Delta\rho = \frac{\sqrt{K^2 + J^2}}{M}$$

κάθε δείκτης $\mu=1, \dots, M$ αντιστοιχίζεται στο διάστημα $[(\mu-1)\Delta\rho, \mu\Delta\rho)$.

Για εκείνα τα v που ισχύει

$$\phi - \frac{\pi}{2} \leq \theta_v \leq \phi + \frac{\pi}{2}, \quad \phi = \text{τοξεφ}\left(\frac{y_P}{x_P}\right)$$

υπολογίζεται η τιμή $\rho = x_p \sin\theta_v + y_p \eta\mu\theta_v$ και ο δείκτης μ του διαστήματος που ανήκει το ρ . Η τιμή $H(\mu, \nu)$ αυξάνεται κατά ένα. Η διαδικασία επαναλαμβάνεται για όλα τα εικονοστοιχεία της εικόνας. Μετά το τέλος της διαδικασίας οι υψηλές τιμές του πίνακα H προσδιορίζουν τις πολικές παραμέτρους ευθειών της εικόνας. Συγκεκριμένα αν $H(\mu, \nu)$ έχει υψηλή τιμή ισχύει ότι

$$\theta = \theta_v + \frac{\Delta\theta}{2} = -\frac{\pi}{2} + \nu\Delta\theta - \frac{\Delta\theta}{2},$$

$$\rho = \rho_\mu + \frac{\Delta\rho}{2} = \mu\Delta\rho - \frac{\Delta\rho}{2}$$

ΚΕΦΑΛΑΙΟ 3

ΕΠΕΞΕΡΓΑΣΙΕΣ ΣΕ ΕΙΚΟΝΕΣ ΑΠΟΧΡΩΣΕΩΝ ΤΟΥ ΓΚΡΙ

3.1.1 Εξομάλωση

Με κατάλληλες τεχνικές είναι δυνατή η βελτίωση της οπτικής εμφάνισης μιας εικόνας ώστε να είναι καλύτερα αντιληπτή από τον άνθρωπο ή κατάλληλη για περαιτέρω επεξεργασία. Συχνά στις εικόνες εμφανίζονται μικροκουκίδες που οφείλονται σε τυχαίους παράγοντες (θόρυβος). Το *φίλτρο της μέσης τιμής* (*median filter*) είναι μια απλή τεχνική εξάλειψης και εξομάλωσης του θορύβου αυτού από ψηφιακές εικόνες αποχρώσεων του γκρι. Σύμφωνα με την τεχνική αυτή από την αρχική εικόνα I παράγεται μία νέα εικόνα I' ίδιων διαστάσεων κάθε εικονοστοιχείο (j,k) της οποίας έχει φωτεινότητα $I'(j,k)$ που είναι η μέση τιμή των τιμών φωτεινότητας μιας γειτονιάς του εικονοστοιχείου (j,k) της εικόνας I . Αν για παράδειγμα κάθε περιοχή S είναι 3×3 τότε το $I'(j_c, k_c)$ δίνεται από την σχέση

$$I'_{jk} = \frac{\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} I_{j-1+m, k-1+n}}{9} \quad (3.1.1.1)$$

Με άλλα λόγια η εικόνα I' είναι το αποτέλεσμα του φιλτραρίσματος της αρχικής εικόνας

$$I \text{ με την μάσκα } W = \frac{1}{9} \begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix}.$$

Είναι δυνατόν να χρησιμοποιηθούν και άλλες μάσκες για την εξομάλυνση μιας εικόνας

όπως για παράδειγμα η $W = \frac{1}{16} \begin{vmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{vmatrix}$. Μάσκες 5x5 ή μεγαλύτερες εξομαλύνουν ακόμα

περισσότερο την εικόνα.

Η μάσκα W_G , $N \times N$ με συντελεστές που δίνονται από τις σχέσεις

$$w_G(j, k) = \frac{G(j, k)}{\sum_{j=-\frac{(N-1)}{2}}^{\frac{(N-1)}{2}} \sum_{k=-\frac{(N-1)}{2}}^{\frac{(N-1)}{2}} G(j, k)}$$

με

$$G(j, k) = \frac{1}{2\pi\sigma^2} e^{-\frac{j^2+k^2}{2\sigma^2}}$$

λέγεται μάσκα Gauss ή φίλτρο Gauss (Gaussian filter) και χρησιμοποιείται ευρέως για την εξομάλυνση της εικόνας. Οι τιμές της για $N=3$ και $\sigma=1$, είναι

$$\begin{matrix} 0.0751 & 0.1238 & 0.0751 \\ 0.1238 & 0.2042 & 0.1238 \\ 0.0751 & 0.1238 & 0.0751 \end{matrix}$$

Οι μάσκες που περιγράψαμε αποτελούν γραμμικά συστήματα που λειτουργούν ως κατωδιαβατά φίλτρα. Ως εκ τούτου επηρεάζουν τις τιμές των περιοχών της εικόνας που εμφανίζονται απότομες μεταβολές τους (ακμές).

Μια άλλη μέθοδος εξομάλυνσης είναι το *φίλτρο της ενδιάμεσης τιμής*. Σύμφωνα με την τεχνική αυτή οι τιμές των εικονοστοιχείων μιας γειτονιάς ταξινομούνται και η τιμή του εικοστοιχείου $I'(j,k)$ είναι η μεσαία από τις τιμές των εικονοστοιχείων της γειτονιάς του $I(j,k)$. Στο Σχ.3.1.1 που ακολουθεί φαίνονται οι τιμές που προκύπτουν από την εξομάλυνση ενός μέρους μιας εικόνας α) με το φίλτρο της μέσης τιμής και β) με το φίλτρο της ενδιάμεσης τιμής.

250	250	250	250	250	250
250	250	250	100	100	250
250	250	50	100	10	250

250	250	50	100	10	250
250	250	100	250	100	250
250	250	250	250	250	250
250	10	250	50	50	250
250	250	250	250	250	250

Τιμές φωτεινότητας της αρχικής εικόνας.

250	250	250	250	250	250
250	228	194	151	173	250
250	206	156	86	130	250
250	189	156	86	147	250
250	211	194	151	190	250
250	207	184	172	189	250
250	223	201	206	206	250
250	250	250	250	250	250

Οι τιμές φωτεινότητας μετά από εξομάλυνση με το φίλτρο της μέσης τιμής.

250	250	250	250	250	250
250	250	250	100	250	250
250	250	100	100	100	250
250	250	100	100	100	250
250	250	250	100	250	250
250	250	250	250	250	250
250	250	250	250	250	250
250	250	250	250	250	250

Οι τιμές φωτεινότητας μετά από εξομάλυνση με το φίλτρο της ενδιάμεσης τιμής.

Σχήμα 3.1.1.

Με την εξομάλυνση γενικά αμβλύνονται οι ακμές της εικόνας. Το φίλτρο της μέσης τιμής απομακρύνει τον θόρυβο χωρίς να επηρεάσει ιδιαίτερα τις ακμές της εικόνας. Το φίλτρο της ενδιάμεσης τιμής απομακρύνει τον θόρυβο και εξομαλύνει τις ακμές της εικόνας.

3.1.2. Εξισορρόπηση του ιστογράμματος

Η διάκριση γειτονικών περιοχών σε εικόνες αποχρώσεων του γκρι, είναι δύσκολη όταν η διαφορά τιμών των αποχρώσεων είναι μικρή. Η αύξηση των χρωματικών αντιθέσεων μιας εικόνας διευκολύνει την διάκριση των περιοχών αυτών. Μια τεχνική για την επίτευξη αυτού του σκοπού είναι η *εξισορρόπηση του ιστογράμματος* της εικόνας (*Histogram equalization*). Σύμφωνα με αυτήν οι τιμές των αποχρώσεων των εικονοστοιχείων μεταβάλλονται έτσι ώστε να αυξηθεί η αντίθεση μεταξύ των περιοχών με διαδοχικές τιμές αποχρώσεων ανάλογα με το ποσοστό του πλήθους των εικονοστοιχείων τους. Συγκεκριμένα, αν G είναι το πλήθος όλων των αποχρώσεων της παλέτας του γκρι, $h(g)$ το κανονικοποιημένο ιστόγραμμα της αρχικής εικόνας I_{kxj} , $g=0,1,\dots,G-1$ και $P(g)$ η συνάρτηση αθροιστικής πιθανότητας του g , προκύπτει μία νέα εικόνα I' ίδιων διαστάσεων με τη I , κάθε εικονοστοιχείο (k,j) της οποίας έχει απόχρωση $I'(k,j) = [(G-1)P(I(k,j))]$. Με άλλα λόγια κάθε τιμής g μιας απόχρωσης της αρχικής εικόνας αντικαθίσταται από μια τιμή $g' = [(G-1)P(g)]$. Υπενθυμίζεται ότι για την $P(g)$ ισχύουν οι σχέσεις

$$P(g) = \sum_{i=0}^g h(i) \text{ άρα και } P(g) = P(g-1) + h(g). \quad (3.1.2.1)$$

Παράδειγμα: Έστω ο ακόλουθος πίνακας τιμών των αποχρώσεων του γκρι μιας ψηφιοποιημένης εικόνας:

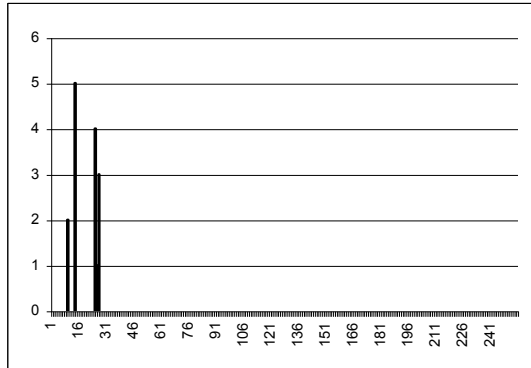
10	10	14	14	14
14	14	25	25	25
25	26	27	27	27

Οι αποχρώσεις g' προκύπτουν σύμφωνα με τους ακόλουθους υπολογισμούς:

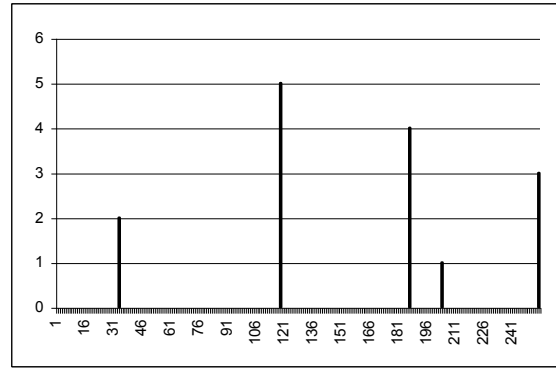
g	$H(g)$	$h(g)$	$P(g)$	g'
10	2	2/15	2/15	$255 \cdot 2/15 = \mathbf{34}$
14	5	5/15	7/15	$255 \cdot 7/15 = \mathbf{119}$
25	4	4/15	11/15	$255 \cdot 11/15 = \mathbf{187}$
26	1	1/15	12/15	$255 \cdot 12/15 = \mathbf{204}$

27	3	3/15	15/15	$255 * 15/15 = 255$
----	---	------	-------	---------------------

Στο Σχ.3.1.2.1 δείχνονται τα ιστογράμματα $h(g)$ και $h(g')$.



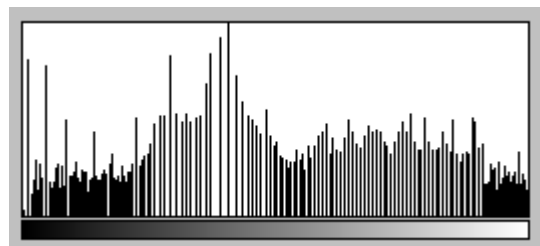
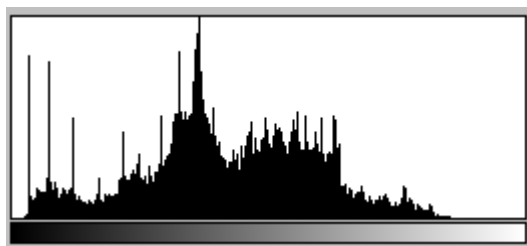
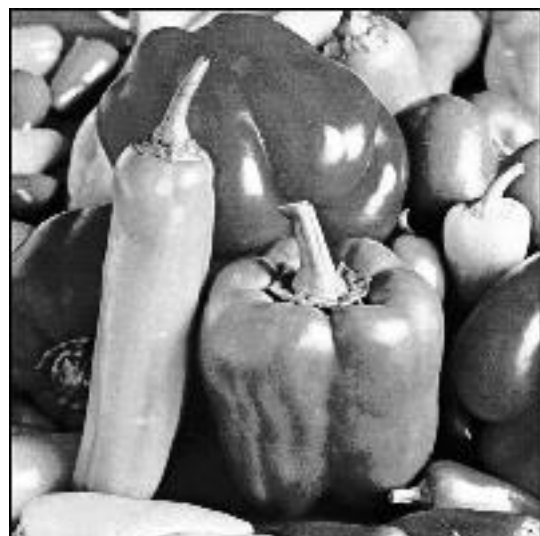
Το ιστόγραμμα $h(g)$



Το ιστόγραμμα $h(g')$

Σχήμα 3.1.2.1

Στο Σχ.3.1.2.2 φαίνεται η εφαρμογή της εξισορρόπησης του ιστογράμματος στην εικόνα με τις πιπεριές



Σχήμα 3.1.2.1.

3.2 Κατωφλίωση

Συχνά τα εικονοστοιχεία ενός αντικειμένου μιας εικόνας παίρνουν τιμές σε ένα μικρό διάστημα αποχρώσεων. Αυτό οδηγεί συνήθως στη δημιουργία ενός τοπικού μέγιστου στην περιοχή του ιστογράμματος της εικόνας. Η εύρεση τέτοιων τοπικών μεγίστων διευκολύνει τον εντοπισμό των αντικειμένων της εικόνας και την απόδοσή της με λιγότερες κύριες αποχρώσεις. Παρακάτω θα περιγράψουμε διάφορες τεχνικές για τον καθορισμό τιμών του πεδίου των αποχρώσεων μεταξύ των οποίων εμφανίζονται τοπικά μέγιστα του ιστογράμματος. Οι τιμές αυτές λέγονται κατώφλια.

Για την μετατροπή μιας εικόνας αποχρώσεων του γκρι σε δυαδική είναι επιθυμητός ο προσδιορισμός των 2 κύριων αποχρώσεών της. Με τον τρόπο αυτό οι αποχρώσεις της εικόνας χωρίζονται σε 2 κλάσεις C_1 , C_2 που κάθε μια αποτελείται από τις αποχρώσεις που βρίσκονται πιο κοντά σε μια από τις 2 κύριες αποχρώσεις. Ο διαχωρισμός αυτός ισοδυναμεί με την εύρεση μιας τιμής κατωφλιού T για την οποία θα ισχύει

$$g \in \begin{cases} C_1 & \text{αν } g < T \\ C_2 & \text{αν } T \leq g \end{cases} \quad (3.2.1)$$

Ακολουθώντας από την αρχική εικόνα I προκύπτει η δυαδική εικόνα I' κάθε εικονοστοιχείο της οποίας δίνεται από τη σχέση

$$I'_{jk} = \begin{cases} 0 & \text{αν } I_{jk} \in C_1 (I_{jk} < T) \\ 1 & \text{αν } I_{jk} \in C_2 (I_{jk} \geq T) \end{cases} \quad (3.2.2)$$

3.2.1 Κατωφλίωση με βάση τη διασπορά

Μια μέθοδος για την εύρεση της τιμής κατωφλιού T είναι η μέθοδος του Otsu. Σύμφωνα με αυτήν η τιμή T πρέπει να είναι τέτοια ώστε η συνολική εσωτερική διασπορά σ_w^2 εντός των κλάσεων να είναι ελάχιστη και η συνολική διασπορά σ_b^2 μεταξύ των κλάσεων μέγιστη. Σύμφωνα με την μέθοδο δημιουργούμε το κανονικοποιημένο ιστόγραμμα $h(g)$ της εικόνας. Η συνάρτηση $h(g)$ δίνει την πιθανότητα να έχει ένα τυχαίο εικονοστοιχείο της εικόνας απόχρωση g .

Έστω T μια τιμή κατωφλιού που διαχωρίζει τις αποχρώσεις σε δύο κλάσεις C_1 ($g < T$) και C_2 ($T \leq g$). Η πιθανότητα ένα τυχαίο εικονοστοιχείο να έχει απόχρωση της κλάσεως C_1 είναι p_1 και δίνεται από τη σχέση

$$p_1 = \sum_{g=0}^{T-1} h(g) \quad (3.2.1.1)$$

όμοια για την κλάση C_2 είναι p_2 και δίνεται από τη σχέση

$$p_2 = \sum_{g=T}^{G-1} h(g) \quad (3.2.1.2)$$

Είναι προφανές ότι ισχύει

$$p_1 + p_2 = 1. \quad (3.2.1.3)$$

Η μέση τιμή μ_1 της πρώτης κλάσης είναι

$$\mu_1 = \frac{\sum_{g=0}^{T-1} h(g)g}{p_1} \quad (3.2.1.4)$$

Η μέση τιμή μ_2 της δεύτερης κλάσης είναι και ισούται με

$$\mu_2 = \frac{\sum_{g=T}^{G-1} h(g)g}{p_2} \quad (3.2.1.5)$$

Η συνολική μέση τιμή μ δίνεται από τη σχέση

$$\mu = \sum_{g=0}^{G-1} h(g)g \quad (3.2.1.6)$$

Από τις τρεις προηγούμενες σχέσεις εύκολα αποδεικνύεται η σχέση

$$\mu = p_1\mu_1 + p_2\mu_2 \quad (3.2.1.7)$$

Η διασπορά της πρώτης κλάσης δίνεται από τη σχέση

$$\sigma_1^2 = \frac{1}{p_1} \sum_{g=0}^{T-1} (g - \mu_1)^2 h(g) \quad (3.2.1.8)$$

και η διασπορά της δεύτερης κλάσης είναι

$$\sigma_2^2 = \frac{1}{p_2} \sum_{g=T}^{G-1} (g - \mu_2)^2 H(g) \quad (3.2.1.9)$$

Η συνολική διασπορά σ^2 είναι

$$\sigma^2 = \sum_{g=0}^{G-1} (g - \mu)^2 h(g) \quad (3.2.1.10)$$

Η συνολική εσωτερική διασπορά των κλάσεων ορίζεται από τη σχέση

$$\sigma_w^2 = p_1\sigma_1^2 + p_2\sigma_2^2 \quad (3.2.1.11)$$

Η συνολική διασπορά μεταξύ των κλάσεων ορίζεται από τη σχέση

$$\sigma_b^2 = p_1(\mu_1 - \mu)^2 + p_2(\mu_2 - \mu)^2 = p_1p_2(\mu_1 - \mu_2)^2 \quad (3.2.1.12)$$

Αναζητούμε τώρα την κατάλληλη τιμή του T ώστε η σ_w^2 να είναι ελάχιστη και σ_b^2 μέγιστη.

Προς τούτο θεωρούμε το πηλίκο

$$\lambda = \frac{\sigma_b^2}{\sigma_w^2} \quad (3.2.1.13)$$

και ζητάμε την τιμή του T για την οποία το λ μεγιστοποιείται.

Για να μειωθούν οι υπολογισμοί θεωρούμε τις σχέσεις

$$\left. \begin{array}{l} \frac{\lambda}{1+\lambda} = \frac{\sigma_b^2}{\sigma_w^2 + \sigma_b^2} \\ \sigma^2 = \sigma_w^2 + \sigma_b^2 \end{array} \right\} \Rightarrow \frac{\lambda}{1+\lambda} = \frac{\sigma_b^2}{\sigma^2} \quad (3.2.1.14)$$

Η σχέση $\sigma^2 = \sigma_w^2 + \sigma_b^2$ αποδεικνύεται στο παράρτημα Α. Επειδή το σ^2 είναι ανεξάρτητο του T και η συνάρτηση $\lambda/(1+\lambda)$ είναι αύξουσα συνεπάγεται πως για να είναι το λ μέγιστο αρκεί να είναι το σ_b^2 μέγιστο. Με τον τρόπο αυτό η τελικά η κατάλληλη τιμή του κατωφλίου T είναι εκείνη που μεγιστοποιεί την ποσότητα

$$\sigma_b^2 = \frac{p_1(\mu_1 - \mu)^2}{1 - p_1} \quad (3.2.1.15)$$

ΠΑΡΑΔΕΙΓΜΑ 3.2.1.1.

Δίνεται ο πίνακας ψηφιακής εικόνας αποχρώσεων του γκρι. Ποια η τιμή κατωφλίωσης της με τη μέθοδο της διασποράς;

210	210	88	100	100
210	200	100	100	210
88	210	210	200	210
100	88	100	210	100
88	100	88	100	100

ΛΥΣΗ:

g	H(g)	h(g)
88	5	0.2
100	10	0.4
200	2	0.08
210	8	0.32

$$\mu = \frac{88 \cdot 5 + 100 \cdot 10 + 200 \cdot 2 + 210 \cdot 8}{25} = 140.8$$

T	μ_1	ρ_1	$\sigma_b^2 = \frac{\rho_1(\mu - \mu_1)^2}{1 - \rho_1}$
88	88	0,2	1936
100	$(88 \times 5 + 100 \times 0,10) / 0,6 = 96$	$0,24 + 0,26 = 0,6$	13824
200	$(88 \times 5 + 100 \times 0,10 + 200 \times 2) / 0,68 = 108,24$	$0,6 + 0,2 = 0,68$	24894

Η μέγιστη τιμή του σ_b^2 προκύπτει για $T=200$ που αποτελεί σύμφωνα με τη μέθοδο η καλλίτερη τιμή για την κατωφλίωση

3.2.2. Κατωφλίωση με βάση την εντροπία

Μια διαφορετική προσέγγιση για την εύρεση κατάλληλης τιμής κατωφλίου T βασίζεται στην έννοια της εντροπίας. Σύμφωνα με την εκδοχή που παρουσίασε ο Karur, η τιμή T χωρίζει το κανονικοποιημένο ιστόγραμμα $h(g)$ σε δυο κλάσεις c_1 και c_2 . Η πρώτη κλάση μπορεί να θεωρηθεί ως πηγή πληροφορίας με σύμβολα $g = 0, \dots, T-1$ και αντίστοιχες πιθανότητες

$$\rho(g) = \frac{h(g)}{P_1} \quad (3.2.2.1)$$

όπου

$$P_1 = \sum_{g=0}^{T-1} h(g) \quad (3.2.2.2)$$

Η μέση πληροφορία (εντροπία) που προκύπτει της c_1 είναι

$$E_1 = - \sum_{g=0}^{T-1} \frac{h(g)}{P_1} \log \left(\frac{h(g)}{P_1} \right) = \log(P_1) - \frac{1}{P_1} \sum_{g=0}^{T-1} h(g) \log(h(g)) \quad (3.2.2.3)$$

Όμοια η μέση πληροφορία που προκύπτει από την c_2 είναι

$$\begin{aligned} E_2 &= - \sum_{g=T}^{255} \frac{h(g)}{P_2} \log \frac{h(g)}{P_2}, P_2 = 1 - P_1 \\ &= \log(1 - P_1) - \frac{1}{1 - P_1} \sum_{g=T}^{255} h(g) \log(h(g)) \end{aligned} \quad (3.2.2.4)$$

Ζητείται η πρώτη κλάση που θα αντιστοιχηθεί στην τιμή ‘0’ και η δεύτερη που θα αντιστοιχηθεί στο ‘1’ να προσφέρουν συνολικά την μεγαλύτερη δυνατή μέση πληροφορία. Δηλαδή η ζητούμενη τιμή του κατωφλίου T_k είναι εκείνη για την οποία $E_1(T)+E_2(T)$ γίνεται μέγιστο.

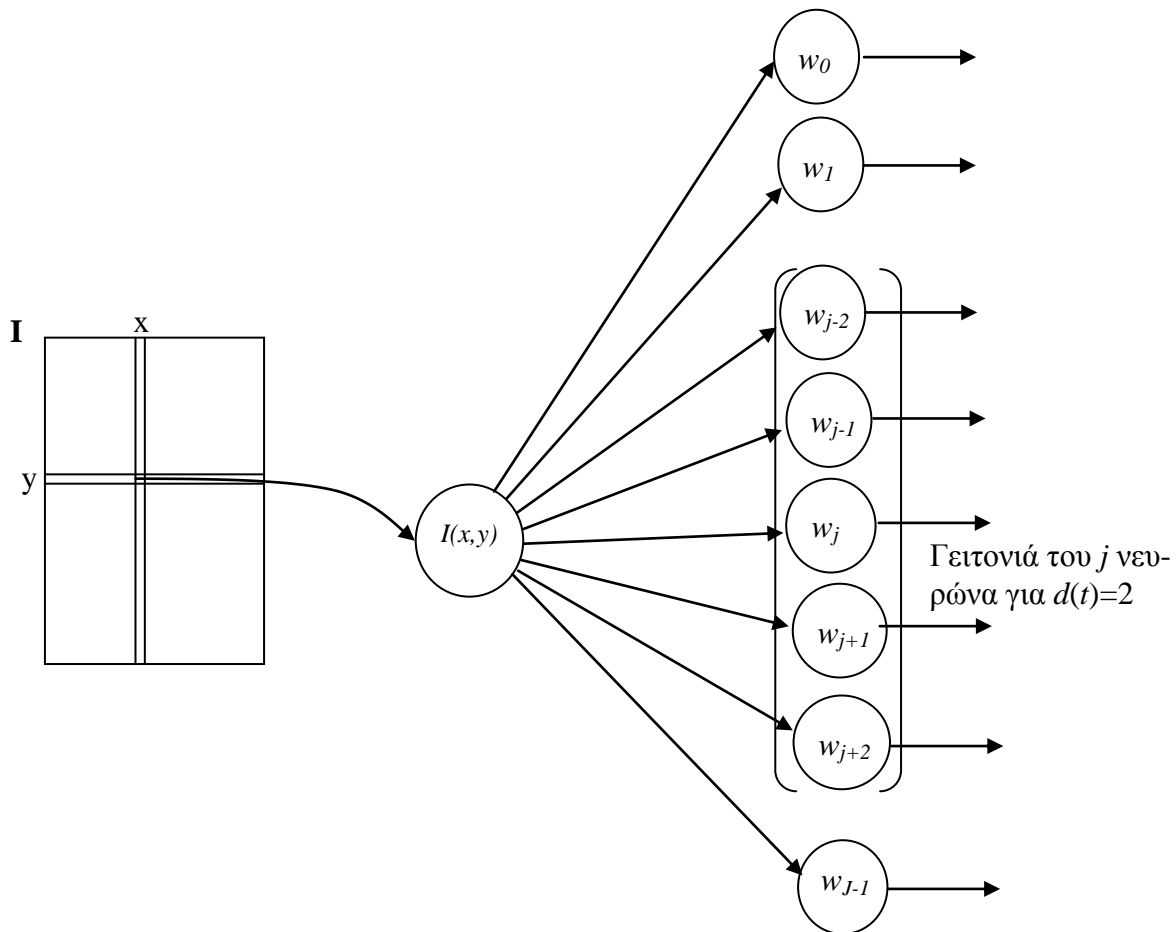
3.2.3. Πολυκατωφλίωση με χρήση νευρωνικού δικτύου

Εάν κάθε αντικείμενο της εικόνας αποτελείται από μία κύρια και κάποιες παραπλήσιες αποχρώσεις του γκρι, τότε δημιουργείται η συσσώρευση των τιμών του ιστογράμματος σε συγκεκριμένες περιοχές. Η επιλογή πολλαπλών κατωφλίων μπορεί να θεωρηθεί ως ένα πρόβλημα εύρεσης τιμών $T(k)$, $k = 1, 2, \dots, J-1$ με σκοπό το μετασχηματισμό της αρχικής εικόνας με L αποχρώσεις του γκρι σε μια νέα με J αποχρώσεις. Ειδικότερα για $T(1) < T(2) < \dots < T(J-1)$, τότε η νέα εικόνα ορίζεται ως

$$I(x,y) = \begin{cases} G(0) \text{ αν } f(x,y) \leq T(1) \\ G(1) \text{ αν } T(1) < f(x,y) \leq T(2) \\ \vdots \\ G(J-1) \text{ αν } T(J-1) < f(x,y) \end{cases} \quad (3.2.3.1)$$

όπου G μία συνάρτηση τιμών των επιπέδων του γκρι. Μία τεχνική πολλαπλών κατωφλίων που βασίζεται στο ιστογράμμο πρέπει να πετυχαίνει έναν ικανοποιητικό διαχωρισμό των διαφόρων περιοχών του ιστογράμματος. Οι μέθοδοι της διασποράς και της εντροπίας είναι δυνατόν να εφαρμοσθούν και στην επίλυση του προβλήματος της πολυκατωφλίωσης, έχουν όμως μεγάλο υπολογιστικό κόστος.

