

Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης
Πολυτεχνική Σχολή
Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών

Ψηφιακή Επεξεργασία Εικόνας

Εργασία 2

Κουκουλέτσου Κατερίνα 10218

Μάιος 2025

Εισαγωγή

Η παρούσα εργασία έχει ως στόχο την υλοποίηση ανίχνευσης ακμών σε ψηφιακές εικόνες, αξιοποιώντας μεθόδους επεξεργασίας όπως οι τελεστές Sobel και Laplacian of Gaussian (LoG), σε συνδυασμό με μια παραλλαγή του μετασχηματισμού Hough ειδικά προσαρμοσμένη για την ανίχνευση κυκλικών περιγραμμάτων.

Αρχικά, ορίζονται οι αντίστοιχες μασκες φίλτρων που αναπαριστούν μαθηματικά τους παραπάνω μετασχηματισμούς, οι οποίες εφαρμόζονται στην εικόνα μέσω της συνάρτησης που υλοποιεί την δισδιάστατη πράξη της συνέλιξης.

Στη συνέχεια, ακολουθεί η εφαρμογή του αλγορίθμου Hough ο οποίος έχει σκοπό να ανιχνεύσει παραμετρικές καμπύλες μέσα στην εικόνα. Ειδικά για την ανίχνευση κύκλων, η μέθοδος αναζητά συνδυασμούς παραμέτρων κέντρου και ακτίνας που συμφωνούν με τα δεδομένα της εικόνας, αξιολογώντας την πιθανότητα ύπαρξης κυκλικών περιγραμμάτων μέσω ενός πίνακα ψηφοφορίας, accumulator.

Αρχικά, επιχειρήθηκε ανίχνευση zero-crossings μέσω pattern matching σε τοπικές περιοχές 3×3 . Ωστόσο, η προσέγγιση αυτή αποδείχθηκε αναποτελεσματική, καθώς η υφή της εικόνας παρουσίαζε τοπικά μοτίβα παρόμοια με εκείνα των ακμών, οδηγώντας σε έντονο ύσρυβο στην έξοδο. Για την αντιμετώπιση αυτού του προβλήματος, δοκιμάστηκε στη συνέχεια η χρήση μεγαλύτερων περιοχών 5×5 , με σκοπό να ενσωματωθεί περισσότερη πληροφορία σε κάθε έλεγχο, επιτρέποντας έτσι την καλύτερη διάκριση υφής από πραγματικές ακμές. Αν και η μέθοδος αυτή παρείχε βελτιωμένα αποτελέσματα, παρουσίασε άλλα προβλήματα, όπως αυτά του σαφή ορισμού και της αντιστοίχισης patterns τόσο αυξημένης διαστασιμότητας. Τελικά, υιοθετήθηκε μια πιο αποτελεσματική προσέγγιση βασισμένη ξανά σε 3×3 περιοχές, αλλά με τη χρήση δύο κατωφλίων: ένα για τη μεταβολή πρόσημου στη συνελιγμένη εικόνα και ένα για τη διαφορά φωτεινότητας στην αρχική grayscale εικόνα. Ο συνδυασμός αυτός απέδωσε ικανοποιητικά αποτελέσματα, επιτρέποντας την απομόνωση πραγματικών ακμών από τον ύσρυβο. Αναλυτική παρουσίαση των δύο υλοποιήσεων και αξιολόγηση των αποτελεσμάτων τους ακολουθεί παρακάτω.

Συνέλιξη

Η συνάρτηση `fir_conv` υλοποιεί τη δισδιάστατη πράξη της συνέλιξης, η οποία ορίζεται μαθηματικά ως εξής:

$$g(n_1, n_2) = \sum_{k_1=-\infty}^{+\infty} \sum_{k_2=-\infty}^{+\infty} h(k_1, k_2) f(n_1 - k_1, n_2 - k_2)$$

Όμως όπως επισημαίνεται και από την εκφώνηση της εργασίας, στην πράξη τόσο η είσοδος f όσο και ο πυρήνας h έχουν πεπερασμένη υποστήριξη, οπότε τα όρια των αθροισμάτων περιορίζονται ως:

$$g(n_1, n_2) = \sum_{k_1=N_{1,\min}}^{N_{1,\min}+N_1-1} \sum_{k_2=N_{2,\min}}^{N_{2,\min}+N_2-1} h(k_1, k_2) f(n_1 - k_1, n_2 - k_2)$$

Κατά την υλοποίηση, πρώτα εφαρμόζεται η αναστροφή (flip) του πυρήνα h ως προς και τις δύο διαστάσεις (η αναστροφή γίνεται και ως προς γραμμές και ως προς στήλες), ώστε να ακολουθείται ο ορισμός της συνέλιξης (σε αντίθεση με την πράξη του correlation όπου δεν γίνεται flip).

Στη συνέχεια, προχύπτει η ανάγκη για padding της αρχικής εικόνας f , ώστε να επιτραπεί η εφαρμογή της μάσκας h σε κάθε pixel της εικόνας, συμπεριλαμβανομένων και των ακραίων pixel. Το padding υπολογίζεται ως εξής:

Έστω ότι η μάσκα έχει διαστάσεις $M_h \times M_w$ και το σημείο αναφοράς της μάσκας είναι $\text{mask_origin} = (p, q)$, δηλαδή ο δείκτης της μάσκας που ευθυγραμμίζεται με το τρέχον pixel της εικόνας. Τότε, το padding που προστίθεται είναι p γραμμές πάνω και $M_h - p - 1$ γραμμές κάτω, καθώς και q στήλες αριστερά και $M_w - q - 1$ στήλες δεξιά. Με αυτόν τον τρόπο, η νέα εικόνα με padding έχει διαστάσεις

$$(H + p + (M_h - p - 1)) \times (W + q + (M_w - q - 1)) = (H + M_h - 1) \times (W + M_w - 1)$$

όπου H και W είναι οι διαστάσεις της αρχικής εικόνας.

Η ύπαρξη αυτού του padding εξασφαλίζει ότι για κάθε pixel (n_1, n_2) της αρχικής εικόνας, μπορεί να οριστεί το αντίστοιχο παράθυρο $M_h \times M_w$ στην padded εικόνα, ώστε να γίνει το στοιχειώδης πολλαπλασιασμός στοιχείου προς στοιχείο με τον αναστραμμένο πυρήνα και να αθροιστούν τα αποτελέσματα.

