

Όραση Υπολογιστών

Εργαστήριο 1

Αναστάσιος Στέφανος Αναγνώστου
03119051

Σπυρίδων Παπαδόπουλος
03119058

7 Απριλίου 2023

1 Μέρος 1ο

1.1 Δημιουργία Εικόνων Εισόδου

Οι εικόνες εισόδου δημιουργούνται διαβάζοντας την δεδομένη εικόνα 'edgetest_23.png' και προσθέτοντας θόρυβο σε αυτήν. Προστίθεται αντιστοίχως θόρυβος με PSNR 20dB και θόρυβος με PSNR 10dB. Αξιοσημείωτο είναι, ότι ο πρώτος θόρυβος είναι στην πραγματικότητα λιγότερος από τον δεύτερον. Το PSNR ορίζεται ως εξής:

$$PSNR = 20 \log_{10} \left(\frac{I_{max} - I_{min}}{\sigma_n} \right) (dB),$$
$$I_{max} = \max_{x,y} I(x,y), I_{min} = \min_{x,y} I(x,y)$$
(1.1)

Φαίνεται ότι από τον ορισμό μπορεί να ληφθεί η μεταβλητότητα του θορύβου, βάσει της οποίας τελικά αυτός ορίζεται. Από την έκφραση της μεταβλητότητας, μάλιστα, φαίνεται ότι, όσο μεγαλύτερο το PSNR σε dB τόσο μικρότερη είναι τελικά η μεταβλητότητα του θορύβου, πράγμα που επαληθεύει τα προαναφερθέντα.

$$\sigma_n = \frac{(I_{max} - I_{min})}{10^{\frac{PSNR}{20}}}$$
(1.2)

1.2 Υλοποίηση Αλγορίθμων Ανίχνευσης Ακμών

Η συνάρτηση που θα αναπτυχθεί εφαρμόζει είτε γραμμική είτε μη γραμμική μέθοδο ανίχνευσης ακμών.

1.2.1 Γραμμική Μέθοδος

Η μεν γραμμική μέθοδος επιχειρεί την προσέγγιση της λαπλασιανής της εξομαλυνμένης εικόνας χρήσει γραμμικών φίλτρων. Συγκεκριμένα, σύμφωνα με την γραμμική μέθοδο:

$$L = \nabla^2 (G_\sigma * I) = (\nabla^2 G_\sigma) * I$$
(1.3)

1.2.2 Μη Γραμμική Μέθοδος

Η δε μη γραμμική μέθοδος επιχειρεί την προσέγγιση της λαπλασιανής της εξομαλυμένης εικόνας χρήσει μορφολογικών τελεστών. Συγκεκριμένα, σύμφωνα με την μη γραμμική μέθοδο:

$$\begin{aligned} I_\sigma &= G_\sigma * I, \\ L &= I_\sigma \oplus B + I_\sigma \ominus B - 2I_\sigma \end{aligned} \quad (1.4)$$

Αμφότερες οι μέθοδοι, στην συνέχεια, προσεγγίζουν τα σημεία μηδενισμού της λαπλασιανής. Το κάνουν αυτό δημιουργώντας από την λαπλασιανή μία δυαδική εικόνα και βρίσκοντας στην συνέχεια το περίγραμμά της. Τα σημεία μηδενισμού είναι τα σημεία στα οποία το περίγραμμα έχει μοναδιαία τιμή.

$$\begin{aligned} X &= (L \geq 0) \\ Y &= (X \oplus B) - (X \ominus B) \\ (i, j) &= \arg |Y| = 1 \end{aligned} \quad (1.5)$$

Επειδή, όμως, το κριτήριο αυτό επιστρέφει και σημεία τα οποία δεν ανήκουν σε πραγματικές ακμές, τελικά επιλέγονται τα σημεία αυτά στα οποία η εξομαλυμένη εικόνα παρουσιάζει μεγάλη κλίση:

$$Y[i, j] = 1 \wedge \|\nabla I_\sigma[i, j]\| > \theta_{edge} \cdot \max_{x, y} \|\nabla I_\sigma\| \quad (1.6)$$

Η ποιότητα της ανίχνευσης ακμών εξαρτάται από την επιλογή των παραμέτρων σ, θ_{edge} , δηλαδή αντιστοίχως της παραμέτρου εξομάλυνσης και του κατωφλίου για την αποδοχή μίας ακμής.

