

Τεχνητή Όραση

ΤΠ 7004

Μάθημα 4^ο: Ανίχνευση Ακμών

Δρ. Θάνος Δεμίρης

Περιεχόμενα 4^{ου} Μαθήματος

- Επανάληψη βασικών σημείων 3^{ου} μαθήματος
- Ανίχνευση ακμών
 - Μαθηματικός ορισμός
 - Τύποι ακμών
 - Παραγωγή
 - Βαθμίδα (Gradient)
 - Roberts
 - Prewitt
 - Sobel
 - Canny
 - LoG

Μορφολογικά φίλτρα εικόνων

- Κύλιση ενός «παραθύρου» επάνω από κάθε σημείο της εικόνας
- Υπολογισμός της νέας τιμής ενός pixel όχι μόνο σε συνάρτηση του αρχικού (σημειακοί μετασχηματισμοί), αλλά μίας γειτονιάς
- Το «πανάθυρο» αποτελεί ένα είδος «γεωμετρικού κανόνα» για τον υπολογισμό της νέας τιμής
- Το σχήμα και το μέγεθος των διαφόρων παραθύρων διαφέρουν

Τυπικά παράθυρα

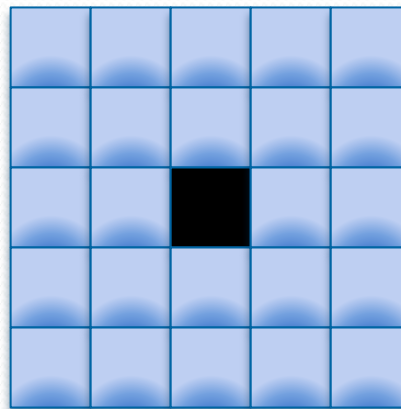
ΓΡΑΜΜΗ



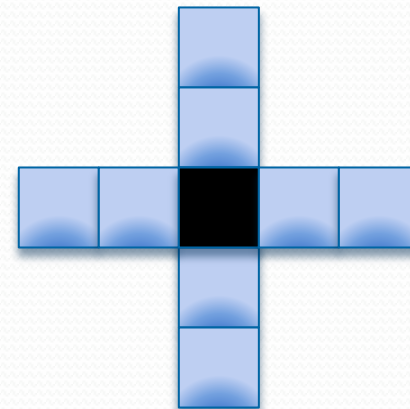
ΣΤΗΛΗ



ΤΕΤΡΑΓΩΝΟ

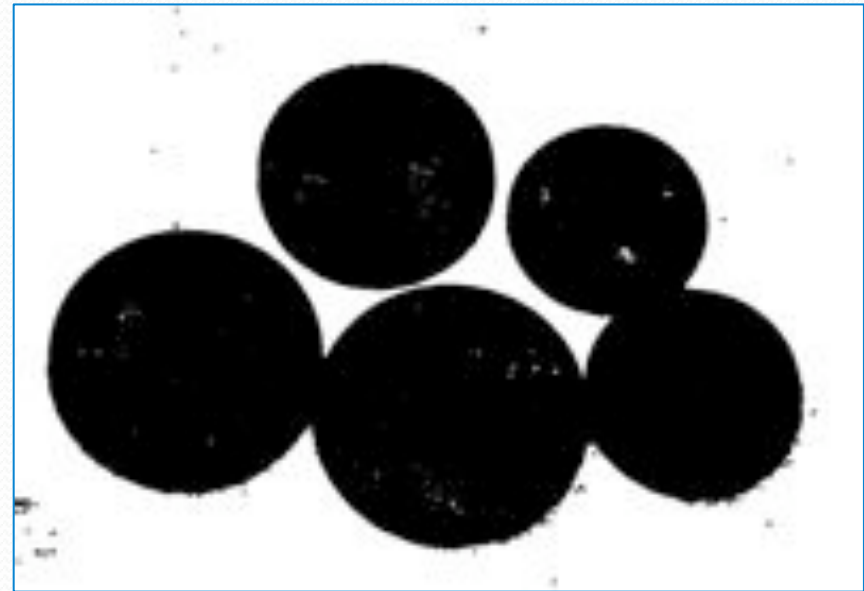


ΣΤΑΥΡΟΣ



$$\mathbf{B}=\text{ROW}(5): \mathbf{B}f(n_1, n_2)=\{f(n_1, n_2-2), f(n_1, n_2-1), f(n_1, n_2), f(n_1, n_2+1), f(n_1, n_2+2)\}$$

Διαστολή (Dilation)



$$g(\vec{n}) = OR[\vec{B}f(\vec{n})]$$

Διάβρωση (Erosion)



$$g(\vec{n}) = \text{AND}[\vec{B}f(\vec{n})]$$

Άνοιγμα και κλείσιμο

Open = dilate(erode)



Close = erode(dilate)



Η διάβρωση και η διαστολή ΔΕΝ ΕΙΝΑΙ αντίθετες!!!