Μία αποτελεσματική προσέγγιση είναι να θεωρήσουμε την εύρεση πολλαπλών κατωφλίων ως πρόβλημα εύρεσης κλάσεων (clustering problem). Ένα κατάλληλο νευρωνικό δίκτυο μπορεί να λύσει αποτελεσματικά προβλήματα εύρεσης κλάσεων. Είναι γνωστό ότι ο κύριος σκοπός ενός νευρωνικού δικτύου *Kohonen* για την δημιουργία ενός αυτό-οργανούμενου πίνακα απεικόνισης χαρακτηριστικών (SOFM: Self Organized Feature Map) είναι η αντιπροσώπευση ενός πολυπληθούς συνόλου διανυσμάτων εισόδου με ένα ολιγομελέστερο σύνολο “πρωτοτύπων ανυσμάτων” ώστε να επιτευχθεί μια καλή προσέγγιση του αρχικού χώρου εισόδου που να ικανοποιεί τα κύρια στατιστικά χαρακτηριστικά του. Θεωρούμε ένα δίκτυο με νευρώνες (κόμβους) όπως στο Σχ.3.2.3.1



Σχήμα 3.2.3.1

- Συγκροτείται το σύνολο εκπαίδευσης του ΑΠΑΧ από τις τιμές των αποχρώσεων των εικονοστοιχείων της εικόνας.
- Θεωρούμε τη μεταβλητή επανάληψης t που παίρνει ακέραιες τιμές από 0 μέχρι μία προκαθορισμένη τελική τιμή T (π.χ. $T=100.000$).
- Θεωρούμε τη μεταβλητή του ρυθμού εκμάθησης $\alpha(t)$.
- Θεωρούμε το μήκος $d(t)$ που ορίζει μια υποπεριοχή (γειτονιά) στην διάταξη των νευρώνων.
- Έστω w_j το βάρος της σύναψης μεταξύ του j νευρώνα του επιπέδου ανταγωνισμού και της εισόδου του.

Εκτελούνται τα παρακάτω βήματα:

Βήμα 1. Αρχικοποιούνται ($t=0$) τα βάρη των συνάψεων $w_j(0)$ με τυχαίες τιμές από 255.

Βήμα 2. Αρχικοποιείται η $\alpha(0)$ με μια μεγάλη τιμή, συνήθως μεταξύ 0.2 και 0.5.

Βήμα 3. Αρχικοποιείται η $d(0)$ με την τιμή $J/2$, που είναι ίση με το μισό του εύρους του κανάβου.

Βήμα 4. Επιλέγεται τιμή απόχρωσης $I(x,y)$ ενός τυχαίου εικονοστοιχείου από το σύνολο εκπαίδευσης.

Βήμα 5. Υπολογίζεται η έξοδος $o_j(t)$ κάθε νευρώνα από τη σχέση

$$o_j(t) = \|I(x,y) - w_j(t)\|. \quad (3.2.3.2)$$

Βήμα 6. Ο νευρώνας c ανακηρύσσεται νικητής εάν ικανοποιείται η συνθήκη

$$o_c(t) = \min\{o_x(t)\}. \quad (3.2.3.3)$$

Εάν οι έξοδοι δύο νευρώνων είναι ίσες, τότε κατά σύμβαση επιλέγεται αυτός με το μικρότερο δείκτη.

Βήμα 7. Τα βάρη w_j των συνάψεων ανανεώνονται σύμφωνα με τις παρακάτω σχέσεις

$$\Delta w_j = \begin{cases} \alpha(t) \cdot (I(x,y) - w_j(t)) & \text{αν } j \in N_c \\ 0 & \text{αν } j \notin N_c \end{cases} \quad (3.2.3.4)$$

$$w_j(t+1) = w_j(t) + \Delta w_j(t) \quad (3.2.3.5)$$

όπου N_c το σύνολο των δεικτών των νευρώνων που βρίσκονται μέσα στην γειτονιά του νικητή νευρώνα και πλευρά $d(t)$.

Βήμα 8. Αυξάνεται η μεταβλητή επανάληψης κατά ένα και αποδίδονται νέες τιμές στις μεταβλητές $\alpha(t)$, $d(t)$ σύμφωνα με τις σχέσεις:

$$\alpha(t) = \alpha(0) \cdot \left(1 - \frac{t}{T}\right) \quad (3.2.3.6)$$

$$d(t) = d(0) \cdot \left(1 - \frac{t}{T}\right) \quad (3.2.3.7)$$

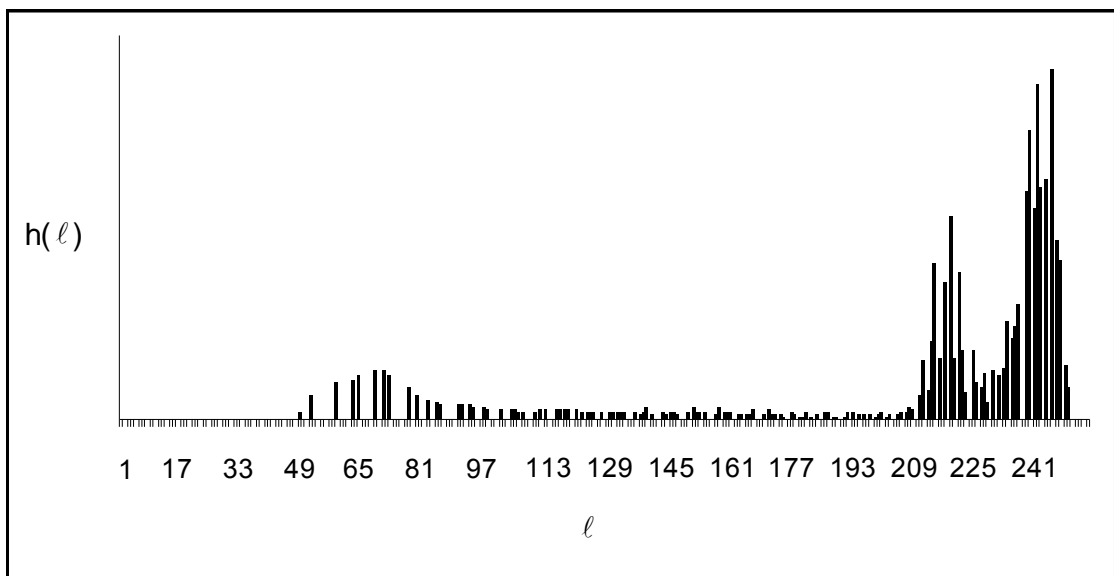
Τα βήματα 4 έως 8 επαναλαμβάνονται έως ότου η μεταβλητή t πάρει τη μέγιστη τελική τιμή T . Είναι φανερό πως οι μεταβλητές $\alpha(t)$ και $d(t)$ συγκλίνουν στο μηδέν καθώς η t τείνει στην τιμή T .

Μετά την εκπαίδευση κάθε άνυσμα εισόδου του ΑΠΑΧ αποδίδεται στον νικητή νευρώνα. Κάθε νευρώνας του επιπέδου εξόδου αντιπροσωπεύει μία ομάδα προτύπων (cluster). Πρότυπα με μεγάλη ομοιότητα αντιπροσωπεύονται από τον ίδιο νευρώνα. Ο χάρτης χαρακτηριστικών του Kohonen οργανώνει τους νευρώνες του επιπέδου ανταγωνισμού με τέτοιο τρόπο ώστε οι ομοιότητες μεταξύ των προτύπων να απεικονίζονται με σχέσεις γειτνίασης επάνω στον κανάβο του επιπέδου ανταγωνισμού

Στο Σχ. 3.2.3.1 φαίνεται η αρχική εικόνα στην οποία θα εφαρμοσθεί πολυκατωφλίωση και στο Σχ. 3.2.3.2.



Σχήμα 3.2.3.1.



Σχήμα 3.2.3.2. Το ιστόγραμμα της εικόνας του Σχ. 3.2.3.1. Μετά την εκπαίδευση ενός ΝΔ *Kohonen* για την πολυκατωφλίωση των νευρώνων προκύπτουν οι τιμές των βαρών που αναγράφονται στον Πιν. 3.2.3.1 και οι τιμές των κατωφλίων του Πιν.3.2.3.2. Οι τιμές αυτές αποτελούν ουσιαστικά ένα μικρότερο πλήθος αποχρώσεων για την απόδοση της αρχικής εικόνας.

W_0	w_1	W_2	w_3	W_4	W_5	W_6	w_7
69	111	153	200	218	231	240	246

Πίνακας 3.2.3.1: Οι τιμές των συντελεστών του ΝΔ μετά την εκπαίδευση.

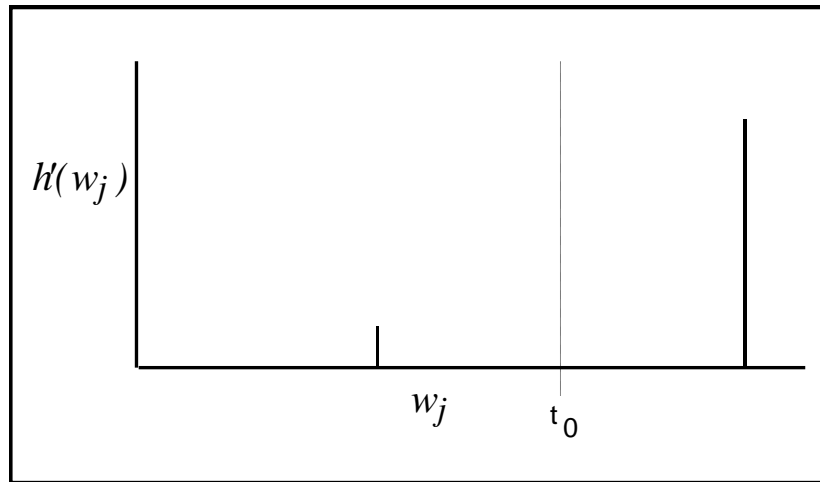
T_0	T_1	T_2	T_3	T_4	T_5	T_6
90	132	176	209	224	235	243

Πίνακας 3.2.3.2. Οι τιμές των κατωφλίων.

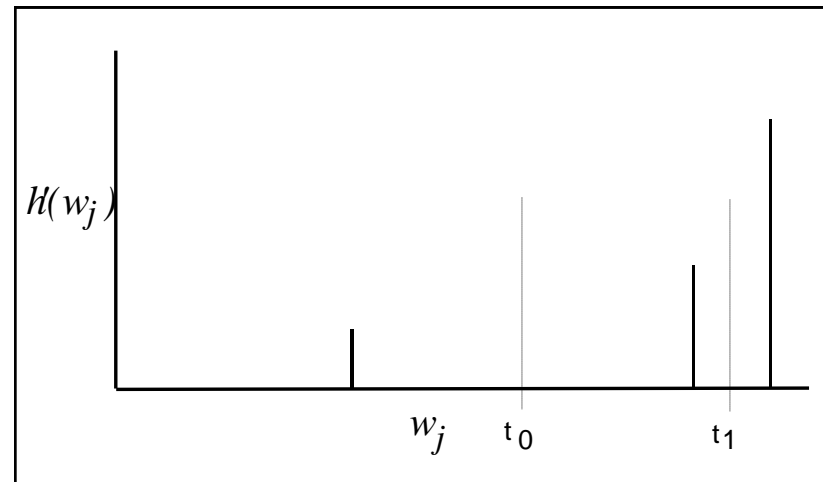
Στον Πιν. 3.2.3.3 φαίνονται οι τιμές των κατωφλίων για διαφορετικό πλήθος (J) νευρώνων του νευρωνικού δικτύου

J	T_0	T_1	T_2	T_3	T_4
2	164				
3	151	227			
4	106	179	229		
5	106	178	224	237	
6	97	143	191	225	239

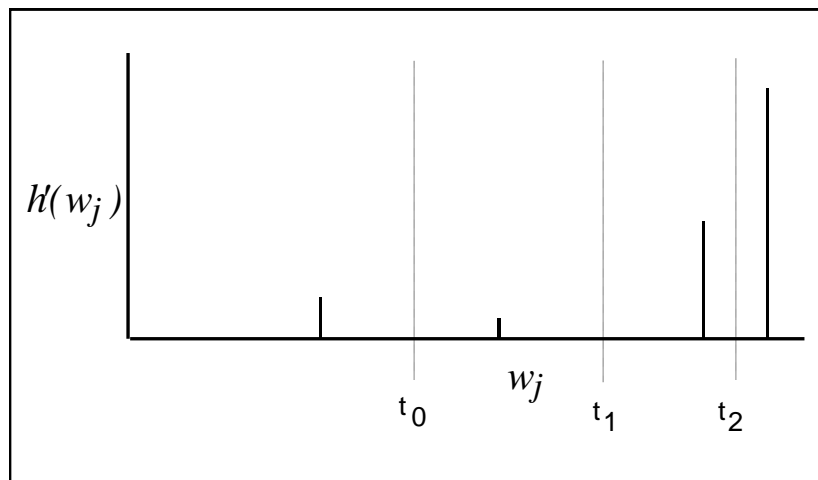
Πίνακας 3.2.3.3. Οι τιμές των κατωφλίων για $J=2, \dots, 6$.



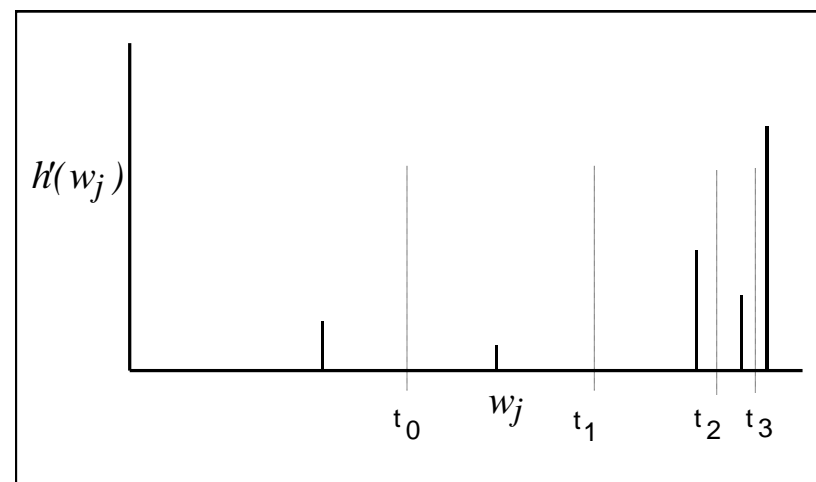
$J=2$



$J=3$



$J=4$

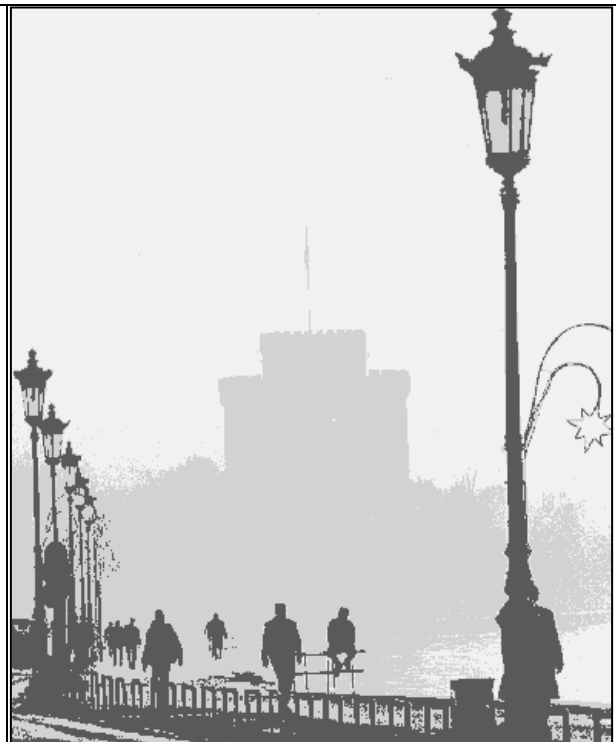


$J=5$

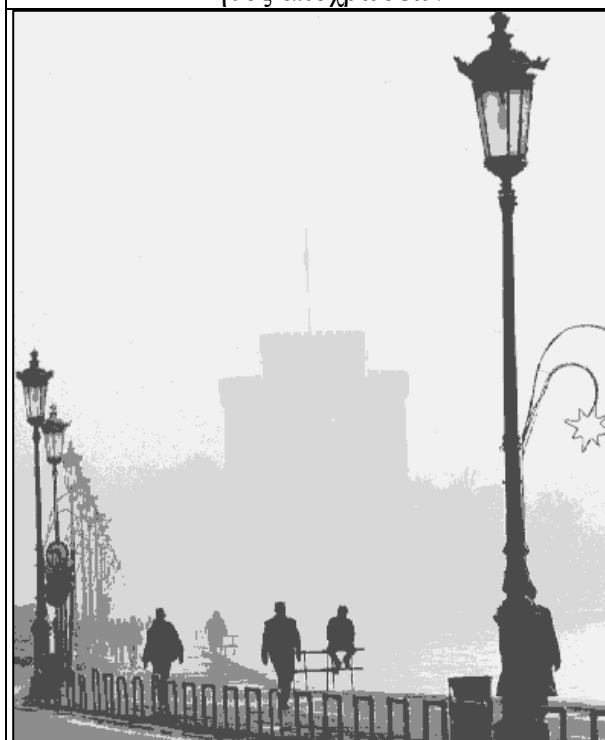
Σχήμα 3.2.3.3. Τα ιστογράμματα και οι αντίστοιχες τιμές καταφλίων για διαφορετικές τιμές του J .



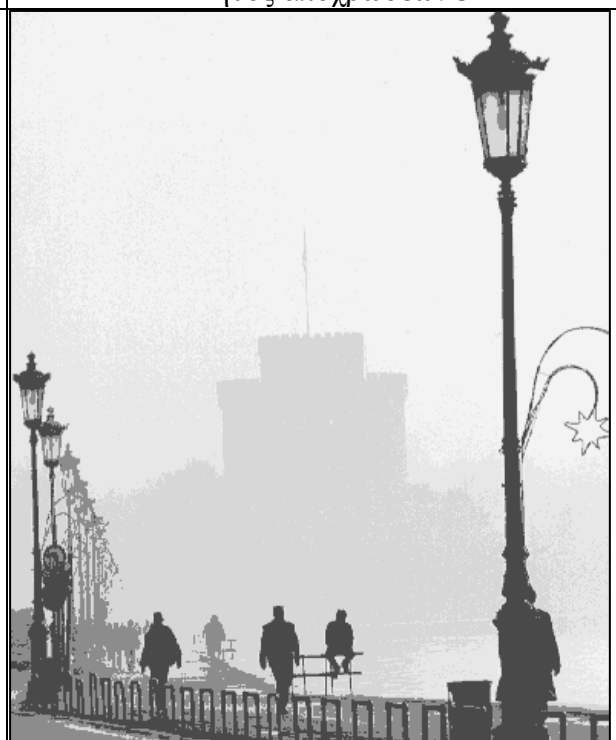
Πλήθος αποχρώσεων 2



Πλήθος αποχρώσεων 3



Πλήθος αποχρώσεων 3



Πλήθος αποχρώσεων 4

Σχήμα 3.2.3.4: Οι εικόνες που προέκυψαν από την πολυκατωφλίωση της εικόνας του Σχ.3.2.3.1 για διαφορετικό πλήθος κατωφλίων και αποχρώσεων.

Η βέλτιστη τιμή του J για ένα δεδομένο σύνολο τιμών $\{2, \dots, J_{max}\}$ ελαχιστοποιεί το σφάλμα προσαρμογής μεταξύ των συναρτήσεων $H(\ell)$ και $H_J(\ell)$. Η συνάρτηση $H(\ell)$ υπολογίζεται από τη σχέση

$$H(\ell) = \frac{h(\ell)}{\sum_{\ell=0}^{L-1} h(\ell)} \quad (3.2.3.8)$$

και είναι συνάρτηση πυκνότητας πιθανότητας επειδή $\sum_{\ell=0}^{L-1} H(\ell) = 1$.

Η πιθανότητα f_j εμφάνισης της κλάσης C_j , η μέση τιμή μ_j της κλάσης και η διασπορά της σ_j^2 δίνονται αντίστοιχα από τις σχέσεις

$$f_j = P_r(C_j) = \sum_{C_j} h(\ell) \quad (3.2.3.9)$$

$$\mu_j = \frac{\sum_{C_j} \ell h(\ell)}{\sum_{C_j} h(\ell)} \quad (3.2.3.10)$$

$$\sigma_j^2 = \frac{\sum_{C_j} (\ell - \mu_j)^2 h(\ell)}{\sum_{C_j} h(\ell)} \quad (3.2.3.11)$$

Η συνάρτηση προσαρμογής $H_J(\ell)$ είναι το άθροισμα

$$H_J(\ell) = \sum_{j=0}^{J-1} f_j H_j(\ell) \quad (3.2.3.12)$$

όπου

$$h_j(\ell) = \frac{1}{\sigma_j \sqrt{2\pi}} \exp\left(-\frac{(\ell - \mu_j)^2}{2\sigma_j^2}\right) \quad (3.2.3.13)$$

Μπορεί να δειχθεί εύκολα ότι η $h_j(\ell)$ είναι επίσης συνάρτηση πυκνότητας πιθανότητας. Το σφάλμα προσαρμογής $|e_J|$ για ένα πλήθος κλάσεων J ορίζεται ως

$$|e_J| = \sum_{\ell=0}^{L-1} [h(\ell) - H_J(\ell)]^2 \quad (3.2.3.14)$$

και υπολογίζεται για $J=2, \dots, J_{max}$. Η τιμή J_o είναι βέλτιστη όταν το αντίστοιχο σφάλμα προσαρμογής $|e_{J_o}|$ ικανοποιεί τη συνθήκη

$$|e_{J_o}| = \text{minimum}\{|e_J|\} \quad (3.2.3.15)$$

Ο Πίν. 3.2.3.3 δείχνει τις τιμές του $|e_J|$ για $J=2, \dots, 6$. Η μικρότερη από αυτές τις τιμές προκύπτει για $J=3$ και θεωρείται ότι είναι η βέλτιστη τιμή J_o . Όπως μπορεί να παρατηρηθεί από το ιστόγραμμα $h(\ell)$ (Σχ. 3.2.3.3), αυτή η τιμή είναι κατάλληλη επειδή αντιστοιχεί στο πλήθος των σημαντικών υψωμάτων του ιστογράμματος.

3.3 Κωδικοποίηση εικόνων αποχρώσεων του γκρι

Για την αναπαράσταση μια ψηφιακής εικόνας J γραμμών και K στηλών με βάθος χρώματος (bit) B bits απαιτούνται $J \times K \times B$ bits. Η κωδικοποίηση μιας συγκεκριμένης ψηφιακής εικόνας με τέτοιο τρόπο δεν είναι η πλέον κατάλληλη όταν ενδιαφέρει το πλήθος των bits που δαπανώνται. Αυτό συχνά συμβαίνει σε περιπτώσεις μετάδοσης ή αποθήκευσης της εικόνας. Με τον όρο *συμπίεση της εικόνας* αναφερόμαστε σε ένα πλήθος τεχνικών που στοχεύουν στην μείωση του πλήθους των bits που απαιτούνται για την κωδικοποίηση της εικόνας. Οι τεχνικές αυτές διακρίνονται σε αυτές που επιτρέπουν απώλεια πληροφορίας και αυτές που δεν επιτρέπουν. Στην πρώτη περίπτωση η πληροφορία που χάνεται δεν θα πρέπει να είναι ζωτικής σημασίας για τον παρατηρητή της εικόνας. Ακολούθως θα περιγραφούν αλγόριθμοι συμπίεσης ψηφιακών εικόνων αποχρώσεων του γκρι.

3.3.1 Κωδικοποίηση Huffman

Σύμφωνα με την τεχνική αυτή κωδικοποιούμε τις αποχρώσεις του γκρι μιας συγκεκριμένης εικόνας με κωδικές λέξεις μεταβλητού μήκους. Μια απόχρωση που εμφανίζεται συχνά (με μεγάλη πιθανότητα) στην εικόνα κωδικοποιείται με λέξη μήκους μικρότερου από άλλες αποχρώσεις που εμφανίζονται σπανιότερα. Το κανονικοποιημένο ιστόγραμμα $h(g)$ των αποχρώσεων της ψηφιακής εικόνας παρέχει τις τιμές της πιθανότητας εμφάνισης $p(g)$ των τιμών της μεταβλητής g των αποχρώσεων της εικόνας διότι $p(g)=h(g)$. Αν $l(g)$ είναι το μήκος της κωδικής λέξης της g απόχρωσης και G το πλήθος των αποχρώσεων, τότε το μέσο μήκος του κώδικα που απαιτείται για την εικόνα είναι

$$\bar{l} = \sum_{g=0}^{G-1} p(g)l(g) = \sum_{g=0}^{G-1} h(g)l(g) \quad (3.3.1.1)$$

Ένας κατάλληλος ευκρινής, μονοσήμαντος και στιγμιαία αποκωδικοποιήσιμος κώδικας που ελαχιστοποιεί το μέσο μήκος είναι ο κώδικας *Huffman*. Αν η εικόνα κωδικοποιηθεί σύμφωνα με αυτόν η τιμή του \bar{l} σύμφωνα με την θεωρία της πληροφορίας θα ικανοποιεί τη σχέση

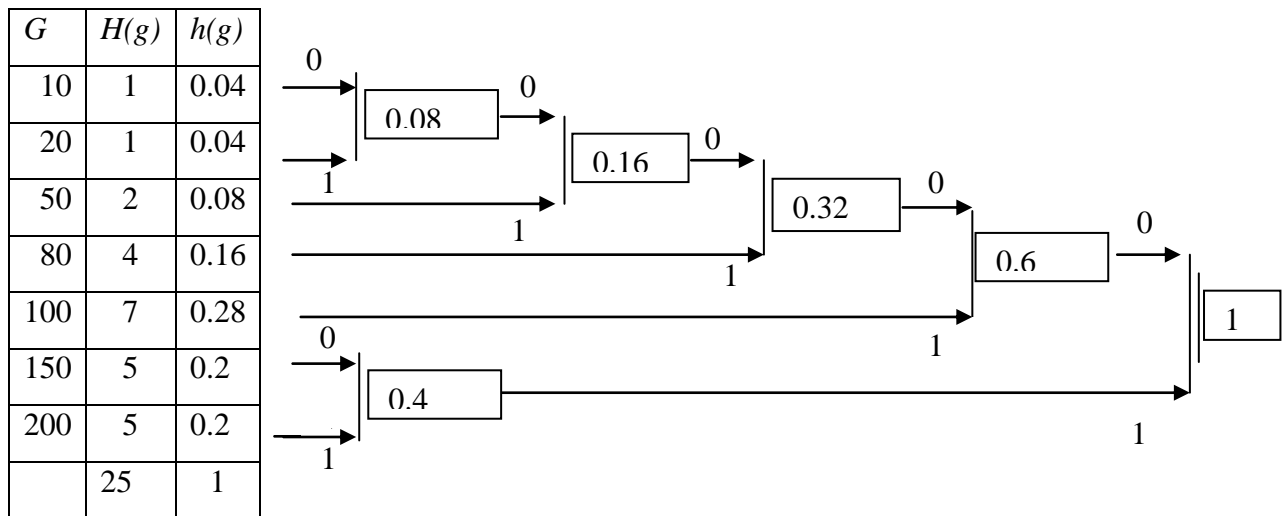
$$H(G) \leq \bar{l} \leq H(G)+1 \quad (3.3.1.2)$$

Όπου $H(G)$ η εντροπία της εικόνας αν την θεωρήσουμε ως πηγή πληροφορίας που δίνεται από την σχέση

$$H(G) = -\sum_{g=0}^{G-1} p(g) \log_2 p(g) = -\sum_{g=0}^{G-1} h(g) \log_2 h(g) \quad (3.3.1.3)$$

Ο εύρεση του κώδικα *Huffman* για μια συγκεκριμένη εικόνα γίνεται με την δημιουργία ενός δυαδικού δένδρου. Τα φύλλα του δένδρου είναι οι τιμές των αποχρώσεων με τις αντίστοιχες πιθανότητες εμφάνισης τους στην εικόνα. Τα φύλλα συνενώνονται σε δυαδικούς κόμβους. Σε κάθε κόμβο αντιστοιχίζεται το άθροισμα των πιθανοτήτων εμφάνισης των παιδιών του. Ο πρώτος κόμβος δημιουργείται από τη συνένωση των δύο αποχρώσεων (φύλλα) με τις μικρότερες πιθανότητες. Κάθε επόμενος κόμβος έχει δύο παιδιά που επιλέγονται από φύλλα ή διαθέσιμους κόμβους που δεν ενώθηκαν και έχουν τις μικρότερες πιθανότητες εμφάνισης.. Στο ακόλουθο παράδειγμα παρουσιάζεται εφαρμογή της κωδικοποίησης σε μια εικόνα αποχρώσεων του γκρι διαστάσεων 4×5 .

100	100	100	100	100
200	10	20	150	100
200	50	50	150	100
200	200	200	150	150
80	80	80	80	80



G	10	20	50	80	100	150	200
Κωδικός	00000	00001	0001	001	01	10	11

Για την εικόνα του παραδείγματος απαιτούνται $1 \times 5 + 1 \times 5 + 2 \times 4 + 4 \times 3 + 7 \times 2 + 5 \times 2 + 5 \times 2 = 54$ bits αντί των $25 \times 3 = 75$ bits που απαιτούνται για κωδικές λέξεις σταθερού μήκους τριών bits ($3 = \lceil \log_2 7 \rceil + 1$). Το μέσος μήκος κωδικολέξης $\bar{l} = 54 / 25 = 2.16$ και το ποσοστό συμπίεσης είναι $(75 - 54) / 75\% = 28\%$

3.3.2 Κωδικοποίηση LZW

Με την τεχνική LZW (Lebel – Ziv – Welch) επιδιώκεται η κωδικοποίηση ακολουθιών από διαδοχικές τιμές εικονοστοιχείων. Προς τούτο για συγκεκριμένη εικόνα δημιουργείται ένα ευρετήριο των ακολουθιών και των κωδικών τους. Για παράδειγμα αν η ακολουθία τιμών 120-30-30 ($3 \times 8 = 24$ bits) κωδικοποιηθεί με τον κωδικό 410 (9 bits) και επαναλαμβάνεται 30 φορές στην εικόνα υπάρχει ένα όφελος $30 \times 15 = 450$ bits. Συνήθως στο ευρετήριο έχουν τοποθετηθεί ήδη οι 256 τιμές των αποχρώσεων του γκρι στις θέσεις-κωδικούς από 0-255. Η εικόνα προσπελαίνεται από πάνω προς τα κάτω και από αριστερά προς τα δεξιά. Εάν μία ακολουθία τιμών διαδοχικών εικονοστοιχείων υπάρχει στο ευρετήριο αυτή αυξάνεται με την τιμή του επόμενου εικονοστοιχείου και η νέα ακολουθία αναζητάται στο ευρετήριο. Αν μία ακολουθία με N τιμές δεν βρεθεί στο ευρετήριο, τότε αυτή καταχωρείται α' αυτό και στην έξοδο του κωδικοποιητή τίθεται ο κωδικός της ακολουθίας των πρώτων $N-1$ τιμών. Στο παρακάτω παράδειγμα φαίνεται η παραπάνω διαδικασία για μια μικρή εικόνα 4×4 .

