



ΠΟΛΥΤΕΧΝΕΙΟ ΚΡΗΤΗΣ
ΣΧΟΛΗ ΜΗΧΑΝΙΚΩΝ ΠΑΡΑΓΩΓΗΣ & ΔΙΟΙΚΗΣΗΣ

Τεχνικές εξαγωγής σημείων ενδιαφέροντος σε τούνελ από σκυρόδεμα

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Δάβαρης Κωνσταντίνος

Επίβλεπων 1: Δουλάμης Αναστάσιος
Επιβλέπων 2: Σταυρουλάκης Γεώργιος

Χανιά, Αύγουστος 2015

Με την ολοκλήρωση αυτής της διπλωματικής εργασίας, θα ήθελα να ευχαριστήσω θερμά τους επιβλέποντες καθηγητές μου, Αναστάσιο Δουλάμη και Γεώργιο Σταυρουλάκη, οι οποίοι μου έδωσαν σημαντικά στοιχεία και πολύτιμη καθοδήγηση κατά τη διάρκεια της εκπόνησης αυτής της εργασίας.

Χανιά, Αύγουστος 2015

Δάβαρης Κωνσταντίνος

Περίληψη

Τον τελευταίο καιρό το ενδιαφέρον της ανίχνευσης ρωγμών και θρυμματισμών στα τούνελ έχει αυξηθεί. Αρκετά τούνελ βρίσκονται σε λειτουργία πολλά χρόνια, ενώ το κόστος κατασκευής ενός νέου τούνελ είναι πολύ υψηλό. Η ύπαρξη πολλών ατυχημάτων, κατά καιρούς, έχει καταστήσει την παρακολούθηση και ανίχνευση της δομικής τους κατάστασης επιτακτική.

Στην παρούσα διπλωματική εργασία, δημιουργήθηκε ένας αλγόριθμος, στον οποίο χρησιμοποιήθηκαν και αξιολογήθηκαν πέντε διαφορετικοί ανιχνευτές που πάρθηκαν από την εργαλειοθήκη της Matlab, με απώτερο σκοπό την ανίχνευση ρωγμών και θρυμματισμών στην εσωτερική επένδυση ενός τούνελ από σκυρόδεμα. Τα δεδομένα που χρησιμοποιήθηκαν ήταν 37 πραγματικές φωτογραφίες από το τούνελ V-S-H της Ελβετίας. Κάθε μια από τις 37 φωτογραφίες, έχει την αντίστοιχη της σχολιαστική εικόνα (annotation image), η οποία εμπεριέχει τις ακριβείς συντεταγμένες των ελαττωματικών σημείων του σκυροδέματος. Οι πραγματικές φωτογραφίες, σε συνδυασμό με τις αντίστοιχες τους σχολιαστικές εικόνες, χρησιμοποιούνται προκειμένου να ανευρεθούν και να διαχωριστούν, τα ελαττωματικά σημεία θρυμματισμών και ρωγμών, από τα σημεία που δεν είναι ελαττωματικά σε δομικό επίπεδο. Τα σημεία αυτά χρησιμοποιούνται για την αξιολόγηση των ανιχνευτών και, κατ' επέκταση, την εξαγωγή συμπερασμάτων. Συμπερασματικά, ο ανιχνευτής SURF δίνει τα πιο εύρωστα αποτελέσματα ως προς την συνολική αξιολόγηση των ελατωμάτων των τούνελ σκυροδέματος. Εν αντιθέσει, ο Harris παρουσιάζει καλύτερη απόδοση στον εντοπισμό των θρυμματισμών ενώ ο BRISK για τις ρωγμές.

Abstract

Recently, the detection of structural defects (cracks and spallings) has been of great interest. Several tunnels have been in operation for many years, while the cost of a new tunnel construction is very high. Occasionally, there have been many accidents, which require the immediate inspection and detection of their structural state.

In this diploma thesis, we developed an algorithm, which uses and assesses five different detectors from the Computer Vision Matlab toolbox. Our ultimate goal is to detect cracks and spallings in the inner concrete lining of a tunnel. The data used are 37 real-time images of the tunnel V-S-H in Switzerland. Each one of the 37 images, has its respective annotation image, which includes the exact coordinates of the defective points of the concrete lining. We use these real-time images combined with their respective annotation images, so as to find and separate the defective points (cracks and spallings), from the structurally non-defective points. These points are used to assess the detectors selected, thus to reach a useful conclusion. In conclusion, SURF descriptor yields the most robust results as regards the total defects of concrete tunnel. On the contrary, Harris presents the best performance in detecting spalling defects while BRISK as for the cracks.

Περιεχόμενα

Περίληψη.....	2
Abstract	3
Περιεχόμενα	4
A Πίνακας εικόνων	6
B Πίνακας Γραφημάτων	9
C Πίνακας πινάκων.....	10
1 Εισαγωγή	11
1.1 Σημερινές ανάγκες	11
1.2 Υπολογιστική Όραση (Computer Vision)	15
1.3 Ανίχνευση χαρακτηριστικών (Feature detection).....	19
1.3.1 Ακμές (Edges).....	20
1.3.2 Γωνίες-Σημεία ενδιαφέροντος (Corners).....	20
1.3.3 Άμορφα σημεία (Blobs) - περιοχές ενδιαφέροντος - ενδιαφέροντα σημεία	21
1.3.4 Κορυφές (Ridges)	21
2 Ανάλυση ανιχνευτών	23
2.1 Ανιχνευτής γωνιών Harris	23
2.2 Αλγόριθμος ελάχιστων ιδιοτιμών Shi-Tomasi	28
2.3 Ανιχνευτής Χαρακτηριστικών με Επιταχυνόμενο έλεγχο Τμημάτων (FAST-Features from Accelerated Segment Test).....	31
2.3.1 Ανίχνευση χαρακτηριστικών χρησιμοποιώντας τον FAST	31
2.3.2 Προσέγγιση μηχανικής εκμάθησης.....	33
2.3.3 Τεχνική μη μέγιστης κατάπνιξης (Non maximum suppression) για την αφαίρεση γειτονικών ακμών.....	35
2.4 Ανιχνευτής και Περιγραφέας Ταχύτερων Υπολογιζομένων Εύρωστων Χαρακτηριστικών (SURF-Speeded Up Robust Features).....	37
2.4.1 Αναγνώριση σημείων ενδιαφέροντος βασισμένα στον Εσσιανό πίνακα (Fast-Hessian Detector).....	39
2.4.2 Τοποθέτηση προσανατολισμού	41
2.4.3 Τα στοιχεία του περιγραφέα	42
2.5 Περιγραφέας Δυναδικών Εύρωστων και Αμετάβλητων Σημείων Ενδιαφέροντος (BRISK - Binary Robust Invariant Scalable Keypoints).....	44

2.5.1	Ανίχνευση σημείων ενδιαφέροντος σε κλίμακα χώρου (Scale-space keypoint detection)	45
2.5.2	Περιγραφή σημείων ενδιαφέροντος (Keypoint description).....	48
3	Υλοποίηση Αλγόριθμου	51
3.1	Δεδομένα.....	51
3.2	Ανάλυση Αλγορίθμου	55
3.3	Αποτελέσματα	63
3.3.1	Ανιχνευτής BRISK	63
3.3.2	Ανιχνευτής Harris	67
3.3.3	Ανιχνευτής ελαχίστων ιδιοτιμών Shi-Tomasi (MinEigen)	70
3.3.4	Ανιχνευτής FAST	74
3.3.5	Ανιχνευτής SURF	77
3.4	Συμπεράσματα	84
	Βιβλιογραφία.....	85
	Παράρτημα.....	87

Α Πίνακας εικόνων

Εικόνα 1.1 : (a) Υπολογισμός τρισδιάστατων μοντέλων περιβάλλοντος από χιλιάδες επικαλυπτόμενες εικόνες. (b) Δημιουργία πυκνών τρισδιάστατων μοντέλων με την χρήση στερεοφωνικής αντιστοίχισης από φωτογραφίες διαφορετικών οπτικών γωνιών. (c) Εντοπισμός κίνησης ενός ανθρώπου που κινείται σε ένα σύνθετο φόντο. (d) Αλγόριθμοι ανίχνευσης προσώπου που προκύπτει από ένα συνδυασμό χαρακτηριστικών από ανθρώπους σε φωτογραφίες.	17
Εικόνα 1.2 : (a) Οπτική αναγνώριση χαρακτήρων(OCR). (b) Επιθεώρηση μηχανών. (c) Λιανεμπόριο. (d) Ιατρική απεικόνιση. (e) Αυτοκίνητα ασφαλείας. (f) Παρακολούθηση(ανάλυση κίνησης αυτοκινήτων).	18
Εικόνα 2.1 : Στην αριστερή εικόνα έχουμε επίπεδη περιοχή χωρίς αλλαγή έντασης. Στην μεσαία εικόνα έχουμε ακμή με καμία μεταβολή κατά μήκος της διεύθυνση της ακμής. Στην δεξιά εικόνα έχουμε γωνία και μεγάλη αλλαγή έντασης σε τουλάχιστον δύο κατευθύνσεις.	23
Εικόνα 2.2 : Δείχνει την αλληλεξάρτηση του μεγέθους R με το είδος των σημείων.	25
Εικόνα 2.3: Υπολογισμός της απόκρισης της γωνίας R	26
Εικόνα 2.4 : Βρίσκουμε τα σημεία με μεγάλο R ($R > \text{ορίου κατωφλίσωσης}$)	26
Εικόνα 2.5 : Κρατάμε τα σημεία με τα τοπικά μέγιστα του R . Αυτά είναι και τα σημεία ενδιαφέροντος	27
Εικόνα 2.6 : Μορφή της συνάρτησης αυτοσυσχετισμού $c(x,y)$ για σημείο που αντιστοιχεί σε γωνία	28
Εικόνα 2.7: Μορφή της συνάρτησης αυτοσυσχετισμού $c(x,y)$ για σημείο που αντιστοιχεί σε ακμή	29
Εικόνα 2.8 : Μορφή της συνάρτησης αυτοσυσχετισμού $c(x,y)$ για σημείο που αντιστοιχεί σε επίπεδη περιοχή	29
Εικόνα 2.9: Αλληλεξάρτηση των ιδιοτιμών λ_1, λ_2 με το είδος των σημείων.	29
Εικόνα 2.10 : Φωτογραφία εισόδου για τον αλγόριθμο ελαχίστων ιδιοτιμών Shi-Tomasi.	30
Εικόνα 2.11 : Τα χαρακτηριστικά που έχουν ανιχνευτεί από τον αλγόριθμο ελαχίστων ιδιοτιμών Shi-Tomasi. Οι τοποθεσίες των χαρακτηριστικών εμφανίζονται με σταυρούς.	30
Εικόνα 2.12: Το ελεγχόμενο σημείο ενδιαφέροντος και τα 16 εικονοστοιχεία εντός του κύκλου.	32
Εικόνα 2.13: Οι 16 τιμές που περικλείουν το εικονοστοιχείο « p » αποθηκευμένες σε μορφή διανύσματος	33
Εικόνα 2.14 : Μια εικόνα με ανιχνευμένα σημεία ενδιαφέροντος. Οι πράσινες τελείες δείχνουν τις μη μέγιστες συγκαλυμμένες ακμές.	36
Εικόνα 2.15 : Από αριστερά προς τα δεξιά: η διακριτοποίηση και η περικοπή της Gaussian 2ης τάξης μερικών παραγώγων στην ψ και $\chi\psi$ κατεύθυνση, και οι προσεγγίσεις χρησιμοποιώντας κουτιά φίλτρα. Οι γκριζες περιοχές είναι ίσες με το μηδέν.	40
Εικόνα 2.16 : Αριστερά: Εντοπισμένα σημεία ενδιαφέροντος για ένα λιβάδι με ηλιοτρόπια. Αυτό το είδος εικόνας δείχνει ξεκάθαρα τη φύση των χαρακτηριστικών των ανιχνευτών που βασίζονται στο πίνακα Hessian. Στο μέσο: Τα είδη κύματος Haar που χρησιμοποιούνται στον περιγρ περιγραφέα SURF. Δεξιά: Λεπτομέρεια εικόνας ενός γκράφιτι, που δείχνει το μέγεθος του παραθύρου του περιγραφέα, σε διαφορετικές κλίμακες.	42
Εικόνα 2.17 : Οι καταχωρήσεις του περιγραφέα όσον αφορά μια υπο-περιοχή αναπαριστούν τη φύση ενός υποκείμενου μοτίβου έντασης. Αριστερά: Σε περίπτωση ομοιογενούς περιοχής, όλες οι τιμές είναι σχετικά χαμηλές. Στο μέσο: Σε παρουσία συχνότητων στην διεύθυνση x , η τιμή του $\sum d_x $ είναι υψηλή, αλλά οι υπόλοιπες παραμένουν χαμηλές. Εάν η ένταση, βαθμιαία, αυξάνεται στην διεύθυνση x , τότε και οι δύο τιμές $\sum d_x$ και $\sum d_x $ είναι υψηλές.	43

Εικόνα 2.18 : Ανίχνευση σημείων ενδιαφέροντος (keypoints) σε κλίμακα χώρου: Ένα σημείο ενδιαφέροντος (δηλαδή προβολή μέγιστου) αναγνωρίζεται σε οκτάβα <i>ci</i> με την ανάλυση 8 σημαντικών γειτονικών βαθμολογιών σε <i>ci</i> , καθώς και στις αντίστοιχες βαθμολογίες-επιδιορθώσεις στα αμέσως-γειτονικά επίπεδα πάνω και κάτω. Και στις τρεις στρώσεις ενδιαφέροντος, η τοπική μέγιστη προεξοχή (local saliency maximum) είναι υπο-εικονοστοιχείο (sub-pixel) επεξεργασμένο πριν η παραβολή 1D τοποθετηθεί κατά μήκος της κλίμακας άξονα για το προσδιορισμό του πραγματικού μεγέθους του σημείου ενδιαφέροντος. Η θέση του σημείου ενδιαφέροντος παρεμβάλλεται ξανά ανάμεσα στα μέγιστα μίας περιοχής (patch) που είναι πιο κοντά προς την καθορισμένη κλίμακα.	47
Εικόνα 2.19 : Το μοτίβο δειγματοληψίας του BRISK για $N = 60$ σημεία: οι μικροί μπλε κύκλοι δηλώνουν τις θέσεις της δειγματοληψίας, οι μεγαλύτεροι και κόκκινοι κύκλοι με τις διακεκομμένες είναι ζωγραφισμένοι με ακτίνα σ που αντιστοιχεί στην τυπική απόκλιση του φίλτρου κανονικής κατανομής (Gaussian) το οποίο και χρησιμοποιείται για την εξομάλυνση των τιμών έντασης στα σημεία δειγματοληψίας. Το μοτίβο εφαρμόζεται με κλίμακα $t = 1$	49
Εικόνα 2.20 : Στα αριστερά της εικόνας εμφανίζονται τα ζεύγη σημείων μικρής απόστασης S . Στα δεξιά της εικόνας εμφανίζονται τα ζεύγη σημείων μεγάλης απόστασης L	50
Εικόνα 3.1 : Παράδειγμα της φωτογραφίας 31 από τις 37 που αντιμετωπίζει ελάττωμα θρυμματισμού μαζί με την annotation εικόνα της.	60
Εικόνα 3.2 : Μετατροπή της εικόνας από έγχρωμη σε γκριζα	61
Εικόνα 3.3 : Χρήση των Medfilt2 και histeq για να φιλτράρουμε και να καθαρίσουμε την γκριζα εικόνα	61
Εικόνα 3.4 : Έχουμε χρησιμοποιήσει στην φωτογραφία 31 τον ανιχνευτή SURF καλώντας την συνάρτηση <code>points = detectSURFFeatures(sgrtunnel)</code> και τον έχουμε βάλει να ψάχνει για 45 σημεία με την εντολή <code>points.selectStrongest(45).Location</code> . Τα σημεία που έχει βρει απεικονίζονται με σταυρούς. Οι πράσινοι σταυροί συμβολίζουν τα σημεία ενδιαφέροντος σε αντίθεση με το μοναδικό κόκκινο σταυρό στην προκειμένη περίπτωση να έχει αστοχήσει.	62
Εικόνα 3.5 : Η καλύτερη φωτογραφία ρωγμών για τον BRISK είναι η 19 όταν βρίσκει 6 σημεία ενδιαφέροντος από τα 10 σημεία ανίχνευσης. Τα κόκκινα είναι τα σημεία στα οποία έχει αστοχήσει.	65
Εικόνα 3.6 : Η χειρότερη φωτογραφία ρωγμών για τον BRISK είναι η 6 και βρίσκει 1 σημείο ενδιαφέροντος για οποιοδήποτε αριθμό σημείων ανίχνευσης	65
Εικόνα 3.7 : Η καλύτερη φωτογραφία θρυμματισμών για τον BRISK είναι η φωτογραφία 23 και βρίσκει το καλύτερο Precision 86,67% όταν ανιχνεύει 15 σημεία δηλαδή τότε βρίσκει 13 σημεία ενδιαφέροντος.	66
Εικόνα 3.8 : Η χειρότερη φωτογραφία θρυμματισμών για τον BRISK είναι η 35 και έχει το χειρότερο Precision 18,57 % όταν βρίσκει 13 σημεία ενδιαφέροντος (φαίνονται στο μπλε πλαίσιο) από τα 70 σημεία ανίχνευσης.	66
Εικόνα 3.9 : Η καλύτερη φωτογραφία ρωγμών για τον Harris είναι η 5 και έχει καλύτερο Precision 100% στα 5 σημεία ανίχνευσης	69
Εικόνα 3.10 : Η χειρότερη φωτογραφία ρωγμών για τον Harris είναι η 17 όπου έχουν αστοχήσει όλα τα σημεία ανίχνευσης. Βλέπουμε ότι η φωτογραφία δεν είναι καλή διότι έχει αποτυπώσει και μέρος με διαφορετική υφή. Όπως φαίνεται ο ανιχνευτής έχει εστιάσει στο μέρος που δεν θέλουμε.	69
Εικόνα 3.11 : Η καλύτερη φωτογραφία θρυμματισμών για τον Harris είναι η 30 και έχει πολλές φορές Precision 100%. Εδώ βλέπουμε όταν έχει 99% Precision στα 100 σημεία ανίχνευσης. Αστοχεί μόνο ο κόκκινος σταυρός.	70

Εικόνα 3.12 : Η χειρότερη φωτογραφία θρυμματισμού για τον Harris είναι η 24 και βρίσκει το μικρότερο Precision 73,33% στα 45 σημεία ανίχνευσης. Εκεί αστοχούν 12 όπως φαίνεται και από τους κόκκινους σταυρούς.....	70
Εικόνα 3.13 : Η καλύτερη φωτογραφία ρωγμών για τον Shi-Tomasi(MinEigen) είναι η 11 και έχει καλύτερο Precision 80% στα 15 σημεία ανίχνευσης. Εκεί αστοχούν 3 με τους κόκκινους σταυρούς.....	72
Εικόνα 3.14 : Η χειρότερη φωτογραφία ρωγμών για τον ανιχνευτή Shi-Tomasi(MinEigen) είναι η 9 όπου δεν βρίσκει κανένα σημείο ανίχνευσης. Η φωτογραφία αποτελείται από 2 μέρη με διαφορετικές υφές. Το πρόβλημα είναι ότι έχει εστιάσει στο μέρος που δεν υπάρχουν τα ελαττώματα ρωγμών.....	73
Εικόνα 3.15 : Η καλύτερη φωτογραφία θρυμματισμών για τον Shi-Tomasi(MinEigen) είναι η 30 και έχει πολλές φορές Precision 100%. Εδώ βλέπουμε όταν έχει 98% Precision στα 100 σημεία ανίχνευσης. Αστοχούν μόνο οι 2 κόκκινοι σταυροί.	73
Εικόνα 3.16 : Η χειρότερη φωτογραφία θρυμματισμών για τον Shi-Tomasi(MinEigen) είναι η 27 και έχει το χειρότερο Precision 60% όταν ψάχνει για 5 σημεία ανίχνευσης. Τότε αστοχούν 2 σημεία με τους κόκκινους σταυρούς.	74
Εικόνα 3.17 : Η καλύτερη φωτογραφία ρωγμών για τον FAST είναι η 11 και έχει το μεγαλύτερο Precision όταν ανιχνεύονται 10 σημεία με Precision 80% . Αστοχούν σε αυτή την περίπτωση 2 σημεία όπως φαίνονται και με τους κόκκινους σταυρούς.	76
Εικόνα 3.18 : Μια από τις χειρότερες φωτογραφίες ρωγμών για τον ανιχνευτή FAST είναι η 2. Η οποία δεν βρίσκει κανένα σημείο ενδιαφέροντος.....	76
Εικόνα 3.19 : Η καλύτερη φωτογραφία θρυμματισμών για τον FAST είναι η 31 και έχει πολλές φορές Precision 100%. Εδώ βλέπουμε όταν έχει 99% Precision στα 100 σημεία ανίχνευσης. Αστοχούν μόνο ένα σημείο(κόκκινος σταυρός) στο πάνω μέρος της φωτογραφίας.	77
Εικόνα 3.20 : Η χειρότερη φωτογραφία θρυμματισμών για τον FAST είναι η 27 και έχει το μικρότερο Precision 68,33% στα 60 σημεία ανίχνευσης διότι βρίσκει 41 σημεία ενδιαφέροντος. Αστοχούν 19 σημεία(κόκκινοι σταυροί).	77
Εικόνα 3.21: Στην πάνω εικόνα φαίνεται η καλύτερη φωτογραφία ρωγμών για τον SURF η οποία είναι η 2 και έχει το μεγαλύτερο Precision 60% όταν ανιχνεύονται 5 σημεία. Αστοχούν σε αυτή την περίπτωση 2 σημεία όπως φαίνονται και με τους κόκκινους σταυρούς. Η κάτω εικόνα δείχνει τα σημεία πιο καθαρά βλέποντας στον μικρόκοσμο της φωτογραφίας.	79
Εικόνα 3.22 : Η χειρότερη φωτογραφία ρωγμών για τον SURF η οποία είναι η 4 και έχει πολύ μικρό μέσο Precision 3,28%. Εδώ βλέπουμε πως αντιδρά στα 100 σημεία ανίχνευσης. Βρίσκει 5 σημεία ενδιαφέροντος με Precision 5%. Τα πράσινα σημεία είναι αυτά που έχουν ευστοχήσει.	80
Εικόνα 3.23 : Η καλύτερη φωτογραφία θρυμματισμού για τον SURF είναι η 31. Έχει μέσο Precision 94,98 % και εμφανίζει 100% Precision σε αρκετές περιπτώσεις. Σε αυτή την εικόνα βλέπουμε όταν έχει το μικρότερο Precision 86,32% στα 95 σημεία ανίχνευσης και βρίσκει 82 σημεία ενδιαφέροντος. Τα 13 κόκκινα σημεία συμβολίζουν τις αστοχίες.....	80
Εικόνα 3.24: Η χειρότερη φωτογραφία θρυμματισμών για τον SURF είναι η 35 με 63,20% μέσο Precision. Εδώ βλέπουμε όταν έχει το μικρότερο Precision 40% και το εμφανίζει στα 100 σημεία ανίχνευσης. Τα 40 πράσινα σημεία είναι τα σημεία που έχει ευστοχήσει.....	81

Β Πίνακας Γραφημάτων

Γράφημα 3-1 : Αξιολόγηση του Brisk ως προς το Precision για κάθε φωτογραφία	64
Γράφημα 3-2 : Η αξιολόγηση του BRISK ως προς το Recall	67
Γράφημα 3-3 : Αξιολόγηση του Harris ως προς το Precision για κάθε φωτογραφία	68
Γράφημα 3-4 : Η αξιολόγηση του Harris ως προς το Recall.....	68
Γράφημα 3-5 : Αξιολόγηση του Shi-Tomasi(MinEigen) ως προς το Precision για κάθε φωτογραφία.....	71
Γράφημα 3-6 : Αξιολόγηση του Shi-Tomasi ως προς το Recall για κάθε φωτογραφία	72
Γράφημα 3-7 : Αξιολόγηση του Fast ως προς το Precision για κάθε φωτογραφία	75
Γράφημα 3-8 : Αξιολόγηση του FAST ως προς το Recall για κάθε φωτογραφία	75
Γράφημα 3-9: Αξιολόγηση του SURF ως προς το Precision για κάθε φωτογραφία.....	78
Γράφημα 3-10: Αξιολόγηση του SURF ως προς το Recall για κάθε φωτογραφία	78
Γράφημα 3-11: Γενική αξιολόγηση των ανιχνευτών στις φωτογραφίες ρωγμών με βάση το Precision	82
Γράφημα 3-12: Γενική αξιολόγηση των ανιχνευτών στις φωτογραφίες θρυμματισμών με βάση το Precision	82
Γράφημα 3-13: Γενική αξιολόγηση των ανιχνευτών για όλες τις φωτογραφίες του δείγματος με βάση το Precision	83
Γράφημα 3-14: Γενική αξιολόγηση των ανιχνευτών για όλες τις φωτογραφίες του δείγματος με βάση την αστάθεια	83

Σ Πίνακας πινάκων

Πίνακας 1-1 : Κατηγοριοποίηση ελαττωμάτων -ρωγμών και θρυμματισμών- σε επενδύσεις από σκυρόδεμα (FHTWA κ.α., 2005).....	12
Πίνακας 3-1 : Δείγμα τεσσάρων φωτογραφιών με ρωγμές μαζί με τις εικόνες που τις απομονώνει (annotation images).	53
Πίνακας 3-2 : Δείγμα τεσσάρων φωτογραφιών με θρυμματισμούς μαζί με τις εικόνες που τους απομονώνουν(annotation images)	54
Πίνακας 3-3 : Οι ανιχνευτές που χρησιμοποιήθηκαν στον αλγόριθμο με τις συναρτήσεις τους από τη βιβλιοθήκη Computer Vision System Toolbox της Matlab.....	55
Πίνακας 3-4 : Μέρος του κώδικα που διαβάζει κάθε εικονοστοιχείο της annotation εικόνας και όταν κάποιο στοιχείο παίρνει την τιμή 255 που συμβολίζει το λευκό, καταχωρεί τις συντεταγμένες του σημείου στον πίνακα matrix	59
Πίνακας 3-5 : Μέρος του κώδικα το οποίο με την χρήση δύο επαναληπτικών βρόγχων for ελέγχει ποιες από τις συντεταγμένες του πίνακα rr ταιριάζουν με τις συντεταγμένες του πίνακα matrix που βρίσκονται τα ελαττώματα. Επίσης αποθηκεύει τ συντεταγμένες στον πίνακα syntetagmenes_shmeiwn.	59
Πίνακας 3-6 : Υπολογισμός χρόνου ολοκλήρωσης του αλγορίθμου με βάση τον ανιχνευτή που έχει επιλεχτεί.	63

1 Εισαγωγή

1.1 Σημερινές ανάγκες

Μια από τις μεγαλύτερες προκλήσεις που αντιμετωπίζουν οι μηχανικοί σήμερα, είναι ο έλεγχος, η εκτίμηση, η συντήρηση¹ και ο ασφαλής χειρισμός της υπάρχουσας δημόσιας υποδομής, όπως τα τούνελ, οι γέφυρες, οι δρόμοι, οι αγωγοί πετρελαίου κ.α. Εξαιτίας της γήρανσης, περιβαλλοντικών παραγόντων, αυξημένου φορτίου, αλλαγής χρήσης, ζημιών από τον άνθρωπο και φυσικούς παράγοντες, ανεπαρκούς συντήρησης και ετεροχρονισμένων επισκευών, οι δημόσιες υποδομές, βαθμιαία, φθείρονται και χρήζουν άμεσου ελέγχου, εκτίμησης και επισκευών.

Η ανάγκη αυτή είναι μεγαλύτερη σε υπόγεια τούνελ μαζικών μεταφορών, από τα οποία ένας μεγάλος αριθμός είναι σε λειτουργία για πάνω από μισό αιώνα. Σε αυτά υπάρχουν εκτεταμένα σημάδια φθοράς, όπως αποδεικνύεται από την αύξηση του ποσοστού των προϋπολογισμών που ξοδεύονται σε έλεγχο και εκτίμηση.

Τα τελευταία χρόνια, η κατάρρευση ενός σημαντικού αριθμού τούνελ², ανέδειξε την ανάγκη για αποτελεσματικότερους τρόπους ελέγχου και εκτίμησης της στατικότητας των τούνελ που βρίσκονται σε λειτουργία. Το κόστος κατασκευής ενός νέου τούνελ είναι πολύ υψηλό³, συνεπώς ο έλεγχος, η εκτίμηση και η επισκευή της υπάρχουσας υποδομής είναι ύψιστης σημασίας. Επιπλέον, στις προσεχείς δύο δεκαετίες, ο ρυθμός επέκτασης των υποδομών μαζικών μεταφορών στην Ευρωπαϊκή Ένωση, δεν θα είναι αντίστοιχος με την αύξηση των απαιτήσεων για μεταφορές, κάνοντας, έτσι, αναγκαία την μεγιστοποίηση του χρόνου λειτουργίας των τούνελ. Συνεπώς, χρειάζεται προληπτική συντήρηση, βασισμένη στον έγκαιρο έλεγχο και την έγκαιρη συντήρηση. Επιπρόσθετα, ο έλεγχος και η εκτίμηση οφείλουν να είναι γρήγορες διαδικασίες, προκειμένου τα τούνελ ή τμήματά τους να μην σταματούν τη λειτουργία τους. Τέλος, οι μηχανικές ώρες για τον έλεγχο και την εκτίμηση των τούνελ ελαττώνονται κατά πολύ.

¹ Συντήρηση: επέμβαση η οποία περιλαμβάνει δομική συντήρηση, επισκευή, επανένταξη και αναβάθμιση.

² Για παράδειγμα, το τούνελ-δρόμος Hanekleiv στην Νορβηγία (NGI et al., 2010) ή το σιδηροδρομικό τούνελ Bosruck στην Αυστρία (Klammer et al., 2008). Παρόμοια είναι η κατάσταση και εκτός Ευρώπης. Για παράδειγμα, το 2006, ένα κομμάτι σκυροδέματος έπεσε στο Fort Point Channel Tunnel της Βοστώνης, σκοτώνοντας ένα άτομο και τραυματίζοντας άλλο ένα, ενώ το 2012 στην Ιαπωνία, εννέα άτομα σκοτώθηκαν, όταν κατέρρευσε η οροφή του τούνελ-αυτοκινητοδρόμου Sasago, συνθλίβοντας τρία οχήματα με κομμάτια σκυροδέματος.

³ Το κόστος κατασκευής ενός νέου τούνελ κυμαίνεται από 120-160 εκατομμύρια ευρώ, ανά χιλιόμετρο.

Στις μέρες μας, ο στατικός έλεγχος των τούνελ εκτελείται, ως επί το πλείστον, μέσω προγραμματισμένων, περιοδικών, οπτικών παρατηρήσεων σε όλο το εύρος του τούνελ. Οι ελεγκτές αναγνωρίζουν τα δομικά ελαττώματα (π.χ. ρωγμή⁴ ή θρυμματισμός⁵), χαρακτηρίζουν τα ελαττώματα αυτά (όπως φαίνεται στον πίνακα 1.1) και τέλος, βασισμένοι στα αναγνωρισμένα ελαττώματα και στη σοβαρότητα τους, κατηγοριοποιούν την επένδυση του τούνελ⁶ (liner) στις κατηγορίες του πίνακα του ANNEX II.

Η διαδικασία αυτή είναι αργή, απαιτεί έντονη εργασία και χαρακτηρίζεται από υποκειμενικότητα όντας εξαρτώμενη από την εμπειρία και την κούραση του ελεγκτή ενώ αυτός εργάζεται σε ένα μη ευχάριστο περιβάλλον, εξαιτίας της σκόνης, της απουσίας του φυσικού φωτός και των άβολων συνθηκών. Επιπλέον, η κατηγοριοποίηση της κατάστασης της επένδυσης του τούνελ (liner) στον ANNEX II, είναι εμπειρική, ημιτελής και στερείται μηχανικής ανάλυσης (engineering analysis). Συνεπώς, δεν είναι αξιόπιστη.⁷

Τύπος Ελαττώματος	Δευτερεύον	Μέτριο	Πρωτεύον
Ρωγμή	μέχρι 0.8mm (πλάτος)	από 0.8mm μέχρι 3.2mm (πλάτος)	πάνω από 3.2mm (πλάτος)
Θρυμματισμός	από 75mm μέχρι 150mm (διάμετρος)	περίπου 150mm (διάμετρος)	πάνω από 150mm (διάμετρος)

Πίνακας 1-1 : Κατηγοριοποίηση ελαττωμάτων -ρωγμών και θρυμματισμών- σε επενδύσεις από σκυρόδεμα (FHTWA κ.α., 2005)

Τα τελευταία χρόνια, διάφορες εξελίξεις στην τεχνολογία, όπως οι υψηλής ανάλυσης, οικονομικές, ψηφιακές φωτογραφικές μηχανές και η διαθεσιμότητα αποθήκευσης πλήθους δεδομένων, έχουν επιτρέψει την ανάπτυξη ενός αριθμού μοντέλων υπολογιστικής όρασης, τα οποία μπορούν να ανιχνεύουν και να αναγνωρίζουν τα χαρακτηριστικά των επενδύσεων, όπως αρμούς και ελαττώματα και να κατηγοριοποιούν τα ελαττώματα σε διαφορετικές κατηγορίες ή να μετρούν το μήκος και τον προσανατολισμό των ρωγμών. Αυτές οι προσεγγίσεις δεν παρέχουν μετρήσεις των ελαττωμάτων με την ακρίβεια που απαιτείται σε μια στατική/δομική εκτίμηση.