Εντοπισμός ορίων



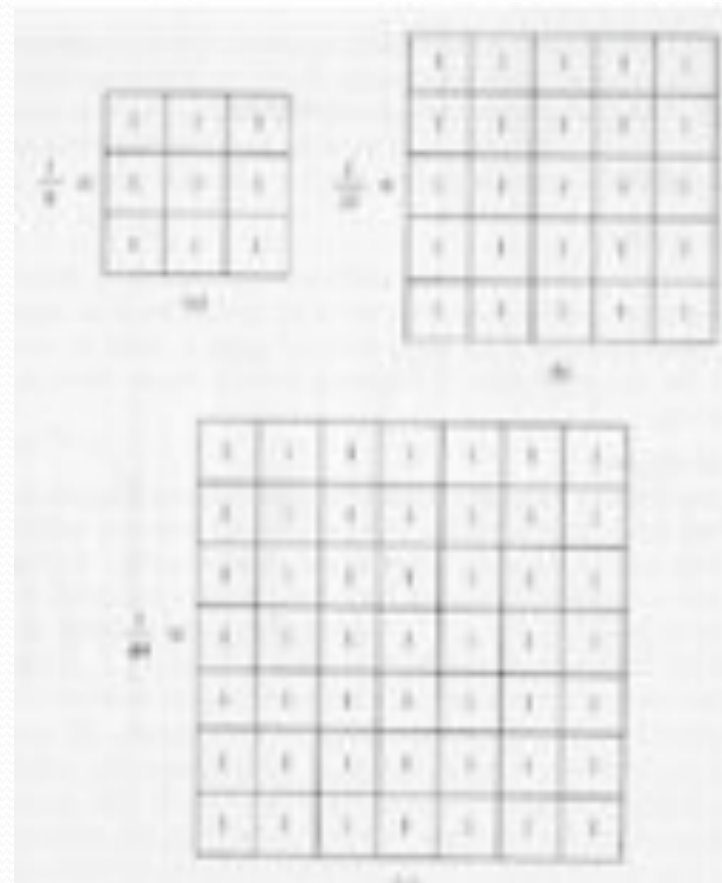
dilate - close =



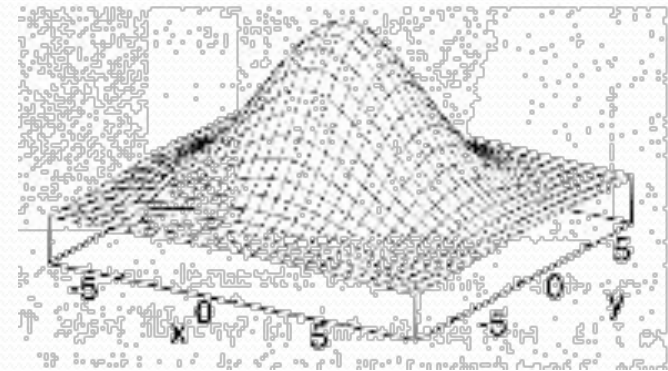
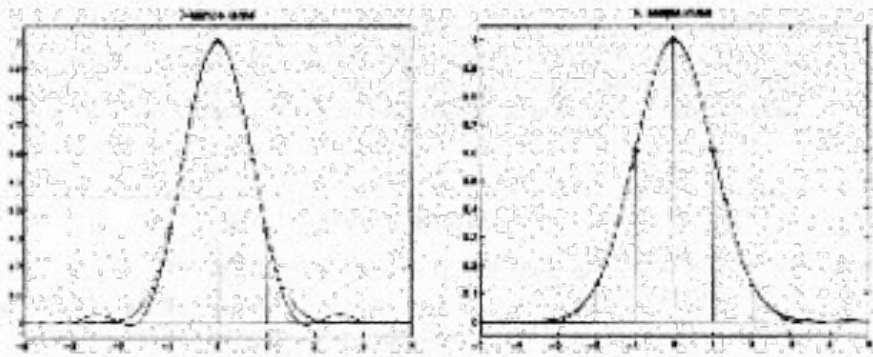
Μάσκες

- Αντίστοιχες των παραθύρων
 - Σε μερικές περιπτώσεις αναφορά και ως δομικά στοιχεία
- Διαφορετικές τιμές σε κάθε σημείο τους
- Διαφορετικό μέγεθος και σχήμα
 - στις συμμετρικές τυπικά το κεντρικό pixel είναι το σημείο αναφοράς
 - στις ασύμμετρες το σημείο αναφοράς μπορεί να είναι οποιοδήποτε (προδηλωμένο) pixel

Φίλτρα μέσων τιμών



Φίλτρα Gauss



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

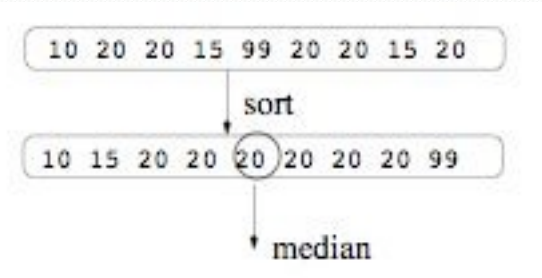
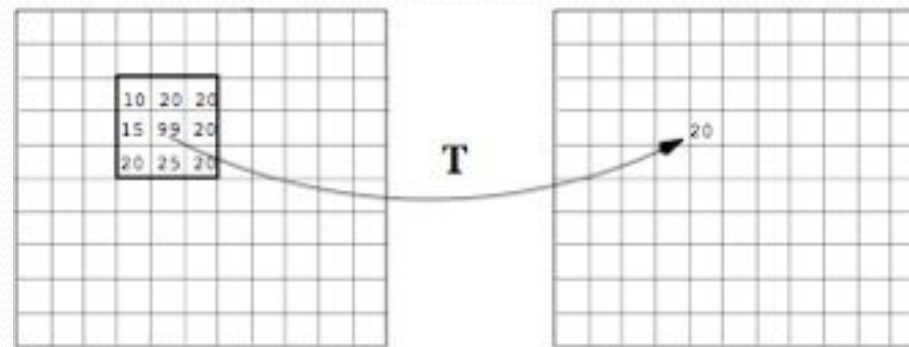
$$\frac{1}{273}$$

1	4	7	4	1
4	16	26	16	4
7	26	41	26	7
4	16	26	16	4
1	4	7	4	1

για $\sigma=1$

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

Φίλτρα διάμεσης τιμής



Επίλογος για τα φίλτρα

- Τα φίλτρα αποτελούν τη συνέλιξη της εικόνας με μία άλλη συνάρτηση
- Ο σχεδιασμός των φίλτρων γίνεται συνήθως με αναλογικά μαθηματικά και μετά μετατρέπεται σε διακριτές τιμές
- Απαιτούνται προσεγγίσεις τιμών κλπ.