10	10	30	30
10	10	30	30
10	10	30	30
10	10	30	30

Ακολουθία τιμών				Τιμή επόμενου εικ/χειού	Έξοδος κωδικο- ποιητή	Ευρετήριο				
						Κωδικός	Ακολουθία			
10				10	10	256:	10	10		
10				30	10	257:	10	30		
30				30	30	258:	30	30		
30				10	30	259:	30	10		
10				10						
10	10			30	256	260:	10	10	30	
30				30						
30	30			10	258	261:	30	30	10	
10				10						
10	10			30						
10	10	30		30	260	262:	10	10	30	30
30				10						
30	10			10	259	263:	30	10	10	
10				30						
10	30			30	257	264:	10	30	30	
30					30					

Είναι σημαντικό να αναφέρουμε ότι κατά την αποκωδικοποίηση των τιμών του συμπίεσμένου αρχείου δεν είναι απαραίτητη η γνώση του ευρετηρίου και ως εκ τούτου δεν είναι αναγκαία η επισύναψή του. Το ευρετήριο δημιουργείται στον αποκωδικοποιητή από τις τιμές που αποκωδικοποιούνται με τον ίδιο τρόπο που δημιουργήθηκε στην φάση της κωδικοποίησης. Ακολούθως παρουσιάζεται η διαδικασία αποκωδικοποίησης των τιμών που δημιουργήθηκαν κατά την φάση κωδικοποίησης του προηγούμενου παραδείγματος.

Έισοδος αποκωδικο- ποιητή	Έξοδος κωδικο-ποιητή			Ευρετήριο			
				Κωδικός		Ακολουθία	
10	10						
10	10						
30	30						
30	30						
256				256:	10	10	
				257:	10	30	
				258:	30	30	
	10	10					
258	30	30					
260				259:	30	10	
				260:	10	10	30
	10	10	30				
259	30	10					
257	10	30					
30	30						

3.3.3 Κωδικοποίηση 2Δ-διακριτού μετασχηματισμό συνημιτόνου

Από όσα αναφέρθηκαν προηγουμένως η διαδικασία της κωδικοποίησης σχετιζόταν άμεσα με τις τιμές των εικονοστοιχείων της εικόνας. Στην ενότητα αυτή θα δούμε τον μετασχηματισμό του διδιάστατου ψηφιακού σήματος με την εύρεση των προβολών του (συντελεστών) σε μία ορθοκανονική βάση, πριν ακολουθήσει η διαδικασία κωδικοποίησης

Ο διδιάστατος μετασχηματισμός συνημιτόνου (2D-DCT: Discrete Cosine Transform) αποτελεί την βάση για συμπίεση μη δυαδικών εικόνων με αποδεκτή απώλεια πληροφορίας. Η τυποποίηση JPEG βασίζεται στις πρώτες εκδόσεις της στον 2D-DCT. Θα αρχίσουμε την παρουσίαση του μετασχηματισμού από την μονοδιάστατη εκδοχή του.

Οι σχέσεις

$$g_0(n) = \frac{1}{\sqrt{N}}, \quad g_k(n) = \sqrt{\frac{2}{N}} \cos\left(\frac{\pi}{2N} k(2n+1)\right) \quad (3.3.3.1)$$

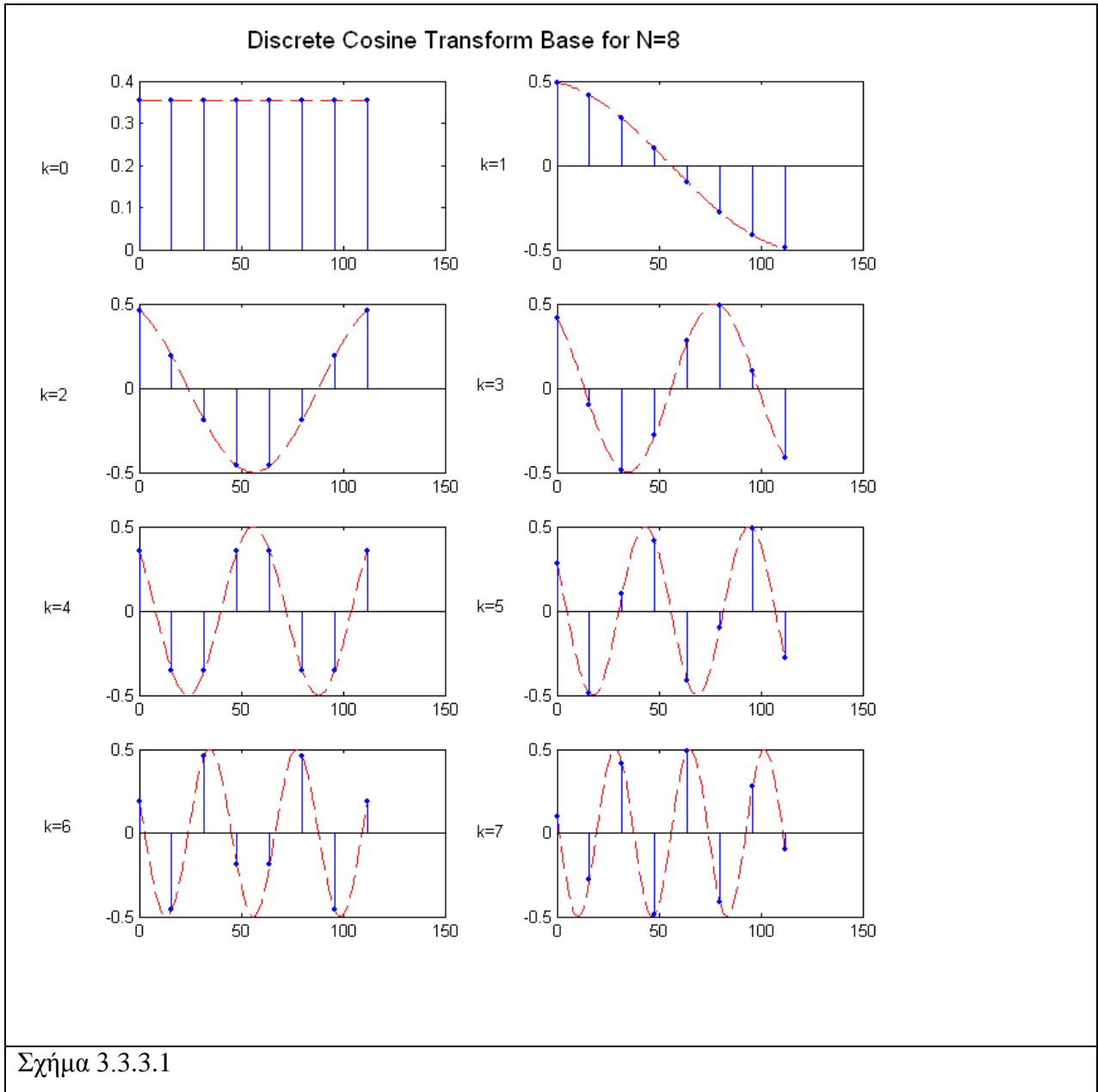
για $k=1..N-1$ και $n=0,..,N-1$ είναι ένας πυρήνας (kernel) δημιουργίας ορθοκανονικών διανυσμάτων για κάθε τιμή του N . Οι συντελεστές c_k για ένα διακριτό σήμα $x(n)$ πεπερασμένου μήκους N (δηλαδή οι προβολές του \mathbf{x} στα \mathbf{g}_k) προκύπτουν από τη σχέση:

$$c_0 = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x[n], \quad c_k = \sqrt{\frac{2}{N}} \sum_{n=0}^{N-1} \cos\left(\frac{\pi}{2N} k \cdot (2n+1)\right) \cdot x[n] \quad (3.3.3.1)$$

Η σχέση αυτή αποτελεί τον μονοδιάστατο διακριτό μετασχηματισμό συνημιτόνου (1D-DCT: Discrete Cosine Transform). Οι τιμές $x(n)$ (δηλαδή το \mathbf{x}) προκύπτουν από τη σχέση

$$x(n) = \frac{c_0}{\sqrt{N}} + \sqrt{\frac{2}{N}} \sum_{k=1}^{N-1} c_k \cdot \cos\left(\frac{\pi}{2N} k(2n+1)\right) \quad (3.3.3.2)$$

Η σχέση αυτή αποτελεί τον αντίστροφο διακριτό μετασχηματισμό του συνημιτόνου. Στο Σχ. 3.3.3.1 δείχνονται τα ανύσματα βάσης του DCT για $N=8$.



Παράδειγμα 3.3.3.1

Για N=3 τα ανύσματα

$$\mathbf{g}_0 = \left[\frac{1}{\sqrt{3}} \quad \frac{1}{\sqrt{3}} \quad \frac{1}{\sqrt{3}} \right], \mathbf{g}_1 = \left[\frac{1}{\sqrt{2}} \quad 0 \quad -\frac{1}{\sqrt{2}} \right], \mathbf{g}_2 = \left[\frac{1}{\sqrt{6}} \quad -\frac{\sqrt{2}}{\sqrt{3}} \quad \frac{1}{\sqrt{6}} \right]$$

προήλθαν από τις σχέσεις

$$g_0(n) = \frac{1}{\sqrt{N}}, \quad g_k(n) = \sqrt{\frac{2}{N}} \cos\left(\frac{\pi}{2N} k(2n+1)\right) \text{ για } k=1..N-1 \text{ και } n=0,..,N-1$$

Και αποτελούν μία ορθοκανονική βάση και το σήμα $\mathbf{x}=[1, 3, 2]^T$ προβάλλεται σ' αυτά με προβολές που δίνουν οι συντελεστές $c_0 = 2\sqrt{3}$, $c_1 = -\sqrt{2}/2$, $c_2 = -\sqrt{6}/2$. Για $n=0,1,2$

$$x[n] = 2\sqrt{3}g_0[n] - \sqrt{2}/2g_1[n] - \sqrt{6}/2g_2[n]$$

Παράδειγμα 3.3.3.2

Για $\mathbf{x}=[1, 3, 2, -1]^T$ η τιμή $N=4$ και για $n=0,1,2,3$ τα ανύσματα βάσης είναι:

$$g_0[n] = \frac{1}{2} \quad \text{H} \quad \mathbf{g}_0=[0.50, 0.50, 0.50, 0.50]^T$$

$$g_1[n] = \sqrt{\frac{2}{N}} \cos\left(\frac{\pi}{2N}(2n+1)\right) \quad \text{H} \quad \mathbf{g}_1=[0.65, 0.27, -0.27, -0.65]^T$$

$$g_2[n] = \frac{1}{2} \cos\left(\frac{\pi}{2N}2 \cdot (2n+1)\right) \quad \text{H} \quad \mathbf{g}_2=[0.50, -0.50, 0.50, -0.50]^T$$

$$g_3[n] = \frac{1}{2} \cos\left(\frac{\pi}{2N}3 \cdot (2n+1)\right) \quad \text{H} \quad \mathbf{g}_3=[0.27, -0.65, 0.65, -0.27]^T$$

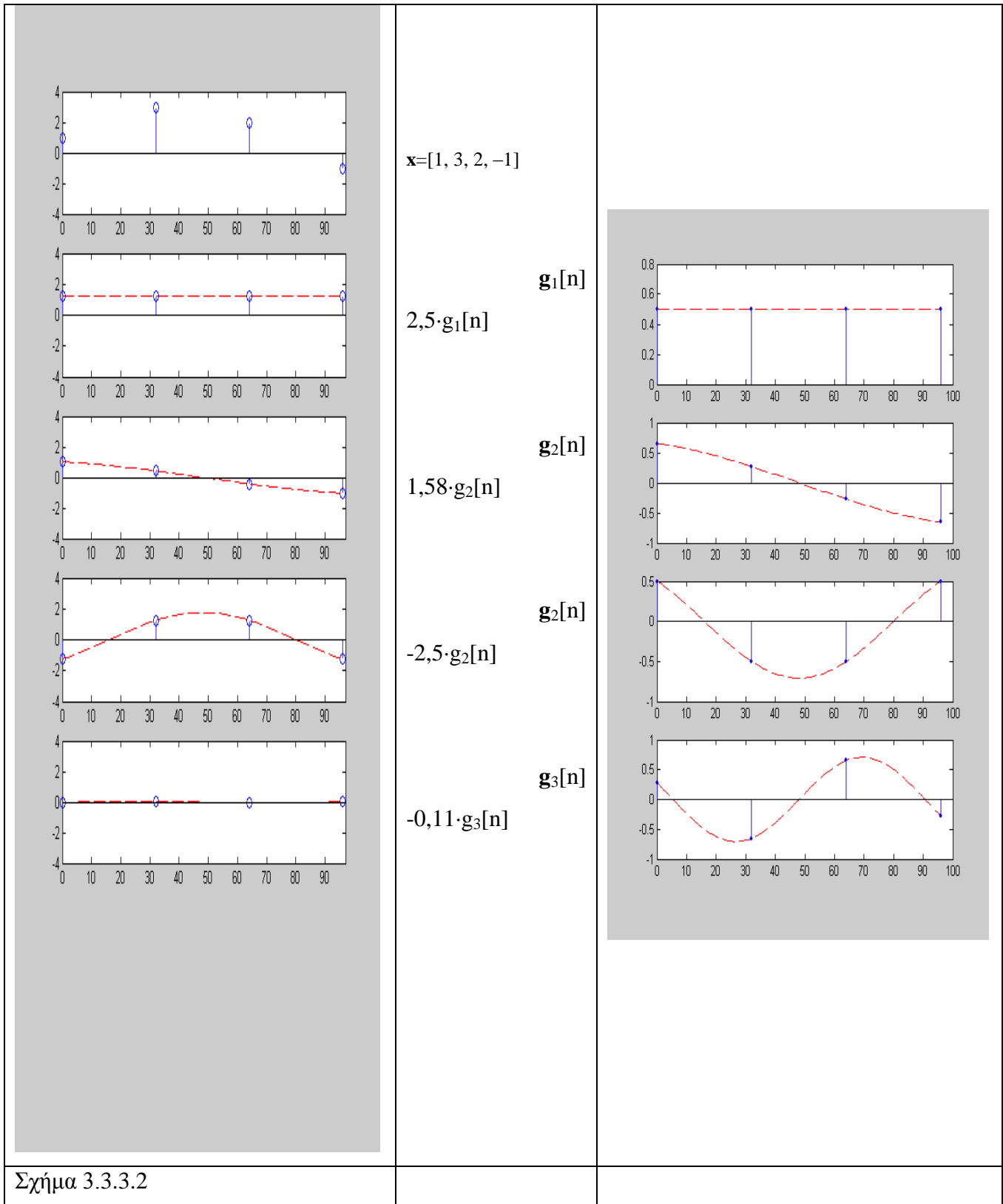
Οι συντελεστές c_k είναι τα στοιχεία του πίνακα

$$\mathbf{C} = \mathbf{G}^T \cdot \mathbf{x} = \begin{bmatrix} 0.50 & 0.50 & 0.50 & 0.50 \\ 0.65 & 0.27 & -0.27 & -0.65 \\ 0.50 & -0.50 & 0.50 & -0.50 \\ 0.27 & -0.65 & 0.65 & -0.27 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \\ 3 \\ -1 \end{bmatrix} = \begin{bmatrix} 2.5 \\ 1.58 \\ -2.5 \\ -0.11 \end{bmatrix}$$

οι συντελεστές c_k προκύπτουν όμοια και από τις (3.3.2.1) Το σήμα ανακτάται από τις σχέσεις (3.3.2.2) ή από τον πίνακα

$$\mathbf{x} = \mathbf{G} \cdot \mathbf{C} = \begin{bmatrix} 0.50 & 0.65 & 0.50 & 0.27 \\ 0.50 & 0.27 & -0.50 & -0.65 \\ 0.50 & -0.27 & 0.50 & 0.65 \\ 0.50 & -0.65 & -0.50 & -0.27 \end{bmatrix} \cdot \begin{bmatrix} 2.5 \\ 1.58 \\ -2.5 \\ -0.11 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 2 \\ -1 \end{bmatrix}$$

Στο Σχ. 3.3.3.2 δείχνονται τα ανύσματα βάσης του DCT για $N=4$, οι συντελεστές c_k του DCT του σήματος $\mathbf{x}=[1, 3, 2, -1]$ και η ανάλυση του στις τέσσερις συνιστώσες $c_k\mathbf{g}_k$. Κάθε τιμή του $x[n]$ είναι $x[n]=2,5g_0[n]+ 1,58g_1[n]-2,5g_2[n]-0,11g_3[n]$.



Για διακριτά σήματα δύο διαστάσεων με μήκη N_1, N_2 τα ανύσματα βάσης για $n_1=0, \dots, N_1-1, n_2=0, \dots, N_2-1, k_1=0, \dots, N_1-1, k_2=0, \dots, N_2-1$ είναι:

$$g_{k_1 k_2}(n_1, n_2) = g_{k_1}(n_1) \cdot g_{k_2}(n_2)$$

ήτοι (3.3.3.3)

$$g_{00}(n_1, n_2) = \frac{1}{\sqrt{N_1 N_2}},$$

$$g_{k_1 0}(n_1, n_2) = \frac{1}{\sqrt{N_2}} \sqrt{\frac{2}{N_1}} \cdot \cos\left(\frac{\pi k_1 (2n_1 + 1)}{2N_1}\right),$$

$$g_{0k_2}(n_1, n_2) = \frac{1}{\sqrt{N_1}} \sqrt{\frac{2}{N_2}} \cdot \cos\left(\frac{\pi k_2 (2n_2 + 1)}{2N_2}\right),$$

$$g_{k_1 k_2}(n_1, n_2) = \sqrt{\frac{2}{N_1}} \cdot \cos\left(\frac{\pi k_1 (2n_1 + 1)}{2N_1}\right) \cdot \sqrt{\frac{2}{N_2}} \cdot \cos\left(\frac{\pi k_2 (2n_2 + 1)}{2N_2}\right)$$

Για ένα δισδιάστατο διακριτό σήμα $x(n_1, n_2)$ με μήκη N_1, N_2 , οι προβολές του στα παραπάνω ορθοκανονικά ανύσματα βάσης είναι:

$$c_{00} = \frac{1}{\sqrt{N_1 N_2}} \sum_{n_2=0}^{N_2-1} \sum_{n_1=0}^{N_1-1} x[n_1, n_2]$$

$$c_{0k_2} = \sqrt{\frac{2}{N_1 N_2}} \sum_{n_2=0}^{N_2-1} \sum_{n_1=0}^{N_1-1} x[n_1, n_2] \cdot \cos\left(\frac{\pi k_2 (2n_2 + 1)}{2N_2}\right)$$

$$c_{k_1 0} = \sqrt{\frac{2}{N_1 N_2}} \sum_{n_2=0}^{N_2-1} \sum_{n_1=0}^{N_1-1} x[n_1, n_2] \cdot \cos\left(\frac{\pi k_1 (2n_1 + 1)}{2N_1}\right)$$

(3.3.3.4)

$$c_{k_1 k_2} = \frac{2}{\sqrt{N_1 N_2}} \sum_{n_2=0}^{N_2-1} \sum_{n_1=0}^{N_1-1} x[n_1, n_2] \cdot \cos\left(\frac{\pi k_1 (2n_1 + 1)}{2N_1}\right) \cdot \cos\left(\frac{\pi k_2 (2n_2 + 1)}{2N_2}\right)$$

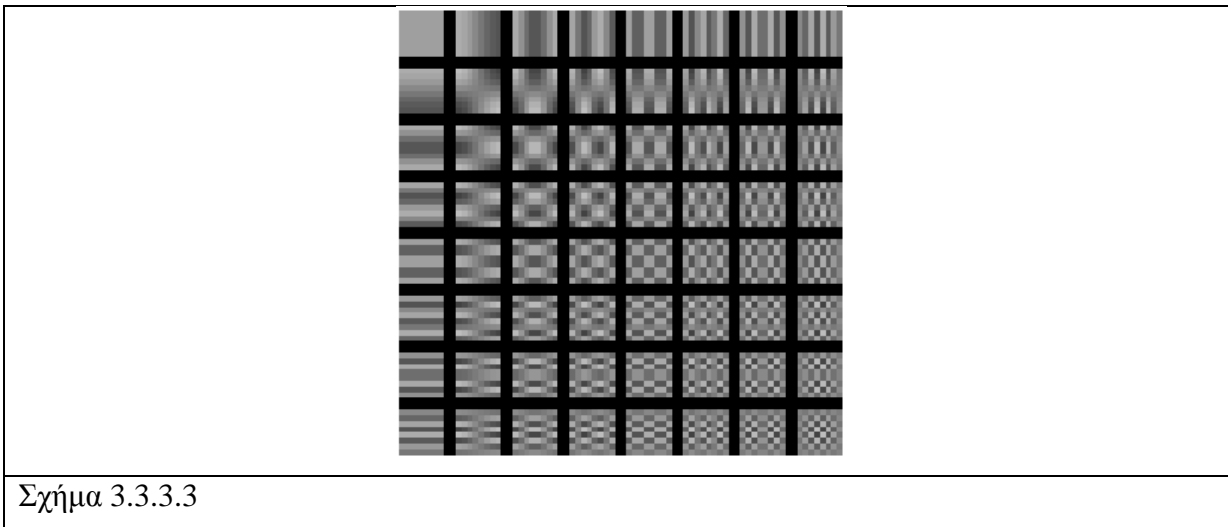
για $k_1, k_2 \neq 0$.

Οι παραπάνω σχέσεις αποτελούν τον δισδιάστατο (2Δ) διακριτό μετασχηματισμό συνημιτόνου (2D-DCT: Discrete Cosine Transform). Το σήμα $x[n_1, n_2]$ ανακτάται σύμφωνα με τη σχέση:

$$x(n_1, n_2) = \sum_{n_2=0}^{N_2-1} \sum_{n_1=0}^{N_1-1} c_{k_1 k_2} \cdot g_{k_1 k_2}(n_1, n_2) \quad (3.3.3.5)$$

που είναι ο αντίστροφος δισδιάστατος (2Δ) διακριτός μετασχηματισμός συνημιτόνου.

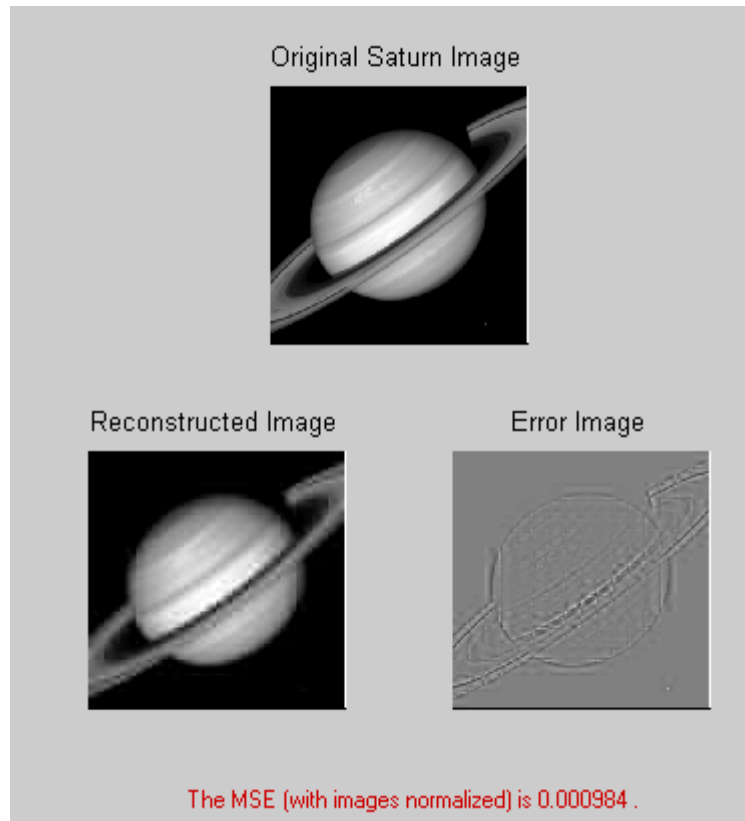
Είναι προφανές από τα παραπάνω αν $M \times N$ είναι τα εικονοστοιχεία της εικόνας τότε $M \times N$ θα είναι και όλοι οι συντελεστές του 2D-DCT αυτής. Για παράδειγμα για μία εικόνα 8×8 , ο μετασχηματισμός οδηγεί στον υπολογισμό 8×8 συντελεστών (προβολών) για κάθε μία από τις 8×8 διακριτές διδιάστατες ακολουθίες βάσης. Κάθε ακολουθία της βάσης έχει τιμές που δίνονται από τις (3.3.3.3) που αν παρασταθούν με τόνους του γκρι προκύπτουν οι εικόνες του Σχ.3.3.3.3.



Μετασχηματισμοί των οποίων η βάση ικανοποιεί την (3.3.3.1) λέγονται διαχωρίσιμοι και οι υπολογισμοί τους μπορούν να αναχθούν σε υπολογισμούς των μονοδιάστατων εκφράσεών τους. Η εφαρμογή του 2D-DCT στην συμπίεση εικόνων αποχρώσεων του γκρι σύμφωνα με την τυποποίηση JPEG (Join Photographic Experts Group) ακολουθεί τα ακόλουθα γενικά βήματα:

- Η εικόνα χωρίζεται σε μη επικαλυπτόμενα τμήματα 8×8 εικονοστοιχείων.
- Οι τιμές των εικονοστοιχείων μετατρέπονται από το διάστημα $[0, 2^L - 1]$ στο διάστημα $[-2^{L-1}, 2^{L-1} - 1]$ με την αφαίρεση από κάθε τιμή του αριθμού 2^{L-1} , όπου L το βάθος χρώματος της εικόνας (bits/pel).
- Για κάθε τμήμα υπολογίζονται οι συντελεστές του 2D-DCT.
- Ακολουθώς οι τιμές κβαντίζονται και τελικά
- Κωδικοποιούνται κατά Huffman.

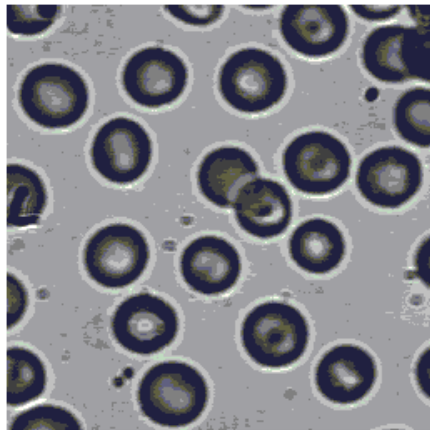
Στο Σχ. 3.3.3.4 η αρχική εικόνα 256×256 pels χωρίζεται σε τμήματα των 8×8 pels και σε κάθε ένα από αυτά εφαρμόζεται ο 2D-DCT, ακολούθως με χρήση μόνο των 8 συντελεστών αναδημιουργείται η εικόνα με σφάλμα όπως φαίνεται στο σχήμα.



Σχήμα 3.3.3.4

3.4 Ανίχνευση ακμών.

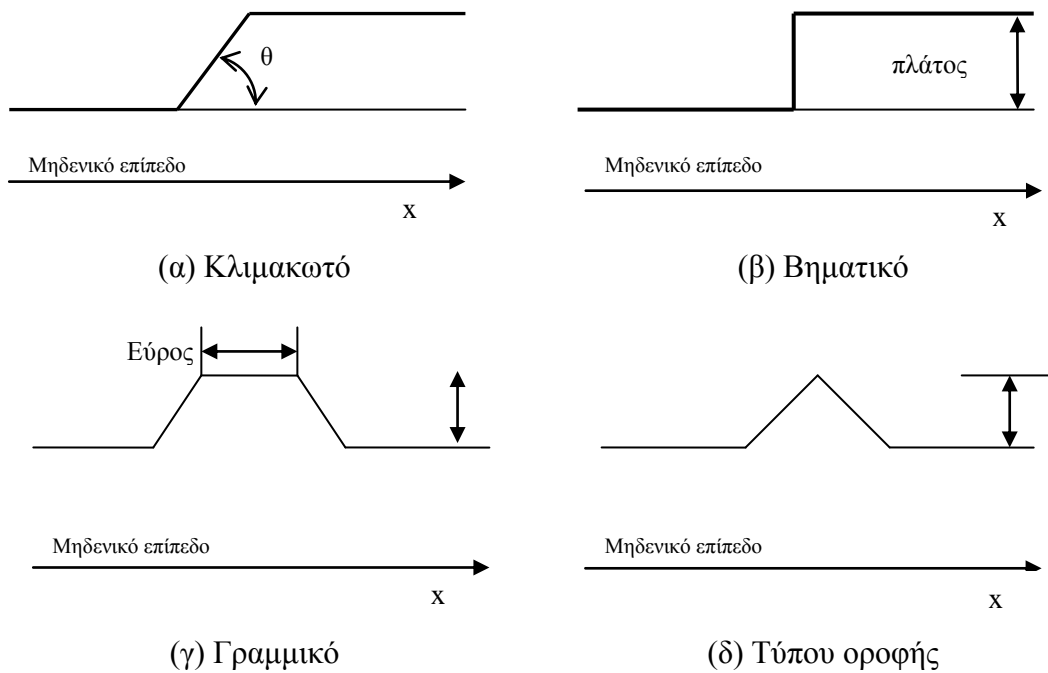
Σε μια εικόνα αποχρώσεων του γκρι υπάρχουν περιοχές εικονοστοιχείων με απότομη αύξηση της φωτεινότητας. Οι περιοχές αυτές βρίσκονται στα όρια των τμημάτων της εικόνας που έχουν σημαντικά διαφορετικές αποχρώσεις. Η ανίχνευση των ορίων αυτών λέγεται *προσδιορισμός των ακμών της εικόνας (edge detection)*. Με άλλα λόγια ακμή είναι μια καμπύλη επί της εικόνας της οποίας οι δυο πλευρές παρουσιάζουν σημαντικές διαφορές της φωτεινότητας ή ενδεχομένως και άλλων χαρακτηριστικών τους όπως η πυκνότητας και η υφή. Η ανίχνευση ακμών είναι μια από τις διαδικασίες ανάλυσης εικόνας που χρησιμοποιούνται ευρύτατα και οι αλγόριθμοι που υπάρχουν στην βιβλιογραφία για την ανίχνευση και βελτίωση τους είναι ίσως οι περισσότεροι από αυτούς άλλων διαδικασιών ανάλυσης εικόνας. Μια ακμή είναι το όριο ενός εικονιζόμενου αντικειμένου και του φόντου ή το όριο μεταξύ επικαλυπτόμενων αντικειμένων (Σχ. 3.4.1).