Στην συνέχεια, με επαναληπτική διαδικασία, για κάθε θέση (i, j) στην έξοδο, ορίζεται το αντίστοιχο region της padded εικόνας και υπολογίζεται η τιμή της εξόδου ως

$$g(i, j) = \sum_{k_1=0}^{M_h-1} \sum_{k_2=0}^{M_w-1} h_{flipped}(k_1, k_2) \cdot \text{padded_img}(i + k_1, j + k_2)$$

όπου $h_{flipped}$ είναι η αναστραμμένη μάσκα.

Η εικόνα εξόδου έχει τις ίδιες διαστάσεις με την αρχική εικόνα $H \times W$ όπως άλλωστε ζητείται και από την εκφώνηση, καθώς η παραπάνω διαδικασία υπολογίζει μια τιμή εξόδου ανά pixel της αρχικής.

Τέλος, ορίζεται η νέα αρχή συστήματος συντεταγμένων της εικόνας εξόδου ως το άθροισμα της αρχής της αρχικής εικόνας και της αρχής της μάσκας, δηλαδή

$$\text{out_origin} = \text{in_origin} + \text{mask_origin}$$

προκειμένου να διατηρηθεί η σωστή γεωμετρική αναφορά.

Τελεστής Sobel

Ο τελεστής Sobel υλοποιεί τη συνέλιξη της εικόνας με δύο διαχριτές μάσκες, κάθε μία εκ των οποίων προσεγγίζει την παράγωγο πρώτης τάξης κατά μία χωρική διάσταση.

Η μάσκα για την κατεύθυνση x_1 (οριζόντια διεύθυνση) είναι:

$$G_{x_1} = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix}$$

Αντίστοιχα, για την κατεύθυνση x_2 (κατακόρυφη διεύθυνση), η μάσκα είναι:

$$G_{x_2} = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

Η εφαρμογή των παραπάνω μασκών στην εικόνα παράγει δύο επιμέρους εικόνες $g_1(n_1, n_2)$ και $g_2(n_1, n_2)$, οι οποίες περιέχουν την κατά προσέγγιση παράγωγο της έντασης της εικόνας κατά x_1 και x_2 αντίστοιχα.

Το συνολικό μέτρο της μεταβολής της έντασης, το οποίο υποδεικνύει την παρουσία ακμής, υπολογίζεται για κάθε σημείο (n_1, n_2) ως το Ευκλείδειο μέτρο του διανύσματος των παραγώγων:

$$g(n_1, n_2) = \sqrt{g_1(n_1, n_2)^2 + g_2(n_1, n_2)^2}$$

Η τιμή $g(n_1, n_2)$ αποτελεί τελικό δείκτη της έντασης ακμής στο αντίστοιχο σημείο της εικόνας.

Τελεστής Laplacian of Gaussian (LoG)

Ο τελεστής Laplacian of Gaussian (LoG) χρησιμοποιείται συνήθως για την ανίχνευση ακμών μέσω του εντοπισμού περιοχών με έντονες αλλαγές στη δεύτερη παράγωγο της έντασης της εικόνας. Θεωρητικά, ο LoG συνδυάζει την εξομάλυνση μέσω ενός Γκαουσιανού φίλτρου με τη Λαπλασιανή παράγωγο δεύτερης τάξης:

$$\nabla^2 f(x_1, x_2) = \frac{\partial^2 f}{\partial x_1^2} + \frac{\partial^2 f}{\partial x_2^2}$$
$$LoG(x_1, x_2) = -\frac{1}{\pi\sigma^4} \left(1 - \frac{x_1^2 + x_2^2}{2\sigma^2}\right) e^{-\frac{x_1^2+x_2^2}{2\sigma^2}}$$

Στην παρούσα υλοποίηση, χρησιμοποιήθηκε διακριτή προσέγγιση του τελεστή μέσω μάσκας συνελίξεως 5×5 , η οποία εφαρμόστηκε στην είσοδο εικόνα ώστε να παραχθεί μια LoG-επεξεργασμένη εικόνα.

Διακριτή Προσέγγιση LoG

$$h = \begin{bmatrix} 0 & 0 & -1 & 0 & 0 \\ 0 & -1 & -2 & -1 & 0 \\ -1 & -2 & 16 & -2 & -1 \\ 0 & -1 & -2 & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 \end{bmatrix}$$

Ανίχνευση Κύκλων με τη Μέθοδο Hough

Η βασική ιδέα του Hough είναι η μεταφορά του προβλήματος ανίχνευσης καμπυλών από το χωρικό επίπεδο σε έναν παραμετρικό χώρο, όπου κάθε σημείο ακμής της εικόνας ”ψηφίζει” για τις παραμέτρους των καμπυλών που πιθανώς περνούν από αυτό. Η συσσώρευση ψήφων σε συγκεκριμένα σημεία του παραμετρικού χώρου (accumulator) υποδεικνύει την παρουσία πιθανών καμπυλών στην εικόνα.

Για την ανίχνευση κύκλων συγκεκριμένα, η μέθοδος Hough προσαρμόζεται ώστε να εντοπίζει καμπύλες της μορφής $(x - a)^2 + (y - b)^2 = \rho^2$, όπου οι παράμετροι (a, b) αντιπροσωπεύουν το κέντρο του κύκλου και ρ την ακτίνα του. Συνεπώς, ο accumulator ορίζεται ως τρισδιάστατος πίνακας διαστάσεων $K \times L \times M$, ο οποίος αντιστοιχεί στις διακριτές τιμές των a , b και ρ σε κατάλληλα επιλεγμένα διαστήματα.

Κατά την ψήφιση, για κάθε σημείο ακμής (x_i, y_i) και για κάθε διακριτή ακτίνα r , υπολογίζονται πιθανές θέσεις κέντρων (a, b) που ικανοποιούν τη γεωμετρία του κύκλου, βάσει προκαθορισμένων τιμών γωνίας $\theta \in [0^\circ, 360^\circ]$. Οι παραγόμενες τιμές (a, b) χαρτογραφούνται στα πλησιέστερα bins του accumulator και αυξάνουν τις αντίστοιχες τιμές ψήφων.