Ανίχνευση Ακμών

Μαθηματικός ορισμός ακμών, αριθμητική προσέγγιση παραγώγων, Prewitt, Sobel, Canny, LoG, DoG

Ορισμός των ακμών

- Μέχρι στιγμής η επεξεργασία βασιζόταν στον εντοπισμό περιοχών της εικόνας με κοινές ιδιότητες (φωτεινότητα κλπ.)
- Ακμές: το πραγματικά ζητούμενο στην κατανόηση της εικόνας (χαμηλή ή πρώιμη όραση)
 - Το σύνολο των σημείων που οριοθετούν δύο ή περισσότερες περιοχές από διαφορετικό υλικό σε μία επιφάνεια στο φυσικό κόσμο
 - Το σύνολο των σημείων που οριοθετούν μία απότομη αλλαγή υφής σε μία φυσική επιφάνεια

Πλήρης ορισμός

- Γεωμετρικές ακμές
 - Όρια αντικειμένων = απότομη διαφοροποίηση στο βάθος, χρώμα, υφή
 - Όρια επιφανειών = ξαφνική διαφοροποίηση στη διεύθυνση, το χρώμα, την υφή
- Μη γεωμετρικές ακμές
 - Σκιές
 - Αντανakλάσεις φωτός
 - Αντανakλάσεις ανάμεσα στα αντικείμενα της σκηνής

Αποτύπωση ακμών σε εικόνες

- Απότομη διαφοροποίηση επιλεγμένων χαρακτηριστικών
 - Κατά κύριο λόγο διαφοροποίηση της φωτεινότητας
 - Εναλλακτικά χαρακτηριστικά (τα οποία όμως τελικά στην εικόνα σχετίζονται με τη φωτεινότητα):
 - Ανακλαστικότητα
 - Χρώμα
 - Υφή
- Οι ακμές στην εικόνα
 - Μπορεί να είναι λανθασμένες («ψευτοακμές», π.χ. όρια σκιάς)
 - Μπορεί να μην έχουν καν αποτυπωθεί (υπερβολική έκθεση φωτός μίας επιφάνειας με γωνίες κλπ.)

Στόχος της ανίχνευσης ακμών

- Δημιουργία απλουστευμένων σχηματικών των αντικειμένων της εικόνας (περιγραφή μόνο βάση των ορίων τους)
- Εξαγωγή σημαντικών χαρακτηριστικών, όπως π.χ. γωνίες, γραμμές, κύκλοι
- Χρήση από συστήματα προχωρημένης (υψηλής) όρασης

Τα βασικά χαρακτηριστικά των ακμών

Κάθετος ακμής (edge normal)

κατεύθυνση ακμής (edge direction)



Ιδανικά προφίλ χαρακτηριστικών ακμών

(a) Κατώφλι (Sudden step edge)

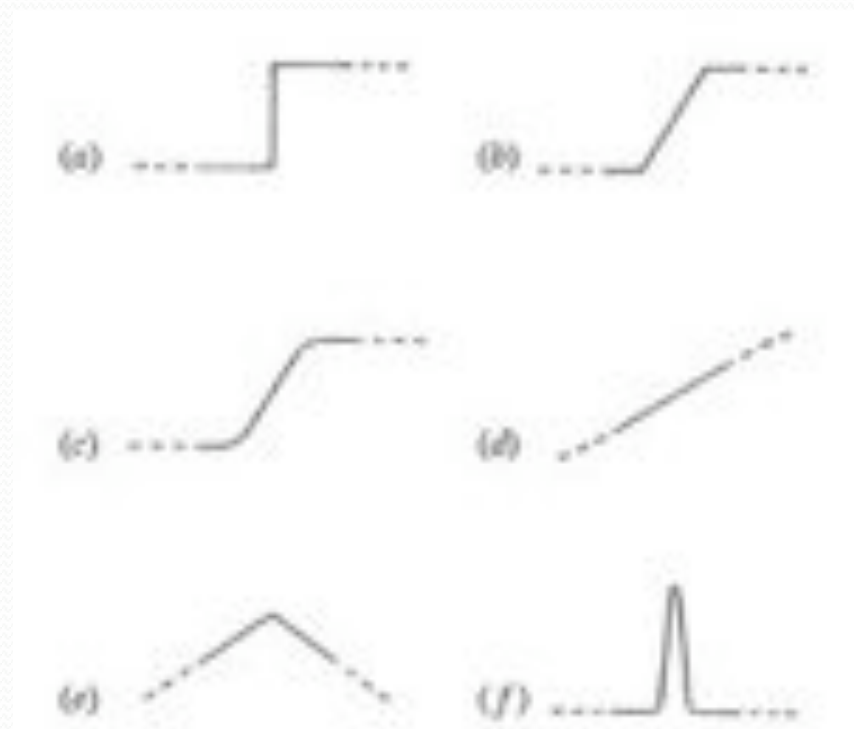
(b) Ράμπα (Ramp edge)