Σχήμα 3.4.1

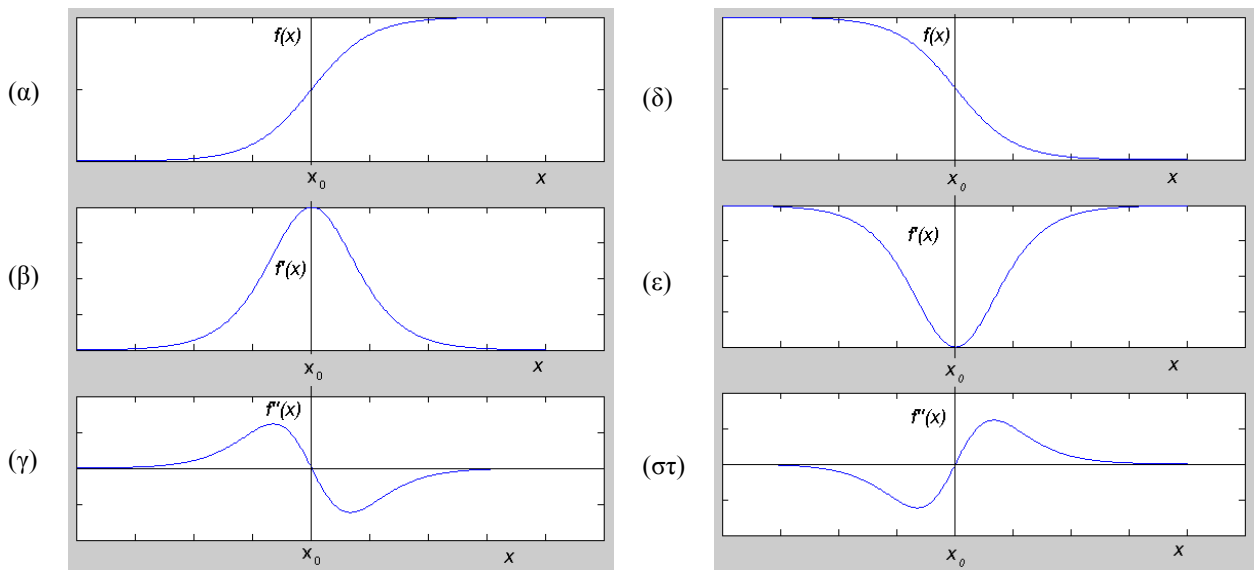
Αυτό σημαίνει ότι αν βρεθούν με ακρίβεια οι ακμές σε μια εικόνα τότε θα εντοπισθούν τα αντικείμενα της εικόνας και είναι δυνατόν να προσδιορισθούν χαρακτηριστικά όπως το εμβαδόν, η περίμετρος ή το σχήμα τους. Η ανίχνευση ακμών είναι εξαιρετικά χρήσιμη εργασία στην ανάλυση των εικόνων διότι μέσω αυτής προσδιορίζονται τα περιγράμματα των αντικειμένων της εικόνας και αποτελεί ένα αρχικό στάδιο τεχνικών τμηματοποίησης της εικόνας. Για τους λόγους αυτούς η εύρεση ακμών είναι βασικό εργαλείο για τον προσδιορισμό και την ταξινόμηση των αντικειμένων που είναι βασικοί στόχοι της μηχανικής όρασης.

Γενικό μαθηματικό μοντέλο ορισμού και ανίχνευσης των ακμών δεν υπάρχει αλλά γίνεται ανάλογα με το πεδίο εφαρμογής, επιλογή από ένα πλήθος μοντέλων και τεχνικών για την ανίχνευση τους. Στο Σχ.3.4.2 φαίνεται η μεταβολή της φωτεινότητας κατά δεδομένη διεύθυνση, ακμών διαφορετικού τύπου ως προς το ύψος, την κλίση και το πλάτος τους.



Σχήμα 3.4.2: Μονοδιάστατα συνεχή μοντέλα ακμών.

Από μαθηματική άποψη αν $f(x,y)$ η συνάρτηση της φωτεινότητας της εικόνας (x, y συνεχείς μεταβλητές - συντεταγμένες σημείου του επιπέδου της εικόνας), η μεταβολή της ως προς μία κατεύθυνση εκφράζεται από την παράγωγο της κατά κατεύθυνση. Ως εκ τούτου η εύρεση των ακμών μπορεί να βασισθεί στην χρήση των παραγώγων της εικόνας. Στο Σχ. 3.4.3 φαίνονται οι γραφικές παραστάσεις της φωτεινότητας, της πρώτης και της δεύτερης παραγώγου της ως προς την x κατεύθυνση.



Σχήμα 3.4.3

Αν το σημείο x_0 είναι σημείο ακμής, η φωτεινότητα παρουσιάζει μέγιστη μεταβολή στο x_0 και συνεπώς η πρώτη παράγωγος έχει ακρότατη τιμή ενώ η δεύτερη παράγωγος μηδενίζεται. Ακολούθως θα δούμε την χρήση των παραγώγων ως βάση για την ανίχνευση των ακμών.

3.4.1 Ανίχνευση ακμών με την χρήση των πρώτων παραγώγων

Αν $f(x,y)$ δισδιάστατη συνάρτηση των συνεχών ανεξάρτητων μεταβλητών x, y , $f(x,y) : \mathbb{R}^2 \rightarrow \mathbb{R}$, οι μερικοί της παράγωγοι $\partial f/\partial x, \partial f/\partial y$ ορίζουν τον πίνακα

$$Df = \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{bmatrix}$$

Αν θεωρήσουμε ορθογώνιους άξονες x', y' με αντίστοιχα μοναδιαία διανύσματα \vec{u}_x και \vec{u}_y , το διάνυσμα κλίσης (*gradient*) της f συμβολίζεται ∇f και δίνεται από τη σχέση

$$\nabla f = \frac{\partial f}{\partial x} \vec{u}_x + \frac{\partial f}{\partial y} \vec{u}_y$$

Αν κινηθούμε κατά την διεύθυνση που ορίζει η γωνία ϕ που δίνεται από σχέση

$$\phi = \text{τοξεφ} \left(\frac{\partial f / \partial x}{\partial f / \partial y} \right)$$

η μεταβολή της $f(x,y)$ είναι μέγιστη και το μέτρο της μεταβολής δίνεται από τη σχέση

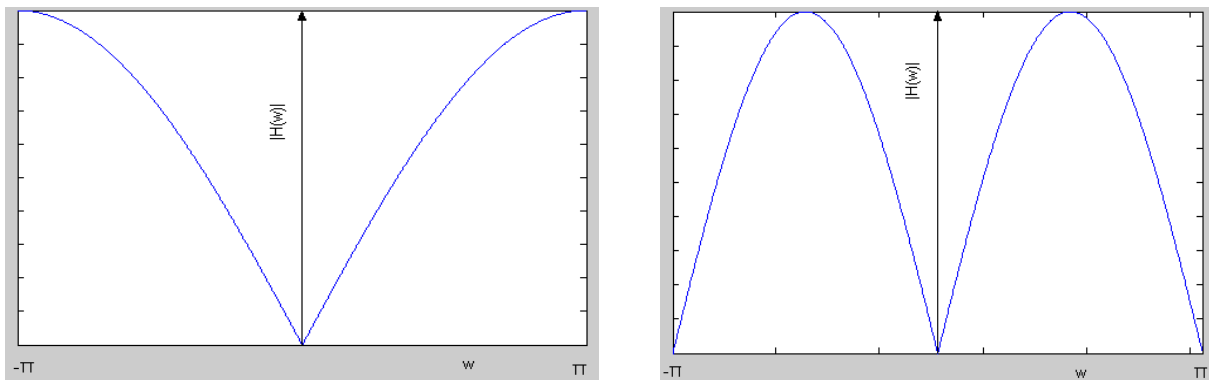
Η κρουστική απόκριση θα είναι τότε

$$h(n) = \delta(n) - \delta(n-1) \text{ ή } h(n) = \delta(n+1) - \delta(n-1)$$

ή με μορφή πίνακα

$$\mathbf{h} = [1, -1] \text{ ή } \mathbf{h} = [1, 0, -1]$$

Η μερική διαφορά $y(n)$ στο σημείο n που είναι σημείο μιας γραμμής ή στήλης της εικόνας θα δίνεται ο από την συνέλιξη $x(n) * h(n)$ δηλαδή την εφαρμογή των αντίστοιχων μασκών. Τα συστήματα αυτά λειτουργούν ως φίλτρα στο πεδίο της συχνότητας. Στο σχήμα 3.4.4 δείχνονται τα διαγράμματα πλάτους των αποκρίσεων συχνότητας των παραπάνω συστημάτων. Πρόκειται για ένα ανωδιαβατό και ένα ζωνοπερατό φίλτρο αντίστοιχα.



Σχήμα 3.4.4

Ο εντοπισμός των σημείων των ακμών μπορεί να βασισθεί στην εύρεση υψηλών τιμών του πίνακα με στοιχεία τις τιμές $|\mathbf{D}(k, j)|$. Υψηλές είναι οι τιμές που είναι μεγαλύτερες από μία καθορισμένη τιμή κατωφλίου T . Συνεπώς στη θέση (k, j) υπάρχει σημείο ακμής εάν $|\mathbf{D}(k, j)| > T$. Μία πρώτη προσέγγιση για την επιλογή μιας καθολικής τιμής κατωφλίου μπορεί να βασισθεί στον μέσο όρο των τιμών $|\mathbf{D}(k, j)|$. Η επιλογή της κατάλληλης τιμής κατωφλίου είναι κρίσιμη. Συχνά η μεταβολή του φωτισμού, η μεταβολή της απόχρωσης του αντικειμένου ή του παρασκηνίου, απαιτούν την χρήση διαφορετικών τοπικών τιμών κατωφλίων. Αν εφαρμόσουμε όλες τις παραπάνω μάσκες σε μία εικόνα 3×3 , και αντιστοιχίσουμε το άθροισμα των αποτελεσμάτων στο κεντρικό εικονοστοιχείο, θα προκύψουν για κάθε κατεύθυνση αντίστοιχα οι τιμές

$$S_k = 1 \cdot I_{(k+1,j+1)} + 2 \cdot I_{(k+1,j)} + 1 \cdot I_{(k+1,j-1)} - 1 \cdot I_{(k-1,j+1)} - 2 \cdot I_{(k-1,j)} - 1 \cdot I_{(k-1,j-1)}$$

$$S_j = 1 \cdot I_{(k+1,j+1)} + 2 \cdot I_{(k+1,j)} + 1 \cdot I_{(k+1,j-1)} - 1 \cdot I_{(k+1,j-1)} - 2 \cdot I_{(k,j-1)} - 1 \cdot I_{(k-1,j-1)}$$

ή με την μορφή μάσκας

$$S_k=1/4 \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad S_j=1/4 \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

Η μάσκα αυτή λέγεται *ανιχνευτής ακμών του Sobel (Sobel edge detector)* και λειτουργεί ως διάνυσμα (S_k, S_j) με τις οικίες τιμές κατεύθυνσης και πλάτους. Πολλές μάσκες έχουν προταθεί από τους ερευνητές για την ανίχνευση ακμών βασισμένες σε διαφορετικές προσεγγίσεις για την εξυπηρέτηση διαφόρων ιδιαιτεροτήτων των εικόνων. Οι πλέον σημαντικοί είναι των

Prewitt

$$P_k=1/3 \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \quad P_j=1/3 \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

Roberts

$$R_k= \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \quad R_j= \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$

Frei - Chen

$$F_k= \frac{1}{2+\sqrt{2}} \begin{bmatrix} 1 & 0 & -1 \\ \sqrt{2} & 0 & -\sqrt{2} \\ 1 & 0 & -1 \end{bmatrix} \quad F_j= \frac{1}{2+\sqrt{2}} \begin{bmatrix} -1 & -\sqrt{2} & -1 \\ 0 & 0 & 0 \\ 1 & \sqrt{2} & 1 \end{bmatrix}$$

Σε κάθε περίπτωση από την αρχική εικόνα παράγεται ένας πίνακας τιμών του μέτρου ή της κλίσης σε κάθε εικονοστοιχείο της εικόνας. Ακολούθως μία τιμή κατωφλίου καθορίζει εκείνα τα εικονοστοιχεία που αντιστοιχούν σε ακμές και παράγουν την δυαδική εικόνα των ακμών.

Η επιλογή της κατάλληλης τιμής κατωφλίου είναι ένα επιπρόσθετο πρόβλημα. Η μεταβολή του φωτισμού, η μεταβολή της απόχρωσης του αντικειμένου ή του παρασκηνίου, απαιτούν την χρήση διαφορετικών τοπικών τιμών κατωφλίων.

ΠΑΡΑΔΕΙΓΜΑ: Δίνεται ο ακόλουθος πίνακας ψηφιακής εικόνας.

$$I = \begin{pmatrix} 10 & 100 & 100 & 20 & 30 \\ 10 & 100 & 110 & 20 & 30 \\ 10 & 100 & 140 & 140 & 30 \\ 10 & 10 & 110 & 20 & 30 \\ 10 & 10 & 100 & 20 & 30 \end{pmatrix}$$

Να ανιχνευθούν οι ακμές του με την κλίση που καθορίζουν οι μάσκες

a)

$$h_1 = \begin{bmatrix} -1 & 1 \end{bmatrix} \quad h_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

b) του Sobel

$$S_{k=} = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad S_{j=} = \begin{bmatrix} -1 & -2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & 1 \end{bmatrix}$$

a)

$D_k(k, j)$ με εφαρμογή της h_1

$D_j(k, j)$ με εφαρμογή της h_2

$$\begin{pmatrix} 10 & 90 & 0 & -80 & 10 \\ 10 & 90 & 10 & -90 & 10 \\ 10 & 90 & 40 & 0 & -110 \\ 10 & 0 & 100 & -90 & 10 \\ 10 & 0 & 90 & -80 & 10 \end{pmatrix} \quad \begin{pmatrix} 10 & 100 & 100 & 20 & 30 \\ 0 & 0 & 10 & 0 & 0 \\ 0 & 0 & 30 & 120 & 0 \\ 0 & -90 & -30 & -120 & 0 \\ 0 & 0 & -10 & 0 & 0 \end{pmatrix}$$

$|\mathbf{D}(k, j)|$

Οι ακμές για $|\mathbf{D}(k, j)| \geq 90$

14.14	134.54	100	82.46	31.62	0	1	1	0	0
10	90	14.14	90	10	0	1	0	1	0
10	90	50	120	110	0	1	0	1	1
10	90	104.4	150	10	0	1	1	1	0
10	0	90.55	80	10	0	0	1	0	0

b) Για τη μάσκα του Sobel

$D_k(k, j)$ με εφαρμογή της S_k

$D_j(k, j)$ με εφαρμογή της S_j

1/4	300	280	-240	-220	-60	, 1/4	120	320	340	180	80
	400	420	-200	-340	-200		0	40	200	280	120
	310	460	10	-380	-320		-90	-180	-90	0	0
	130	420	70	-340	-200		-90	-220	-290	-280	-120
	30	280	30	-220	-60		-30	-140	-250	-180	-80

$|\mathbf{D}(k, j)|$

Οι ακμές για

$|\mathbf{D}(k, j)| \geq 75$

1/4	323.11	425.21	416.17	284.25	100.00	1	1	1	0	0
	400.00	421.90	282.84	440.45	233.24	1	1	0	1	0
	322.80	493.96	90.55	380.00	320.00	1	1	0	1	1
	158.11	474.13	298.33	440.45	233.24	0	1	0	1	0
	42.43	313.05	251.79	284.25	100.00	0	1	0	0	0

3.4.2 Ανίχνευση ακμών με την χρήση του τελεστή Laplace

Η εύρεση των περιοχών που το μέτρο της κλίσης (άνυσμα πρώτων μερικών παραγώγων) έχει τοπικά μεγάλες τιμές μπορεί βασισθεί στην εύρεση περιοχών όπου η δεύτερες μερικές

παράγωγοι της συνάρτησης της εικόνας μηδενίζονται. Σε μία ψηφιακή εικόνα η δεύτερη παράγωγος κατά την οριζόντια διεύθυνση βασίζεται στην διπλή διαδοχική εφαρμογή του τελεστή ∇_{1K} . Συγκεκριμένα η $\frac{\partial^2 f(x, y)}{\partial x^2}$ αντικαθίσταται από την διαφορά

$$\nabla_{1K}(k+1,j) - \nabla_{1K}(k,j) = I(k+1,j) - I(k,j) - (I(k,j) - I(k-1,j)) = I(k+1,j) - 2I(k,j) + I(k-1,j)$$

Όμοια η $\frac{\partial^2 f(x, y)}{\partial y^2}$ αντικαθίσταται από την διαφορά

$$\nabla_{1J}(k,j+1) - \nabla_{1J}(k,j) = I(k,j+1) - I(k,j) - (I(k,j) - I(k,j-1)) = I(k,j+1) - 2I(k,j) + I(k,j-1)$$

Μπορούμε να εφαρμόσουμε τα παραπάνω σε μία μάσκα 3 X 3 με τιμές

$$\nabla^2 = \begin{bmatrix} & & \\ & 4 & \\ & & \end{bmatrix}$$

Η μάσκα αυτή είναι μία έκφραση του τελεστή Laplace μιας συνάρτησης $f(x,y)$

$$\nabla^2 f(x, y) = \frac{\partial^2 f(x, y)}{\partial x^2} + \frac{\partial^2 f(x, y)}{\partial y^2}$$

στο πίνακα $I(k,j)$. Οι ακμές βρίσκονται στα σημεία με τιμή μηδέν και εκατέρωθεν των οποίων υπάρχουν ετερόσημες τιμές. Επειδή μια ψηφιακή εικόνα αποτελείται από δείγματα τιμών, η μια μηδενική τιμή της $\nabla^2 I(j,k)$ πέφτει σπάνια σε συγκεκριμένο εικονοστοιχείο. Εκείνο που συμβαίνει κυρίως είναι σε δύο γειτονικές, οριζόντια ή κάθετα θέσεις, να εμφανίζονται ετερόσημες τιμές. Ένας απλός τρόπος να παράγουμε την εικόνα των ακμών είναι θεωρούμε σημείο ακμής κάθε στοιχείο $\nabla^2 I(j,k)$ όταν

- $\nabla^2 I(j,k) \cdot \nabla^2 I(j,k+1) < 0$ ή
- $\nabla^2 I(j,k) \cdot \nabla^2 I(j+1,k) < 0$ ή
- $\nabla^2 I(j,k) = 0$ και $\nabla^2 I(j,k-1) \cdot \nabla^2 I(j,k+1) < 0$ ή $\nabla^2 I(j-1,k) \cdot \nabla^2 I(j+1,k) < 0$

Η ύπαρξη ακμών δεν εξετάζεται στα εξωτερικά στοιχεία της εικόνας

Για παράδειγμα με την εφαρμογή των ανωτέρω οι ακμές της εικόνας I είναι τα σημεία με την τιμή 1 στο πίνακα A .

$$I = \begin{vmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 10 & 10 & 1 & 1 \\ 1 & 10 & 10 & 10 & 1 \\ 1 & 1 & 10 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{vmatrix} \quad \nabla^2 I = \begin{vmatrix} -2 & 8 & 8 & -1 & -2 \\ 8 & -18 & -18 & 18 & -1 \\ 8 & -18 & 0 & -27 & 8 \\ -1 & 18 & -27 & 18 & -1 \\ -2 & -1 & 8 & -1 & -2 \end{vmatrix}$$

$$A = \begin{vmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{vmatrix}$$

Η χρήση του τελεστή Laplace πλεονεκτεί, διότι εντοπίζονται τα τοπικά μέγιστα της πρώτης παραγώγου και προκύπτουν λεπτές ακμές. Μειονεκτεί όμως, διότι παρουσιάζει ευαισθησία στο θόρυβο και στις μικρές αυξομειώσεις της έντασης της φωτεινότητας. Για το λόγο αυτό εφαρμόζεται σε συνδυασμό με τον ∇_1 , τεχνικές τοπικής διασποράς και τεχνικές αφαίρεσης θορύβου.

Για την αφαίρεση του θορύβου χρησιμοποιείται ευρύτατα η μάσκα εξομάλυνσης του Gauss. Στην εικόνα που προκύπτει από την εξομάλυνση εφαρμόζεται η μάσκα Laplace. Επειδή οι δύο μάσκες είναι γραμμικά συστήματα ανεξάρτητα από την μετατόπιση με κρουστικές αποκρίσεις h_G και h_L αντίστοιχα η τελική έξοδος θα προκύπτει από την διαδοχική εφαρμογή του δισδιάστατου συνελκτικού αθροίσματος σύμφωνα με την σχέση

$$I' = h_L ** (h_G ** I) = (h_L ** h_G) ** I$$

Την μάσκα που αντιστοιχεί σε ένα σύστημα με κρουστική απόκριση $h_L ** h_G$ ονομάζουμε μάσκα LoG και αποτελεί την διακριτή έκδοση εφαρμογής του τελεστή Laplace σε δισδιάστατη Gaussian συνάρτηση. Ακολούθως φαίνονται οι τιμές μιας μάσκας Log 5 X 5

$$W_{LoG} = \begin{vmatrix} 0 & 0.0751 & 0.1238 & 0.0751 & 0 \\ 0.0751 & -0.0528 & -0.1410 & -0.0528 & 0.0751 \\ 0.1238 & -0.1410 & -0.3214 & -0.1410 & 0.1238 \\ 0.0751 & -0.0528 & -0.1410 & -0.0528 & 0.0751 \\ 0 & 0.0751 & 0.1238 & 0.0751 & 0 \end{vmatrix}$$

Στην γενική περίπτωση η εφαρμογή του τελεστή Laplace στην δισδιάστατη Gaussian συνάρτηση

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

ικανοποιεί την σχέση

$$\nabla^2 G(x, y) = -\frac{1}{\pi\sigma^4} e^{-\frac{x^2+y^2}{2\sigma^2}} \left(1 - \frac{x^2 + y^2}{2\sigma^2}\right)$$

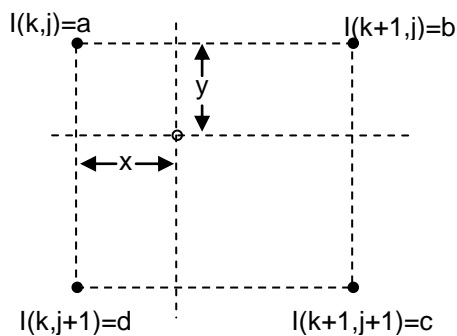
Με βάση όσα αναπτύχθηκαν παραπάνω έχουν προταθεί τεχνικές ανίχνευσης των ακμών που αντιμετωπίζουν προβλήματα όπως του θορύβου, της μεταβολής του φωτισμού, της συνέχειας των ακμών κ.α. Από τις πλέον επιτυχημένες και διαδεδομένες είναι ο ανιχνευτής ακμών του *Canny* (*Canny edge detector*).

3.5 Μεταβολή μεγέθους – Παρεμβολή τιμών

Η σμίκρυνση, η μεγέθυνση, η περιστροφή και άλλοι μετασχηματισμοί της εικόνας απαιτούν τον υπολογισμό της φωτεινότητας σε σημεία άλλα από αυτά που ορίζει ο αρχικός κάναβος των συντεταγμένων (k,j). Οι συνηθέστερες μέθοδοι που χρησιμοποιούνται για το σκοπό αυτό είναι:

- του κοντινότερου γείτονα (nearest neighbor)
- της δι-γραμμικής παρεμβολής (bi-linear interpolation)
- της δι-κυβικής παρεμβολής (bi-cubic interpolation)

Σύμφωνα με την μέθοδο του κοντινότερου γείτονα η τιμή στην θέση (x,y) είναι ίση με την τιμή του κοντινότερου προς αυτήν εικονοστοιχείου. Για παράδειγμα στο σχήμα 3.5.1. η τιμή της φωτεινότητας στη θέση (x,y) είναι $I(x,y) = I(k,j)$.



Σχήμα 3.5.1

Στην μέθοδο της δι-γραμμικής παρεμβολής υπολογίζεται η εξίσωση της ευθείας που διέρχεται από τα σημεία $I(k,j)$ και $I(k+1,j)$. Η εξίσωση της οριζόντιας αυτής ευθείας είναι:

$$I(x,j) = a+x \cdot (b-a)$$

Όμοια η εξίσωση της ευθείας που διέρχεται από τα σημεία $I(k,j+1)$ και $I(k+1,j+1)$ θα είναι

$$I(x,j+1) = d+x \cdot (c-d)$$

Η εξίσωση της ευθείας που διέρχεται από τα σημεία $I(x,j)$ και $I(x,j+1)$ είναι:

$$I(x,y) = I(x,j)+y \cdot [I(x,j+1)-I(x,j)]=a+(b-a) \cdot x+(d-a) \cdot y+x \cdot y \cdot (c+a-d-b)$$

Το αποτέλεσμα δεν αλλάζει αν η παρεμβολή ξεκινήσει από τις κάθετες ευθείες.

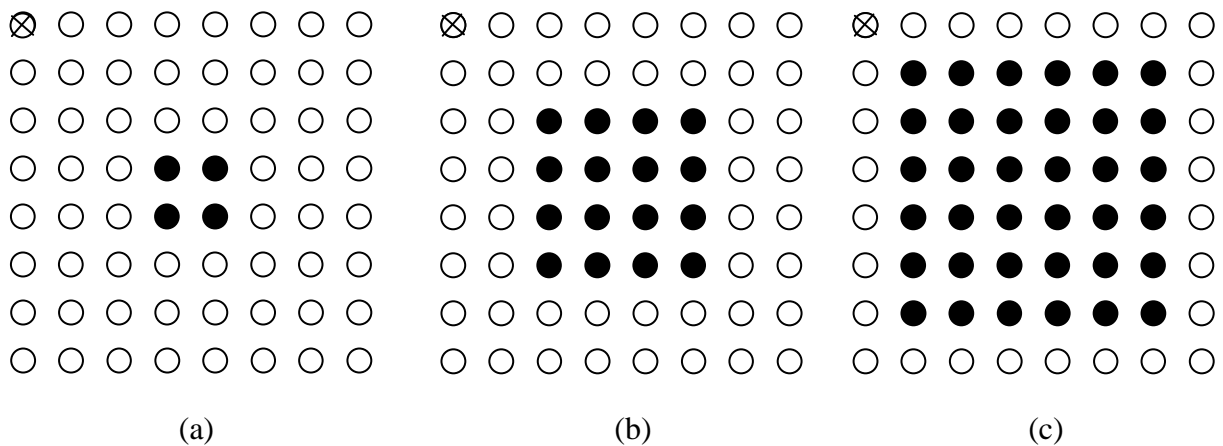
ΚΕΦΑΛΑΙΟ 4

ΜΟΡΦΟΛΟΓΙΑ ΨΗΦΙΑΚΩΝ ΔΥΑΔΙΚΩΝ ΕΙΚΟΝΩΝ

Μία προσπάθεια ακριβούς ερμηνείας του όρου μορφολογία μπορεί να οδηγήσει αρχικά στον ορισμό ότι αυτή ασχολείται με τη μορφή και τη δομή ενός αντικειμένου και τις αλληλοσυσχετίσεις μεταξύ των μερών του. Ειδικά, η ψηφιακή μορφολογία αποσκοπεί στη περιγραφή και την ανάλυση της μορφής ενός ψηφιακού αντικειμένου. Η μορφολογία έχει σχέση με το σχήμα και η ψηφιακή μορφολογία είναι ένας τρόπος να αναλύσουμε το σχήμα ενός ψηφιακού αντικειμένου. Στην παρούσα εργασία θα ασχοληθούμε με μορφολογικές πράξεις που αφορούν δυαδικές ψηφιακές εικόνες. Υπενθυμίζουμε στον αναγνώστη ότι μία ψηφιακή δυαδική εικόνα είναι ουσιαστικά ένα δυσδιάστατο διακριτό σήμα που παίρνει μόνο δύο τιμές. Η μετατροπή μιας εικόνας σε ψηφιακή μορφή ουσιαστικά απαιτεί την μετατροπή ενός δυσδιάστατου αναλογικού σήματος σε ψηφιακό και απαιτεί τις διαδικασίες της δειγματοληψίας και του κβαντισμού. Σ' αυτήν την περίπτωση, κάθε εικονοστοιχείο (picture element, pixel) μπορεί να χρωματιστεί με ένα από δύο χρώματα (συνήθως άσπρο ή μαύρο). Για κάθε εικονοστοιχείο απαιτείται ένα bit πληροφορίας, π.χ. με τιμή μηδέν (0) για το μαύρο και ένα (1) για λευκό. Οι εικόνες των εγγράφων που αποτελούνται μόνο από το χρώμα του χαρτιού και της μελάνης αναπαρίστανται σε δυαδική ψηφιακή μορφή.