Ανίχνευση Zero-Crossing με Pattern Matching 5x5

Μετά την εφαρμογή του Laplacian of Gaussian (LoG) στον αρχικό πίνακα εικόνας, επιχειρήθηκε αρχικά η τεχνική ανίχνευσης ακμών μέσω **pattern matching** σε περιοχές 3×3 . Ωστόσο, διαπιστώθηκε ότι η προσέγγιση αυτή δεν απέδιδε ικανοποιητικά αποτελέσματα. Ο κύριος λόγος είναι πως οι μικρές περιοχές 3×3 δεν περιέχουν αρκετές πληροφορίες για να γίνει σαφής διάκριση μεταξύ πραγματικών ακμών και θορύβου, ειδικά σε εικόνες με έντονη υφή. Πολλά μοτίβα θορύβου παρουσίαζαν υψηλή ομοιότητα με τις ακμές, οδηγώντας σε ψευδώς θετικές ανιχνεύσεις.

Μετάβαση σε περιοχές 5×5 Για την αντιμετώπιση του προβλήματος αυτού, υιοθετήθηκε η προσέγγιση της ανίχνευσης σε μεγαλύτερες περιοχές 5×5 , καθώς αυτές περιέχουν περισσότερο τοπικό χωρικό πλαίσιο και επιτρέπουν καλύτερη κατανόηση της δομής της εικόνας. Η αύξηση της περιοχής παρατήρησης, ωστόσο, οδηγεί εκθετικά σε περισσότερους πιθανούς συνδυασμούς δυαδικών μοτίβων. Για να αντιμετωπιστεί αυτό:

- Ορίστηκε μια μικρή συλλογή από αντιροσωπευτικά **βασικά μοτίβα ακμών** (`base_patterns`).
- Με τη βοήθεια της συνάρτησης `generate_rotations`, παράγονται αυτόμata οι 4 περιστροφές κάθε μοτίβου κατά 0° , 90° , 180° , και 270° .

Έτσι, αποφεύγεται η ανάγκη για ρητό ορισμό όλων των πιθανών μοτίβων, μειώνοντας σημαντικά την πολυπλοκότητα της υλοποίησης. Για την εφαρμογή εφαρμογή του Pattern Matching, αρχικά εφαρμόζεται στον πίνακα εισόδου ένας 5×5 πυρήνας LoG (Laplacian of Gaussian) μέσω συνέλιξης. Το αποτέλεσμά της αποθηκεύεται στο `conv_result`. Από αυτό, δημιουργείται ένας **δυαδικός χάρτης πρόσημου** (`sign_map`), που περιέχει τιμές 1 για θετικά στοιχεία και 0 για αρνητικά. Η σύγκριση γίνεται σε περιοχές 5×5 του `sign_map`, με τα περιστρεφόμενα μοτίβα.

Για να θεωρηθεί μια περιοχή ως πιθανή ακμή:

1. Υπολογίζεται η μέγιστη διαφορά τιμών (`max_diff`) στο `conv_result` της περιοχής. Αν είναι μικρότερη από το κατώφλι `threshold`, η περιοχή παραλείπεται, καθώς θεωρείται ότι δεν περιέχει σημαντικές μεταβολές.
2. Ελέγχεται η ομοιότητα της περιοχής με κάθε μοτίβο. Ο αριθμός των αντιστοιχίσεων (ίδιες τιμές ανά θέση) συγκρίνεται με το μέγιστο δυνατό (25 pixels). Αν το ποσοστό των όμοιων στοιχείων υπερβαίνει το `match_ratio`, τότε η περιοχή χαρακτηρίζεται ως ακμή.

Ρόλος των παραμέτρων

- **Threshold** (`threshold`): Ρυθμίζει την ευαισθησία του ανιχνευτή στις έντονες μεταβολές φωτεινότητας. Αν μια περιοχή έχει χαμηλή διακύμανση (δηλ. μικρή μέγιστη διαφορά τιμών), απορρίπτεται ως μη σημαντική.
- **Match Ratio** (`match_ratio`): Ελέγχει το πόσο αυστηρή είναι η απαίτηση ομοιότητας μεταξύ του μοτίβου της περιοχής και των πρότυπων μοτίβων. Π.χ. τιμή 0.8 σημαίνει ότι 80% των στοιχείων πρέπει να συμφωνούν.

Η τελική εικόνα εξόδου (`out_img`) περιέχει τιμές 1 σε θέσεις όπου εντοπίστηκαν αξιόπιστες ακμές και 0 αλλού.

Ανίχνευσης Zero-Crossing με Έλεγχο Luminance

Στη δεύτερη υλοποίηση της ανίχνευσης ακμών με τον τελεστή LoG, αξιοποιούνται πληροφορίες τόσο από την εικόνα που έχει προκύψει μετά τη συνέλιξη όσο και από την αρχική grayscale εικόνα.

Η νέα εκδοχή κώδικα βασίζεται στην παρατήρηση ότι παρόλο που οι περιοχές υφής μπορούν να παράγουν αλλαγές πρόσημου στο αποτέλεσμα της συνέλιξης (zero-crossings), αυτές οι αλλαγές συνήθως συνοδεύονται από μικρές μεταβολές φωτεινότητας στην grayscale εικόνα. Αντίθετα, οι ακμές αντικειμένων —όπως η μετάβαση από το λευκό φόντο στην πορτοκαλί περίμετρο της μπάλας— προκαλούν σημαντικότερες διαφορές φωτεινότητας. Συνεπώς, γίνεται η εξής δοκιμή: κάθε φορά που εντοπίζεται πιθανό zero-crossing σε περιοχή 3×3 της συνελιγμένης εικόνας, ελέγχεται επιπλέον αν η αντίστοιχη περιοχή στην αρχική grayscale εικόνα παρουσιάζει επαρκή ένταση μεταβολής luminance. Συγκεκριμένα ελέγχεται αν η διαφορά της μέγιστης τιμής luminance της 3×3 περιοχής από την ελάχιστης, ξεπερνάει το κατώφλι luminance threshold. Αν η διαφορά αυτή είναι μικρή, η αλλαγή απορρίπτεται ως πιθανός θόρυβος. Διαφορετικά αποθηκεύεται ως ακμή.

Ο αλγόριθμος επομένως εφαρμόζει δύο ελέγχους.

1. Αρχικά, έλεγχο μεταβολής πρόσημου στη συνελιγμένη εικόνα LoG με βάση κατώφλι threshold, που εκφράζει την ελάχιστη απαιτούμενη διαφορά τιμών εντός της περιοχής για να ληφθεί υπόψη.
2. Δεύτερο έλεγχο μεταβολής φωτεινότητας του ίδιου region στην grayscale εικόνα, μέσω παραμέτρου `lum_thresh`, που προσδιορίζει τη μικρότερη αποδεκτή διαφορά φωτεινότητας για να θεωρηθεί η μετάβαση ουσιαστική.

