Πανεπιστήμιο Πατρών

Τμήμα Μηχανικών Η//Υ και Πληροφορικής

Εφαρμογές Ψηφιακής Επεξεργασίας Σημάτων

Αναφορά Άσκησης 3

Όνομα: ΦΩΤΕΙΝΗ

Επώνυμο:ΑΓΓΕΛΑΚΗ

Α.Μ.:1067540

Πατρα Δεκέμβριος 2022

1.

* Ανίχνευση Ακροτάτων στον Χώρο Κλίµακας(Scale-space extrema detection): Γραμμες: 20-68
* Εύρεση Τοπικών Σηµείων-Κλειδιών(Keypoint localization): Γραμμες: 69-116
* ∆ιόρθωση Κατεύθυνσης(Orientation assignmen): Γραμμές 190-260.
* Περιγραφέας Σηµείων-Κλειδιών(Keypoint descriptor generation): γραμμές 261- 344

2.

* Ανίχνευση Ακροτάτων στον Χώρο Κλίµακας(Scale-space extrema detection): Yλοποιεί την ανίχνευση ακραίων διαστημάτων κλίμακας χρησιμοποιώντας τη διαφορά της πυραμίδας Gaussian (DoG). Λαμβάνει μια εικόνα εισόδου και δημιουργεί μια πυραμίδα θολών εικόνων με φίλτρα Gauss ποικίλης κλίμακας. Στη συνέχεια, για κάθε επίπεδο στην πυραμίδα, η διαφορά του Gaussian υπολογίζεται αφαιρώντας το τρέχον επίπεδο από το προηγούμενο επίπεδο και τα άκρα (δηλαδή, μέγιστα και ελάχιστα) ανιχνεύονται στον τρισδιάστατο χώρο συγκρίνοντας κάθε pixel στους 26 γείτονές του σε παρακείμενα επίπεδα.
* Εύρεση Τοπικών Σηµείων-Κλειδιών(Keypoint localization): Αρχικά υπολογίζει τον αριθμό των πιθανών ακρών επαναλαμβάνοντας σε κάθε επίπεδο στην πυραμίδα και υπολογίζοντας τον αριθμό των pixel στο επίπεδο. Στη συνέχεια, αναζητά κάθε εικονοστοιχείο στον χάρτη DoG για να βρει άκρα συγκρίνοντας κάθε pixel με τους 26 γείτονές του σε γειτονικά επίπεδα στο χώρο κλίμακας.

Για κάθε ανιχνευόμενο άκρο, ο κώδικας αποθηκεύει την οκτάβα, την κλίμακα, τη θέση του και αν είναι μέγιστο ή ελάχιστο. Στη συνέχεια, μετατρέπει τη θέση του εικονοστοιχείου στην αντίστοιχη θέση στην αρχική εικόνα, λαμβάνοντας υπόψη τον παράγοντα κλιμάκωσης που εφαρμόζεται σε κάθε οκτάβα. Τέλος, ο κώδικας σχεδιάζει τα εντοπισμένα σημεία-κλειδιά (δηλαδή τα ακραία) στην αρχική εικόνα.

* ∆ιόρθωση Κατεύθυνσης(Orientation assignmen): Γραμμές 190-260.

Για κάθε εντοπισμένο σημείο κλειδιού, επιλέγεται μια περιοχή γύρω από το σημείο κλειδιού με βάση την κλίμακα του σημείου κλειδιού.

Το μέγεθος της κλίσης και ο προσανατολισμός υπολογίζονται για κάθε εικονοστοιχείο εντός της περιοχής.

Ένα σταθμισμένο ιστόγραμμα προσανατολισμών δημιουργείται δεσμεύοντας τις τιμές προσανατολισμού και σταθμίζοντας κάθε τιμή με το αντίστοιχο μέγεθος κλίσης και ένα βάρος Gauss.

Η υψηλότερη κορυφή στο ιστόγραμμα επιλέγεται ως ο κύριος προσανατολισμός για το σημείο-κλειδί και οποιεσδήποτε άλλες κορυφές μεγαλύτερες από το 80% της κύριας κορυφής καταγράφονται επίσης ως πρόσθετοι προσανατολισμοί.

Το σημείο κλειδιού αντιγράφεται για κάθε πρόσθετο προσανατολισμό και οι τιμές προσανατολισμού και άλλες παράμετροι του σημείου κλειδιού ενημερώνονται ανάλογα.