Η επιστήμη της ψηφιακής μορφολογίας είναι σχετικά πρόσφατη και ξεκινά από τότε που οι ηλεκτρονικοί υπολογιστές έκαναν εφικτή μια τέτοια προσπάθεια. Από την άλλη τα μαθηματικά που απαιτούνται είναι μόνο η θεωρία συνόλων που είναι μια γνωστή επιστημονική περιοχή. Η βασική ιδέα που «κρύβεται» κάτω από την ψηφιακή μορφολογία είναι ότι οι εικόνες αποτελούνται από εικονοστοιχεία (pixels, picture-elements) σε ομάδες των οποίων συγκεκριμένες μορφολογικές πράξεις μπορούν να οδηγήσουν στην αναγνώριση και στην καταμέτρηση των σχημάτων στα οποία ανήκουν.

Οι δυαδικές μορφολογικές λειτουργίες ορίζονται σε εικόνες δύο αποχρώσεων, (bi-level, δυαδικές) δηλαδή εικόνες που αποτελούνται από μαύρα ή άσπρα pixel μόνο. Για παράδειγμα στην εικόνα (a) του σχήματος 1.. το σύνολο των μαύρων pixel διαμορφώνει ένα τετραγωνικό αντικείμενο. Το αντικείμενο στο σχήμα (b) είναι επίσης τετραγωνικό, αλλά είναι μεγαλύτερο προς όλες τις κατευθύνσεις κατά ένα pixel. Προέκυψε από το τετράγωνο (a) μαυρίζοντας κάθε λευκό pixel του που έχει ένα τουλάχιστον γειτονικό με μαύρο χρώμα.. Αυτό με όρους ψηφιακής μορφολογίας ισοδυναμεί με μια απλή ψηφιακή διαστολή και ονομάζεται έτσι επειδή αναγκάζει το αρχικό αντικείμενο να γίνει μεγαλύτερο. Το σχήμα (c) δείχνει το αποτέλεσμα της διαστολής του σχήματος (b) κατά ένα ακόμη pixel, το οποίο είναι το ίδιο διαστέλλοντας την (a) κατά δύο pixels.



Σχήμα 1.

Προβλήματα όπως η εξαγωγή μαύρων ή λευκών ανεπιθύμητων στιγμάτων λόγω θορύβου, ο εντοπισμός συγκεκριμένων μορφών στην εικόνα, ο έλεγχος ύπαρξης συγκεκριμένης υφής και η αποκατάσταση ατελειών σε περιγράμματα και μορφές βρίσκουν ικανοποιητικές λύσεις με την εφαρμογή μορφολογικών τεχνικών.

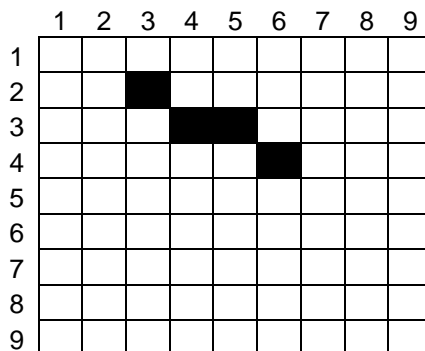
Οι βασικότερες μορφολογικές πράξεις που θα αναλύσουμε παρακάτω είναι οι:

- Μεταφορά (Translation)
- Ανάκλαση (Reflection)
- Συμπλήρωμα (Complement)
- Τομή (Intersection)
- Ένωση (Union)
- Διαφορά (Difference)

- Διαστολή (Dilation)
- Διάβρωση (Erosion)
- Άνοιγμα (Opening)
- Κλείσιμο (Closing)

Πριν περάσουμε στην αναλυτική περιγραφή των επόμενων κεφαλαίων θα αναφέρουμε εδώ βασικές μαθηματικές έννοιες και σημειογραφία που θα χρησιμοποιηθούν. Οι ορισμοί των μορφολογικών πράξεων βασίζονται στην θεωρία συνόλων [ΠΑΡΑΡΤΗΜΑ Β]. Για την μαθηματική περιγραφή των ανωτέρω μορφολογικών πράξεων, θα χρησιμοποιήσουμε βασικές έννοιες της θεωρίας συνόλων. Το διατεταγμένο ζεύγος (k,j) , $k, j \in Z$ των συντεταγμένων ενός εικονοστοιχείου μιας ψηφιακής, δυαδικής εικόνας θα αποτελεί στοιχείο του συνόλου στα οποία θα αναφερθούμε. Το διατεταγμένο αυτό ζεύγος θα απεικονίζεται με ένα έντονο πεζό γράμμα

Για παράδειγμα το σύνολο A των εικονοστοιχείων που αποτελούν την μορφή M του παρακάτω σχήματος



Σχήμα 2.

είναι:

$$A = \{(3,2), (4,3), (5,3), (6,4)\} \text{ (αναγραφή)}$$

ή

$$A = \{c | c \in M\} \text{ (περιγραφή)}.$$

4.1 Μεταφορά (Translation)

Η μεταφορά του συνόλου \mathbf{A} στο σημείο \mathbf{x} ορίζεται ως:

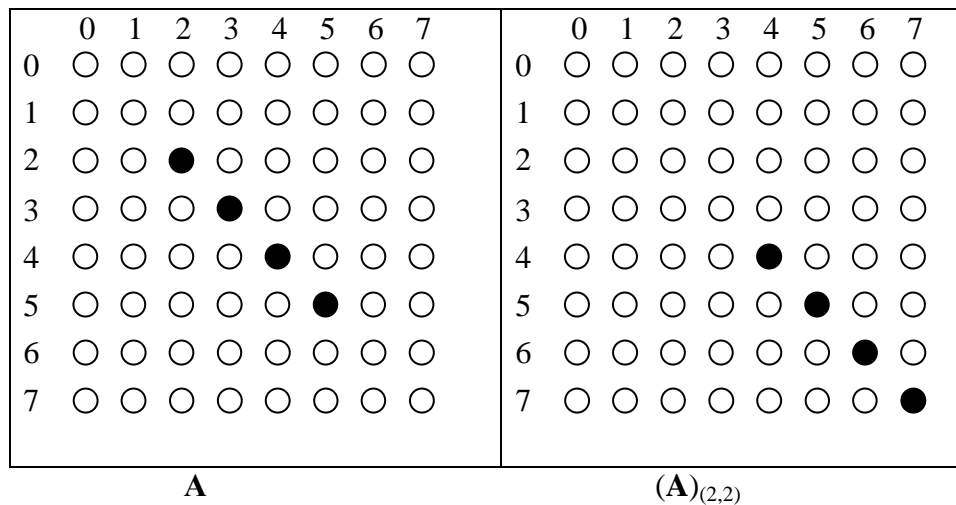
$$(\mathbf{A})_{\mathbf{x}} = \{ \mathbf{c} \mid \mathbf{c} = \mathbf{a} + \mathbf{x}, \mathbf{a} \in \mathbf{A} \} \quad (1)$$

Για παράδειγμα, εάν το \mathbf{x} είναι $(2,2)$, τότε το πρώτο (πάνω αριστερά) pixel του $\mathbf{A}_{\mathbf{x}}$ θα ήταν $(2,2) + (2,2) = (4,4)$. Σ' αυτήν την περίπτωση όλα τα εικονοστοιχεία μετατοπίζονται κατά δύο (2) σειρές και κατά δύο (2) στήλες. Στο ακόλουθο Σχήμα 3 φαίνεται η εφαρμογή της προηγούμενης πράξης μεταφοράς κατά $(2,2)$, μιας μορφής που περιγράφεται από το σύνολο:

$$\mathbf{A} = \{ (2,2), (3,3), (4,4), (5,5) \}$$

και το αποτέλεσμα της

$$(\mathbf{A})_{(2,2)} = \{ (4,4), (5,5), (6,6), (7,7) \}.$$



Σχήμα 3.

4.2 Ανάκλαση (Reflection)

Η ανάκλαση του συνόλου \mathbf{A} ορίζεται ως:

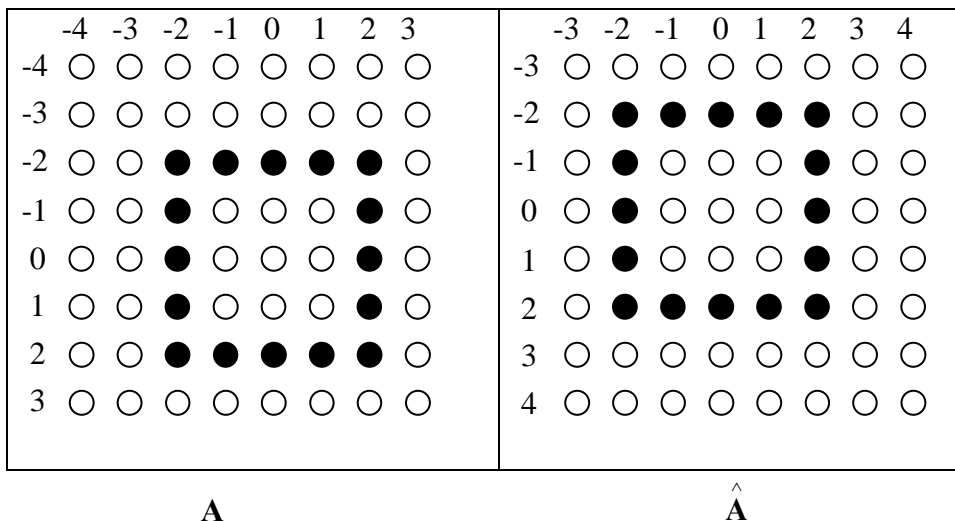
$$\hat{\mathbf{A}} = \{\mathbf{c} \mid \mathbf{c} = -a, a \in \mathbf{A}\} \quad (2)$$

Η ανάκλαση έχει ως αποτέλεσμα την περιστροφή του αντικειμένου \mathbf{A} κατά 180° γύρω από την αρχή των αξόνων. Με τον τρόπο αυτό παράγεται μια κατοπτρική εικόνα της αρχικής. Αν για παράδειγμα θεωρήσουμε την μορφή του Σχήματος 4 τότε:

$$\mathbf{A} = \{ (2,2), (2,1), (2,0), (2,-1), (2,-2), (1,2), (1,-2), (0,2), (0,-2), (-1,2), (-1,-2), (-2,2), (-2,1), (-2,0), (-2,-1), (-2,-2) \}$$

και

$$\hat{\mathbf{A}} = \{ (-2,-2), (-2,-1), (-2,0), (-2,1), (-2,2), (-1,-2), (-1,2), (0,-2), (0,2), (1,-2), (1,2), (2,-2), (2,-1), (2,0), (2,1), (2,2) \}$$



Σχήμα 4.

4.3 Συμπλήρωμα (Complement)

Το συμπλήρωμα του συνόλου A είναι το σύνολο των pixel που δεν ανήκουν στο A και ορίζεται ως:

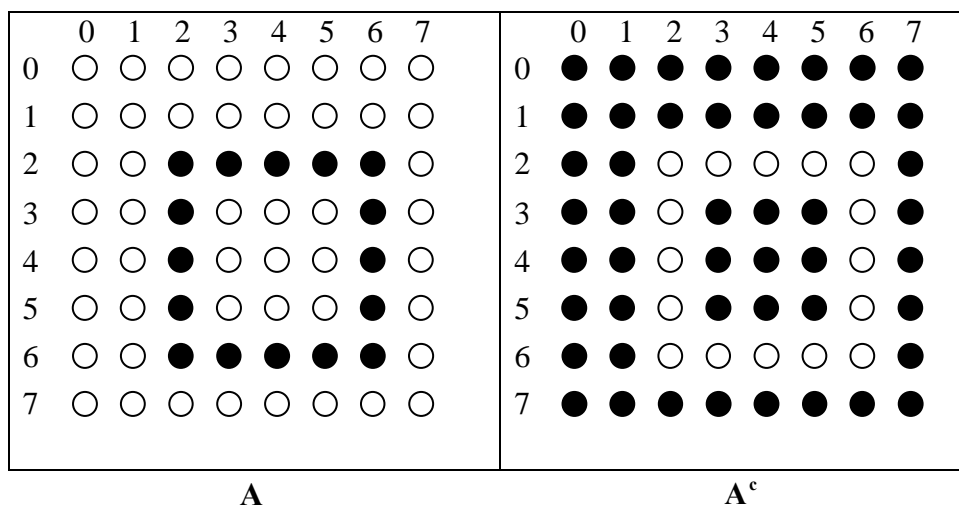
$$A^c = \{c \mid c \notin A\} \quad (3)$$

Αν για παράδειγμα θεωρήσουμε την μορφή του Σχήματος 5 τότε

$A = \{ (2,2), (3,2), (4,2), (5,2), (6,2), (2,3), (2,4), (2,5), (2,6), (6,3), (6,4), (6,5), (6,6), (3,6), (4,6), (5,6) \}$

και

$A^c = \{ (0,0), (0,1), (0,2), (0,3), (0,4), (0,5), (0,6), (0,7), (1,0), (1,1), (1,2), (1,3), (1,4), (1,5), (1,6), (1,7), (2,0), (2,1), (2,7), (3,0), (3,1), (3,3), (3,4), (3,5), (3,7), (4,0), (4,1), (4,3), (4,4), (4,5), (4,7), (5,0), (5,1), (5,3), (5,4), (5,5), (5,7), (6,0), (6,1), (6,7), (7,0), (7,1), (7,2), (7,3), (7,4), (7,5), (7,6), (7,7) \}$



Σχήμα 5

4.4 Τομή (Intersection)

Η τομή των δύο συνόλων **A** και **B** είναι το σύνολο των στοιχείων που ανήκουν και στα δύο, (στο **A** και στο **B**) και ορίζεται ως:

$$\mathbf{A \cap B = \{c \mid c ((c \in A) \wedge (c \in B))\}} \quad (4)$$

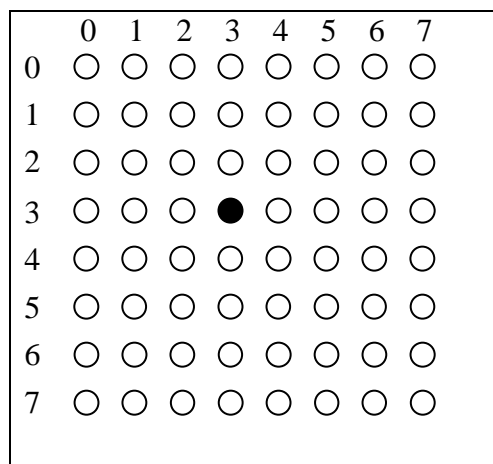
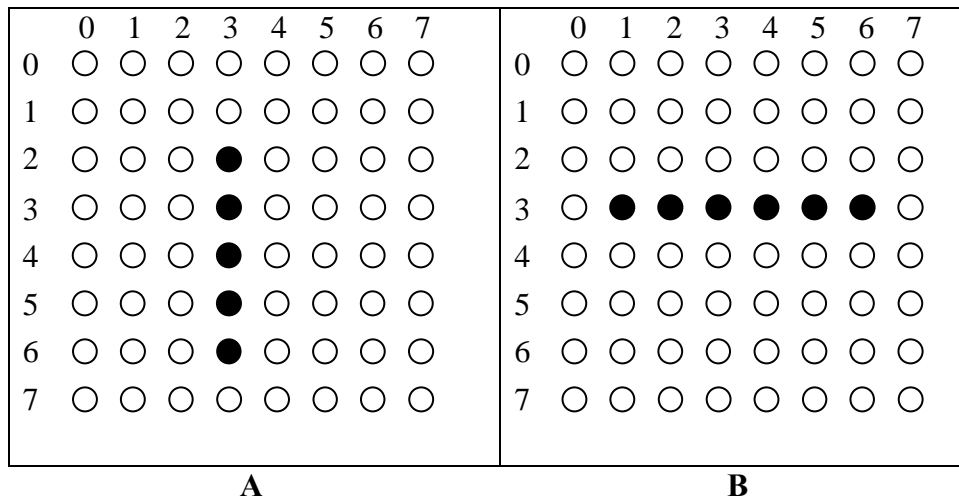
Αν για παράδειγμα θεωρήσουμε το Σχήμα 6 τότε:

$$\mathbf{A = \{ (2,3), (3,3), (4,3), (5,3), (6,3) \}}$$

$$\mathbf{B = \{ (3,1), (3,2), (3,3), (3,4), (3,5), (3,6) \}}$$

και το αποτέλεσμα της

$$\mathbf{A \cap B = \{ (3,3) \}}$$



A ∩ B

Σχήμα 6.

4.5 Ένωση (Union)

Η ένωση δύο συνόλων **A** και **B** είναι το σύνολο των pixel που ανήκουν ή στο **A** ή στο **B** ή και στα δύο και ορίζεται ως:

$$\mathbf{A \cup B = \{c \mid c ((c \in A) \vee (c \in B))\}} \quad (5)$$

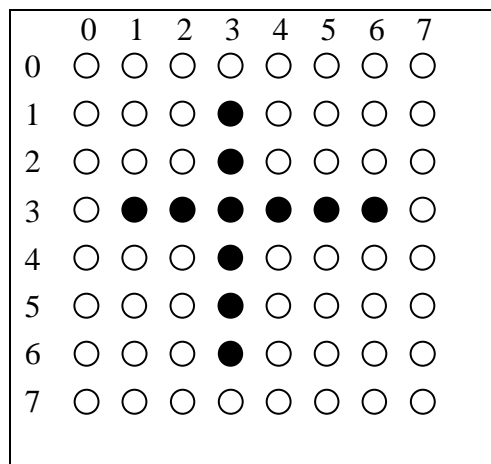
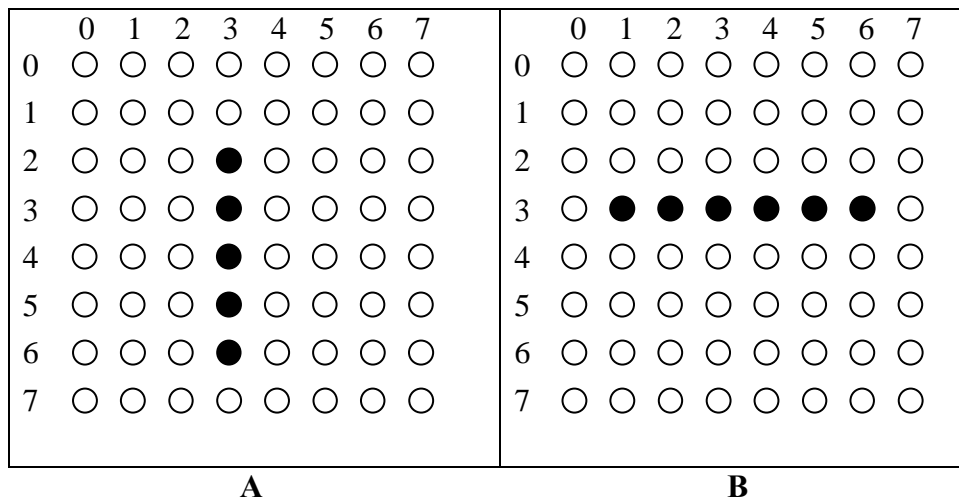
Αν για παράδειγμα θεωρήσουμε τις εικόνες του Σχήματος 7 τότε:

$$\mathbf{A = \{ (2,3), (3,3), (4,3), (5,3), (6,3) \}}$$

$$\mathbf{B = \{ (3,1), (3,2), (3,3), (3,4), (3,5), (3,6) \}}$$

και το αποτέλεσμα της

$$\mathbf{A \cup B = \{ (2,3), (3,3), (4,3), (5,3), (6,3) (3,1), (3,2), (3,4), (3,5), (3,6) \}}$$



A ∪ B

Σχήμα 7

4.6 Διαφορά (Difference)

Η διαφορά μεταξύ του συνόλου **A** και **B** είναι:

$$\mathbf{A - B} = \{c \mid c ((c \in \mathbf{A}) \wedge (c \notin \mathbf{B}))\} \tag{6}$$

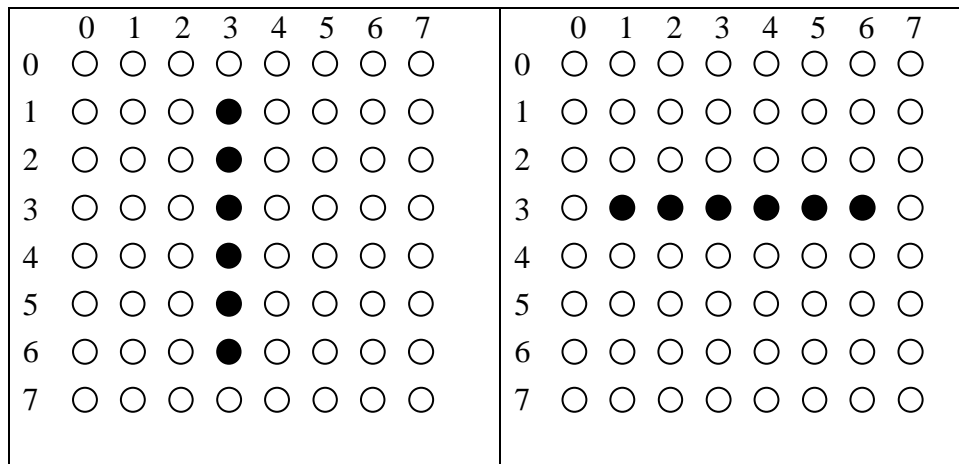
η οποία είναι το σύνολο των pixel που ανήκουν στο **A** αλλά όχι και στο **B**. Αυτό είναι ακριβώς η τομή του **A** με το συμπλήρωμα του **B** ή $\mathbf{A \cap B^c}$. Αν για παράδειγμα θεωρήσουμε τις εικόνες του Σχήματος 8 τότε :

$$\mathbf{A} = \{ (2,3), (3,3), (4,3), (5,3), (6,3) \}$$

$$\mathbf{B} = \{ (3,1), (3,2), (3,3), (3,4), (3,5), (3,6) \}$$

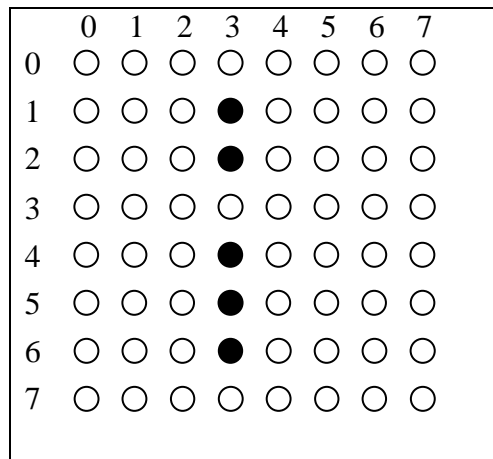
και το αποτέλεσμα της

$$\mathbf{A - B} = \{ (2,3), (4,3), (5,3), (6,3) \}$$



A

B



A – B

Σχήμα 8.

4.7 Διαστολή (Dilation)

Η διαστολή ορίζεται ως:

$$\mathbf{A} \oplus \mathbf{B} = \{ \mathbf{c} \mid \mathbf{c} = \mathbf{a} + \mathbf{b}, \mathbf{a} \in \mathbf{A}, \mathbf{b} \in \mathbf{B} \} \quad (7)$$

όπου το **A** περιγράφει την αρχική εικόνα και το **B** είναι ένα δεύτερο σύνολο pixels που επενεργεί επί του **A** για να παράγει το τελικό αποτέλεσμα. Το σύνολο **B** ονομάζεται στοιχείο δόμησης (Σ.Δ., SE: Structuring Element) ή δομικό στοιχείο ή πυρήνας (kernel) και καθορίζει την επίδραση της διαστολής στην αρχική εικόνα. Όσο πιο μεγάλο είναι το Δ., τόσο πιο «έντονη» θα είναι η επίδρασή του.

Για να γίνει πιο κατανοητό, υποθέτουμε ότι το **A** είναι το σύνολο του σχήματος 9(a) και το **B** είναι το σύνολο $\{ (0,0) (0,1) \}$.

Το αποτέλεσμα είναι: $\mathbf{A}_{(0,0)}$

$$\begin{aligned} (3,3) + (0,0) &= (3,3) & (3,4) + (0,0) &= (3,4) \\ (4,3) + (0,0) &= (4,3) & (4,4) + (0,0) &= (4,4) \end{aligned}$$

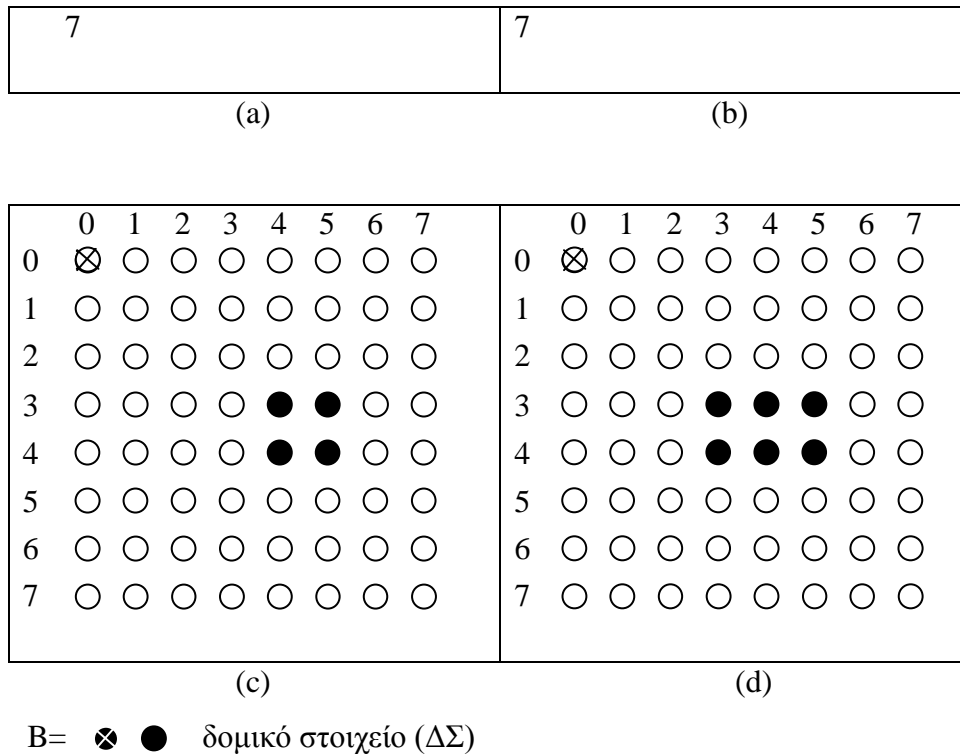
Το αποτέλεσμα του $\mathbf{A}_{(0,1)}$ είναι:

$$\begin{aligned} (3,3) + (0,1) &= (3,4) & (3,4) + (0,1) &= (3,5) \\ (4,3) + (0,1) &= (4,4) & (4,4) + (0,1) &= (4,5) \end{aligned}$$

Το σύνολο (c) είναι το αποτέλεσμα της διαστολής του συνόλου **A**, χρησιμοποιώντας το Δ. **B**.

Το Σχήμα 9 επεξηγεί αυτήν τη λειτουργία, την επίδραση της διαστολής.

A=	0	1	2	3	4	5	6	7		0	1	2	3	4	5	6	7	
0	⊗	○	○	○	○	○	○	○		0	⊗	○	○	○	○	○	○	
1	○	○	○	○	○	○	○	○		1	○	○	○	○	○	○	○	
2	○	○	○	○	○	○	○	○		2	○	○	○	○	○	○	○	
3	○	○	○	●	●	○	○	○		3	○	○	○	●	●	○	○	
4	○	○	○	●	●	○	○	○		4	○	○	○	●	●	○	○	
5	○	○	○	○	○	○	○	○		5	○	○	○	○	○	○	○	
6	○	○	○	○	○	○	○	○		6	○	○	○	○	○	○	○	
	○	○	○	○	○	○	○	○			○	○	○	○	○	○	○	



Σχήμα 9 Διαστολή του συνόλου **A** από το σύνολο **B**

(a) Σύνολο **A** (εικόνα εισαγωγής)

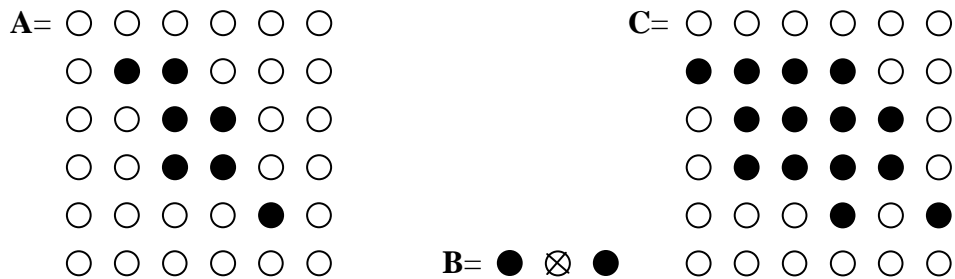
(b) Το αποτέλεσμα $A_{(0,0)}$

(c) Το αποτέλεσμα $A_{(0,1)}$

(d) Η ένωση των δύο συνόλων (Σχήμα 9(b) και 9(c)) είναι το αποτέλεσμα της διαστολής (εικόνα εξαγωγής)

Τα pixel που μαρκάρονται με ένα “X”, είτε άσπρο είτε μαύρο, αντιπροσωπεύουν το όρισμα της κάθε εικόνας. Η θέση του ορίσματος είναι σημαντική. Στο παραπάνω παράδειγμα, εάν το όρισμα του **B** ήταν δεξιά των δύο pixel, τότε το **B** θα ήταν $\{(0,-1) (0,0)\}$.

