ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΟΡΑΣΗ ΥΠΟΛΟΓΙΣΤΩΝ

(2020-2021)

2ο Σύνολο Αναλυτικών Ασκήσεων

Ονοματεπώνυμο:

- Χρήστος Τσούφης

Αριθμός Μητρώου:

- 03117176

Στοιχεία Επικοινωνίας:

- el17176@mail.ntua.gr

Άσκηση 2.1 (Ανίχνευση Ακμών)

(a) *Ζητείται η λύση της άσκησης 10.17(α) από Κεφ. 10 (έκδοση 6/2018).*

Εκφώνηση: (Ronse) (a) Έστω ότι χρησιμοποιούνται ως φίλτρα h_e, h_o στο ενεργειακό μοντέλο (10.114) οι δύο Gabor συναρτήσεις (10.115). Ζητείται η διερεύνηση για το αν η προκύπτουσα ανίχνευση ακμών είναι αναλλοίωτη σε προσθήκη μιας σταθερής φωτεινότητας στην περίπτωση 1Δ γραμμικής ακμής, δηλ. f(x) = δ(x), και στην περίπτωση βηματικής ακμής, δηλ. f(x) = $\delta^{(-1)}(x)$.

Λύση:

(a) Δίνονται τα φίλτρα h_e, h_o που χρησιμοποιεί το ενεργειακό μοντέλο στις εξισώσεις (10.114) και (10.115) και ζητάει να παρθούν δύο περιπτώσεις ιδανικών ακμών (μια γραμμική και μια βηματική ακμή) και να διερευνηθεί αν η προκύπτουσα ανίχνευση ακμών είναι αναλλοίωτη.

Επιλέγεται μια άρτια συνάρτηση για κρουστική απόκριση που είναι μια Gaussian με cos και μια περιττή συνάρτηση που θα είναι μια Gaussian με sin. Αυτές οι δύο θα χρησιμοποιηθούν στα μοντέλα των δύο φίλτρων εξ. (10.114) και είναι οι εξής: $h(x) = G_\sigma(x)\cos(\omega_c x)$ & $h_q(x) = G_\sigma(x)\sin(\omega_c x)$. Το ενεργειακό αυτό μοντέλο, με απλά λόγια λέει ότι γίνεται συνέλιξη με την είσοδο και μετά τετραγωνίζονται οι αποκρίσεις των δύο φίλτρων και στη συνέχεια προστίθενται.

Έστω ότι η συνάρτηση εισόδου είναι f(x) = δ(x), δηλαδή μια γραμμική ακμή η οποία είναι κεντραρισμένη στο x = 0. Τότε, υπολογίζεται η απόκριση του ενεργειακού μοντέλου για τη συνάρτηση f και για την συνάρτηση f + c (c = σταθερά). Σύμφωνα με το ενεργειακό μοντέλο, η θέση της ακμής βρίσκεται εκεί που έχει μέγιστο το ενεργειακό μοντέλο και ζητείται το κατά πόσο η εύρεση του σημείου της ακμής παραμένει αναλλοίωτη.

Ένα κοινό χαρακτηριστικό που έχουν και τα δύο μοντέλα είναι ότι τα φίλτρα είναι σε quadrature. Και στα δύο μοντέλα αποδεικνύεται ότι τα σημεία μεγίστου της ενέργειας E(x) συμπίπτουν με τα σημεία μεγίστης συμφωνίας τάσης.

Το κυριότερο πλεονέκτημα που έχει το ενεργειακό μοντέλο είναι ότι μπορεί να ανιχνεύσει και βηματικές ακμές (π.χ. $\delta^{(-1)}(x)$) και ακμές τύπου line (π.χ. δ(x)), όπου το πρώτο είδος ταιριάζει με το περιττό φίλτρο ($\psi_o(f) = f*h_o$) ενώ το δεύτερο είδος ταιριάζει με το άρτιο φίλτρο ($\psi_e(f) = f*h_e$).

Επιπροσθέτως, μπορούν να οριστούν δύο φίλτρα h_e, h_o με κάποια παράμετρο κλίμακας ή ακόμα και να διαφοροποιηθούν οι κλίμακες των βηματικών από εκείνες των γραμμικών ακμών. Οι Perona & Malik ανέλυσαν περισσότερο αυτό το μοντέλο και απέδειξαν ότι το μη-γραμμικό ενεργειακό φίλτρο E(x) μπορεί να ανιχνεύσει μεικτές ακμές. Αντιθέτως, αν χρησιμοποιηθεί ξεχωριστά μια οποιαδήποτε συλλογή γραμμικών φίλτρων υπάρχει ένας συνδυασμός c_1, c_2 στο ανωτέρω μοντέλο μεικτής ακμής ο οποίος δεν ανιχνεύεται από γραμμικά φίλτρα.

Επιπλέον σημαντικό θέμα είναι ότι ο τελεστής $\mathcal{E}$ ανίχνευσης ακμών είναι ταυτοδύναμος. Ο Ronse το ανέλυσε θεωρητικά και για γραμμικούς τελεστές και για μη-γραμμικούς ενεργειακούς τελεστές και βρήκε τις εξής συνθήκες για τα δύο φίλτρα h_e, h_o του ενεργειακού μοντέλου ώστε η ανίχνευση ακμών από αυτό να είναι αναλλοίωτη σε προσθήκη μιας σταθερής φωτεινότητας στην f(x) ή μιας γραμμικής ράμπας: $\int h_e(x)dx = \int x h_o(x)dx = 0$. Και πρότεινε το ακόλουθο ζεύγος φίλτρων που διατηρεί το διπλό αναλλοίωτο:

$$h_e(x) = G_\sigma(x)[\cos(\omega_c x) - \exp(-\omega_c^2\sigma^2/2)] \;\&\; h_o(x) = G_\sigma(x)[\sin(\omega_c x) - (\omega_c x)\exp(-\omega_c^2\sigma^2/2)]$$