1.3 Αξιολόγηση των Αποτελεσμάτων Ανίσχνευσης Ακμών

Για να αξιολόγηση των αποτελεσμάτων της συνάρτησης, πρέπει να οριστεί ένα ποιοτικό κριτήριο. Εν προκειμένω ορίζεται ο μέσος όρος μεταξύ των αληθώς ανισχνευθεισών ακμών και των ψευδών ανισχνευθεισών ακμών. Φυσικά, είναι αναγκαία η πληροφορία των αληθινών ακμών. Αυτές βρίσκονται από την αυθεντική, μη θορυβημένη εικόνα, χρήσει ενώ κατωφλίου ως εξής:

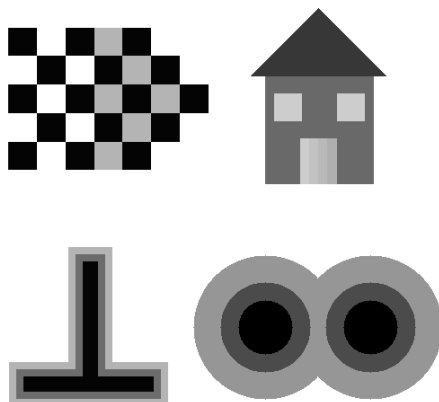
$$M = (I \oplus B) - (I \ominus B) \Rightarrow T = (M > \theta_{real}) \quad (1.7)$$

Δεδομένων, λοιπόν, των αληθινών ακμών, το ποιοτικό κριτήριο διατυπώνεται ως:

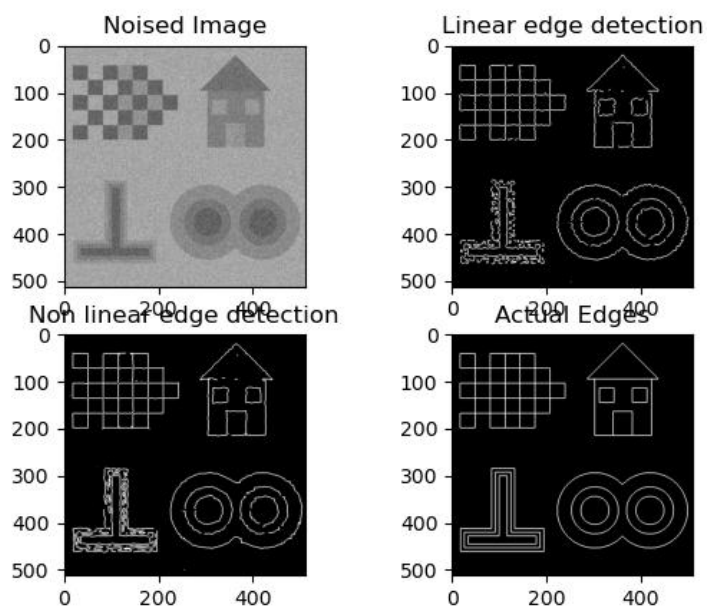
$$\begin{aligned} C &= \frac{Pr(D|T) + Pr(T|D)}{2}, \\ Pr(T|D) &= \frac{\|D \cap T\|}{\|T\|} \end{aligned} \quad (1.8)$$

1.4 Αποτελέσματα Εφαρμογής της Συνάρτησης

Παρατίθενται τα αποτελέσματα εφαρμογής της μεθόδου στην εικόνας edgetest_23.png.