Στο επόμενο παράδειγμα, εξετάζουμε το αντικείμενο και το Σ.Δ. που παρουσιάζεται στο Σχήμα 10. Σ' αυτήν την περίπτωση, το όρισμα του ΣΔ **B** είναι ένα άσπρο ριxel, που υπονοεί ότι το όρισμα δεν συμπεριλαμβάνεται στο σύνολο **B**.



Σχήμα 10 Διαστολή από ένα Σ.Δ. στο οποίο δεν περιλαμβάνεται το όρισμα.

Το σύνολο **A** ισούται με { (1,1) (1,2) (2,2) (2,3) (3,2) (3,3) (4,4) }
 και το Σ.Δ **B** ισούται με { (0,-1) (0,1) }

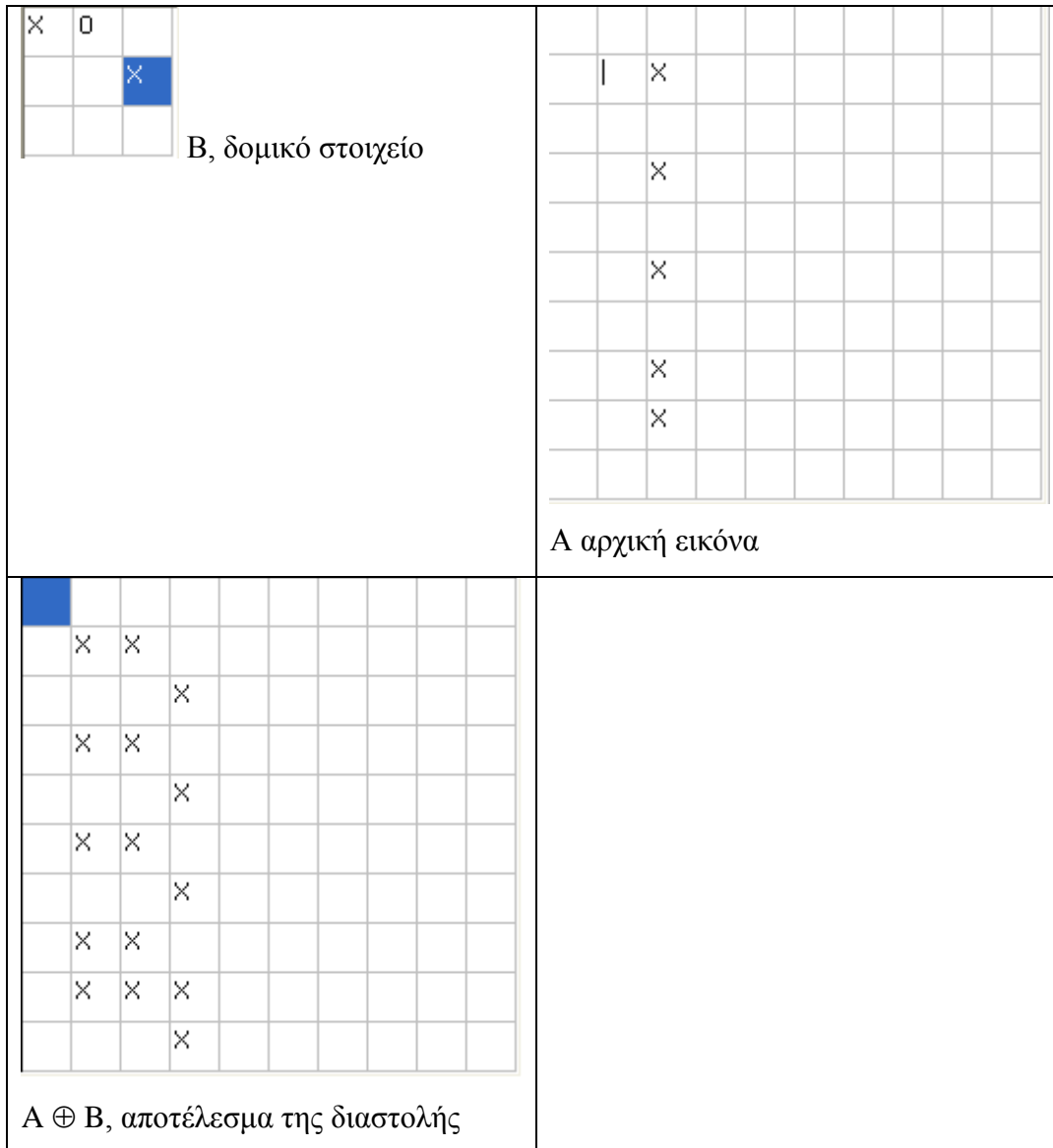
Το αποτέλεσμα του $A_{(0,-1)}$ είναι:

- (1,1) + (0,-1) = (1,0)
- (1,2) + (0,-1) = (1,1)
- (2,2) + (0,-1) = (2,1)
- (2,3) + (0,-1) = (2,2)
- (3,2) + (0,-1) = (3,1)
- (3,3) + (0,-1) = (3,2)
- (4,4) + (0,-1) = (4,3)

Το αποτέλεσμα του $A_{(0,1)}$ είναι:

- (1,1) + (0,1) = (1,2)
- (1,2) + (0,1) = (1,3)
- (2,2) + (0,1) = (2,3)
- (2,3) + (0,1) = (2,4)
- (3,2) + (0,1) = (3,3)
- (3,3) + (0,1) = (3,4)
- (4,4) + (0,1) = (4,5)

Επειδή ισχύει ότι $\mathbf{a+b = b+a}$ συμπεραίνουμε ότι η διαστολή είναι αντιμεταθετική πράξη. Με βάση αυτό μπορούμε να την ερμηνεύσουμε λέγοντας ότι το δομικό στοιχείο μεταφέρεται «κρατώντας» το από το όρισμα σε κάθε θέση που ορίζουν τα εικονοστοιχεία της αρχικής εικόνας. Στο παράδειγμα του Σχήματος 11 φαίνεται η ερμηνεία αυτή.



Σχήμα 11.

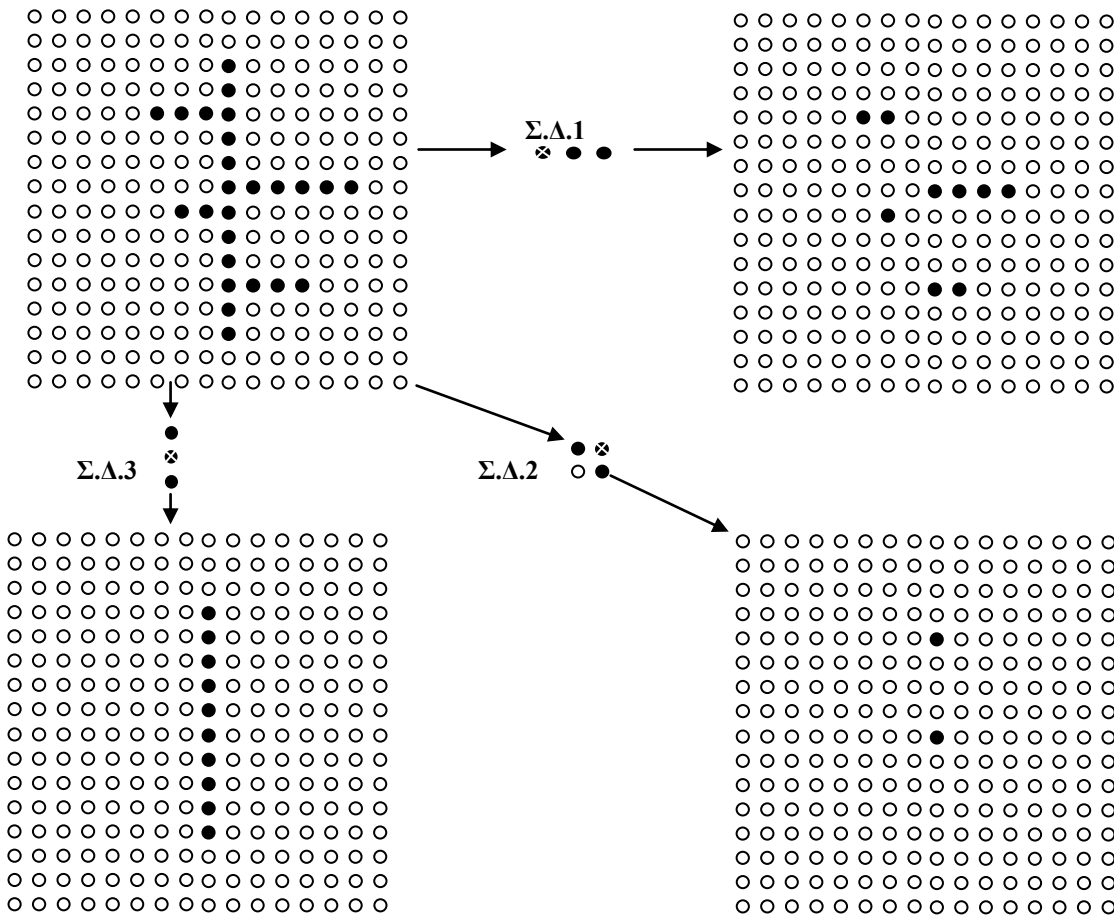
4.8 Διάβρωση (Erosion)

Εάν η διαστολή προσθέτει pixel σ' ένα αντικείμενο ή το καταστεί μεγαλύτερο, τότε η διάβρωση θα το καταστήσει μικρότερο.

Η διάβρωση ορίζεται ως:

$$\mathbf{A} \ominus \mathbf{B} = \{ \mathbf{c} \mid (\mathbf{B})_c \subseteq \mathbf{A} \} \quad (8)$$

Συνεπώς, είναι το σύνολο όλων των pixel στα οποία αν μεταφερθεί το στοιχείο δόμησης **B** τα εικονοστοιχεία της μεταφοράς είναι στοιχεία του **A**. Οποιαδήποτε pixels του συνόλου **A** που δεν ταιριάζουν με το σχέδιο (πατρόν) το οποίο καθορίζεται από τα μαύρα pixel του Σ.Δ., δεν θα ανήκουν στο αποτέλεσμα. Δηλαδή, τα μαύρα pixel του Σ.Δ. πρέπει να αντιστοιχούν στα μαύρα pixel της εικόνας εισαγωγής, προκειμένου να εμφανιστεί το όρισμα στο αποτέλεσμα (εικόνα εξαγωγής). Όσο πιο μεγάλο είναι το Σ.Δ., τόσο πιο ακραία είναι η επίδρασή του. Το παρακάτω παράδειγμα θα μας βοηθήσει να γίνει πιο κατανοητή η παράξη της διάβρωσης.



Σχήμα 12: Παραδείγματα διάβρωσης από 3 διαφορετικά στοιχεία δόμησης (Σ.Α.).

Το πρώτο ΣΔ αποτελείται από 3 μαύρα pixels. Το όρισμα μόλις «πατήσει» στο πρώτο μαύρο pixel που θα βρει (π.χ. $j=2, k=8$), θα ελέγξει αν το «πατρών» του «ταιριάζει», δηλαδή αν από τα δεξιά του έχει άλλα 2 μαύρα pixel. Αν δεν έχει, τότε το pixel που ταυτίζεται με το όρισμα δεν θα εμφανιστεί στο αποτέλεσμα, όπως στο παράδειγμα. Στην αντίθετη περίπτωση, το πατρών ταιριάζει και το pixel θα εμφανιστεί στην αντίστοιχη θέση της εικόνας εξαγωγής. Στη συγκεκριμένη περίπτωση, είναι το pixel της γραμμής $j=4$ και στήλης $k=5$.

Το δεύτερο ΣΔ αποτελείται από 4 pixels, από τα οποία το ένα είναι άσπρο. Το όρισμα μόλις ‘πατήσει’ στο μαύρο pixel της θέσης $j=4$ και $k=8$, θα ελέγξει αν υπάρχουν μαύρα pixel από αριστερά και από κάτω. Διαγωνίως δεν μας ενδιαφέρει γιατί το pixel είναι άσπρο. Στη συγκεκριμένη περίπτωση, το πατρών ταιριάζει, συνεπώς το αντίστοιχο pixel της θέσης του ορίσματος θα εμφανιστεί στην αντίστοιχη θέση στην εικόνα εξαγωγής.

Το τρίτο ΣΔ ενεργεί με ανάλογο τρόπο με το πρώτο ΣΔ αλλά κατά την κάθετη διεύθυνση

4.9 Άνοιγμα (Opening)

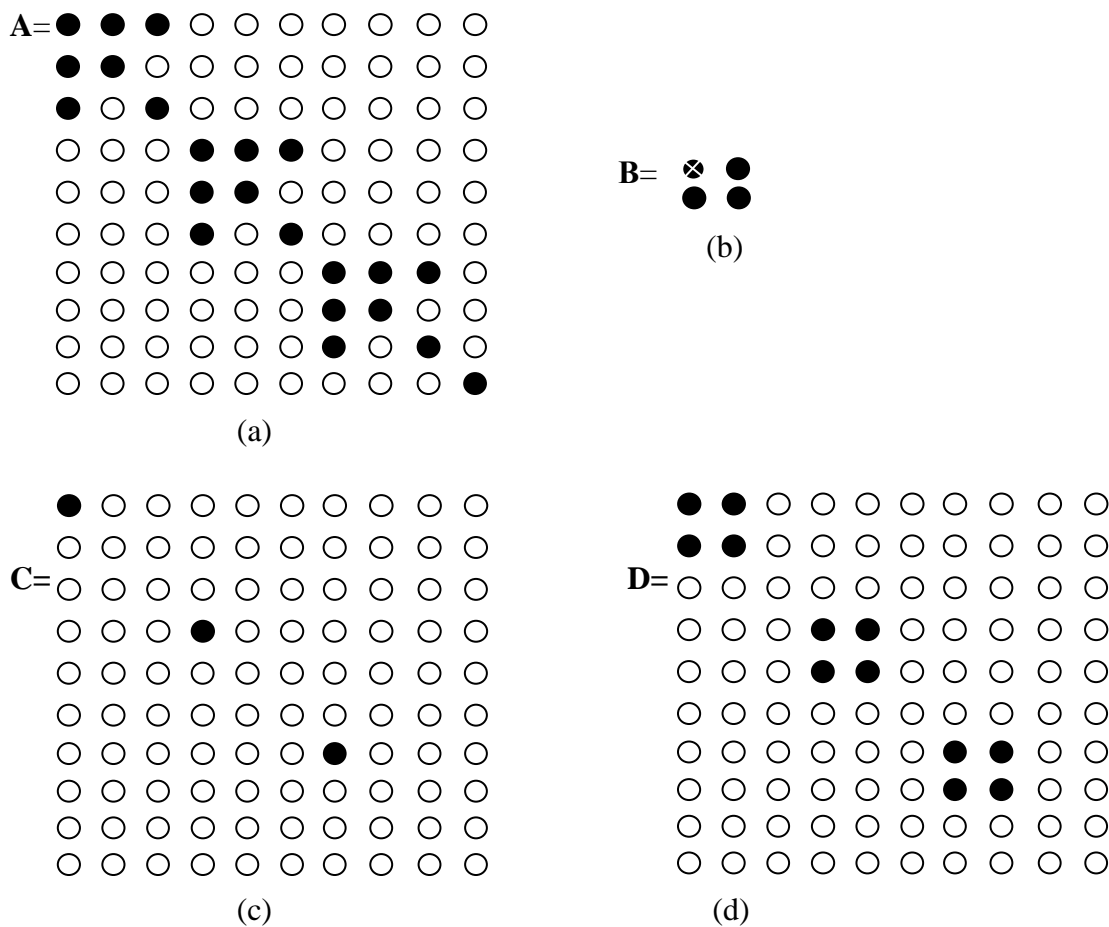
Όπως έχουμε δει, η διαστολή επεκτείνει τα στοιχεία μιας εικόνας και η διάβρωση τα συρρικνώνει. Η εφαρμογή μιας διάβρωσης πάνω σε μια εικόνα, ακολουθούμενη αμέσως από μια διαστολή χρησιμοποιώντας το ίδιο δομικό στοιχείο αναφέρεται ως λειτουργία *Opening*. Ο όρος *Opening* είναι περιγραφικός και τείνει "να ανοίξει" τα μικρά χάσματα ή τα διαστήματα μεταξύ αντικειμένων που αγγίζονται σε μια εικόνα. Το *Opening* γενικά λειαίνει το περίγραμμα ενός αντικειμένου, σπάει τους στενούς ισθμούς, και αποβάλλει τις λεπτές προεξοχές. Αυτή η επίδραση παρατηρείται ευκολότερα όταν χρησιμοποιείται ένα απλό στοιχείο δόμησης

Το *opening* του συνόλου **A** με δομικό στοιχείου το **B**, συμβολίζεται $A \circ B$, και ορίζεται ως

$$A \circ B = \{(A \ominus B) \oplus B\} \tag{9}$$

Κατά συνέπεια, το *opening* του **A** από το **B** είναι η διάβρωση του **A** από **B**, ακολουθούμενη από μια διαστολή χρησιμοποιώντας το ίδιο στοιχείο δόμησης **B**.

Για παράδειγμα ας εξετάζουμε το αντικείμενο και το Σ.Δ. που παρουσιάζεται στο Σχήμα 13.



Σχήμα 13

Σχήμα 13. Opening του συνόλου **A** από το σύνολο **B**

- (a) Σύνολο **A** (εικόνα εισαγωγής)
- (b) Στοιχείο δόμησης **B** ($\Sigma\Delta$)
- (c) Διάβρωση του συνόλου **A** από το $\Sigma\Delta$ **B**
- (d) Διαστολή του συνόλου **C** από το $\Sigma\Delta$ **B**(εικόνα εξαγωγής)

Η λειτουργία *opening* ικανοποιεί τις ακόλουθες ιδιότητες:

- (a) $A \circ B$ είναι ένα υποσύνολο (υποεικόνα) του A
- (b) Εάν C είναι ένα υποσύνολο του D , τότε $C \circ B$ είναι ένα υποσύνολο του $D \circ B$
- (c) $(A \circ B) \circ B = A \circ B$.

4.10 Κλείσιμο (Closing)

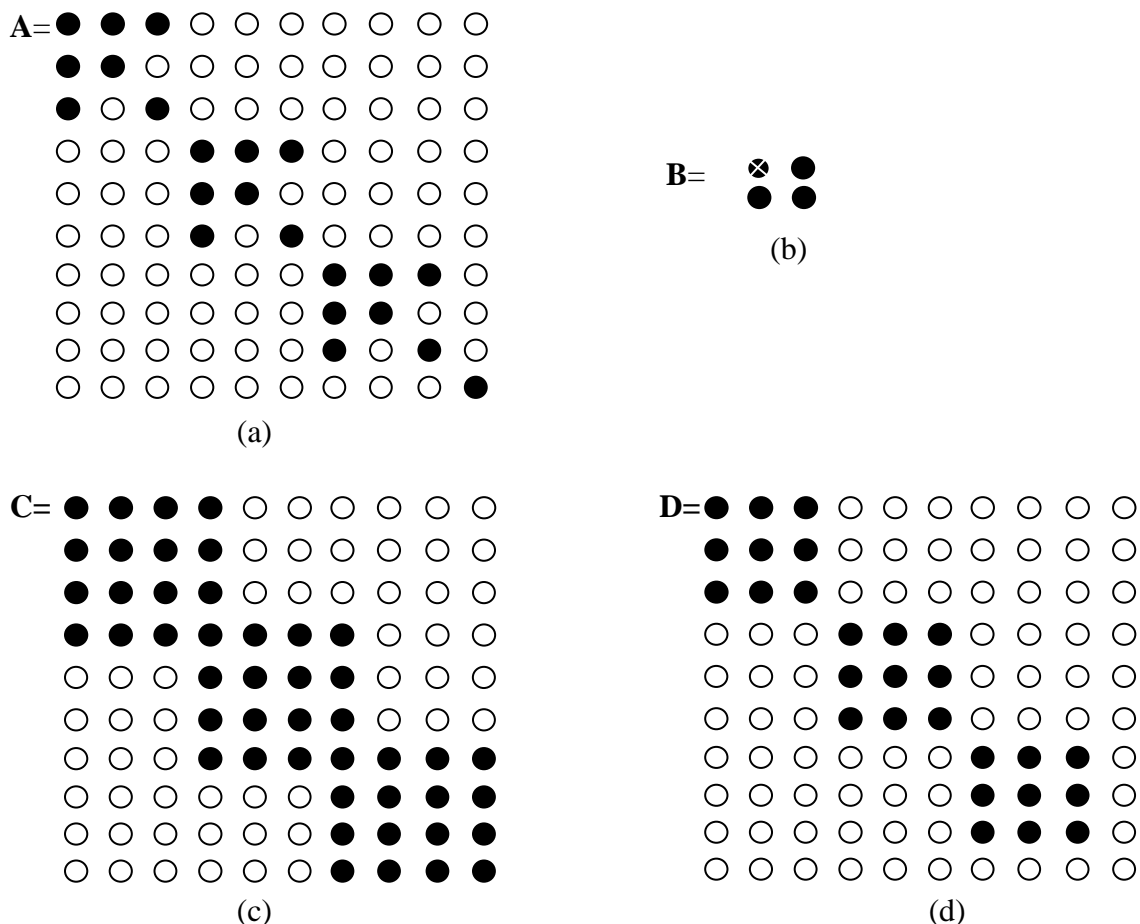
Η λειτουργία *closing* είναι παρόμοια με την *Opening* εκτός από ότι η διαστολή εκτελείται πρώτα, ακολουθούμενη από μια διάβρωση χρησιμοποιώντας το ίδιο δομικό στοιχείο. Εάν ένα *Opening* δημιουργεί μικρά κενά στην εικόνα, ένα *Closing* γεμίζει, ή "κλείνει" τα κενά. Το *closing* γενικά τείνει να λειάνει τα τμήματα των περιγραμμάτων αλλά, σε αντιδιαστολή με το *Opening*, απαλείφει τα λεπτά σπασίματα και τους μακριούς λεπτούς κόλπους, αποβάλλει τις μικρές τρύπες, και γεμίζει τα κενά στο περίγραμμα.

Το *closing* του συνόλου **A** με δομικό στοιχείο το **B**, συμβολίζεται $\mathbf{A} \bullet \mathbf{B}$, και ορίζεται ως

$$\mathbf{A} \bullet \mathbf{B} = \{(\mathbf{A} \oplus \mathbf{B}) \ominus \mathbf{B}\} \tag{10}$$

Δηλαδή, το *closing* του συνόλου **A** από το δομικό στοιχείο **B** είναι απλά η διαστολή του **A** από το **B**, ακολουθούμενη από τη διάβρωση του αποτελέσματος κατά **B**.

Για παράδειγμα ας εξετάζουμε το αντικείμενο και το Σ.Δ. που παρουσιάζεται στο Σχήμα 14.



Σχήμα 14.

Σχήμα 14. Closing του συνόλου **A** από το σύνολο **B**

- (a) Σύνολο **A** (εικόνα εισαγωγής)
- (b) Στοιχείο δόμησης **B** ($\Sigma\Delta$)
- (c) Διαστολή του συνόλου **A** από το $\Sigma\Delta$ **B**
- (d) Διάβρωση του συνόλου **C** από το $\Sigma\Delta$ **B**(εικόνα εξαγωγής)

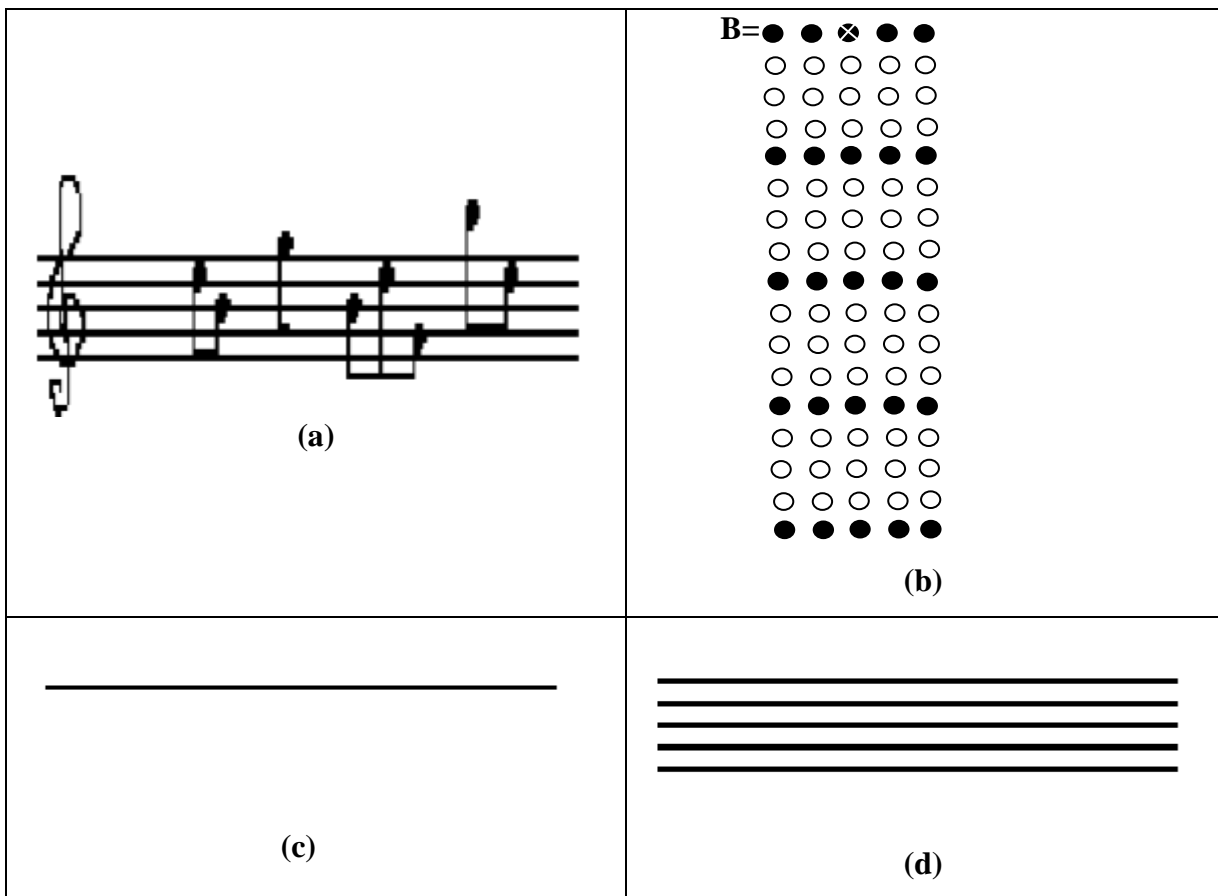
Η λειτουργία **closing** ικανοποιεί τις ακόλουθες ιδιότητες:

- (a) A είναι ένα υποσύνολο του (υποεικόνα) of $A \bullet B$
- (b) Εάν C είναι ένα υποσύνολο του D , τότε $C \bullet B$ είναι ένα υποσύνολο του $D \bullet B$
- (c) $(A \bullet B) \bullet B = A \bullet B$

Οι πρώτες 6 μέθοδοι βρίσκουν εφαρμογή στην απλή επεξεργασία εικόνων, ενώ οι 4 τελευταίες (Dilation , Erosion, Opening, Closing) αποτελούν τη βάση για την μορφολογική επεξεργασία εικόνων.

4.11 ΕΦΑΡΜΟΓΕΣ –ΠΑΡΑΔΕΙΓΜΑΤΑ

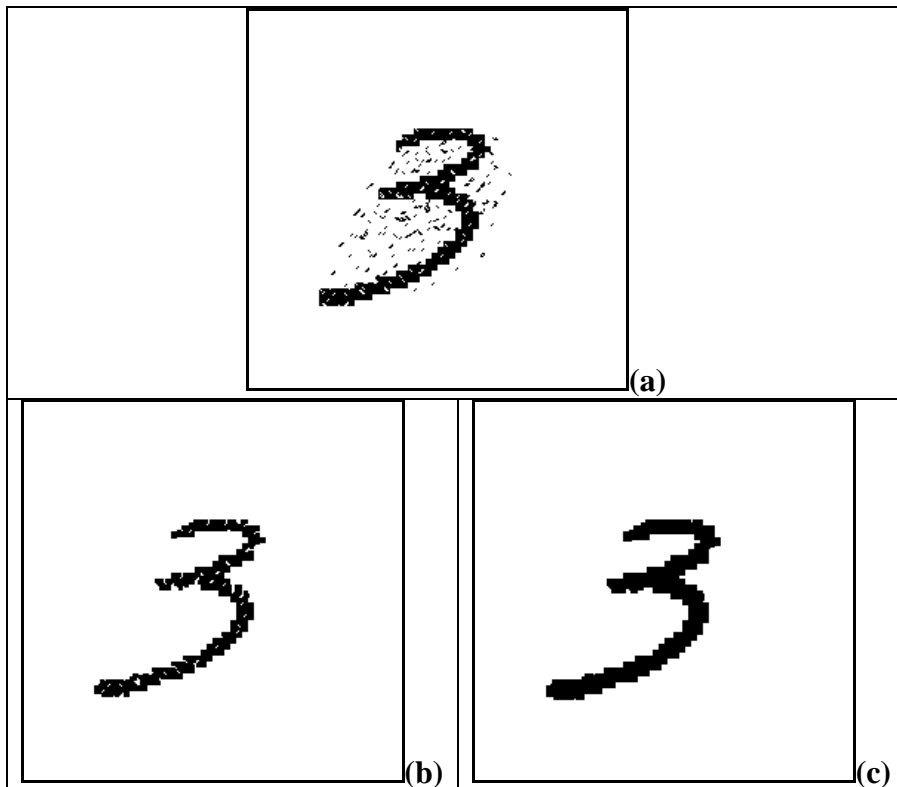
Το παράδειγμα [1] είναι ένα πραγματικό παράδειγμα opening με ένα σύνθετο στοιχείο δόμησης. Η εικόνα διαβρώθηκε (αφήνοντας μόνο μια οριζόντια γραμμή), και έπειτα υπεστεί διαστολή από το ίδιο στοιχείο δόμησης. Αυτό που διαβρώθηκε σε αυτήν την περίπτωση είναι όλα τα τμήματα της εικόνας που δεν είναι ίσιες γραμμές, τα οποία μετέπειτα η διαστολή τα αποκαθιστά. Το οποίο, βεβαίως όπως γνωρίζουμε, είναι η λειτουργία του opening.