⁴ Ένας ολικός ή μη διαχωρισμός του σκυροδέματος σε δύο ή περισσότερα τμήματα, ως αποτέλεσμα σπασίματος ή ραγίσματος. Προκαλείται από εφελκυστικές τάσεις, που υπερβαίνουν το όριο του εφελκυσμού του σκυροδέματος.

⁵ Ξεφλούδισμα και απώλεια υλικού σκυροδέματος, από μια εκτεθειμένη επιφάνεια, που προκαλείται από οξείδωση του ενσωματωμένου μεταλλικού οπλισμού ή από παγετό, τη δράση του αλατιού (σε παραθαλάσσιες περιοχές. Είναι μια περίπου κυκλική ή οβάλ εσοχή στο σκυρόδεμα.

⁶ Μόνιμη ή προσωρινή κάλυψη της επιφάνειας της πέτρας ή του χώματος, στον τοίχο μιας εκσκαφής ενός τούνελ. Εδώ, γίνεται αναφορά στην μόνιμη επένδυση σκυροδέματος ενός τούνελ.

⁷ Το τούνελ Sasago, το οποίο κατέρρευσε στην Ιαπωνία το 2012 σκοτώνοντας 9 ανθρώπους είχε υποβληθεί σε οπτικό έλεγχο δύο μήνες νωρίτερα του συμβάντος, χωρίς να έχουν εντοπιστεί ελαττώματα.

Επιπλέον, στερούνται μαθησιακών ικανοτήτων, που θα μπορούσαν να επιτρέψουν έναν πιο αξιόπιστο και γρήγορο έλεγχο των τούνελ.

Σε όλες τις παραπάνω περιπτώσεις (οπτικός έλεγχος ή μοντέλα υπολογιστικής όρασης), όταν γίνεται η αναγνώριση των περιοχών ενδιαφέροντος, γίνονται πιο λεπτομερείς μετρήσεις των ελαττωμάτων, σε ένα επόμενο βήμα, μέσω μη καταστρεπτικών ή καταστρεπτικών μεθόδων, που παρέχουν την απαιτούμενη εισαγωγή/πληροφορία (input) για την στατική/δομική ανάλυση. Αυτή η προσέγγιση, όσον αφορά την εκτίμηση, είναι αργή και απαιτεί έντονη εργασία και ακριβό εξοπλισμό. Επιπλέον, συχνά η διαδικασία πρέπει να γίνεται σε πολύ κοντινή απόσταση και απαιτεί πρόσβαση, στα επιθυμητά σημεία, με σκαλωσιές. Αυτό καθυστερεί την όλη διαδικασία και καθιστά αναγκαίο τον αποκλεισμό τουλάχιστον μιας λωρίδας κυκλοφορίας.

Η εξάπλωση των αυτοματοποιημένων, ταχύτερων, ελάχιστα ή καθόλου επεμβατικών στις υπηρεσίες των τούνελ, αξιόπιστων λύσεων ελέγχου και εκτίμησης των τούνελ με ένα πέρασμα, οι οποίες μπορούν να συνδυάσουν τον έλεγχο και τη λεπτομερή, δομική-στατική εκτίμηση, μπορεί να έχει σημαντικό αντίκτυπο στην ρομποτική αγορά. Η οποία αφορά τα ρομποτικά συστήματα (robotics market), τη βιομηχανία έλεγχου των τούνελ και την κοινωνία μας:

Τα συστήματα αυτά θα αυξήσουν την ανταγωνιστικότητα της βιομηχανίας ελέγχου των τούνελ, αυξάνοντας την ταχύτητα και την αξιοπιστία των ελέγχων. Επιπλέον, δίνοντας τη δυνατότητα εκτιμήσεων, θα ελαχιστοποιήσουν τη χρήση ελεγκτών τούνελ, οι οποίοι είναι δυσεύρετοι, ενώ θα βελτιώσουν τις συνθήκες εργασίας για αυτούς τους ελεγκτές και θα μειώσουν το κόστος ελέγχου και εκτίμησης.

Η ασφάλεια των επιβατών θα αυξηθεί, κάτι που αφορά ένα μεγάλο αριθμό ανθρώπων (π.χ. σε ένα δίκτυο μετρό, μπορούν να μετακινούνται 60.000 επιβάτες ανά ώρα). Επιπλέον, ο χρόνος στη διάρκεια του οποίου τα τούνελ θα παραμένουν κλειστά για έλεγχο και λεπτομερή εκτίμηση, θα μειωθεί. Κάτι τέτοιο έχει σημαντικές οικονομικές συνέπειες. Τα τούνελ είναι βασικά στοιχεία του Εθνικού και –Ευρωπαϊκού Δικτύου Μεταφορών (National and Trans European Transport Network) (π.χ. η τιμή των αγαθών που περνούν το τούνελ Frejus, μεταξύ Ιταλίας και Γαλλίας, ανέρχεται στα 100 δις ευρώ, ετησίως) και η χωρητικότητα κάποιων τούνελ, ακόμα και χωρίς αποκλεισμό των λωρίδων για έλεγχο και εκτίμηση, είναι ανεπαρκής.

Υπάρχει μια ευρεία αγορά για ένα ολοκληρωμένο, αυτοματοποιημένο, ρομποτικό σύστημα για τον έλεγχο των τούνελ που εξυπηρετούν μεταφορές και την εκτίμηση τους: Το μήκος των τούνελ μαζικών μεταφορών που λειτουργούν στην Ευρώπη είναι πάνω από 15.000 χιλιόμετρα

(Haack, 2004) και πολλά από αυτά απαιτούν άμεσο έλεγχο και εκτίμηση (π.χ. τμήματα του μετρό του Λονδίνου ή του Παρισιού, τα οποία μετρούν πάνω από 100 χρόνια λειτουργίας). Παρόμοια είναι η κατάσταση εκτός Ευρώπης. Για παράδειγμα το 16% του δικτύου μετρό των ΗΠΑ, αξιολογείται ως υποτυπώδες ή ανεπαρκές (ASCE, 2009). Επιπλέον, τα συστατικά στοιχεία ενός τέτοιου ρομποτικού συστήματος για την εκτίμηση των ορατών ελαττωμάτων, μπορεί να είναι χρήσιμα για τον έλεγχο ενός μεγάλου εύρους αστικών υποδομών, όπως γέφυρες, τοιχία αντιστήριξης κ.α., τα οποία απαιτούν επίσης έλεγχο του μηχανολογικού εξοπλισμού.

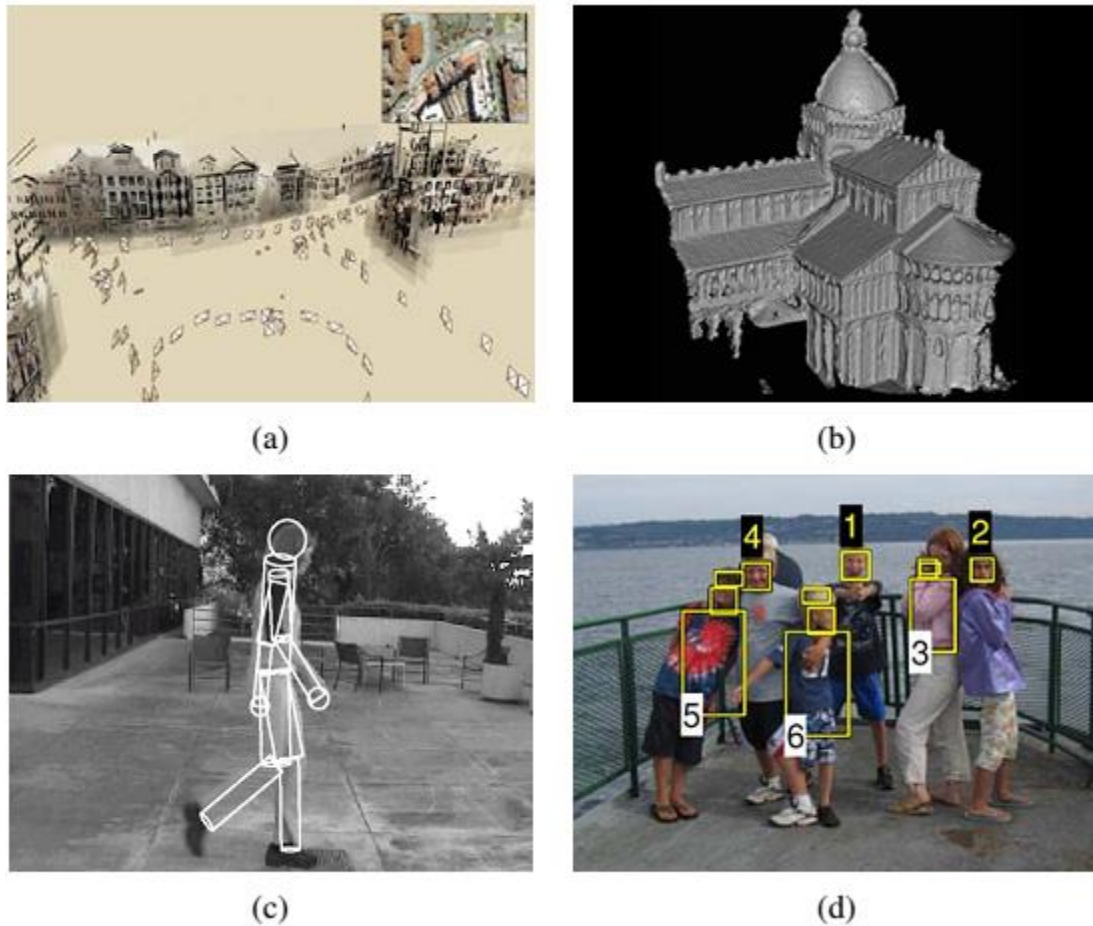
Τα παραπάνω εξηγούν γιατί η βιομηχανία κατασκευής τούνελ απαιτεί ένα αυτοματοποιημένο και ολοκληρωμένο σύστημα ελέγχου και εκτίμησης των τούνελ με ένα πέρασμα και γιατί η βιομηχανία ρομποτικών συστημάτων αντιμετωπίζει ένα τέτοιο σύστημα ως μια ελκυστική, πιθανή αγορά. Στην συγκεκριμένη διπλωματική συγκρίνουμε πέντε διαφορετικούς ανιχνευτές που ίσως μπορούσαν οι ίδιοι, επεκτάσεις αυτών ή συνδυασμός τους να χρησιμοποιηθούν σαν σύστημα ανίχνευσης ενός ρομποτικού συστήματος.

1.2 Υπολογιστική Όραση (Computer Vision)

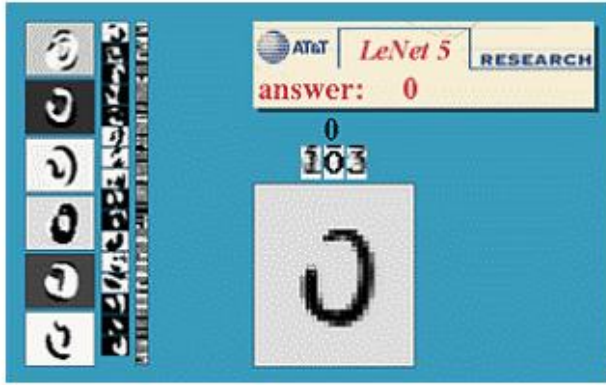
Η υπολογιστική όραση (ή μηχανική όραση ή τεχνητή όραση) συνιστά ένα επιστημονικό πεδίο της τεχνητής νοημοσύνης, στόχος του οποίου είναι η αλγοριθμική αναπαραγωγή της αίσθησης της όρασης, είτε σε ηλεκτρονικό υπολογιστή, είτε σε ρομπότ. Στους υπολογιστές το οπτικό ερέθισμα είναι ένα σύνολο από ψηφιακά δεδομένα, όπως ψηφιακές εικόνες ή ακολουθίες εικόνων (video), όψεις από πολλαπλές κάμερες, πολυδιάστατες εικόνες από ιατρικό σαρωτή. Στην υπολογιστική όραση έχουν αναπτυχθεί πάρα πολλές έρευνες. Ταυτόχρονα, αναπτύσσονται πολλές μαθηματικές τεχνικές που στοχεύουν στην ανάκτηση τρισδιάστατων σχημάτων και την εμφάνιση αντικειμένων σε εικόνες. Αυτή την στιγμή έχουμε πολλές αξιόπιστες τεχνικές για τον ακριβή υπολογισμό τρισδιάστατων μοντέλων περιβάλλοντος από χιλιάδες επικαλυπτόμενες εικόνες (εικόνα 1.1.a). Επιπρόσθετα, έχουμε την δυνατότητα όταν έχουμε φωτογραφίες από πολλές οπτικές γωνίες ενός αντικειμένου να δημιουργήσουμε πυκνά τρισδιάστατα μοντέλα με την χρήση στερεοφωνικής αντιστοίχισης (εικόνα 1.1.b). Μπορούμε να εντοπίσουμε την κίνηση ενός ανθρώπου που κινείται σε ένα σύνθετο φόντο (εικόνα 1.1.c) αλλά και να ανιχνεύουμε ένα συνδυασμό χαρακτηριστικών από ανθρώπους σε φωτογραφίες που έχει ως κύριο σκοπό την ταυτοποίηση (εικόνα 1.1.d). Σύμφωνα και με τον Szeliski ενώ υπάρχουν όλες αυτές οι εξελίξεις υπάρχουν και πολλές δυσκολίες. Αυτό συμβαίνει διότι η όραση υπολογιστών είναι αντικειμενοστραφής φύσεως. Δηλαδή πολλές φορές χρειαζόμαστε να ανακτήσουμε πληροφορίες οι οποίες παραμένουν άγνωστες για εμάς ή δεν έχουμε καθορίσει πλήρως την λύση τους. Στην όραση υπολογιστών προσπαθούμε να περιγράψουμε τον κόσμο που βλέπουμε από μία ή και περισσότερες φωτογραφίες ώστε να μπορέσουμε να ανοικοδομήσουμε ιδιότητες όπως το σχήμα, ο φωτισμός και οι διανομές χρωμάτων. Ένα πολύ βασικό συστατικό πολλών εφαρμογών στην όραση υπολογιστών είναι η ανίχνευση χαρακτηριστικών (feature detection). Ο στόχος της παρούσας διπλωματικής είναι η ανίχνευση ρωγμών και θρυμματισμών σε τούνελ από σκυρόδεμα. Όμως η όραση υπολογιστών όμως έχει στο δυναμικό της πολλές άλλες εφαρμογές όπως :

- Οπτική αναγνώριση χαρακτήρων(OCR) (εικόνα 1.2.a)
- Επιθεώρηση μηχανών : Γρήγορος έλεγχος τμημάτων μίας μηχανής χρησιμοποιώντας στερεοσκοπική όραση με εξειδικευμένο φωτισμό για την διασφάλιση της ποιότητας της (π.χ. έλεγχος των ανοχών των πτερών ενός αεροσκάφους ή και μέρη ενός αυτοκινήτου (εικόνα 1.2.b))

- Λιανεμπόριο : Αναγνώριση αντικειμένων σε ένα ταμείο (εικόνα 1.2.c)
- Κατασκευή τρισδιάστατων μοντέλων κτιρίων : Πλήρως αυτοματοποιημένη κατασκευή τρισδιάστατων μοντέλων από αεροφωτογραφίες που χρησιμοποιούνται σε συστήματα όπως το Bing Maps
- Ιατρική απεικόνιση : Προεγχειρητική και διεγχειρητική απεικόνιση (εικόνα 1.2.d)
- Αυτοκίνητα ασφαλείας : ανίχνευση απροσδόκητων εμποδίων όπως πεζοί στον δρόμο (εικόνα 1.2.e)
- Ταυτοποίηση κίνησης : Συγχώνευση εικόνων που έχουν παραχθεί από υπολογιστή(CGI) με ζωντανό υλικό δράσης από βίντεο για την ανίχνευση χαρακτηριστικών σημείων, τα οποία προβλέπουν την μορφή του περιβάλλοντος (εφαρμόζεται πολύ στο Χόλιγουντ)
- Καταγραφή της κίνησης (mocap) : Χρήση ανακλαστικών δεικτών που παρατηρούνται από πολλαπλές κάμερες ώστε να καταγραφεί η κίνηση των ηθοποιών για τα κινούμενα σχέδια
- Παρακολούθηση : Παρακολούθηση εισβολέων, ανάλυση κίνησης αυτοκινήτων (εικόνα 1.2.f)
- Αναγνώριση δακτυλικών αποτυπωμάτων και άλλων βιομετρικών στοιχείων : Για τον αυτόματο έλεγχο πρόσβασης και τις ιατροδικαστικές εφαρμογές



Εικόνα 1.1 : (a) Υπολογισμός τρισδιάστατων μοντέλων περιβάλλοντος από χιλιάδες επικαλυπτόμενες εικόνες. (b) Δημιουργία πυκνών τρισδιάστατων μοντέλων με την χρήση στερεοφωνικής αντιστοίχισης από φωτογραφίες διαφορετικών οπτικών γωνιών. (c) Εντοπισμός κίνησης ενός ανθρώπου που κινείται σε ένα σύνθετο φόντο. (d) Αλγόριθμοι ανίχνευσης προσώπου που προκύπτει από ένα συνδυασμό χαρακτηριστικών από ανθρώπους σε φωτογραφίες.



(a)



(b)



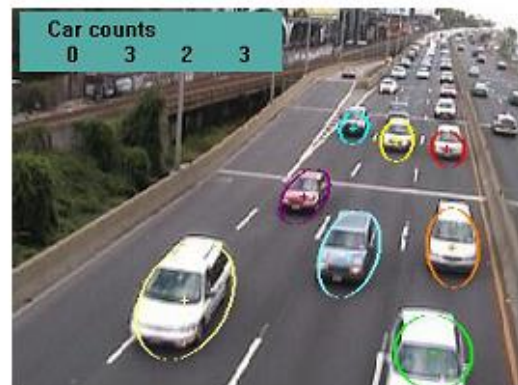
(c)



(d)



(e)



(f)

Εικόνα 1.2 : (a) Οπτική αναγνώριση χαρακτήρων(OCR). (b) Επιθεώρηση μηχανών. (c) Λιανεμπόριο. (d) Ιατρική απεικόνιση. (e) Αυτοκίνητα ασφαλείας. (f) Παρακολούθηση(ανάλυση κίνησης αυτοκινήτων).

1.3 Ανίχνευση χαρακτηριστικών (Feature detection)

Στην όραση υπολογιστών (computer vision) και στην επεξεργασία εικόνας η έννοια της ανίχνευσης χαρακτηριστικών παραπέμπει σε μεθόδους που στοχεύουν στον υπολογισμό αφηρημένων πληροφοριών μιας φωτογραφίας, λαμβάνοντας τοπικές αποφάσεις για κάθε σημείο της εικόνας ανεξάρτητα αν το κάθε σημείο έχει τα χαρακτηριστικά που μας ενδιαφέρουν. Τα προκύπτοντα χαρακτηριστικά θα είναι υποσύνολα του τομέα της εικόνας (image domain), συχνά με τη μορφή απομονωμένων σημείων, συνεχής καμπύλες ή συνδεδεμένες περιοχές. Δεν υπάρχει καθολική ή ακριβής ορισμός του τι συνιστά ένα χαρακτηριστικό, και ο ακριβής ορισμός εξαρτάται συχνά από το πρόβλημα ή το είδος της εφαρμογής. Στην προκειμένη περίπτωση το πρόβλημα-χαρακτηριστικό που αντιμετωπίζουμε είναι ο εντοπισμός ρωγμών και θρυμματισμών σε τούνελ από σκυρόδεμα. Με βάση τα παραπάνω, ένα χαρακτηριστικό ορίζεται ως ένα ενδιαφέρον κομμάτι της εικόνας. Γενικότερα τα χαρακτηριστικά χρησιμοποιούνται ως σημείο εκκίνησης για πολλούς αλγόριθμους όρασης των υπολογιστών. Δεδομένου ότι τα χαρακτηριστικά χρησιμοποιούνται ως σημείο εκκίνησης και κύρια αρχέτυπα για αλγόριθμους, συνολικά ο αλγόριθμος θα είναι τις περισσότερες φορές τόσο καλός όσο ανιχνευτής των χαρακτηριστικών. Κατά συνέπεια, η επιθυμητή ιδιότητα για έναν ανιχνευτή χαρακτηριστικών είναι η επαναληψιμότητα, αυτό σημαίνει ότι ούτως ή άλλως το ίδιο χαρακτηριστικό θα πρέπει να ανιχνεύεται σε δύο ή περισσότερες διαφορετικές εικόνες της ίδιας σκηνής. Η ανίχνευση χαρακτηριστικών είναι συνήθως η πρώτη ενέργεια που πραγματοποιείται σε μία εικόνα. Εξετάζει κάθε εικονοστοιχείο (pixel) ώστε να παρατηρηθεί αν υπάρχει το συγκεκριμένο χαρακτηριστικό που ψάχνουμε. Σε περίπτωση που αυτή η ενέργεια είναι απλά ένα μέρος του αλγορίθμου, τότε ο αλγόριθμος θα εξετάσει και θα περιγράψει (feature description) τα μέρη της εικόνας που έχουν αυτά τα χαρακτηριστικά. Η εικόνα που εισάγεται εξομαλύνεται τις περισσότερες φορές με φίλτρο του Gauss στην χωροκλίμακα αναπαράστασης.

Πολλοί αλγόριθμοι υπολογιστικής όρασης χρησιμοποιούν την ανίχνευση χαρακτηριστικών για αρχικό βήμα, αυτό έχει ως αποτέλεσμα να έχουν αναπτυχθεί πολλοί ανιχνευτές χαρακτηριστικών σημείων. Οι οποίοι διαφέρουν στην επαναληψιμότητα και την υπολογιστική πολυπλοκότητα ανάλογα με το είδος των χαρακτηριστικών που ανιχνεύονται. Οι ανιχνευτές χωρίζονται συνήθως στις ακόλουθες ομάδες.

1.3.1 Ακμές (Edges)

Ακμές είναι τα σημεία όπου υπάρχει ένα όριο (ή μια ακμή) μεταξύ δύο περιοχών της εικόνας. Σε γενικές γραμμές, μία ακμή μπορεί να είναι σχεδόν αυθαίρετου σχήματος, και μπορεί να περιλαμβάνουν διασταυρώσεις. Στην πράξη, τα άκρα ορίζονται συνήθως ως σύνολα σημείων της εικόνας που έχουν μια ισχυρή κλίση μεγέθους. Επιπλέον, ορισμένοι αλγόριθμοι συνδέουν αυτά τα σημεία μαζί ώστε να σχηματιστεί μια πιο ολοκληρωμένη περιγραφή μίας ακμής. Αυτοί οι αλγόριθμοι τοποθετούν συνήθως ορισμένους περιορισμούς σχετικά με τις ιδιότητες μιας ακμής, όπως το σχήμα, την ομαλότητα, και τις τιμές της κλίσης. Σε τοπικό επίπεδο, τα άκρα έχουν μονοδιάστατη δομή.

1.3.2 Γωνίες-Σημεία ενδιαφέροντος (Corners)

Μια γωνία μπορεί να οριστεί ως η τομή δύο ακμών. Μια γωνιά μπορεί επίσης να οριστεί ως ένα σημείο για το οποίο υπάρχουν δύο κυρίαρχες και διαφορετικές κατευθύνσεις ακμών σε μια τοπική γειτονιά του σημείου. Ένα σημείο ενδιαφέροντος είναι ένα σημείο το οποίο βρίσκεται σε μία καλά ορισμένη θέση της εικόνας και μπορεί να ανιχνευτεί. Αυτό σημαίνει ότι ένα σημείο ενδιαφέροντος μπορεί να είναι μια γωνία, αλλά μπορεί επίσης να έχει και διαφορετικές μορφές. Για παράδειγμα, θα μπορούσε να είναι ένα απομονωμένο σημείο τοπικής μέγιστης ή ελάχιστης έντασης, ένα σημείο που βρίσκεται στο τέλος μίας γραμμής καθώς και ένα σημείο μίας καμπύλης όπου η καμπυλότητα είναι τοπικά μέγιστη.

Οι όροι γωνίες και τα σημεία ενδιαφέροντος χρησιμοποιούνται κάπως εναλλακτικά και αναφέρονται όπως τα χαρακτηριστικά σε μια εικόνα, η οποία έχει μια τοπική διδιάστατη δομή. Το όνομα «Γωνιά» προέκυψε από τους πρώτους αλγορίθμους που χρησιμοποιήθηκαν για την ανίχνευση άκρων, μετά από αυτό ακολούθησε η ανάλυση των άκρων και η γρήγορη αλλαγή κατεύθυνσης τους δηλαδή η εύρεση γωνιών. Αυτοί οι αλγόριθμοι αναπτύχθηκαν τότε, έτσι ώστε η κατηγορηματική ανίχνευση άκρων δεν ήταν πλέον αναγκαία. Για παράδειγμα, βλέποντας τα υψηλά επίπεδα της καμπύλης της κλίσης της εικόνας, παρατηρήθηκε ότι οι λεγόμενες και ως γωνίες ανιχνεύονταν και σε μέρη τα οποία δεν ήταν γωνίες με την παραδοσιακή έννοια (για παράδειγμα μπορεί να ανιχνευθεί ένα μικρό φωτεινό σημείο σε σκούρο φόντο). Τα σημεία αυτά συχνά είναι γνωστά ως σημεία ενδιαφέροντος, αλλά ο όρος γωνιά χρησιμοποιείται από την

παράδοση. Μία από τις μεθόδους ανίχνευσης ακμών που χρησιμοποιείται και στην παρούσα διπλωματική είναι ο ανιχνευτής Harris-Stephens.

1.3.3 Άμορφα σημεία (Blobs) - περιοχές ενδιαφέροντος - ενδιαφέροντα σημεία

Στην όραση υπολογιστών, οι μέθοδοι ανίχνευσης άμορφης μάζας σε μία ψηφιακή εικόνα στοχεύουν εντοπισμό περιοχών που διαφέρουν ως προς τις ιδιότητες όπως για παράδειγμα το χρώμα και η φωτεινότητα. Όλα τα σημεία μίας άμορφης μάζας είναι παρόμοια μεταξύ τους αν όχι όμοια. Παρέχουν μια συμπληρωματική περιγραφή των δομών της εικόνας από την άποψη των περιοχών, σε αντίθεση με τις γωνίες που είναι περισσότερο σημειακές. Ωστόσο αυτοί οι περιγραφείς μπορεί να εμπεριέχουν συχνά ένα επιθυμητό σημείο (ένα τοπικό μέγιστο της απόκρισης του χειριστή ή ένα κέντρο βάρους). Πράγμα που σημαίνει ότι πολλοί τέτοιοι ανιχνευτές μπορεί επίσης να θεωρηθούν ως χειριστές σημείου ενδιαφέροντος. Οι ανιχνευτές άμορφων σημείων μπορεί να ανιχνεύσουν περιοχές σε μια εικόνα που είναι πολύ λεία για να ανιχνευθεί από έναν ανιχνευτή γωνιών. Για παράδειγμα αν μειώναμε το μέγεθος μίας εικόνας και εκτελούσαμε ανίχνευση γωνιών, ο ανιχνευτής θα ανταποκρινόταν σε σημεία τα οποία είναι απότομα. Αντίθετα αυτά τα σημεία θα ήταν ομαλά στο κανονικό μέγεθος της εικόνας. Υπάρχουν πολλά κίνητρα για τη μελέτη και την ανάπτυξη ανιχνευτών άμορφης μάζας. Ένας βασικός λόγος είναι η παροχή συμπληρωματικών πληροφοριών σχετικά με τις περιοχές σημείων. Κάτι το οποίο δεν παρέχεται από τους ανιχνευτές ακμών ή ανιχνευτές γωνιών. Οι μέθοδοι ανίχνευσης άμορφων σημείων έχουν πολλές εφαρμογές όπως η αναγνώριση αλλά ακόμα και η παρακολούθηση αντικειμένων. Άλλωστε μία άμορφη μάζα θα μπορούσε να είναι ένα αμάξι ή ακόμα και ένας άνθρωπος.

1.3.4 Κορυφές (Ridges)

Για επιμήκη αντικείμενα, η έννοια των κορυφογραμμών είναι ένα φυσικό εργαλείο. Ο υπολογισμός ενός περιγραφέα κορυφογραμμής σε μια εικόνα με αποχρώσεις του γκρι μπορεί να θεωρηθεί ως μια γενίκευση ενός ενδιάμεσου άξονα. Από πρακτική άποψη, μια κορυφογραμμή μπορεί να θεωρηθεί ως μια μονοδιάστατη καμπύλη που αντιπροσωπεύει έναν άξονα συμμετρίας. Επιπρόσθετα ένα χαρακτηριστικό γνώρισμα είναι ότι το πλάτος της τοπικής κορυφής συνδέεται με κάθε σημείο της κορυφής. Είναι γεγονός ότι είναι πιο δύσκολο αλγοριθμικά να εξαχθούν τα

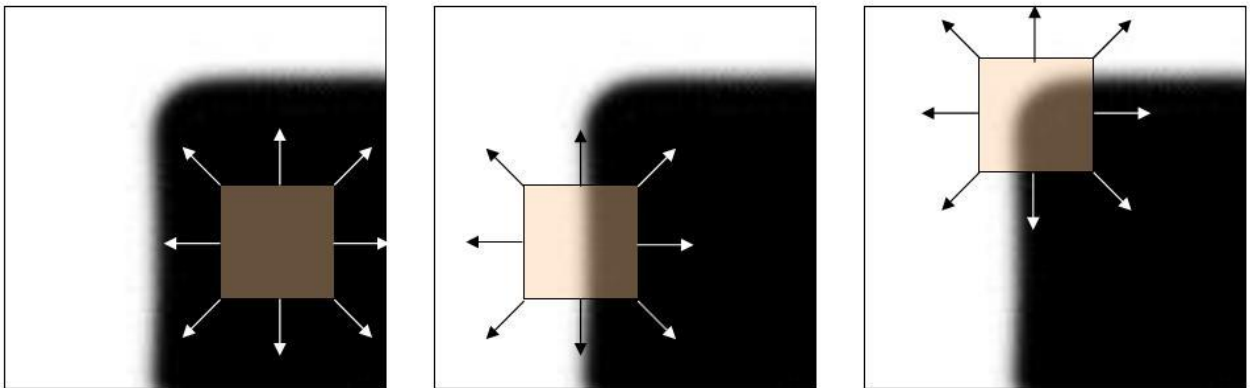
χαρακτηριστικά γνωρίσματα μίας κορυφογραμμής από τις γενικές κατηγορίες εικόνων με αποχρώσεις του γκρι παρά ακμών, γωνιών και άμορφων χαρακτηριστικών. Παρ' όλα αυτά, χρησιμοποιούνται συχνά για την εξαγωγή σημείων από αεροφωτογραφίες δρόμων και αιμοφόρων αγγείων σε ιατρικές εικόνες.

2 Ανάλυση ανιχνευτών

Σε αυτό το κεφάλαιο θα αναλυθεί το θεωρητικό υπόβαθρο των 5 διαφορετικών ανιχνευτών που θα χρησιμοποιηθούν στο τελικό αλγόριθμο της παρούσας διπλωματικής. Αυτοί οι ανιχνευτές έχουν πολλές εφαρμογές, όμως εδώ θα χρησιμοποιηθούν μόνο για την εύρεση ρωγμών και θρυμματισμών σε τούνελ από σκυρόδεμα.

2.1 Ανιχνευτής γωνιών Harris

Ο ανιχνευτής γωνιών Harris, προτάθηκε από τους Harris και Stephens το 1988 και καθορίζει αν ένα σημείο βρίσκεται σε γωνία, σε ακμή ή αν είναι ένα σημείο περιοχής με ομοιόμορφη ένταση. Ένα σημείο θα μπορούσε να αναγνωριστεί αν είναι γωνία κοιτώντας το από ένα μικρό παράθυρο (εικόνα 2.1). Μετατοπίζοντας το παράθυρο σε οποιαδήποτε κατεύθυνση θα πρέπει να μας δώσει μεγάλη αλλαγή στην ένταση σε τουλάχιστον δύο κατευθύνσεις.



Εικόνα 2.1 : Στην αριστερή εικόνα έχουμε επίπεδη περιοχή χωρίς αλλαγή έντασης. Στην μεσαία εικόνα έχουμε ακμή με καμία μεταβολή κατά μήκος της διεύθυνση της ακμής. Στην δεξιά εικόνα έχουμε γωνία και μεγάλη αλλαγή έντασης σε τουλάχιστον δύο κατευθύνσεις.