Σχήμα 1: Η δεδομένη εικόνα εισόδου

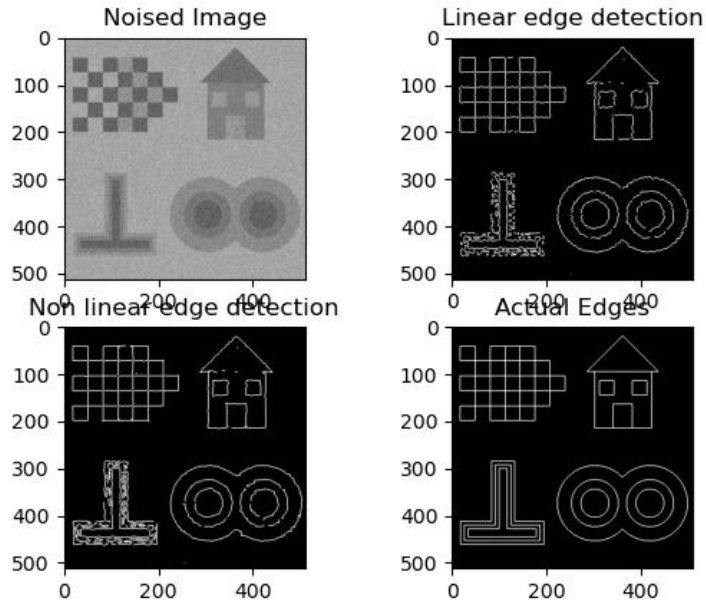


Σχήμα 2: Θόρυβος PSNR 10dB

Όταν η εικόνα 1 θορυβήθηκε με θόρυβο 10 dB PSNR, τα αποτελέσματα ήταν αυτά που φαίνονται στην 2 και οι βαθμολογίες στο κριτήριο ποιότητας οι ακόλουθες:

$$\begin{aligned} C_{10db-lin} &= 0.6182010279466628 \\ C_{10db-non} &= 0.7362110886293316 \end{aligned} \quad (1.9)$$

Ενώ όταν η εικόνα θορυβήθηκε με θόρυβο 20 dB PSNR, τα αποτελέσματα ήταν:



Σχήμα 3: Θόρυβος PSNR 20dB

Με κριτήριο ποιότητας:

$$\begin{aligned} C_{20db-lin} &= 0.9375887192342924 \\ C_{20db-non} &= 0.9671201094187654 \end{aligned} \quad (1.10)$$

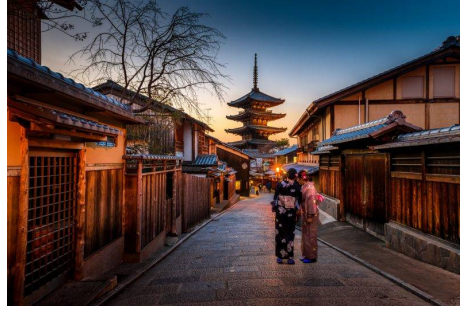
Τα αποτελέσματα επιτεύχθηκαν με τις ακόλουθες παράμετρους:

$$\sigma = [3 \quad 1.5], \theta = [0.2 \quad 0.2], \theta_{real} = 0,08 \quad (1.11)$$

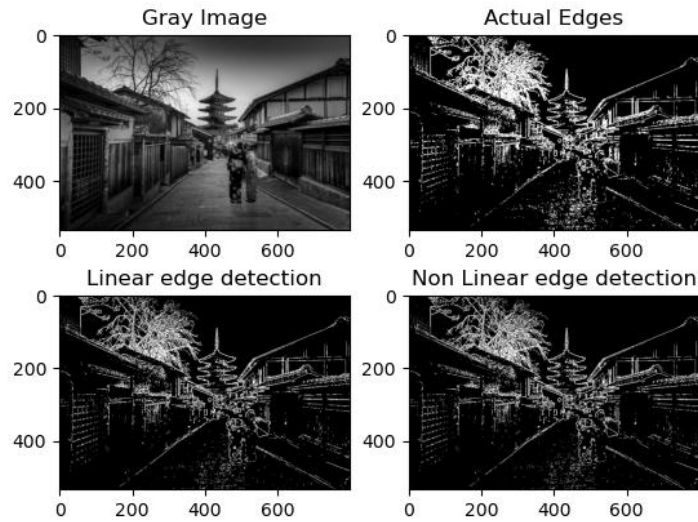
Παρατηρείται ότι η επίδραση του θορύβου είναι καθοριστική, καθώς στην εικόνα με τον λιγότερο θόρυβο επιτυγχάνεται καλύτερη βαθμολογία στο ποιοτικό κριτήριο. Επίσης, φαίνεται ότι η μη γραμμική μέθοδος πετυχαίνει καλύτερα αποτελέσματα σε κάθε περίπτωση.

1.5 Εφαρμογή των Αλγορίθμων Ανίχνευσης Ακμών σε Πραγματικές Εικόνες

Η ίδια συνάρτηση χρησιμοποιήθηκε και για την ανίχνευση ακμών στην πραγματική εικόνα του Kyoto, όπως φαίνεται στην εικόνα 4.



