

Όραση Υπολογιστών
Εργαστήριο 1

Αναστάσιος Στέφανος Αναγνώστου
03119051

Σπυρίδων Παπαδόπουλος
03119058

7 Απριλίου 2023

Περιεχόμενα

1	Μέρος 1ο	3
1.1	Δημιουργία Εικόνων Εισόδου	3
1.2	Υλοποίηση Αλγορίθμων Ανίχνευσης Ακμών	3
1.2.1	Γραμμική Μέθοδος	3
1.2.2	Μη Γραμμική Μέθοδος	3
1.3	Αξιολόγηση των Αποτελεσμάτων Ανίχνευσης Ακμών	4
1.4	Αποτελέσματα Εφαρμογής της Συνάρτησης	4
1.5	Εφαρμογή των Αλγορίθμων Ανίχνευσης Ακμών σε Πραγματικές Εικόνες	7
2	Μέρος 2ο	8
2.1	Ανίχνευση Γωνιών	8
2.2	Πολυκλιμακωτή Ανίχνευση Γωνιών	10
2.3	Ανίχνευση Blobs	11
2.4	Πολυκλιμακωτή Ανίχνευση Blobs	13
2.5	Επιτάχυνση Ανίχνευσης Blobs	16
2.6	Επιτάχυνση Πολυκλιμακωτής Ανίχνευση Blobs	18
3	Μέρος 3ο	19
3.1	Ταίριασμα Εικόνων υπό Περιστροφή και Αλλαγή Κλίμακας	19
3.2	Κατηγοριοποίηση Εικόνων	20
4	Παράρτημα	23
4.1	Επιτάχυνση χρήσει ολοκληρωτικών εικόνων και box filters	23

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:

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

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.

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

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:

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

1.2.2 Μη Γραμμική Μέθοδος

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

$$I_\sigma = G_\sigma * I, \quad (1.4)$$
$$L = I_\sigma \oplus B + I_\sigma \ominus B - 2I_\sigma$$

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.

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

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

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

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 κατωφλιού ως εξής:

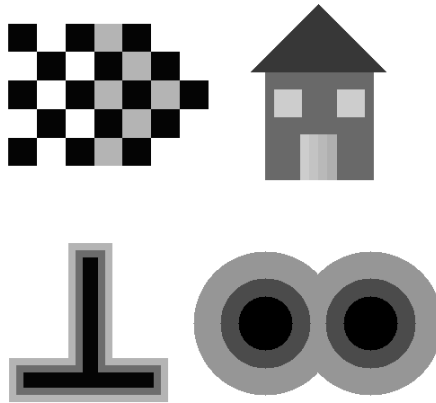
$$M = (I \oplus B) - (I \ominus B) \Rightarrow T = (M > \theta_{real}) \tag{1.7}$$

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

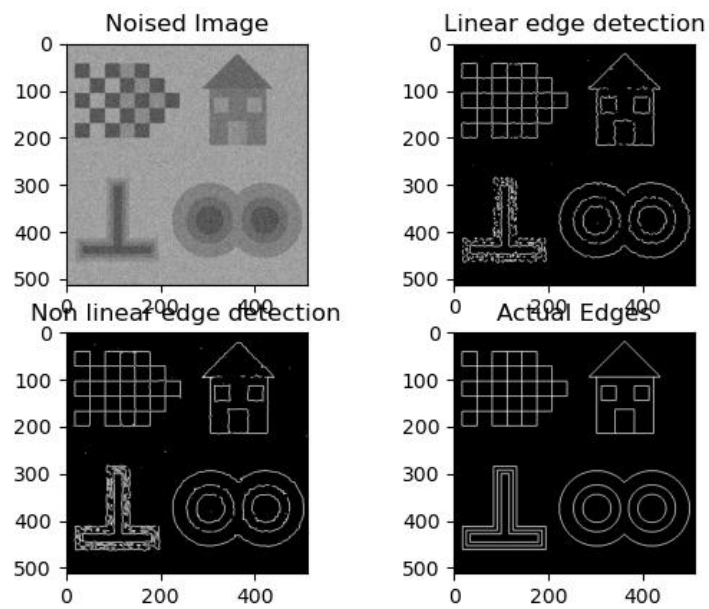
$$\begin{aligned}
C &= \frac{Pr(D|T) + PR(T|D)}{2}, \\
Pr(T|D) &= \frac{\|D \cap T\|}{\|T\|}
\end{aligned} \tag{1.8}$$