Συγκεκριμένα ο ανιχνευτής γωνιών Harris είναι βασισμένος στην τοπική συνάρτηση αυτοσυσχέτισης (local autocorrelation function), η οποία περιγράφει την τοπική αλλαγή φωτεινότητας σε μία τοπική “γειτονιά” ενός σημείου (x,y) . Δεδομένου μιας μετατόπισης $(\Delta x, \Delta y)$ και του σημείου (x,y) η συνάρτηση αυτοσυσχετισμού ορίζεται ως εξής :

$$c(x, y) = \sum_{(x_i, y_i) \in W} [I(x_i, y_i) - I(x_i + \Delta x, y_i + \Delta y)]^2 \quad (2.1)$$

Όπου $I(x, y)$ υποδηλώνει την ένταση της εικόνας και (x_i, y_i) τα σημεία σε ένα παράθυρο W με κέντρο το (x, y) . Η ένταση των σημείων του παραθύρου W εξομαλύνεται με την χρήση Γκαουσιανού φίλτρου με κέντρο το (x, y) :

$$G(x_i, y_i) = e^{-[(x_i - x)^2 + (y_i - y)^2] / 2\sigma^2} \quad (2.2)$$

Το μέγεθος του Γκαουσιανού φίλτρου καθορίζει το ισοζύγιο μεταξύ της ακρίβειας εντοπισμού του σημείου ενδιαφέροντος και της αποφυγής του θορύβου της εικόνας. Η συνάρτηση αυτοσυσχετισμού προσεγγίζεται, από το πρώτο μέρος του αναπτύγματος Taylor :

$$c(x, y) = \sum [\Delta x \ \Delta y] M \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \quad (2.3)$$

Το M είναι ο πίνακας αυτοσυσχέτισης της εικόνας και ορίζεται ως εξής:

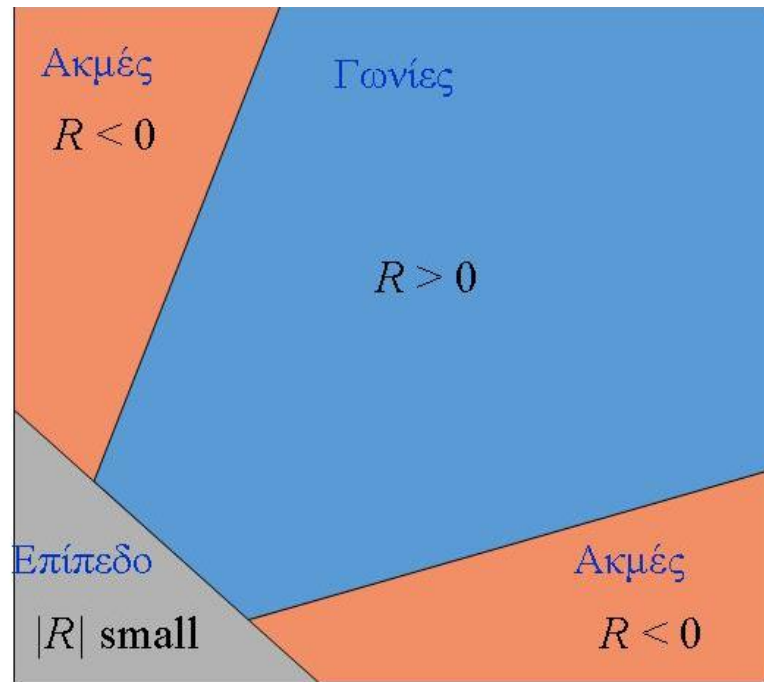
$$M = \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \quad (2.4)$$

Όπου I_x και I_y είναι οι μερικοί παράγωγοι της έντασης της εικόνας στις κατευθύνσεις x και y αντίστοιχα. Οι Harris και Stephens για να αποφύγουν το υπολογιστικό κόστος προτείνουν τη χρήση ενός μεγέθους που λέγεται απόκριση της γωνίας (cornerness response) που περιλαμβάνει την ορίζουσα (determinant) και το ίχνος (trace) του πίνακα της αυτοσυσχέτισης. Η απόκριση της γωνίας σε ένα σημείο (x_0, y_0) ορίζεται :

$$R(x_0, y_0) = \det M - k \cdot (\text{trace } M)^2 \quad (2.5)$$

Το k είναι μία μικρή θετική παράμετρος, η οποία ορίζει την ευαισθησία του ανιχνευτή και παίρνει τιμές κοντά στο 0,05. Η απόκριση R (εικόνα 2.2) :

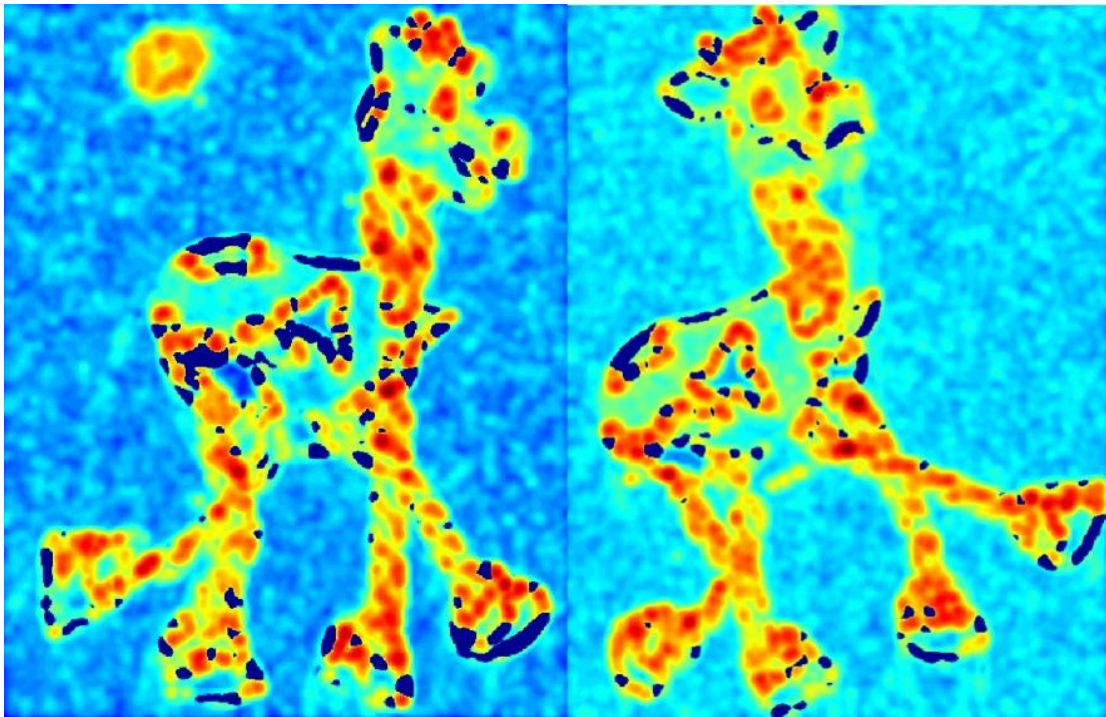
- Παίρνει μεγάλες θετικές τιμές στα σημεία που είναι γωνίες
- Αρνητικές τιμές στα σημεία που είναι ακμές
- Τιμές που τείνουν προς το μηδέν σε σημεία που είναι σε ομαλές και επίπεδες περιοχές



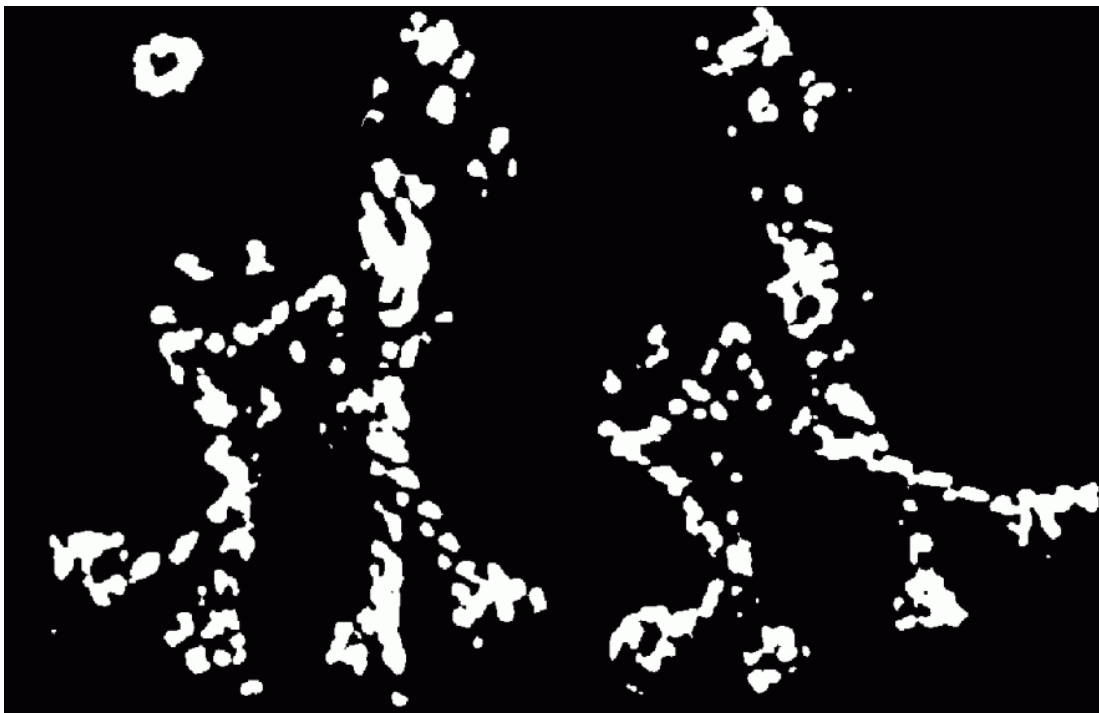
Εικόνα 2.2 : Δείχνει την αλληλεξάρτηση του μεγέθους R με το είδος των σημείων

Συνοπτικά τα βήματα του αλγορίθμου είναι :

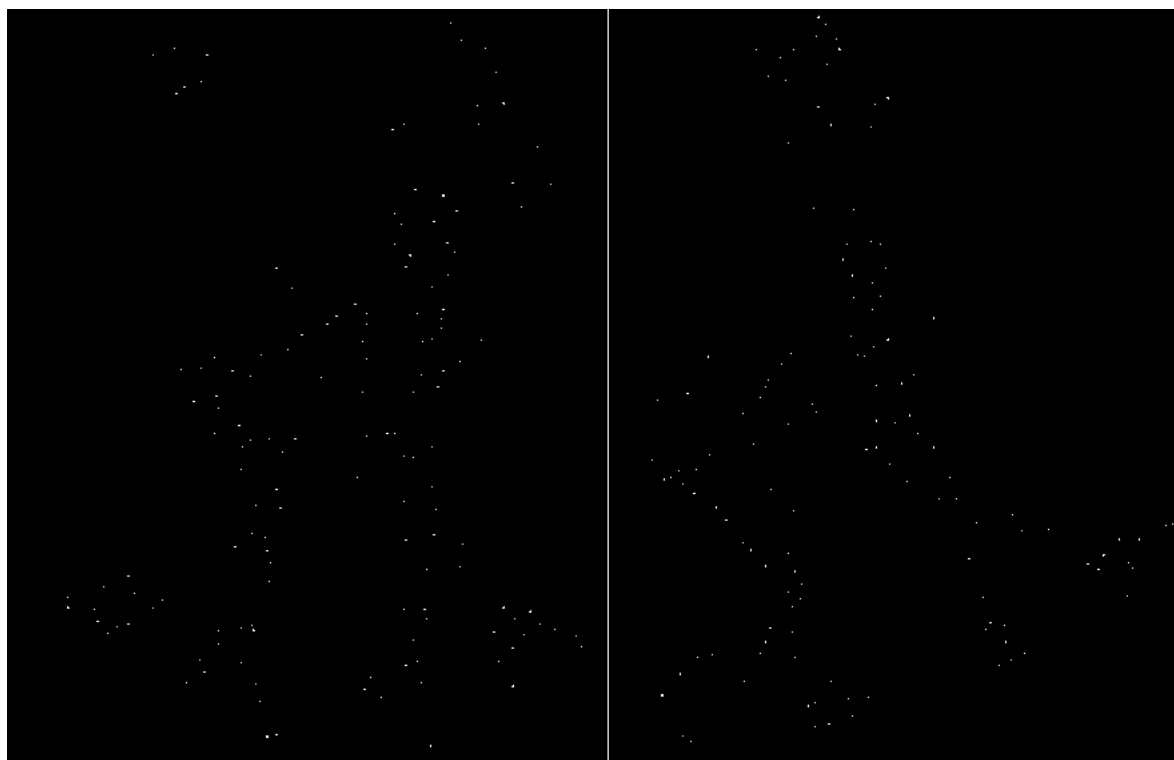
1. Υπολογισμοί των I_x και I_y , δηλαδή των μερικών παραγώγων της έντασης της εικόνας στις κατευθύνσεις x και y αντίστοιχα.
2. Υπολογισμοί των I_x^2 , I_y^2 , $I_x I_y$ για κάθε εικονοστοιχείο.
3. Υπολογισμοί των αθροισμάτων των προηγούμενων τετραγώνων $\sum I_x^2$, $\sum I_y^2$, $\sum I_x I_y$ (τα οποία είναι τα ορίσματα του πίνακα M) με την χρήση ενός Γκαουσιανού φίλτρου εξομάλυνσης.
4. Υπολογισμός της απόκριση της γωνίας R με τον τύπο:
$$R(x_0, y_0) = \det M - k \cdot (\text{trace } M)^2$$
 (εικόνα 2.3).
5. Βρίσκουμε τα σημεία με μεγάλη τιμή του R , μεγαλύτερη και από το όριο κατωφλίσωσης (threshold) δηλαδή ($R > \text{threshold}$) (εικόνα 2.4)
6. Βρίσκουμε τα τοπικά μέγιστα του R , αυτά είναι και τα σημεία ενδιαφέροντος που ψάχνουμε (εικόνα 2.5).



Εικόνα 2.3: Υπολογισμός της απόκρισης της γωνίας R.



Εικόνα 2.4 : Βρίσκουμε τα σημεία με μεγάλο R ($R >$ ορίου κατωφλίου)



Εικόνα 2.5 : Κρατάμε τα σημεία με τα τοπικά μέγιστα του R. Αυτά είναι και τα σημεία ενδιαφέροντος

2.2 Αλγόριθμος ελάχιστων ιδιοτιμών Shi-Tomasi

Ο αλγόριθμος ελαχίστων ιδιοτιμών (minimum eigenvalue algorithm), δημιουργήθηκε από τους Shi και Tomasi το 1993 και είναι μια περαιτέρω ανάπτυξη του ανιχνευτή γωνιών Harris. Η μικρή διαφορά μεταξύ αυτών των δύο ανιχνευτών έγκειται στη λειτουργία βαθμολόγησης που χρησιμοποιείται για να αποφασίσει αν το εικονοστοιχείο μιας εικόνας είναι γωνία. Σε αντίθεση με τον ανιχνευτή γωνιών Harris, ο αλγόριθμος ελάχιστων ιδιοτιμών δίνει έμφαση στον υπολογισμό των ιδιοτιμών λ_1, λ_2 της απόκρισης γωνίας R όπου:

$$R(x_0, y_0) = \det M - k \cdot (\text{trace } M)^2 \quad (2.5)$$

$$\text{Det}(M) = \lambda_1 \cdot \lambda_2$$

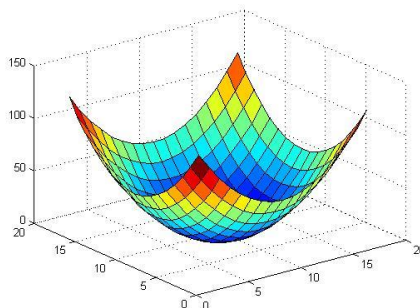
$$\text{Trace}(M) = \lambda_1 + \lambda_2$$

Με αυτό τον τρόπο μετατρέπει την απόκριση γωνίας R ως εξής :

$$R = \min(\lambda_1, \lambda_2)$$

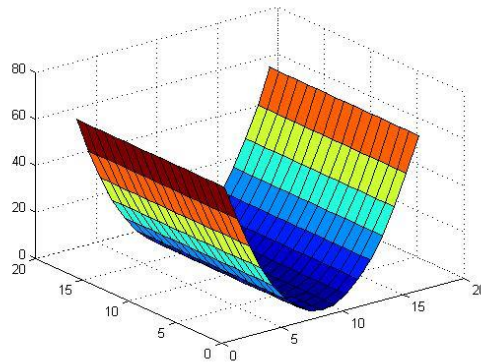
Δηλαδή ανάλογα με τις ιδιοτιμές λ_1, λ_2 εξάγουμε τις ακόλουθες περιπτώσεις (εικόνα 2.9) :

- Αν και οι δύο ιδιοτιμές είναι μεγάλες τότε οι μετατοπίσεις σε οποιαδήποτε κατεύθυνση έχουν μεγάλες μεταβολές έντασης. Η συνάρτηση αυτοσυσχετισμού $c(x,y)$ παίρνει την μορφή της εικόνας 2.6. Οπότε έχουμε σημείο που αντιστοιχεί σε γωνία, το οποίο είναι και ένα καλό χαρακτηριστικό (good feature to track) σύμφωνα με τους Shi και Tomasi.



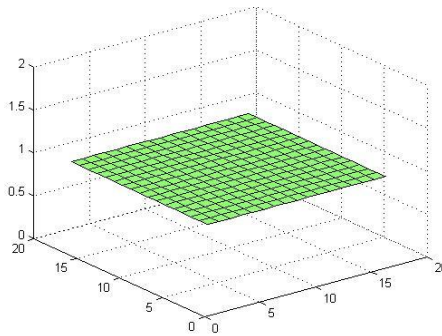
Εικόνα 2.6 : Μορφή της συνάρτησης αυτοσυσχετισμού $c(x,y)$ για σημείο που αντιστοιχεί σε γωνία

- Αν η μία ιδιοτιμή είναι πολύ μεγαλύτερη από την άλλη, τότε η συνάρτηση $c(x,y)$ παίρνει την μορφή ράχης (εικόνα 2.7). Η περίπτωση αυτή αντιστοιχεί σε ακμή.

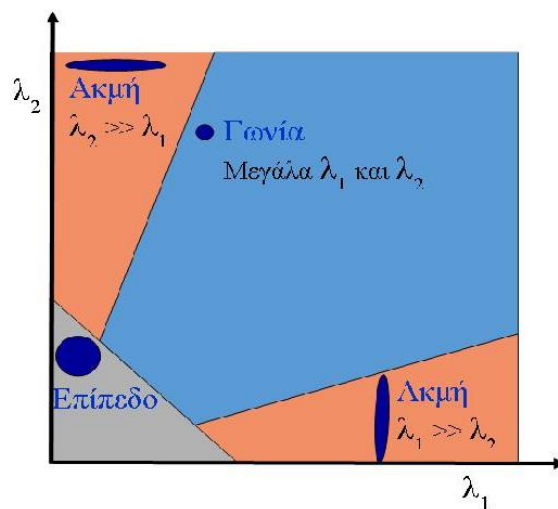


Εικόνα 2.7: Μορφή της συνάρτησης αυτοσυσχετισμού $c(x,y)$ για σημείο που αντιστοιχεί σε ακμή

- Αν και οι δύο ιδιοτιμές έχουν πολύ μικρές τιμές, τότε έχουμε πολύ μικρές μεταβολές σε οποιαδήποτε κατεύθυνση. Η συνάρτηση $c(x,y)$ παίρνει την μορφή της εικόνας 2.8. Αυτό έχει ως αποτέλεσμα να μιλάμε για επίπεδη περιοχή (flat region).



Εικόνα 2.8 : Μορφή της συνάρτησης αυτοσυσχετισμού $c(x,y)$ για σημείο που αντιστοιχεί σε επίπεδη περιοχή

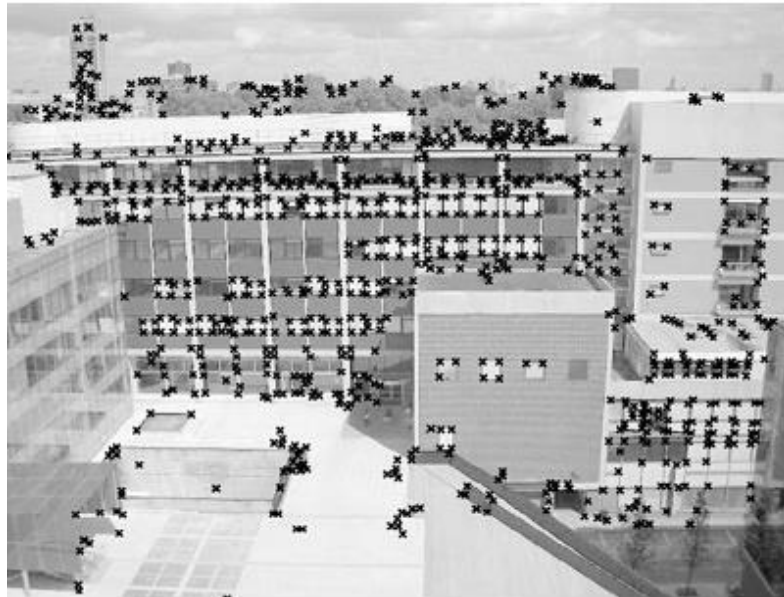


Εικόνα 2.9: Αλληλεξάρτηση των ιδιοτιμών λ_1, λ_2 με το είδος των σημείων

Τα βήματα του αλγορίθμου ελαχίστων ιδιοτιμών ταιριάζουν με αυτά του ανιχνευτή Harris, με την μόνη διαφορά τους στον υπολογισμό του μεγέθους R .



Εικόνα 2.10 : Φωτογραφία εισόδου για τον αλγόριθμο ελαχίστων ιδιοτιμών Shi-Tomasi



Εικόνα 2.11 : Τα χαρακτηριστικά που έχουν ανιχνευτεί από τον αλγόριθμο ελαχίστων ιδιοτιμών Shi-Tomasi. Οι τοποθεσίες των χαρακτηριστικών εμφανίζονται με σταυρούς

2.3 Ανιχνευτής Χαρακτηριστικών με Επιταχυνόμενο έλεγχο Τμημάτων (FAST-Features from Accelerated Segment Test)

Ο FAST είναι ένας αλγόριθμος που δημιουργήθηκε από τους Edward Rosten και Tom Drummond το 2005 για την αναγνώριση σημείων ενδιαφέροντος σε μια εικόνα. Ένα σημείο ενδιαφέροντος σε μια εικόνα είναι ένα εικονοστοιχείο (pixel), το οποίο έχει μια καθορισμένη θέση και μπορεί να ανιχνευτεί χωρίς ανωμαλίες κατά την εκτέλεση. Τα σημεία ενδιαφέροντος έχουν υψηλό περιεχόμενο τοπικής πληροφορίας και ιδανικά, θα έπρεπε να επαναλαμβάνονται μεταξύ διαφορετικών εικόνων. Η ανίχνευση ενός σημείου ενδιαφέροντος έχει εφαρμογές στο ταίριασμα εικόνων, στην αναγνώριση αντικειμένων, στον εντοπισμό κ.α.

Η ιδέα της ανίχνευσης ενός σημείου ενδιαφέροντος ή μιας ανίχνευσης ακμών⁸ δεν είναι καινούρια. Υπάρχουν διάφοροι γνωστοί αλγόριθμοι, όπως ο αλγόριθμος ανίχνευσης ακμών Moravec, οι αλγόριθμοι ανίχνευσης ακμών Harris & Stephens, ο ανιχνευτής ακμών SUSAN.

Η ιδέα πίσω από τον αλγόριθμο FAST είναι η ανάπτυξη ενός ανιχνευτή σημείων ενδιαφέροντος, που θα χρησιμοποιείται σε εφαρμογές σε πραγματικό χρόνο (use in real time frame rate applications), όπως η χρήση SLAM (simultaneous localization and mapping) σε ένα κινητό ρομπότ, οι οποίες έχουν περιορισμένους υπολογιστικούς πόρους.

2.3.1 Ανίχνευση χαρακτηριστικών χρησιμοποιώντας τον FAST

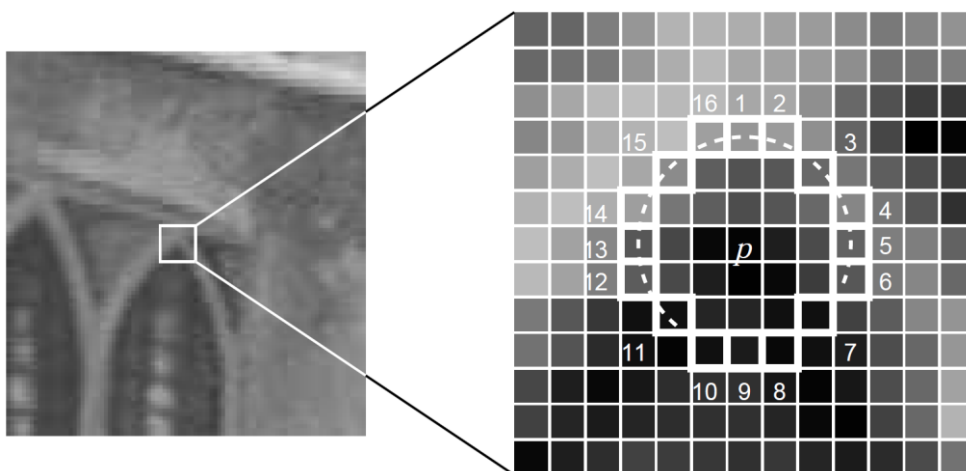
Επεξήγηση αλγορίθμου:

1. Επιλέγουμε ένα εικονοστοιχείο «p» στην εικόνα. Θεωρούμε την ένταση αυτού του εικονοστοιχείου IP. Αυτό είναι το εικονοστοιχείο το οποίο πρέπει να αναγνωριστεί ως ένα σημείο ενδιαφέροντος ή όχι (εικόνα 2.12).
2. Θέτουμε μια τιμή με όριο έντασης T (το 20% του εικονοστοιχείου).
3. Θεωρούμε έναν κύκλο Bresenham (midpoint circle algorithm) ακτίνας 3 και 16 εικονοστοιχείων που θα περικλείει το εικονοστοιχείο «p»

⁸ Και οι δύο έννοιες –ανίχνευση σημείων ενδιαφέροντος και ανίχνευση ακμών- χρησιμοποιούνται στη θεωρία ως ταυτόσημες.

4. Τα «N» γειτονικά εικονοστοιχεία από τα 16, πρέπει να είναι κατά I_p πάνω ή κάτω από την τιμή T , εάν το εικονοστοιχείο πρέπει να ανιχνευτεί ως σημείο ενδιαφέροντος (στο παράδειγμα της εικόνας 2.12 έχει χρησιμοποιηθεί $N=12$).
5. Για να γίνει γρήγορος ο αλγόριθμος, πρώτα συγκρίνουμε την ένταση των εικονοστοιχείων 1,5,9,13 του κύκλου με το I_p . Τουλάχιστον τρία από αυτά τα τέσσερα εικονοστοιχεία πρέπει να ικανοποιούν το κριτήριο του ορίου, ώστε να υφίσταται σημείο ενδιαφέροντος (εικόνα 2.12).
6. Εάν τουλάχιστον τρεις από τις τέσσερις τιμές των εικονοστοιχείων-11 ,15 ,19 ,13- δεν είναι πάνω ή κάτω από $I_p + T$, τότε το «p» δεν είναι σημείο ενδιαφέροντος (ακμή). Σε αυτήν την περίπτωση απορρίπτουμε το εικονοστοιχείο «p», ως πιθανό σημείο ενδιαφέροντος. Αντίθετα, εάν τουλάχιστον τρία από τα εικονοστοιχεία είναι πάνω ή κάτω από $I_p + T$, τότε ελέγχουμε τα 16 εικονοστοιχεία και, επιπλέον, ελέγχουμε εάν τα 12 γειτονικά εικονοστοιχεία εμπίπτουν στο κριτήριο.
7. Επαναλαμβάνουμε τη διαδικασία για όλα τα εικονοστοιχεία της εικόνας.

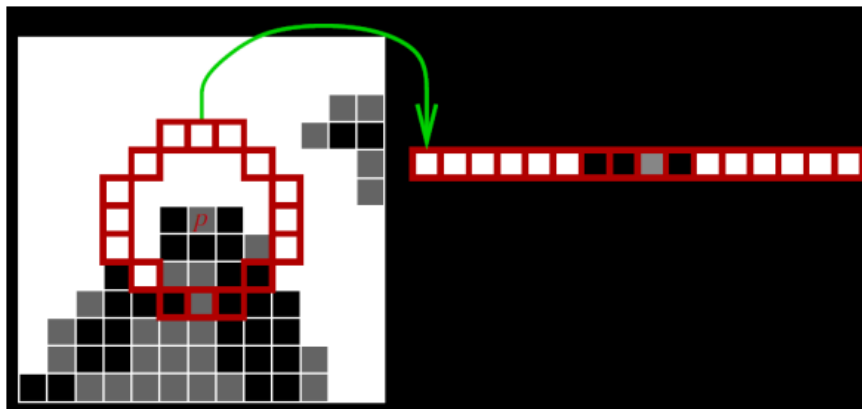
Ο αλγόριθμος παρουσιάζει μερικούς περιορισμούς. Αρχικά, για $N < 12$, ο αλγόριθμος δεν λειτουργεί πάντα αποτελεσματικά, επειδή όταν $N < 12$, ο αριθμός των σημείων ενδιαφέροντος που ανιχνεύονται είναι πολύ υψηλός. Επιπλέον, η σειρά στην οποία μπαίνουν τα 16 εικονοστοιχεία, καθορίζει την ταχύτητα του αλγορίθμου. Μια προσέγγιση μηχανικής εκμάθησης (machine learning approach) έχει προστεθεί στον αλγόριθμο για την αντιμετώπιση αυτών των προβλημάτων.



Εικόνα 2.12: Το ελεγχόμενο σημείο ενδιαφέροντος και τα 16 εικονοστοιχεία εντός του κύκλου.

2.3.2 Προσέγγιση μηχανικής εκμάθησης

1. Επιλέγουμε ένα πακέτο εικόνων για εκμάθηση.
2. Σε κάθε εικόνα «τρέχουμε» τον αλγόριθμο FAST, για να ανιχνεύσουμε τα σημεία ενδιαφέροντος, επιλέγοντας ένα εικονοστοιχείο κάθε φορά και αξιολογώντας τα 16 εικονοστοιχεία του κύκλου.
3. Για κάθε εικονοστοιχείο «p», αποθηκεύουμε τα 16 γειτονικά του εικονοστοιχεία, ως διάνυσμα.(εικόνα 2.13)



Εικόνα 2.13: Οι 16 τιμές που περικλείουν το εικονοστοιχείο «p» αποθηκευμένες σε μορφή διανύσματος

4. Επαναλαμβάνουμε τη διαδικασία για όλα τα εικονοστοιχεία σε όλες τις εικόνες. Αυτό είναι το διάνυσμα P, το οποίο περιέχει τα δεδομένα για την εκμάθηση. (το p αναφέρεται στο εικονοστοιχείο και το P στο διάνυσμα)
5. Κάθε τιμή (π.χ. τιμή x για κάθε ένα από τα 16 εικονοστοιχεία) στο διάνυσμα, μπορεί να λάβει τρεις καταστάσεις. Σκοτεινότερη, φωτεινότερη ή παρόμοια με το «p».

Μαθηματικά:

$$S_{p \rightarrow x} = \begin{cases} d, & I_{p \rightarrow x} \leq I_p - t & \text{(darker)} \\ s, & I_p - t < I_{p \rightarrow x} < I_p + t & \text{(similar)} \\ b, & I_p + t \leq I_{p \rightarrow x} & \text{(brighter)} \end{cases} \quad (2.6)$$

Όπου $S_{p \rightarrow x}$ είναι η κατάσταση, $I_{p \rightarrow x}$ είναι η ένταση του εικονοστοιχείου x και t είναι ένα όριο.