Σχήμα 4: Η πραγματική φωτογραφία του Kyoto



Σχήμα 5: Ανίχνευση ακμών στην πραγματική φωτογραφία του Kyoto

Η επεξεργασία έγινε με τις παραμέτρους:

$$\sigma = 0.3, \theta = 0.2, \theta_{real} = 0.23 \quad (1.12)$$

και πέτυχε τις βαθμολογίες:

$$\begin{aligned} C_{lin} &= 0.818604135617401 \\ C_{non} &= 0.8188562865181751 \end{aligned} \quad (1.13)$$

Τα δε οπτικά αποτελέσματα φαίνονται στην εικόνα 5. Τόσο από την βαθμολογία στο ποιοτικό κριτήριο όσο και οπτικά, φαίνεται ότι αμφότερες η γραμμική και η μη γραμμική μέθοδος έδωσαν πολύ καλά αποτελέσματα.

2 Μέρος 2ο

Στο σημείο αυτό της εργασίας θα επιχειρηθεί η ανίχνευση διαφόρων χαρακτηριστικών, τόσο σε μία μόνο κλίμακα όσο και σε πολλαπλές κλίμακες.

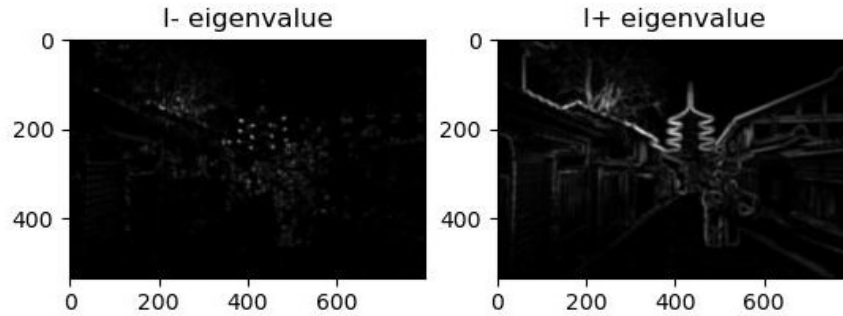
2.1 Ανίχνευση Γωνιών

Πρώτα επιχειρείται ανίχνευση γωνιών. Θα παρουσιαστεί η μεθοδολογία και ταυτοχρόνως τα βήματα της επεξεργασίας, για να γίνει σαφής.

Αρχικά υπολογίζεται ο δομικός τανυστής \mathbf{J} και οι ιδιοτιμές του.

$$\begin{aligned} J_1(x, y) &= G_\rho * \left(\frac{\partial I_\sigma}{\partial x} \cdot \frac{\partial I_\sigma}{\partial x} \right) (x, y) \\ J_2(x, y) &= G_\rho * \left(\frac{\partial I_\sigma}{\partial x} \cdot \frac{\partial I_\sigma}{\partial y} \right) (x, y) \\ J_3(x, y) &= G_\rho * \left(\frac{\partial I_\sigma}{\partial y} \cdot \frac{\partial I_\sigma}{\partial y} \right) (x, y) \\ \lambda_\pm(x, y) &= \frac{1}{2} \left(J_1 + J_2 \pm \sqrt{(J_1 - J_3)^2 + 4J_2^2} \right) \end{aligned} \quad (2.1)$$

Οι ιδιοτιμές περιέχουν χρήσιμη πληροφορία σχετικά με τις ακμές και τις γωνίες της εικόνας. Παρακάτω, στην εικόνα 6, δίνεται η απεικόνισή τους ως γκριζες εικόνες.



Σχήμα 6: Απεικόνιση των ιδιοτιμών της εικόνας ως γκριζες εικόνες

Φαίνεται ότι η εικόνα η αντιστοιχούσα στην ιδιοτιμή λ_- αποτυπώνει έντονα τις απολήξεις του πύργου, δηλαδή τις αιχμηρές γωνίες, ενώ η εικόνα αντιστοιχούσα στην ιδιοτιμή λ_+ αποτυπώνει έντονα τα περιγράμματα, δηλαδή τις ακμές.

Στην συνέχεια, εξάγεται ένα κριτήριο γωνιότητας συναρτήσει των ιδιοτιμών και μίας σταθεράς k :

$$R(x, y) = \lambda_- \lambda_+ - k \cdot (\lambda_- + \lambda_+)^2 \quad (2.2)$$

Τέλος, επιλέγονται ως γωνίες τα σημεία αυτά τα οποία μεγιστοποιούν το κριτήριο εντός τετραγωνικών παραθύρων και αποδίδουν στο κριτήριο γωνιότητας τιμή μεγαλύτερη από ένα κατώφλι.