1.4 Αποτελέσματα Εφαρμογής της Συνάρτησης

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



Σχήμα 1: Η δεδομένη εικόνα εισόδου

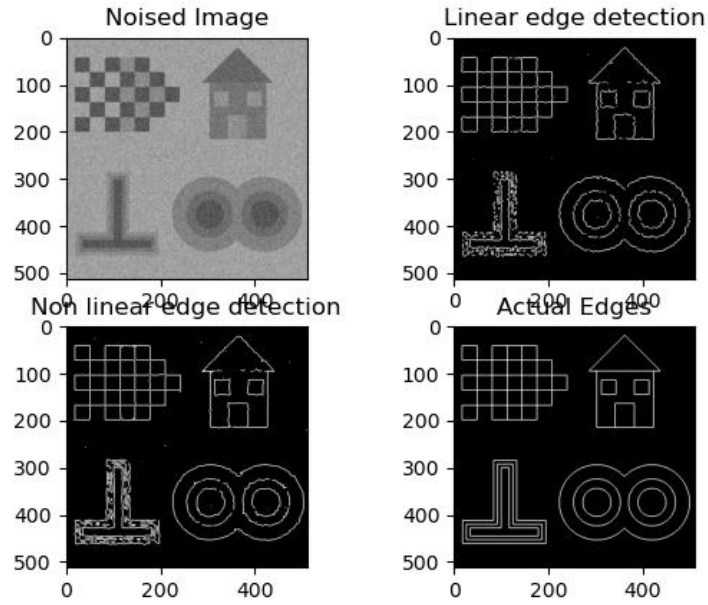


Σχήμα 2: Θόρυβος PSNR 10dB

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:

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

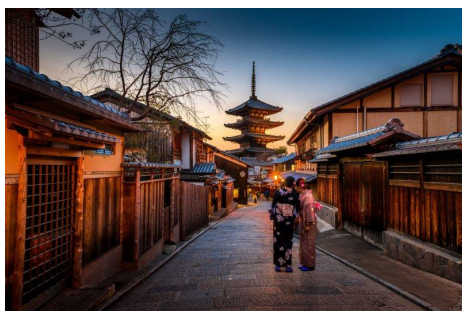
Τα αποτελέσματα επιτεύχθηκαν με τις ακόλουθες παράμετρους:

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

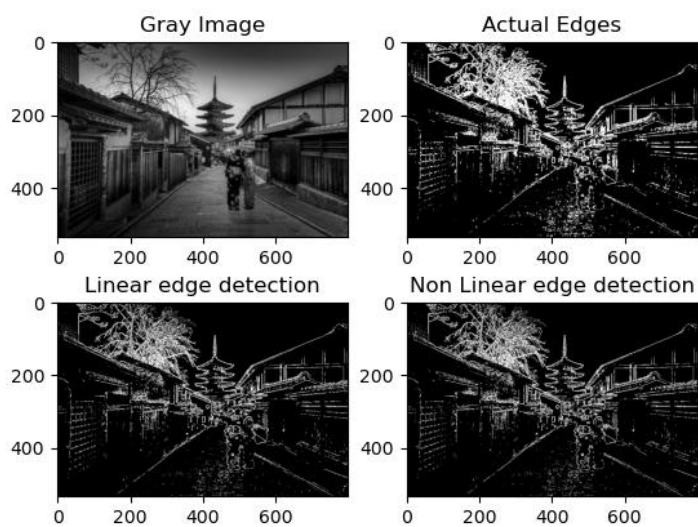
Παρατηρείται ότι η επίδραση του θορύβου είναι καθοριστική, καθώς στην εικόνα με τον λιγότερο θόρυβο επιτυγχάνεται καλύτερη βαθμολογία στο ποιοτικό κριτήριο. Επίσης, φαίνεται ότι η μη γραμμική μέθοδος πετυχαίνει καλύτερα αποτελέσματα σε κάθε περίπτωση.

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.

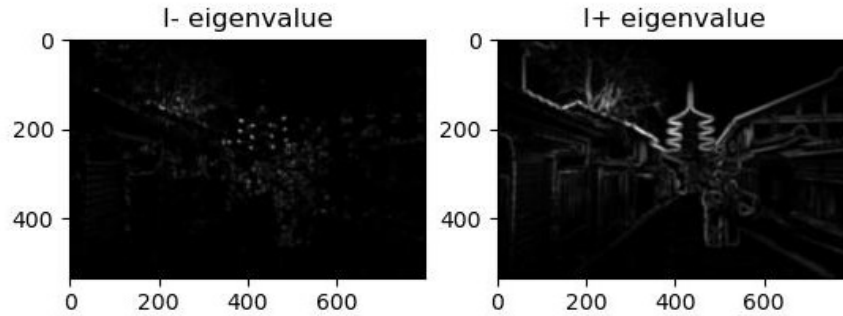
2.1 Ανίχνευση Γωνιών

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

Αρχικά υπολογίζεται ο δομικός ταυστής \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 γκρίζες εικόνες.



Σχήμα 6: Απεικόνιση των ιδιοτιμών της εικόνας ως γκρίζες 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 επιλεχθούν 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 σχεπτικό 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

Η ανίχνευση blobs αναφέρεται γενικώς στην ανίχνευση περιοχών με κάποια ομοιο-
γένεια. Οι περιοχές αυτές βρίσκονται χρήσει των μερικών παραγώγων δευτέρας
τάξεως της εξομαλυσμένης εικόνας. Συγκεκριμένα, δεδομένου του Χεσιανού πίνα-
κα:

$$H(x, y) = \begin{bmatrix} L_{xx}(x, y, \sigma) & L_{xy}(x, y, \sigma) \\ L_{xy}(x, y, \sigma) & L_{yy}(x, y, \sigma) \end{bmatrix} \quad (2.6)$$

Το κριτήριο για τον εντοπισμό περιοχών ομοιογένειας είναι απλώς ο μηδενισμός
της ορίζουσας:

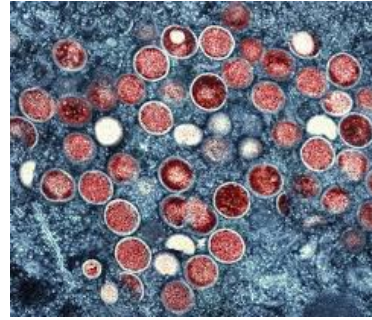
$$\begin{aligned} R(x, y) &= \det |H(x, y)| \\ (x, y) &= \arg \min_{x, y} (R(x, y) = 0) \end{aligned} \quad (2.7)$$

Επομένως, αναπτύσσεται μεθοδολογία πλήρως ανάλογη με την μεθοδολογία για
την ανίχνευση ακμών. Δηλαδή, εξομαλύνεται η εικόνα, υπολογίζεται ο Χεσιανός
πίνακας και το κριτήριο $R(x, y)$, εντοπίζονται τα σημεία μηδενισμού και επιλέγονται
αυτά τα οποία έχουν τιμή στην εικόνα μεγαλύτερη από κάποιο κατώφλι.

Η μέθοδος εφαρμόζεται στην εικόνα U_p , από την ομώνυμη ταινία της Pixar,
και σε μία εικόνα ανθρωπίνων κυττάρων. Πρώτη θα εξεταστεί η εφαρμογή στην
εικόνα U_p .



(α') Η εικόνα U_p



(β') Η εικόνα ανθρωπίνων κυττάρων

Σχήμα 9: Οι αρχικές εικόνες για την ανίχνευση blobs

Παρακάτω, στην εικόνα 10 φαίνονται τα αποτελέσματα για παραμέτρους $\sigma = 2.5, \theta = 0.005$

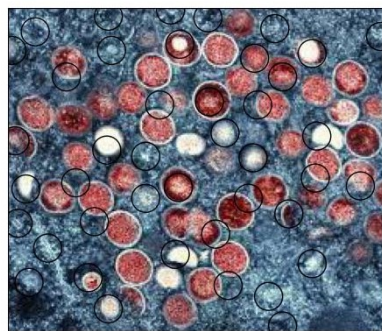


Σχήμα 10: Ανίχνευση blobs στην εικόνα Up

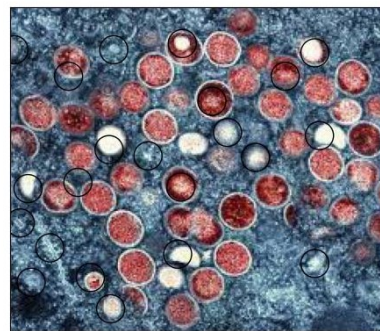
Παρατηρείται, ότι αν και το κατώφλι για την αποδοχή της περιοχής ομοιογένειας είναι πολύ χαμηλό, τα αποτελέσματα είναι πολύ ικανοποιητικά. Συγκεκριμένα, εντοπίζονται οι περιοχές στα μπαλόνια οι οποίες αντανακλούν το φως του ηλίου και έχουν χαρακτηριστική λευκή λάμψη. Επιπλέον, εντοπίζονται οι περιοχές μεταξύ των μπαλονιών στις οποίες υπάρχει σκιά.

Εξετάζεται τώρα η εφαρμογή της μεθόδου στην εικόνα των κυττάρων.

Τα αποτελέσματα για τις ίδιες παραμέτρους με πριν είναι αυτά που φαίνονται στην εικόνα 11α΄



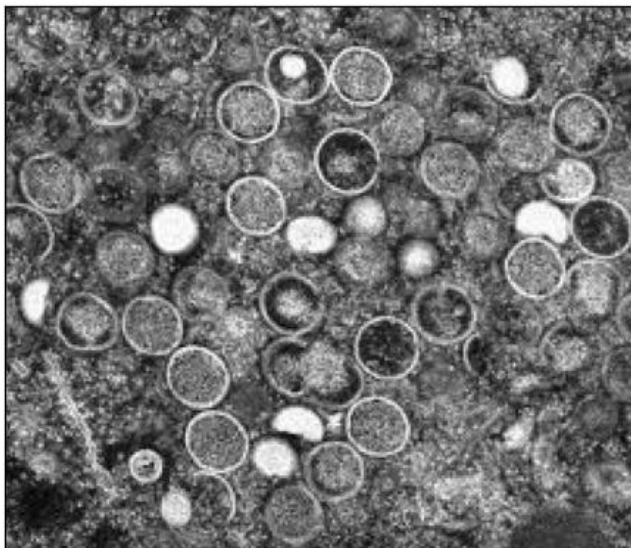
(α΄) Ανίχνευση blobs στην εικόνα ανθρω-
πίνων κυττάρων χαμηλό κατώφλι



(β΄) Ανίχνευση blobs στην εικόνα ανθρω-
πίνων κυττάρων υψηλό κατώφλι

Παρατηρείται, ότι αν και ανιχνεύονται επιτυχώς μερικές από τις λευκές περιοχές

ομοιογένειας, ανιχνεύονται επίσης ανεπιθυμητές περιοχές. Αυτό συμβαίνει, γιατί το περιβάλλον της εικόνας είναι τραχύ, με αποτέλεσμα οποιαδήποτε ομοιογένεια να ξεχωρίζει. Αντιθέτως, ο ουρανός προηγουμένως είναι πολύ ομαλός και ομοιογενής, επομένως δεν ξεχωρίζει. Για αυτόν τον λόγο, προσαρμόζεται η παράμετρος $\theta \leftarrow 0.25$. Τα νέα αποτελέσματα φαίνονται στην εικόνα 11β', όπου πλέον ανιχνεύονται σαφώς λιγότερες περιοχές ομοιογένειας, αλλά ορθότερες.