6. Βάσει των καταστάσεων ολόκληρο το διάνυμα P , θα υποδιαιρεθεί σε τρία υποσύνολα, P_d , P_s , P_b .
7. Ορίζουμε μια μεταβλητή K_p , η οποία είναι αληθής εάν το « p » είναι σημείο ενδιαφέροντος και ψευδής εάν το « p » δεν είναι σημείο ενδιαφέροντος.
8. Χρησιμοποιούμε τον αλγόριθμο ID3 (δέντρο αποφάσεων κλάσης) για να εξετάσουμε κάθε υποσύνολο χρησιμοποιώντας την μεταβλητή K_p , ώστε να μάθουμε αν το σημείο είναι γωνία
9. Ο αλγόριθμος ID3 λειτουργεί με την αρχή της ελαχιστοποίησης της εντροπίας. Εξετάζουμε τα 16 εικονοστοιχεία, με τέτοιο τρόπο, ώστε να βρούμε εάν είναι σημείο ενδιαφέροντος με τον ελάχιστο αριθμό αναζητήσεων. Δηλαδή, επιλέγουμε το εικονοστοιχείο x , το οποίο έχει την περισσότερη πληροφορία για το εικονοστοιχείο « p ». Η εντροπία για το σύνολο P μπορεί να παρουσιαστεί μαθηματικά ως:

$$H(P) = (c + \bar{c}) \log_2(c + \bar{c}) - c \log_2 c - \bar{c} \log_2 \bar{c}$$

where $c = |\{p|K_p \text{ is true}\}|$ (number of corners)

and $\bar{c} = |\{p|K_p \text{ is false}\}|$ (number of non corners)

(2.7)

10. Αναδρομικά εφαρμόζουμε αυτήν την ελαχιστοποίηση της εντροπίας στα τρία υποσύνολα.
11. Τερματίζουμε τη διαδικασία, όταν η εντροπία ενός υποσυνόλου είναι μηδενική.
12. Η εντολή επερώτησης(querying order), η οποία μαθαίνεται από το δέντρο απόφασης, μπορεί να χρησιμοποιηθεί για ταχύτερη ανίχνευση και σε άλλες εικόνες.

2.3.3 Τεχνική μη μέγιστης κατάπνιξης (Non maximum suppression) για την αφαίρεση γειτονικών ακμών

Η ανίχνευση πολλαπλών σημείων ενδιαφέροντος γειτονικών το ένα με το άλλο, είναι ένα από τα άλλα προβλήματα της αρχικής έκδοσης του αλγορίθμου. Αυτό μπορεί να αντιμετωπιστεί εφαρμόζοντας τεχνική της μη μέγιστης κατάπνιξης μετά την ανίχνευση των σημείων ενδιαφέροντος.

1. Υπολογίζουμε τα αποτελέσματα της συνάρτησης (score function) V για κάθε ένα από τα ανιχνευμένα σημεία. Τα αποτελέσματα της συνάρτησης (score function) V ορίζονται ως το άθροισμα της απόλυτης διαφοράς μεταξύ των εικονοστοιχείων εντός του γειτονικού τόξου και του κεντρικού εικονοστοιχείου.
2. Θεωρούμε δύο γειτονικά σημεία ενδιαφέροντος και συγκρίνουμε τις τιμές τους V .
3. Απορρίπτουμε το ένα με τη χαμηλότερη τιμή V .

Η όλη διαδικασία μπορεί να συνοψιστεί μαθηματικά ως εξής:

$$V = \max \begin{cases} \sum (\text{pixel values} - p) & \text{if } (\text{value} - p) > t \\ \sum (p - \text{pixel values}) & \text{if } (p - \text{value}) > t \end{cases} \quad (2.8)$$

Όπου, p είναι το κεντρικό εικονοστοιχείο, t είναι το όριο για την ανίχνευση και οι τιμές των εικονοστοιχείων (pixel values) αντιστοιχούν στα N γειτονικά εικονοστοιχεία του κύκλου.

Τα αποτελέσματα της συνάρτησης μπορούν να οριστούν και με άλλους τρόπους. Το σημείο κλειδί, εδώ, είναι να ορίσουμε μια ευρετική συνάρτηση η οποία μπορεί να συγκρίνει δύο γειτονικές ανιχνευμένες ακμές και να απορρίπτει την, συγκριτικά, μη σημαντική.



Εικόνα 2.14 : Μια εικόνα με ανιχνευμένα σημεία ενδιαφέροντος. Οι πράσινες τελείες δείχνουν τις μη μέγιστες συγκαλυμμένες ακμές.

2.4 Ανιχνευτής και Περιγραφέας Ταχύτερων Υπολογιζομένων Εύρωστων Χαρακτηριστικών (SURF-Speeded Up Robust Features)

Ο ανιχνευτής και περιγραφέας SURF δημιουργήθηκε από τους Herbert Bay , Andreas Ess, Tinne Tuytelaars και Luc Van Gool το 2006. Χρησιμοποιείται ευρέως σε ποικίλες εφαρμογές του τομέα της υπολογιστικής όρασης όπως η αναγνώριση αντικειμένων, ανθρώπων ή πρόσωπα, να κατασκευάζει 3D σκηνές και να εξαγάγει σημεία ενδιαφέροντος. Ο SURF έχει βασιστεί στα πιο δυνατά σημεία των σημαντικότερων ανιχνευτών και περιγραφέων πριν από αυτόν. Εν προκειμένω χρησιμοποιεί τον Γρήγορο Εσσιανό πίνακα(Fast Hessian matrix) για τον ανιχνευτή και ένα περιγραφέα που βασίζεται στην κατανομή. Αυτό οδηγεί σε ένα συνδυασμό καινοτόμου ανίχνευσης(detection), περιγραφής(description) και ταιριαστών βημάτων (matching steps) για την αντιστοιχία σημείων. Συνοπτικά τα βήματα του αλγορίθμου, τα οποία θα αναλυθούν και παρακάτω εκτενέστερα είναι τα εξής:

1. Τα σημεία ενδιαφέροντος επιλέγονται σε ξεχωριστές τοποθεσίες στην εικόνα, όπως γωνίες, άμορφα σημεία, διασταυρώσεις. Η πιο πολύτιμη ιδιότητα ενός ανιχνευτή σημείων ενδιαφέροντος είναι η επαναληψιμότητα του, δηλαδή εάν βρίσκει ακριβώς τα ίδια σημεία ενδιαφέροντος, ανεξαρτήτου οπτικής γωνίας. Για την αναγνώριση σημείων ενδιαφέροντος ο SURF χρησιμοποιεί μία ακέραια προσέγγιση της ορίζουσας του Εσσιανού(Hessian) πίνακα ανίχνευσης και την συνδυάζει με την χρήση εικόνων ολοκληρώματος(integral images), ώστε να μειωθεί ο χρόνος υπολογισμού
2. Στη συνέχεια, αναγκαία είναι η τοποθέτηση προσανατολισμού. Αυτό επιτυγχάνεται με τον υπολογισμό των αποκρίσεων των κυμάτων(Haar). Η ευρύτερη περίμετρος (γειτονιά) κάθε σημείου ενδιαφέροντος αναπαριστάται από ένα χαρακτηριστικό διάνυσμα (feature vector). Ο περιγραφέας πρέπει να είναι διακριτικός και ανθεκτικός στον «θόρυβο» της φωτογραφίας, σε σφάλματα εντοπισμού και γεωμετρικές ή φωτομετρικές παραμορφώσεις.

3. Τέλος, η ολοκλήρωση του αλγορίθμου SURF είναι η δημιουργία του περιγραφέα. Αυτό υλοποιείται με την κατασκευή τετράγωνης περιοχής, επικεντρωμένης γύρω από το σημείο ενδιαφέροντος. Τα διανύσματα του περιγραφέα αντιστοιχίζονται εν μέσω διαφορετικών εικόνων. Το «ταίριασμα» ή αλλιώς αντιστοιχία, συχνά, βασίζεται σε μια απόσταση μεταξύ των διανυσμάτων, όπως η απόσταση Mahalanobis ή την Ευκλείδεια απόσταση. Η διάσταση του περιγραφέα έχει άμεσο αντίκτυπο στο χρόνο που κάνει και συνεπώς, ένας χαμηλότερος αριθμός διαστάσεων είναι επιθυμητός.
4. Όταν δουλεύουμε με τοπικά χαρακτηριστικά, ένα ζήτημα που πρέπει να διευθετηθεί είναι ο απαιτούμενος βαθμός μη μεταβλητότητας σε σχέση με την κλίμακα και την περιστροφή της εικόνας. Προφανώς, αυτό εξαρτάται από τις γεωμετρικές και φωτομετρικές παραμορφώσεις, οι οποίες, με τη σειρά τους, καθορίζονται από τις διαφορετικές οπτικές (changes in viewing conditions). Ο SURF εστιάζει σε αμετάβλητους ανιχνευτές και περιγραφείς ως προς την κλίμακα και την περιστροφή της εικόνας.

Ο SURF για να μειώσει τον χρόνο υπολογισμού χρησιμοποιεί τις εικόνες ολοκληρώματος (integral images). Το ολοκλήρωμα της εικόνας S πρόκειται για έναν πίνακα ίδιας διάστασης με την εικόνα εισόδου I . Η είσοδος μίας εικόνας ολοκληρώματος $S(x,y)$ σε μία τοποθεσία (x,y) αντιπροσωπεύει το άθροισμα όλων των εικονοστοιχείων στην εικόνα εισόδου I μέσα σε μία ορθογώνια σχηματισμένη περιοχή από την αρχή και το εικονοστοιχείο (x,y) . Η σχέση που περιλαμβάνει τα παραπάνω είναι η εξής :

$$S(x,y) = \sum_{i=0}^x \sum_{j=0}^y I(i,j) \quad (2.10)$$

Με αυτό τον τρόπο ο χρόνος υπολογισμού είναι ανεξάρτητος από το μέγεθος της περιοχής. Αυτό έχει ως αποτέλεσμα τα τετραγωνικά φίλτρα μεγάλου μεγέθους που χρησιμοποιεί ο SURF να μην επηρεάζουν την ταχύτητα και να απαιτούνται μόνο τέσσερεις αξιολογήσεις στην γωνία του ορθογωνίου.

2.4.1 Αναγνώριση σημείων ενδιαφέροντος βασισμένα στον Εσσιανό πίνακα

(Fast-Hessian Detector)

Ο SURF χρησιμοποιεί έναν ανιχνευτή άμορφων σημείων(blob) βασισμένο στον Εσσιανό (Hessian) πίνακα για να βρει τα σημεία ενδιαφέροντος. Η ορίζουσα του Εσσιανού πίνακα χρησιμοποιείται ως μέτρο της τοπικής αλλαγής γύρω από το σημείο και τα σημεία που επιλέγονται είναι εκείνα στα οποία η ορίζουσα είναι μέγιστη. Ο SURF χρησιμοποιεί την ορίζουσα του Εσσιανού πίνακα και για την επιλογή της κλίμακας. Λαμβάνοντας υπόψη ένα σημείο $p = (x, y)$ σε μια εικόνα I , ο Εσσιανός πίνακας $H(p, \sigma)$, όπου p το σημείο και σ η κλίμακα, ορίζεται ως εξής:

$$H(p, \sigma) = \begin{pmatrix} L_{xx}(p, \sigma) & L_{xy}(p, \sigma) \\ L_{xy}(p, \sigma) & L_{yy}(p, \sigma) \end{pmatrix} \quad (2.11)$$

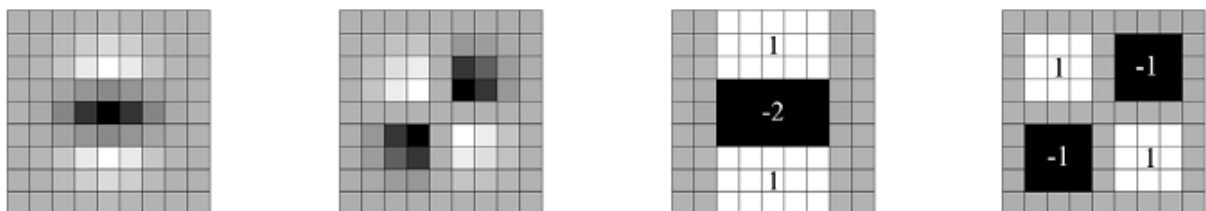
Όπου $L_{xx}(p, \sigma), L_{xy}(p, \sigma), L_{yy}(p, \sigma)$ είναι οι παράγωγοι 2^{ης} τάξης της εικόνας κλίμακας γκρι (grayscale image). Οι Gaussians είναι βέλτιστες για την ανάλυση της κλίμακας χώρου. Στην πράξη όμως χρειάζονται να διακριτοποιούνται και να περικόπτονται (εικόνα 2.15 αριστερό ήμισυ). Εφόσον τα Gaussian φίλτρα είναι μη-ιδανικά σε κάθε περίπτωση, πιέζεται ακόμη περισσότερο η προσέγγιση με κουτιά φίλτρων (εικόνα 2.15 δεξιό ήμισυ). Αυτές οι κατά προσέγγιση 2^{ης} τάξης Gaussian παράγωγοι, μπορούν να αξιολογηθούν πολύ γρήγορα χρησιμοποιώντας ολοκληρώματα εικόνας ανεξάρτητα από το μέγεθος.

Το φίλτρο κουτί μεγέθους 9×9 στην εικόνα 2.15 είναι μια προσέγγιση ενός Gaussian με $\sigma = 1,2$ και αντιπροσωπεύει τη χαμηλότερη κλίμακα (υψηλότερη χωρική ανάλυση). Οι προσεγγίσεις έχουν δηλωθεί ως D_{xx}, D_{xy} και D_{yy} . Τα βάρη που εφαρμόζονται στις ορθογώνιες περιοχές διατηρούνται απλά ώστε να υπάρχει υπολογιστική αποδοτικότητα. Παρόλα αυτά, είναι αναγκαίο να ισορροπήσουν τα σχετικά βάρη με τον Εσσιανό πίνακα με

$$\frac{|L_{xy}(1.2)|_F |D_{xx}(9)|_F}{|L_{xx}(1.2)|_F |D_{xy}(9)|_F} = 0,912.. \cong 0,9 \quad (2.12)$$

όπου $|x|_F$ είναι η Frobenius νόρμα. Αυτό οδηγεί στην σχέση: $\det(H_{\text{approx}}) = D_{xx}D_{yy} - (0.9D_{xy})^2$. Επιπλέον, οι αποκρίσεις του φίλτρου είναι κανονικοποιημένες ως προς την αλλαγή κλίμακας σε σχέση με το μέγεθος της μάσκας. Αυτό καθορίζει μία σταθερή νόρμα Frobenius για κάθε μέγεθος του φίλτρου. Οι κλίμακες χώρου υλοποιούνται συνήθως ως πυραμίδες εικόνας. Οι εικόνες είναι επανειλημμένα εξομαλυμένες με Gaussian και υπο-δειγματοληπτικές ώστε να

επιτευχθεί υψηλότερο επίπεδο της πυραμίδας. Λόγω της χρήσης φίλτρων σε συνδυασμό με τα ολοκληρώματα εικόνων, δεν είναι αναγκαίο να εφαρμόζεται επαναληπτικά το ίδιο φίλτρο στην έξοδο του προηγούμενου φιλτραρισμένου στρώματος αλλά να εφαρμόζονται φίλτρα οποιοδήποτε μεγέθους με την ίδια ταχύτητα κατεύθυνση στην αρχική εικόνα. Επομένως, η κλίμακα χώρου αναλύεται με κλιμακωτό μέγεθος φίλτρου παρά με την επαναληπτική μείωση του μεγέθους της εικόνας. Στον SURF, το χαμηλότερο επίπεδο του χώρου κλίμακας λαμβάνεται από την έξοδο 9×9 φίλτρου με κλίμακα $\sigma=1,2$ (που αντιστοιχεί σε Gaussian παραγώγους με $\sigma = 1,2$). Τα ακόλουθα στρώματα λαμβάνονται με φιλτράρισμα της εικόνας με σταδιακά μεγαλύτερες μάσκες λαμβάνοντας υπόψη τη διακριτή φύση των εικόνων ολοκληρώματος και τη συγκεκριμένη δομή των φίλτρων. Συγκεκριμένα, αυτό έχει ως αποτέλεσμα σε φίλτρα μεγέθους 9×9 , 15×15 , 21×21 , 27×27 , και σε μεγαλύτερες κλίμακες, το βήμα μεταξύ των διαδοχικών μεγεθών του φίλτρου να αναβαθμίζεται αναλόγως. Για κάθε νέα οκτάβα η αύξηση του μεγέθους του φίλτρου διπλασιάζεται (πηγαίνοντας από 6 σε 12 και στη συνέχεια 24). Ταυτόχρονα, τα διαστήματα δειγματοληψίας για την εξαγωγή των σημείων ενδιαφέροντος μπορεί να διπλασιαστούν. Δεδομένου ότι οι αναλογίες της διάταξης του φίλτρου παραμένουν σταθερές μετά την κλιμάκωση, οι κατά προσέγγιση παράγωγοι Gaussian κλιμακώνονται αναλόγως. Έτσι, για παράδειγμα, το φίλτρο 27×27 αντιστοιχεί $\sigma = 3 \times 1,2 = 3,6 = s$. Προκειμένου να εντοπισθούν τα σημεία ενδιαφέροντος στην εικόνα, εφαρμόζεται η τεχνική της μη μέγιστης κατάπνιξης (non maximum suppression) σε μια γειτονιά $3 \times 3 \times 3$. Στη συνέχεια τα μέγιστα της ορίζουσας του Εσσιανού πίνακα παρεμβάλλονται στη κλίμακα και στο χώρο της εικόνας με την μέθοδο που προτείνει ο Brown. Η παρεμβολή της κλίμακας χώρου είναι ιδιαίτερα σημαντική στην περίπτωση μας. Διότι η διαφορά της κλίμακας μεταξύ των πρώτων στρωμάτων κάθε οκτάβας είναι σχετικά μεγάλη. Η εικόνα 2.16(αριστερά) δείχνει ένα παράδειγμα με εντοπισμένα σημεία ενδιαφέροντος χρησιμοποιώντας τον ανιχνευτή (Fast-Hessian).

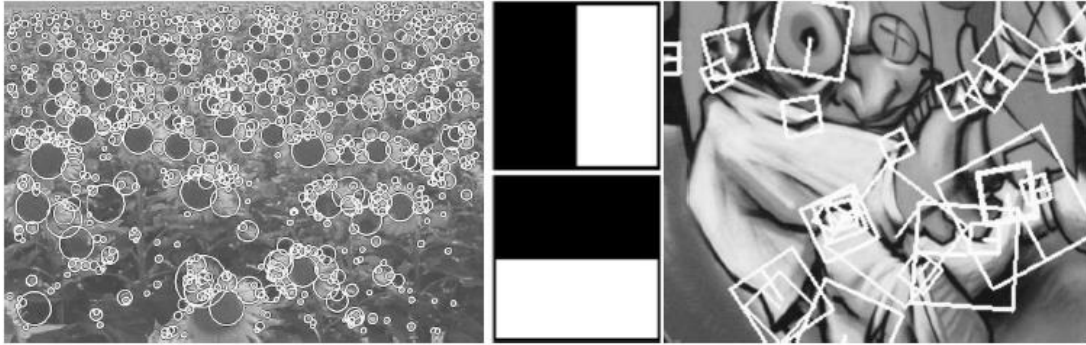


Εικόνα 2.15 : Από αριστερά προς τα δεξιά: η διακριτοποίηση και η περικοπή της Gaussian 2ης τάξης μερικών παραγώγων στην ψ και $\chi\psi$ κατεύθυνση, και οι προσεγγίσεις χρησιμοποιώντας κουτιά φίλτρα. Οι γκριζες περιοχές είναι ίσες με το μηδέν.

2.4.2 Τοποθέτηση προσανατολισμού

Ο περιγραφέας SURF βασίζεται σε κοινές ιδιότητες με έναν επίσης πολύ γνωστό περιγραφέα τον SIFT, όμως με πολύ λιγότερη πολυπλοκότητα. Το πρώτο βήμα αποτελείται από τη δημιουργία ενός προσανατολισμού, βασισμένου σε πληροφορία από μια κυκλική περιοχή γύρω από το σημείο ενδιαφέροντος. Τότε, κατασκευάζεται μια τετράγωνη περιοχή ευθυγραμμισμένη με τον επιλεγμένο προσανατολισμό και ο περιγραφέας SURF αποσπάται από αυτήν. Προκειμένου να μην μεταβάλλεται στην περιστροφή, αναγνωρίζουμε έναν προσανατολισμό για τα σημεία ενδιαφέροντος. Για αυτό το σκοπό, πρώτα υπολογίζουμε τις αποκρίσεις κύματος Haar στις διευθύνσεις χ και ψ , όπως φαίνεται στην εικόνα 2.16, και αυτό σε μια κυκλική γειτονιά ακτίνας $6s$ γύρω από το σημείο ενδιαφέροντος, όπου s είναι η κλίμακα του σημείου ενδιαφέροντος, όπως εντοπίστηκε. Επιπλέον, το βήμα δειγματοληψίας είναι κλίμακα εξαρτώμενη και επιλεγμένη να είναι s . Τέλος, οι αποκρίσεις του κύματος (wavelet) υπολογίζονται στην τρέχουσα κλίμακα s . Ανάλογα, σε υψηλές κλίμακες το μέγεθος των κυμάτων είναι μεγάλο. Συνεπώς, χρησιμοποιούμε ξανά ακέραιες φωτογραφίες για γρήγορο φιλτράρισμα. Χρειάζονται μόνο έξι χειρισμοί για να υπολογίσουμε την απόκριση στις χ και ψ διευθύνσεις, σε κάθε κλίμακα. Το μήκος της πλευράς του κύματος είναι $4s$.

Όταν οι αποκρίσεις των κυμάτων υπολογιστούν και επιβαρυνθούν με ένα Gaussian ($\sigma = 2.5s$) επικεντρωμένο στο σημείο ενδιαφέροντος, οι αποκρίσεις εμφανίζονται ως διανύσματα σε ένα χώρο, με οριζόντια δύναμη απόκρισης κατά μήκος της τετμημένης και την κατακόρυφη δύναμη απόκρισης κατά μήκος της τεταγμένης. Ο επικρατέστερος προσανατολισμός υπολογίζεται από τον υπολογισμό του αθροίσματος όλων των αποκρίσεων μέσα σε ένα περιστρεφόμενο κυκλικό τμήμα που καλύπτει γωνία $\pi/3$ γύρω από το σημείο ενδιαφέροντος. Σε κάθε θέση που δείχνει αυτό το κυκλικό τμήμα, οι αποκρίσεις στους άξονες x και y αθροίζονται και φτιάχνουν ένα νέο διάνυσμα. Το μεγαλύτερο από αυτά τα διανύσματα, καθορίζει και τον κυρίαρχο προσανατολισμό.



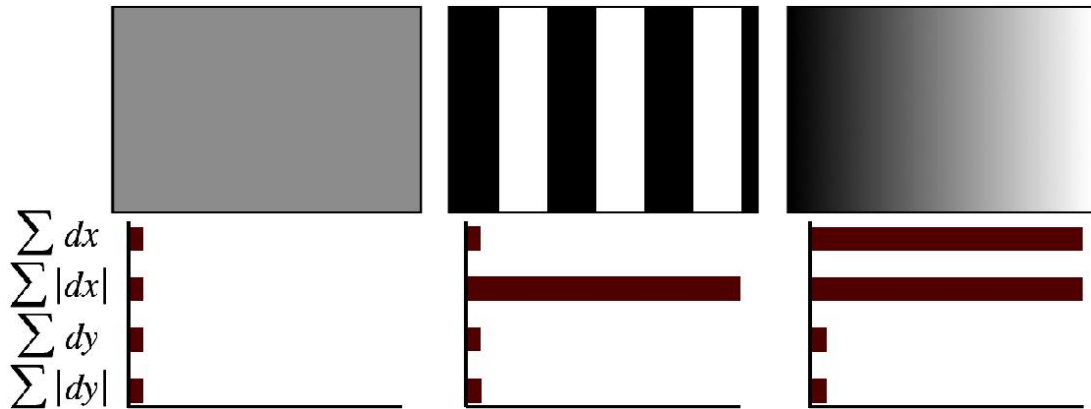
Εικόνα 2.16 : Αριστερά: Εντοπισμένα σημεία ενδιαφέροντος για ένα λιβάδι με ηλιοτρόπια. Αυτό το είδος εικόνας δείχνει ξεκάθαρα τη φύση των χαρακτηριστικών των ανιχνευτών που βασίζονται στο πίνακα Hessian. Στο μέσο: Τα είδη κύματος Haar που χρησιμοποιούνται στον περιγρ περιγραφέα SURF. Δεξιά: Λεπτομέρεια εικόνας ενός γκράφιτι, που δείχνει το μέγεθος του παραθύρου του περιγραφέα, σε διαφορετικές κλίμακες.

2.4.3 Τα στοιχεία του περιγραφέα

Για την ολοκλήρωση του περιγραφέα, το πρώτο βήμα αποτελείται από την κατασκευή μιας τετράγωνης περιοχής επικεντρωμένης στο σημείο ενδιαφέροντος και προσανατολισμένης ανάλογα τον προσανατολισμό που επιλέχτηκε προηγουμένως. Το μέγεθος αυτού του παραθύρου είναι $20s$ (βλέπε παράδειγμα στην εικόνα 2.17). Η περιοχή χωρίζεται σε μικρότερες τετράγωνες υπο-περιοχές μεγέθους 4×4 . Αυτό διατηρεί σημαντική χωρική πληροφορία. Για κάθε υπο-περιοχή, υπολογίζουμε μερικά απλά χαρακτηριστικά σε δείγματα σημείων μεγέθους 5×5 . Πιο απλά, αποκαλούμε dx την απόκριση του κύματος Haar στην οριζόντια διεύθυνση και dy την απόκριση του κύματος Haar στην κατακόρυφη διεύθυνση. Εδώ, τα «οριζόντια» και «κάθετα» καθορίζονται σε σχέση με τον επιλεγμένο προσανατολισμό του σημείου ενδιαφέροντος. Για την αύξηση της στιβαρότητας ως προς τις γεωμετρικές παραμορφώσεις και τα σφάλματα εντοπισμού, οι αποκρίσεις dx και dy πρώτα με μια Gaussian με $\sigma = 3.3s$ επικεντρωμένη και πάλι στο σημείο ενδιαφέροντος.

Τότε, οι αποκρίσεις των κυμάτων dx και dy αθροίζονται πάνω σε κάθε υπο-περιοχή και σχηματίζουν το πρώτο σύνολο καταχωρήσεων στο διάνυσμα του περιγραφέα. Προκειμένου να εφαρμόσουμε την πληροφορία για πολικότητα των αλλαγών στην ένταση, εξάγουμε το άθροισμα των απόλυτων τιμών των αποκρίσεων, $|dx|$ και $|dy|$. Συνεπώς, κάθε υπο-περιοχή έχει ένα διάνυσμα-περιγραφέα v τεσσάρων διαστάσεων για την υποκείμενη του δομή $v = (\sum dx, \sum dy, \sum |dx|, \sum |dy|)$. Αυτό έχει ως αποτέλεσμα ένα διάνυσμα περιγραφέα για όλες τις υπό-περιοχές που είναι 4×4 με μήκος 64. Η διαδικασία αυτή, παραμένει αναλλοίωτη στην περιστροφή, αλλαγή

κλίμακας, φωτεινότητας ακόμα και μετά την κανονικοποίηση. Η εικόνα 2.17 δείχνει τις ιδιότητες του περιγραφέα για τρία διαφορετικά μοτίβα έντασης εικόνων εντός μιας υπο-περιοχής. Κάποιος μπορεί να φανταστεί συνδυασμούς τέτοιων τοπικών μοτίβων έντασης από έναν χαρακτηριστικό περιγραφέα.



Εικόνα 2.17 : Οι καταχωρήσεις του περιγραφέα όσον αφορά μια υπο-περιοχή αναπαριστούν τη φύση ενός υποκείμενου μοτίβου έντασης. Αριστερά: Σε περίπτωση ομοιογενούς περιοχής, όλες οι τιμές είναι σχετικά χαμηλές. Στο μέσο: Σε παρουσία συχνοτήτων στην διεύθυνση x, η τιμή του $\sum |d_x|$ είναι υψηλή, αλλά οι υπόλοιπες παραμένουν χαμηλές. Εάν η ένταση, βαθμιαία, αυξάνεται στην διεύθυνση x, τότε και οι δύο τιμές $\sum d_x$ και $\sum |d_x|$ είναι υψηλές.

2.5 Περιγραφέας Δυναδικών Εύρωστων και Αμετάβλητων Σημείων Ενδιαφέροντος (BRISK - Binary Robust Invariant Scalable Keypoints)

Ο περιγραφέας BRISK δημοσιεύτηκε από τους Stefan Leutenegger, Margarita Chli και Roland Y. Siegwart το 2011. Η αποτελεσματική και αποδοτική παραγωγή σημείων ενδιαφέροντος από μία φωτογραφία είναι ένα καλά μελετημένο πρόβλημα το οποίο αποτελεί βάση πολλών εφαρμογών στην όραση υπολογιστών (computer vision). Ο BRISK είναι ένας υψηλού επιπέδου ποιότητας οπτικός περιγραφέας με χαμηλό υπολογιστικό κόστος δηλαδή τον χαρακτηρίζει η υψηλή ταχύτητα. Η αποσύνθεση μιας εικόνας σε τοπικές περιοχές ενδιαφέροντος ή «χαρακτηριστικά» είναι μια ευρέως εφαρμοσμένη τεχνική υπολογιστικής όρασης που χρησιμοποιείται για να μειώσει την πολυπλοκότητα, αξιοποιώντας παράλληλα τις τοπικές ιδιότητες εμφάνισης. Ο ανιχνευτής σημείων ενδιαφέροντος λειτουργεί ομαλά σε κάθε μετατροπή της εικόνας και βρίσκει τις σημαντικότερες περιοχές της εικόνας οι οποίες ανιχνεύονται επαναληπτικά από κάθε οπτική γωνία. Επίσης αιχμαλωτίζει τις πιο σημαντικές και ξεχωριστές πληροφορίες που περιέχονται στις πιο εμφανείς περιοχές που έχουν ανιχνευτεί έτσι ώστε η ίδια δομή να αναγνωριστεί αν εντοπιστεί. Επιπλέον, πάνω από την εκπλήρωση αυτών των ιδιοτήτων για να επιτευχθεί η επιθυμητή ποιότητα των σημείων ενδιαφέροντος, η ταχύτητα της ανίχνευσης και της περιγραφής πρέπει επίσης να βελτιστοποιηθεί και να προσαρμοστεί με τους περιορισμούς του χρόνου.

Τα βασικά στάδια του BRISK είναι η δυνατότητα ανίχνευσης χαρακτηριστικών (feature detection), η σύνθεση του περιγραφέα (descriptor composition) και η δυνατότητα αντιστοίχισης σημείων ενδιαφέροντος (keypoint matching). Ο ανιχνευτής BRISK μπορεί να συνδυαστεί με οποιοδήποτε άλλο περιγραφέα σημείων ενδιαφέροντος και αντίστροφα. Αυτό μας δίνει την δυνατότητα να έχουμε καλύτερες επιδόσεις για τις ενέργειες που θέλουμε να πραγματοποιήσουμε. Τα στάδια ανίχνευσης των χαρακτηριστικών ή αλλιώς των σημείων ενδιαφέροντος είναι :

1. Δημιουργία χωροκλίμακας
2. Υπολογισμός του μέτρου Fast προβάλλοντας το στην χωροκλίμακα (εικόνα 2.18)
3. Το επίπεδο των εικονοστοιχείων υποβάλλεται στην τεχνική μη μέγιστης κατάπνιξης (non-maximal suppression)

4. Υπολογισμός θέσης υπο-εικονοστοιχείου(αυτό το εικονοστοιχείο είναι εικονικό) επεξεργασμένο πριν η παραβολή 1D τοποθετηθεί κατά μήκος της κλίμακας άξονα
5. Προσδιορισμός του πραγματικού μεγέθους του σημείου ενδιαφέροντος

Τα στάδια περιγραφής των σημείων ενδιαφέροντος είναι:

1. Δημιουργία μοτίβου(pattern) γύρω από τα σημεία ενδιαφέροντος (εικόνα 2.19)
2. Διαχωρισμός στα ζεύγη εικονοστοιχείων σε μικρής και μεγάλης απόστασης ζευγάρια
3. Υπολογισμός της τοπικής κλίση μεταξύ των ζευγών μεγάλων αποστάσεων
4. Άθροισμα των κλίσεων για να προσδιοριστεί ο προσανατολισμός των χαρακτηριστικών
5. Περιστροφή στα ζεύγη μικρών αποστάσεων χρησιμοποιώντας προσανατολισμό
6. Κατασκευή δυαδικού περιγραφέα από τα ζεύγη μικρών αποστάσεων που έχουν περιστραφεί

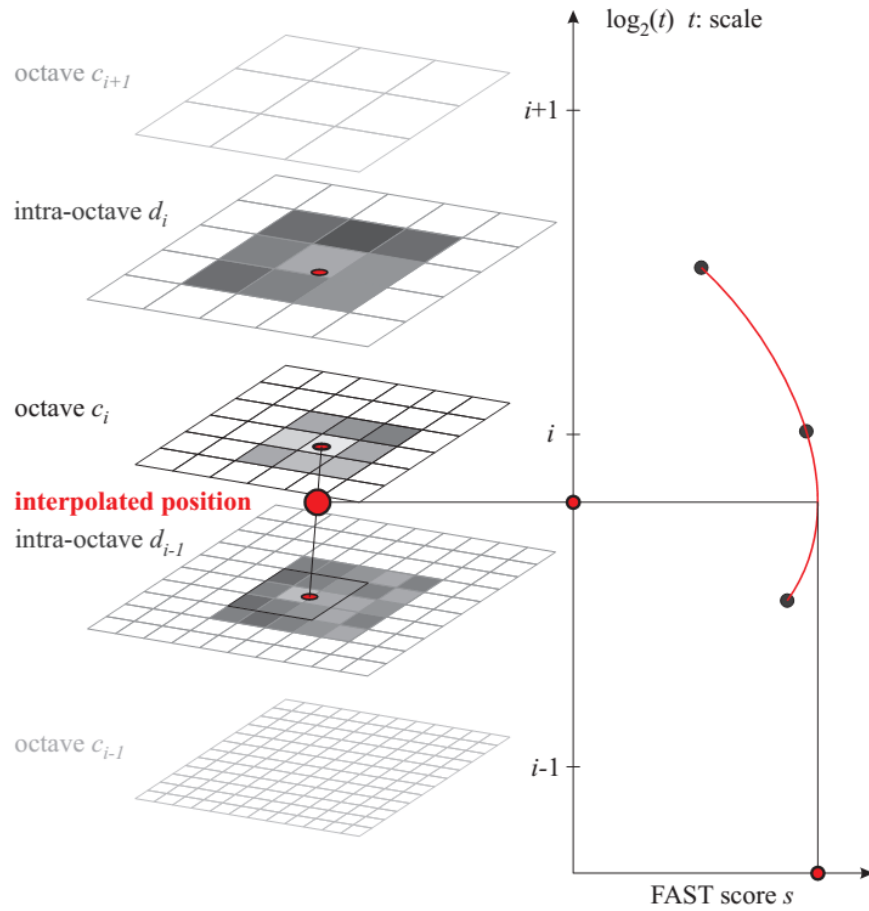
2.5.1 Ανίχνευση σημείων ενδιαφέροντος σε κλίμακα χώρου (Scale-space keypoint detection)

Τα σημεία που μας ενδιαφέρουν προσδιορίζονται μέσα από την εικόνα και τις διαστάσεις της κλίμακας, χρησιμοποιώντας ένα σημαντικό κριτήριο. Προκειμένου να αυξηθεί η αποτελεσματικότητα των υπολογισμών, τα σημεία ενδιαφέροντος ανιχνεύονται σε στρώματα οκτάβας της πυραμίδας εικόνας, καθώς και σε ενδιάμεσα στρώματα. Η θέση και η έκταση κάθε σημείου ενδιαφέροντος λαμβάνεται από ένα συνεχή τομέα δια μέσου κατάλληλης τετραγωνικής συνάρτησης.