Το προκύπτον σύνολο σημείων κλειδιών με πολλαπλούς προσανατολισμούς επιστρέφεται.

* Περιγραφέας Σηµείων-Κλειδιών(Keypoint descriptor generation):

Διαιρεί την περιοχή γύρω από κάθε σημείο κλειδιού σε ένα πλέγμα 4x4 και υπολογίζοντας ένα ιστόγραμμα προσανατολισμών κλίσης σε κάθε κελί πλέγματος. Κάθε ιστόγραμμα έχει 8 θέσεις που αντιστοιχούν σε γωνίες μεταξύ 0 και 360 μοιρών, με κάθε κάδο να καλύπτει μια περιοχή 45 μοιρών. Το μέγεθος της κλίσης χρησιμοποιείται ως βάρος για κάθε κάδο προσανατολισμού.

Ο τελικός περιγραφέας είναι μια συνένωση των 128 κάδων ιστογράμματος (δηλαδή 4x4x8), η οποία στη συνέχεια κανονικοποιείται σε μονάδα μήκους. Τέλος, οποιοδήποτε στοιχείο στον περιγραφέα του οποίου η απόλυτη τιμή είναι μεγαλύτερη από 0,2 περιορίζεται στο 0,2 και ο περιγραφέας κανονικοποιείται ξανά στο μοναδιαίο μήκος.Οι παραγόμενοι περιγραφείς αποθηκεύονται σε έναν πίνακα που ονομάζεται "χαρακτηριστικό", όπου κάθε στήλη αντιστοιχεί σε έναν περιγραφέα για ένα σημείο κλειδιού. Η μεταβλητή "extr\_volume" καθορίζει τον αριθμό των σημείων κλειδιών που εξήχθησαν από την εικόνα.

3. Ο αριθμός των επιπέδων που χρησιμοποιούνται ανά οκτάβα σε αυτόν τον αλγόριθμο είναι 3. Αυτό μπορεί να φανεί στη γραμμή level=3 όπου το μεταβλητό επίπεδο έχει πάρει την τιμή 3. Αυτό σημαίνει ότι κάθε οκτάβα θα αποτελείται από 3 εικόνες και η διαφορά των Gaussian θα είναι υπολογίζεται μεταξύ γειτονικών επιπέδων σε κάθε οκτάβα για να δημιουργηθεί η πυραμίδα κλίμακας-διαστήματος.

4. Για να εντοπίσουµε την ϐέλτιστη, µε ακρίβεια µεγαλύτερη από αυτήν που µας προσφέρει η οµοιόµορϕη διακριτοποίηση του άξονα της κλάµακας, ενός σηµείου κλειδιού θα μπορούσαμε να χρησιμοποιήσουμε έναν αλγόριθμο βελτιστοποίησης. Οι αλγόριθμοι αυτοί μπορούν να βρουν το ελάχιστο ή το μέγιστο μιας συνάρτησης με μεγαλύτερη ακρίβεια από την ομοιόμορφη διακριτοποίηση. Θα μπορούσε να γίνει χρήση των αλγορίθμων:

* Gradient descent: Αυτός ο αλγόριθμος χρησιμοποιεί τη διαβάθμιση της συνάρτησης για να ενημερώνει επαναληπτικά τη θέση του σημείου κλειδιού έως ότου η διαβάθμιση γίνει μηδέν, υποδεικνύοντας ότι έχει επιτευχθεί ένα τοπικό ελάχιστο ή μέγιστο.

Πλεονεκτήματα: Απλό στην εφαρμογή και υπολογιστικά αποδοτικό για προβλήματα χαμηλών διαστάσεων. Μπορεί να χειριστεί θορυβώδεις ή στοχαστικές αντικειμενικές συναρτήσεις.