(b) *Ζητείται η λύση της άσκησης 10.19 από Κεφ. 10 (έκδοση 6/2018).*

Εκφώνηση: ***(a)*** *Έστω μια 2Δ έγχρωμη εικόνα f(x, y) = (R(x, y), B(x, y), G(x, y)) και οι δύο μερικές παράγωγοι της $f_x = (R_x, B_x, G_x)$, $f_y = (R_y, B_y, G_y)$, όπου ισχύει ότι $u_x = \partial u/ \partial x$, $u_y = \partial u/ \partial y$. Επίσης, έστω η Di Zenzo μήτρα G_f στην εξ. (10.120) και ο ρυθμός μεταβολής $F(\theta) = |df|^2$ στην εξ. (10.123). Ζητείται να αποδειχθεί ότι, σε κάθε σημείο (x, y) η διεύθυνση θ_{max} του μεγίστου ρυθμού μεταβολής είναι λύση της εξ. (10.124).*

(b) *Ως αριθμητικό παράδειγμα, να θεωρηθεί μια έγχρωμη εικόνα f με σταθερά B, οι τρεις συνιστώσες R, G, B δεν μεταβάλλονται κατά μήκος του άξονα y, και οι συνιστώσες R, G έχουν τον ίδιο απόλυτο ρυθμό μεταβολής κατά μήκος του άξονα x αλλά με αντίθετο πρόσημα. Ζητείται η εύρεση των διευθύνσεων θ_{max} και θ_{min} του μεγίστου και ελαχίστου αντίστοιχα ρυθμού μεταβολής, καθώς και η ένταση ακμών.*

Λύση:

(b) Το gradient της f είναι ο τανυστής G = $\begin{bmatrix} g_{xx} & g_{xy} \\ g_{xy} & g_{yy} \end{bmatrix}$, $g_{xx} = f_x \cdot f_x$, $g_{xy} = f_x \cdot f_y$, $g_{yy} = f_y \cdot f_y$

Σε κάθε σημείο (x, y) για την εύρεση της κατεύθυνσης θ και της απόλυτης τιμής του μεγίστου ρυθμού μεταβολής της f αρκεί να βρεθεί η θ_{max} που μεγιστοποιείται η συνάρτηση:

$F(\theta) = |df|^2 = g_{xx}\cos^2(\theta) + 2g_{xy}\cos(\theta)\sin(\theta) + g_{yy}\sin^2(\theta)$

Η τιμή της θ_{max} προκύπτει ως λύση της εξίσωσης $\theta = (1/2)\arctan(2g_{xy}/(g_{xx} - g_{yy}))$

Ο υπολογισμός γίνεται με βάση την εξίσωση 10.123 η οποία ορίζει το F(θ). Η συνάρτηση της εικόνας είναι διανυσματική συνάρτηση, δηλαδή σε κάθε pixel (x, y) υπάρχει ένα διάνυσμα 3Δ. Η συνάρτηση αυτή παίρνει το μέτρο αυτού του διανύσματος. Στην εξίσωση αυτή η παράγωγος πρέπει κάπως να αλλάξει μιας και είναι διανυσματική η συνάρτηση, οπότε γίνεται λόγος για ένα διαφορικό που θα έχει κάποια ευκλείδεια νόρμα η οποία δίνεται (βλ. θεωρία γύρω από τις εξισώσεις 10.120 κλπ. δηλ. ενότητα 10.6.7). Αυτή η νόρμα του διαφορικού στο τετράγωνο, δίνει την ένταση των ακμών.

(a) Αρχικά, παραγωγίζεται η F: $F'(\theta) = 2g_{xy}\cos 2\theta - (g_{xx} - g_{yy})\sin 2\theta$

Θέτοντας $dF/d\theta = 0$ προκύπτει $\tan(2\theta_o) = [2g_{xy}/(g_{xx} - g_{yy})]$ (*)

Η εξίσωση αυτή έχει δύο λύσεις σε ορθογώνιες διευθύνσεις θ_o και $\theta_o + (\pi/2)$, εκ των οποίων η μία είναι η θ_{max}.

(b) Από τις τρεις υποθέσεις για τα έγχρωμα κανάλια, ισχύει ότι:

$B_x = B_y = 0$, $R_y = G_y = B_y = 0$, $R_x = -G_x$

Οπότε, προκύπτει ότι $f_y = 0$, $g_{xy} = g_{yy} = 0$, $g_{xx} = 2|R_x|^2$.

Άρα, οι λύσεις της (*) είναι $\theta_o = 0, \pi/2$, από τις οποίες $\theta_{max} = 0$ και $\theta_{min} = \pi/2$.

Επίσης, η ένταση ακμών $\sqrt{F(\theta_{max})} = \sqrt{2}|R_x|$.

<u>Άσκηση 2.2 (Γενικευμένα Φίλτρα Εικόνων σε Πλέγματα)</u>

(a) *Να αποδειχθεί ότι το ακτινικό (radial) opening* $a_{rad}(X) = \bigcup_{0\leq\theta\leq\pi} X \circ L_\theta$, $X \subseteq \mathbb{R}^2$, *όπου* L_θ *είναι ένα ευθύγραμμο τμήμα σταθερού μήκους σε κατεύθυνση θ ως προς την οριζόντια, δηλ. ένα lattice opening, δηλ. αυξάνον, μη-επεκτατικό και ταυτοδύναμο.*

(b) *Να αποδειχθεί ότι το area opening* $a_n(X)$ *ψηφιακών σχημάτων* $X \subseteq \mathbb{Z}^2$ *με κατώφλι εμβαδού* $n \geq 2$ *μπορεί να υπολογισθεί και ως ένωση απλών openings* $a_n(X) = \bigcup_{card(A)=n} X \circ A$ *με όλα τα συνεκτικά υποσύνολα A ενός* $n\times n$ *τετραγώνου που αποτελούνται από n σημεία και δεν είναι μετατοπίσεις το ένα του άλλου.*

<u>Λύση</u>:

(a) *Θεώρημα 1*: Το ακτινικό (radial) opening: $a_{rad}(X) = \bigcup_{0\leq\theta\leq\pi} X \circ L_\theta$, $X \subseteq \mathbb{R}^2$ είναι ένα lattice opening.