Σχήμα 12: Η εικόνα των κυττάρων γκρίζα

Αξιοσημείωτο είναι το γεγονός ότι οι ερυθρές περιοχές γενικώς δεν ανιχνεύονται, σίγουρα όχι στον ίδιο βαθμό με τις λευκές. Αυτό συμβαίνει γιατί η επεξεργασία γίνεται στην γκρίζα μορφή της εικόνας, επομένως η χρωματική διαφορά απαλείφεται, όπως φαίνεται στην εικόνα 12. Εφόσον η υφή είναι η ίδια, ανεξαρτήτως χρώματος, γίνεται σαφές ότι δεν μπορούν να διακριθούν στον ίδιο βαθμό με τις λευκές. Εκτός από αυτό, συμπεραίνεται ότι σε τραχύ περιβάλλον πρέπει να αυξηθεί το κατώφλι αποδοχής σημείων.

2.4 Πολυκλιμακωτή Ανίχνευση Blobs

Για την βελτίωση της μεθόδου χρησιμοποιούνται τώρα πολλαπλές κλίμακες, σε πλήρη αναλογία με την πολυκλιμακωτή ανίχνευση γωνιών.

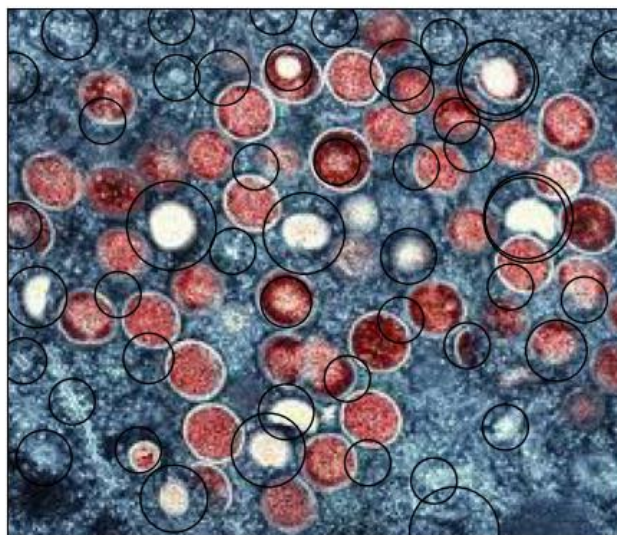
Στην εικόνα 13 φαίνονται τα αποτελέσματα στην εικόνα Up για τις τιμές παραμέτρων $\sigma = 3, \theta = 0.05, s = 1.1, N = 8$. Παρατηρείται ότι εντοπίζονται blobs ποικίλων μεγεθών, στα οποία περιλαμβάνονται τόσο ολόκληρα μπαλόνια όσο και απλές αντανακλάσεις φωτός, όπως προηγουμένως.



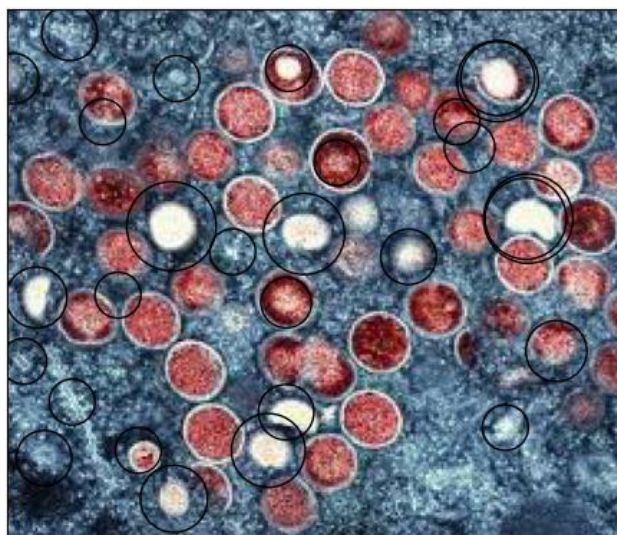
Σχήμα 13: Πολυκλιμακωτή ανίχνευση blobs στην εικόνα Up

Στην εικόνα 14α' φαίνεται το αποτέλεσμα της πολυκλιμακωτής ανίχνευσης blobs για τις τιμές παραμέτρων $\sigma = 3, \theta = 0.05, s = 1.1, N = 8$. Η ανίχνευση γίνεται σε 8 κλίμακες του $\sigma = 3$ με παράγοντα το 1.1 και χαμηλό κατώφλι. Αντιστοίχως, στην εικόνα 14β' φαίνεται το αποτέλεσμα της πολυκλιμακωτής ανίχνευσης blobs για τις τιμές παραμέτρων $\sigma = 3, \theta = 0.15, s = 1.1, N = 8$. Δηλαδή, η ανίχνευση γίνεται σε 8 κλίμακες του $\sigma = 3$ και με υψηλότερο κατώφλι, με παράγοντα το 1.1.

