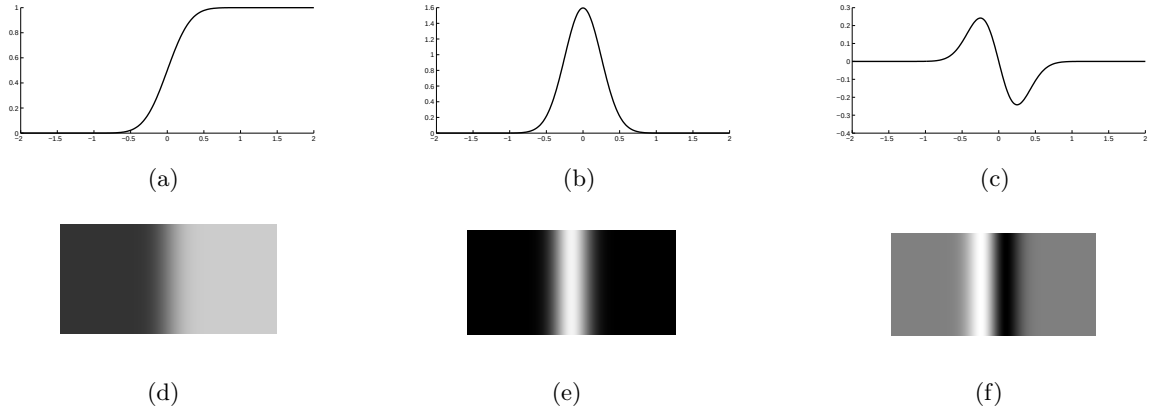
Ανίχνευση ακμών

Γεώργιος Τζιρίτας
Τμήμα Επιστήμης Υπολογιστών
Πανεπιστήμιο Κρήτης

Φεβρουάριος 2021

1 Εισαγωγή

Η ανίχνευση ακμών σκοπεύει στον εντοπισμό σημείων με σημαντική μεταβολή στη φωτεινή ένταση και συνακόλουθα στην εύρεση των φυσικών αιτίων που προκαλούν τις μεταβολές. Οι μεταβολές της φωτεινής έντασης μπορεί να οφείλονται σε ασυνέχειες βάθους, σε ασυνέχειες στην επιφάνεια ενός αντικειμένου, γεωμετρικές ή φωτομετρικές. Επομένως τα σημεία ακμής εντοπίζονται εκεί όπου η μεταβολή της φωτεινής έντασης είναι τοπικά μέγιστη. Ένα σύνολο συνδεόμενων σημείων ακμής αποτελεί ένα περίγραμμα. Σε κάθε σημείο ενός περιγράμματος μπορεί να ορισθεί ο προσανατολισμός του σε σχέση με το σύστημα συντεταγμένων του επιπέδου της εικόνας. Η φωτεινή ένταση παρουσιάζει τη μέγιστη μεταβολή στην κάθετη στο περίγραμμα κατεύθυνση. Με αυτή την έννοια η ανίχνευση ακμής ισοδυναμεί με τον εντοπισμό μιας μονοδιάστατης μεταβολής. Στο Σχήμα 1 δίδεται μια οριζόντια βαθμιαία μεταβολή και δείχνεται πώς η πρώτη ή η δεύτερη παράγωγος μπορούν να χρησιμοποιηθούν για την ανίχνευση και τον εντοπισμό της μεταβολής.



Σχήμα 1: Επάνω δίδονται η συνάρτηση μιας βαθμιαίας μεταβολής, η πρώτη και η δεύτερη παράγωγός της. Κάτω δίδονται μια εικόνα με οριζόντια βαθμιαία μεταβολή, η πρώτη οριζόντια παραγώγιση και η δεύτερη οριζόντια παραγώγιση.

