

Ψηφιακή Επεξεργασία Εικόνας

Εργασία 3

Βασίλειος Αραϊλόπουλος
varailop@ece.auth.gr
AEM: 9475

Ιούλιος 2021

1 Gaussian Space Scale

Γενικά, σκοπός της εργασίας αυτής ήταν η υλοποίηση του πρώτου μέρους του αλγορίθμου SIFT για την εύρεση των ιδιαζόντων σημείων (salient points) μιας εικόνας. Πρώτο ζητούμενο αυτής της διαδικασίας είναι η δημιουργία των Gaussian Space Scales.

1.1 Κατασκευή του δισδιάστατου Gaussian φίλτρου

Η συνάρτηση που κατασκευάζει το φίλτρο δέχεται ως είσοδο την διάσταση K και την τυπική απόκλιση σ . Θεωρώ ως δεδομένο ότι το K που δίνεται στην συνάρτηση είναι μονός αριθμός ώστε το κέντρο του να βρίσκεται σε ακέραια θέση. Δηλαδή για μέγεθος φίλτρου ίσο με 5 το κέντρο του φίλτρου πρέπει να είναι στο $\text{floor}(5/2) + 1 = 3$, άρα στην θέση του πίνακα (3,3). Αν το K που δοθεί είναι ζυγός αριθμός κατασκευάζεται το φίλτρο με μέγεθος ίσο με τον επόμενο μονό αριθμό (π.χ. για $K = 4$ θα κατασκευαστεί το φίλτρο 5x5).

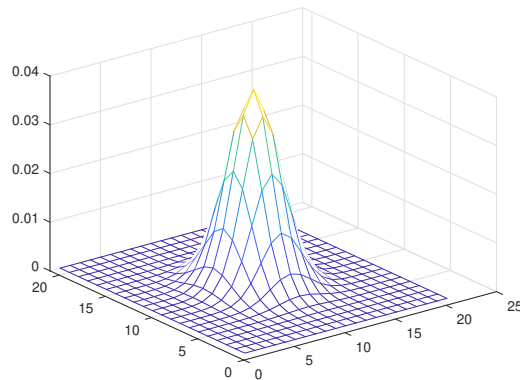


Figure 1: Το φίλτρο για $K = 21$ και $\sigma = 2$

1.2 Κατασκευή του Gaussian Space Scale

Σκοπός της συνάρτησης αυτής είναι η κατασκευή της Γκαουσιανής πυραμίδας και των Difference of Gaussians (DoGs) για να προετοιμαστεί το επόμενο στάδιο που είναι η εύρεση των τοπικών minima και maxima. Αναλυτικότερα, μία διπλή *for* τρέχει για κάθε octave και level που εισάγεται στην συνάρτηση. Για ένα octave βρίσκονται για όλα τα level του τα αντίστοιχα space scales σύμφωνα με την διαδικασία που περιγράφεται στην εκφώνηση, δηλαδή ένα space scale προκύπτει από το προηγούμενο του με συνέλιξη με το φίλτρο που έχει τυπική απόκλιση $\sqrt{2}\sigma'$ όπου σ' είναι η τυπική απόκλιση του φίλτρου που χρησιμοποιήθηκε για να προκύψει το προηγούμενο space scale. Δύο διαδοχικά space scales αφαιρούνται και δίνουν ένα DoG.

Όταν αλλάζει το octave, υποδειγματοληπτείται το τρίτο κατά σειρά space scale της προηγούμενης εικόνας. Αυτό είναι σωστό μόνο όταν ο συντελεστής k που πολλαπλασιάζεται με το σ' είναι ίσος με $\sqrt{2}$. Επίσης, όταν είναι να υπολογιστεί το πρώτο space scale ενός octave γίνεται συνέλιξη με το φίλτρο που έχει τυπική απόκλιση $2^{i-1}\sigma$ όπου i είναι ο αύξων αριθμός του octave. Τα space scales και τα DoGs αποθηκεύονται, όπως ζητείται και από την εκφώνηση, σε δύο cell arrays και τα αποτελέσματα που προκύπτουν φαίνονται παρακάτω. Να επισημανθεί ότι για να γίνουν διακριτά τα DoGs χρησιμοποιείται η συνάρτηση της MATLAB rescale που κανονικοποιεί τις τιμές στο διάστημα (0,1).