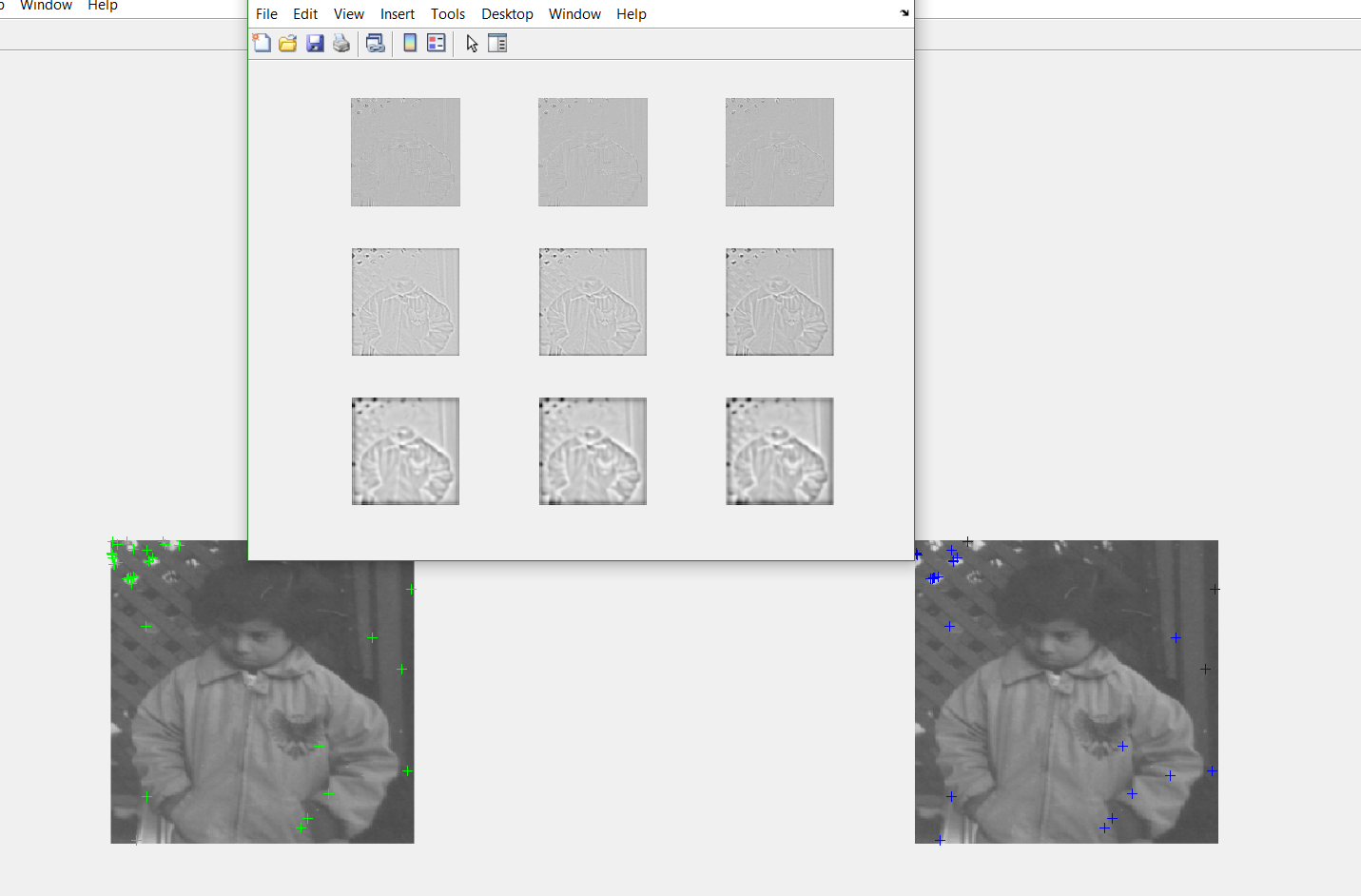
Μειονεκτήματα: Μπορεί να συγκλίνει σε τοπικό βέλτιστο αντί για καθολικό βέλτιστο. Μπορεί να απαιτεί προσεκτική ρύθμιση του ρυθμού εκμάθησης.