(c) Smooth ramp edge

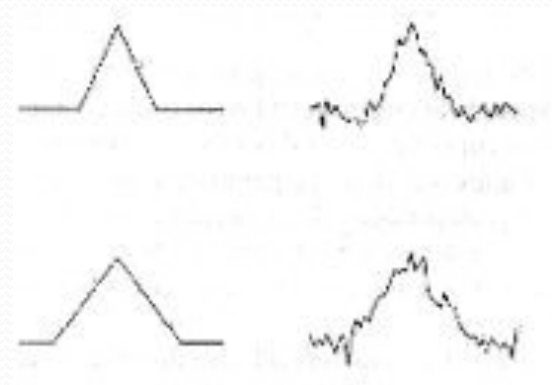
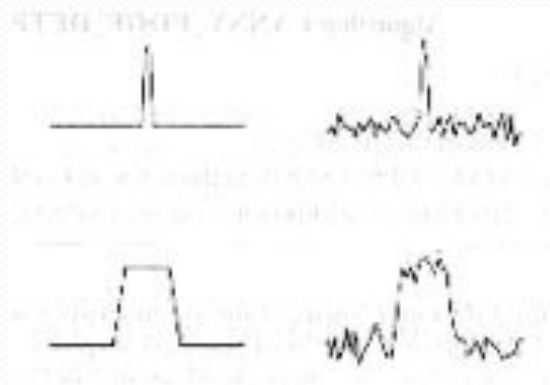
(d) Planar edge

(e) Στέγη (Roof edge)

(f) Ακμή Γραμμής (Line edge)



Κανονικά προφίλ ακμών



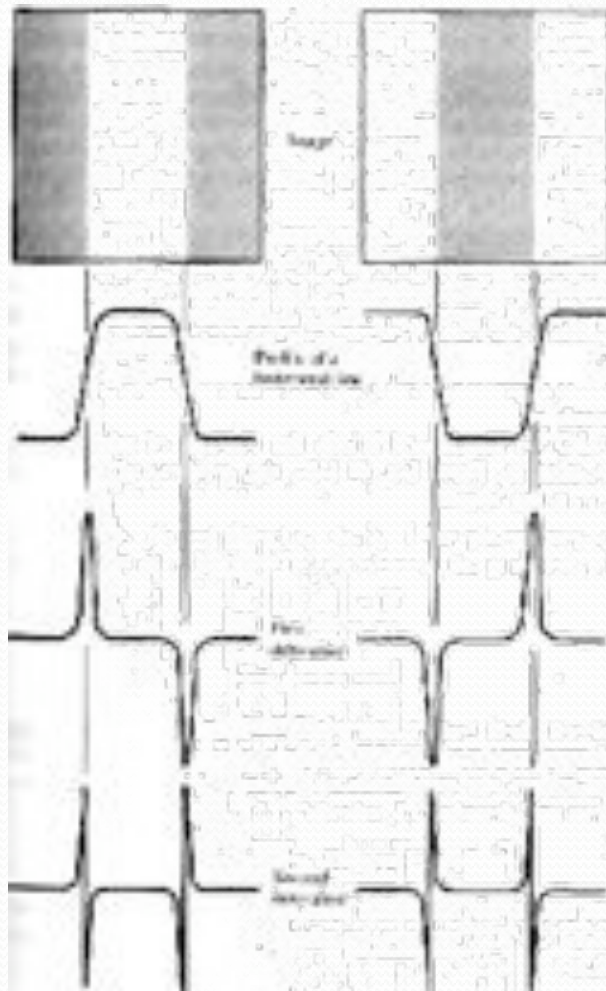
Η γενική διαδικασία ανίχνευσης ακμών

1. Φιλτράρισμα (Smoothing): καταπολέμηση του θορύβου χωρίς αλλοίωση των πραγματικών ακμών
2. Βελτιώσεις (φίλτρα για τη βελτίωση της αντίθεσης)
3. Ανίχνευση των σημείων των ακμών
4. Ακριβής εντοπισμός των σημείων των ακμών (και αφαίρεση λανθασμένων, με «αδυνάτισμα» και «σύνδεση»)

Ανίχνευση

- Η εικόνα είναι ένα δισδιάστατο σήμα
- Σε ένα σήμα ο εντοπισμός των ξαφνικών αλλαγών γίνεται με
 - Τον εντοπισμό των μέγιστων και ελάχιστων στην πρώτη παράγωγο
 - Τον εντοπισμό των σημείων τομής του μηδέν στη δεύτερη παράγωγο

Παραγοντοποίηση εικόνων



Η προσέγγιση της πρώτης παραγώγου

$$f'(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} \approx f(x+1) - f(x) \quad (h=1)$$

mask: [-1 1]