Συμπεράσματα:

- Είναι ξεκάθαρο ότι στα space scales όσο ανεβαίνουν τα levels τόσο πιο έντονο γίνεται το blur της εικόνας και όσο ανεβαίνουν τα octaves τόσο μειώνεται η ανάλυση της εικόνας και από ένα σημείο και μετά δεν υπάρχει νόημα να γίνεται υποδειγματοληψία, γιατί καταστρέφεται η εικόνα (όπως φαίνεται και από της εικόνες που υπάρχουν 7 octaves).
- Στα DoGs οι ακμές των αντικειμένων γίνονται πιο έντονες και ευδιάκριτες.

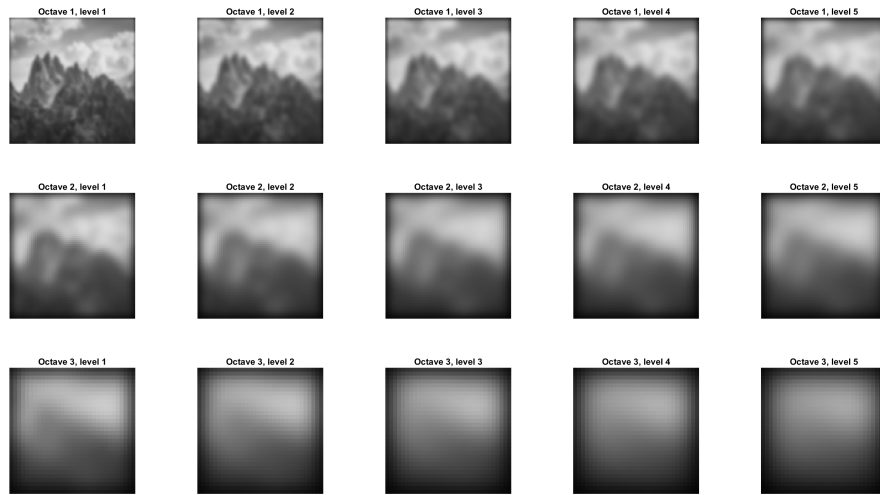


Figure 2: Τα space scales της εικόνας mountains για 3 octaves και 5 levels

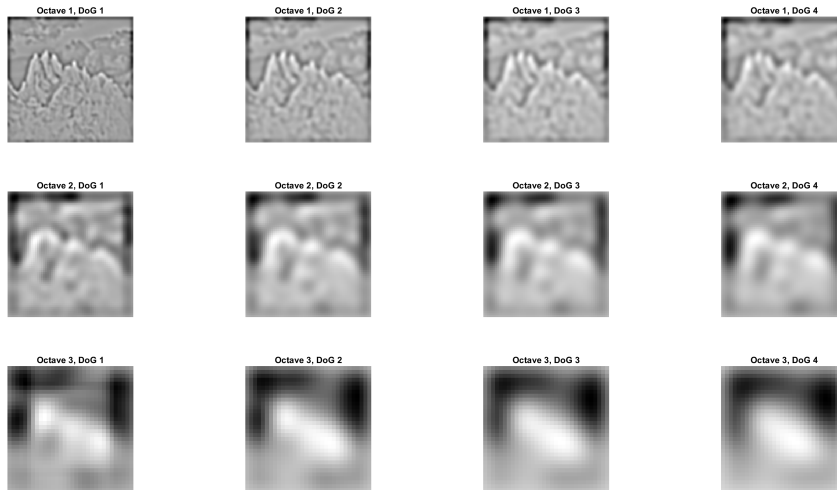