Παράδειγμα 1. (a) Εικόνα εισαγωγής (b) Στοιχείο δόμησης (c) Αποτέλεσμα διάβρωσης της εικόνας (a) από το ΣΔ B (d) Αποτέλεσμα διαστολής της εικόνας (c) από το ίδιο Σ.Δ.

Το παράδειγμα 2 επεξηγεί μια αρκετά κοινή χρήση της λειτουργίας του *opening* που είναι η αφαίρεση θορύβου. Στο στάδιο της *διάβρωσης* της λειτουργίας του *opening* θα αφαιρεθούν τα απομονωμένα pixels στα όρια των αντικειμένων και στο στάδιο της *διαστολής* θα αποκατασταθούν τα περισσότερα από τα pixels χωρίς να επανέλθουν εικονοστοιχεία του θορύβου. Η διαδικασία φαίνεται να είναι επιτυχής στην αφαίρεση των μαύρων pixels θορύβου, αλλά δεν αφαιρεί τα άσπρα. Αυτό που παράγεται από το *opening* της εικόνα 2(a) είναι μια εικόνα απαλλαγμένη από θόρυβο της οποίας το αποτέλεσμα φαίνεται στην εικόνα 2(b)

Στο ίδιο παράδειγμα παρουσιάζεται μια εφαρμογή του *closing* για να αποκαταστήσει τα σπασίματα της μορφής. Κάνοντας *closing* την εικόνα 2.(b) αποκαθίστανται πολλά από τα σπασίματα, όχι όμως όλα. Αξίζει να τονίσουμε ότι στις πραγματικές εικόνες είναι σπάνιο η οποιαδήποτε τεχνική να παρέχει το τέλει αποτέλεσμα..

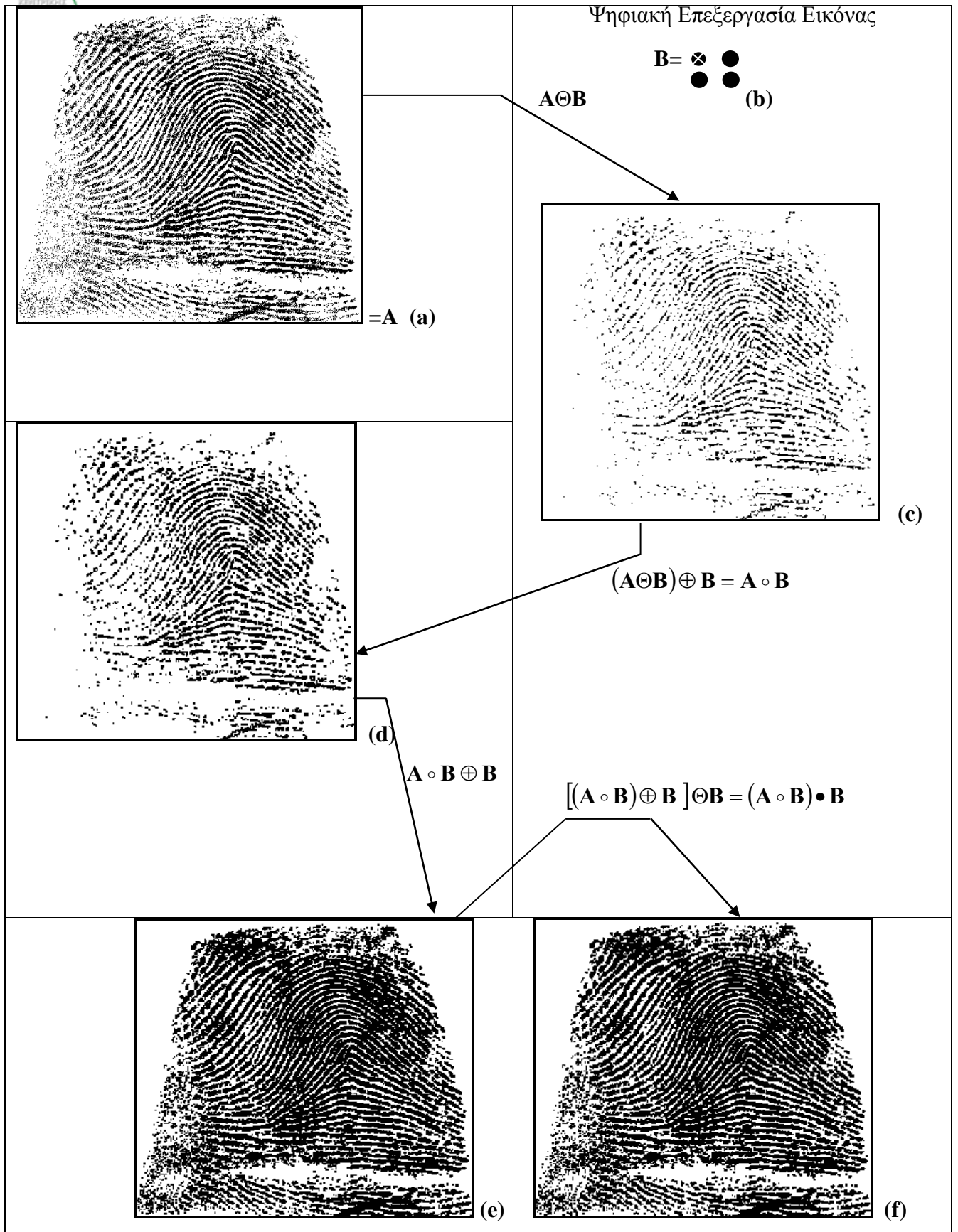


Παράδειγμα 2. (a) Εικόνα εισαγωγής (b) Opening στο A (c) Closing του Opening

Οι μορφολογικές λειτουργίες μπορούν να χρησιμοποιηθούν για να κατασκευάσουν φίλτρα παρόμοια με τα χωρικά φίλτρα. Η δυαδική εικόνα (a) του παραδείγματος 3 παρουσιάζει ένα τμήμα ενός δακτυλικού αποτυπώματος που αλλοιώνεται από θόρυβο μαύρων ή λευκών pixels. Η μορφή αυτή θορύβου ονομάζεται συχνά θόρυβος αλάτι-πιπέρι. Ο θόρυβος αυτός εμφανίζεται με σκοτεινά στοιχεία στο φωτεινό υπόβαθρο και και φωτεινά στοιχεία στα σκοτεινά μέρη του δακτυλικού αποτυπώματος. Ο στόχος είναι να αποβληθεί ο θόρυβος με όσο το δυνατόν λιγότερη παραμόρφωση. Ένα μορφολογικό φίλτρο που αποτελείται από ένα *opening* ακολουθούμενο από ένα *closing* μπορεί να χρησιμοποιηθεί για να επιτευχθεί αυτός ο στόχος.

Στην εικόνα 3(b) παρουσιάζεται το δομικό στοιχείο που χρησιμοποιούμε. Στη συνέχεια του παραδείγματος παρουσιάζονται βήμα προς βήμα τα στάδια του φιλτραρίσματος. Στην εικόνα 3(c) φαίνεται το αποτέλεσμα της *διάβρωσης* του A από το στοιχείο δόμησης. Ο θόρυβος πιπέρι αποβλήθηκε πλήρως στο στάδιο της *διάβρωσης* του *opening* επειδή σε αυτήν την περίπτωση όλα τα στοιχεία του θορύβου είναι μικρότερα από το στοιχείο δόμησης. Το μέγεθος των λευκών εικονοστοιχείων του θορύβου (αλάτι) που περιέχονται μέσα στο δακτυλικό αποτύπωμα συχνά αυξάνονται σε μέγεθος. Ο λόγος είναι ότι αυτά τα στοιχεία είναι στα εσωτερικά όρια των καμπυλών και μεγαλώνουν καθώς το αντικείμενο διαβρώνεται. Αυτή η μεγέθυνση αντιμετωπίζεται με την εκτέλεση μιας *διαστολής* στην εικόνα 3(c). Στην εικόνα 3(d) παρουσιάζεται το τελικό αποτέλεσμα. Τα τμήματα θορύβου που περιλήφθηκαν στο δακτυλικό αποτύπωμα μειώθηκαν στο μέγεθος ή διαγράφηκαν εντελώς.

Στην εικόνα 3(d) η καθαρή επίδραση του *opening* είναι να αποβάλει ουσιαστικά όλα τα τμήματα θορύβου και στο υπόβαθρο και στο ίδιο το δακτυλικό αποτύπωμα. Εντούτοις, νέα κενά μεταξύ των κορυφογραμμών του δακτυλικού αποτυπώματος δημιουργούνται. Για να αντιμετωπιστεί αυτή η ανεπιθύμητη επίδραση, εκτελούμε μια *διαστολή* πάνω στο *opening*, όπως φαίνεται στην εικόνα 3(e). Τα περισσότερα από τα σπασίματα αποκαταστάθηκαν, αλλά οι κορυφογραμμές πυκνώθηκαν. Η κατάσταση μπορεί να θεραπευθεί ικανοποιητικά από μια *διάβρωση*. Το αποτέλεσμα παρουσιάζεται στην εικόνα 3(f) που αποτελεί το *closing* του *opening* της εικόνα 3(d). Αυτό το τελικό αποτέλεσμα είναι εντυπωσιακά καθαρό από θορύβους, αλλά έχει το μειονέκτημα ότι μερικές από τις κορυφογραμμές δεν αποκαταστάθηκαν πλήρως διότι περιέχουν σπασίματα.



Παράδειγμα [3] (a) Εικόνα εισαγωγής (b) Δομικό στοιχείο (c) Διαβρωμένη εικόνα (d) Opening στο A (e) Διαστολή του Opening (f) Closing του Opening

ΕΡΓΑΣΤΗΡΙΑΚΕΣ ΑΣΚΗΣΕΙΣ ΣΕ MATLAB
ΠΛΗΡΟΦΟΡΙΕΣ ΣΧΕΤΙΚΑ ΜΕ ΤΟ ΑΡΧΕΙΟ ΓΡΑΦΙΚΩΝ .

Εντολή *imfinfo*(FILENAME , FMT) επιστρέφει μια δομή (πχ INFO) του οποίου

πεδία περιέχουν πληροφορίες σχετικά με μια εικόνα που περιέχεται σε ένα αρχείο

FILENAME: είναι μια συμβολοσειρά που καθορίζει το όνομα

του αρχείου **FMT:** είναι μια συμβολοσειρά που καθορίζει τη μορφή του αρχείου . Το αρχείο πρέπει να είναι στον τρέχοντα κατάλογο ή ένα κατάλογο στο κατάλογο MATLAB .

Αν η *imfinfo* δεν μπορεί να βρει το αρχείο με το όνομα **FILENAME** , ψάχνει για ένα αρχείο που ονομάζεται **FILENAME.FMT** .

Οι πιθανές τιμές της **FMT** περιέχονται σε μητρώο , το οποίο είναι προσβάσιμο μέσω της εντολής **IMFORMATS** .

Αν παραληφθεί η παράμετρος **FMT** η *imfinfo* επιχειρεί να συναγάγει τη μορφή του αρχείου από το περιεχόμενό του.

imfinfo(URL , ...) διαβάζει την εικόνα από μια διεύθυνση URL στο Internet

Η διεύθυνση URL πρέπει να περιλαμβάνει τον τύπο πρωτοκόλλου (π.χ. , " http://") .

Το σύνολο των πεδίων εξαρτάται από τον εκάστοτε φάκελο και την μορφή του . Ωστόσο, τα πρώτα εννέα πεδία είναι πάντα τα ίδια. Αυτά τα κοινά πεδία είναι :

Filename: συμβολοσειρά που περιέχει το όνομα του αρχείου

FileModDate: συμβολοσειρά που περιέχει την ημερομηνία τροποποίησης του αρχείου

FileSize: Ένας ακέραιος αριθμός που δείχνει το μέγεθος του αρχείου σε bytes

Format: ένα string που περιέχει τη μορφή του αρχείου (π.χ. , αρχεία JPEG και TIFF)

FormatVersion: Μια συμβολοσειρά ή αριθμός που καθορίζει τη μορφή του αρχείου

Width: ένας ακέραιος που υποδεικνύει το πλάτος της εικόνας σε pixels

Height: Ένας ακέραιος που υποδεικνύει το ύψος της εικόνας σε pixels

<p>BitDepth: Ένας ακέραιος που δείχνει τον αριθμό των bits ανά εικονοκύτταρο</p> <p>ColorType: Μια συμβολοσειρά που υποδεικνύει τον τύπο της εικόνας. Μπορεί να περιλαμβάνουν, αλλά δεν περιορίζεται σε 'TrueColor' για μια truecolor (RGB) εικόνα , 'GrayScale' , για μια εικόνα αποχρώσεων γκρι , ή 'indexed' για εικόνα που κωδικοποιεί το περιεχόμενο με δεικτοδότηση</p> <p>Εάν το όνομα αρχείου περιέχει ετικέτες Exif (μόνο για JPEG και TIFF) , τότε η INFO struct μπορεί επίσης να περιέχει 'DigitalCamera' ή 'GPSInfo'.</p> <p>Η τιμή του πεδίου 'DelayTime' αρχείων GIF μετριέται σε εκατοστά του δευτερολέπτου .</p>	
<pre>INFO = imfinfo('C:\Documents and Settings\Administrator\Desktop\DIPenviropmet\imagegray\lenag.bmp'); INFO</pre>	
<pre>Filename: [1x81 char] FileModDate: '12-Ιουν-2009 18:48:58' FileSize: 17462 Format: 'bmp' FormatVersion: 'Version 3 (Microsoft Windows 3.x)' Width: 128 Height: 128 BitDepth: 8 ColorType: 'indexed' FormatSignature: 'BM' NumColormapEntries: 256 Colormap: [256x3 double] RedMask: [] GreenMask: [] BlueMask: [] ImageDataOffset: 1078 BitmapHeaderSize: 40 NumPlanes: 1 CompressionType: 'none' BitmapSize: 16384 HorzResolution: 0 VertResolution: 0 NumColorsUsed: 256 NumImportantColors: 256</pre>	
<p>Η εντολή imread διαβάζει μια εικόνα από ένα αρχείο γραφικών . A = imread (FILENAME , FMT) Το αρχείο της εικόνας (αποχρώσεων του γκρι ή έγχρωμης) καθορίζεται από την συμβολοσειρά FILENAME. Εάν το αρχείο δεν βρίσκεται στον τρέχοντα κατάλογο ή σε έναν κατάλογο του MATLAB , πρέπει να καθορίζεται η πλήρης διαδρομή .</p>	<pre>I = imread('lenag.bmp');</pre>

<p>Η συμβολοσειρά FMT καθορίζει τη μορφή του αρχείου από το πρότυπο της επέκτασης του αρχείου . Για παράδειγμα « gif » για Graphics Interchange.</p> <p>Η τιμή επιστροφής A είναι ένας πίνακας που περιέχει τα δεδομένα της εικόνας . Εάν το αρχείο περιέχει μια ασπρόμαυρη εικόνα , το A είναι ένας πίνακας M x N. Αν το αρχείο περιέχει μια εικόνα TrueColor , το A είναι πίνακας M x N x 3. Για ένα αρχείο TIFF που περιέχει έγχρωμες εικόνες που χρησιμοποιούν το χρωματικό χώρο CMYK , το A είναι ένας πίνακας M x N x 4.</p>	
<p>ΜΟΡΦΕΣ ΑΡΧΕΙΩΝ</p> <p>BMP -- Windows Bitmap</p> <p>CUR -- Cursor File</p> <p>GIF -- Graphics Interchange Format</p> <p>HDF -- Hierarchical Data Format</p> <p>ICO -- Icon File See CUR.</p> <p>JPEG -- Joint Photographic Experts Group</p> <p>JPEG 2000 - Joint Photographic Experts Group 2000</p> <p>Note: Indexed JPEG 2000 images are not supported.</p> <p>PBM -- Portable Bitmap</p> <p>PCX -- Windows Paintbrush</p> <p>PGM -- Portable Graymap</p> <p>PNG -- Portable Network Graphics</p> <p>PPM -- Portable Pixmap</p> <p>RAS -- Sun Raster</p> <p>TIFF -- Tagged Image File Format</p>	

imshow(image, [low high]) : εμφάνιση εικόνας στο ενεργό παράθυρο γραφικών (ή σε νέο, αν δεν υπάρχει). Το πρώτο όρισμα είναι ο πίνακας που περιέχει τις τιμές φωτεινότητας. Το δεύτερο, αν υπάρχει, είναι ένα διάνυσμα δύο τιμών, όπου η πρώτη είναι η ελάχιστη τιμή φωτεινότητας (που απεικονίζεται ως μαύρο) και η δεύτερη είναι η μέγιστη (που απεικονίζεται ως λευκό).



[h, w]=size(A) : επιστρέφει τις διαστάσεις του πίνακα A που παίρνει ως όρισμα (όπου h το ύψος ή πλήθος γραμμών και w το πλάτος ή πλήθος στηλών).

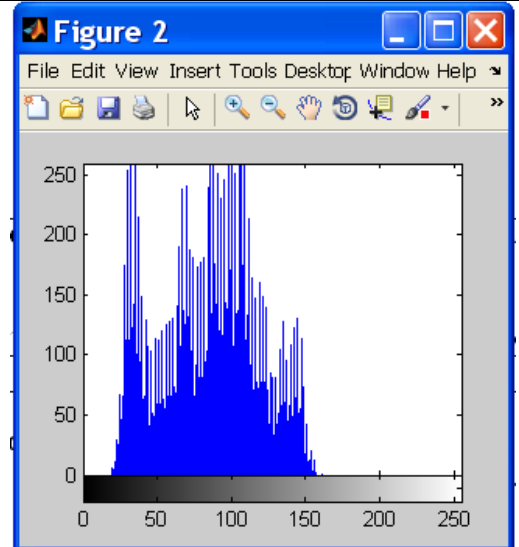
B=double(A) : μετατρέπει τα στοιχεία του πίνακα A σε κινητής υποδιαστολής διπλής ακρίβειας (κι ενδεχομένως τα αναθέτει σε νέο πίνακα).

Εμφάνιση ιστόγραμματος των δεδομένων εικόνας.

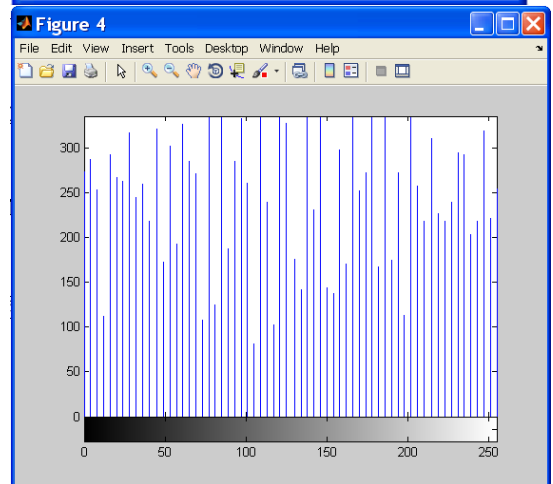
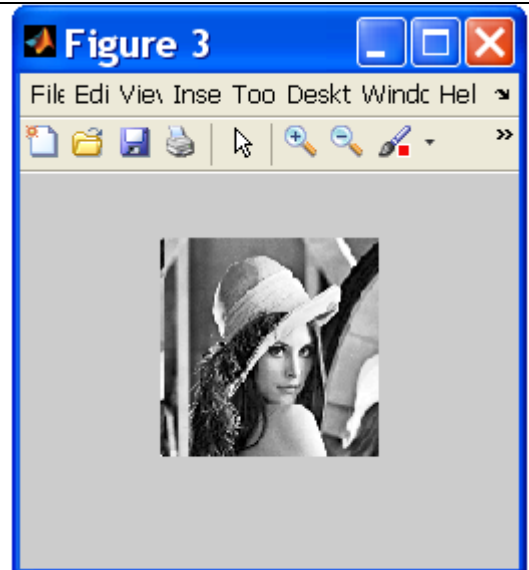
Η εντολή ***imhist*** (I) εμφανίζει ένα ιστόγραμμα των αποχρώσεων της εικόνας I. Οι κλάσεις προσδιορίζονται από τον τύπο της εικόνας. Αν η εικόνα είναι ασπρόμαυρη οι κλάσεις είναι 256 ως προεπιλεγμένη τιμή. Αν είναι μια δυαδική εικόνα οι κλάσεις είναι 2.

Το πλήθος των κλάσεων N μπορεί να ορισθεί ως παράμετρος,
imhist (I, N)

plot(Y): σχεδιάζει ένα διάγραμμα των τιμών του διανύσματος Y ως προς τους δείκτες του.



histeq(I).....
 $E_q = \text{histeq}(I)$
 $\text{imshow}(E_q)$
 $\text{imhist}(E_q)$



Description

$I = \text{ind2gray}(X, \text{map})$ converts the image X with colormap map to a grayscale image I . ind2gray removes the hue and saturation information from the input image while retaining the luminance.

fspecial

Δημιουργεί προκαθορισμένα φίλτρα δύο διαστάσεων

Σύνταξη

$h = \text{fspecial}(\text{τύπος})$

$h = \text{fspecial}(\text{τύπος}, \text{παράμετροι})$

Περιγραφή

$h = \text{fspecial}(\text{τύπος})$ δημιουργεί ένα

δισδιάστατο φίλτρο h του καθορισμένου τύπου. Η `fspecial` επιστρέφει τον πίνακα h ώστε αυτός να χρησιμοποιηθεί στην πράξη της συσχέτισης και να χρησιμοποιηθεί με την εντολή `imfilter` Ο *τύπος* είναι μια συμβολοσειρά που έχει μία από τις τιμές αυτές .

'average'

'disk'

'gaussian'

'laplacian'

'log'

'motion'

'prewitt'

'sobel'

'unsharp'

$h = \text{fspecial}$ (τύπος , παράμετροι) δέχεται το φίλτρο που καθορίζεται από τον τύπο και πρόσθετες παραμέτρους τροποποίησης αντίστοιχες με τον τύπο του φίλτρου που επιλέγεται. Εάν παραληφθούν οι παράμετροι, η `fspecial` χρησιμοποιεί προεπιλεγμένες τιμές.

Η λίστα που ακολουθεί δείχνει τη σύνταξη για κάθε τύπο φίλτρου.



$h = \text{fspecial}$ ('average' , $hsize$) επιστρέφει ένα φίλτρο μέσου όρου του μεγέθους $hsize$. Η προεπιλεγμένη τιμή για $hsize$ είναι [3 3] .


$h = \text{fspecial}$ ('disk' , $radius$) επιστρέφει ένα κυκλικό φίλτρο μέσου όρου (`pillbox`) στο τετράγωνο πλέγμα πλευράς $2 * radius + 1$. Η προεπιλεγμένη ακτίνα είναι 5 .

$h = \text{fspecial}$ ('gaussian', $hsize$, $sigma$) επιστρέφει ένα περιστροφικά συμμετρικό Gaussian φίλτρο μεγέθους $hsize$ με τυπική απόκλιση $sigma$. Η προεπιλεγμένη τιμή για $hsize$ είναι [3 3] ? Η προεπιλεγμένη τιμή για το σίγμα είναι 0,5.

<p>h = fspecial ('laplacian' , alpha) επιστρέφει ένα 3 X 3 φίλτρο που προσεγγίζει τον Λαπλασιανό τελεστή. Η παράμετρος alpha ελέγχει τη μορφή του τελεστή και πρέπει να είναι μεταξύ 0.0 και 1.0. Η προεπιλεγμένη τιμή για της είναι 0.2.</p> <p>h = fspecial ('log', hsize , sigma) επιστρέφει ένα περιστροφικά συμμετρικό πίνακα του Laplacian of Gaussian φίλτρου, μεγέθους hsize με τυπική απόκλιση sigma. Η προεπιλεγμένη τιμή για την hsize είναι [5 5] και για την sigma 0.5.</p> <p>h = fspecial ('motion', len , theta) επιστρέφει ένα φίλτρο για την προσέγγιση της γραμμική κίνηση μιας κάμερας με γωνία theta βαθμών (αριστερόστροφα) . Η προεπιλεγμένη τιμή της len είναι 9 pixels και η προεπιλεγμένη τιμή της theta είναι 0 , η οποία αντιστοιχεί σε μια οριζόντια κίνηση εννέα pixels .</p> <p>h = fspecial ('prewitt') επιστρέφει το 3 X 3 φίλτρο [1 1 1 0 0 0 -1 -1 -1]</p> <p>h = fspecial ('sobel') επιστρέφει το 3X3 φίλτρο [1 2 1 0 0 0 -1 -2 -1]</p> <p>h = fspecial ('unsharp', alpha) επιστρέφει 3 X 3 φίλτρο. Η fspecial δημιουργεί το αρνητικό φίλτρο Laplacian φίλτρου με παράμετρο alpha με τιμές από 0 έως 1. Η προεπιλεγμένη τιμή είναι 0.2.</p>	
<p>imfilter</p> <p>B = imfilter (A , H) B = imfilter (A, H , option1 , option2 , ...)</p>	


<p>περιγραφή</p> <p>$B = \text{imfilter}(A, H)$ φιλτράρει το πολυδιάστατο πίνακα A με τον πολυδιάστατο H. Το αποτέλεσμα B έχει το ίδιο μέγεθος τον A.</p> <p>Κάθε στοιχείο της εξόδου B υπολογίζεται χρησιμοποιώντας διπλής ακρίβειας κινητή υποδιαστολή. Αν ο A είναι ένας ακέραιος ή λογική σειρά, στη συνέχεια, τα στοιχεία εξόδου που υπερβαίνουν το εύρος του τύπου <code>integer</code> περικόπτονται, και οι κλασματικές τιμές στρογγυλοποιούνται.</p> <p>Επιλογές για τις τιμές έξω από τα όρια του πίνακα A</p> <p>'symmetric' Οι τιμές συμπληρώνονται κατοπτρικά.</p> <p>'replicate' Κάθε τιμή προκύπτει από την γειτονική της.</p> <p>'circular' Οι τιμές συμπληρώνονται θεωρώντας ότι ο πίνακας είναι περιοδικός</p> <p>Αν δεν ορισθεί οι παράμετρος οι τιμές έξω από όρια του A θεωρούνται μηδενικές.</p> <p>Παράμετροι για τις διαστάσεις του πίνακα εξόδου</p> <p>'same' Ο πίνακας εξόδου έχει το ίδιο μέγεθος με τον A. Το ίδιο συμβαίνει αν δεν δοθεί η παράμετρος.</p> <p>'full' Ο πίνακας εξόδου έχει το μέγεθος της συνέλιξης A με τον h.</p> <p>Επιλογές Συσχέτισης και Συνέλιξης</p> <p>'corr'</p>	
--	--


'conv'	
Imnoise ...	
	