Απόδειξη: Θα δειχθεί ότι είναι αυξάνον, μη-επεκτατικό και ταυτοδύναμο.

Αυξάνον: Έστω $X \leq Y$. Αφού ο τελεστής $\circ$ είναι αυξάνον, θα ισχύει ότι:

$X \circ L_\theta \leq Y \circ L_\theta \Rightarrow \bigcup_\theta X \circ L_\theta \leq \bigcup_\theta Y \circ L_\theta \Leftrightarrow a_{rad}(X) \leq a_{rad}(Y)$

Μη-επεκτατικό: Αφού ο τελεστής $\circ$ είναι μη-επεκτατικός, θα ισχύει ότι:

$X \leq X \circ L_\theta \Rightarrow \bigcup_\theta X \circ L_\theta \leq \bigcup_\theta X \Leftrightarrow a_{rad}(X) \leq X$

Ταυτοδύναμο: Ισχύει ότι:

$a_{rad}(a_{rad}(X)) = \bigcup_{\theta_1}(\bigcup_{\theta_2} X \circ L_{\theta_2}) \circ L_{\theta_1} = \bigcup_{\theta_1}((X \circ L_{\theta_1}) \circ L_{\theta_1}) \bigcup_{\theta_1\neq\theta_2}(X \circ L_{\theta_2})$

Όμως, ο τελεστής $\circ$ είναι ταυτοδύναμος, δηλ. $(X \circ L_\theta) \circ L_\theta = X \circ L_\theta$

Οπότε, η παραπάνω σχέση, γράφεται:

$a_{rad}(a_{rad}(X)) = \bigcup_{\theta_1}(X \circ L_{\theta_1}) \bigcup_{\theta_1\neq\theta_2}(X \circ L_{\theta_2}) \geq a_{rad}(X)$

Όμως, επειδή αποδείχθηκε πριν ότι ο τελεστής είναι μη επεκτατικός, τότε πρέπει να ισχύει η ισότητα, δηλ. $a_{rad}(a_{rad}(X)) = a_{rad}(X)$

(b) *Θεώρημα 2*: Το opening εμβαδού μπορεί να υπολογιστεί και ως: $a_n(X) = \bigcup_{card(A)=n} X \circ A$, όπου A είναι όλα τα υποσύνολα n σημείων ενός τετραγώνου n×n όπου κανένα δεν αποτελεί μετατόπιση ενός άλλου.

Απόδειξη: Το opening εμβαδού ορίζεται ως: $a_n(X) = \coprod_{area(X_j) \geq n} X_j$

Θα δειχθεί ότι: $a_n(X) = \bigcup_{card(A)=n} X \circ A$

Όμως, $X = \coprod_i X_i$

Αρκεί, να δειχθεί ότι: $a_n(X) = \bigcup_{card(A)=n}(\coprod_i X_i) \circ A = \coprod_i \bigcup_{card(A)=n} X_i \circ A$

Ισχύει ότι: $X_i \circ A = (X_i \ominus A) \oplus A$

Όμως, $X \ominus A = \{z: X_{+z} \subseteq A\}$

Το A έχει area n, συνεπώς είναι προφανές ότι αν $area(X_i) < n$, τότε: $X_i \ominus A = \emptyset$

Άρα, αρκεί να δειχθεί ότι: $a_n(X) = \coprod_{area(X_i) \geq n} \bigcup_{card(A)=n} X_i \circ A$

Το X_i όμως είναι connected component, αφού το X προκύπτει από disjoint union των X_i. Αφού το X_i είναι connected component και έχει $area(X_i) \geq n$, μπορεί να γραφτεί ως: $X_i = \bigcup_{area(X_k)=n} X_k$

Συνεπώς, αρκεί να δειχθεί ότι:

$a_n(X) = \coprod_{area(X_i) \geq n} \bigcup_{card(A)=n} \bigcup_{area(X_k)=n} X_k \circ A = \coprod_{area(X_i) \geq n} \bigcup_{area(X_k)=n} \bigcup_{card(A)=n} X_k \circ A$

Αφού προκύπτουν όλα τα δυνατά connected σχήματα μήκους n από το A, θα υπάρχει $A_i = X_k$ $\forall k$ και από την μη-επεκτατικότητα του opening προκύπτει ότι: $\bigcup_{card(A)=n} X_k \circ A = X_k$.

Έτσι, $a_n(X) = \coprod_{area(X_i) \geq n} \bigcup_{area(X_k)=n} X_k = \coprod_{area(X_i) \geq n} X_i$

Άσκηση 2.3 (Scale Spaces Μορφολογικών Φίλτρων και Μετ/σμός Κλίσης)

Έστω μια 2Δ πραγματική εικόνα f(x, y). Έστω, επίσης, η συνάρτηση που δημιουργείται από πολύ-κλιμακωτές σταθμισμένες dilations, δηλ. max-plus συνελίξεις της αρχικής εικόνας f(x, y) με μια πολυκλιμακωτή παραβολή $k_t(x, y)$, $t \geq 0$: $\delta(x, y, t) = f(x, y) \oplus k_t(x, y)$, $k_t(x, y) \triangleq -\frac{x^2+y^2}{2t}$

(a) *Να αποδειχθεί ότι η πολυκλιμακωτή παραβολή ικανοποιεί τον ακόλουθο νόμο άθροισης κλιμάκων:*

$k_t \oplus k_s = k_{t+s}$, $t, s \geq 0$

Υπόδειξη: να χρησιμοποιηθεί μετ/σμός κλίσης.