Figure 3: Τα DoGs της εικόνας mountains για 3 octaves και 5 levels

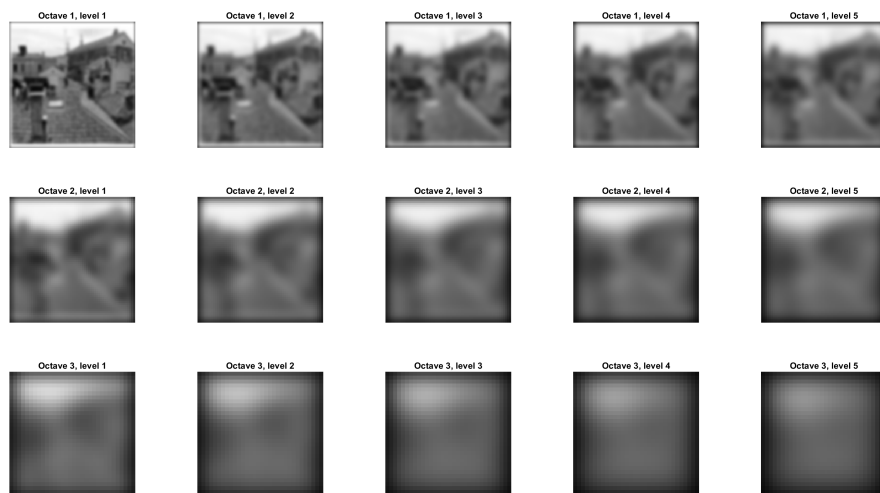


Figure 4: Τα space scales της εικόνας roofs για 3 octaves και 5 levels

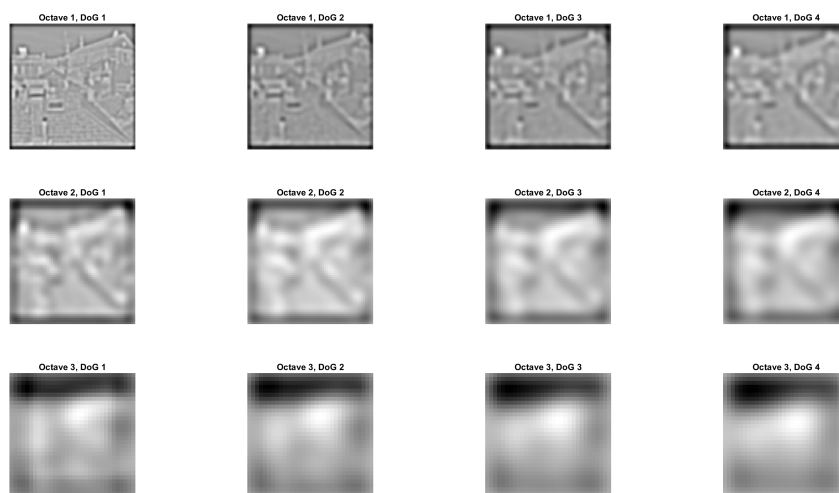


Figure 5: Τα DoGs της εικόνας roofs για 3 octaves και 5 levels