Σε δισδιάστατες εικόνες η ανίχνευση των σημείων ακμής μπορεί να γίνει με τη βοήθεια είτε του διανύσματος των δύο πρώτων μερικών παραγώγων της φωτεινής έντασης, που ονομάζεται κλίση, είτε του αθροίσματος των δύο δεύτερων μερικών παραγώγων της φωτεινής έντασης, που καλείται

λαπλασιανός (Laplacian) τελεστής. Ο τελεστής αυτός είναι ισοτροπικός, δηλαδή η απόκρισή του είναι ανεξάρτητη από τον προσανατολισμό του περιγράμματος. Η ισοτροπικότητα είναι σημαντική ιδιότητα για την ακρίβεια του εντοπισμού των περιγραμμάτων. Ωστόσο οι τελεστές παραγώγισης είναι ευαίσθητοι σε θόρυβο και γι΄ αυτό απαιτείται η εξομάλυνση των τιμών της εικόνας, που οδηγεί σε απώλεια της ακρίβειας του εντοπισμού. Γενικά απαιτείται επίσης μια ύστερη επεξεργασία για την απαλοιφή δομικά εσφαλμένων ακμών. Τέλος, μπορεί να ακολουθεί μια διεργασία σύνδεσης τμημάτων από γραμμές ακμών για να προκύψουν κλειστά περιγράμματα που ορίζουν τμήματα της εικόνας.

2 Μέθοδος της κλίσης

Το διάνυσμα της κλίσης της φωτεινής έντασης, $I(x, y)$, έχει σα συνιστώσες τις δύο πρώτες μερικές παραγώγους

$$\nabla I = \begin{bmatrix} I_x \\ I_y \end{bmatrix} \tag{1}$$

Για την ανίχνευση των ακμών χρησιμοποιείται το μέτρο του διανύσματος της κλίσης

$$\|\nabla I\| = \sqrt{I_x^2 + I_y^2}. \tag{2}$$

Αποδεικνύεται ότι το μέτρο του διανύσματος της κλίσης για συνεχή σήματα ικανοποιεί την ιδιότητα της ισοτροπικότητας. Η κατεύθυνση του διανύσματος της κλίσης δίδεται από τη γωνία

$$\theta = \arctan \frac{I_y}{I_x}. \tag{3}$$

Η κλίση της φωτεινής έντασης σε μια δοσμένη κατεύθυνση, ϕ, είναι

$$I_\phi = I_x \cos\phi + I_y \sin\phi. \tag{4}$$

Η κλίση της φωτεινής έντασης είναι μέγιστη στην κατεύθυνση του διανύσματος της κλίσης, και μηδενική στην κάθετη σ΄ αυτήν κατεύθυνση. Σημεία ακμής ανιχνεύονται ακριβώς στην κατεύθυνση που η κλίση είναι μέγιστη, οπότε είναι ίση με το μέτρο του διανύσματος της κλίσης, $\|\nabla I\|$. Τελικά τα σημεία ακμής ανιχνεύονται ως τα τοπικά μέγιστα του μέτρου του διανύσματος της κλίσης, με τα τοπικά μέγιστα να ευρίσκονται στην κατεύθυνση αυτού του ίδιου διανύσματος.

Για την εφαρμογή της μεθόδου της κλίσης απαιτείται επομένως ο αριθμητικός υπολογισμός των πρώτων μερικών παραγώγων της φωτεινής έντασης. Εφόσον πρόκειται για μερική παραγώγιση, το χρησιμοποιούμενο φίλτρο είναι μονοδιάστατο. Ωστόσο λόγω της ύπαρξης θορύβου, είναι αναγκαία επιπλέον η χρήση ενός βαθυπερατού φίλτρου στην άλλη κατεύθυνση, ως προς αυτή της παραγώγισης. Έτσι στην περίπτωση της οριζόντιας παραγώγισης, πραγματοποιείται ολοκλήρωση στην κατακόρυφη κατεύθυνση, και αντίστροφα. Τα δύο φίλτρα χρησιμοποιούνται ξεχωριστά, ώστε το 2-Δ φίλτρο που προκύπτει να είναι διαχωρίσιμο. Ας είναι $g(.)$ η κρουστική απόκριση του φίλτρου παραγώγισης, και $h(.)$ η κρουστική απόκριση του φίλτρου ολοκλήρωσης. Το φίλτρο παραγώγισης $g(.)$ είναι πάντοτε αντισυμμετρικό

$$g(n) = -g(-n),$$

ενώ θα πρέπει το φίλτρο εξομάλυνσης $h(.)$ να είναι συμμετρικό, για να μην προκαλεί μετατόπιση,