Η μεθοδολογία ανίχνευσης του BRISK είναι εμπνευσμένη από τον ανιχνευτή AGAST⁹, ο οποίος είναι ουσιαστικά μία επέκταση του σημερινού δημοφιλή FAST, ο οποίος έχει αποδειχθεί ότι είναι πολύ αποδοτικός για την εξαγωγή χαρακτηριστικών. Με στόχο την επίτευξη αναλλοίωτης κλίμακας η οποία είναι πολύ σημαντική για την εύρεση υψηλής ποιότητας σημείων ενδιαφέροντος, ερευνούν τα μέγιστα όχι μόνο στο επίπεδο της εικόνας αλλά και στο επίπεδο

⁹ E. Mair, G. D. Hager, D. Burschka, M. Suppa, and G. Hirzinger. Adaptive and generic corner detection based on the accelerated segment test. In *Proceedings of the European Conference on Computer Vision (ECCV)*, 2010.

χώρου χρησιμοποιώντας τον FAST ως μέτρο για την σημαντικότητα. Παρά την διακριτοποίηση των αξόνων της κλίμακας σε χονδροειδή διαστήματα σε αντίθεση με άλλους ανιχνευτές υψηλής απόδοσης (πχ. Fast-Hessian) ο ανιχνευτής BRISK υπολογίζει το πραγματικό μέγεθος κάθε σημείου ενδιαφέροντος στη συνεχή κλίμακα χώρου(scale-space). Στο πλαίσιο που θέτει ο BRISK, τα πυραμιδοειδή επίπεδα που αναφέρονται σε κλίμακα χώρου αποτελούνται από n οκτάβες(octaves) c_i και από n ενδο-οκτάβες(inta-octaves) d_i για $i = \{0, 1, \dots, n - 1\}$, τυπικά παίρνουμε για $n=4$. Οι οκτάβες σχηματίζονται με σταδιακή δειγματοληψία της αρχικής εικόνας(που αντιστοιχεί σε c_0). Κάθε ενδο-οκτάβα βρίσκεται μεταξύ των στρωμάτων c_i και c_{i+1} όπως απεικονίζεται στην εικόνα 2.18. Η πρώτη ενδο-οκτάβα d_0 λαμβάνεται με την αρχική μείωση της δειγματοληψίας της αρχικής εικόνας c_0 κατά ένα συντελεστή 1.5, ενώ το υπόλοιπο των στρωμάτων των ενδο-οκτάβων προέρχεται από διαδοχικές δειγματοληψίες στο μισό. Ως εκ τούτου, εάν το t υποδηλώνει κλίμακα τότε ισχύει $t(c_i)=2^i$ και $t(d_i)=2^i * 1.5$. Οι AGAST και FAST έχουν διαφορετικά σχήματα στις μάσκες για την ανίχνευση σημείων κλειδιών. Στον BRISK χρησιμοποιείται περισσότερο η μάσκα 9-16, η οποία απαιτεί κατ 'ουσίαν τουλάχιστον 9 διαδοχικά εικονοστοιχεία στο κύκλο των 16 εικονοστοιχείων (16-pixel circle) ώστε να είναι αρκετά πιο φωτεινά ή σκούρα από το κεντρικό εικονοστοιχείο για την εκπλήρωση του κριτηρίου FAST.



Εικόνα 2.18 : Ανίχνευση σημείων ενδιαφέροντος (keypoints) σε κλίμακα χώρου: Ένα σημείο ενδιαφέροντος (δηλαδή προβολή μέγιστου) αναγνωρίζεται σε οκτάβα c_i με την ανάλυση 8 σημαντικών γειτονικών βαθμολογιών σε c_i , καθώς και στις αντίστοιχες βαθμολογίες-επιδιορθώσεις στα αμέσως-γειτονικά επίπεδα πάνω και κάτω. Και στις τρεις στρώσεις ενδιαφέροντος, η τοπική μέγιστη προεξοχή (local saliency maximum) είναι υπο-εικονοστοιχείο (sub-pixel) επεξεργασμένο πριν η παραβολή 1D τοποθετηθεί κατά μήκος της κλίμακας άξονα για το προσδιορισμό του πραγματικού μεγέθους του σημείου ενδιαφέροντος. Η θέση του σημείου ενδιαφέροντος παρεμβάλλεται ξανά ανάμεσα στα μέγιστα μίας περιοχής (patch) που είναι πιο κοντά προς την καθορισμένη κλίμακα.

2.5.2 Περιγραφή σημείων ενδιαφέροντος (Keypoint description)

Ένα σχέδιο δειγματοληψίας αποτελείται από σημεία που βρίσκονται σε κατάλληλης κλίμακας ομόκεντρους κύκλους και εφαρμόζονται στην ευρύτερη περιφέρεια κάθε σημείου ενδιαφέροντος με σκοπό να ανακτηθούν οι γκρι τιμές. Τέλος το προγραμματισμένο πρότυπο δειγματοληψίας BRISK λαμβάνει τα αποτελέσματα από τις κατά ζεύγη συγκρίσεις της φωτεινότητας που συναρμολογούνται στο δυαδικό BRISK περιγραφέα. Τέλος ο προς κατεύθυνση BRISK δειγματολειτουργείται με σκοπό να παρέχει συγκρίσεις ως προς την φωτεινότητα των εικόνων. Όταν δημιουργηθούν τα σημεία ενδιαφέροντος του BRISK μπορούν να είναι πολύ αποτελεσματικά χάρη στη δυαδική φύση του περιγραφέα.

Δειγματοληψία μοτίβου και εκτίμηση περιστροφής (Sampling Pattern and Rotation Estimation) :

Ο περιγραφέας BRISK χρησιμοποιεί ένα μοτίβο για την δειγματοληψία της γειτονιάς του σημείου ενδιαφέροντος. Το μοτίβο (εικόνα 2.19) ορίζει τις θέσεις N που ισαπέχουν οι ομόκεντροι κύκλοι με το σημείο ενδιαφέροντος. Για να αποφευχθούν οι παραμορφώσεις όταν παίρνουμε τα δείγματα της έντασης της εικόνας σε ένα σημείο p_i του μοτίβου, εφαρμόζεται μία Gaussian εξομάλυνση με τυπική απόκλιση σ_i ανάλογη της απόστασης του αντίστοιχου κύκλου. Μετά την τοποθέτηση και κλιμάκωση του μοτίβου σύμφωνα με το σημείο ενδιαφέροντος k της εικόνας, εξετάζουμε ένα από τα διαθέσιμα $N \cdot (N - 1) / 2$ ζεύγη σημείων δειγματοληψίας έστω ένα (p_i, σ_i) . Οι εξομαλυσμένες τιμές έντασης σε αυτά τα σημεία οι οποίες είναι $I(p_i, \sigma_i)$ και $I(p_j, \sigma_j)$ χρησιμοποιούνται για να εκτιμηθεί η τοπική κλίση $g(p_i, p_j)$ από τον τύπο :

$$g(p_i, p_j) = (p_j - p_i) \cdot \frac{I(p_j, \sigma_j) - I(p_i, \sigma_i)}{\|p_j - p_i\|^2} \quad (2.13)$$

Το σύνολο A από όλα τα ζεύγη σημείων της δειγματοληψίας ορίζεται :

$$A = \{ (p_i, p_j) \in \mathbb{R}^2 \times \mathbb{R}^2 \mid i < N \wedge j < i \wedge i, j \in \mathbb{N} \} \quad (2.14)$$

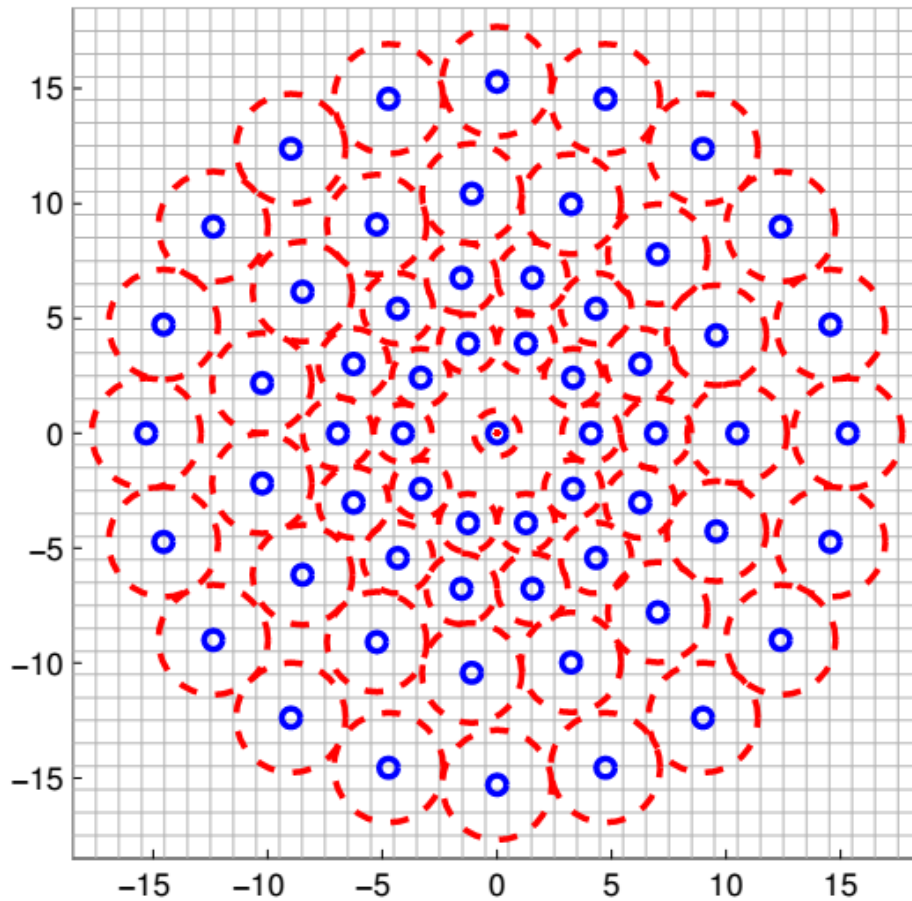
Επίσης ορίζεται ένα υποσύνολο από ζεύγη μικρής απόστασης S και ένα άλλο μεγάλης απόστασης L (εικόνα 2.20) :

$$\begin{aligned} S &= \{ (p_i, p_j) \in A \mid \|p_j - p_i\| < \delta_{max} \} \subseteq A \\ L &= \{ (p_i, p_j) \in A \mid \|p_j - p_i\| > \delta_{min} \} \subseteq A \end{aligned} \quad (2.15)$$

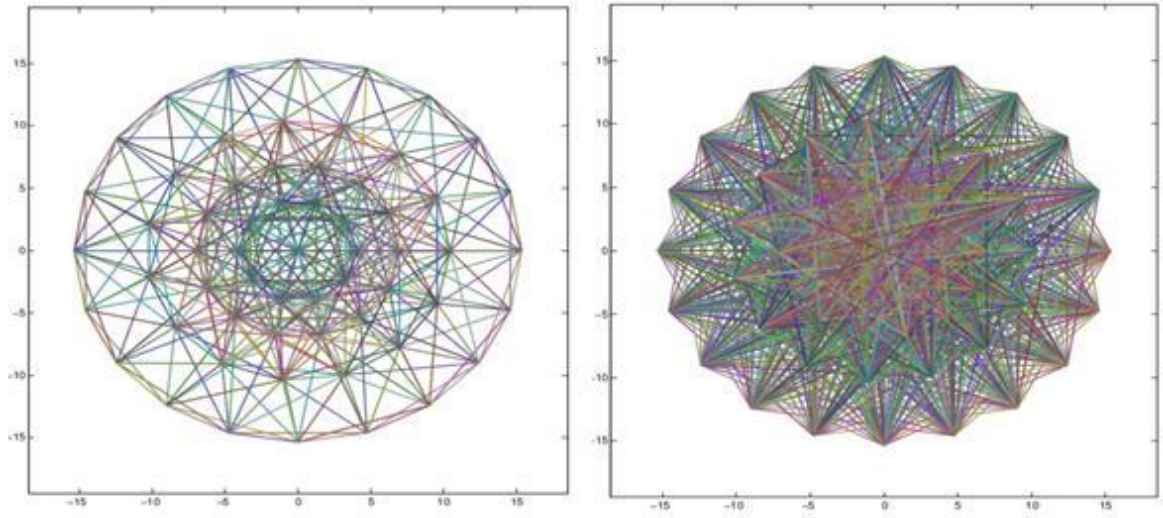
Τα όρια κατωφλίωσης (threshold) καθορίζονται στο $\delta_{\max} = 9,75t$ και στο $\delta_{\min} = 13,67t$ (όπου t είναι η κλίμακα του k). Επαναλαμβάνοντας για τα ζεύγη σημείων του L , εκτιμάται ότι η γενική κατεύθυνση του χαρακτηριστικού μοτίβου για το σημείο ενδιαφέροντος k είναι :

$$g = \begin{pmatrix} g_x \\ g_y \end{pmatrix} = \frac{1}{L} \cdot \sum_{(p_i, p_j) \in L} g(p_i, p_j) \quad (2.16)$$

Τα ζεύγη μεγάλων αποστάσεων που χρησιμοποιούνται για αυτό τον υπολογισμό, βασίζονται στην παραδοχή ότι οι τοπικές κλίσεις εξουδετερώνουν η μία την άλλη και ως αποτέλεσμα να μην είναι αναγκαίες στον καθορισμό ολικής κλίσης (global gradient determination).



Εικόνα 2.19 : Το μοτίβο δειγματοληψίας του BRISK για $N = 60$ σημεία: οι μικροί μπλε κύκλοι δηλώνουν τις θέσεις της δειγματοληψίας, οι μεγαλύτεροι και κόκκινοι κύκλοι με τις διακεκομμένες είναι ζωγραφισμένοι με ακτίνα σ που αντιστοιχεί στην τυπική απόκλιση του φίλτρου κανονικής κατανομής (Gaussian) το οποίο και χρησιμοποιείται για την εξομάλυνση των τιμών έντασης στα σημεία δειγματοληψίας. Το μοτίβο εφαρμόζεται με κλίμακα $t = 1$.



Εικόνα 2.20 : Στα αριστερά της εικόνας εμφανίζονται τα ζεύγη σημείων μικρής απόστασης S . Στα δεξιά της εικόνας εμφανίζονται τα ζεύγη σημείων μεγάλης απόστασης L .

Για τον σχηματισμό της περιστροφής και την ομαλοποίηση της κλίμακας του περιγραφέα, ο BRISK εφαρμόζει μοτίβο δειγματοληψίας (sampling pattern) περιστρεφόμενο κατά $\alpha = \arctan 2(g_y, g_x)$ γύρω από το σημείο ενδιαφέροντος k . Οι δυαδικές τιμές του διανύσματος του περιγραφέα d_k συναρμολογούνται εκτελώντας όλες τις συγκρίσεις των εντάσεων μικρής απόστασης από τα ζευγάρια σημείων $(p_i^a, p_j^a) \in S$ ώστε κάθε bit b να αντιστοιχεί σε:

$$b = \begin{cases} 1, I(p_j^a, \sigma_j) > I(p_i^a, \sigma_i) \\ 0, I(p_j^a, \sigma_j) \leq I(p_i^a, \sigma_i) \end{cases} \quad \forall (p_i^a, p_j^a) \in S \quad (2.17)$$

3 Υλοποίηση Αλγόριθμου


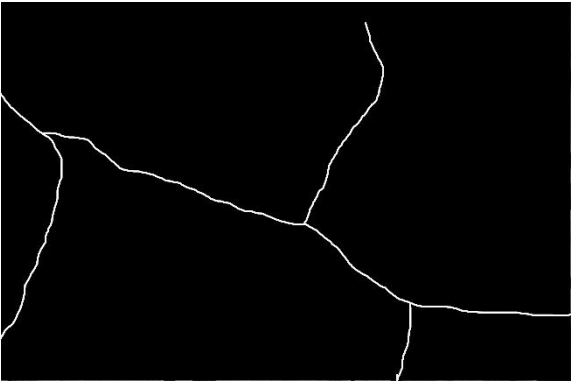

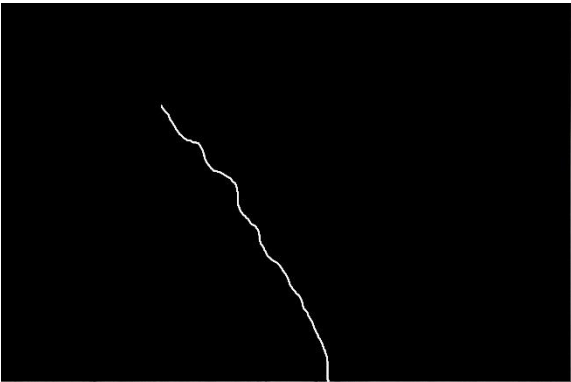

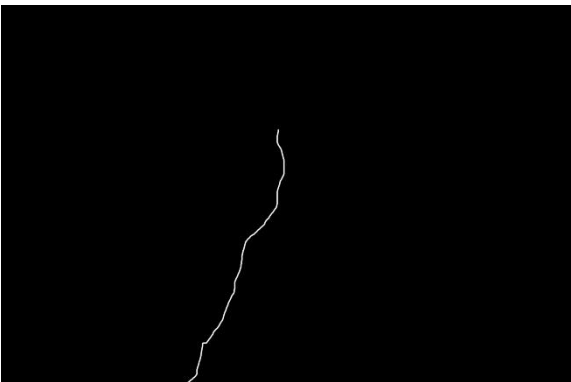

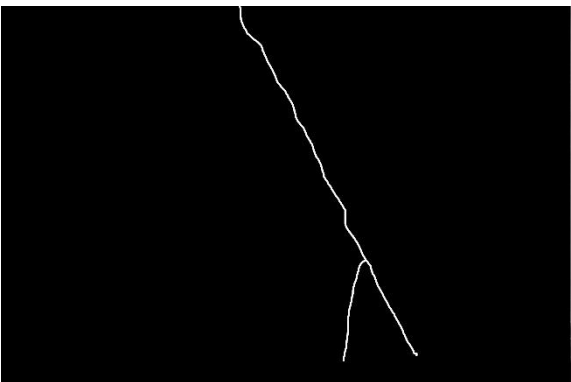
Ο αλγόριθμος που δημιουργήθηκε είχε ως σκοπό την εύρεση σημείων ενδιαφέροντος που παραπέμπουν στην συγκεκριμένη διπλωματική σε ρωγμές και θρυμματισμούς σε τούνελ από σκυρόδεμα. Τα σημεία αυτά δοκιμάστηκαν να βρεθούν με πέντε διαφορετικούς ανιχνευτές σημείων. Αυτοί είναι ο Harris, ο MinEigen, ο FAST, ο Brisk και ο SURF. Από τα αποτελέσματα που βρήκαμε αξιολογήθηκαν οι ανιχνευτές ως προς τον χρόνο, το Precision και το Recall. Σε αυτό το κεφάλαιο θα δούμε τι δεδομένα χρησιμοποιήθηκαν για τον αλγόριθμο, πως υλοποιήθηκε ο αλγόριθμος και τι αποτελέσματα έβγαλε για τον κάθε ανιχνευτή.

3.1 Δεδομένα

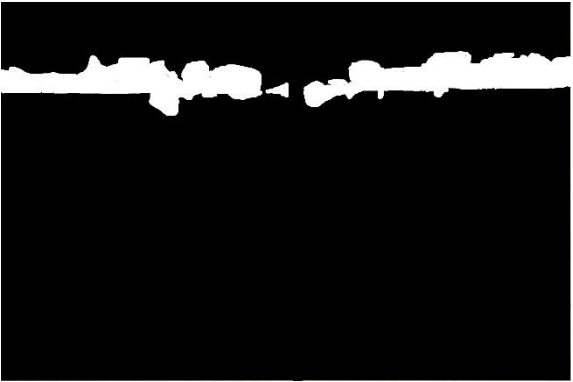

Γενικότερα η συλλογή ποιοτικών δεδομένων από τα τούνελ έχει πολλές δυσκολίες λόγω ποικίλων ειδών προβλημάτων που αντιμετωπίζουν οι εγκαταστάσεις. Τέτοιες δυσκολίες μπορεί να είναι θέματα που αφορούν τις περιβαλλοντικές συνθήκες όπως η ισχυρή ροή αέρα, η υγρασία και το στέγνωμα περιοχών που έχουν λιμνάζοντα νερά στις επιφάνειες τους. Επιπρόσθετα, παρατηρείται πολλές φορές χαμηλός και μεταβαλλόμενος φωτισμός σε ορισμένες περιοχές με αποτέλεσμα να απεικονίζονται θολές φωτογραφίες λόγω της χαμηλής ταχύτητας του κλείστρου της φωτογραφικής μηχανής. Καθώς επίσης και ο τεχνητός φωτισμός της οροφής αποκρύπτει πολλές σημαντικές πληροφορίες για το άνω μέρος των επενδύσεων του τούνελ, ενώ η διαφορετική υφή των εικόνων λόγω των διαφορετικών υλικών και του χρωματισμού μπορούν να δυσκολέψουν την διαδικασία ανίχνευσης των ελαττωμάτων. Ακόμη ένας από τους παράγοντες που μπορεί να επηρεάσει την ποιότητα των δεδομένων είναι τα διερχόμενα οχήματα κατά την διάρκεια της λειτουργίας στο τούνελ λόγω των δονήσεων που δημιουργούν. Όλα αυτά τα προβλήματα είναι αναγκαίο να ληφθούν υπόψη και να αντιμετωπισθούν στο μεγαλύτερο βαθμό που γίνεται ώστε να λαμβάνουμε όσο το δυνατόν καλύτερα δεδομένα για την διευκόλυνση των αλγορίθμων που αναγνωρίζουν διαφόρων ειδών ελαττώματα. Στην προκειμένη περίπτωση έχουμε ελαττώματα ρωγμών και θρυμματισμών.

Τα δεδομένα που χρησιμοποιήθηκαν για την υλοποίηση του συγκεκριμένου αλγορίθμου πάρθηκαν από το τούνελ V-S-H. Το V-S-H βρίσκεται στο Flums της Ελβετίας και

δραστηριοποιείται στον τομέα της έρευνας για την ασφάλεια και την προστασία υπόγειων υποδομών. Τα δεδομένα αυτά αποτελούνται αρχικά από 37 φωτογραφίες διαστάσεων 4288 x 2848 και έχουν τραβηχτεί από μία φωτογραφική μηχανή μοντέλου Nikon D300S. Οι 37 φωτογραφίες απεικονίζουν δύο ειδών ελαττώματα τα οποία είναι ρωγμές και θρυμματισμοί. Οι φωτογραφίες αυτές είναι αποθηκευμένες με μορφή jpeg και έχουν ονομαστεί im1 έως im37 ώστε να μπορούν να διαβάζονται επαναληπτικά στον αλγόριθμο. Οι 20 πρώτες απεικονίζουν μόνο ελαττώματα ρωγμών ενώ οι 17 τελευταίες μόνο θρυμματισμών. Ο σκοπός του αλγορίθμου ήταν η εύρεση σημείων ενδιαφέροντος που παραπέμπουν σε αυτά τα ελαττώματα. Για να βρεθεί όμως η αποτελεσματικότητα των ανιχνευτών για αυτά τα σημεία ήταν αναγκαίες οι ακριβείς συντεταγμένες των ελαττωμάτων αυτών ώστε να παρατηρηθούν ποια σημεία αποτύχανε στον στόχο και ποια σημεία ήταν μέσα στις περιοχές που ήταν προς αναζήτηση. Αυτές οι συντεταγμένες αποθηκεύτηκαν έπειτα από σχετική μελέτη του σκυροδέματος από πολιτικούς μηχανικούς. Η διαδικασία έγινε με χειροκίνητο τρόπο μέσω του προγράμματος Adobe Photoshop CS5.1 Windows σε ασπρόμαυρες φωτογραφίες (annotation images) ίδιων διαστάσεων 4288 x 2848 με τις φωτογραφίες που τραβήχτηκαν ώστε να ταυτίζονται τα σημεία μεταξύ τους όταν θα γίνεται ο έλεγχος για την ευστοχία των σημείων. Το λευκό χρώμα αυτών των φωτογραφιών απεικονίζει τα ελαττώματα του σκυροδέματος σε αντίθεση με το μαύρο. Οπότε πέρα από τις 37 φωτογραφίες που έχουμε για να ανιχνεύονται τα ελαττώματα έχουμε και άλλες 37 φωτογραφίες (annotation images) που απεικονίζουν με ακρίβεια τις ρωγμές και τους θρυμματισμούς. Ομοίως και αυτές οι φωτογραφίες έχουν αποθηκευτεί με όνομα ima1 έως ima37 για να μπορούν να διαβάζονται και αυτές επαναληπτικά και να ταυτίζονται με τις φωτογραφίες που είναι προς ανίχνευση. Ένα δείγμα από τα δεδομένα που χρησιμοποιήθηκαν σε αυτό τον αλγόριθμο φαίνεται στον Πίνακα 3-1 και στον Πίνακα 3-2. Σε αυτούς τους πίνακες απεικονίζονται μερικά από τα ζεύγη των φωτογραφιών που χρησιμοποιήθηκαν. Το αριστερό μέρος από τα ζεύγη αποτελεί τις εικόνες από την φωτογραφική μηχανή ενώ το δεξιό μέρος τις επεξεργασμένες εικόνες (annotation images) που απεικονίζουν με ακρίβεια τα ελαττώματα της φωτογραφίας χρησιμοποιώντας το λευκό χρώμα.

Εικόνες ρωγμών	Σχολιασμός εικόνων ρωγμών (annotation images)
	
	
	
	

Πίνακας 3-1 : Δείγμα τεσσάρων φωτογραφιών με ρωγμές μαζί με τις εικόνες που τις απομονώνει (annotation images).

Εικόνες θρυμματισμού	Σχολιασμός εικόνων θρυμματισμού (annotation images)
	
	
	
	

Πίνακας 3-2 : Δείγμα τεσσάρων φωτογραφιών με θρυμματισμούς μαζί με τις εικόνες που τους απομονώνουν (annotation images)

3.2 Ανάλυση Αλγορίθμου

Ο αλγόριθμος υλοποιήθηκε στο περιβάλλον της Matlab και οι ανιχνευτές των σημείων ενδιαφέροντος πάρθηκαν από τη βιβλιοθήκη Computer Vision System Toolbox (πίνακας 3-3).

Ανιχνευτές	Τύποι χαρακτηριστικών	Ονόματα συναρτήσεων
Harris	Γωνίες	detectHarrisFeatures
Shi-Tomasi	Γωνίες	detectMinEigenFeatures
FAST	Γωνίες	detectFASTFeatures
BRISK	Γωνίες	detectBRISKFeatures
SURF	Άμορφα σημεία(Blobs)	detectSURFFeatures

Πίνακας 3-3 : Οι ανιχνευτές που χρησιμοποιήθηκαν στον αλγόριθμο με τις συναρτήσεις τους από τη βιβλιοθήκη Computer Vision System Toolbox της Matlab.

Κάθε φωτογραφία που απεικονίζει τα ελαττώματα του σκυροδέματος είναι ένας δισδιάστατος πίνακας και κάθε τιμή αυτού του πίνακα απεικονίζει ένα από τα εικονοστοιχεία της εικόνας. Για να μπορέσουν να διαβαστούν επαναληπτικά από τον αλγόριθμο οι εικόνες ή αλλιώς οι πίνακες χρησιμοποιήθηκε η εντολή `cell`. Αυτή η εντολή αποθηκεύει σε ένα κελί όλους τους πίνακες επαναληπτικά από τον φάκελο που εμπεριέχει τις φωτογραφίες. Στη συνέχεια με την χρήση της εντολής `imread` διαβάζουμε κάθε φωτογραφία. Οι ανιχνευτές για να μπορέσουν να λειτουργήσουν δέχονται ως είσοδο εικόνες με αποχρώσεις του γκριζου (grayscale images) και για αυτό το λόγο επιλέγεται η εντολή `rgb2gray` ώστε να μετατραπούν οι εικόνες από έγχρωμες σε γκριζες. Στη συνέχεια για να φιλτράρουμε και να καθαρίσουμε την γκριζα εικόνα επιλέγονται οι εντολές `medfilt2` και `histeq`. Με αντίστοιχο τρόπο με την εντολή `imread` διαβάζουμε τις εικόνες που έχουν απομονώσει τα ελαττώματα. Σε κάθε μια από αυτές τις φωτογραφίες απομονώνουμε το λευκό χρώμα που αντιπροσωπεύει το εκάστοτε ελάττωμα (ρωγμή ή θρυμματισμός). Χρησιμοποιώντας δύο επαναληπτικούς βρόγχους διαβάζουμε κάθε εικονοστοιχείο της εικόνας και όταν κάποιο στοιχείο παίρνει την τιμή 255 που συμβολίζει το λευκό, αποθηκεύουμε τις συντεταγμένες του σημείου στον πίνακα `matrix`. Το επόμενο βήμα είναι να επιλέξουμε με ποιον από τους πέντε ανιχνευτές θέλουμε να ψάξουμε να βρούμε τα σημεία που μας ενδιαφέρουν. Οι συναρτήσεις φαίνονται στον Πίνακα 3-3. Σε αυτές τις συναρτήσεις έχουμε την δυνατότητα να επιλέξουμε τον αριθμό των σημείων που θέλουμε να ψάξουμε στην εκάστοτε φωτογραφία. Για αυτό το λόγο βάζουμε ένα επαναληπτικό βρόγχο με αρχική τιμή 5, τελική τιμή 100 και βήμα 5 ώστε να έχουμε δείγμα για το πώς συμπεριφέρεται ο ανιχνευτής. Τις συντεταγμένες των σημείων που βρίσκει ο ανιχνευτής που έχουμε επιλέξει τις

αποθηκεύουμε στον πίνακα *ir*. Έπειτα με την χρήση δύο επαναληπτικών βρόγχων ελέγχουμε αν οι συντεταγμένες κάθε σημείου που έχει βρει ο ανιχνευτής ταιριάζουν με τις συντεταγμένες των ελαττωμάτων. Αφού οι συντεταγμένες έχουν αποθηκευτεί στους πίνακες *ir* και *matrix* αντίστοιχα κάνουμε έναν έλεγχο ισότητας σε κάθε ζευγάρι των συντεταγμένων των δύο πινάκων βάζοντας και μία απόκλιση 120 εικονοστοιχείων διότι και τότε θα είμαστε κοντά στο ελάττωμα που ψάχνεται. Οι συντεταγμένες του πίνακα *ir* που ικανοποιούν την συνθήκη αυτή είναι και τα σημεία που βρίσκονται πάνω σε μία ρωγμή ή θρυμματισμό, άρα είναι και τα σημεία που μας ενδιαφέρουν. Οι συντεταγμένες αυτές αποθηκεύονται σε έναν αναλυτικό πίνακα τεσσάρων διαστάσεων που λέγεται *syntetagmenes_shmeiwn*. Αυτός ο πίνακας έχει την δυνατότητα να αποθηκεύει τις συντεταγμένες των σημείων ενδιαφέροντος ανάλογα με τον αριθμό των σημείων που έχουμε βάλει να ψάχνει ο ανιχνευτής (έχουμε βάλει να ψάχνει από 5 έως 100 σημεία με βήμα 5) για οποιαδήποτε φωτογραφία από τις 37 που έχουμε επιλέξει. Επίσης έχουμε αποθηκεύσει σε ένα πίνακα δύο διαστάσεων των αριθμό των σημείων αυτών και λέγεται *ar_shm_cracks*. Στον οποίο οι στήλες αντιπροσωπεύουν τις εικόνες και οι γραμμές το δείγμα των σημείων που έχουμε βάλει να ψάχνει ο κάθε ανιχνευτής.