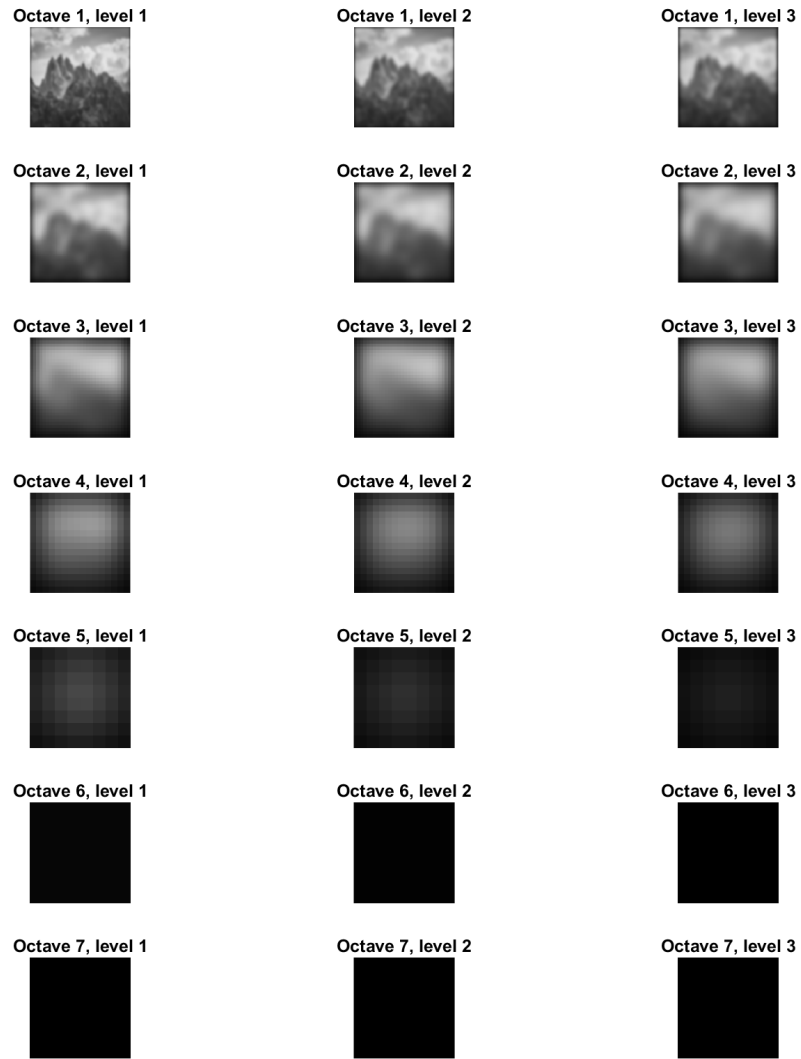


Figure 6: Τα space scales της εικόνας mountains για 7 octaves και 3 levels

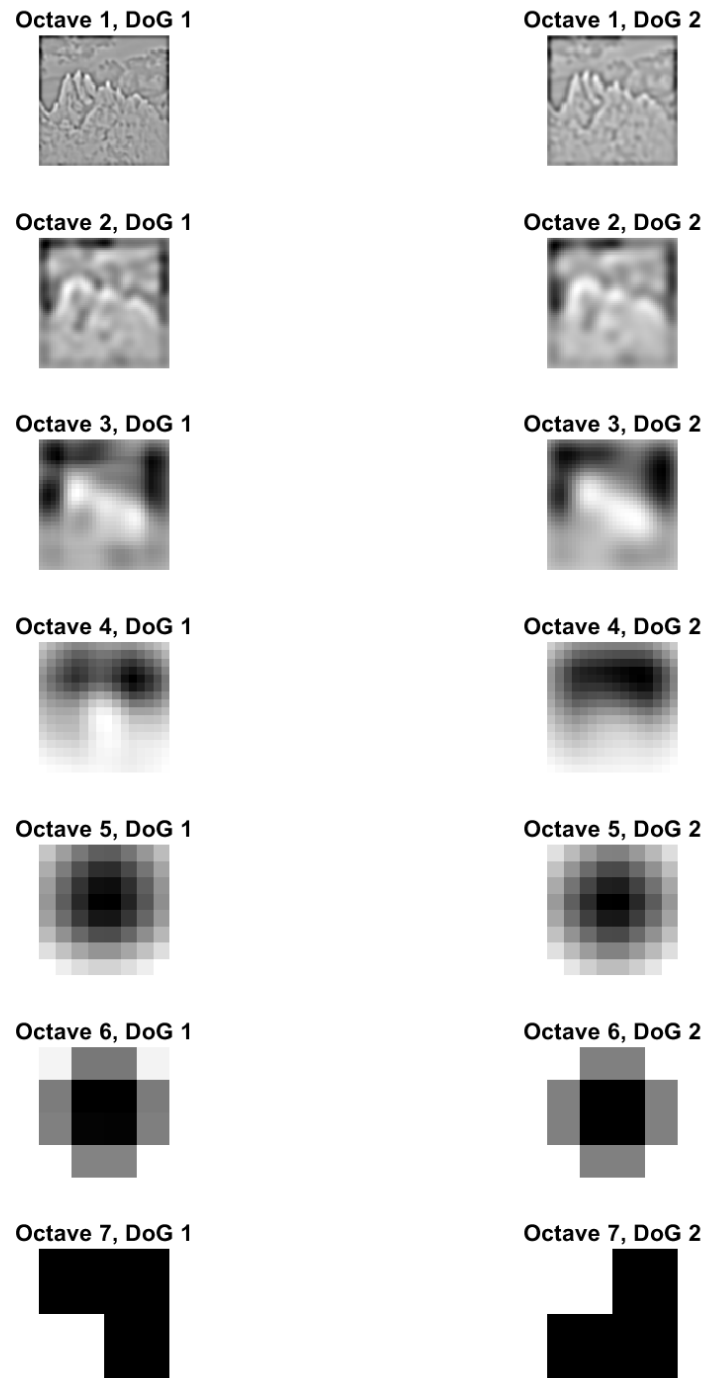


Figure 7: Τα DoGs της εικόνας mountains για 7 octaves και 3 levels

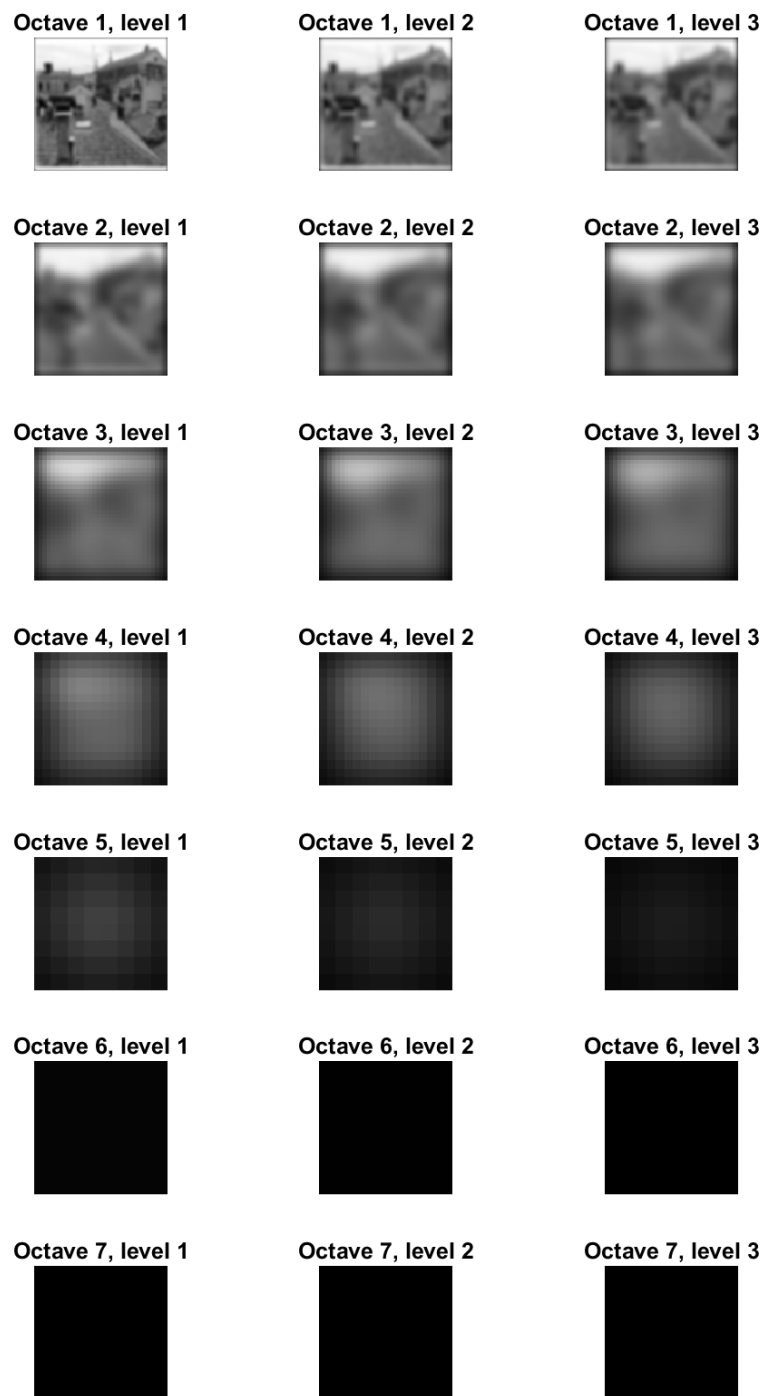
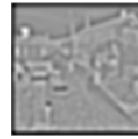