$$h(n) = h(-n).$$

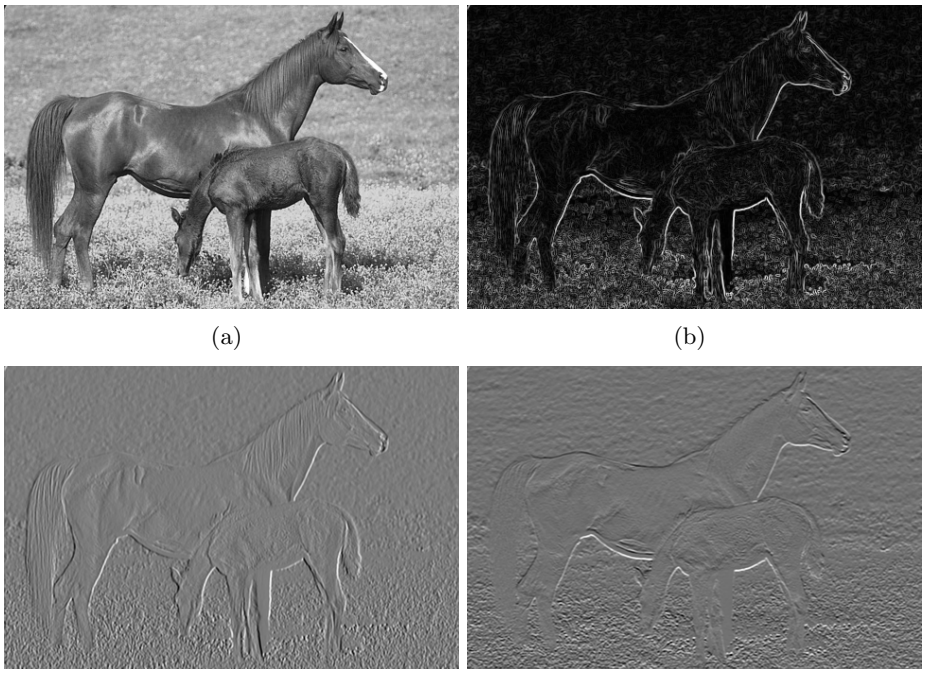
Παρουσιάζονται στη συνέχεια μερικά φίλτρα παραγώγισης και ολοκλήρωσης με πεπερασμένη κρουστική απόκριση. Το απλούστερο αντισυμμετρικό φίλτρο παραγώγισης είναι το ακόλουθο

$$g(n) = \begin{cases} 1/2 & n = -1 \\ -1/2 & n = 1 \\ 0 & |n| \neq 1 \end{cases} \tag{5}$$

Σε αντιστοιχία με την ίδια έκταση ως φίλτρο εξομάλυνσης μπορεί να ληφθεί το ακόλουθο

$$h(m) = \begin{cases} 1/4 & |m| = 1 \\ 1/2 & m = 0 \\ 0 & |m| > 1 \end{cases} \tag{6}$$

Το 2-Δ φίλτρο που προκύπτει είναι γνωστό ως το φίλτρο Sobel. Προκύπτει με αυτό τον τρόπο



(a) (b)

(c) (d)

Σχήμα 2: Χρήση του φίλτρου Sobel για την ανίχνευση ακμών. Επάνω δίδεται η εικόνα και το μέτρο του διανύσματος της κλίσης. Κάτω αριστερά δίδεται η οριζόντια παραγώγιση και κάτω δεξιά η κατακόρυφη.

ένα ζεύγος φίλτρων. Όταν οριζόντια χρησιμοποιείται το φίλτρο παραγώγισης, κατακόρυφα χρησιμοποιείται το φίλτρο εξομάλυνσης, και αντίστροφα. Αποτελέσματα με χρήση αυτού του ζεύγους φίλτρων δίδονται στο Σχήμα 2. Αν είναι

$$h(m) = \begin{cases} 1/3 & |m| \leq 1 \\ 0 & |m| > 1 \end{cases} \tag{7}$$

πρόκειται για το φίλτρο Prewitt.