S_1			17	17	13	17	17	21	24	24	24	24
S_1	\ominus	\oplus	0	0	0	0	4	4	0	0	0	0

(a) S_1 is an upward step edge

S_2			24	24	24	24	24	17	17	17	17	17
S_2	\ominus	\oplus	0	0	0	0	-7	-7	0	0	0	0

(b) S_2 is a downward step edge

S_3			0	17	17	17	17	17	21	24	24	24
S_3	\ominus	\oplus	0	0	0	0	0	0	4	7	0	0

(c) S_3 is an upward ramp

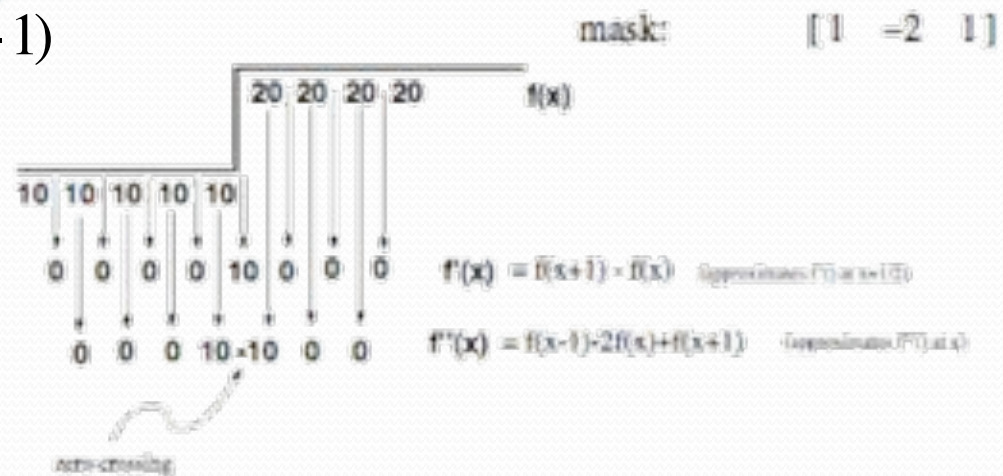
S_4			17	17	17	17	17	17	17	17	17	17
S_4	\ominus	\oplus	0	0	0	0	0	-7	0	0	0	0

Η προσέγγιση της δεύτερης παραγώγου

$$f''(x) = \lim_{h \rightarrow 0} \frac{f'(x+h) - f'(x)}{h} \approx f'(x+1) - f'(x) = f(x+2) - 2f(x+1) + f(x)$$

(h = 1)

$$f''(x) = f(x+1) - 2f(x) + f(x-1)$$



Παραδείγματα 2^{ης} παραγώγου για χαρακτηριστικές γωνίες

παιδί $M = [-1, 2 -1]$

α		0	15	10	12	12	20	21	20	24	24
β_1	α	8	0	1	0	-12	12	0	3	0	0

(a) β_1 is an optimal step

β_1		21	24	21	24	12	12	12	12	12	12
β_2	α	2	0	3	0	12	12	0	3	0	0

(b) β_2 is a dominant step

β_2		12	12	12	12	12	21	21	24	24	24
β_3	α	2	0	0	3	0	0	0	3	0	0

(c) β_3 is an optimal step

β_3		10	12	10	12	24	21	12	-3	12	12
β_4	α	0	0	0	-12	24	-12	0	0	0	0

Ανίχνευση ακμών με χρήση βαθμίδας (gradient)

$$\nabla f = \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{pmatrix}$$

$$\text{magn}(\nabla f) = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2} = \sqrt{M_x^2 + M_y^2} \approx |M_x| + |M_y|$$

$$\text{dir}(\nabla f) = \tan^{-1}\left(\frac{M_x}{M_y}\right)$$

Οι προσεγγίσεις στον υπολογισμό της βαθμίδας

$$\frac{\partial f}{\partial x} = \lim_{h \rightarrow 0} \frac{f(x+h, y) - f(x, y)}{h}$$

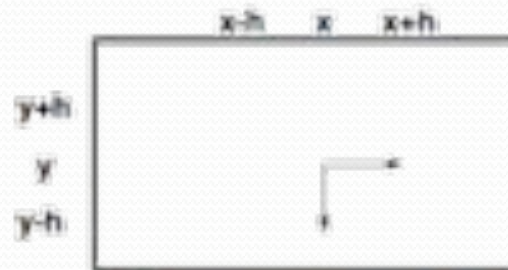
$$\frac{\partial f}{\partial y} = \lim_{h \rightarrow 0} \frac{f(x, y+h) - f(x, y)}{h}$$

απλούστευση

$$\frac{\partial f}{\partial x} = \frac{f(x+h_x, y) - f(x, y)}{h_x} = f(x+1, y) - f(x, y), (h_x=1)$$

$$\frac{\partial f}{\partial y} = \frac{f(x, y+h_y) - f(x, y)}{h_y} = f(x, y+1) - f(x, y), (h_y=1)$$

χρήση συντεταγμένων εικόνας



$$\frac{\partial f}{\partial x} = f(i, j+1) - f(i, j)$$

$$\frac{\partial f}{\partial y} = f(i-1, j) - f(i, j) \text{ or } \frac{\partial f}{\partial y} = f(i, j) - f(i+1, j)$$

Παράδειγμα υπολογισμού

z1	z2	z3
z4	z5	z6
z7	z8	z9

-1	1
----	---

1
-1

$$\frac{\partial I}{\partial x} = z_6 - z_5, \quad \frac{\partial I}{\partial y} = z_5 - z_8$$

$$\text{magn}(\nabla I) = \sqrt{(z_6 - z_5)^2 + (z_5 - z_8)^2}$$

Ανίχνευση Roberts

$$\frac{\partial f}{\partial x} = f(i, j) - f(i+1, j+1)$$

$$\frac{\partial f}{\partial y} = f(i+1, j) - f(i, j+1)$$

$$M_x = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad M_y = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

Ανίχνευση Prewitt και Sobel (1968)

