

ΠΛΗ 418: Μηχανική Όραση
Assignment 2: Blob detection

Ομάδα 34

Ζήσης Αρέστης	2009030053
Νίκου Γεώργιος-Νεκτάριος	2016030125

Στην δεύτερη άσκηση κληθήκαμε να προσομοιώσουμε τη λειτουργία του αλγορίθμου Scale-Invariant Feature Transform (SIFT).

Αρχικά ανοίγουμε τις εικόνες και τις αποθηκεύουμε σε grayscale μορφή με double τιμές. Έπειτα ορίζουμε τις παραμέτρους του scale space. Ύστερα από πειραματισμό αλλά και από τις υποδείξεις της άσκησης καταλήξαμε στις τιμές:

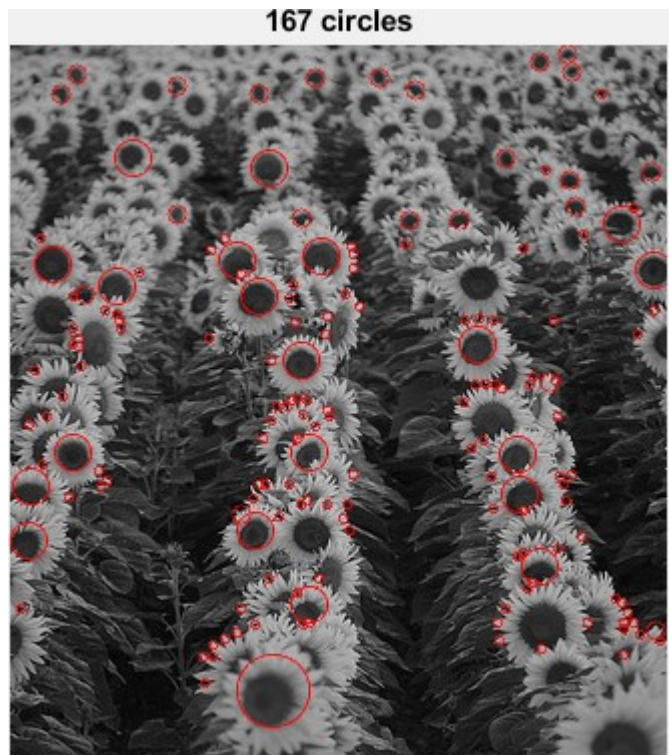
factor(k) = 2, number of levels = 15, threshold = 0.02 και sigma = 1.6

Για τις τρεις πρώτες τιμές κάναμε διάφορους πειραματισμούς για να καταλήξουμε στο αποτέλεσμα. Για τον αριθμό των επιπέδων της πυραμίδας παρατηρήσαμε ότι για >17 επίπεδα ο αλγόριθμος αργεί υπερβολικά να ολοκληρωθεί. Επίσης για threshold >0.05 παρατηρήσαμε λίγα extrema ενώ για threshold < 0.01 είδαμε μεγάλο αριθμό κύκλων να εμφανίζεται. Οπότε καταλήξαμε σε μια ενδιάμεση τιμή (0.02) στην οποία εμφανιζόταν ικανοποιητικός αριθμός κύκλων και δεν υπήρχε μεγάλη επικάλυψη. Όσο αναφορά το σ (standard deviation) παρατηρήσαμε ότι για $\sigma < 0.6$ παίρνουμε ελάχιστους έως μηδενικούς κύκλους γιαυτό και διαλέξαμε να του δώσουμε τιμή 1.6 όπως και προτείνεται από τις διαλέξεις.