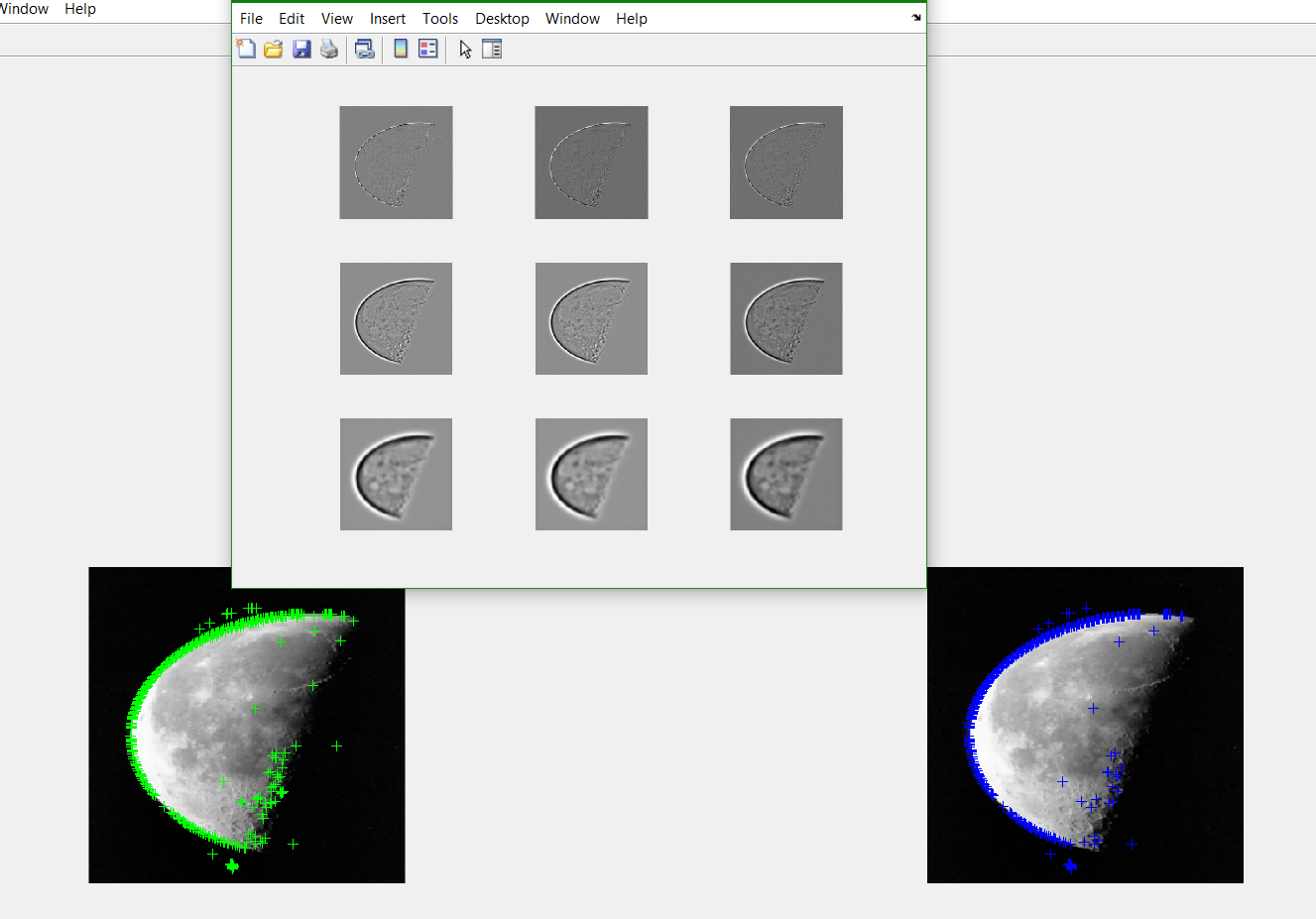
* Newton's method: χρησιμοποιεί τη διαβάθμιση και τον πίνακα Hessian της συνάρτησης για να ενημερώσει επαναληπτικά τη θέση του σημείου κλειδιού