Figure 8: Τα space scales της εικόνας roofs για 7 octaves και 3 levels

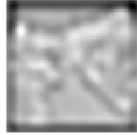
Octave 1, DoG 1



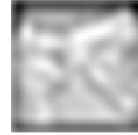
Octave 1, DoG 2



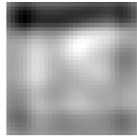
Octave 2, DoG 1



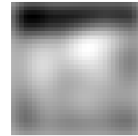
Octave 2, DoG 2



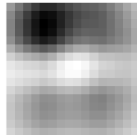
Octave 3, DoG 1



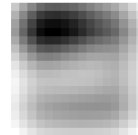
Octave 3, DoG 2



Octave 4, DoG 1



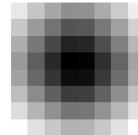
Octave 4, DoG 2



Octave 5, DoG 1



Octave 5, DoG 2



Octave 6, DoG 1



Octave 6, DoG 2



Octave 7, DoG 1



Octave 7, DoG 2



Figure 9: Τα DoGs της εικόνας roofs για 7 octaves και 3 levels

2 Εντοπισμός των salient keypoints

Σε αυτό το κομμάτι γίνεται το δεύτερο στάδιο της εύρεσης των τοπικών μεγίστων και ελαχίστων για να γίνει ο εντοπισμός των ιδιαζόντων σημείων.

2.1 Αφιλτράριστα σημεία

Σε αυτό το κομμάτι ελέγχονται όλα τα σημεία των DoGs που έχουν 26 γείτονες και εξετάζεται αν η τιμή τους είναι η μεγαλύτερη ή η μικρότερη μεταξύ των γειτόνων του. Αν βρεθεί ένα τέτοιο σημείο αποθηκεύεται σε μια σειρά του πίνακα keypoints το διάνυσμα (o,s,m,n) . Το o είναι ο αύξων αριθμός του octave, το s είναι ο αύξων αριθμός του DoG και m,n οι συντεταγμένες του σημείου στο DoG. Η έξοδος της συνάρτησης είναι δηλαδή ένας πίνακας $l \times 4$, όπου l είναι ο αριθμός των keypoints που βρέθηκαν.

2.2 Φιλτραρισμένα σημεία

Στην συνάρτηση αυτή απορρίπτονται τα σημεία που έχουν χαμηλή αντίθεση. Ελέγχονται όλα τα σημεία που έχουν βρεθεί από την προηγούμενη συνάρτηση και κρατιούνται όλα εκείνα που οι τιμές των σημείων, από το DoG που ανήκουν, ικανοποιούν την ανισότητα $|w| > pt$.

Αποτελέσματα και συμπεράσματα

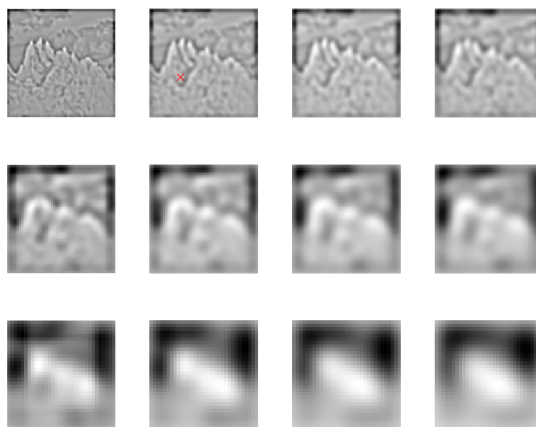


Figure 10: Τα DoGs της εικόνας mountains μαζί με τα αντίστοιχα keypoints



Figure 11: Η εικόνα mountains μαζί με τα αντίστοιχα φιλτραρισμένα keypoints

Στις εικόνες 10 και 11 είναι τα αποτελέσματα που προκύπτουν βάζοντας τα δεδομένα της εκφώνησης. Είναι δηλαδή για μέγεθος φίλτρου 7 και τυπική απόκλιση $\sqrt{2}$. Για την εικόνα roofs δεν βρίσκεται κανένα salient point.

Μεγαλώνοντας, όμως, το μέγεθος του φίλτρου σε 11 έχουμε λίγα περισσότερα keypoints.