(b) *Να αποδειχθεί ότι η scale-space συνάρτηση δ(x, y, t) ικανοποιεί την ακόλουθη μερική διαφορική εξίσωση: $\frac{\partial \delta}{\partial t}(x, y, t) = \frac{1}{2}||\nabla\delta||^2$ με αρχική συνθήκη δ(x, y, 0) = f(x, y), χρησιμοποιώντας τις εξής δύο μεθόδους και βρίσκοντας ότι οδηγούν στο ίδιο αποτέλεσμα:*

(b1) *Οι εξισώσεις (11.18) και (11.24) του Κεφ. 11 (ΟΥ 2005), αφού επεκταθούν για 2Δ εικόνες, χρησιμοποιώντας την ιδιότητα additive semigroup των πολυκλιμακωτών dialations με κοίλες συναρτήσεις από το μέρος (a) και ανάπτυξη Taylor της διαφοράς μεταξύ των dilations σε κλίμακες t και t + dt με χωρικές παραγώγους.*

(b2) *Μετ/σμό Κλίσης: Εξ. (11.42).*

Λύση:

(a) Είναι προφανές ότι η συνάρτηση $k(x, y) = -\frac{x^2+y^2}{2}$ έχει κυρτό υπογράφο και συνεπώς, από το θεώρημα του βιβλίου, υπάρχει η ιδιότητα της υποομάδας ως προς την πράξη ⊕ για τις συναρτήσεις της οικογένειας $k_t(x, y) = tk(x/t, y/t)$. Επομένως, $k_t \oplus k_s = k_{t+s}$.

Εναλλακτικά, θα μπορούσε να δειχθεί ότι:

$k_t \oplus k_s = \bigvee_{(a,b)\in\mathbb{R}^2} k_t(x, y) - a\cdot s\cdot k(a/r) - b\cdot s\cdot k(b/r) = k_{t+s}$

(b1) Ισχύει ότι: $\delta(x, y, s) = (f\oplus k_s)(x, y) = \bigvee_{(a,b)\in\mathbb{R}^2} f(x + a, y + b) - k_s(x, y)$

Είναι: $\frac{\partial \delta}{\partial s}(x, y, s) = \lim_{r\to 0^+} \frac{\delta(x,y,s+r)-\delta(x,y,s)}{r}$

Όμως, αφού $k_s \oplus k_r = k_{s+r}$ θα ισχύει ότι: $\delta(x, y, s+r) = \delta(x, y, s)\oplus k_r(x, y)$

Άρα, $\frac{\partial \delta}{\partial s}(x, y, s) = \lim_{r\to 0^+} \frac{\delta(x,y,s)\oplus k_r(x,y)-\delta(x,y,s)}{r} = \lim_{r\to 0^+} \frac{\bigvee_{(a,b)\in\mathbb{R}^2} \delta(x+a,y+b,s)-k_r(x,y)-\delta(x,y,s)}{r}$

Έπειτα, το ανάπτυγμα κατά Taylor της δ(x+a, y+b, s) γύρω από το σημείο (x, y) θα είναι:

$\delta(x+a, y+b, s) = \delta(x, y, s) + \frac{\partial \delta}{\partial x}(x, y, s)\cdot a + \frac{\partial \delta}{\partial y}(x, y, s)\cdot b + ||a, b|| \circ (||a, b||)$

Όμως, επειδή τα a, b είναι μικρές μετατοπίσεις, στο όριο μπορεί να αγνοηθεί ο όρος $||a, b|| \circ (||a, b||)$

Έτσι, $\frac{\partial\delta}{\partial s}(x, y, s) = \lim_{r\to 0^+} \frac{\bigvee_{(a,b)\in\mathbb{R}^2} \delta_x \cdot a + \delta_y \cdot b - \frac{a^2+b^2}{2r}}{r}$

Αφού ο αριθμητής είναι διαχωρίσιμη συνάρτηση ως προς τις μεταβλητές a, b η έκφραση μπορεί να γραφτεί ισοδύναμα:

$$\frac{\partial\delta}{\partial s}(x, y, s) = \lim_{r\to 0^+} \left(\bigvee_a \frac{\delta_x \cdot a - \frac{a^2}{2r}}{r} + \bigvee_b \frac{\delta_y \cdot b - \frac{b^2}{2r}}{r}\right) = \lim_{r\to 0^+} \left(\bigvee_a \frac{\delta_x \cdot a \cdot 2 \cdot r - a^2}{2r^2} + \bigvee_b \frac{\delta_y \cdot b \cdot 2 \cdot r - b^2}{2r^2}\right)$$

Αναζητείται το a για το οποίο μεγιστοποιείται η συνάρτηση $g(a) = \delta_x \cdot a \cdot 2 \cdot r - a^2$

Τα πιθανά ακρότατα είναι τα a για τα οποία: $g'(a) = 0 \Leftrightarrow \delta_x \cdot 2 \cdot r - 2 \cdot a = 0 \Leftrightarrow a = \delta_x \cdot r$

Για να βρεθεί αν όντως υπάρχει μέγιστο, υπολογίζεται η δεύτερη παράγωγος: $g''(a) = -2 < 0$

Αφού $g''(a) < 0$, η g' είναι πάντα φθίνουσα άρα δεξιά από το σημείο μηδενισμού της είναι αρνητική και αριστερά είναι θετική. Οπότε, η g είναι αύξουσα αριστερά από το πιθανό ακρότατο και φθίνουσα αριστερά, δηλαδή όντως το πιθανό ακρότατο είναι μέγιστο.