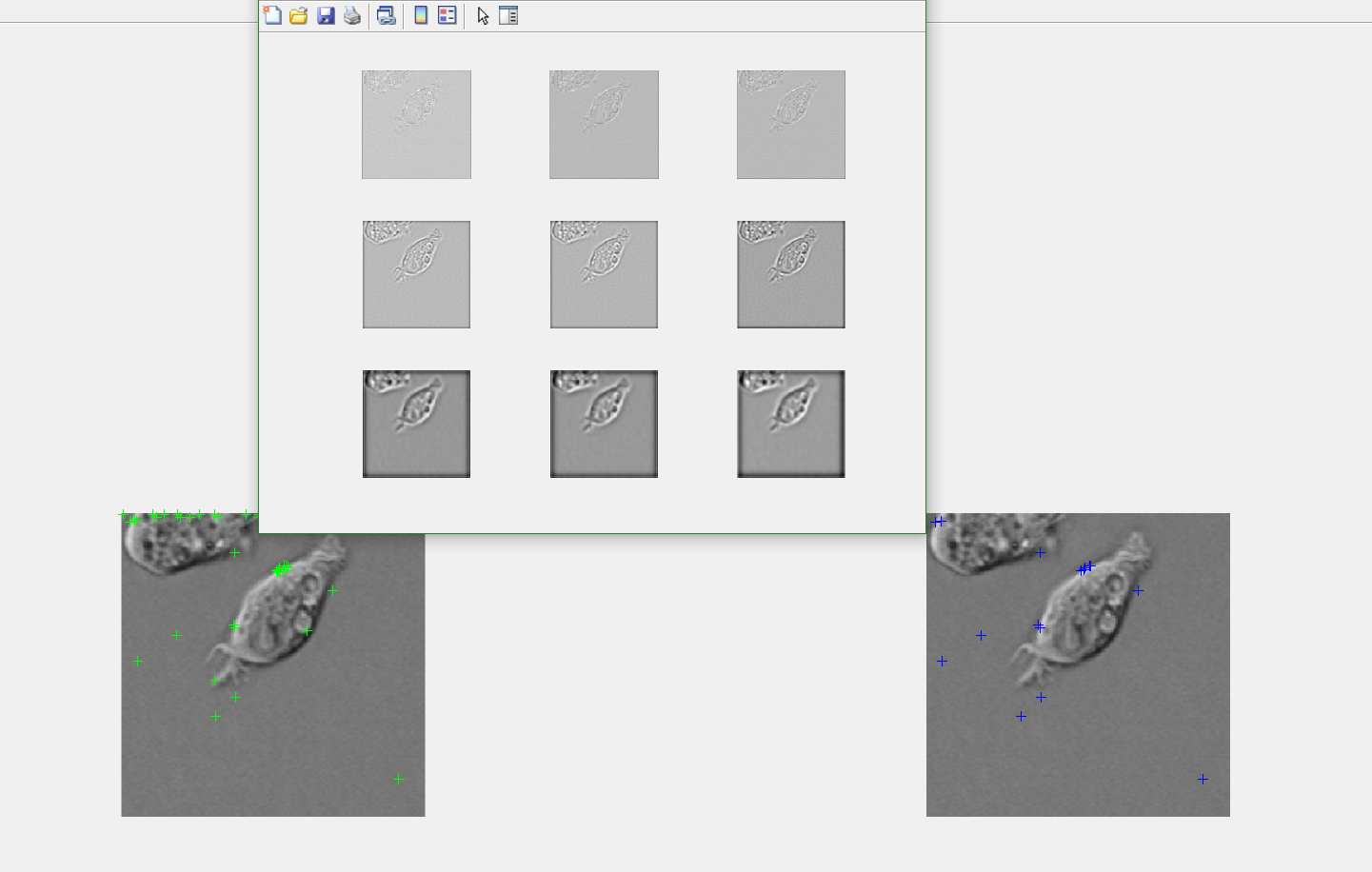
Πλεονεκτήματα: Μπορεί να συγκλίνει γρήγορα στο παγκόσμιο βέλτιστο, ειδικά για κυρτά προβλήματα με καλή συμπεριφορά. Μπορεί να χειριστεί προβλήματα υψηλών διαστάσεων.