Στην δε εικόνα 14 παρατηρείται, ότι μειώνοντας το κατώφλι, όπως και στην μονοκλιμακωτή ανίχνευση, απορρίπτονται επιτυχώς κάποια ανεπιθύμητα blobs. Μάλιστα, αυτό γίνεται εν προκειμένω με μεγαλύτερη επιτυχία, καθώς εξακολουθούν να ανιχνεύονται όλες οι λευκές περιοχές. Δεν απορρίπτεται, δηλαδή, κατά λάθος, κάποιο επιθυμητό blob. Επομένως, συμπεραίνεται ότι η επέκταση της μεθόδου σε πολλαπλές κλίμακες είναι ευεργετική.



(α') Χαμηλό κατώφλι



(β') Υψηλό κατώφλι

Σχήμα 14: Πολυκλιμακωτή ανίχνευση blobs σε κύτταρα

2.5 Επιτάχυνση Ανίχνευσης Blobs

Η παραπάνω μέθοδος μπορεί να επιταχυνθεί ¹ χρήσει ολοκληρωτικών εικόνων και box filter. Συγκεκριμένα, ο υπολογισμός των μερικών παραγώγων μέχρι στιγμής απαιτούσε την συνέλιξη φίλτρων με την εκάστοτε εικόνα, γεγονός που τον καθιστούσε υπολογιστικά ακριβό.

Οι ολοκληρωτικές εικόνες είναι μία τεχνική η οποία επιτρέπει των υπολογισμό συνέλιξεων με την εικόνα σε χρόνο ανεξάρτητο του πυρήνα και εξαρτώμενο μόνο από το μέγεθος της εικόνας. Αυτό συμβαίνει, γιατί, χρήσει ολοκληρωτικών εικόνων, το άθροισμα των pixels μίας ορθογώνιας περιοχής μπορεί να υπολογιστεί σε σταθερό χρόνο. Για να επιτευχθεί, όμως, η επιτάχυνση αυτή, χρησιμοποιείται μία προσέγγιση της παραγώγου. Επομένως, πρόκειται για προσεγγιστική μέθοδο και αναμένονται αποκλίσεις από τα προηγούμενα αποτελέσματα.

Στην εικόνα 15 φαίνονται τα αποτελέσματα της απόπειρας επιτάχυνσης της ανίχνευσης blobs στις ίδιες εικόνες όπως προηγουμένως, στην οποία χρησιμοποιήθηκαν οι ίδιες ακριβώς παράμετροι όπως και στην απλή μέθοδο.

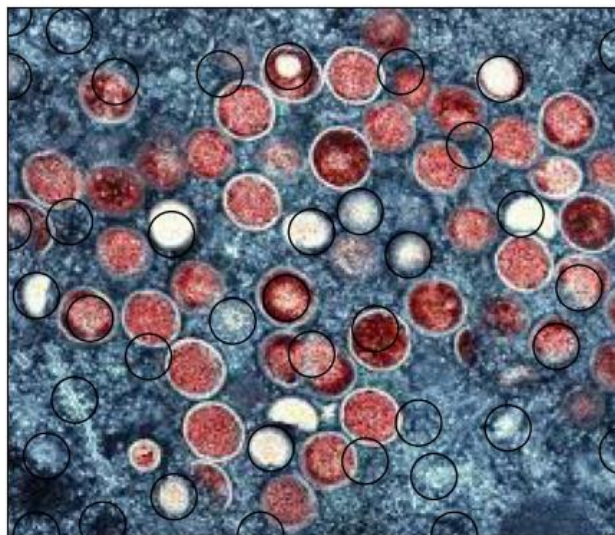
Στην εικόνα 15α', όπου χρησιμοποιείται μικρή τιμή της παραμέτρου σ , τα αποτελέσματα είναι πολύ ικανοποιητικά, καθώς εντοπίζονται επιτυχώς οι αντανακλάσεις φωτός επάνω στα μπαλόνια και, φυσικά, δεν αντιμετωπίζονται εσφαλμένες περιοχές ομοιογένειας.

Στην δε εικόνα 15β', όπου χρησιμοποιείται μεγαλύτερη τιμή της παραμέτρου σ , τα αποτελέσματα δεν είναι το ίδιο ικανοποιητικά. Συγκεκριμένα, παρατηρείται ότι ανιχνεύονται μερικά ανεπιθύμητα blobs, ωστόσο ανιχνεύονται και μερικά ορθώς. Επομένως, είναι εμφανές το γεγονός ότι η επιτάχυνση οφείλεται σε προσέγγιση.

¹Σημειώνεται ότι η παρατηρούμενη επιτάχυνση εξαρτάται από την υλοποίηση της απλής μεθόδου ανίχνευσης blobs. Βλέπε 4.1.