Αντίστοιχα, προκύπτει: $b = \delta_y \cdot r$

Άρα, $\frac{\partial\delta}{\partial s}(x, y, s) = \lim_{r\to 0^+} \left(\frac{\delta_x^2 \cdot 2r^2 - \delta_x^2 \cdot r^2}{2r^2} + \frac{\delta_y^2 \cdot 2r^2 - \delta_y^2 \cdot r^2}{2r^2}\right) \Leftrightarrow \frac{\partial\delta}{\partial s}(x, y, s) = \frac{1}{2}(\delta_x^2 + \delta_y^2) = \frac{||\nabla\delta||^2}{2}$

(b2) *Θεώρημα*: Η scale-space συνάρτηση δ(x, y, t) ικανοποιεί τη μερική διαφορική εξίσωση:

$\frac{\partial\delta}{\partial t}(x, y, t) = K_v\left(\frac{\partial\delta}{\partial x}, \frac{\partial\delta}{\partial y}\right)$

Παρακάτω θα υπολογιστεί ο μετασχηματισμός κλίσης $K_v(a, b)$.

Είναι $K_v(a, b) = \bigvee_{x\in\mathbb{R}^2} k(x, y) - a_1 x - a_2 y$

Αρκεί να μεγιστοποιηθεί η συνάρτηση: $g(x, y) = -\frac{x^2+y^2}{2} - a_1 x - a_2 y$

Αφού η συνάρτηση είναι διαχωρίσιμη, μπορούν ισοδύναμα να μεγιστοποιηθούν οι συναρτήσεις:

$g_1(x) = -\frac{x^2}{2} - a_1 x$

$g_2(x) = -\frac{y^2}{2} - a_2 y$

Η συνάρτηση $g_1(x)$ έχει πιθανά ακρότατα στα σημεία (x, y) για τα οποία ισχύει:

$g_1'(x) = 0 \Leftrightarrow x = -a_1$

Αφού όμως, $g''(x) = -1 < 0$, ακριβώς όπως και πριν προκύπτει ότι το σημείο είναι όντως τοπικό μέγιστο.

Ακριβώς αντίστοιχα προκύπτει ότι: $y = -a_2$

Άρα: $K_v(a, b) = -\frac{a_1^2+a_2^2}{2} + a_1^2 + a_2^2 = \frac{a_1^2}{2} + \frac{a_2^2}{2}$

Σύμφωνα με το αρχικό θεώρημα, θα ισχύει ότι: $\frac{\partial\delta}{\partial t} = K_v\left(\frac{\partial\delta}{\partial x}, \frac{\partial\delta}{\partial y}\right) = \frac{1}{2}\left(\frac{\partial\delta^2}{\partial x} + \frac{\partial\delta^2}{\partial y}\right) = \frac{||\nabla\delta||^2}{2}$

<u>Άσκηση 2.4 (Γενικευμένοι Γραμμικοί Χώροι Κλίμακας)</u>

Στην εκφώνηση της άσκησης αναλύεται η θεωρία. Ζητείται να βρεθεί το φ(x) και τα multiscale kernels kt(x) σε closed analytic form για τις περιπτώσεις:

(a) *p = 1*

(b) *p = 2*

(c) *p = ∞*

(d) *Να εξηγηθεί η σημασία των δύο ελεύθερων παραμέτρων α και p στην γενική περίπτωση.*

<u>Λύση</u>:

Η απάντηση δίνεται με βάση το 2ο paper και αντί για p χρησιμοποιείται το σύμβολο α. Ο σχολιασμός για τις παραμέτρους a και p γίνεται παράλληλα με την ανάλυση της κάθε περίπτωσης.

Σύμφωνα με αυτό λοιπόν, ερευνάτε το πως η συναρτησιακή μορφή των scale-space φίλτρων καθορίζεται από τον αριθμό των a priori συνθηκών. Συγκεκριμένα, αν υποτεθεί ότι τα scale-space φίλτρα είναι γραμμικά, ισοτροπικά, συνελικτικά φίλτρα, τότε δύο συνθήκες (recursivity & scale-invariance) καταλήγουν να περιορίσουν την οικογένεια αυτών των φίλτρων σε εκείνα που βασίζονται μόνο σε μια παράμετρο η οποία καθορίζει το ποιοτικό σχήμα του φίλτρου. Τα Gaussian filters ανταποκρίνονται σε μια συγκεκριμένη τιμή αυτής της παραμέτρου. Για άλλες τιμές τα φίλτρα επιδεικνύουν ένα πιο σύνθετο pattern.

Από το paper (pg. 694) προκύπτει ότι μπορεί να προκύψει το “unscaled” kernel φ θέτοντας t = 1.

Οπότε, $\varphi(x) = k_t(x) = \frac{1}{2}\int_{\mathbb{R}} \cos(\omega x) e^{-a|\omega|^{\alpha}} d\omega$ και $\psi(t) = t^{1/a}$

(a) Για την τιμή α = 1, θα ισχύει ότι ο αντίστροφος Fourier είναι εφικτός σε αναλυτική κλειστή μορφή. Ένα τέτοιο παράδειγμα αποτελεί το εξής:

Για α = 1, προκύπτει $\widetilde{\varphi}(\omega) = e^{-|\omega|}$ που είναι ο Fourier Μετασχηματισμός του Cauchy density.

Έτσι, το φ και οι kernel functions θα είναι: $\varphi(x) = \frac{1}{\pi}\frac{1}{1+x^2}$ και $k_t(x) = \frac{1}{2}\frac{t}{t^2+x^2}$

(b) Για την τιμή α = 2, προκύπτει ο Gaussian kernel:

$\varphi(x) = \frac{1}{2}\int_{\mathbb{R}} \cos(\omega x) e^{-a\omega^2} d\omega = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-x^2/2\sigma^2}$, όπου $\sigma^2 = 1/2\alpha$ και $\psi(t) = \sqrt{t}$

Αυτό το αποτέλεσμα είναι γνωστό στη βιβλιογραφία αλλά από την ανάλυση του paper μπορεί να προκύψει ότι η recursivity & η scale-invariance δεν αρκούν για να διαχωρίσουν τον Gaussian kernel.