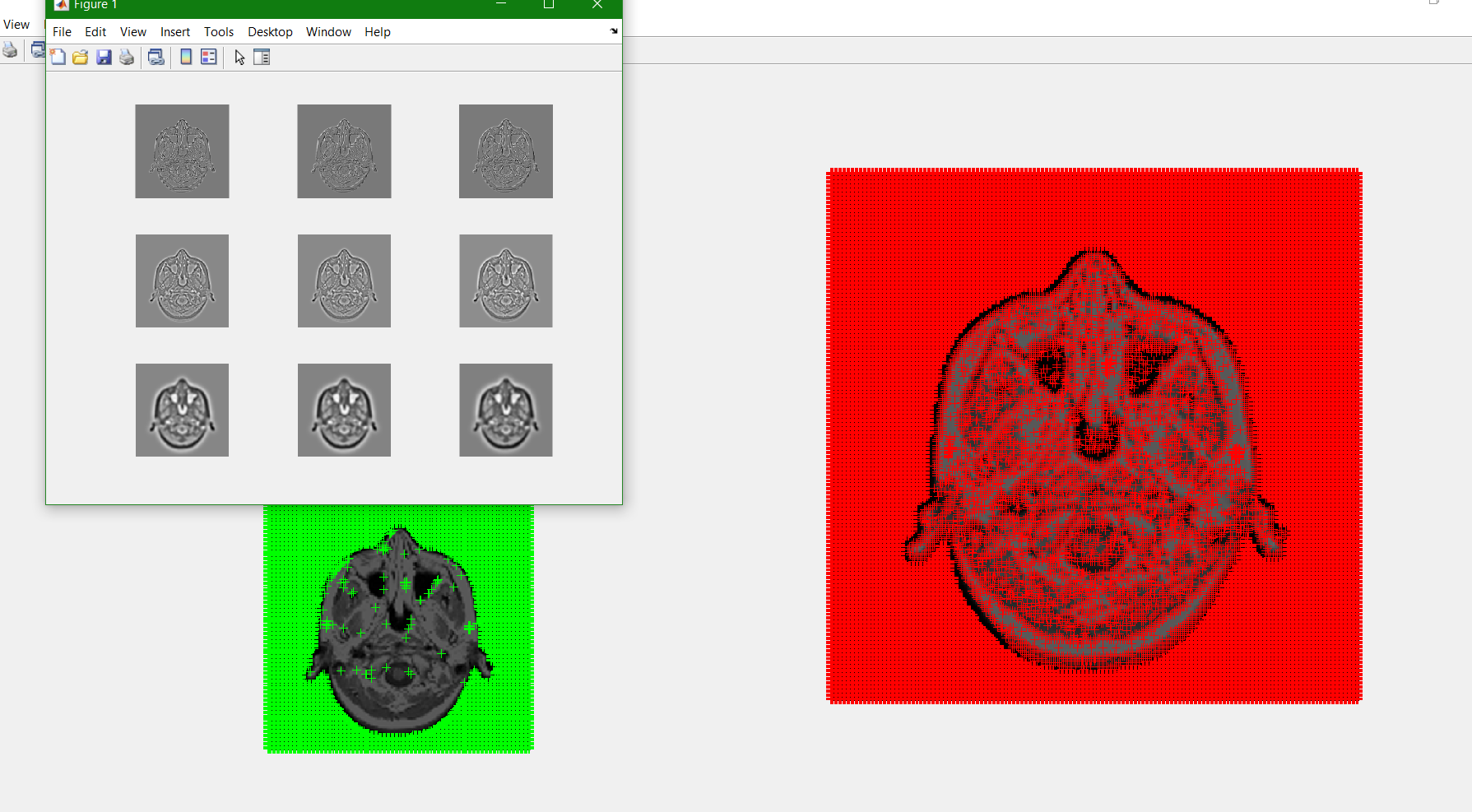
Μειονεκτήματα: Υπολογιστικά ακριβός υπολογισμός του πίνακα Hessian. Μπορεί να αποτύχει εάν το Hessian δεν είναι θετικό οριστικό ή εάν το αντίστροφο δεν είναι υπολογίσιμο.