Αν λόγω θορύβου απαιτείται ένα φίλτρο ολοκλήρωσης με πιο τονισμένο το βαθυπερατό χαρακτήρα, αρκεί να πάρουμε τη συνέλιξη του (6) με τον εαυτό του, που δίνει το εξής αποτέλεσμα

$$h(m) = \begin{cases} 1/16 & |m| = 2 \\ 1/4 & |m| = 1 \\ 6/16 & m = 0 \\ 0 & |m| > 2 \end{cases}$$

Αντίστοιχο φίλτρο παραγώγισης με την ίδια έκταση μπορεί να ορισθεί το

$$g(n) = \begin{cases} 1/8 & n = -2 \\ 1/4 & n = -1 \\ -1/4 & n = 1 \\ -1/8 & n = 2 \\ 0 & n = 0 \ \text{ είτε } \ |n| > 2 \end{cases}$$

Το πρόβλημα της σχεδίασης του φίλτρου παραγώγισης μπορεί να τεθεί σαν ένα πρόβλημα βελτιστοποίησης ορισμένων κριτηρίων που να εκφράζουν την καλή ανίχνευση ακμών με παρουσία θορύβου. Το πρόβλημα τέθηκε αρχικά από τον Canny [1] και λύθηκε για ένα συνεχές μονοδιάστατο σήμα, όπου το μοντέλο της ασυνέχειας δίδεται από μια βηματική συνάρτηση, στην οποία προστίθεται λευκός θόρυβος με μηδενική μέση τιμή και διασπορά σ^2. Ας υποθέσουμε ότι το σήμα έχει ως εξής

$$f(x) = \begin{cases} A, & x \geq 0 \\ 0, & x < 0 \end{cases} \tag{8}$$

Ζητείται το φίλτρο παραγώγισης που ικανοποιεί τρία κριτήρια: καλής ανίχνευσης, μικρής τοπικής απόκλισης και μικρής παρενόχλησης από πολλαπλά μέγιστα. Η καλή ανίχνευση μετριέται με τη σηματοθορυβική σχέση στην έξοδο του φίλτρου στο σημείο της ασυνέχειας

$$\mathcal{R} = \frac{|\int_{-\infty}^{\infty} g(-x) f(x) dx|}{\sigma \sqrt{\int_{-\infty}^{\infty} g^2(x) dx}}. \tag{9}$$

Η σηματοθορυβική σχέση πρέπει να είναι η μέγιστη δυνατή. Για τη βηματική συνάρτηση (8) το κριτήριο της σηματοθορυβικής σχέσης δίδει, για $A > 0$,

$$\mathcal{R} = \frac{A}{\sigma} \frac{|\int_{0}^{\infty} g(-x) dx|}{\sqrt{\int_{-\infty}^{\infty} g^2(x) dx}}. \tag{10}$$

Η ασυνέχεια πρέπει επίσης να ανιχνεύεται όσο γίνεται πλησιέστερα στο σημείο της πραγματικής ασυνέχειας του σήματος. Έχει προταθεί το ακόλουθο κριτήριο καλού εντοπισμού

$$\mathcal{L} = \frac{|\int_{-\infty}^{\infty} f'(-x) g'(x) dx|}{\sigma \sqrt{\int_{-\infty}^{\infty} g'^2(x) dx}}. \tag{11}$$

Για τη βηματική συνάρτηση (8) το κριτήριο καλού εντοπισμού δίδει

$$\mathcal{L} = \frac{A}{\sigma} \frac{|g'(0)|}{\sqrt{\int_{-\infty}^{\infty} g'^2(x) dx}}. \tag{12}$$

Αποδεικνύεται ότι το γινόμενο των δύο κριτηρίων $\mathcal{RL}$ είναι αναλλοίωτο σε μια αλλαγή κλίμακας των φίλτρων. Αυτό σημαίνει ότι ένα πιο ομαλό φίλτρο παραγώγισης βελτιώνει τη σηματοθορυβική σχέση και κατά τον ίδιο παράγοντα χειροτερεύει την ακρίβεια εντοπισμού.

Τα δύο ανωτέρω κριτήρια δεν αρκούν για την καλή ανίχνευση των ακμών, γιατί δε λαμβάνουν υπόψη ότι μια βηματική μεταβολή, μπορεί να προκαλέσει πολλαπλά μέγιστα μετά την παραγώγιση λόγω της παρουσίας θορύβου. Αποδεικνύεται ότι η μέση απόσταση μεταξύ δύο μεγίστων στην έξοδο του φίλτρου είναι

$$\mathcal{D} = 2\pi\sqrt{\frac{\int_{-\infty}^{\infty} g'^2(x)dx}{\int_{-\infty}^{\infty} g''^2(x)dx}}. \tag{13}$$