(c) Ένα ακόμη παράδειγμα για το οποίο το αναλυτικό αποτέλεσμα είναι διαθέσιμο αφορά την οριακή τιμή α = ∞.

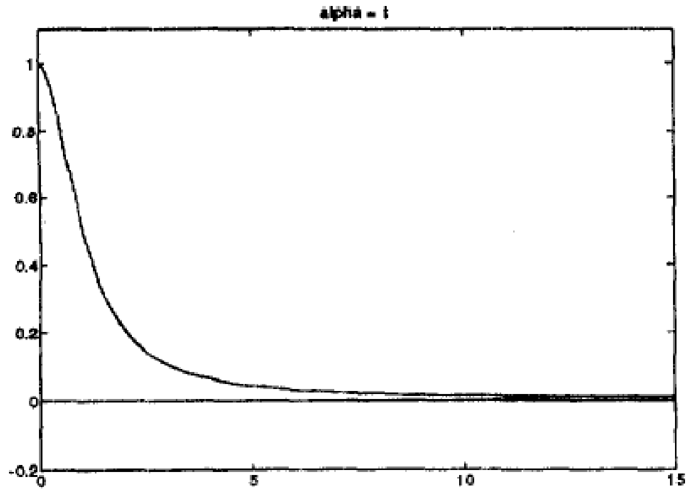
Από την αρχική σχέση του φ, προκύπτει ότι: $\widetilde{\varphi}(\omega) \to \begin{cases} 1, \alpha\nu\ |\omega| < 1 \\ 0, \alpha\nu\ |\omega| > 1 \end{cases}$ καθώς $\alpha \to \infty$.

Αξίζει να σημειωθεί ότι μέχρι ένα παράγοντα 1/π, αυτός είναι ο Fourier Μετασχηματισμός του sinc(x).

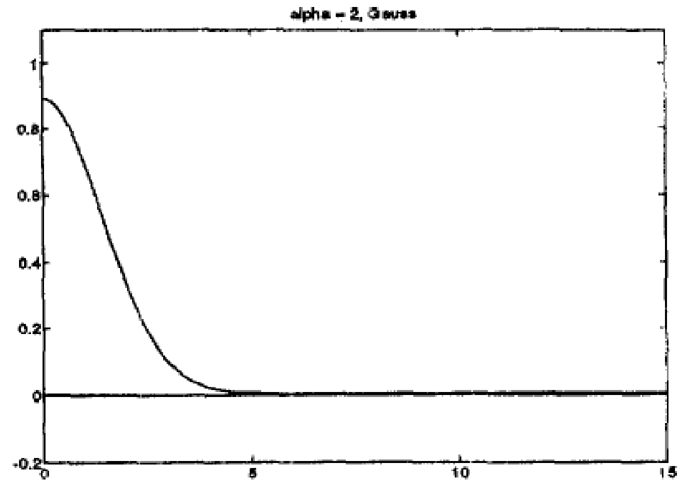
(d) Για τις περισσότερες τιμές του α, πρέπει να ληφθεί υπόψιν ο numerical Fourier inversion για να διατηρηθεί μια αναλυτική περιγραφή του αντίστοιχου kernel φίλτρου. Ωστόσο, η functional μορφή του Fourier Μετασχηματισμού παρέχει μερικές global πληροφορίες σχετικά με τον kernel. Είναι αναμενόμενο ότι για α > 2, το φ είναι τουλάχιστον δύο φορές διαφορίσιμο και ότι φ''(0) = 0. Αυτό υπονοεί ότι τη δεύτερη φορά η φ εξαλείφεται και επομένως το φ δεν μπορεί να είναι παντού θετικό. Έτσι, όταν α > 2, ο φ kernel έχει zero-crossings.

Παρακάτω φαίνονται και οι σχετικές εικόνες:

Για α = 1:



Για α = 2:



Άσκηση 2.5 (Ανάλυση Σχήματος με Καμπυλότητα)

(a) *Αν μετασχηματιστεί μια ομαλή καμπύλη $\vec{C}(p)$ σε μια νέα καμπύλη $\overrightarrow{C^*}(p)$ μέσω μετ/σμού ομοιόμορφης κλιμάκωσης, δηλ. $\vec{C}(p) \mapsto \overrightarrow{C^*}(p) = \lambda\vec{C}(p)$, $\lambda > 0$ να αποδειχθεί ότι το μήκος τόκου s και η καμπυλότητα κ θα υποστούν τις εξής αλλαγές: $s \mapsto s^* = \lambda s$, $\kappa(s) \mapsto \kappa^*(s^*) = \kappa(s)/\lambda$.*

(b) *Έστω μια επίπεδη ομαλή καμπύλη με παραμετρική παράσταση $\vec{C}(p) = (x(p), y(p)) \in \mathbb{R}^2$, $p \in [0, 1]$, όπου (x(p), y(p)) είναι οι καρτεσιανές συντεταγμένες των σημείων της. Έστω s = s(p) το μήκος τόξου, και έστω (r(p), φ(p)) οι πολικές συντεταγμένες των σημείων της.*

(b1) *Να αποδειχθεί ότι, σε πολικές συντεταγμένες η καμπυλότητα της καμπύλης ισούται με*

$$\kappa(p) = \frac{2(r')^2\varphi' + r^2(\varphi')^3 + rr'\varphi'' - rr''\varphi'}{[(r')^2 + r^2(\varphi')^2]^{3/2}}$$

(b2) *Ζητείται η εφαρμογή του αποτελέσματος (b1) στην καμπύλη $r(p) = \alpha\cdot\varphi(p)$, α = constant > 0, $0 \le \varphi \le \infty$, $\varphi' > 0$, που είναι γνωστή ως 'σπιράλ του Αρχιμήδη'. Επίσης, να βρεθεί η καμπυλότητά της (με όσο το δυνατόν απλούστερη έκφραση). Τέλος, ζητείται να σχεδιασθεί προσεγγιστικά η καμπύλη για την περίπτωση $\alpha = 1/\pi$, $\varphi(p) = 3\pi p$, $0 \le p \le 1$.*