Για την σύγκριση των ανιχνευτών και την αξιολόγηση των αποτελεσμάτων χρησιμοποιήθηκαν στην πορεία δύο μέτρα αξιολόγησης η ακρίβεια (*Precision*) και το μέτρο της ευαισθησίας (*Recall*). Αυτά ορίζονται ως εξής :

Precision : Πρόκειται για το ποσοστό των δεδομένων που ταξινομήθηκαν σωστά σε σχέση με όλα τα δεδομένα που ταξινομήθηκαν στην εκάστοτε κλάση.

$$Precision = \frac{TP}{TP + FP} \quad (3.1)$$

Recall : Είναι το ποσοστό των σωστά ταξινομημένων δεδομένων σε μια κλάση σε σχέση με το σύνολο των δεδομένων που ανήκουν σε αυτή.

$$Recall = \frac{TP}{TP + FN} \quad (3.2)$$

Όπου *TP* : Ο αριθμός των πραγματικών και θετικών ορισμένων σημείων

FP : Ο αριθμός των ψεύτικων και θετικών ορισμένων σημείων

FN : Ο αριθμός των ψεύτικων και αρνητικών ορισμένων σημείων

Οπότε με βάση αυτά υπολογίσαμε το Precision και το Recall κάθε φωτογραφίας για όσα σημεία έχουμε βάλει να ψάχνουν οι ανιχνευτές από τις σχέσεις :

$$Precision(t, z) = \frac{ar_shm_cracks(t, z)}{size(rr, 1)}$$

$$Recall(t, z) = \frac{ar_shm_cracks(t, z)}{size(matrix, 1)}$$

Όπου $ar_shm_cracks(t, z)$ είναι ο αριθμός των σημείων ενδιαφέροντος που έχει βρει ένας από τους πέντε ανιχνευτές για όσα σημεία ($t \times \text{βήμα}$) έχουμε βάλει να ανιχνεύσει για μία εικόνα z από τις 37 που έχουμε ως δεδομένα (Παράδειγμα το $ar_shm_cracks(4, 19)$ είναι ο αριθμός των σημείων ενδιαφέροντος ενός ανιχνευτή που τον έχουμε βάλει να ψάχνει για $4 \times 5 = 20$ σημεία στην φωτογραφία 19). Το $size(rr, 1)$ είναι το μέγεθος των σημείων που έχει βρει ο αλγόριθμος είτε έχουν πέσει πάνω στο ελάττωμα της εικόνας είτε όχι. Ενώ το $size(matrix, 1)$ είναι ο αριθμός όλων των εικονοστοιχείων της φωτογραφίας που έχουν ελαττώματα. Στην συνέχεια υπολογίζουμε το μέγιστο, το ελάχιστο και το μέσο Precision και Recall κάθε φωτογραφίας από 5 έως 100 σημεία με βήμα 5 προς ανίχνευση. Επιπλέον, υπολογίζουμε το μέσο Precision και μέσο Recall όλων των φωτογραφιών για κάθε ανιχνευτή ώστε να συγκρίνουμε τους ανιχνευτές συνολικά ως προς την απόδοσή τους.

Συνοπτικά τα στάδια με τις εντολές του αλγορίθμου είναι :

1. `Imread` : Διαβάζουμε όλες τις φωτογραφίες και τις annotation εικόνες τους επαναληπτικά (εικόνα 3.1).
2. `Rgb2gray` : Μετατρέπουμε την φωτογραφία από έγχρωμη σε γκρίζα ώστε να μπορούν τις διαβάσουν οι ανιχνευτές (εικόνα 3.2).
3. `Medfilt2` και `histeq` : Για να φιλτράρουμε και να καθαρίσουμε την γκρίζα εικόνα (εικόνα 3.3).
4. Με δύο επαναληπτικούς βρόγχους `for` : Διαβάζουμε κάθε εικονοστοιχείο της annotation εικόνας και όταν κάποιο στοιχείο παίρνει την τιμή 255 που συμβολίζει το λευκό, καταχωρούμε τις συντεταγμένες του σημείου στον πίνακα `matrix` (Πίνακας 3-4).

5. Θέτουμε πόσα σημεία θέλουμε να ψάξει ο κάθε ανιχνευτής. Στην προκειμένη περίπτωση θέλουμε να ψάξει από 5 έως 100 σημεία με βήμα 5. Επιλέγουμε με ποιον ανιχνευτή θέλουμε να ψάξουμε για σημεία ενδιαφέροντος καλώντας την αντίστοιχη συνάρτηση (εικόνα 3.4).
6. Αποθηκεύουμε τις συντεταγμένες των σημείων που βρίσκει ο ανιχνευτής στον πίνακα `π`.
7. Με την χρήση δύο επαναληπτικών βρόγχων `for` : Ελέγχουμε ποιες από τις συντεταγμένες του πίνακα `π` ταιριάζουν με τις συντεταγμένες του πίνακα `matrix` που βρίσκονται τα ελαττώματα (Πίνακας 3-5).
8. Αποθηκεύουμε τις συντεταγμένες στον πίνακα `syntagmenes_shmeiwn` και ειδικότερα τον αριθμό των σημείων αυτών στον πίνακα `ar_shm_cracks`.
9. Υπολογίζουμε τα μέτρα αξιολόγησης `Precision` και `Recall` για κάθε φωτογραφία στον εκάστοτε ανιχνευτή.
10. Υπολογίζουμε τα μέσα μέτρα αξιολόγησης για όλες τις φωτογραφίες σε κάθε ανιχνευτή ώστε να συγκρίνουμε τους ανιχνευτές συνολικά ως προς την απόδοσή τους.

```

row=1;
[m,n]=size(anoatation);
for i=1:m
    for j=1:n
        if anoatation(i,j)==255;
            matrix(row,1)=j;
            matrix(row,2)=i;
            row=row+1;
        end
    end
end
end

```

Πίνακας 3-4: Μέρος του κώδικα που διαβάζει κάθε εικονοστοιχείο της annotation εικόνας και όταν κάποιο στοιχείο παίρνει την τιμή 255 που συμβολίζει το λευκό, καταχωρεί τις συντεταγμένες του σημείου στον πίνακα matrix

```

a=1;
w=1;
rr=round(points.selectStrongest(x).Location);
myindex=ones(length(rr),1)*(-1);

for i=1:size(matrix,1)
    for w=1:size(rr,1)

        if ((rr(w,1) < matrix(i,1)+120) && (rr(w,1) >
matrix(i,1)-120)) && ((rr(w,2) < matrix(i,2)+120) &&
(rr(w,2) > matrix(i,2)-120)) && (myindex(w) == -1)

            shmeia(a,1,t)=rr(w,1);

syntagmenes_shmeiwn(a,1,t,z)=shmeia(a,1,t);
            shmeia(a,2,t)=rr(w,2);

syntagmenes_shmeiwn(a,2,t,z)=shmeia(a,2,t);
            a=a+1;
            myindex(w)=1;

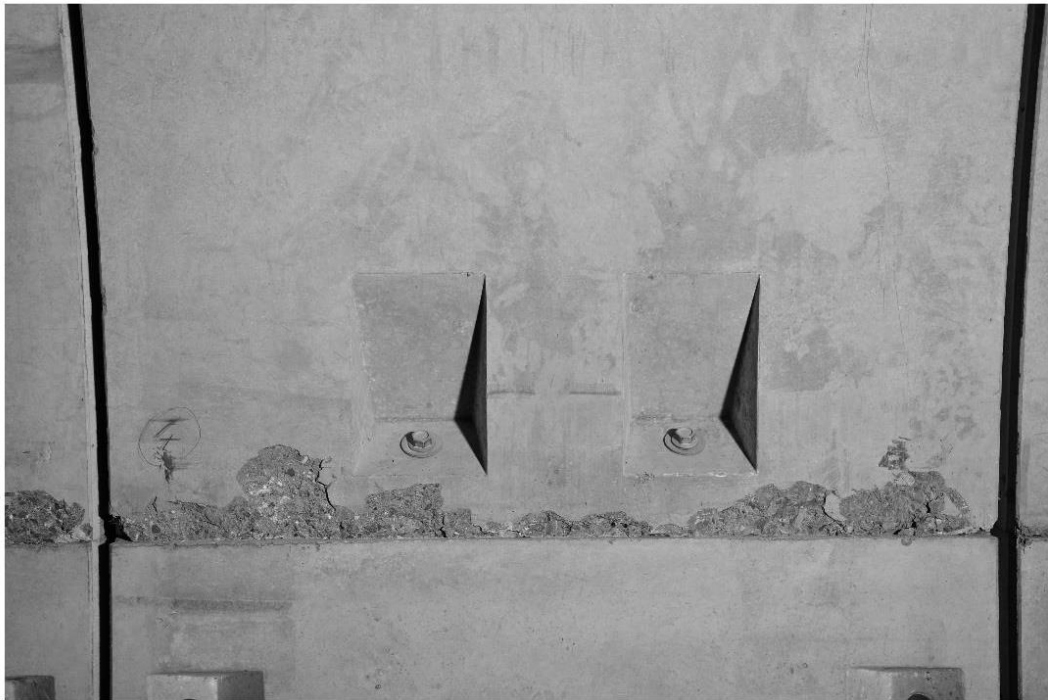
        end
    end
end
end

```

Πίνακας 3-5 : Μέρος του κώδικα το οποίο με την χρήση δύο επαναληπτικών βρόγχων for ελέγχει ποιες από τις συντεταγμένες του πίνακα rr ταιριάζουν με τις συντεταγμένες του πίνακα matrix που βρίσκονται τα ελαττώματα. Επίσης αποθηκεύει τ συντεταγμένες στον πίνακα syntagmenes_shmeiwn.



Εικόνα 3.1 : Παράδειγμα της φωτογραφίας 31 από τις 37 που αντιμετωπίζει ελάττωμα θρυμματισμού μαζί με την annotation εικόνα της.



Εικόνα 3.2 : Μετατροπή της εικόνας από έγχρωμη σε γκρίζα



Εικόνα 3.3 : Χρήση των Medfilt2 και histeq για να φιλτράρουμε και να καθαρίσουμε την γκρίζα εικόνα



Εικόνα 3.4 : Έχουμε χρησιμοποιήσει στην φωτογραφία 31 τον ανιχνευτή SURF καλώντας την συνάρτηση `points = detectSURFFeatures(sgrtunnel)` και τον έχουμε βάλει να ψάχνει για 45 σημεία με την εντολή `points.selectStrongest(45).Location`. Τα σημεία που έχει βρει απεικονίζονται με σταυρούς. Οι πράσινοι σταυροί συμβολίζουν τα σημεία ενδιαφέροντος σε αντίθεση με το μοναδικό κόκκινο σταυρό στην προκειμένη περίπτωση να έχει αστοχήσει.

3.3 Αποτελέσματα

Σε αυτή την φάση της διπλωματικής παρατίθενται τα αποτελέσματα του αλγορίθμου ως προς τον χρόνο και την αποτελεσματικότητα του ανάλογα με τον ανιχνευτή που επιλέχθηκε να ανιχνεύσει τις φωτογραφίες για ρωγμές και θρυμματισμούς. Στον παρακάτω πίνακα 3-6 φαίνεται ο χρόνος που έκανε ο αλγόριθμος για κάθε ανιχνευτή ξεχωριστά για τις 20 φωτογραφίες ρωγμών, για τις 17 εικόνες θρυμματισμού και συνολικά για όλες τις φωτογραφίες. Όπως φαίνεται και από τον πίνακα 3-6 οι FAST και BRISK ήταν ταχύτεροι από τους υπόλοιπους ανιχνευτές με 6436,7 δευτερόλεπτα και 6492,51 αντίστοιχα. Σε αντίθεση ο αλγόριθμος ελαχίστων ιδιοτιμών (MinEigen) έκανε τον περισσότερο χρόνο με διαφορά κάνοντας τον αλγόριθμο να χρειαστεί 9880,3 δευτερόλεπτα για να ολοκληρωθεί. Αυτή η καθυστέρηση είναι πιθανόν να οφείλεται στον υπολογισμό των ιδιοτιμών της απόκρισης της γωνίας R.

Αλγόριθμοι αναγνώρισης σημείων ενδιαφέροντος	Χρόνος αλγορίθμου t(sec) για 20 εικόνες ρωγμών	Χρόνος αλγορίθμου t(sec) για 17 εικόνες θρυμματισμού	Συνολικός χρόνος αλγορίθμου t(sec)
BRISK	563,51	5929	6492,51
FAST	505,7	5931	6436,7
HARRIS	1360,65	6237,98	7598,63
MinEigen	2434,93	7445,37	9880,3
SURF	1128,38	6297,06	7425,44

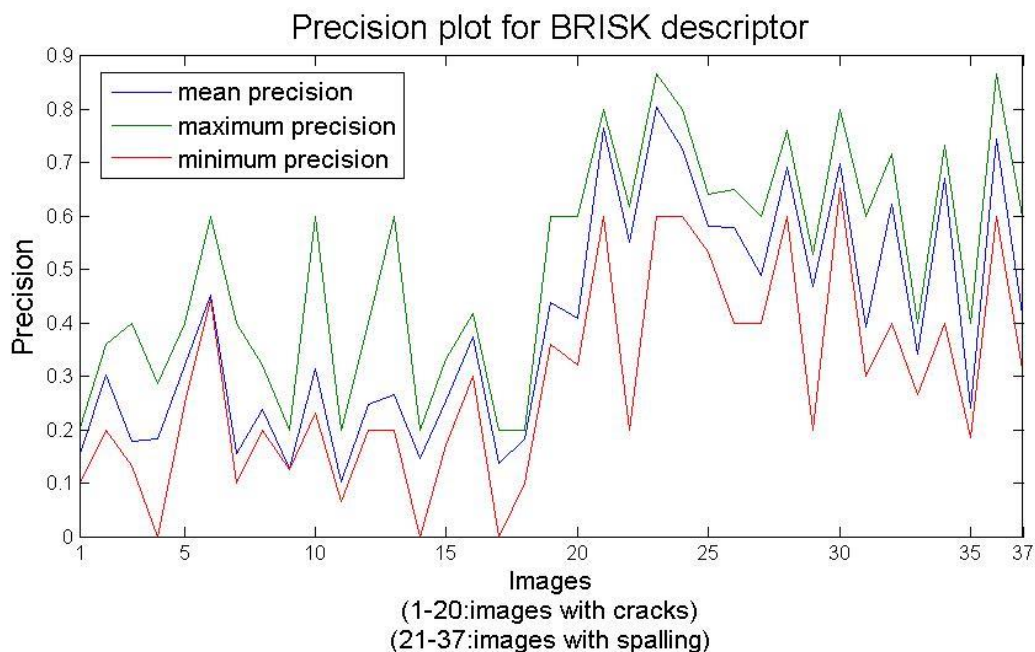
Πίνακας 3-6 : Υπολογισμός χρόνου ολοκλήρωσης του αλγορίθμου με βάση τον ανιχνευτή που έχει επιλεγεί.

Στη συνέχεια θα παρατεθούν γραφήματα με την αξιολόγηση των ανιχνευτών για κάθε φωτογραφία. Για αυτή την αξιολόγηση όπως είχαμε αναφέρει χρησιμοποιούνται τα μέτρα αξιολόγησης Precision και Recall.

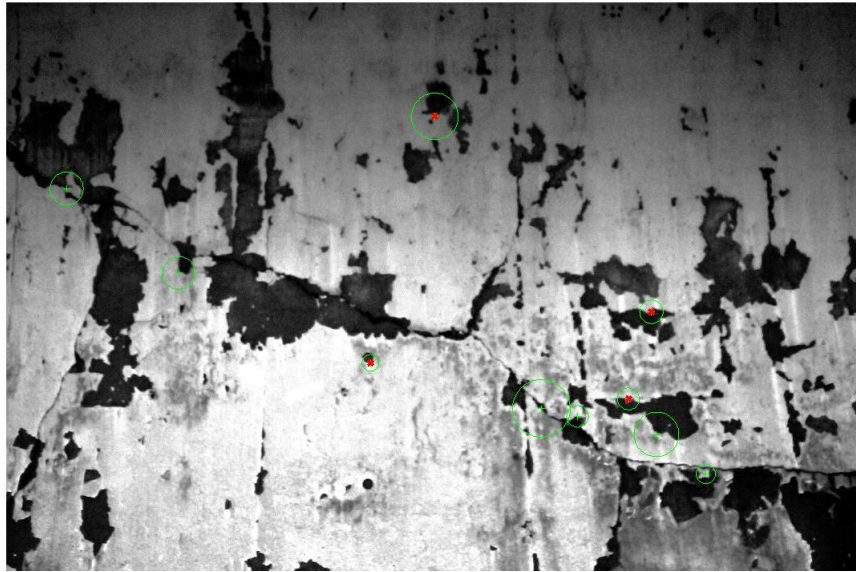
3.3.1 Ανιχνευτής BRISK

Στο γράφημα 3-1 φαίνεται η αξιολόγηση του Brisk ως προς το μέσο, μέγιστο και ελάχιστο Precision για κάθε φωτογραφία εφόσον τον έχουμε βάλει να ψάχνει από 5 έως 100 με βήμα 5 σημεία ενδιαφέροντος. Παρατηρούμε οι μετρήσεις να αυξάνονται από την φωτογραφία 20 και μετά διότι οι φωτογραφίες αυτές έχουν ελαττώματα θρυμματισμού. Αυτά τα ελαττώματα έχουν μεγαλύτερες επιφάνειες οπότε είναι και πιο εύκολο για τον Brisk αλλά και για τους άλλους ανιχνευτές που θα δούμε στην πορεία να τις αναγνωρίσουν. Η καλύτερη φωτογραφία με ρωγμές

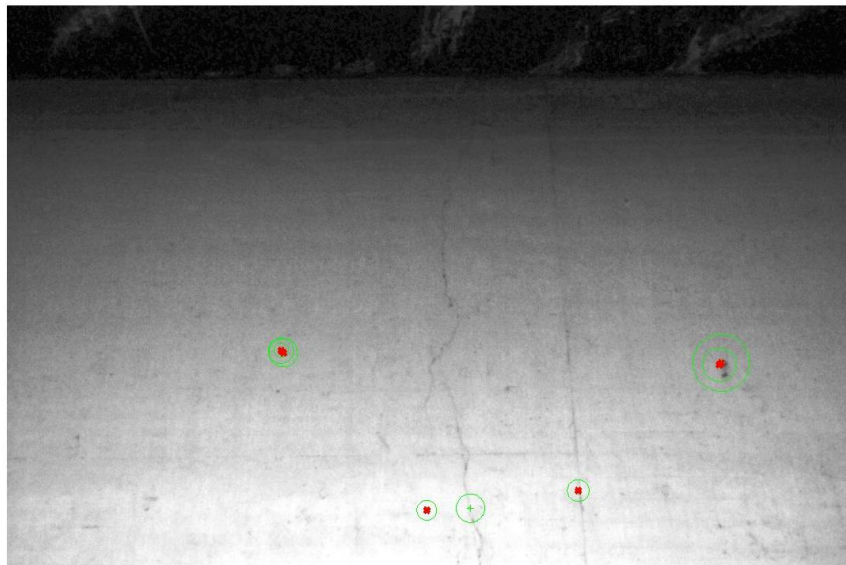
(φωτογραφίες 1-20) για τον Brisk που βρήκε και τα περισσότερα σημεία είναι η φωτογραφία 19 (εικόνα 3.5) με μέσο Precision 43,83% . Το καλύτερο Precision το είχε όταν έψαχνε μόνο για 10 σημεία ο ανιχνευτής ενώ στην συνέχεια υπήρχε μια σταθεροποίηση στις τιμές που έπαιρνε. Η φωτογραφία 6 παρόλο που είχε πολύ μεγάλη τιμή στο Precision δεν την λαμβάνουμε υπόψη διότι μετά από κάποια στιγμή ο ανιχνευτής σταμάτησε να βρίσκει σημεία. Η φωτογραφία που βρήκε τα λιγότερα σημεία είναι η 9 (εικόνα 3.6) και βρήκε μόλις ένα σημείο με μέσο precision 12,88%. Η καλύτερη φωτογραφία που είχε ελαττώματα θρυμματισμών (φωτογραφίες 21-37) είναι η φωτογραφία 23 (εικόνα 3.7) με μέσο Precision 80,49% και βρίσκει το καλύτερο Precision 86,67% όταν ανιχνεύει 15 σημεία δηλαδή τότε βρίσκει 13 σημεία ενδιαφέροντος. Η φωτογραφία αυτή συμπεριφέρεται ομαλά ως προς την ανίχνευση και βρίσκει 100 σημεία ανίχνευσης 80 σημεία ενδιαφέροντος. Η χειρότερη φωτογραφία θρυμματισμού είναι η 35 (εικόνα 3.8) με μέσο Precision 23,95% και έχει το μικρότερο Precision όταν ψάχνει για 70 σημεία με τιμή 18,57%. Μετά αντιμετωπίζει μία μικρή άνοδο και φτάνει 24% όταν ψάχνει για 100 σημεία.



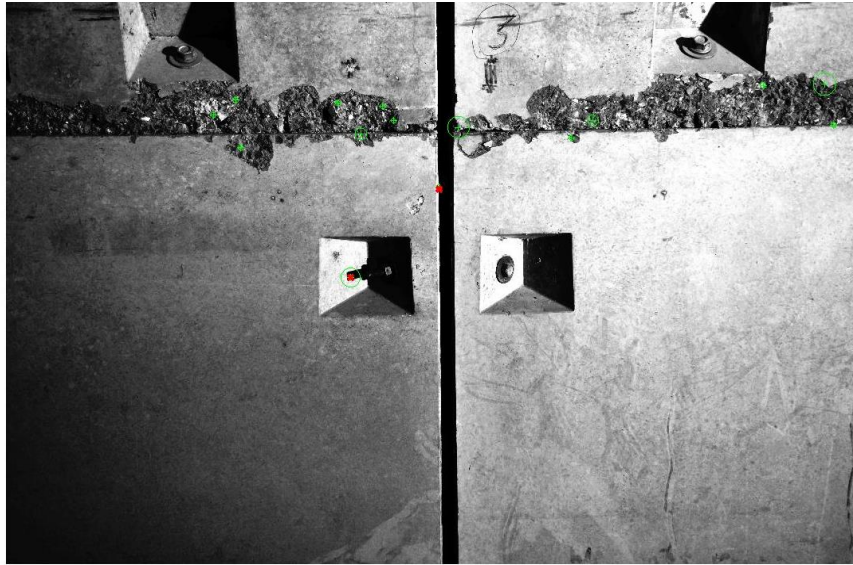
Γράφημα 3-1 : Αξιολόγηση του Brisk ως προς το Precision για κάθε φωτογραφία



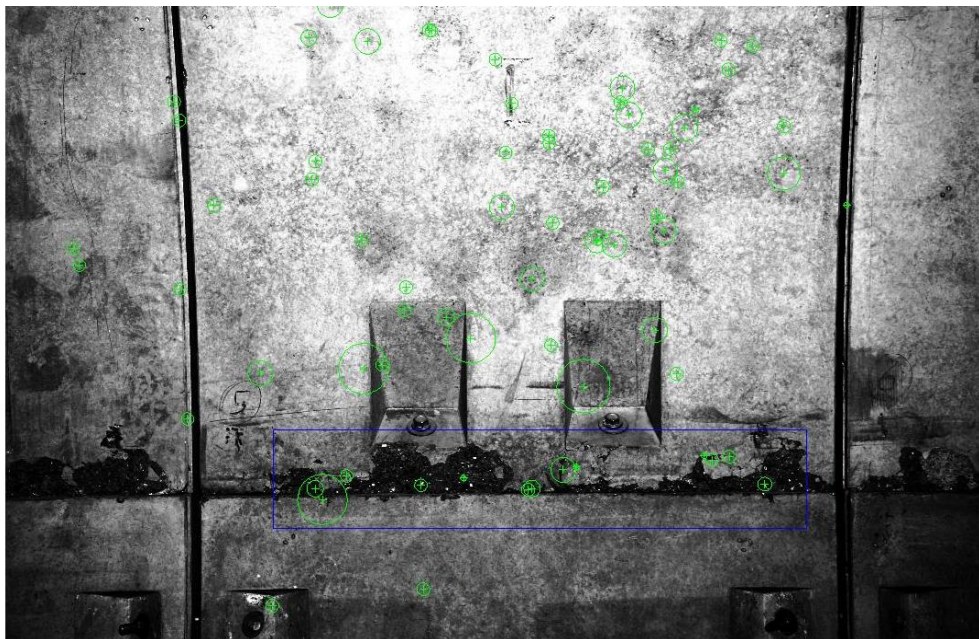
Εικόνα 3.5 : Η καλύτερη φωτογραφία ρωγμών για τον BRISK είναι η 19 όταν βρίσκει 6 σημεία ενδιαφέροντος από τα 10 σημεία ανίχνευσης. Τα κόκκινα είναι τα σημεία στα οποία έχει αστοχήσει.



Εικόνα 3.6 : Η χειρότερη φωτογραφία ρωγμών για τον BRISK είναι η 6 και βρίσκει 1 σημείο ενδιαφέροντος για οποιοδήποτε αριθμό σημείων ανίχνευσης



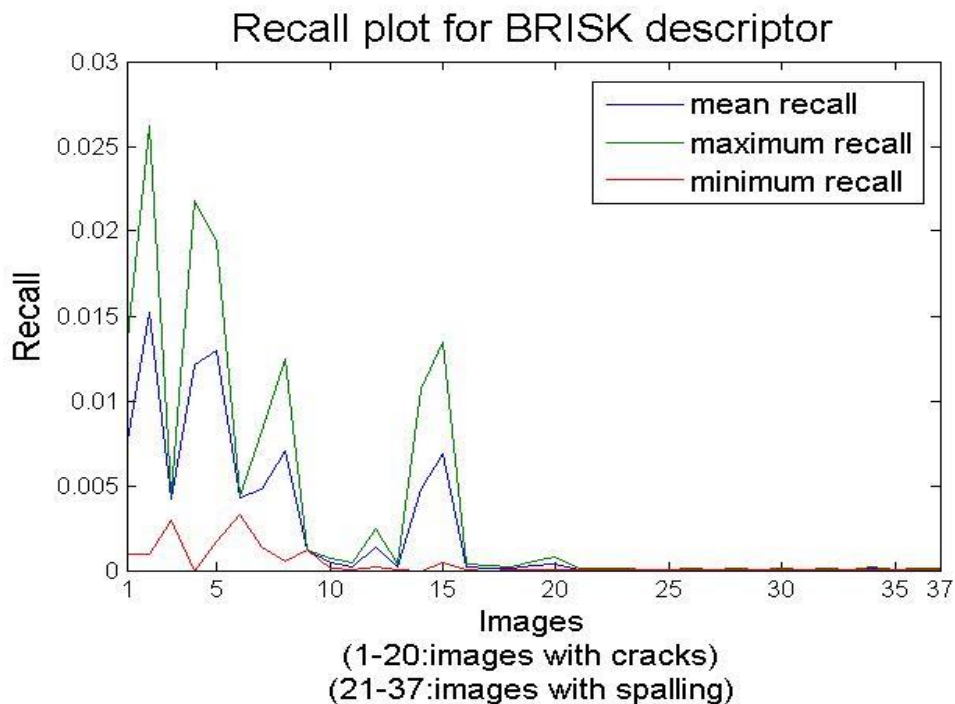
Εικόνα 3.7 : Η καλύτερη φωτογραφία θρυμματισμών για τον BRISK είναι η φωτογραφία 23 και βρίσκει το καλύτερο Precision 86,67% όταν ανιχνεύει 15 σημεία δηλαδή τότε βρίσκει 13 σημεία ενδιαφέροντος.



Εικόνα 3.8 : Η χειρότερη φωτογραφία θρυμματισμών για τον BRISK είναι η 35 και έχει το χειρότερο Precision 18,57 % όταν βρίσκει 13 σημεία ενδιαφέροντος(φαίνονται στο μπλε πλαίσιο) από τα 70 σημεία ανίχνευσης

Στο γράφημα 3-2 παρακάτω αξιολογείται ο BRISK ως προς το Recall. Παρατηρούμε πολύ μικρές τιμές στο γράφημα διότι τα σημεία που ανιχνεύουμε είναι πάρα πολύ λιγότερα σε σχέση με τον αριθμό των σημείων των ρωγμών και ειδικότερα από τα σημεία των θρυμματισμών. Για

αυτό βλέπουμε και οι μετρήσεις να μειώνονται απότομα από την φωτογραφία 20 και μετά, σε αντίθεση με το γράφημα του Precision. Γενικότερα δεν μπορούμε με αυτό το μέτρο να συγκρίνουμε τις εικόνες μεταξύ τους διότι η κάθε φωτογραφία διαφέρει ως προς τον αριθμό των ελαττωμάτων τις. Σε περίπτωση όμως που είχαμε μεγαλύτερη υπολογιστική ισχύ το μέτρο αυτό θα έπαιρνε διαφορετική και μεγαλύτερη διάσταση αφού θα είχαμε την δυνατότητα να βάζουμε τους αλγορίθμους να συγκρίνονται για περισσότερα σημεία.

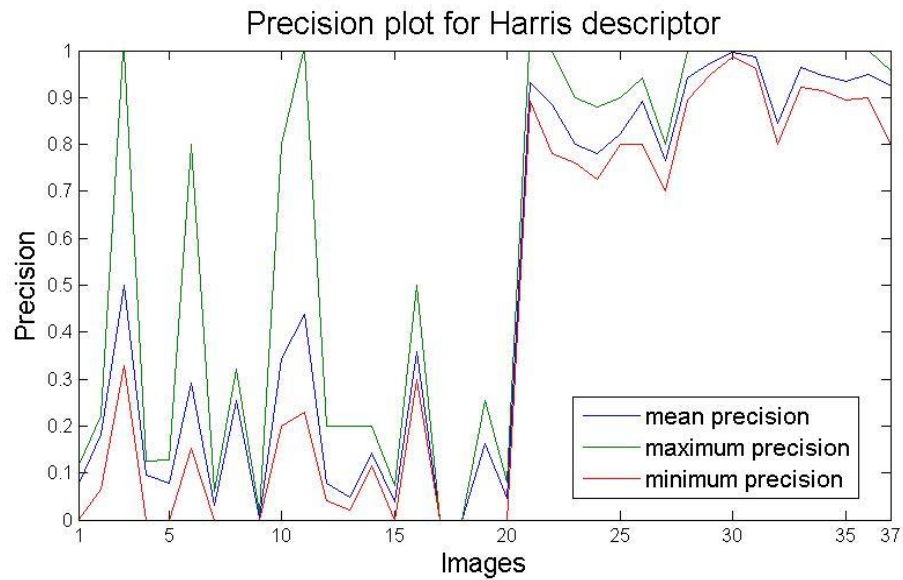


Γράφημα 3-2 : Η αξιολόγηση του BRISK ως προς το Recall

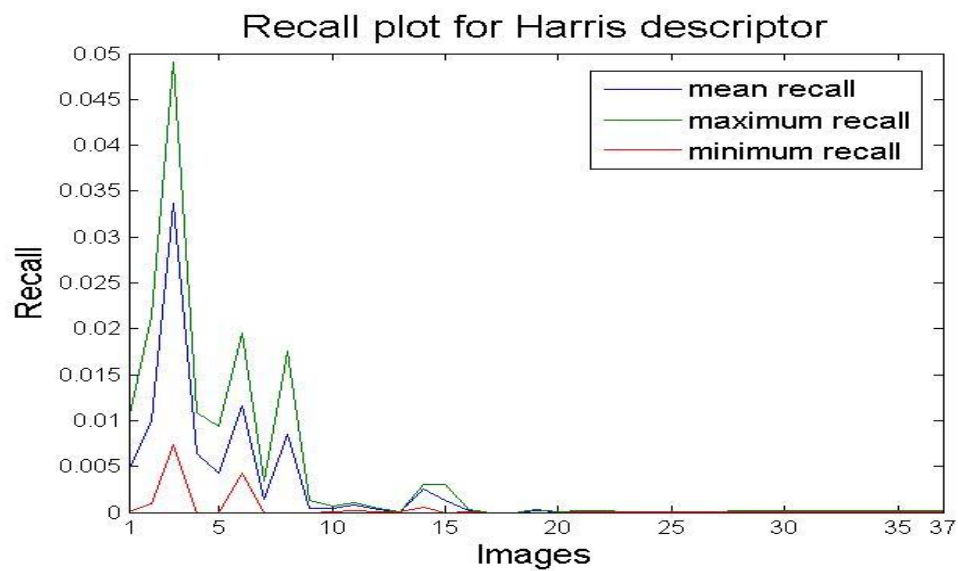
3.3.2 Ανιχνευτής Harris

Στο γράφημα 3-3 θα δούμε πως συμπεριφέρεται ο ανιχνευτής Harris αντίστοιχα ως προς το Precision. Η φωτογραφία 3 (εικόνα 3.9) εμφανίζει μεγαλύτερη αποτελεσματικότητα με μέσο Precision 50,09% από τις φωτογραφίες που έχουν ως ελαττώματα ρωγμές και βρίσκει το καλύτερο Precision με 100% όταν ψάχνει για 5 σημεία. Μετά ακολουθεί πτωτική πορεία και φτάνει στα 100 σημεία να έχει 33%. Η χειρότερες φωτογραφίες με ρωγμές είναι οι φωτογραφίες 17 (εικόνα 3.10) και 18 εκεί ο αλγόριθμος δεν βρίσκει ούτε ένα σημείο ενδιαφέροντος. Η καλύτερη φωτογραφία θρυμματισμού είναι η 30 (εικόνα 3.11) με μέσο Precision με 99,65%. Εκεί υπάρχουν πολλές φορές περιπτώσεις που ο ανιχνευτής έχει Precision 100%. Στα 100 σημεία ανίχνευσης βρίσκει 99 σημεία ενδιαφέροντος. Η χειρότερη φωτογραφία θρυμματισμού

είναι η 24 (εικόνα 3.12) με μέσο Precision 78,02% και βρίσκει το μικρότερο Precision 73,33% στα 45 σημεία ανίχνευσης και βρίσκει σε αυτή την περίπτωση 33 σημεία.