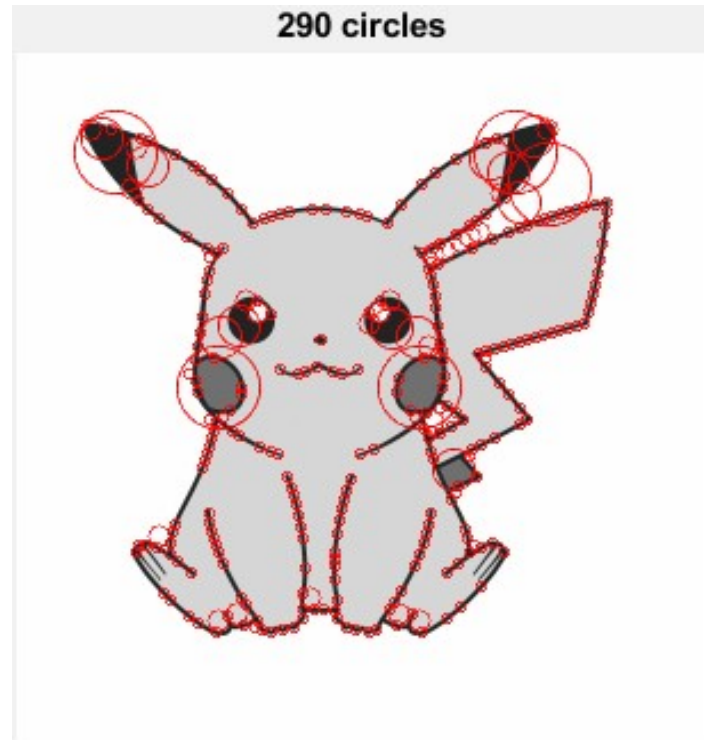
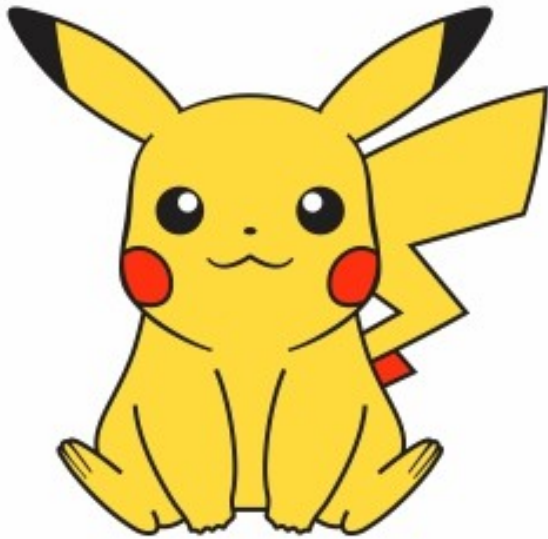
Για την δημιουργία του scale space ορίζουμε έναν τρισδιάστατο πίνακα καθώς και το Laplacian of Gaussian φίλτρο. Το φίλτρο με περιττό μέγεθος 5x5 παράχθηκε με την fspecial και το κανονικοποιούμε με το σ^2 . Ύστερα με ένα βρόχο κατασκευάζουμε το scale pyramid, κάνοντας downsampling την εικόνα μας κατά $1/k$, την φιλτράρουμε και την κανονικοποιούμε και έπειτα την επαναφέρουμε στο αρχικό μέγεθος. Τέλος την αποθηκεύουμε στον τρισδιάστατο πίνακα, ο οποίος κρατάει τις τιμές όλων των επιπέδων.

Αφού δημιουργήθηκε το scale space μέσω της εντολής ordfilt2 βρίσκουμε τα τοπικά extrema σε κάθε επίπεδο του. Έπειτα βρίσκουμε τα maxima σε γειτονικές περιοχές σε όλο το scale space μέσω της συνάρτησης findExtrema. Συγκρίνουμε τις τιμές του scale space, όπως έχει διαμορφωθεί, με το threshold και παράγουμε έναν πίνακα με binary τιμές, όπου για 1 αποτυπώνονται τα σημεία ενδιαφέροντος. Τέλος με βάση τον τελευταίο πίνακα υπολογίζουμε τους κύκλους και τους εμφανίζουμε μέσω της show_all_circles.

Αποτελέσματα πειράματος



Όσον αφορά τις εικόνες που μας δόθηκαν, για την πρώτη εικόνα παρατηρούμε ότι το threshold δεν είναι τόσο αποδοτικό καθώς εμφανίζονται πολύ λίγοι κύκλοι. Μια πιθανή μείωση του threshold ή αύξηση των επιπέδων θα παρήγαγε καλύτερα αποτελέσματα για την πρώτη εικόνα, αλλά η επιλογή τους έγινε βάσει όλων των εικόνων. Μια διαφορετική επιλογή θα οδήγούσε σε μεγάλη αύξηση των key points στις υπόλοιπες εικόνες, κάτι το οποίο δεν είναι θεμιτό. Για την δεύτερη εικόνα βλέπουμε αρκετούς κύκλους στο εσωτερικό των ηλιοτροπίων και μικρότερους στα φύλλα τους. Το αποτέλεσμα θεωρούμε πως είναι ικανοποιητικό διότι έχουν βρεθεί αρκετά key points στα περισσότερα λουλούδια.



Οι εικόνες που διαλέξαμε εμείς είναι ένα γήπεδο μπάσκετ και μια εικόνα του pikachu. Στην πρώτη ο αλγόριθμος εντοπίζει μεγάλο αριθμό σημείων ενδιαφέροντος, ιδιαίτερα στα σημεία με μεγάλη φωτεινότητα και έντονα χρώματα. Στην δεύτερη αποτυπώνονται πολλά key points στα edges του σχεδίου, σκιαγραφώντας σχεδόν την φιγούρα, όπως και στα έντονα σημεία εσωτερικά. Τέλος θεωρούμε την απόδοση του αλγορίθμου πάνω στις δύο αυτές εικόνες αρκετά καλή διότι εντοπίζει αρκετά από τα αναμενόμενα σημεία ενδιαφέροντος.