(α') Αποτέλεσμα στην εικόνα U_p



(β') Αποτέλεσμα στην εικόνα κυττάρων

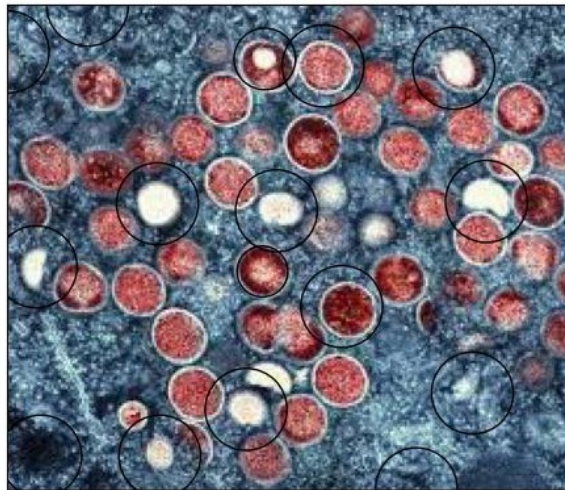
Σχήμα 15: Προσέγγιση ανίχνευσης blobs με ολοκληρωτικές εικόνες

2.6 Επιτάχυνση Πολυκλιμακωτής Ανίχνευση Blobs

Τα ίδια συμπεράσματα προκύπτουν και για την πολυκλιμακωτή ανίχνευση.



(α') Αποτέλεσμα στην εικόνα Up



(β') Αποτέλεσμα στην εικόνα κυττάρων

Σχήμα 16: Προσέγγιση πολυκλιμακωτής ανίχνευσης blobs με ολοκληρωτικές ει-
κόνες

3 Μέρος 3ο

3.1 Ταίριασμα Εικόνων υπό Περιστροφή και Αλλαγή Κλίμακας

Σε αυτό το μέρος επιχειρείται το ταίριασμα μεταξύ εικόνων και μετασχηματισμών αυτών. Συγκεκριμένα, 3 αυθεντικές εικόνες υπόκεινται σε περιστροφές ($-20^\circ, -10^\circ, 0^\circ, 10^\circ, 20^\circ$) και σε κλιμακώσεις (0.6, 0.8, 1.0, 1.2). Δεδομένης μίας εικόνας αναφοράς, επιχειρείται η αντιστοίχιση των τοπικών περιγραφητών της με αυτούς των μετασχηματισμένων εικόνων.

Σημειώνεται, ότι οι ανιχνευτές είναι ως εξής:

Πίνακας 1: Αντιστοίχιση μονοκλιμακωτών και πολυκλιμακωτών μεθόδων.

Μονοκλιμακωτή	Πολυκλιμακωτή
CornerDetection	HarrisLaplacian
BlobDetection	HessianLaplacian
-	BoxLaplacian

Παρακάτω παρατίθενται τα αποτελέσματα ανά περιγραφητή ανά ανιχνευτή για τις 3 εικόνες.

```
featuresSURF, CornerDetection : [0.003 0.002 0.097]
featuresSURF, HarrisLaplacian : [0.001 0.002 0.002]
featuresSURF, BlobDetection : [0.027 0.01 0.001]
featuresSURF, HessianLaplacian : [0.002 0.001 0.001] BEST
featuresSURF, BoxLaplacian : [0.073 0.086 0.003]
featuresHOG, CornerDetection : [0.186 0.351 0.285]
featuresHOG, HarrisLaplacian : [0.202 0.137 0.717]
featuresHOG, BlobDetection : [0.186 0.1 0.154]
featuresHOG, HessianLaplacian : [0.162 0.114 0.253]
featuresHOG, BoxLaplacian : [0.856 0.283 0.185]
```

Σχήμα 17: Σφάλματα Κλίμακας

Παρατηρείται, ότι οι περιγραφητές SURF αποδίδουν κατά πολύ καλύτερα από τους περιγραφητές HOG. Αυτό είναι αναμενόμενο, αφού εξάγουν περιστροφικά και ως προς την κλίμακα ανεξάρτητους περιγραφητές.

Με BEST επισημειώνεται το ζευγάρι περιγραφητή και ανιχνευτή το οποίο έχει το ελάχιστο μέσο σφάλμα στις 3 εικόνες. Παρατηρείται, τόσο στην γωνία όσο και στην κλίμακα είναι το ζευγάρι πολυκλιμακωτής ανίχνευσης blobs με SURF. Αυτό είναι αναμενόμενο, αφού τα blobs είναι γενικές περιοχές ομοιομορφίας και δεν αλλοιώνονται τόσο με την περιστροφή, σε αντίθεση με τις γωνίες, οι οποίες

```

featuresSURF, CornerDetection : [ 1.968  0.318 12.909]
featuresSURF, HarrisLaplacian : [0.077 0.169 0.099]
featuresSURF, BlobDetection : [7.759 0.229 0.054]
featuresSURF, HessianLaplacian : [0.064 0.11  0.05 ] BEST
featuresSURF, BoxLaplacian : [ 7.707 11.362  0.463]
featuresHOG, CornerDetection : [22.619 19.199 23.699]
featuresHOG, HarrisLaplacian : [24.426 25.926 25.471]
featuresHOG, BlobDetection : [ 7.231 13.674 27.219]
featuresHOG, HessianLaplacian : [25.439 18.598 13.546]
featuresHOG, BoxLaplacian : [16.661 23.906 14.184]

```

Σχήμα 18: Σφάλματα Γωνίας

μπορούν να συμπέσουν πιο εύκολα με άλλες γωνίες και να επικρατήσει σύγχυση. Αξιοσημείωτο είναι, ότι το ζευγάρι αυτό είναι και από τα πιο χρονοβόρα, ενώ η αντίστοιχη υλοποίηση χρήσει ολοκληρωτικών εικόνων και box filters (ΒοξΛαπλασιαν) είναι ταχύτερη, αλλά λιγότερο ακριβής, πράγμα αναμενόμενο αφού πρόκειται για προσέγγιση.