Σημείωση Καθώς η εκφώνηση δεν διευκρίνιζε πολλά για την ανίχνευση των zero-crossings, δεν ήταν γνωστό τι επιτρέπεται να κάνουμε και τι όχι. Από τη μία υπάρχει η σκέψη ότι ίσως έπρεπε να έχει χρησιμοποιηθεί αποκλειστικά η εικόνα μετά την εφαρμογή της μάσκας και η εξαγωγή πληροφορίας από την αρχική grayscale να θεωρηθεί εκτός πλαισίου.

Από την άλλη, η grayscale εικόνα αποτελεί ήδη δεδομένο εισόδου στη διαδικασία εφαρμογής του τελεστή LoG. Κατά συνέπεια, δεν είναι παράλογο να αξιοποιηθεί περισσότερο.

Πάντως, επειδή έγινε προσπάθεια να εφαρμοστεί λογική κατωφλίου (`threshold`) απευθείας στην εικόνα της συνέλιξης και σε περιοχή 3×3 αλλά και σε περιοχή 5×5 οπως παρουσιάστηκε πάνω, σημειώνεται ότι η προσέγγιση απέτυχε να αποδώσει καλό αποτέλεσμα. Αυτό ίσως εξηγείται αν σκεφτεί κανείς ότι μετά την συνέλιξη η τιμή του κάθη πίξελ έχει προκύψει από συνδυασμό (άθροισμα) του ίδιου και των γειτονικών πίξελ πολλαπλασιασμένων με κάποιον "συντελεστή βαρύτητας" (τιμή του αντίστοιχου στοιχείου μάσκας) το καθένα. Επομένως μετά την εφαρμογή της μάσκας, χάνεται η "οξύτητα" στις μεγάλες μεταβολές τιμών, οι τοπικές μεταβολές "εξομαλύνονται" και χάνεται μέρος της έντονης διαφοροποίησης που χαρακτηρίζει τις πραγματικές ακμές. Αυτό καθιστά πολύ πιο δύσκολο τον καθορισμό ενός κατωφλίου που να διαχωρίζει ικανοποιητικά ακμές από θόρυβο, καθώς το δυναμικό εύρος των διαφορών έχει μειωθεί.

Ρύθμιση Παραμέτρων

Τέλος επισημαίνεται ότι κατά την εκτέλεση κώδικα έγινε επιλογή του εύρους ζώνης και του αριθμού των bins. Καθώς και ορισμός του R_{max} και του V_{min} .

Επιλογή Εύρους Ζώνης

Πρώτον, επιλέχθηκε ένα πιο στενό εύρος ζώνης για τα ρ και άλλων παραμέτρων ώστε να περιοριστεί η αναζήτηση εντός του περιοχής ενδιαφέροντος. Η ρύθμιση των παραμέτρων βασίστηκε στην παρατήρηση της εικόνας, όπου διαπιστώθηκε οπτικά η περίπου θέση και διάμετρος της κυκλικής μπάλας του μπάσκετ.

Ενώ αρχικά θεωρήθηκε ότι η στοχευμένη επιλογή των ορίων για τις παραμέτρους θα μπορούσε να θεωρηθεί υπερβολικά καθοδηγούμενη προς το αποτέλεσμα, και ότι το σωστό θα ήταν ο αλγόριθμος να μπορεί να ανιχνεύσει τον σωστό κύκλο ακόμη και με ευρύτερα διαστήματα τιμών, διαπιστώθηκε στην πράξη ότι κάτι τέτοιο ήταν ανεφικτό λόγω του υπολογιστικού κόστους. Ο αριθμός των δυνατών συνδυασμών παραμέτρων αυξάνεται εκθετικά με το μέγεθος των διαστημάτων και την ανάλυση του πλέγματος (δηλαδή τα K, L, M), καθιστώντας την εκτέλεση του αλγορίθμου μη πρακτική σε συνήθεις υπολογιστικές υποδομές, όπως ένας προσωπικός φορητός υπολογιστής. Ως εκ τούτου, η περιορισμένη και στοχευμένη επιλογή των παραμετρικών διαστημάτων κρίθηκε αναγκαία για τη λειτουργικότητα της μεθόδου.