$$R(x, y) > \theta_{corn} \cdot R_{max} \quad (2.3)$$

Το αποτέλεσμα εφαρμογής αυτών των συνθηκών στην εικόνα-κριτήριο R , για τις παραμέτρους $\sigma = 2, \rho = 2.5, k = 0.05, \theta_{corn} = 0.1$ είναι το ακόλουθο:



Σχήμα 7: Ανίχνευση γωνιών στην εικόνα Kyoto

Φαίνεται ότι έχουν εντοπιστεί με επιτυχία πολλές γωνίες, ειδικά οι αιχμηρές απολήξεις του πύργου αλλά και οι ορθές γωνίες οι σχηματιζόμενες από τους πασσάλους των κτιρίων. Ωστόσο, έχουν αναγνωριστει ως γωνίες και ανεπιθύμητα σημεία, όπως τα κλαδιά του δέντρου. Επομένως, συμπεραίνεται ότι η μέθοδος είναι αποτελεσματική αλλά έχει περιθώριο βελτίωσης.

2.2 Πολυκλιμακωτή Ανίχνευση Γωνιών

Ένα σημαντικό μειονέκτημα της προηγούμενης μεθόδου ήταν ότι περιοριζόταν σε μία μόνο κλίμακα, αφού δεχόταν μόνο ένα ζευγάρι παραμέτρων (σ, ρ) . Η ιδέα είναι να εφαρμοσθεί η ίδια μέθοδος N φορές, κλιμακώνοντας γεωμετρικά κάθε φορά τις παραμέτρους με μία παραμέτρο κλίμακας s . Δηλαδή, η επεξεργασία θα γίνεται με τις παραμέτρους:

$$\begin{aligned}\sigma_0, \sigma_1, \dots, \sigma_{N-1} &= s^0 \sigma_0, s^1 \sigma_0, \dots, s^{N-1} \sigma_0 \\ \rho_0, \rho_1, \dots, \rho_{N-1} &= s^0 \rho_0, s^1 \rho_0, \dots, s^{N-1} \rho_0\end{aligned}\quad (2.4)$$

Λόγω των πολλαπλών κλιμάκων, θα επιλεγθούν πολλά σημεία τα οποία έχουν την ίδια πιθανότητα σφάλματος όπως προηγουμένως. Για αυτόν τον λόγο, επιλέγονται τελικά αυτά τα σημεία-γωνίες τα οποία μεγιστοποιούν κάποιο κριτήριο σε μία περιοχή κλιμάκων. Εν προκειμένω, το κριτήριο επιλέγεται να είναι η κανονικοποιημένη Λαπλασιανή της Γκαουσιανής.

$$|LoG(x, y, \sigma_i)| = \sigma_i^2 |L_x x(x, y, \sigma_i) + L_y y(x, y, \sigma_i)| \quad (2.5)$$

Ουσιαστικά, το σκεπτικό είναι ότι μία γωνία οφείλει να εντοπίζεται τουλάχιστον σε μία μικρή περιοχή κλιμάκων, όχι μόνο σε μία διακεκριμένη κλίμακα. Το αποτέλεσμα της πολυκλιμακωτής μεθόδου στην ίδια εικόνα φαίνεται παρακάτω:



Σχήμα 8: Πολυκλιμακωτή ανίχνευση γωνιών στην εικόνα Kyoto

Παρατηρείται, ότι εντοπίζονται λιγότερες γωνίες από,τι προηγουμένως. Επι-
θεωρώντας την εικόνα, παρατηρείται ότι έπαψαν να εντοπίζονται κάποιες ψεύτικες
γωνίες, κυρίως αυτές στα ψιλά κλαδιά του δέντρου, παρασέρνοντας όμως και λίγες
πραγματικές, όπως αυτές στον πύργο. Ωστόσο, εξακολουθούν να εντοπίζονται
εν πολλοίς οι πραγματικές γωνίες, και μάλιστα με μεγαλύτερη βεβαιότητα, αφού
περικλείονται από κύκλους διαφόρων κλιμάκων. Συμπερασματικά, η μέθοδος αυτή
είναι σαφώς βελτίωση της προηγούμενης, με αντίτιμο μεγαλύτερη υπολογιστική
πολυπλοκότητα.

2.3 Ανίχνευση Blobs

2.4 Πολυκλιμακωτή Ανίχνευση Blobs

2.5 Επιτάχυνση Ανίχνευσης Blobs

2.6 Επιτάχυνση Πολυκλιμακωτής Ανίχνευση Blobs