Λύση:

Στο μάθημα έχει ειπωθεί ότι η συνάρτηση καμπυλότητας παραμένει αναλλοίωτη αν κάνει μια περιστροφή και μετατόπιση. Τώρα μπαίνει ένας τρίτος Μ/Σ, η κλιμάκωση, και ζητείται να βρεθεί τι θα συμβεί στη συνάρτηση καμπυλότητας εάν παρθεί μια νέα καμπύλη που προκύπτει από την προηγούμενη με μια ομοιόμορφη κλιμάκωση (δηλ. εάν ληφθούν υπόψιν όλα τα διανύσματα των θέσεων της αρχικής καμπύλης και πολλαπλασιαστούν με μια σταθερή θετική παράμετρο λ).

(a) Ομοιόμορφη κλιμάκωση καμπύλης: $\vec{C}(p) = (x(p), y(p)) \mapsto \overrightarrow{C^*}(p) = (x^*(p), y^*(p)) = \lambda\vec{C}(p)$ με $\lambda > 0$.

Άρα, $x^* = \lambda x$ και $y^* = \lambda y$.

Επομένως, $\frac{dx^*}{dp} = \lambda x'$, $\frac{dy^*}{dp} = \lambda y'$

$u^*(p) = \sqrt{(\frac{dx^*}{dp})^2 + (\frac{dy^*}{dp})^2} = \lambda\sqrt{(x')^2 + (y')^2} = \lambda u(p)$, όπου $x' = \frac{dx}{dp}$ και $y' = \frac{dy}{dp}$

Οπότε, $s^*(p) = \int_0^q u^*(q)dq = \lambda\int_0^q u(q)dq = \lambda s(p)$

Ομοίως, αποδεικνύεται ότι $\theta^*(p) = \theta(p)$ και άρα $\kappa(p) = \frac{d\theta^*}{ds^*} = \frac{d\theta}{\lambda ds} = \frac{\kappa}{\lambda}$

(b) Σε καρτεσιανές συντεταγμένες, η καμπυλότητα δίνεται από τη σχέση: $k(s) = \frac{y''x' - x''y'}{[(x')^2 + (y')^2]^{3/2}}$

Άρα: $k_*(s) = \frac{y_*''x_*' - x_*''y_*'}{[(x_*')^2 + (y_*')^2]^{3/2}}$

Όμως, αφού η καμπύλη έχει κλιμακωθεί κατά λ, κατά λ έχουν κλιμακωθεί και οι τετμημένες και τεταγμένες της.

Έτσι, $x_* = \lambda x$, $y_* = \lambda y \Rightarrow x'_* = \lambda x'$, $x''_* = \lambda x''$, $y'_* = \lambda y'$, $y''_* = \lambda y''$

Και, $k_*(s) = \frac{\lambda^2(y''x'-x''y')}{[\lambda^2(x')^2+(y')^2]^{3/2}} = \frac{k}{\lambda}$

Για το μήκος του τόξου, ισχύει ότι: $s = \int_a^b \sqrt{1+(\frac{dy}{dx})^2}dx$

Έτσι, $s_* = \int_a^b \sqrt{1+(\frac{dy_*}{dx_*})^2}dx_*$

Όμως, $dx_* = \lambda dx$, $dy_* = \lambda dy$

Συνεπώς, $s_* = \lambda s$

(b1) Είναι γνωστό ότι σε καρτεσιανές συντεταγμένες είναι: $k(p) = \frac{y''x'-x''y'}{[(x')^2+(y')^2]^{3/2}}$

Επίσης, $x = r\cos\varphi$, $y = r\sin\varphi$

Από τα παραπάνω προκύπτει: $x' = r'\cos\varphi - r'\sin\varphi\cdot\varphi'$

$x'' = r''\cos\varphi + r'(-\sin\varphi)\varphi' - r'\sin\varphi\cdot\varphi' - r(\sin\varphi\cdot\varphi')' = r''\cos\varphi - 2r'\varphi'\sin\varphi - r[(\varphi')^2\cos\varphi + \sin\varphi\cdot\varphi''] =$

$= r''\cos\varphi - 2r'\sin\varphi\cdot\varphi' - r\cos\varphi(\varphi')^2 - r\sin\varphi\cdot\varphi''$

Ακόμη, $y' = r'\sin\varphi + r\cos\varphi\cdot\varphi'$

Επίσης, $y'' = r''\sin\varphi + r'\cos\varphi\cdot\varphi' + r'\cos\varphi\cdot\varphi' + r(\cos\varphi\cdot\varphi')' =$

$= r''\sin\varphi + 2r'\cos\varphi\cdot\varphi' + r(-\sin\varphi\cdot(\varphi')^2 + \varphi''\cos\varphi) =$

$= r''\sin\varphi + 2r'\cos\varphi\cdot\varphi' - r\sin\varphi\cdot(\varphi')^2 + r\cos\varphi\cdot\varphi''$

Έτσι, $y''x' = (r''\sin\varphi + 2r'\cos\varphi\cdot\varphi' - r\sin\varphi\cdot(\varphi')^2 + r\cos\varphi\cdot\varphi'')\cdot(r'\cos\varphi - r\sin\varphi\cdot\varphi') =$

$= r''\cdot r'\cdot\sin\varphi\cdot\cos\varphi - r''\cdot r\cdot\sin^2\varphi\cdot\varphi' + 2\cdot(r')^2\cdot\cos^2\varphi\cdot(\varphi') - 2\cdot r'r\cdot\cos\varphi\cdot\sin\varphi\cdot(\varphi')^2 - r'\cdot r\cdot\sin\varphi\cdot\cos\varphi\cdot(\varphi')^2 + r^2\cdot\sin^2\varphi\cdot(\varphi')^3 + r'\cdot r\cdot\cos^2\varphi\cdot\varphi'' - r^2\cdot\sin\varphi\cdot\cos\varphi\cdot\varphi'\cdot\varphi''$