<pre>h = fspecial('average') imshow(imfilter(J,h))</pre>	<pre>h = 0.1111 0.1111 0.1111 0.1111 0.1111 0.1111 0.1111 0.1111 0.1111</pre> 
<pre>h = fspecial('gaussian',3,0.7) imshow(imfilter(J,h))</pre>	<pre>h=[0.0439 0.1217 0.0439 0.1217 0.3377 0.1217 0.0439 0.1217 0.0439]</pre>

	
<p>medfilt2</p> <p>2-D median filtering Syntax</p> <p>B = medfilt2(A, [m n]) B = medfilt2(A) B = medfilt2(A, 'indexed', ...) B = medfilt2(..., padopt)</p> <p>Περιγραφή</p> <p>Η εντολή εφαρμόζει στην εικόνα την μη γραμμική επεξεργασία της ενδιάμεσης τιμής και είναι πιο αποτελεσματική για την αφαίρεση κρουστικού θορύβου (αλάτι-πιπέρι) σε σχέση με την γραμμική ως προς την διατήρηση των ακμών.</p> <p>B = medfilt2(A, [m n]) εφαρμόζει το φίλτρο της ενδιάμεσης τιμής στον πίνακα A σε δύο διαστάσεις. Κάθε pixel της εξαγόμενης εικόνας περιέχει την ενδιάμεση τιμή σε μια m X n γειτονιά γύρω από το αντίστοιχο φίλτρο της αρχικής. Η medfilt2 συμπληρώνει με μηδενικά τις θέσεις εκτός της αρχικής εικόνας.</p> <p>B = medfilt2(A), η γειτονιά είναι 3 X 3.</p>	

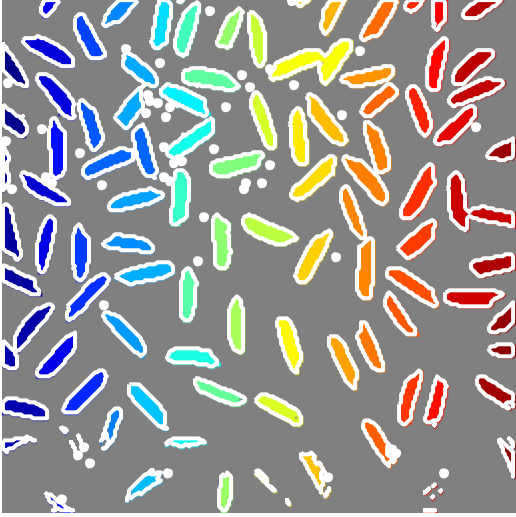

<pre>imshow (medfilt2 (J))</pre>	
<pre>h1 = fspecial('prewitt') E1 = imfilter(I,h1); h2 = h1' E2 = imfilter(I,h2); E = double(E1.^2+E2.^2); E = E.^0.5 T = mean(mean(E)) Edges = E>2*T imshow(Edges)</pre>	<pre>h1 = 1 1 1 0 0 0 -1 -1 -1 h2 = 1 0 -1 1 0 -1 1 0 -1</pre> 
<p>edge</p> <p>Προσδιορίζει ακμές σε μια εικόνα αποχρώσεων του γκρι</p> <p>Σύνταξη</p> <pre>BW = edge (I)</pre> <pre>BW = edge (I, 'sobel')</pre> <pre>BW = edge (I, 'sobel', thresh)</pre> <pre>BW = edge (I, 'sobel', thresh, direction)</pre> <pre>[BW, thresh] = edge (I, 'sobel', ...)</pre> <pre>BW = edge (I, 'prewitt')</pre> <pre>BW = edge (I, 'prewitt', thresh)</pre> <pre>BW =</pre>	

<pre> edge(I, 'prewitt', thresh, direction) [BW, thresh] = edge(I, 'prewitt', ...) BW = edge(I, 'roberts') BW = edge(I, 'roberts', thresh) [BW, thresh] = edge(I, 'roberts', ...) BW = edge(I, 'log') BW = edge(I, 'log', thresh) BW = edge(I, 'log', thresh, sigma) [BW, threshold] = edge(I, 'log', ...) BW = edge(I, 'zerocross', thresh, h) [BW, thresh] = edge(I, 'zerocross', ...) BW = edge(I, 'canny') BW = edge(I, 'canny', thresh) BW = edge(I, 'canny', thresh, sigma) [BW, threshold] = edge(I, 'canny', ...) </pre>	
--	--

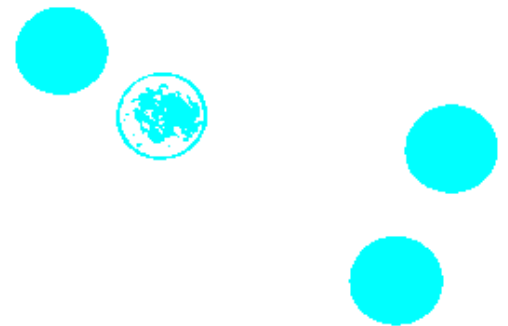
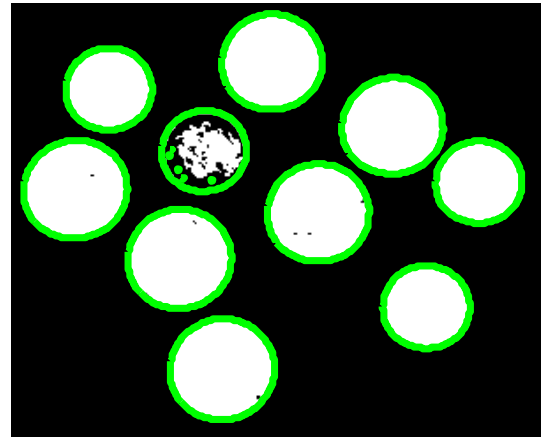
<p>im2bw</p> <p>Για την μετατροπή εικόνας σε δυαδική εικόνα με καθορισμό του κατώφλιου</p> <p><code>BW = im2bw(I, level)</code></p> <p>περιγραφή</p> <p><code>BW = im2bw (I, threshold)</code> μετατρέπει μια εικόνα αποχρώσεων του γκρι I σε μία δυαδική εικόνα. Στην εικόνα εξόδου BW όλα τα εικονοστοιχεία της εικόνας εισόδου με φωτεινότητα μεγαλύτερη από threshold έχουν την τιμή 1 (λευκό) και όλα τα άλλα pixels την τιμή 0 (μαύρο). Καθορίστε το κατώφλι (threshold) στην κλίμακα [0,1] Για να υπολογίσουμε την τιμή κατώφλιου, μπορείτε να χρησιμοποιήσετε την εντολή <code>graythresh</code>. Εάν δεν καθορίσετε το επίπεδο, <code>im2bw</code> χρησιμοποιεί την τιμή 0,5.</p>	
---	---

<p>graythresh</p> <p>Υπολογίζει την τιμή κατωφλίσωσης για την δυαδικοποίηση μιας εικόνας αποχρώσεων του γκρι χρησιμοποιώντας την μεθοδο του Otsu</p> <p>Σύνταξη</p> <p>$T = \text{graythresh}(I)$</p>	 <p>T = 80</p>
<p>Bwlabel</p> <p>$L = \text{bwlabel}(BW, n)$</p> <p>επιστρέφει ένα μητρώο L, του ίδιου μεγέθους όπως το BW, που περιέχει ετικέτες για τα συνδεδεμένα αντικείμενα (connected components) του BW. Η μεταβλητή n μπορεί να έχει μία τιμή είτε 4 είτε 8 και ορίζει την συνδετικότητα. Αν το όρισμα n παραλειφθεί, η τυπική τιμή είναι 8.</p> <p>Τα στοιχεία της L είναι τιμές ακέραιες μεγαλύτερες ή ίσες με μηδέν. Τα εικονοστοιχεία με τιμή 0 είναι το φόντο. Τα εικονοστοιχεία με τιμή 1 συνθέτουν ένα αντικείμενο, τα εικονοστοιχεία με την ένδειξη 2 συνθέτουν ένα δεύτερο αντικείμενο και ούτω καθεξής.</p> <p>$[L, \text{num}] = \text{bwlabel}(BW, n)$ επιστρέφει στο num το πλήθος των συνδεδεμένων αντικειμένων που βρέθηκαν στο BW.</p>	
<p>Bwarea</p> <p>Εμβαδό της περιοχής των αντικειμένων σε δυαδική εικόνα</p> <p>σύνταξη</p>	

<p>total = bwarea (BW)</p> <p>περιγραφή</p> <p>total = bwarea (BW) εκτιμά το εμβαδό των αντικειμένων σε μια δυαδική BW εικόνα, αντιστοιχεί χονδρικά στο συνολικό αριθμό των pixels των αντικειμένων της εικόνας, αλλά ενδέχεται να μην είναι ακριβώς το ίδιο, επειδή διαφορετικά πρότυπα pixels σταθμίζονται διαφορετικά.</p> <p>Η BW μπορεί να είναι μητρώο με αριθμητικές ή λογικές τιμές. Μη μηδενικές τιμές εικονοστοιχείων θεωρούνται ότι ανήκουν σε αντικείμενα.</p> <p>Αλγόριθμος</p> <p>Η bwarea υπολογίζει το εμβαδόν όλων των pixels σε μια εικόνα αθροίζοντας τις περιοχές του κάθε pixel στην εικόνα. Η επιφάνεια ενός μεμονωμένου pixel προσδιορίζεται με την εξέταση της 2 X 2 γειτονιάς του. Υπάρχουν έξι διαφορετικά σχέδια, το καθένα αντιπροσωπεύει μια διαφορετική περιοχή :</p> <p>Μοτίβα με μηδέν pixels (περιοχή = 0)</p> <p>Μοτίβα με ένα εικονοστοιχείο (περιοχή = 1/4)</p> <p>Μοτίβα με δύο παρακείμενα pixels (περιοχή = 1/2)</p> <p>Μοτίβα με δύο διαγώνια pixels (περιοχή = 3/4)</p> <p>Μοτίβα με τρία pixels (περιοχή = 7/8)</p> <p>Μοτίβα με τέσσερα σε pixels (περιοχή = 1).</p>	
<p>Bwboundaries</p> <p>σύνταξη</p> <p>B = bwboundaries (BW) B = bwboundaries (BW , conn)</p>	<pre>% ΠΑΡΑΔΕΙΓΜΑ I = imread('rice.png'); BW = im2bw(I, graythresh(I)); [B,L] = bwboundaries(BW, 'noholes'); imshow(label2rgb(L, @jet, [.5 .5 .5])) hold on for k = 1:length(B) boundary = B{k}; plot(boundary(:,2), boundary(:,1), 'w',</pre>

<p><code>B = bwboundaries (BW , conn , επιλογές)</code> <code>[B , L] = bwboundaries (...)</code> <code>[B , L , N , A] = bwboundaries (...)</code></p> <p>περιγραφή</p> <p><code>B = bwboundaries (BW)</code> ανιχνεύει τα εξωτερικά όρια των αντικειμένων, καθώς επίσης και τα όρια των οπών μέσα σε αυτά τα αντικείμενα, της δυαδικής εικόνας <code>BW</code>. Η <code>BW</code> πρέπει να είναι μια δυαδική εικόνα όπου τα pixels με μηδενικές τιμές ανήκουν σε κάποιο αντικείμενο και τα pixels με μηδενικές τιμές ανήκουν στο υπόβαθρο .</p> <p><code>B = bwboundaries</code> επιστρέφει τη δομή <code>B</code> που είναι μια <code>PX1</code> συστοιχία κελιών , όπου <code>P</code> είναι ο αριθμός των αντικειμένων και οπών. Κάθε κελί στη συστοιχία περιέχει ένα <code>QX2</code> μητρώο. Κάθε σειρά στο μητρώο περιέχει τις σειρές και στήλες που είναι οι συντεταγμένες ενός οριακού pixel. <code>Q</code> είναι ο αριθμός των pixels ορίου για την αντίστοιχη περιοχή .</p>	<pre>'LineWidth', 2) end</pre> 
<p>Εφαρμογή</p> <pre>close all; I=imread('coins.png'); imshow(I) BW = im2bw(I,graythresh(I)); figure;imshow(BW) hold on [LA, num] = bwlabel(BW); objboundaries=bwboundaries(BW); for k = 1:num b = objboundaries{k}; plot(b(:,2),b(:,1),'g','LineWidth',3); end figure; for k = 1:num area(k)=bwarea(LA==k); end areathr = 1.5*mean(area) x=find(area<areathr); [row col]=size(x); R=zeros(size(LA)); for k=1:col R = R+(LA==x(k)); end</pre>	

```
RGB = label2rgb(R);  
imshow(RGB)
```



`bwmorph`

Morphological operations on binary images
Syntax

```
BW2 = bwmorph(BW,operation)  
BW2 = bwmorph(BW,operation,n)
```

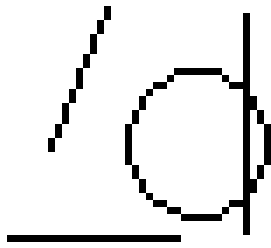
Description

`BW2 = bwmorph(BW,operation)` applies a specific morphological operation to the binary image `BW`.

<p>BW2 = bwmorph(BW,operation,n) applies the operation n times. n can be Inf, in which case the operation is repeated until the image no longer changes.</p> <p>operation is a string that can have one of the values listed below.</p>	
---	--

ΕΝΤΟΠΙΣΜΟΣ ΕΥΘΥΓΡΑΜΜΩΝ ΤΜΗΜΑΤΩΝ, Ο ΑΠΛΟΣ ΜΕΤΑΣΧΗΜΑΤΙΣΜΟΣ ΤΟΥ HOUGH	
<p>Hough transform Σύνταξη [H, theta, rho] = hough(BW) [H, theta, rho] = hough(BW, ParameterName, ParameterValue)</p> <p>Περιγραφή</p> <p>[H, theta, rho] = hough(BW) υπολογίζεται ο τυπικός μετασχηματισμός του Hough (HT) της δυαδικής εικόνας BW. Χρησιμοποιείται για την ανίχνευση ευθυγράμμων τμημάτων μιας εικόνας. Η συνάρτηση επιστρέφει στο πίνακα H, τον πίνακα του HT (πεδίο (ρ,θ)). Η Παράμετρος theta σε μοίρες και η rho χρησιμοποιούνται ως ορίσματα.</p> <p>[H, theta, rho] = hough(BW, ParameterName, ParameterValue). Η παράμετρος 'RhoResolution' προσδιορίζει το βήμα κβαντισμού του άξονα των rho. Τυπική τιμή είναι το 1.</p> <p>Η παράμετρος 'Theta', προσδιορίζει τα εύρος και το βήμα της γωνίας θ, π.χ. -90:1:89, δηλαδή από -90 έως 89 μοίρες με βήμα 1 μοίρα</p>	
<p>houghpeaks</p> <p>Βρίσκει μεγάλες τιμές – κορυφές (peaks) στον πίνακα του μετασχηματισμού Hough Συντακτικό</p>	

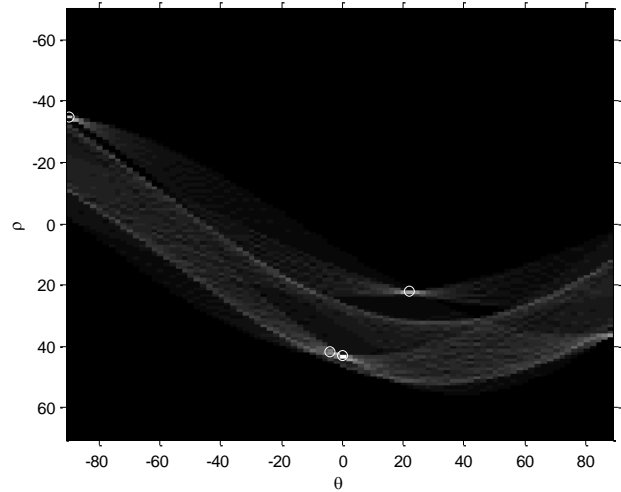
<p>peaks = houghpeaks(H, numpeaks) peaks = houghpeaks(..., param1, val1, param2, val2)</p> <p>Περιγραφή peaks = houghpeaks(H, numpeaks). Η τιμή της παραμέτρου numpeaks το μέγιστο πλήθος των κορυφών που θα εντοπισθούν. Η εξ ορισμού τιμή είναι 1.</p> <p>Η συνάρτηση επιστρέφει τις κορυφές σε ένα πίνακα $Q \times 2$ όπου Q το πλήθος των κορυφών που εντοπίστηκαν Q είναι μικρότερο του numpeaks. Ο πίνακας περιέχει τις γραμμές και τις στήλες των κορυφών.</p> <p>peaks = houghpeaks(..., param1, val1, param2, val2) specifies parameter/value pairs, listed in the following table. Parameter names can be abbreviated, and case does not matter. Με την παράμετρο 'Threshold' μπορούμε επιπρόσθετα να προσδιορίσουμε ελάχιστη τιμή ώστε ένα στοιχείο του πίνακα του HT να θεωρηθεί κορυφή</p>	
<p>houghlines</p> <p>Εξάγει τα ευθύγραμμα τμήματα μιας εικόνας βασισμένοι στα αποτελέσματα που προέκυψαν από την εφαρμογή του μετασχηματισμού του Hough</p> <p>Σύνταξη</p> <p>lines = houghlines(BW, theta, rho, peaks) lines = houghlines(..., param1, val1, param2, val2)</p> <p>Περιγραφή BW η δυαδική εικόνα theta και rho τα ανύσματα (μονοδιάστατοι πίνακες) που επιστράφηκαν από την συνάρτηση hough. peaks και είναι οι συντεταγμένες γραμμής και στήλης των</p>	

<p>‘κελιών’ του HT και θα χρησιμοποιηθούν για την ανίχνευση των ευθυγράμμων τμημάτων.</p> <p>Η lines είναι πίνακας στοιχείων δομής που είναι τα point1, point2, theta, rho, όπου point1, point2 οι συντεταγμένες των άκρων του ευθυγράμμου τμήματος που εντοπίστηκε σε απόσταση rho και γωνία theta.</p> <p>Με την παράμετρο 'FillGap' ορίζεται η μέγιστη απόσταση δύο συνευθειακών ευθυγράμμων τμημάτων (με το ίδιο ρ και θ) ώστε αυτά να συνενωθούν σε ένα.</p> <p>Με την παράμετρο 'MinLength' ορίζεται περιορισμός ελάχιστου μήκους για τα ευθύγραμμα τμήματα</p>	
<p>Παράδειγμα</p>	
<pre>clear, close all RGB = imread('C:\Documents and Settings\Administrator\My Documents\ΨΕΕ\NewLabMaterial\hought est1.bmp', 'bmp'); % Display the original image. subplot(3,1,1); imshow(RGB); % Convert to intensity. I = rgb2gray(RGB); BW = im2bw(I)-1; [H,T,R] = hough(BW, 'RhoResolution',1, 'Theta', -90:2:88); % Display the Hough matrix. subplot(3,1,2); %imshow(H, [], 'XData', T, 'YData', R, 'f it'); imshow(H, [], 'XData', T, 'YData', R, 'In itialMagnification', 'fit'); xlabel('\theta'), ylabel('\rho'); axis on, axis normal, hold on; P = houghpeaks(H,4, 'threshold',ceil(0.5 *max(H(:))))); plot(T(P(:,2)),R(P(:,1)), 'o', 'color ', 'white'); T(P(:,2)) R(P(:,1)) lines = houghlines(BW, T, R, P, 'FillGap',5, 'Mi</pre>	<div style="text-align: center;">  <p>Η αρχική εικόνα</p> </div>

```

nLength', 7);
BG = zeros(size(BW));
subplot(3,1,3);imshow(BW);hold on
for k = 1:length(lines)
    xy = [lines(k).point1;
        lines(k).point2];
plot(xy(:,1),xy(:,2), 'LineWidth', 2,
'Color', 'green');
end

```



Το πεδίο (ρ, θ)



Η εικόνα των ευθυγράμμων τμημάτων

Εφαρμογή - Διόρθωση κλίσης εγγράφου


```

clear, close all
RGB =
imread('imrotated.bmp','bmp');
imshow(RGB);

% Convert to intensity.
I = rgb2gray(RGB);

BW = im2bw(I)-1;
[H,T,R] =
hough(BW,'RhoResolution',1,'Theta',
-90:2:88);

% Display the Hough matrix.

imshow(H,[],'XData',T,'YData',R,'f
it');

figure;
imshow(H,[],'XData',T,'YData',R,'In
itialMagnification','fit');
xlabel('\theta'), ylabel('\rho');
axis on, axis normal, hold on;

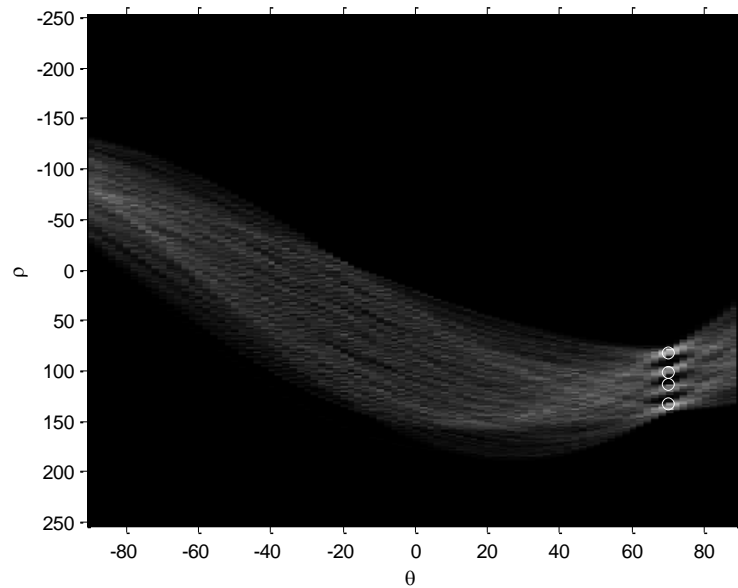
P =
houghpeaks(H,4,'threshold',ceil(0.5
*max(H(:)))));
plot(T(P(:,2)),R(P(:,1)),'o','color
','white');
T(P(:,2))
R(P(:,1))

lines =
houghlines(BW,T,R,P,'FillGap',5,'Mi
nLength',7);
BG = zeros(size(BW));
figure;imshow(BW);hold on
for k = 1:length(lines)
    xy = [lines(k).point1;
lines(k).point2];
plot(xy(:,1),xy(:,2),'LineWidth',2,
'Color','green');
end
thindex = max(P(:,2));
%T(max(P(:,2)))
figure; imshow(
imrotate(RGB,T(thindex)-90) );

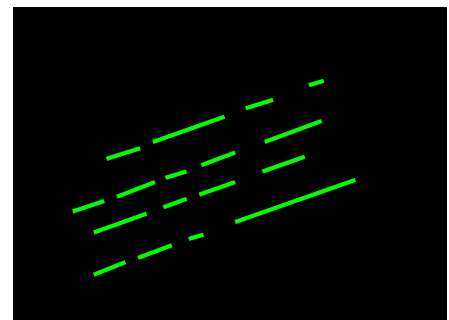
```

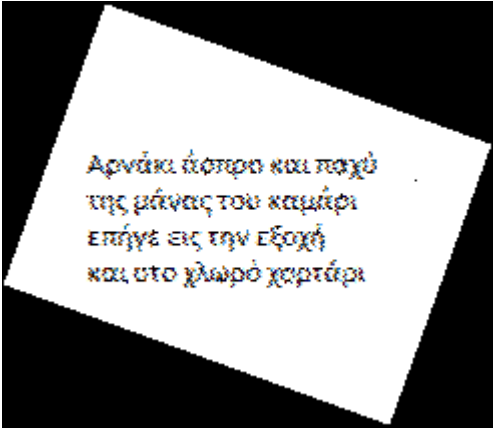
Αρνάκι άσπρο και παχύ
της μάνας του καμάρι
επήγε εις την εξοχή
και στο χλωρό χαρτάρι

Η εικόνα του κεκλιμένου εγγράφου



Το πεδίο (ρ, θ) του μετασχηματισμού Hough



	
<p>imrotate Rotate image.</p> <p>Σύνταξη</p> <pre>B = imrotate(A, ANGLE, METHOD)</pre> <p>Περιστρέφει την εικόνα του πίνακα A κατά ANGLE</p> <p>Μοίρες ανθρωλογιακά χρησιμοποιώντας μέθοδο παρεμβολής που καθορίζεται από την παράμετρο METHOD που παίρνει τις ακόλουθες τιμές</p> <pre>'nearest'</pre> <pre>'bilinear'</pre> <pre>'bicubic'</pre> <p>με τυπική τιμή την 'nearest'. Για να περιστρέψουμε την εικόνα ωρολογιακά καθορίζουμε μια αρνητική γωνία.</p>	<p>Το έγγραφο μετά την διόρθωση της κλίσης</p>

ΠΑΡΑΡΤΗΜΑ Α

$$\begin{aligned}
 \sigma_B^2 &= \rho_1(\mu_1 - \rho_1\mu_1 - \rho_2\mu_2)^2 + \rho_2(\mu_2 - \rho_1\mu_1 - \rho_2\mu_2)^2 = \rho_1(\mu_1(1 - \rho_1) - \rho_2\mu_2)^2 + \rho_2(\mu_2(1 - \rho_2) - \rho_1\mu_1)^2 = \\
 &= \rho_1(\mu_1\rho_2 - \rho_2\mu_2)^2 + \rho_2(\mu_2\rho_1 - \rho_1\mu_1)^2 = \rho_1\rho_2^2(\mu_1 - \mu_2)^2 + \rho_2\rho_1^2(\mu_2 - \mu_1)^2 = \rho_1\rho_2(\mu_1 - \mu_2)^2(\rho_1 + \rho_2) = \\
 &= \rho_1\rho_2(\mu_1 - \mu_2)
 \end{aligned}$$

$$\begin{aligned}
 \sigma^2 &= \sum_{g=0}^{G-1} (g - \mu)^2 H(g) = \sum_{g=0}^{G-1} (g^2 + \mu^2 - 2g\mu)H(g) = \sum_{g=0}^{G-1} g^2 H(g) + \mu^2 \sum_{g=0}^{G-1} H(g) - 2\mu \sum_{g=0}^{G-1} H(g)g = \\
 &= \sum_{g=0}^{G-1} g^2 H(g) + \mu^2 - 2\mu\mu = \sum_{g=0}^{G-1} g^2 H(g) - \mu^2
 \end{aligned}$$

$$\begin{aligned}
 \sigma_\omega^2 + \sigma_b^2 &= \rho_1\sigma_1^2 + \rho_2\sigma_2^2 + \rho_1(\mu_1 - \mu)^2 + \rho_2(\mu_2 - \mu)^2 = \\
 &= \frac{\rho_1 \sum_{g=0}^{T-1} (g - \mu_1)^2 H(g)}{\rho_1} + \frac{\rho_2 \sum_{g=T}^{G-1} (g - \mu_2)^2 H(g)}{\rho_2} + \rho_1(\mu_1^2 + \mu^2 - 2\mu\mu_1) + \rho_2(\mu_2^2 + \mu^2 - 2\mu\mu_2) = \\
 &= \sum_{g=0}^{T-1} (g^2 + \mu_1^2 - 2g\mu_1)H(g) + \sum_{g=T}^{G-1} (g^2 + \mu_2^2 - 2g\mu_2)H(g) + \rho_1\mu_1^2 + \rho_1\mu^2 - 2\mu\mu_1\rho_1 + \rho_2\mu_2^2 + \rho_2\mu^2 + 2\mu\mu_2\rho_2 = \\
 &= \sum_{g=0}^{T-1} g^2 H(g) + \sum_{g=0}^{T-1} \mu_1^2 H(g) - \sum_{g=0}^{T-1} 2g\mu_1 H(g) + \sum_{g=T}^{G-1} g^2 H(g) + \sum_{g=T}^{G-1} \mu_2^2 H(g) - \sum_{g=T}^{G-1} 2g\mu_2 H(g) + \\
 &\quad + \rho_1\mu_1^2 + \rho_2\mu_2^2 + \mu^2(\rho_1 + \rho_2) - 2\mu(\mu_1\rho_1 + \mu_2\rho_2) = \\
 &= \sum_{g=0}^{G-1} g^2 H(g) + \mu_1^2 \sum_{g=0}^{T-1} H(g) - 2\mu_1 \sum_{g=0}^{T-1} gH(g) + \mu_2^2 \sum_{g=T}^{G-1} H(g) - 2\mu_2 \sum_{g=T}^{G-1} gH(g) + \rho_1\mu_1^2 + \rho_2\mu_2^2 + \mu^2 - 2\mu^2 = \\
 &= \sum_{g=0}^{G-1} g^2 H(g) + \mu_1^2\rho_1 - 2\mu_1^2\rho_1 + \mu_2^2\rho_2 - 2\mu_2^2\rho_2 + \rho_1\mu_1^2 + \rho_2\mu_2^2 - \mu^2 = \sum_{g=0}^{G-1} g^2 H(g) - \mu^2 = \sigma^2
 \end{aligned}$$

ΠΑΡΑΡΤΗΜΑ Β

Για να συμβολίσουμε ένα σύνολο στα Μαθηματικά, χρησιμοποιούμε ένα από τα κεφαλαία γράμματα του Ελληνικού ή του Λατινικού αλφάβητου, ενώ για τα στοιχεία του χρησιμοποιούμε τα μικρά γράμματα αυτών.

Για να δηλώσουμε ότι το x είναι στοιχείο του συνόλου A , γράφουμε

$$x \in A$$

Και διαβάζουμε «το x ανήκει στο A », ενώ για να δηλώσουμε ότι το x δεν είναι στοιχείο του συνόλου A γράφουμε

$$x \notin A$$

Και διαβάζουμε «το x δεν ανήκει στο A »

Για να παραστήσουμε ένα σύνολο χρησιμοποιούμε συνήθως δύο τρόπους:

α) Παράσταση με αναγραφή των στοιχείων του.

$$\text{Π.χ. } A = \{2, 4, 6\} \quad B = \{1, 2, 3, \dots, 100\} \quad \Gamma = \{1, 1/2, 1/3, 1/4, \dots\}$$

β) Παράσταση με περιγραφή των στοιχείων του.

π.χ. $\{x \in \mathbb{R} / x > 0\}$ δηλαδή το σύνολο των πραγματικών αριθμών που έχουν την ιδιότητα να είναι θετικοί.

Γενικά, αν από ένα σύνολο Ω επιλέξουμε εκείνα τα στοιχεία του, που έχουν μια ορισμένη ιδιότητα I , φτιάχνουμε ένα νέο σύνολο που συμβολίζεται :

$$\{x \in \Omega / x \text{ έχει την ιδιότητα } I\}$$

και διαβάζεται το σύνολο των $x \in \Omega$, όπου x έχει την ιδιότητα I .

Ένα σύνολο A λέγεται υποσύνολο ενός συνόλου B όταν κάθε στοιχείο του A είναι και στοιχείο του B και συμβολίζεται: $A \subseteq B$.

ΒΙΒΛΙΟΓΡΑΦΙΑ

Algorithms for image processing and computer vision J.R.Parker ISBN:0-471-14056-2
Εκδόσεις : Wiley Computer Publishing

Digital Image Processing Rafael C. Gonzalez & Richard E. Woods ISBN: 0-13-168728-X
Εκδόσεις:Pearson Education. Inc

Ψηφιακή Επεξεργασία & Ανάλυση Εικόνας Ν. Παπαμάρκος Καθηγητής Δ.Π.Θ. ISBN
960-387-352-7 Εκδόσεις: Β.Γκιούρδας Εκδοτική.