Υποθετική γειτονιά

$$\begin{array}{ccc} a_0 & a_1 & a_2 \\ a_7 & [i, j] & a_3 \\ a_6 & a_5 & a_4 \end{array} \quad \begin{array}{l} M_x = (a_2 + ca_3 + a_4) - (a_0 + ca_7 + a_6) \\ M_y = (a_6 + ca_5 + a_4) - (a_0 + ca_1 + a_2) \end{array}$$

Η σταθερά c τονίζει τη σημασία της κεντρικής σειράς.
Θέτοντας την ίση με 1 προκύπτει η ανίχνευση Prewitt:

$$M_x = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \quad M_y = \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

Θέτοντας την ίση με 2 προκύπτει η ανίχνευση Sobel:

$$M_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad M_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

Παραδείγματα (χωρίς φιλτράρισμα)

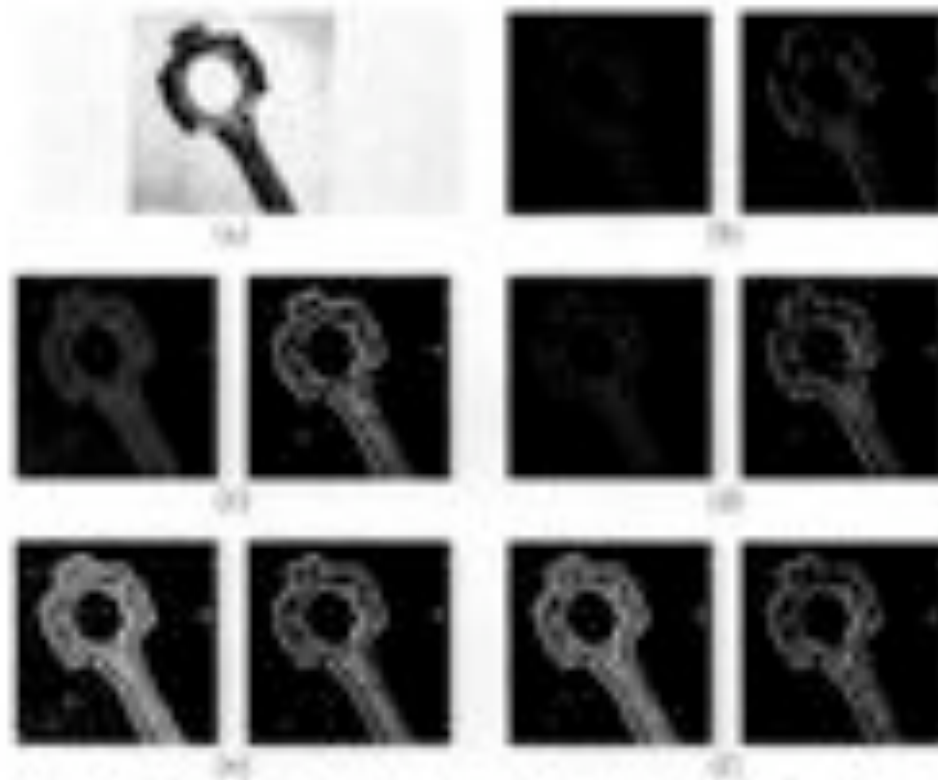


Figure 1.6: A comparison of various edge detectors without blurring. (a) Original image. (b) Gradient gradient using 1×3 and 3×1 kernels, $T = 64$. (c) Gradient using 1×3 kernel, $T = 64$. (d) Roberts cross operator, $T = 64$. (e) Sobel operator, $T = 64$. (f) Prewitt operator, $T = 64$.

Παραδείγματα (με φιλτράρισμα)