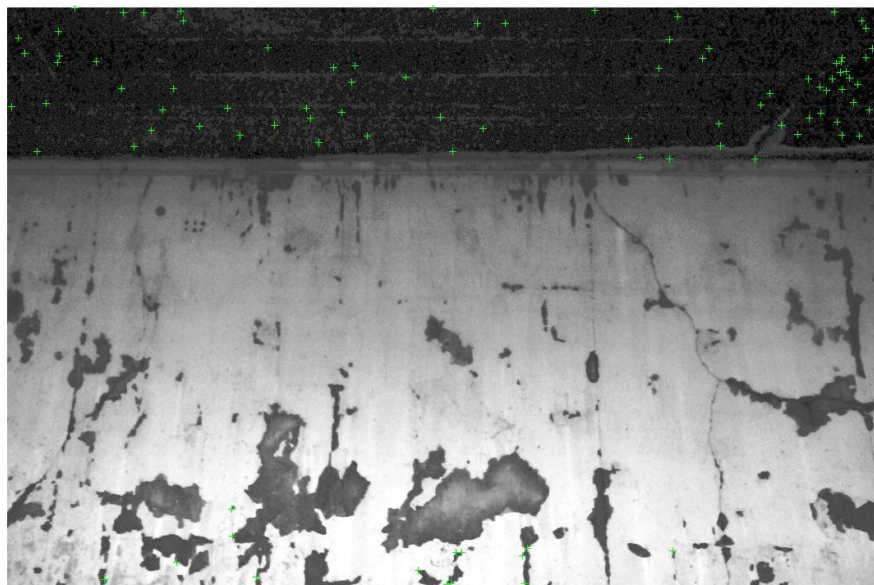
Γράφημα 3-3 : Αξιολόγηση του Harris ως προς το Precision για κάθε φωτογραφία



Γράφημα 3-4 : Η αξιολόγηση του Harris ως προς το Recall



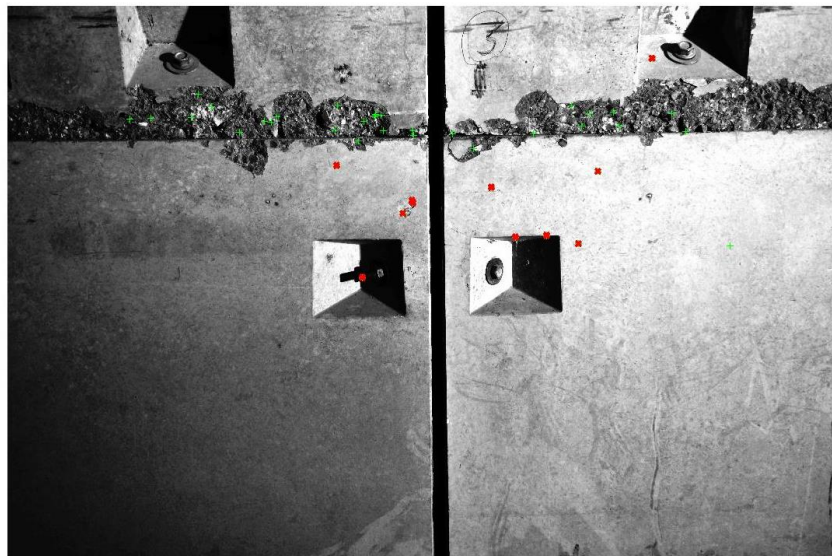
Εικόνα 3.9 : Η καλύτερη φωτογραφία ρωγμών για τον Harris είναι η 5 και έχει καλύτερο Precision 100% στα 5 σημεία ανίχνευσης



Εικόνα 3.10 : Η χειρότερη φωτογραφία ρωγμών για τον Harris είναι η 17 όπου έχουν αστοχήσει όλα τα σημεία ανίχνευσης. Βλέπουμε ότι η φωτογραφία δεν είναι καλή διότι έχει αποτυπώσει και μέρος με διαφορετική υφή. Όπως φαίνεται ο ανιχνευτής έχει εστιάσει στο μέρος που δεν θέλουμε.



Εικόνα 3.11 : Η καλύτερη φωτογραφία θρυμματισμών για τον Harris είναι η 30 και έχει πολλές φορές Precision 100%. Εδώ βλέπουμε όταν έχει 99% Precision στα 100 σημεία ανίχνευσης. Αστοχεί μόνο ο κόκκινος σταυρός.

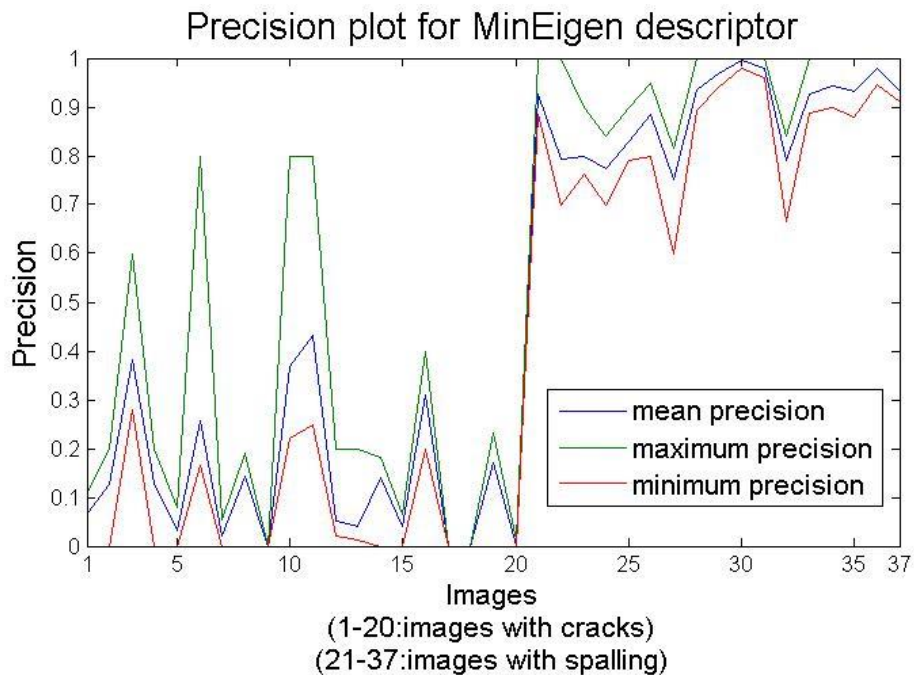


Εικόνα 3.12 : Η χειρότερη φωτογραφία θρυμματισμού για τον Harris είναι η 24 και βρίσκει το μικρότερο Precision 73,33% στα 45 σημεία ανίχνευσης. Εκεί αστοχούν 12 όπως φαίνεται και από τους κόκκινους σταυρούς

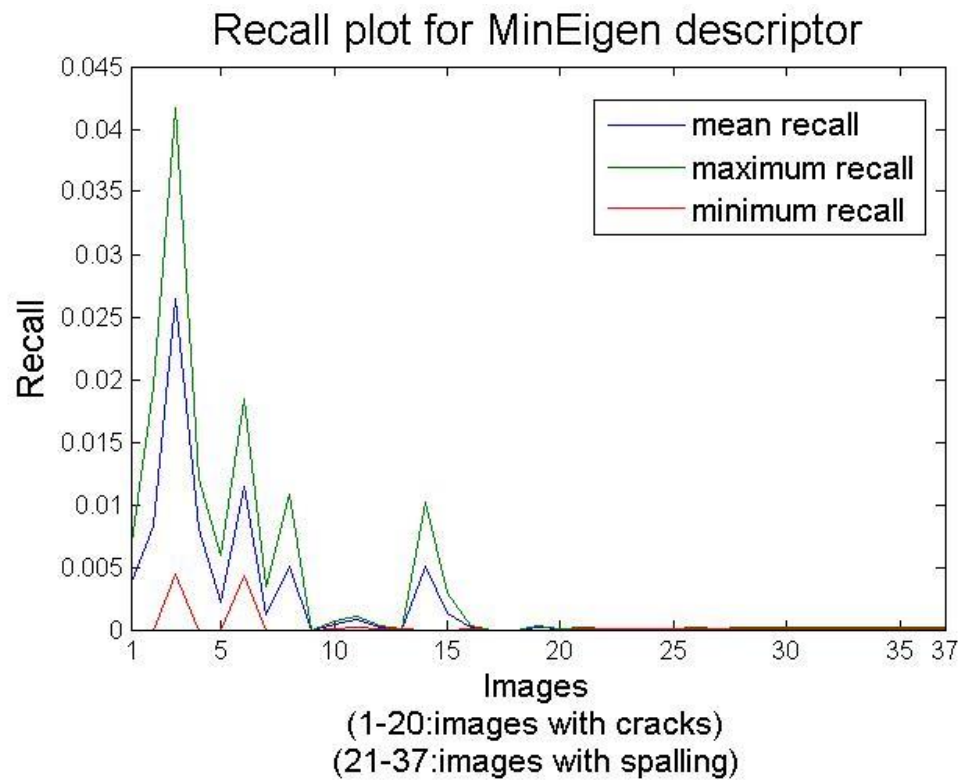
3.3.3 Ανιχνευτής ελαχίστων ιδιοτιμών Shi-Tomasi (MinEigen)

Στον ανιχνευτή ελαχίστων ιδιοτιμών Shi-Tomasi (MinEigen) παρατηρούμε από το γράφημα 3-5 ότι το μεγαλύτερο μέσο Precision από τις φωτογραφίες ρωγμών το έχει η φωτογραφία 11 (εικόνα 3.13) με 43,12%. Εμφανίζει το μέγιστο Precision 80% όταν βρίσκει 12 σημεία

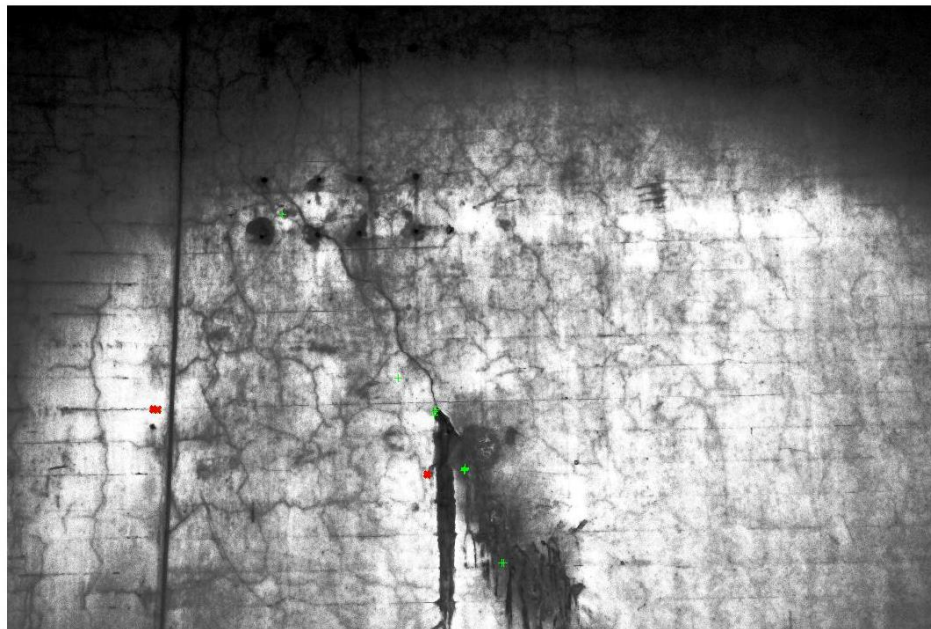
ενδιαφέροντος από τα 15 σημεία ανίχνευσης. Μετά έχει σταδιακή πτώση η τιμή της και φτάνει να βρίσκει 25 σημεία ενδιαφέροντος από τα 100 σημεία ανίχνευσης. Η χειρότερες φωτογραφίες με ρωγμές είναι οι 9 (βλέπε εικόνα 3.14), 17 και 18 που δεν ανιχνεύουν κανένα σημείο που να μας ενδιαφέρει. Η καλύτερη φωτογραφία θρυμματισμού είναι η 30 (εικόνα 3.15) με μέσο Precision 99,67%. Έχει σε πολλές περιπτώσεις 100% Precision και κάνει ανίχνευση 100 σημείων βρίσκει 98 σημεία ενδιαφέροντος. Δεν είναι τυχαίο ότι και στον Harris εμφανίζεται η ίδια φωτογραφία ως καλύτερη εικόνα θρυμματισμού αφού οι ανιχνευτές Shi-Tomasi και Harris διαφέρουν πολύ λίγο μεταξύ τους. Η χειρότερη φωτογραφία θρυμματισμού είναι η 27 (εικόνα 3.16) η οποία έχει μέσο Precision 75,07%. Εμφανίζει το μικρότερο Precision 60% στα 5 σημεία ανίχνευσης βρίσκοντας 3 σημεία ενδιαφέροντος.



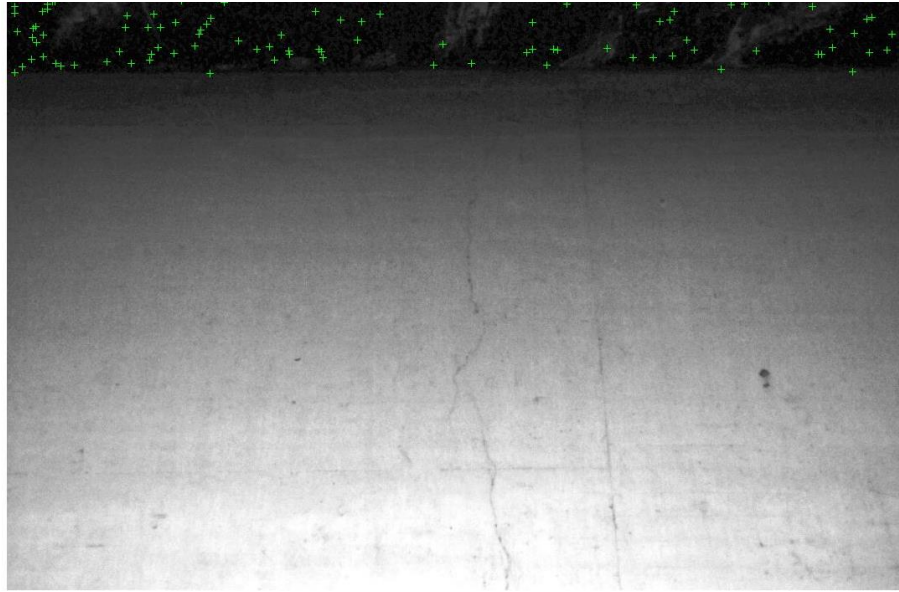
Γράφημα 3-5 : Αξιολόγηση του Shi-Tomasi(MinEigen) ως προς το Precision για κάθε φωτογραφία



Γράφημα 3-6 : Αξιολόγηση του Shi-Tomasi ως προς το Recall για κάθε φωτογραφία



Εικόνα 3.13 : Η καλύτερη φωτογραφία ρωγμών για τον Shi-Tomasi(MinEigen) είναι η 11 και έχει καλύτερο Precision 80% στα 15 σημεία ανίχνευσης. Εκεί αστοχούν 3 με τους κόκκινους σταυρούς



Εικόνα 3.14 : Η χειρότερη φωτογραφία ρωγμών για τον ανιχνευτή Shi-Tomasi(MinEigen) είναι η 9 όπου δεν βρίσκει κανένα σημείο ανίχνευσης. Η φωτογραφία αποτελείται από 2 μέρη με διαφορετικές υφές. Το πρόβλημα είναι ότι εστιάζει στο μέρος που δεν υπάρχουν τα ελαττώματα ρωγμών



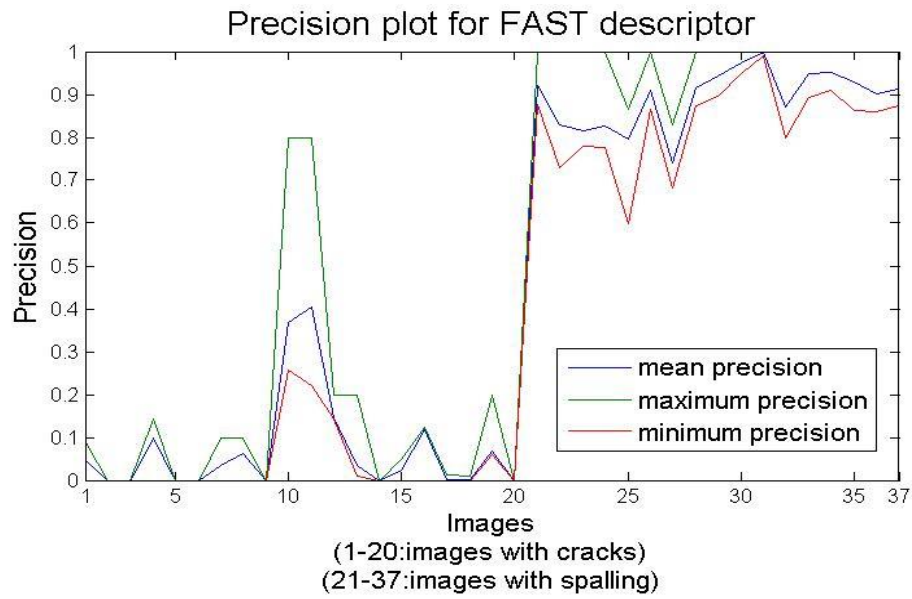
Εικόνα 3.15 : Η καλύτερη φωτογραφία θρυμματισμών για τον Shi-Tomasi(MinEigen) είναι η 30 και έχει πολλές φορές Precision 100%. Εδώ βλέπουμε όταν έχει 98% Precision στα 100 σημεία ανίχνευσης. Αστοχούν μόνο οι 2 κόκκινοι σταυροί.



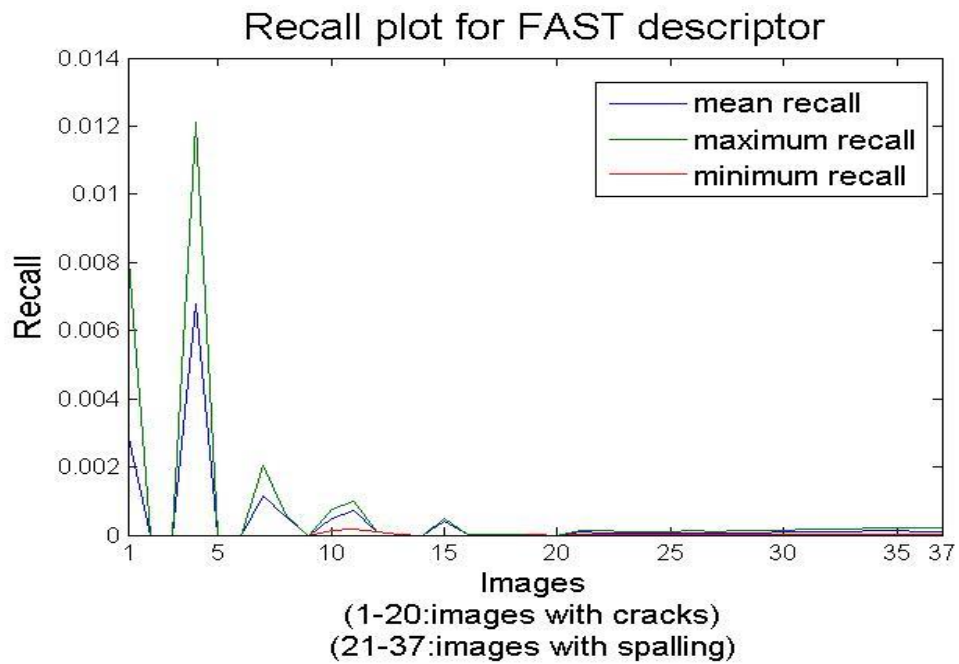
Εικόνα 3.16 : Η χειρότερη φωτογραφία θρυμματισμών για τον Shi-Tomasi(MinEigen) είναι η 27 και έχει το χειρότερο Precision 60% όταν ψάχνει για 5 σημεία ανίχνευσης. Τότε αστοχούν 2 σημεία με τους κόκκινους σταυρούς.

3.3.4 Ανιχνευτής FAST

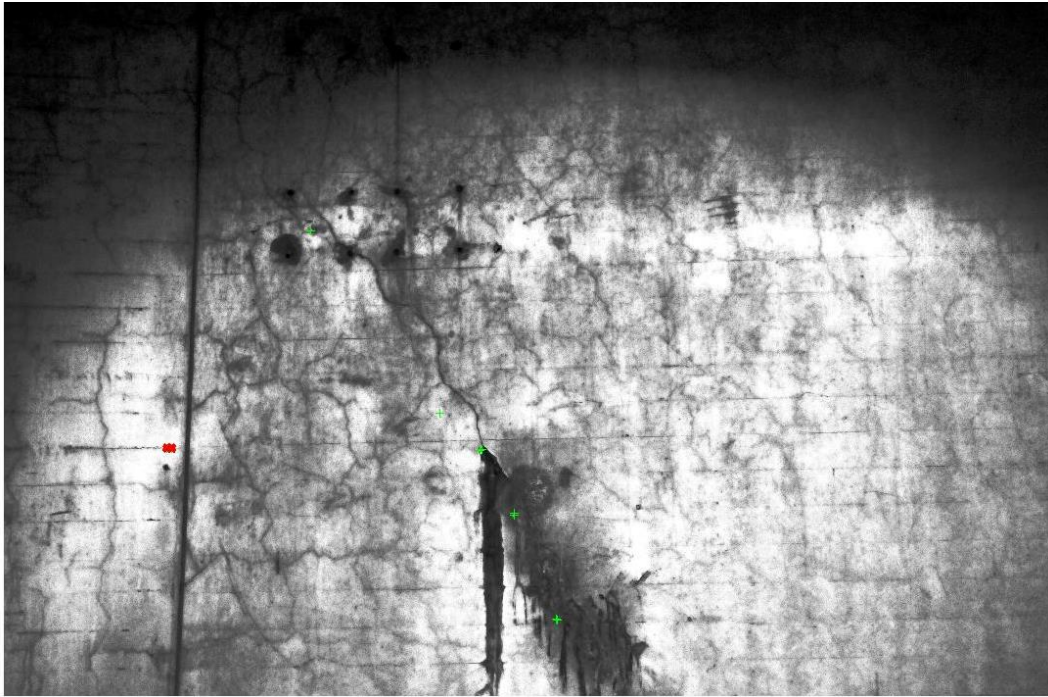
Σύμφωνα με το γράφημα 3-7 ο ανιχνευτής FAST για τις φωτογραφίες ρωγμών έχει το μεγαλύτερο μέσο Precision 40,39% στην φωτογραφία 11 (εικόνα 3.17). Αυτή η φωτογραφία έχει μεγαλύτερο Precision όταν ανιχνεύονται 5 ή 10 σημεία με Precision 80%. Το μέτρο αξιολόγησης αρχίζει να έχει φθίνουσα πορεία μετά και φτάνει την μικρότερη τιμή του στα 100 σημεία με 22%. Η χειρότερες φωτογραφίες ρωγμών για τον FAST είναι οι 2(βλέπε εικόνα 3.18),3,5,6,9,14 και 20 διότι δεν ανιχνεύουν κανένα σημείο ενδιαφέροντος. Η καλύτερη φωτογραφία θρυμματισμών είναι η 31 (εικόνα 3.19) με 99% μέσο Precision. Η οποία έχει υψηλή αποτελεσματικότητα και έχει 100% στις πιο πολλές περιπτώσεις ανίχνευσης. Στα 100 σημεία ανίχνευσης βρίσκει 99 σημεία. Η χειρότερη φωτογραφία για τις φωτογραφίες θρυμματισμού είναι η εικόνα 27 (εικόνα 3.20) με μέσο Precision 74,15%. Βρίσκει τη μικρότερη τιμή του Precision με 68,33% στα 60 σημεία ανίχνευσης διότι βρίσκει 41 σημεία ενδιαφέροντος.



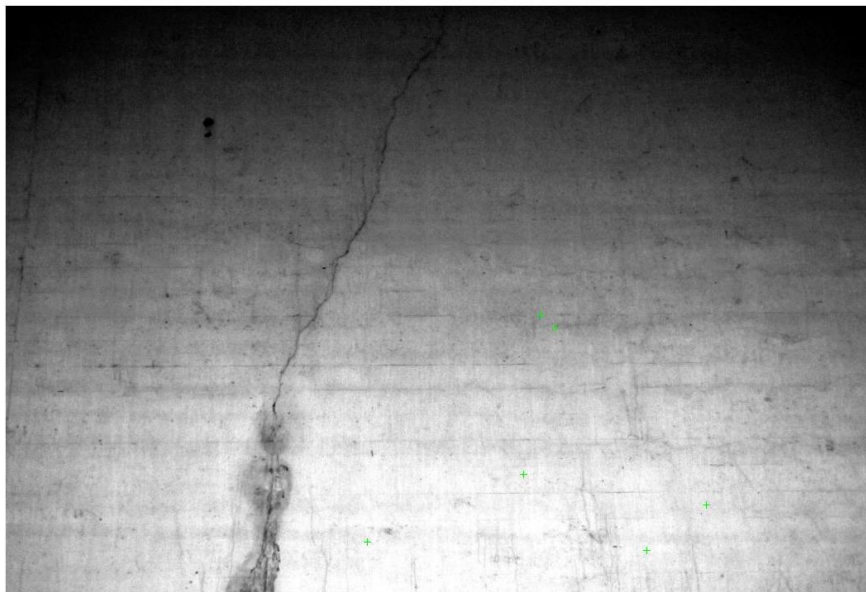
Γράφημα 3-7 : Αξιολόγηση του Fast ως προς το Precision για κάθε φωτογραφία



Γράφημα 3-8 : Αξιολόγηση του FAST ως προς το Recall για κάθε φωτογραφία



Εικόνα 3.17 : Η καλύτερη φωτογραφία ρωγμών για τον FAST είναι η 11 και έχει το μεγαλύτερο Precision όταν ανιχνεύονται 10 σημεία με Precision 80% . Αστοχούν σε αυτή την περίπτωση 2 σημεία όπως φαίνονται και με τους κόκκινους σταυρούς.



Εικόνα 3.18 : Μια από τις χειρότερες φωτογραφίες ρωγμών για τον ανιχνευτή FAST είναι η 2. Η οποία δεν βρίσκει κανένα σημείο ενδιαφέροντος.



Εικόνα 3.19 : Η καλύτερη φωτογραφία θρυμματισμών για τον FAST είναι η 31 και έχει πολλές φορές Precision 100%. Εδώ βλέπουμε όταν έχει 99% Precision στα 100 σημεία ανίχνευσης. Αστοχούν μόνο ένα σημείο(κόκκινος σταυρός) στο πάνω μέρος της φωτογραφίας.

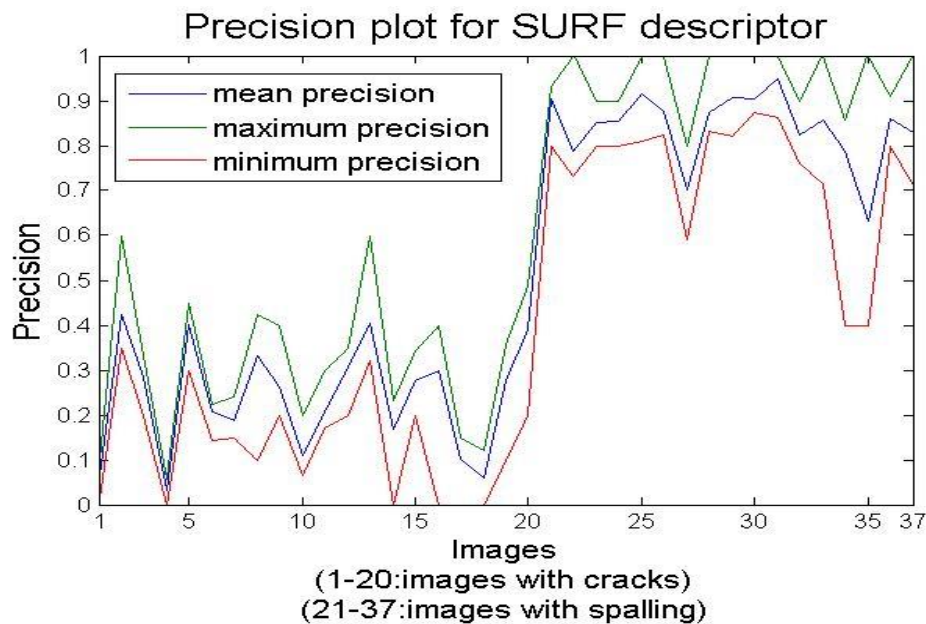


Εικόνα 3.20 : Η χειρότερη φωτογραφία θρυμματισμών για τον FAST είναι η 27 και έχει το μικρότερο Precision 68,33% στα 60 σημεία ανίχνευσης διότι βρίσκει 41 σημεία ενδιαφέροντος. Αστοχούν 19 σημεία(κόκκινοι σταυροί).