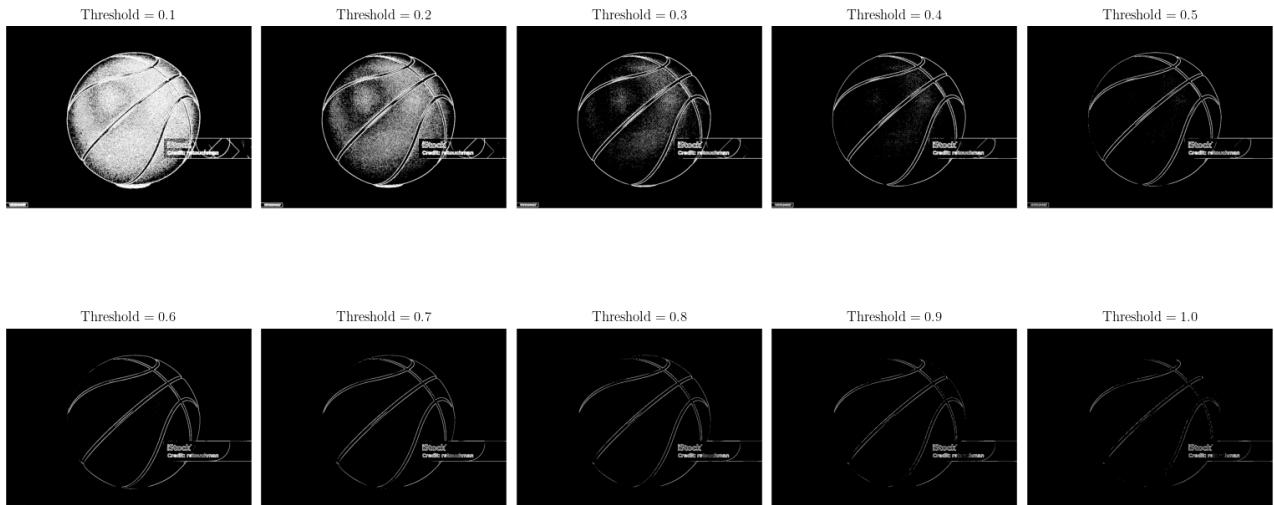
Επιλογή Αριθμού Bins

Προς αυτήν την κατεύθυνση, η αναζήτηση στα διαστήματα τιμών οδήγησε στην ανίχνευση πολλών σχεδόν ομόκεντρων κύκλων με ελάχιστα διαφορετικές ακτίνες, καθώς οι ψήφοι διασκορπίζονταν σε γειτονικά bins. Δηλαδή ακόμη και για την ανίχνευση ενός και μόνο κύκλου, το αποτέλεσμα περιελάμβανε πλήθος ομόκεντρων κύκλων με ελαφρώς διαφορετικές ακτίνες. Αυτό το φαινόμενο οφείλοταν πιθανά στο γεγονός ότι η πραγματική παράμετρος του κύκλου ενδέχεται να μην ευθυγραμμίζεται ακριβώς με κάποιο από τα διακριτά bins του accumulator. Συνεπώς, οι ψήφοι διασπείρονται σε γειτονικά bins, δημιουργώντας πολλαπλές τοπικές κορυφές αντί για μία έντονη και μοναδική. Για την επίλυση του προβλήματος, αυξήθηκε ο αριθμός των bins.

Διαγράμματα Sobel και LoG

Στην παρακάτω εικόνα παρουσιάζεται η εφαρμογή του τελεστή Sobel για διάφορες τιμές κατωφλίου. Παρατηρείται ότι όσο αυξάνεται το κατώφλι, μειώνεται η παρουσία θορύβου υφής. Ωστόσο, ταυτόχρονα παρατηρείται η εμφάνιση κενών/χασμάτων στη συνέχεια της ακμής της ίδιας της μπάλας. Η επιλογή κατάλληλης τιμής κατωφλίου αποτελεί trade-off μεταξύ καταστολής θορύβου και διατήρησης της συνέχειας των ακμών.

Επιλέχθηκε η εικόνα με κατώφλι 0.4 γιατί αν και εμφανίζει θόρυβο στο εσωτερικό της μπάλας είναι η μεγαλύτερη τιμή του threshold για την οποία διατηρείται πλήρως η συνέχεια της ακμής της εξωτερικής περιφέρειας. Παρ' όλα αυτά, και οι τιμές 0.3, 0.5 ή 0.6 θα μπορούσαν να δώσουν αξιόλογα αποτελέσματα, ανάλογα με τη ρύθμιση εύρους ακτίνας, θέσης κέντρου και V_{min} στον Hough Transform.

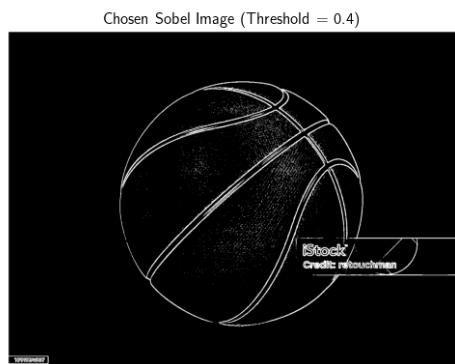
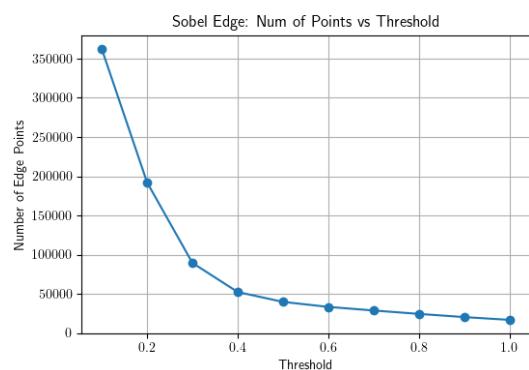


Η επιλογή έγινε με τη σκέψη ότι ο θόρυβος που προκύπτει από την υφή βρίσκεται εντός της μπάλας, δηλαδή εσωτερικά του τελικού επιθυμητού κύκλου. Συνεπώς, οι "ψεύτικοι" κύκλοι που θα μπορούσαν λανθασμένα να ανιχνευτούν, έχουν διάμετρο πολύ μικρότερη από αυτήν του κύριου κύκλου ενώ παράλληλα -λόγω της φύσης του θορύβου- δεν σχηματίζουν πλήρεις κυκλικές δομές. Με βάση αυτές τις δύο παρατηρήσεις, έγινε η υπόθεση ότι θα παράγουν πολύ πιο λίγες ψήφους στον accumulator του Hough.

Αντίθετα, η ακμή της εξωτερικής περιφέρειας της μπάλας, έχοντας πλήρη συνέχεια και πολύ μεγαλύτερη ακτίνα, συντίθεται από περισσότερα σημεία ακμής και συνεπώς συγκεντρώνει σημαντικά περισσότερες ψήφους. Έτσι θεωρήθηκε ότι, με την κατάλληλη ρύθμιση του αριθμού ψήφων (V_{min}), θα μπορούσε να επιτευχθεί η ανίχνευση μόνο του επιθυμητού εξωτερικού κύκλου και η απόρριψη λανθασμένα ανιχνευμένων εσωτερικών κύκλων.

Όπως φαίνεται από το διάγραμμα, παρατηρείται μία σαφής φθίνουσα σχέση μεταξύ της τιμής του κατωφλίου και του αριθμού ανιχνεύμενων σημείων. Καθώς το κατώφλι αυξάνεται, μειώνεται σταδιακά ο αριθμός των ανιχνεύμενων ακμών, γεγονός που υποδηλώνει ότι απορρίπτονται οι πιο

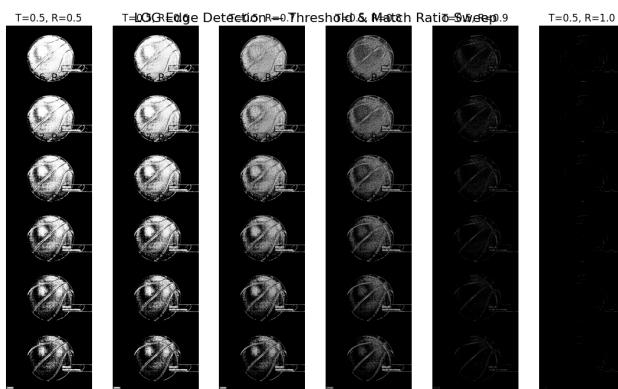
αδύναμες (λιγότερο έντονες) ακμές. Αυτό το αποτέλεσμα είναι και το αναμενόμενο, καθώς ένα υψηλότερο κατώφλι λειτουργεί ως αυστηρότερο φίλτρο, κρατώντας μόνο τα πιο ισχυρά σημεία μεταβολής στη φωτεινότητα. Από δίπλα φαίνεται σε μεγέθυνση η εικόνα που τελικά επιλέχθη.