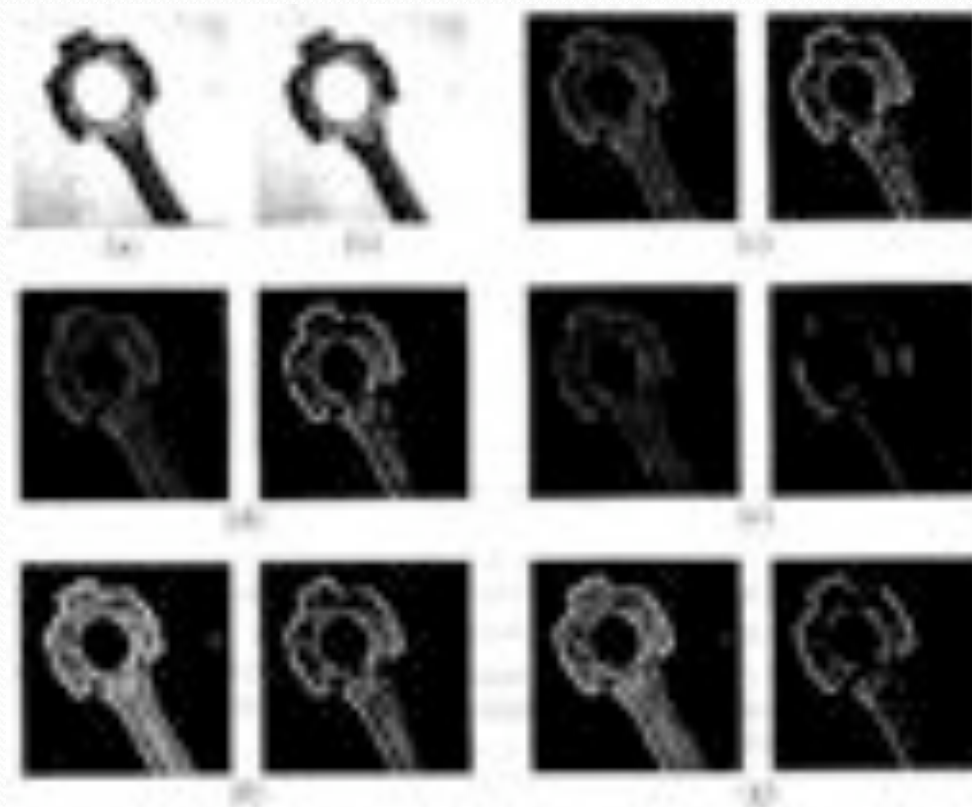


Figure 14. A comparison of various edge detectors. (a) Original image. (b) Filtered image. (c) Sobel gradient using 1×1 and 3×3 masks, $T = 22$. (d) Gabor using 3×3 masks, $T = 44$. (e) Roberts cross operator, $T = 66$. (f) Scharr operator, $T = 110$. (g) Prewitt operator, $T = 110$.

Και ποια είναι μία καλή μέθοδος;

- Ο τελεστής Canny...
 - Καλό μαθηματικό υπόβαθρο (απόδειξη ότι η πρώτη παράγωγος της Γκαουσιανής – Gaussian – ως τελεστής βελτιστοποιεί τη σχέση σήματος-θορύβου- *signal-to-noise ratio* SNR – και πληροί τις προϋποθέσεις θέσης)
 - Λιγότερο ένας τελεστής και περισσότερο μία μέθοδος που αποτελείται από σειρά βημάτων

Ο αλγόριθμος

1. Compute f_x and f_y

$$f_x = \frac{\partial}{\partial x} (f * G) = f * \frac{\partial}{\partial x} G = f * G_x$$

$$f_y = \frac{\partial}{\partial y} (f * G) = f * \frac{\partial}{\partial y} G = f * G_y$$

$G(x, y)$ is the Gaussian function

$G_x(x, y)$ is the derivative of $G(x, y)$ with respect to x : $G_x(x, y) = \frac{-x}{\sigma^2} G(x, y)$

$G_y(x, y)$ is the derivative of $G(x, y)$ with respect to y : $G_y(x, y) = \frac{-y}{\sigma^2} G(x, y)$

2. Compute the gradient magnitude

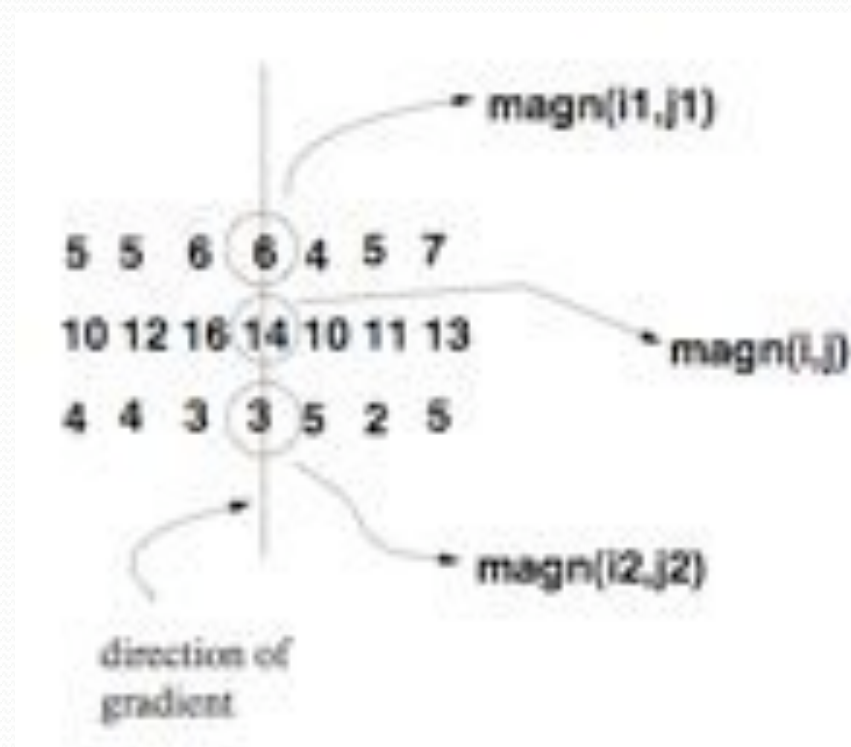
$$\text{mag}(f_x, f_y) = \sqrt{f_x^2 + f_y^2}$$

3. Apply non-maxima suppression.

4. Apply hysteresis thresholding/edge linking.

Αποκλεισμός μη μέγιστων τιμών (non-maximum suppression)

- Εντοπισμός των τοπικών μέγιστων τιμών
- Οι φαρδιές γραμμές πρέπει να περιοριστούν ώστε να παραμείνει η ακμή ως ένδειξη μεγάλης διαφοροποίησης και τίποτε άλλο
- Όλες οι τιμές κατά μήκος της βαθμίδα εκτός της μέγιστης απορρίπτονται



Διαβάθμιση με υστέρηση

- Παρά τον αποκλεισμό των μη μέγιστων τιμών, οι τοπικά μέγιστες τιμές λόγω θορύβου εξακολουθούν να υφίστανται
- Η χρήση κατωφλιών επιφέρει τα γνωστά προβλήματα
 - Με υψηλή τιμή κατωφλιού «καταστρέφονται» και οι κανονικές ακμές
 - Με χαμηλή τιμή κατωφλιού δεν καταπολεμάται όλος ο θόρυβος
- Πιθανή διέξοδος: δύο κατώφλια...
 - Χαμηλό: t_l
 - Υψηλό: t_h

$$t_h \approx 2t_l$$

Διαβάθμιση με υστέρηση (2)