Με δοσμένη την απαίτηση ικανοποίησης πολλαπλών κριτηρίων, έχει προταθεί ο ορισμός του βέλτιστου φίλτρου παραγώγισης ως εκείνου που μεγιστοποιεί το γινόμενο $\mathcal{RL}$, για δοσμένη αρχική σηματοθορυβική σχέση $\frac{A}{\sigma}$, με μία συνθήκη για το $\mathcal{D}$. Κάτω από αυτούς τους όρους είναι δυνατή η λύση του προβλήματος βελτιστοποίησης για την περίπτωση του σήματος (8) και για πεπερασμένη ή άπειρη κρουστική απόκριση. Για την περίπτωση μιας πεπερασμένης, και οπωσδήποτε αντισυμμετρικής, κρουστικής απόκρισης, εκτεινόμενης στο διάστημα $[-M, M]$, ζητείται επομένως, για διάφορες τιμές του $\mathcal{D}/M$, η μεγιστοποίηση του

$$\mathcal{R}_0\mathcal{L}_0 = \frac{|g'(0)||\int_{-M}^{0} g(x)dx|}{\sqrt{\int_{-M}^{M} g^2(x)dx \int_{-M}^{M} g'^2(x)dx}} \tag{14}$$

Για χαμηλές τιμές του $\mathcal{D}/M(\approx 0.1)$, η λύση που προκύπτει μπορεί να προσεγγισθεί από

$$g(x) = \begin{cases} 1/M^2 & -M \leq x < 0 \\ -1/M^2 & 0 < x \leq M \\ 0 & x = 0 \text{ είτε } |x| > M \end{cases} \tag{15}$$

Το αντίστοιχο φίλτρο ολοκλήρωσης είναι

$$h(x) = \begin{cases} \frac{M-|x|}{M^2} & |x| \leq M \\ 0 & |x| > M \end{cases} \tag{16}$$

Για υψηλές τιμές του $\mathcal{D}/M(\approx 1.5)$, η λύση που προκύπτει μπορεί να προσεγγισθεί από την παράγωγο μιας συνάρτησης Gauss

$$g(x) = -\frac{x}{\alpha_2^3\sqrt{2\pi}} e^{-\frac{x^2}{2\alpha_2^2}}. \tag{17}$$

Το φίλτρο αυτό επιτυγχάνει $\mathcal{R}_0\mathcal{L}_0 = \sqrt{\frac{8}{3\pi}}$. Προφανώς το φίλτρο αυτό έχει άπειρη κρουστική απόκριση και μπορεί να υλοποιηθεί, είτε κατά προσέγγιση από φίλτρο πεπερασμένης έκτασης, είτε με αναδρομικές σχέσεις, ή ακόμα στο χώρο των συχνοτήτων.

Αν το φίλτρο ζητηθεί κατευθείαν να έχει άπειρη κρουστική απόκριση, η λύση του προβλήματος βελτιστοποίησης δίδει [2]

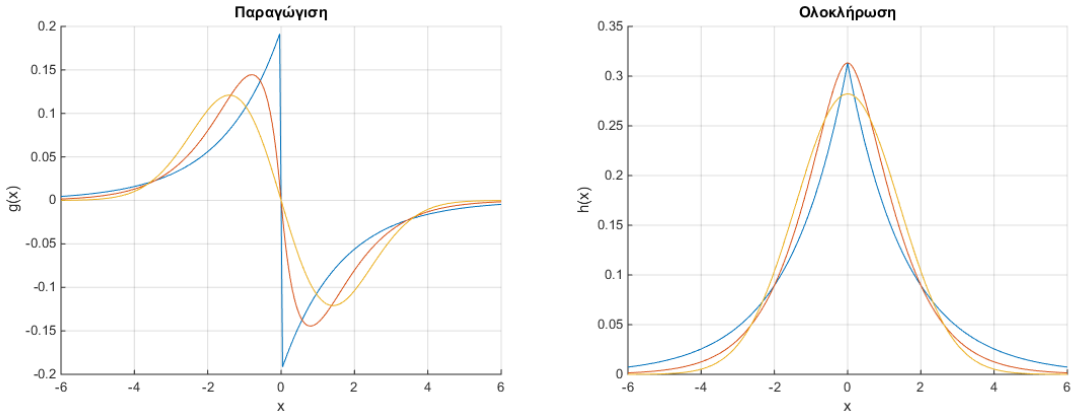
$$g(x) = -\frac{1}{4}\alpha_1^3 x e^{-\alpha_1|x|}. \tag{18}$$