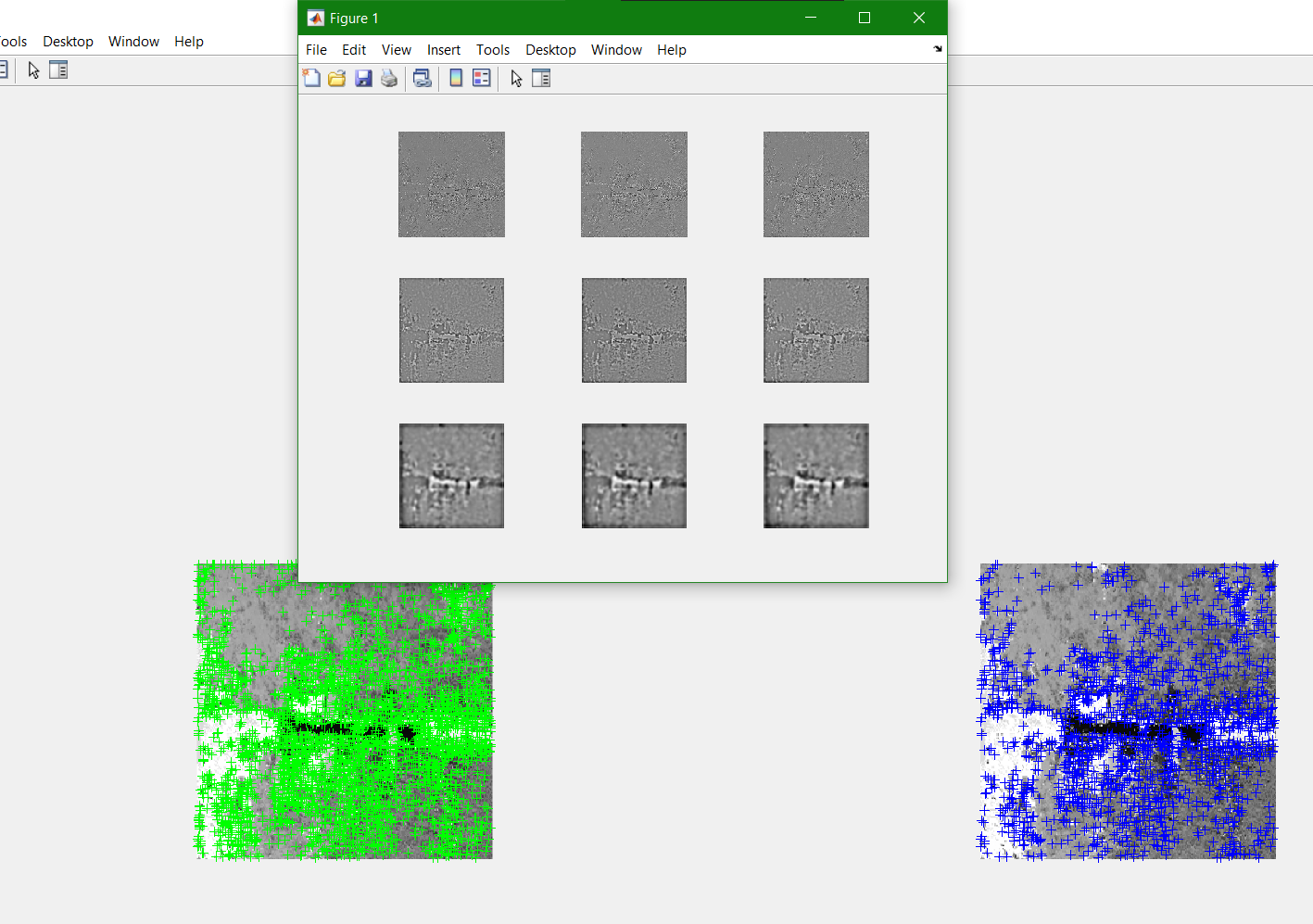
5.Οι 5 εικόνες των οποίων έγινε χρήση είναι οι: "Pout.tif", "Moon.tif", "Cell.tif", ""Mri.tif", "Canoe.tif" οι οποίες βρέθηκαν στο Image Processing Toolbox της MATLAB. Σε όλες τις εικόνες το αποτέλεσμα του κώδικα ήταν να αναγνωρίζει σημεία-κλειδιά που είναι πιθανό να είναι διακριτικά και ανθεκτικά σε αλλαγές στην οπτική γωνία ή στο φωτισμό.











6.

7. Η RANSAC, είναι μια επαναληπτική μέθοδος για την εκτίμηση ενός μαθηματικού μοντέλου από ένα σύνολο δεδομένων που περιέχει ακραίες τιμές. Ο αλγόριθμος RANSAC λειτουργεί με τον εντοπισμό των ακραίων τιμών σε ένα σύνολο δεδομένων και την εκτίμηση του επιθυμητού μοντέλου χρησιμοποιώντας δεδομένα που δεν περιέχουν ακραίες τιμές.