- Δημιουργούμε 2 εικόνες με τα δύο κατώφλια $I_1(i,j)$ καθώς και $I_2(i,j)$
- Συμπληρώνουμε τις ακμές στην εικόνα με το υψηλό κατώφλι με τη βοήθεια της εικόνας με το χαμηλότερο κατώφλι
 - Για κάθε κενό που ανακαλύπτουμε στην εικόνα 1 κοιτάζουμε στην εικόνα δύο στη γειτονιά των 8
 - Συμπληρώνουμε μέχρι να προκύψει σύνδεση με άλλη ακμή (αν όχι απορρίπτουμε)

Η δεύτερη παράγωγος

- Εναλλακτικά στη μέγιστη τιμή της πρώτης παραγώγου, μπορεί να χρησιμοποιηθεί και η μηδενοδιέλευση (ή διασταύρωση μηδενικού σημείου, zero crossing) της 2^{ης} παραγώγου
- Ένας πιθανός τρόπος υπολογισμού της σε δύο διαστάσεις είναι με τη βοήθεια της Λαπλασιανής:

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

- Approximating $\nabla^2 f$:

$$\frac{\partial^2 f}{\partial x^2} = f(i, j+1) - 2f(i, j) + f(i, j-1)$$

$$\frac{\partial^2 f}{\partial y^2} = f(i+1, j) - 2f(i, j) + f(i-1, j)$$

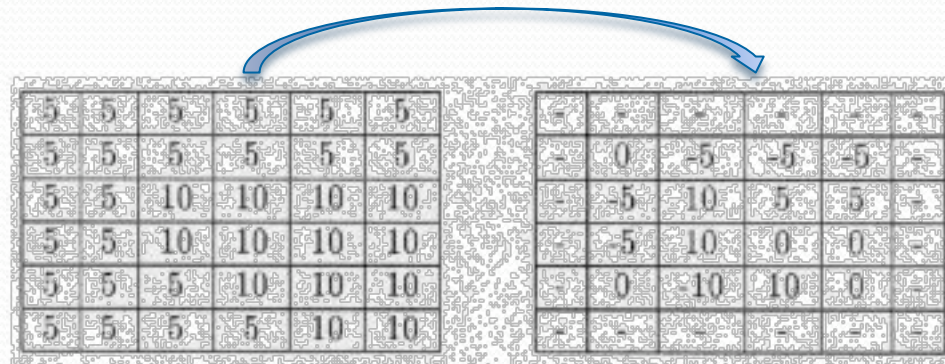
$$\nabla^2 f = -4f(i, j) + f(i, j+1) + f(i, j-1) + f(i+1, j) + f(i-1, j)$$

Η Λαπλασιανή για εικόνες

z1	z2	z3
z4	z5	z6
z7	z8	z9

$$\nabla^2 f = -4z_5 + (z_2 + z_4 + z_6 + z_8)$$

0	1	0
1	-4	1
0	1	0



Ιδιότητες Λαπλασιανής

- + Ισότροπη
- + Μικρές απαιτήσεις υπολογιστικά (μία μάσκα)
- Δε δίνει πληροφορία για την κατεύθυνση των ακμών
- Πιο ευαίσθητη στο θόρυβο (με τη διπλή παραγωγή)



Laplacian of Gaussian (LoG)

- Απλά πρώτα βελτιώνουμε την εικόνα με ένα φίλτρο Gauss

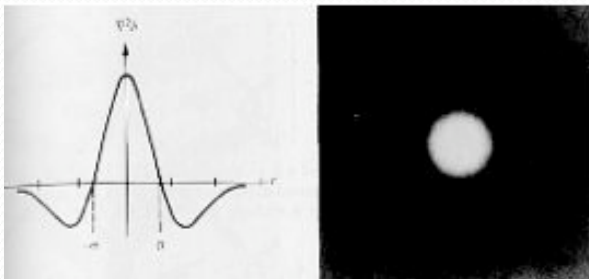
$$G(x,y) = e^{-\frac{x^2+y^2}{2\sigma^2}}$$

(σ determines the degree of smoothing, mask size increases with σ)

$$\nabla^2[f(x,y) * G(x,y)] = \nabla^2 G(x,y) * f(x,y)$$



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



Και τι κερδίσαμε;

- Σε σχέση με τη βαθμίδα:
 - Η βαθμίδα εντοπίζει κυρίως έντονες ακμές
 - Η βαθμίδα επηρεάζεται από θόρυβο
 - Η LoG έχει καλύτερη απόδοση στην εύρεση σωστής θέσης, αλλά και όταν οι ακμές δεν είναι ιδιαίτερα έντονες (και τα δύο χάρη στη μηδενοδιέλευση)

Λίγα τελικά σχόλια για την ανίχνευση ακμών

- Το επίκεντρο της τεχνητής όρασης
- Σπάνια μπορεί κάποιος να βασιστεί σε έναν τελεστή για να λύσει το πρόβλημά του
 - Προσπαθούμε να εξασφαλίσουμε στο χώρο της λήψης της κατάλληλες συνθήκες (π.χ. Φωτισμό, απόσταση από φακό κλπ.) που θα επιτρέψουν ευκολότερο προσδιορισμό
- Συχνά απαιτείται συνδυασμός περισσότερων τεχνικών (πολλές φορές ευρετικών) και φυσικά υψηλής όρασης για να λυθεί ένα πρόβλημα εντοπισμού