Υψηλές τιμές του α_1 αντιστοιχούν σε χαμηλές τιμές του $\mathcal{D}$, και αντίστροφα. Το αντίστοιχο φίλτρο ολοκλήρωσης είναι

$$h(x) = \frac{1}{4}\alpha_1(1 + \alpha_1|x|)e^{-\alpha_1|x|}. \tag{19}$$

Άλλο χρησιμοποιούμενο ζεύγος φίλτρων με άπειρη κρουστική απόκριση είναι το ακόλουθο [4]

$$g(x) = -\frac{\alpha_0^2}{2}\text{sign}(x)e^{-\alpha_0|x|} \quad \text{και} \quad h(x) = \frac{\alpha_0}{2}e^{-\alpha_0|x|} \tag{20}$$



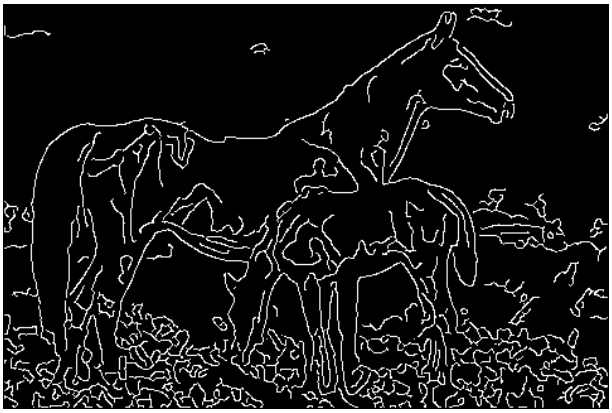
(a) (b)

Σχήμα 3: Φίλτρα παραγώγισης και ολοκλήρωσης.

με παρόμοιο ρόλο για το α_0. Όλα τα παραπάνω φίλτρα χρησιμοποιούνται φυσικά σε διακριτή μορφή, ανεξάρτητα αν είναι πεπερασμένης ή άπειρης κρουστικής απόκρισης. Για τα φίλτρα άπειρης κρουστικής απόκρισης υπάρχουν υλοποιήσεις μέσω αναδρομικών σχέσεων.

Στο Σχήμα 3(a) δίδονται τα παραπάνω φίλτρα παραγώγισης, με τέτοιες τιμές για τις παραμέτρους, ώστε η σηματοθορυβική σχέση στην έξοδο του φίλτρου να είναι ίδια. Συγκεκριμένα θα είναι για την παράγωγο της συνάρτησης Gauss, $\alpha_2 = \sqrt{2}$, για το φίλτρο της Εξίσωσης (18), που δίδεται με διακοπτόμενη γραμμή, $\alpha_1 = \sqrt{\pi/2}$, και για το φίλτρο της Εξίσωσης (20), $\alpha_0 = \frac{1}{2}\sqrt{\frac{\pi}{2}}$. Τα αντίστοιχα φίλτρα ολοκλήρωσης δίδονται στο Σχήμα 3(b).

Στο Σχήμα 4 δίδεται το αποτέλεσμα χρήσης της παραγώγου της συνάρτησης Gauss (φίλτρο Canny) με $\alpha_2 = 1$ και με ανίχνευση που χρησιμοποιεί δύο κατώφλια και την τεχνική της υστέρησης. Με αυτή την τεχνική κατά την εφαρμογή του δεύτερου και μικρότερου κατωφλιού λαμβάνεται επίσης υπόψη η συνδεσιμότητα των σημείων ακμής.



Σχήμα 4: Αποτέλεσμα ανίχνευσης ακμών με τη μέθοδο του Canny.

3 Μέθοδος δευτέρων παραγώγων

Η διέλευση από το μηδέν των τιμών του αθροίσματος των δύο δευτέρων παραγώγων δίδει επίσης σημεία ακμής της φωτεινής έντασης [3]. Ο τελεστής που δίδει αυτό το άθροισμα, ονομάζεται λαπλασιανός (Laplacian), και ορίζεται ως ακολούθως

$$\Delta I = I_{xx} + I_{yy}. \tag{21}$$

Αποδεικνύεται ότι ο λαπλασιανός τελεστής είναι ισοτροπικός, αρκεί να είναι: $I_{xy} = I_{yx}$.