3.2 Κατηγοριοποίηση Εικόνων

Σε αυτό το μέρος επιχειρείται η κατηγοριοποίηση εικόνων βάσει των εξαχθέντων χαρακτηριστικών τους. Συγκεκριμένα, δοκιμάζονται συνδυασμοί περιγραφητών και ανιχνευτών σημείων ενδιαφέροντος για την εξαγωγή χαρακτηριστικών από καθεμία εικόνα καθεμίας κλάσης. Στην συνέχεια, διαχωρίζονται σε σύνολα εκπαίδευσης και επικύρωσης και δημιουργούνται οι ανάλογες ετικέτες ταξινόμησης.

Για την ταξινόμηση πρέπει να μετατραπούν οι τοπικοί περιγραφητές σε καθολικά διανύσματα. Αυτό επιτυγχάνεται μέσω της αναπαράστασης Bag of Visual Words. Συγκεκριμένα, δημιουργείται ένα οπτικό 'λεξικό' συνενώνοντας τους περιγραφητές του συνόλου εκπαίδευσης και εφαρμόζοντας τον αλγόριθμο συσταδοποίησης kmeans σε ένα υποσύνολο αυτών. Έτσι δημιουργούνται κέντρα / αντιπρόσωποι χαρακτηριστικών. Άρα, το επόμενο βήμα είναι να μετρηθεί η ευκλίδεια απόσταση καθενός περιγραφητή του συνόλου εκπαίδευσης και επικύρωσης από τα κέντρα αυτά και να επιλεγεί η ελάχιστη. Τέλος, κατασκευάζεται το ιστόγραμμα της εικόνας βάσει της συχνότητας εμφάνισης των οπτικών λέξεων και κανονικοποιείται.

Η ίδια η κατηγοριοποίηση πραγματοποιείται από έναν ταξινομητή SVM.

Τα αποτελέσματα των δοκιμών, χρήσει της προμεταγλωττισμένης συνάρτησης (500 ομάδες kmeans) BagOfWords είναι τα ακόλουθα:

BlobDetection featuresHOG: 66.345
BlobDetection featuresSURF: 55.862
BoxLaplacian featuresHOG: 61.517
BoxLaplacian featuresSURF: 53.379
CornerDetection featuresHOG: 63.724
CornerDetection featuresSURF: 51.862
HarrisLaplacian featuresHOG: 66.483
HarrisLaplacian featuresSURF: 59.310
HessianLaplacian featuresHOG: 63.586
HessianLaplacian featuresSURF: 57.103

Σχήμα 19: Αποτελέσματα Ταξινόμησης (Προμεταγλωττισμένες Συναρτήσεις, 500 ομάδες)

Τα αποτελέσματα των δοκιμών, χρήσει υλοποιημένης συνάρτησης από τους γράφοντες (1000 ομάδες kmeans) BagOfWords είναι τα ακόλουθα:

BlobDetection featuresHOG: 64.552
BlobDetection featuresSURF: 54.345
BoxLaplacian featuresHOG: 62.207
BoxLaplacian featuresSURF: 52.552
CornerDetection featuresHOG: 61.379
CornerDetection featuresSURF: 56.966
HarrisLaplacian featuresHOG: 65.103
HarrisLaplacian featuresSURF: 57.793
HessianLaplacian featuresHOG: 65.517
HessianLaplacian featuresSURF: 59.586

Σχήμα 20: Αποτελέσματα Ταξινόμησης (Πρωτότυπες Συναρτήσεις, 1000 ομάδες)

Τα αποτελέσματα των δοκιμών, χρήσει υλοποιημένης συνάρτησης από τους γράφοντες (500 ομάδες kmeans) BagOfWords είναι τα ακόλουθα:

BlobDetection featuresHOG: 65.931
 BlobDetection featuresSURF: 53.655
 BoxLaplacian featuresHOG: 65.931
 BoxLaplacian featuresSURF: 54.483
 CornerDetection featuresHOG: 64.690
 CornerDetection featuresSURF: 54.621
 HarrisLaplacian featuresHOG: 64.828
 HarrisLaplacian featuresSURF: 58.207
 HessianLaplacian featuresHOG: 64.690
 HessianLaplacian featuresSURF: 56.552

Σχήμα 21: Αποτελέσματα Ταξινόμησης (Πρωτότυπες Συναρτήσεις, 500 ομάδες)

Γενικώς, παρατηρείται, ότι οι περιγραφητές HOG αποδίδουν σταθερά καλύτερα από τους περιγραφητές SURF. Παρατηρείται, επίσης, ότι οι πολυκλιμακωτοί ανιχνευτές δεν αποδίδουν σταθερά καλύτερα από τους μονοκλιμακωτούς. Συγκεκριμένα, φαίνεται ότι οι περιγραφητές HOG δεν ευνοούνται ιδιαίτερα από την πολυκλιμακωτή ανάλυση, σε αντίθεση με τους SURF οι οποίοι ευνοούνται κατά πολύ, όπως φαίνεται στον πίνακα 2 εν συντομία². Από κάποιες μικρές διακυμάνσεις δεν μπορούν να εξαχθούν με βεβαιότητα συμπεράσματα, καθώς τα αποτελέσματα εξαρτώνται και από την τυχαία επιλογή χαρακτηριστικών στον χμεανς. Συγκεκριμένα, μπορεί κάποια τυχαία επιλογή χαρακτηριστικών να οδήγησε στον σχηματισμό καλύτερων ομάδων.