Αντίστοιχα, $x''y' = (r''\cdot\cos\varphi - 2\cdot r'\cdot\sin\varphi\cdot\varphi' - r\cdot\cos\varphi\cdot(\varphi')^2 - r\cdot\sin\varphi\cdot\varphi'')\cdot(r'\sin\varphi + r\cos\varphi\cdot\varphi') =$

$= r''\cdot r'\cdot\cos\varphi\cdot\sin\varphi + r''\cdot r\cdot\cos^2\varphi\cdot\varphi' - 2\cdot(r')^2\cdot\sin^2\varphi\cdot\varphi' - 2\cdot r'\cdot r\cdot\cos\varphi\cdot\sin\varphi\cdot(\varphi')^2 - r'\cdot r\cdot\cos\varphi\cdot\sin\varphi\cdot(\varphi')^2 - r^2\cdot\cos^2\varphi\cdot(\varphi')^3 - r'\cdot r\cdot\sin^2\varphi\cdot\varphi'' - r^2\cdot\cos\varphi\cdot\sin\varphi\cdot\varphi'\cdot\varphi''$

Επίσης, $(x')^2 + (y')^2 = (r'\cdot\cos\varphi - r\cdot\sin\varphi\cdot\varphi')^2 + (r'\cdot\sin\varphi + r\cdot\cos\varphi\cdot\varphi')^2 =$

$= (r')^2\cdot\cos^2\varphi - 2\cdot r'\cdot r\cdot\cos\varphi\cdot\sin\varphi\cdot\varphi' + r^2\cdot\sin^2\varphi\cdot\varphi' + (r')^2\cdot\sin^2\varphi + 2\cdot r'\cdot r\cdot\sin\varphi\cdot\cos\varphi\cdot\varphi' + r^2\cdot\cos^2\varphi\cdot(\varphi')^2 =$

$= (r')^2 + r^2(\varphi')^2$

Σύμφωνα με τα παραπάνω λοιπόν, προκύπτει:

$$k(p) = \frac{y''x'-x''y'}{[(x')^2+(y')^2]^{3/2}} = \frac{2(r')^2\varphi' + r^2(\varphi')^3 + rr'\varphi'' - rr''\varphi'}{[(r')^2+r^2(\varphi')^2]^{3/2}}$$

(b2) $k(p) = \frac{2(r')^2\varphi' + r^2(\varphi')^3 + rr'\varphi'' - rr''\varphi'}{[(r')^2 + r^2(\varphi')^2]^{3/2}}$

Ισχύει ότι: r(p) = a·φ(p) , r'(p) = a·φ'(p) , r''(p) = a·φ''(p)

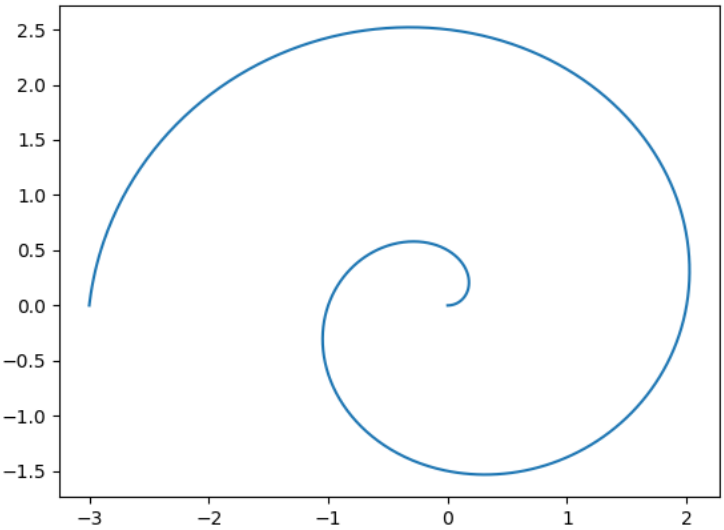
Επομένως: $k(p) = \frac{2a^2(\varphi')^2\varphi' + r^2(\varphi')^3 + \alpha^2\varphi\varphi'\varphi'' - a^2\varphi\varphi''\varphi'}{[(r')^2 + r^2(\varphi')^2]^{3/2}} = \frac{2a^2(\varphi')^3 + a^2\varphi(\varphi')^3}{[a^2(\varphi')^2\cdot(1+\varphi^2)]^{3/2}} = \frac{a^2(\varphi')^3\cdot(2+\varphi)}{a^3(\varphi')^3\cdot(1+\varphi^2)^{3/2}} \Rightarrow$

$k(p) = \frac{2+\varphi}{a\cdot(1+\varphi^2)^{3/2}}$

Μελέτη περιπτώσεων:

$$\begin{cases} r(p) = \frac{1}{\pi}\varphi(p) \\ \varphi(p) = 3\pi p, p \in [0,1] \end{cases} \Leftrightarrow \begin{cases} r(p) = 3p, p \in [0,1] \\ \varphi(p) = 3\pi p, p \in [0,1] \end{cases}$$

Η γραφική παράσταση φαίνεται ακολούθως:



<u>*Resources*</u>*:*

1. *C. Ronse, "On idempotence and related requirements in edge detection," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 15, no. 5, pp. 484-491, May 1993, doi: 10.1109/34.211468.*
2. *E. J. Pauwels, L. J. van Gool, P. Fiddelaers and T. Moons, "An extended class of scale-invariant and recursive scale space filters," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 17, no. 7, pp. 691-701, July 1995, doi: 10.1109/34.391411.*
3. *Maragos_CV_Book2018_chapter**