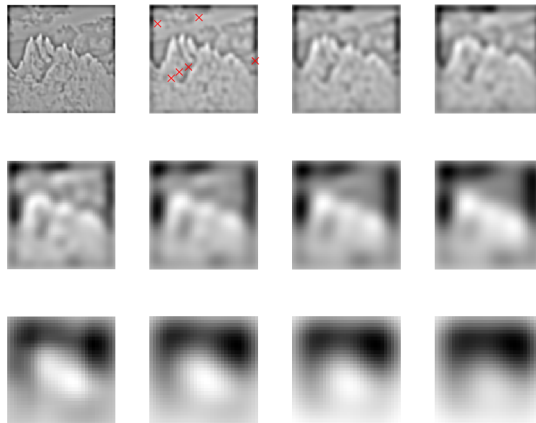


Figure 12: Τα DoGs της εικόνας mountains μαζί με τα αντίστοιχα keypoints



Figure 13: Η εικόνα mountains μαζί με τα αντίστοιχα φιλτραρισμένα keypoints

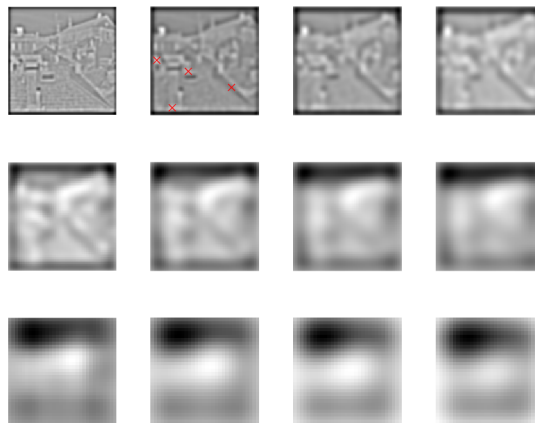


Figure 14: Τα DoGs της εικόνας roofs μαζί με τα αντίστοιχα keypoints



Figure 15: Η εικόνα roofs μαζί με τα αντίστοιχα φιλτραρισμένα keypoints

Βλέπουμε, όμως ότι και πάλι δεν είναι αρκετό για να βρεθούν τα keypoints των εικόνων.

Τώρα θα χρησιμοποιήσω την συνάρτηση της MATLAB imresize και θα κάνω τις εικόνες να έχουν μέγεθος 512x512.

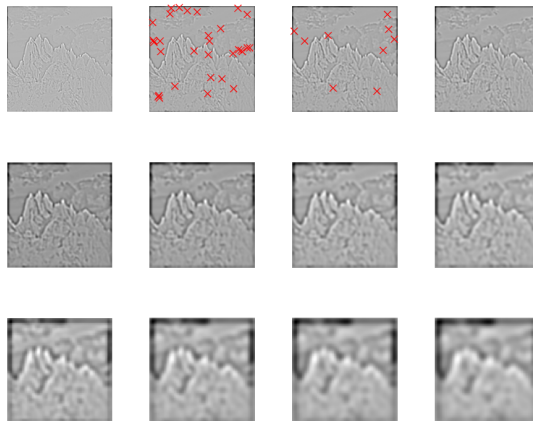


Figure 16: Τα DoGs της εικόνας mountains μαζί με τα αντίστοιχα keypoints



Figure 17: Η εικόνα mountains μαζί με τα αντίστοιχα φιλτραρισμένα keypoints

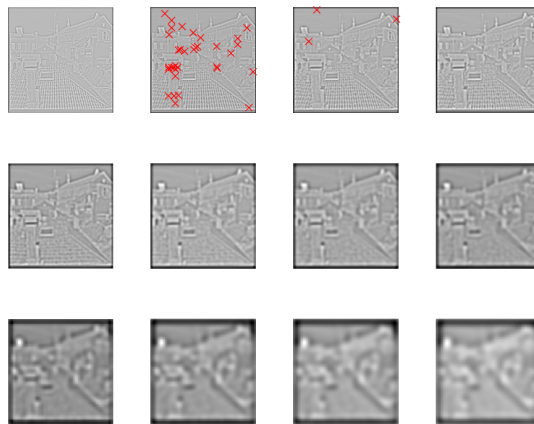


Figure 18: Τα DoGs της εικόνας roofs μαζί με τα αντίστοιχα keypoints



Figure 19: Η εικόνα roofs μαζί με τα αντίστοιχα φιλτραρισμένα keypoints

Για τις εικόνες 16 - 23, όπου το μέγεθος του φίλτρου είναι 7 και το $\sigma = \sqrt{2}$, είναι φανερό ότι ο αλγόριθμος αρχίζει και εφαρμόζεται καλύτερα. Ειδικά στην εικόνα με τις σκεπές αρχίζουν και σημαδεύονται σημεία στις σκεπές και στην καμινάδα μπροστά (εικόνα 23).