Λειτουργεί επιλέγοντας τυχαία ένα υποσύνολο των σημείων δεδομένων (που ονομάζεται ελάχιστο δείγμα) και προσαρμόζοντας ένα μοντέλο σε αυτά τα σημεία. Στη συνέχεια, το μοντέλο αξιολογείται σε σχέση με τα υπόλοιπα σημεία δεδομένων και τυχόν σημεία δεδομένων που είναι συμβατά με το μοντέλο (που ονομάζονται inliers) ταξινομούνται ως τέτοια. Αυτή η διαδικασία επαναλαμβάνεται πολλές φορές, με διαφορετικά υποσύνολα δεδομένων, για να βρεθεί το μοντέλο που έχει τον μεγαλύτερο αριθμό εσωτερικών στοιχείων.

Στην όραση υπολογιστή, το RANSAC χρησιμοποιείται ως μια ισχυρή προσέγγιση για την εκτίμηση του θεμελιώδους πίνακα στη στερεοφωνική όραση, για την εύρεση της κοινότητας μεταξύ δύο συνόλων σημείων για την ανίχνευση αντικειμένων βάσει χαρακτηριστικών και την καταχώρηση διαδοχικών καρέ βίντεο για σταθεροποίηση βίντεο.

8. Το SIFT, το GLOH και το SURF είναι όλοι αλγόριθμοι που χρησιμοποιούνται για την εξαγωγή και αντιστοίχιση χαρακτηριστικών σε εφαρμογές υπολογιστικής όρασης. Μερικές διαφορές τους είναι οι εξής:

* Το SIFT και το SURF χρησιμοποιούν διαφορετικές μεθόδους για τον εντοπισμό σημείων κλειδιών. Το SIFT χρησιμοποιεί μια προσέγγιση Διαφοράς Gaussian (DoG) για να προσδιορίσει τα σημεία-κλειδιά που είναι τοπικά άκρα στον χώρο κλίμακας, ενώ το SURF χρησιμοποιεί την ορίζουσα του πίνακα Hessian για να προσδιορίσει τα σημεία-κλειδιά. Το GLOH, από την άλλη πλευρά, χρησιμοποιεί την ίδια μέθοδο ανίχνευσης σημείων κλειδιού με το SIFT.
* Το SURF είναι ταχύτερο από το SIFT και μπορεί να επιτύχει συγκρίσιμα αποτελέσματα σε πολλές εφαρμογές, αλλά μπορεί να είναι λιγότερο ανθεκτικό στις αλλαγές στις συνθήκες φωτισμού.
* Το SIFT είναι υπολογιστικά ακριβό, ειδικά για μεγάλες εικόνες ή σύνολα δεδομένων, λόγω των πολύπλοκων μεθόδων ανίχνευσης σημείων κλειδιού και υπολογισμού του περιγραφέα..
* Το GLOH είναι επίσης ταχύτερο από το SIFT σε ορισμένες περιπτώσεις λόγω της απλούστερης μεθόδου υπολογισμού του περιγραφέα.

9. Η χρονική πολυπλοκότητα των SIFT, GLOH και SURF εξαρτάται από διάφορους παράγοντες, όπως το μέγεθος και την πολυπλοκότητα των εικόνων εισόδου, τον αριθμό των σημείων-κλειδιών που εντοπίστηκαν και τις λεπτομέρειες υλοποίησης του αλγορίθμου.

SIFT χρονική πολυπλοκότητα: O(N^2 log N), όπου N είναι ο αριθμός των σημείων κλειδιών που εντοπίστηκαν στην εικόνα.

GLOH χρονική πολυπλοκότητα: O(N log N), όπου N είναι ο αριθμός των σημείων κλειδιών που ανιχνεύονται στην εικόνα.

SURF χρονική πολυπλοκότητα: O(N log N), όπου N είναι ο αριθμός των σημείων κλειδιών που ανιχνεύονται στην εικόνα.

10. ΒΙΒΛΙΟΓΡΑΦΊΑ: <https://www.mathworks.com/discovery/ransac.html#:~:text=Random%20sample%20consensus%2C%20or%20RANSAC,that%20does%20not%20contain%20outliers>.

<https://www.analyticsvidhya.com/blog/2019/10/detailed-guide-powerful-sift-technique-image-matching-python/>

<https://docs.opencv.org/4.x/da/df5/tutorial_py_sift_intro.html>

chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/http://www.micc.unifi.it/delbimbo/wp-content/uploads/2011/10/slide\_corso/A32\_keypoint\_detectors.pdf

<https://docs.opencv.org/3.4/df/dd2/tutorial_py_surf_intro.html>

<https://chat.openai.com/chat>