Πίνακας 2: Αντιστοίχιση μονοκλιμακωτών και πολυκλιμακωτών μεθόδων για SURF περιγραφητή.

Ομάδες	Μονοκλιμακωτή	Επίδοση	Πολυκλιμακωτή	Επίδοση	Διαφορά
500	CornerDetection	54.621	HarrisLaplacian	59.310	4.689
500	BlobDetection	55.862	HessianLaplacian	57.103	1.241
1000	CornerDetection	56.966	HarrisLaplacian	59.310	2.344
1000	BlobDetection	54.345	HessianLaplacian	59.586	5.241

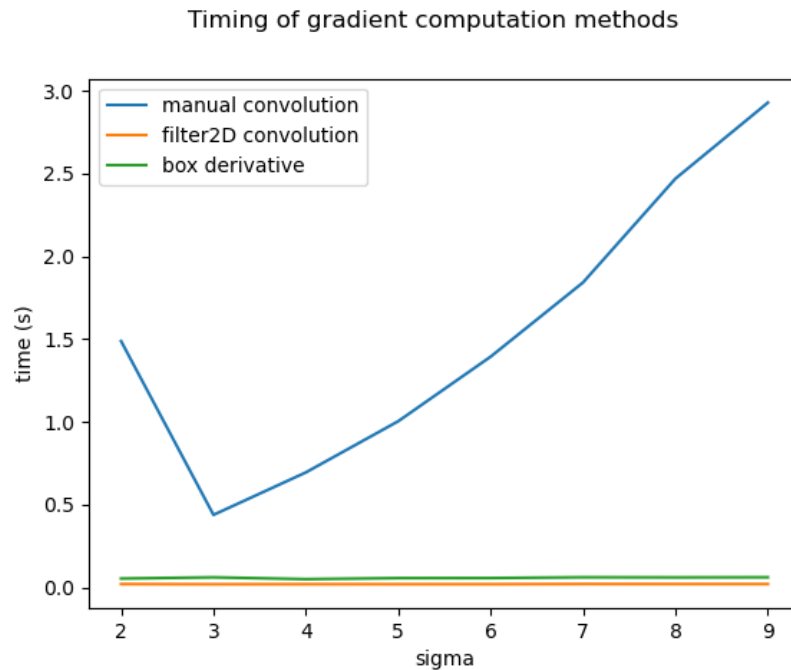
Συμπερασματικά, φαίνεται ότι την καλύτερη επίδοση την έχουν, σταθερά, οι περιγραφητές HOG, οι οποίοι αποδίδουν κατά 5 με 10 ποσοστιαίες μονάδες καλύτερα από τους SURF. Σχετικά με τους ανιχνευτές, οι διαφορές στα αποτελέσματα είναι πολύ μικρές.

²Για να είναι δίκαια η σύγκριση, έγινε μεταξύ των πρωτότυπων υλοποιήσεων

4 Παράρτημα

4.1 Επιτάχυνση χρήσει ολοκληρωτικών εικόνων και box filters

Η παρατηρούμενη επιτάχυνση εξαρτάται και από την υλοποίηση της απλής μεθόδου ανίχνευσης blobs. Συγκεκριμένα, αν ο υπολογισμός των μερικών παραγώγων είχε γίνει χρήσει 'χειροποίητης' συνάρτησης συνελίξεως δύο διαστάσεων, τότε η επιτάχυνση γίνεται αμέσως εμφανής. Ωστόσο, αν ο υπολογισμός μερικών παραγώγων είχε γίνει χρήσει έτοιμης συνάρτησης συνελίξεως από την βιβλιοθήκη `cv2`, τότε η επιτάχυνση δεν θα είναι εμφανής. Αυτό συμβαίνει γιατί, κατόπιν μετρήσεων για διάφορες τιμές τις παραμέτρου σ , παρατηρήθηκε ότι η ταχύτητα της συνάρτησης `cv2.filter2D` ήταν πρακτικά σταθερή. Το ίδιο παρατηρήθηκε και για την υλοποίηση χρήσει ολοκληρωτικών εικόνων και box filters. Η δε 'χειροκίνητη' υλοποίηση της διαδικασίας συνελίξεως εμφάνισε τετραγωνική χρονική εξάρτηση από την παράμετρο σ , όπως αναμενόταν. Τα συμπεράσματα παρουσιάζονται συνοπτικά στο ακόλουθο γράφημα 22. Αξιοσημείωτο είναι ότι, για την δίκαιη σύγκριση των αποτελεσμάτων, χρησιμοποιήθηκε ο Just-in-Time μεταγλωττιστής Numba για την βελτιστοποίηση των βρόχων `for` στην υλοποίηση της διαδικασίας συνελίξεως. Παρόλα αυτά, η διαφορά στην απόδοση παραμένει σημαντική.



Σχήμα 22: Σύγκριση ταχύτητας διαφόρων υλοποιήσεων υπολογισμού μερικών παραγώγων