Απλά φίλτρα πεπερασμένης κρουστικής απόκρισης που χρησιμοποιούνται για την υλοποίηση του λαπλασιανού τελεστή είναι τα ακόλουθα:

$$g(m,n) = \begin{cases} -4 & m^2+n^2=0 \\ 1 & m^2+n^2=1 \\ 0 & m^2+n^2>1 \end{cases} \tag{22}$$

και

$$g(m,n) = \begin{cases} -4 & m^2+n^2=0 \\ 1/2 & m^2+n^2=1 \text{ είτε } m^2+n^2=2 \\ 0 & m^2+n^2>2 \end{cases} \tag{23}$$

Ωστόσο αυτά τα φίλτρα χαρακτηρίζονται από ενίσχυση του θορύβου.

Για τη μείωση του θορύβου προτείνεται η χρήση ενός βαθυπερατού φίλτρου Gauss, του οποίου η κρουστική απόκριση, σε συνεχείς συντεταγμένες είναι

$$h(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}. \tag{24}$$

Η όλη επεξεργασία είναι ισοδύναμη με τη συνέλιξη του σήματος της φωτεινής έντασης με το φίλτρο που προκύπτει από την εφαρμογή του λαπλασιανού τελεστή στην κρουστική απόκριση Gauss, που δίδει το ακόλουθο αποτέλεσμα

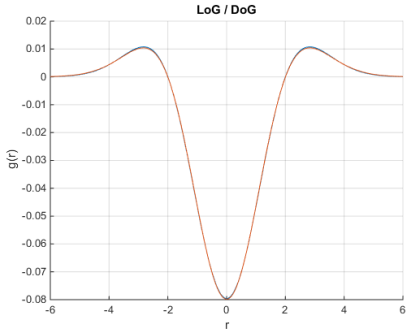
$$g(r) = \frac{1}{\pi\sigma^4}\left(\frac{r^2}{2\sigma^2} - 1\right) e^{-\frac{r^2}{2\sigma^2}} \tag{25}$$

όπου $r = \sqrt{x^2+y^2}$. Αποδεικνύεται επίσης ότι η παραπάνω κρουστική απόκριση μπορεί να προσεγγισθεί από μία διαφορά δύο συναρτήσεων Gauss,

$$g(r) \propto \frac{1}{2\pi\sigma_1^2\sigma^2} e^{-\frac{r^2}{2\sigma_1^2}} - \frac{1}{2\pi\sigma_2^2\sigma^2} e^{-\frac{r^2}{2\sigma_2^2}}, \tag{26}$$

όπου $\sigma_1^2 = 1.28\sigma^2$ και $\sigma_2^2 = 0.8\sigma^2$.

Στο Σχήμα 5 δίδεται τόσο η λαπλασιανή της συνάρτησης Gauss, όσο και η προσέγγισή της με τη διαφορά δύο συναρτήσεων Gauss (διαφορετικά χρώματα). Επειδή η συνάρτηση Gauss δύο μεταβλητών είναι διαχωρίσιμη, η διαφορά δύο συναρτήσεων Gauss απαιτεί την υλοποίηση δύο διαχωρίσιμων φίλτρων. Η συνάρτηση Gauss μιας μεταβλητής προσεγγίζεται σε διακριτή μορφή, είτε από ένα φίλτρο πεπερασμένης κρουστικής απόκρισης, είτε από δύο αναδρομικά φίλτρα.



(a)

(b)

Σχήμα 5: Λαπλασιανή της συνάρτησης Gauss και αποτέλεσμα του φίλτρου για τον εντοπισμό των διελεύσεων από το μηδέν.

4 Βιβλιογραφία

[1] J. Canny, A computational approach to edge detection, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 8, No 6, pp. 679–698, Nov. 1986.

[2] R. Deriche, Using Canny' s criteria to derive a recursively implemented optimal edge detector, *International Journal of Computer Vision*, Vol. 1, No. 2, pp. 167–187, 1987.

[3] D. Marr and E. Hildreth, Theory of edge detection, *Proceedings of the Royal Society of London. Series B, Biological Sciences*, pp. 187–217, 1980.

[4] J. Shen and S. Castan, An optimal linear operator for step edge detection, *CVGIP: Graphical Models and Image Processing*, Vol. 54, No. 2, pp. 112–133, March 1992.