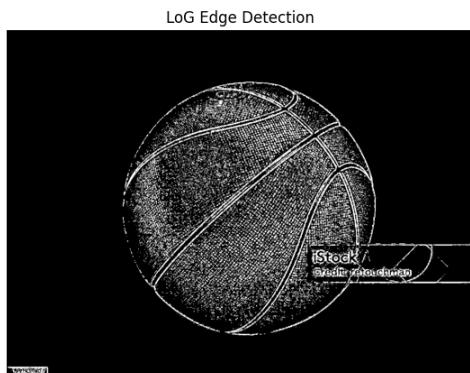
LoG - Πρώτη Εκδοχή

Ομοίως και στην περίπτωση εφαρμογής του τελεστή Laplacian of Gaussian. Ανάμεσα στο δίλημμα της διατήρησης της συνέχειας του περιγράμματος με ταυτόχρονη παρουσία ενισχυμένου εσωτερικού όγκου, και της μείωσης του όγκου με κόστος τη διακοπή της ακμής του αντικειμένου, κρίθηκε προτιμότερη η πρώτη επιλογή.

Η διατήρηση ενός συνεχούς και πλήρους περιγράμματος ως προτίμηση σημαντικότερη, ενώ ο εσωτερικός όγκος —αν και αυξημένος— αναμένεται να έχει περιορισμένη και εύκολα ελεγχόμενη επίδραση, καθώς ούτε σχηματίζει επαρκώς καθορισμένες κυκλικές δομές ούτε μιλάμε για μεγάλες τιμές ακτίνας ώστε να προκαλέσει ψευδείς ανιχνεύσεις. Παρατίθεται plot που δείχνει την δυαδική εικόνα για όλους τους δυνατούς συνδυασμούς match ratio και threshold και με βάση το οποίο plot έγινε και η επιλογή.



Τελικά η επιλεγμένη εικόνα για $matchratio = 0.8$ και $threshold = 0.6$ είναι η



Σύγκριση Sobel και LoG

Συγκρίνοντας τα αποτελέσματα του τελεστή Sobel με εκείνα του τελεστή Laplacian of Gaussian (LoG), παρατηρείται ότι ο LoG εμφανίζει μεγαλύτερη ευαισθησία στον όγκο που προέρχεται από την υφή της εικόνας. Συγκεκριμένα, ανιχνεύει μεγάλο αριθμό ακμών που δεν αντιστοιχούν σε ουσιαστικές μεταβολές της δομής του αντικειμένου, αλλά σε μικροδομές και λεπτομέρειες της υφής.

Αυτό πιθανόν οφείλεται στο γεγονός ότι ο LoG λειτουργεί ως τελεστής δεύτερης παραγώγου,

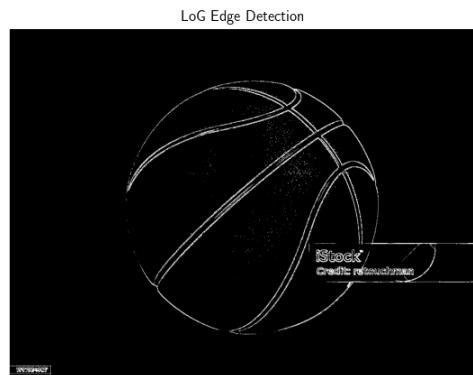
γεγονός που τον καθιστά πιο ευαίσθητο σε μικρές διακυμάνσεις φωτεινότητας σε σχέση με τον Sobel, ο οποίος βασίζεται σε πρώτες παραγώγους. Επιπλέον, ο LoG παρέχει μεγαλύτερη ακρίβεια στον εντοπισμό ακμών, όμως αυτή η ακρίβεια έρχεται εις βάρος της ανθεκτικότητας στον θόρυβο.

Επομένως, ενώ ο LoG μπορεί να ανιχνεύσει λεπτότερες και πιο αδύναμες ακμές, καθίσταται λιγότερο κατάλληλος για εικόνες με έντονη υφή ή θόρυβο, εφόσον δυσκολεύεται να διακρίνει τις σημαντικές ακμές από τις ανεπιθύμητες.

LoG - Δεύτερη Εκδοχή

Η εικόνα που προκύπτει από τη δεύτερη υλοποίηση του αλγορίθμου ανίχνευσης zero-crossings παρουσιάζει σαφώς μειωμένο θόρυβο, με τις παραμέτρους να έχουν οριστεί ως: threshold = 0.1 και luminance threshold = 0.15.