Αλλάζοντας το μέγεθος του φίλτρου στο 11 έχουμε:

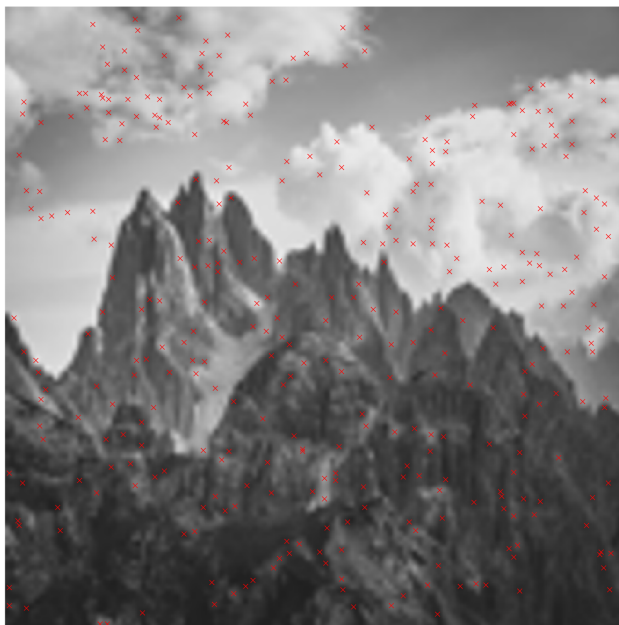


Figure 20: Η εικόνα mountains μαζί με τα αντίστοιχα φιλτραρισμένα keypoints



Figure 21: Η εικόνα roofs μαζί με τα αντίστοιχα φιλτραρισμένα keypoints

Σε αυτήν την περίπτωση έχει βελτιωθεί πολύ το αποτέλεσμα. Στην εικόνα με τα βουνά υπάρχουν μερικά σημεία στις κορυφογραμμές και στα περιγράμματα των σύννεφων. Στην εικόνα με τις σκεπές υπάρχουν σημεία στην καμινάδα και αρκετά σημεία στα παράθυρα και στα όρια των σκεπών.

Τέλος, στην περίπτωση της εικόνας με την Lena για μέγεθος φίλτρου 9 και $\sigma = \sqrt{2}$ έχουμε ένα πολύ καλό αποτέλεσμα. Τα περισσότερα σημεία εντοπίζονται στο φόντο της εικόνας κυρίως στον στύλο πίσω αριστερά.

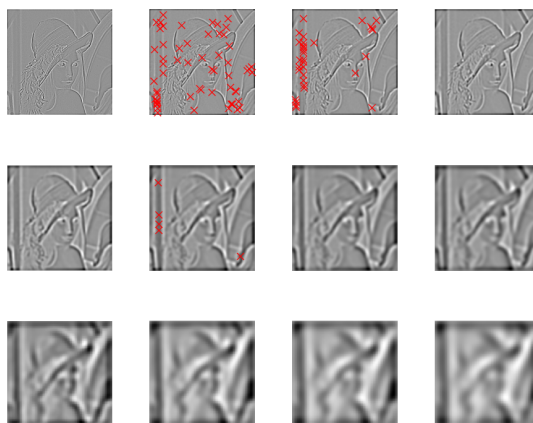


Figure 22: Τα DoGs της εικόνας Lena μαζί με τα αντίστοιχα keypoints



Figure 23: Η εικόνα Lena μαζί με τα αντίστοιχα φιλτραρισμένα keypoints

Παρατηρώντας τα αποτελέσματα μου βλέπω ότι τα περισσότερα ιδιάζοντα σημεία προκύπτουν από τα DoGs του πρώτου octave. Εντοπίζονται σημεία και από το 2ο octave αλλά είναι πολύ λίγα. Από το 3ο όμως δεν εντοπίστηκε σε κανένα DoG.