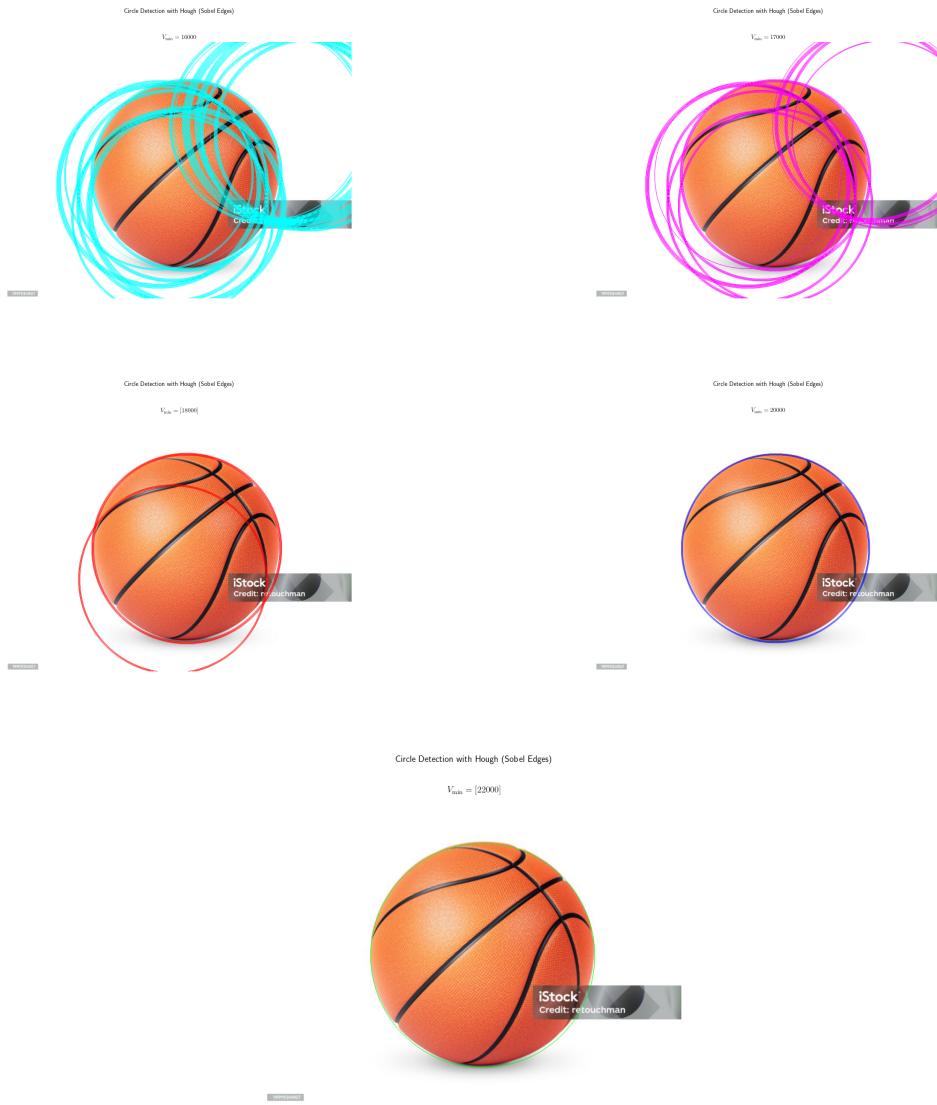
Δοκιμάστηκαν επίσης και άλλοι συνδυασμοί παραμέτρων, οι οποίοι παρήγαγαν επίσης παρόμοιες ικανοποιητικές εκδοχές με περισσότερο ή λιγότερο βαθμό θορύβου και με συνεχή ή ασυνεχή απεικόνιση της κύριας ακμής αντίστοιχα. Δηλαδή για κάποια εύρη τιμών κοντά στις τιμές που επιλέχθηκαν, τα αποτελέσματα της εικόνας ήταν σχετικά όμοια και άρα επίσης ικανοποιητικά.



Σύγχριση Sobel και LoG

Συγχρίνοντας πάλι Sobel και LoG, φαίνεται αυτήν την φορά πως τα αποτελέσματά τους είναι πολύ πιο κοντά. Παρόλ' αυτά, το Sobel και πάλι είναι προτιμότερο καθώς δεν έχει ασυνέχεια της ακμής.

Sobel και Hough Transform

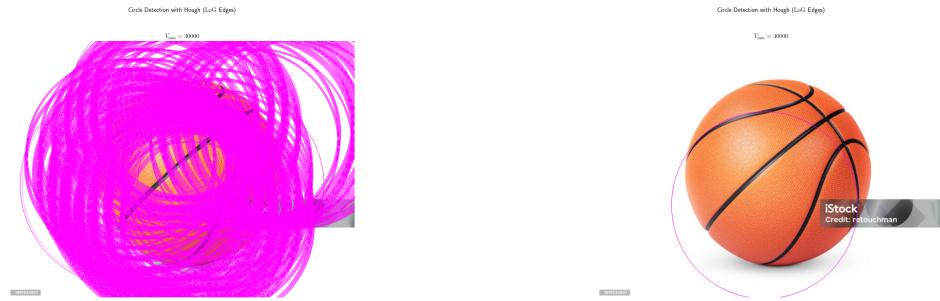


Μπορεί να μην είναι απόλυτα εμφανές αλλά στην τελευταία εικλονα φαίνεται ένας μόνο ανιχνευμένος κύκλος με πράσινο χρώμα, ενώ στην προτελευταία είναι 3 ομόκεντροι με ελάχιστα διαφορετικές ακτίνες με μπλε χρώμα.

Όπως φαίνεται από τα αποτελέσματα, η αύξηση του V_{\min} οδηγεί σε μείωση του αριθμού των ανιχνευμένων κύκλων. Συγκεκριμένα, για $V_{\min} = 2000$ εντοπίζονται οι τρεις μπλε κύκλοι γύρω από την περιοχή της πραγματικής ακμής, ενώ για $V_{\min} = 22000$ διατηρείται μόνο ο πράσινος — ο σωστός. Το γεγονός αυτό ενδέχεται να οφείλεται στην κατανομή των ψήφων σε γειτονικά bins του accumulator, με παρόμοια κέντρα και ακτίνες, εξαιτίας της εγγενώς διαχριτής φύσης του χώρου παραμέτρων.

Ο αλγόριθμος φαίνεται να ανταποκρίνεται επιτυχώς, καθώς με την κατάλληλη ρύθμιση του κατωφλίου ψήφων V_{\min} , εντοπίζεται μόνο ο επιθυμητός κύκλος. Για παράδειγμα, με τιμές $R_{\min} = 390$ και $R_{\max} = 430$ και $V_{\min} = 22000$, επιτυγχάνεται καθαρός και ακριβής εντοπισμός της ακμής του κύκλου, με εξάλειψη ψευδών ανιχνεύσεων που προκαλούνται από θόρυβο ή ελαφρώς παρεμφερή τοπικά χαρακτηριστικά.

LoG και Hough Transform - Version 1

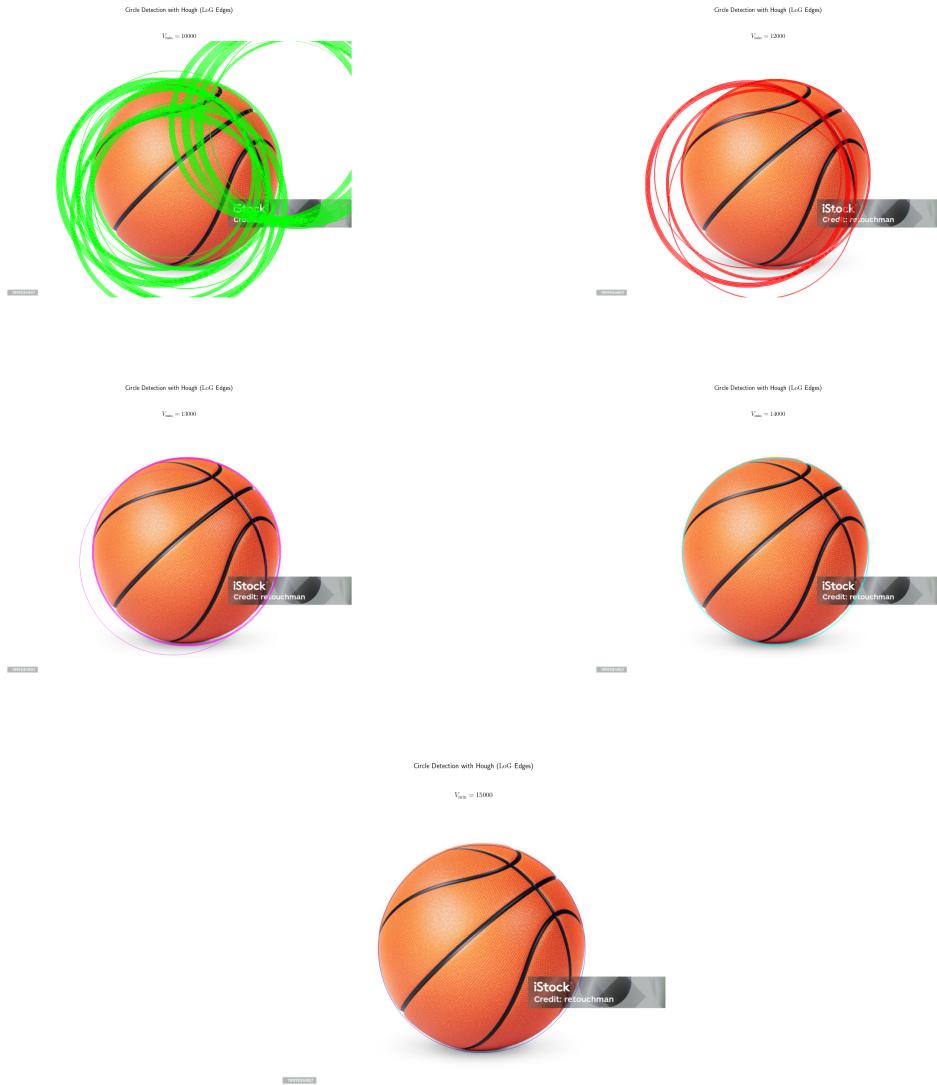


Παρουσιάζονται μόνο 2 σενάρια V_{min} για την πρώτη εκδοχή του LoG. Αρχικά γιατί το Version 1 δεν είναι η βέλτιστη λύση που βρέθηκε και κατα δεύτερον γιατί ο υπολογιστικός χρόνος για το Version 1 είναι τεράστιος.

Φαίνεται να μην λειτουργεί ικανοποιητικά καλά. Πιθανότατα γιατί η επιλογή μεγέθους για το pattern matching (5×5) είναι αρκετά δύσκολο να προσδιοριστεί σωστά. Παρόλο που η επιλογή αυτή εξασφαλίζει μια πιο ευρεία τοπική πληροφορία, ενδέχεται να χάνεται ακρίβεια στον εντοπισμό ακμών λόγω υπερβολικής γενίκευσης. Ακόμα και με τους μηχανισμούς που εισήχθησαν όπως η περιστροφή του μοτίβου και το match ratio, είναι πιθανό ορισμένες περιπτώσεις ακμών να μην αναγνωρίζονται ως τέτοιες, οδηγώντας σε απώλεια ψήφων που διαφορετικά θα συνέβαλλαν σε πιο ακριβή και κεντραρισμένη ανίχνευση του κύκλου.

Ότι προσπάθεια και αν έγινε — είτε με αλλαγές στα εύρη τιμών είτε με διάφορες ρυθμίσεις παραμέτρων — δεν ήταν δυνατό να εντοπιστεί που οφείλεται το offset του κυκλικού περιγράμματος. Πέρα υπόθεση είναι ότι το πρόβλημα οφείλεται στον θόρυβο της δυαδικής εικόνας, και συνεπώς σχετίζεται με τον τρόπο ορισμού των patterns κατά την ανίχνευση ακμών και δεν είναι κάτι που μπορεί να διορθωθεί με parameter tuning. Ισως μάλιστα να είναι θέμα του θορύβου της επικέτας iStock και όχι μόνο του θορύβου της υφής καθώς στην δυαδική εικόνα διαγράφεται σχετικά έντονα και αυτή.

LoG και Hough Transform - Version 2



Ο αλγόριθμος φαίνεται να δουλεύει πολύ καλά και καταφέρνει να ανιχνεύσει τη σωστή ακμή με κατάλληλη ρύθμιση του V_{min} . Αυτό είναι και αναμενόμενο, αφού από την εικόνα της δεύτερης εκδοχής του LoG φαίνεται πόσο πιο καθαρή είναι από ψόρυβο. Από τα πρώτα δύο plots παρατηρείται ότι οι κύκλοι που απορρίπτονται πρώτοι καθώς αυξάνεται το V_{min} είναι αυτοί που βρίσκονται πιο εκτός, δηλαδή έχουν μεγαλύτερη απόσταση από την πραγματική ακμή. Αντίθετα, παραμένουν οι κύκλοι που είναι πιο κοντά στην ακμή.

Οι ακτίνες των κύκλων που εντοπίζονται είναι παρόμοιες, κάτι που βγάζει νόημα αφού το εύρος ακτινών έχει περιοριστεί αρκετά (390–430). Μπορεί να μη φαίνεται καθαρά, αλλά στην εικόνα με $V_{min} = 14000$ ανιχνεύονται 2–3 ίδιοι κύκλοι, πιθανότατα λόγω διασποράς των ψήφων σε γειτονικά bins. Στην εικόνα με $V_{min} = 15000$ απομένει μόνο ένας κύκλος, ο σωστός.

Τελικά Αποτελέσματα

Οι δύο καλύτερες εικόνες σε μεγένθυση και με linewidth 2 για να φαίνεται καθαρά το κυκλικό περίγραμμα.

Circle Detection with Hough (Sobel Edges)

$V_{\min} = 22000$



1991934907

Circle Detection with Hough (LoG Edges)

$V_{\min} = 15000$



1991934907