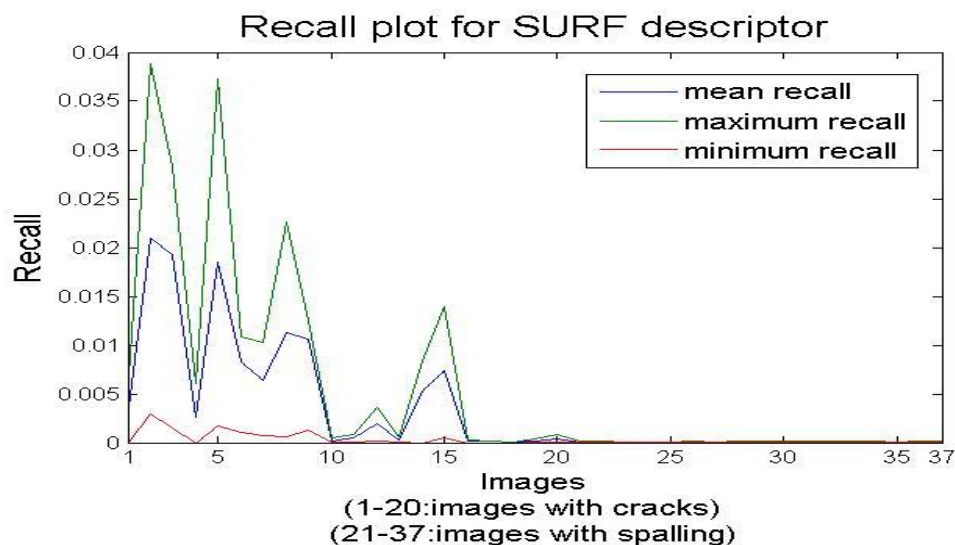
3.3.5 Ανιχνευτής SURF

Τέλος ο ανιχνευτής SURF σύμφωνα με το γράφημα 3-9 βρίσκει στις φωτογραφίες ρωγμών μεγαλύτερο μέσο Precision 42,34% στην φωτογραφία 2 (εικόνα 3.21). Εκεί έχει μεγαλύτερο Precision με τιμή 60% όταν βρίσκει 3 σημεία ενδιαφέροντος από τα 5 σημεία ανίχνευσης. Η

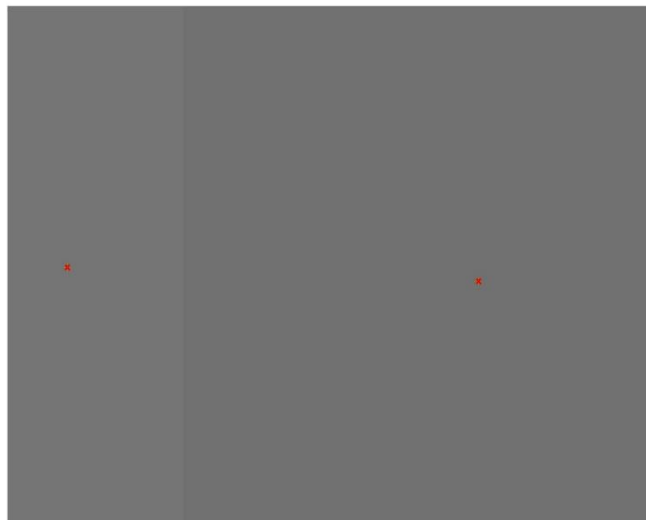
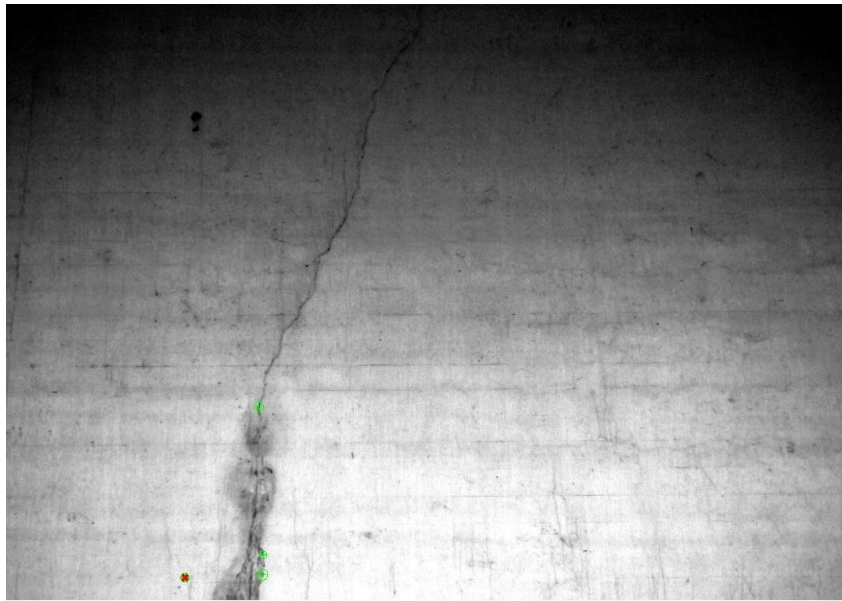
χειρότερη φωτογραφία με ρωγμές είναι η εικόνα 4 (εικόνα 3.22) με μέσο Precision 3,28%. Οι τιμές του Precision είναι 0% ή πολύ χαμηλές για διαφορετικό αριθμό σημείων ανίχνευσης. Η καλύτερη φωτογραφία θρυμματισμού είναι η 31 (εικόνα 3.23) με μέσο Precision 94,98%. Εμφανίζει 100% Precision σε αρκετές περιπτώσεις και έχει το μικρότερο Precision 86,32% όταν ανιχνεύει 95 σημεία και βρίσκει 82 σημεία ενδιαφέροντος. Η χειρότερη είναι η φωτογραφία 35 (εικόνα 3.24) με 63,20% μέσο Precision. Σε αυτή την περίπτωση έχουμε 100% Precision για σημεία ανίχνευσης 5 και 10. Όμως η ακρίβεια αρχίζει και έχει φθίνουσα πορεία με αποτέλεσμα να φτάνει το Precision στο 40% για τα 100 σημεία ανίχνευσης.



Γράφημα 3-9: Αξιολόγηση του SURF ως προς το Precision για κάθε φωτογραφία



Γράφημα 3-10: Αξιολόγηση του SURF ως προς το Recall για κάθε φωτογραφία



Εικόνα 3.21: Στην πάνω εικόνα φαίνεται η καλύτερη φωτογραφία ρωγμών για τον SURF η οποία είναι η 2 και έχει το μεγαλύτερο Precision 60% όταν ανιχνεύονται 5 σημεία. Αστοχούν σε αυτή την περίπτωση 2 σημεία όπως φαίνονται και με τους κόκκινους σταυρούς. Η κάτω εικόνα δείχνει τα σημεία πιο καθαρά βλέποντας στον μικρόκοσμο της φωτογραφίας.



Εικόνα 3.22 : Η χειρότερη φωτογραφία ρωγμών για τον SURF η οποία είναι η 4 και έχει πολύ μικρό μέσο Precision 3,28%. Εδώ βλέπουμε πως αντιδρά στα 100 σημεία ανίχνευσης. Βρίσκει 5 σημεία ενδιαφέροντος με Precision 5%. Τα πράσινα σημεία είναι αυτά που έχουν ευστοχήσει.

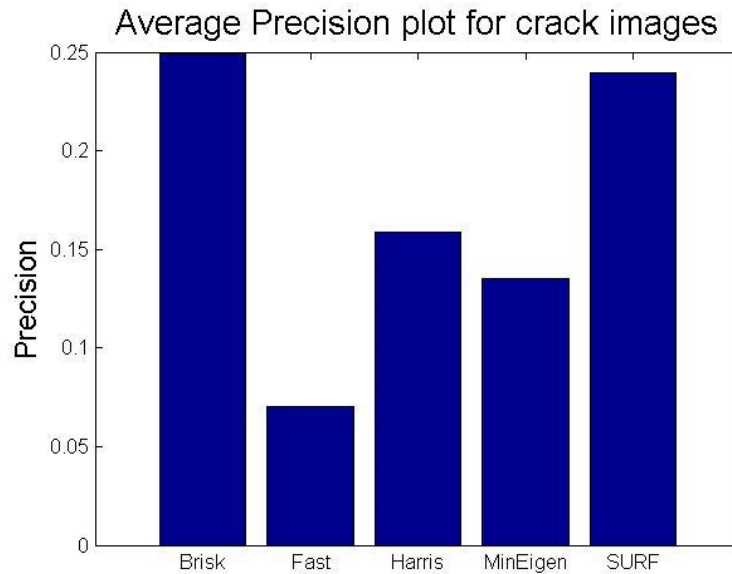


Εικόνα 3.23 : Η καλύτερη φωτογραφία θρυμματισμού για τον SURF είναι η 31. Έχει μέσο Precision 94,98 % και εμφανίζει 100% Precision σε αρκετές περιπτώσεις. Σε αυτή την εικόνα βλέπουμε όταν έχει το μικρότερο Precision 86,32% στα 95 σημεία ανίχνευσης και βρίσκει 82 σημεία ενδιαφέροντος. Τα 13 κόκκινα σημεία συμβολίζουν τις αστοχίες.



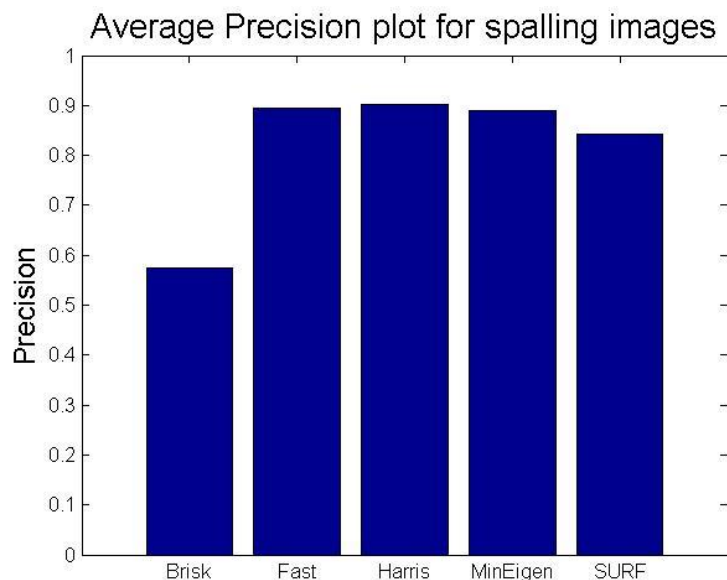
Εικόνα 3.24: Η χειρότερη φωτογραφία θρυμματισμών για τον SURF είναι η 35 με 63,20% μέσο Precision. Εδώ βλέπουμε όταν έχει το μικρότερο Precision 40% και το εμφανίζει στα 100 σημεία ανίχνευσης. Τα 40 πράσινα σημεία είναι τα σημεία που έχει ευστοχήσει.

Για να μπορέσουμε να συγκρίνουμε τους ανιχνευτές μεταξύ τους έπρεπε να πάρουμε τα συνολικά μέτρα αξιολόγησης για τον κάθε ανιχνευτή από όλες τις φωτογραφίες. Αυτό έγινε υπολογίζοντας τον μέσο όρο από τα μέσα μέτρα αξιολόγησης του κάθε ανιχνευτή. Επίσης κρίθηκε αναγκαίο πέρα από την αξιολόγηση των ανιχνευτών για όλες τις φωτογραφίες του δείγματος, να υπολογισθεί και η αξιολόγηση των ανιχνευτών ανάλογα με το ελάττωμα(ρωγμές ή θρυμματισμοί) που έχουν. Στο γράφημα 3-11 φαίνεται η αξιολόγηση των ανιχνευτών με βάση το Precision για τις φωτογραφίες που έχουν μόνο ρωγμές. Ο BRISK αναδείχθηκε κορυφαίος για τα ελαττώματα των ρωγμών με Precision 24,92%. Σε αντίθεση με τον FAST που είχε πολύ μικρή αξιολόγηση με 7,06% .



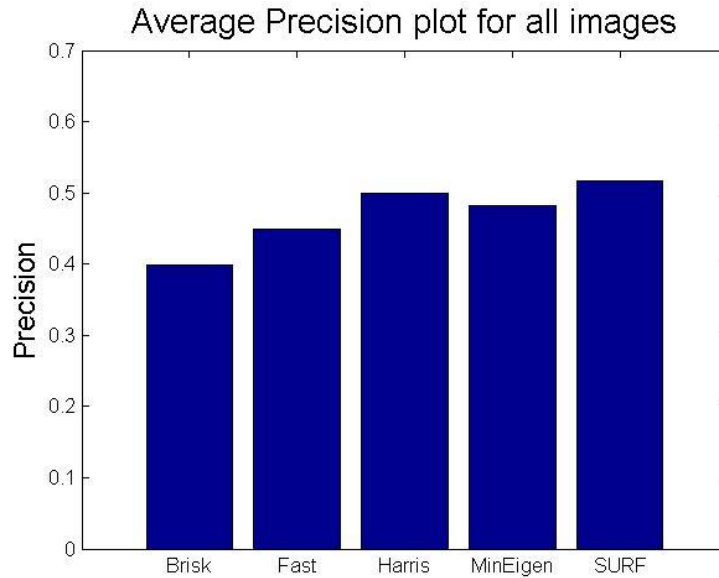
Γράφημα 3-11: Γενική αξιολόγηση των ανιχνευτών στις φωτογραφίες ρωγμών με βάση το Precision

Το γράφημα 3-12 παραθέτει την αξιολόγηση των ανιχνευτών στις φωτογραφίες θρυμματισμών με βάση το Precision. Παρατηρούμε σε αυτή την περίπτωση πολύ μεγαλύτερη αποτελεσματικότητα και αυτό οφείλεται όπως είχαμε προαναφέρει στο γεγονός ότι οι θρυμματισμοί έχουν μεγαλύτερη επιφάνεια από ότι οι ρωγμές και η υφή τους είναι ευδιάκριτη. Ο καλύτερος ανιχνευτής θρυμματισμών είναι ο Harris με Precision 90,19%. Ενώ ο χειρότερος με διαφορά από τους υπόλοιπους τέσσερεις είναι ο BRISK με τιμή 57,57%.



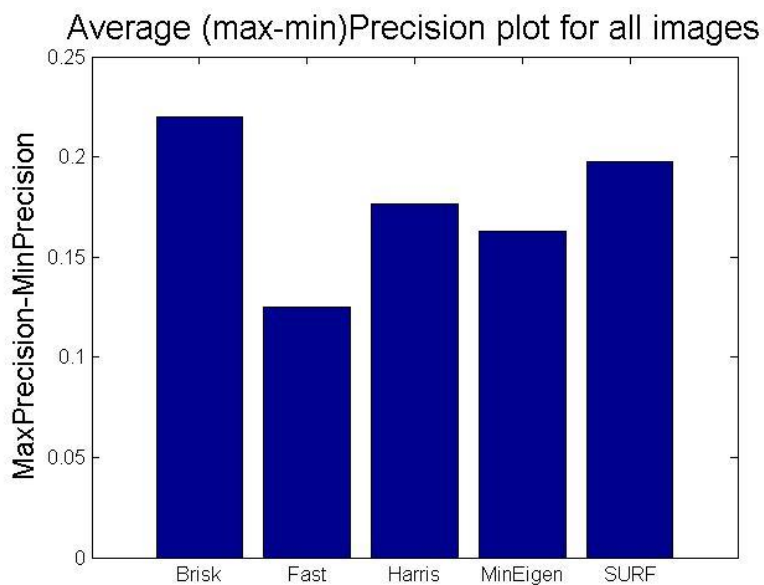
Γράφημα 3-12: Γενική αξιολόγηση των ανιχνευτών στις φωτογραφίες θρυμματισμών με βάση το Precision

Στο γράφημα 3-13 έχουμε την αξιολόγηση των ανιχνευτών για όλο το δείγμα των φωτογραφιών με βάση το Precision. Ο καλύτερος ανιχνευτής είναι ο SURF με 51,64 % ενώ ο χειρότερος ο BRISK με 39,92%.



Γράφημα 3-13: Γενική αξιολόγηση των ανιχνευτών για όλες τις φωτογραφίες του δείγματος με βάση το Precision

Επιπρόσθετα στο γράφημα 3-14 παρατηρείται η αστάθεια των ανιχνευτών υπολογίζοντας αρχικά το μέσο όρο των μέγιστων και των ελάχιστων του μέσου Precision κάθε φωτογραφίας. Και στην συνέχεια υπολογίζοντας την διαφορά τους. Την μεγαλύτερη αστάθεια την έχει ο BRISK με την τιμή αυτού του μεγέθους να φτάνει στο 22%. Ενώ την μικρότερη την είχε ο FAST με 12,48%.



Γράφημα 3-14: Γενική αξιολόγηση των ανιχνευτών για όλες τις φωτογραφίες του δείγματος με βάση την αστάθεια

3.4 Συμπεράσματα

Ο στόχος της παρούσας διπλωματικής εργασίας ήταν η ανίχνευση ρωγμών και θρυμματισμών σε τούνελ από σκυρόδεμα, με τη χρήση πέντε διαφορετικών ανιχνευτών από την εργαλειοθήκη της Matlab. Τα δεδομένα που χρησιμοποιήθηκαν ήταν 37 φωτογραφίες ρωγμών και θρυμματισμών, καθώς και 37 αντίστοιχες σχολιαστικές εικόνες (annotation images), που εμπεριέχουν με ακρίβεια, με την χρήση του λευκού χρώματος, τις ακριβείς συντεταγμένες των ελαττωμάτων που μας ενδιαφέρουν. Τα δεδομένα αυτά πάρθηκαν από το τούνελ V-S-H, που βρίσκεται στην Ελβετία.

Ύστερα από δοκιμές, αποφασίστηκε το είδος των φίλτρων που χρησιμοποιήθηκε στον τελικό αλγόριθμο. Πρόκειται για φίλτρα επεξεργασίας και τον καθαρισμού (αφαίρεση θορύβου) των φωτογραφιών, προκειμένου αυτές να αποτελέσουν κατάλληλες εισαγωγές στους ανιχνευτές, για το καλύτερο τελικό επιθυμητό αποτέλεσμα. Τα φίλτρα που χρησιμοποιήθηκαν είναι το μέσο φίλτρο medfilt2 και histeq. Για την αξιολόγηση των πέντε ανιχνευτών, υπολογίστηκαν στον τελικό αλγόριθμο, δύο μέτρα αξιολόγησης: η ακρίβεια (Precision) και το μέτρο της ευαισθησίας (Recall).

Τα αποτελέσματα του αλγορίθμου με βάση το μέτρο αξιολόγησης Precision έδειξαν, ότι είναι προτιμότερο, σε ρωγμές, να χρησιμοποιείται ο ανιχνευτής BRISK και σε θρυμματισμούς ο ανιχνευτής Harris. Συνολικά, σε όλες τις φωτογραφίες ο καλύτερος ανιχνευτής ήταν ο SURF.

Το ζήτημα της ανίχνευσης ελαττωμάτων σε τούνελ μπορεί να έχει ποικίλες προεκτάσεις. Θα μπορούσε να εξεταστεί η δυνατότητα εφαρμογής ενός συνδυασμού ανιχνευτών σε τούνελ, των οποίων η επένδυση, δεν αποτελείται μόνο από σκυρόδεμα, αλλά και από άλλα υλικά και ποικίλες αποχρώσεις.

Επίσης, με τη χρήση καλύτερων φωτογραφικών μηχανών, αλλά και με τη συμβολή αποτελεσματικότερων μηχανισμών φιλτραρίσματος, θα ήταν δυνατή η εισαγωγή ποιοτικότερων δεδομένων στα συστήματα ανίχνευσης.

Η διενέργεια ανίχνευσης ελαττωματικών σημείων, είτε σε τούνελ, είτε σε άλλες κατασκευές, θα μπορούσε να έχει καλύτερα αποτελέσματα, με τη χρήση νευρωνικών δικτύων, τα οποία, βάσει της μηχανικής εκμάθησης, «εκπαιδεύουν», με κατάλληλο τρόπο, τον εκάστοτε αλγόριθμο, προκειμένου να ενισχυθεί η αποτελεσματικότητα της διαδικασίας.

Βιβλιογραφία

- [1] Edge detection, Tony Lindeberg, Encyclopedia of Mathematics
- [2] J. Canny (1986) "A computational approach to edge detection", IEEE Trans. Pattern Analysis and Machine Intelligence, vol 8, pages 679-714.
- [3] Shapiro, Linda and George C. Stockman (2001). *Computer Vision*, p. 257. Prentice Books, Upper Saddle River.
- [4] https://en.wikipedia.org/wiki/Ridge_detection
- [5] Richard Szeliski, βιβλίο Computer Vision: Algorithms and Applications σελίδες (3-10), (2010)
- [6] Chris Harris, Mike Stephens, A Combined Corner and Edge Detector, Proceedings of the 4th Alve5y Vision Conference, (1988).
- [7] M. Gevrekci, B.K. Gunturk, Illumination robust interest point detection, Comput. Vis. Image Understand. (2009)
- [8] Leow Wee Kheng, Computer Vision and Pattern Recognition
- [9] Davide Scaramuzza, Lecture 05 Feature Extraction 2
- [10] Frank Nielsen, Harris-Stephens' combined corner/edge detector
- [11] J. Shi, C. Tomasi, "Good features to track," in *1994 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1994. Proceedings CVPR'94.*, pp. 593–600, 1994.
- [12] Anil Bharath, Maria Petrou, Next Generation Artificial Vision Systems : Reverse Engineering the Human Visual System, (2008).
- [13] Voser Michael, Real-Time Visual Object Discovery for Tracking Using Motion and Depth Data, ETH, (2011)
- [14] Rosten Edward., Drummond Tom., *Machine learning for high speed corner detection*, στο Ales Leonardis, Horst Bischof, Axel Pinz (eds.), *Computer Vision - ECCV 2006_9th European Conference on Computer Vision*, vol. 1, Springer Publications, Berlin, 2006, σελ.: 430–443.
- [15] Deepak Geetha Viswanathan, Features from Accelerated Segment Test (FAST)

- [16] Rosten Edward., Drummond Tom., “Fusing points and lines for high performance tracking”, στο Proc. Int. Conf. Computer Vision, October 2005, vol. 2, pp. 1508–1511.
- [17] Elmar Mair, Gregory D. Hager, Darius Burschka, Michael Suppa, και Gerhard Hirzinger, Adaptive and Generic Corner Detection Based on the Accelerated Segment Test
- [18] https://en.wikipedia.org/wiki/Midpoint_circle_algorithm
- [19] https://en.wikipedia.org/wiki/Features_from_accelerated_segment_test
- [20] H. Bay, T. Tuytelaars, L. Van Gool. Surf: Speeded Up Robust Features(SURF). (2006)
- [21] P.A. Viola and M.J. Jones. Rapid Object Detection using a Boosted Cascade of Simple Features. pages 511 – 518. (2001)
- [22] H. Bay, T. Tuytelaars, , Andreas Ess, L. Van Gool. Surf: Speeded Up Robust Features(SURF),(revised publication) (2008)
- [23] Jacob Toft Pedersen, Study group SURF: Feature detection & description, (2011)
- [24] Ville Kangas, COMPARISON OF LOCAL FEATURE DETECTORS AND DESCRIPTORS FOR VISUAL OBJECT CATEGORIZATION,(2011)
- [25] Stefan Leutenegger, Margarita Chli and Roland Y. Siegwart, BRISK: Binary Robust Invariant Scalable Keypoints,(2011)
- [26] ROBotic System with Intelligent Vision and Control for Tunnel Structural INSPECTion and Evaluation "(2013-08-06)
- [27] INFLUENCE OF STONES' SIZE ON THE COLLAPSE OF MASONRY BRIDGES Borja Conde , Georgios A. Drosopoulos , Belen Riveiro, Maria E. Stavroulaki and Georgios E. Stavroulakis (8th GRACM International Congress on Computational Mechanics Volos, 12 July – 15 July 2015)
- [28] Doulamis, A. (2010b). “Dynamic Tracking Re-Adjustment: A Method for Automatic Tracking Recovery in Complex Visual Environments,” Multimedia Tools and Applications, 50 (1), pp. 49-73.

Παράρτημα

Αλγόριθμος με επιλογή τον BRISK

```
tunnel=cell(1,37);
for z=1:37
tunnel{z}=imread(sprintf('im%d.jpg',z));
sgrtunnel=rgb2gray(tunnel{z});
x=0;%arithmos strongest shmeiwn
sgrtunnel = histeq(sgrtunnel);
sgrtunnel = medfilt2(sgrtunnel);
figure, imshow(sgrtunnel);
anot=cell(1,37);
anot{z}=imread(sprintf('ima%d.jpg',z));
anotation=rgb2gray(anot{z});
row=1;
[m,n]=size(anotation);
for i=1:m %pinakas matrix pou emperiexei ta cracks/spalling ths annotation
eikonas
for j=1:n
if anotation(i,j)==255;
matrix(row,1)=j;
matrix(row,2)=i;
row=row+1;
end
end
end
t=1;
strongest100_vima5=5:5:100;
for x=strongest100_vima5

%%
points = detectBRISKFeatures(sgrtunnel,'MinQuality',0.1);
figure('Name','brisk'), imshow(sgrtunnel); hold on;
plot(points.selectStrongest(x));

%%
a=1;
w=1;
rr=round(points.selectStrongest(x).Location);
myindex=ones(length(rr),1)*(-1);

for i=1:size(matrix,1) %syntetagmenes tw n shmeiwn pou vriskoun oi detectors
mesa sthn perioxh pou yparxoun ta cracks/spalling
for w=1:size(rr,1)

if ((rr(w,1) <matrix(i,1)+120) && (rr(w,1)> matrix(i,1)-120)) && ((rr(w,2)
<matrix(i,2)+120) && (rr(w,2)> matrix(i,2)-120)) && (myindex(w)==-1)
shmeia(a,1,t)=rr(w,1);
syntetagmenes_shmeiwn(a,1,t,z)=shmeia(a,1,t);
shmeia(a,2,t)=rr(w,2);
```



```

syntetagmenes_shmeiwn(a,2,t,z)=shmeia(a,2,t);
a=a+1;
myindex(w)=1;

end
end

end
logiko = exist('shmeia','var');
if (logiko) == 0
ar_shm_cracks(t,z)=0;
precision(t,z)=0;
recall(t,z)=0;
end
if (logiko) == 1
ar_shm_cracks(t,z)=size(shmeia,1);
precision(t,z)=ar_shm_cracks(t,z)/size(rr,1);
recall(t,z)=ar_shm_cracks(t,z)/size(matrix,1);
end
t=t+1;
end
figure('Name','Brisk Precision plot for 100 strongest points')
plot(strongest100_vima5,precision(:,z))
figure('Name','Brisk Recall plot for 100 strongest points')
plot(strongest100_vima5,recall(:,z))
tempp=mean(recall);
mean_recall(z)=tempp(z);

tempp=mean(precision);
mean_precision(z)=tempp(z);

tempp=min(recall);
min_recall(z)=tempp(z);

tempp=min(precision);
min_precision(z)=tempp(z);

tempp=max(recall);
max_recall(z)=tempp(z);

tempp=max(precision);
max_precision(z)=tempp(z);
clearvars -except syntetagmenes_shmeiwn ar_shm_cracks precision recall
tunnel anot mean_recall mean_precision min_recall min_precision max_recall
max_precision
end
mean_precision_cracks=mean_precision(1:20);
mean_precision_spalling=mean_precision(21:37);

max_precision_cracks=max_precision(1:20);
max_precision_spalling=max_precision(21:37);

min_precision_cracks=min_precision(1:20);
min_precision_spalling=min_precision(21:37);

mean_recall_cracks=mean_recall(1:20);
mean_recall_spalling=mean_recall(21:37);

max_recall_cracks=max_recall(1:20);
max_recall_spalling=max_recall(21:37);

```

```
min_recall_cracks=min_recall(1:20);
min_recall_spalling=min_recall(21:37);

meanprecision_brisk=mean(mean_precision);
meanprecision_brisk_cracks=mean(mean_precision_cracks);
meanprecision_brisk_spalling=mean(mean_precision_spalling);
meanrecall_brisk=mean(mean_recall);
meanrecall_brisk_cracks=mean(mean_recall_cracks);
meanrecall_brisk_spalling=mean(mean_recall_spalling);
```

Αλγόριθμος με επιλογή του Harris

```
tunnel=cell(1,37);
for z=1:37
    tunnel{z}=imread(sprintf('im%d.jpg',z));
    sgrtunnel=rgb2gray(tunnel{z});
    x=0;%arithos strongest shmeiwn
    sgrtunnel = histeq(sgrtunnel);
    sgrtunnel = medfilt2(sgrtunnel);
    figure, imshow(sgrtunnel);
    anot=cell(1,37);
    anot{z}=imread(sprintf('ima%d.jpg',z));
    anotation=rgb2gray(anot{z});
    row=1;
    [m,n]=size(anotation);
    for i=1:m %pinakas matrix pou emperiexei ta cracks/spalling ths
        annotation eikonas
            for j=1:n
                if anotation(i,j)==255;
                    matrix(row,1)=j;
                    matrix(row,2)=i;
                    row=row+1;
                end
            end
        end
    end
    t=1;
    strongest100_vima5=5:5:100;
    for x=strongest100_vima5

        %%
        points = detectHarrisFeatures(sgrtunnel);
        figure('Name','harris'),imshow(sgrtunnel); hold on;
        plot(points.selectStrongest(x));
        %%
        a=1;
        w=1;
        rr=round(points.selectStrongest(x).Location);
        myindex=ones(length(rr),1)*(-1);

        for i=1:size(matrix,1) %syntetagmenes tw n shmeiwn pou vriskoun oi
            detectors mesa sthn perioxh pou ypaxoun ta cracks/spalling
                for w=1:size(rr,1)
```

```

        if ((rr(w,1) < matrix(i,1)+120) && (rr(w,1) > matrix(i,1)-120))
&& ((rr(w,2) < matrix(i,2)+120) && (rr(w,2) > matrix(i,2)-120))
&& (myindex(w)==-1)
            shmeia(a,1,t)=rr(w,1);
            syntetagmenes_shmeiwn(a,1,t,z)=shmeia(a,1,t);
            shmeia(a,2,t)=rr(w,2);
            syntetagmenes_shmeiwn(a,2,t,z)=shmeia(a,2,t);
            a=a+1;
            myindex(w)=1;

        end
    end

end

logiko = exist('shmeia','var');
if (logiko) == 0
    ar_shm_cracks(t,z)=0;
    precision(t,z)=0;
    recall(t,z)=0;
end
if (logiko) == 1
    ar_shm_cracks(t,z)=size(shmeia,1);
    precision(t,z)=ar_shm_cracks(t,z)/size(rr,1);
    recall(t,z)=ar_shm_cracks(t,z)/size(matrix,1);
end
t=t+1;
end
figure('Name','Harris Precision plot for 100 strongest points')
plot(strongest100_vima5,precision(:,z))
figure('Name','Harris Recall plot for 100 strongest points')
plot(strongest100_vima5,recall(:,z))
tempp=mean(recall);
mean_recall(z)=tempp(z);

tempp=mean(precision);
mean_precision(z)=tempp(z);

tempp=min(recall);
min_recall(z)=tempp(z);

tempp=min(precision);
min_precision(z)=tempp(z);

tempp=max(recall);
max_recall(z)=tempp(z);

tempp=max(precision);
max_precision(z)=tempp(z);
clearvars -except syntetagmenes_shmeiwn ar_shm_cracks precision recall
tunnel anot mean_recall mean_precision min_recall min_precision max_recall
max_precision
end
mean_precision_cracks=mean_precision(1:20);
mean_precision_spalling=mean_precision(21:37);

max_precision_cracks=max_precision(1:20);
max_precision_spalling=max_precision(21:37);

min_precision_cracks=min_precision(1:20);
min_precision_spalling=min_precision(21:37);

```

```

mean_recall_cracks=mean_recall(1:20);
mean_recall_spalling=mean_recall(21:37);

max_recall_cracks=max_recall(1:20);
max_recall_spalling=max_recall(21:37);

min_recall_cracks=min_recall(1:20);
min_recall_spalling=min_recall(21:37);

meanprecision_harris=mean(mean_precision);
meanprecision_harris_cracks=mean(mean_precision_cracks);
meanprecision_harris_spalling=mean(mean_precision_spalling);
meanrecall_harris=mean(mean_recall);
meanrecall_harris_cracks=mean(mean_recall_cracks);
meanrecall_harris_spalling=mean(mean_recall_spalling);

```

Αλγόριθμος με επιλογή τον Shi-Tomasi(MinEigen)

```

tunnel=cell(1,37);
for z=1:37
    tunnel{z}=imread(sprintf('im%d.jpg',z));
    sgrtunnel=rgb2gray(tunnel{z});
    x=0;%arithos strongest shmeiwn
    sgrtunnel = histeq(sgrtunnel);
    sgrtunnel = medfilt2(sgrtunnel);
    figure, imshow(sgrtunnel);
    anot=cell(1,17);
    anot{z}=imread(sprintf('ima%d.jpg',z));
    anotation=rgb2gray(anot{z});
    row=1;
    [m,n]=size(anotation);
    for i=1:m %pinakas matrix pou emperiexei ta cracks/spalling ths
        annotation eikonas
            for j=1:n
                if anotation(i,j)==255;
                    matrix(row,1)=j;
                    matrix(row,2)=i;
                    row=row+1;
                end
            end
        end
    end
    t=1;
    strongest100_vima5=5:5:100;
    for x=strongest100_vima5

        %%
        points = detectMinEigenFeatures(sgrtunnel);
        figure('Name','mineigen'),imshow(sgrtunnel); hold on;
        plot(points.selectStrongest(x));

        %%
        a=1;
        w=1;
        rr=round(points.selectStrongest(x).Location);
        myindex=ones(length(rr),1)*(-1);
    end
end

```

```

for i=1:size(matrix,1) %syntetagmenes tw n shmeiwn pou vriskoun oi
detectors mesa sthn perioxh pou yparxoun ta cracks/spalling
    for w=1:size(rr,1)

        if ((rr(w,1) <matrix(i,1)+120) && (rr(w,1)> matrix(i,1)-120))
&& ((rr(w,2) <matrix(i,2)+120) && (rr(w,2)> matrix(i,2)-120))
&&(myindex(w)==-1)

            shmeia(a,1,t)=rr(w,1);
            syntetagmenes_shmeiwn(a,1,t,z)=shmeia(a,1,t);
            shmeia(a,2,t)=rr(w,2);
            syntetagmenes_shmeiwn(a,2,t,z)=shmeia(a,2,t);
            a=a+1;
            myindex(w)=1;

        end
    end

end

logiko = exist('shmeia','var');
if (logiko) == 0
    ar_shm_cracks(t,z)=0;
    precision(t,z)=0;
    recall(t,z)=0;
end
if (logiko) == 1
    ar_shm_cracks(t,z)=size(shmeia,1);
    precision(t,z)=ar_shm_cracks(t,z)/size(rr,1);
    recall(t,z)=ar_shm_cracks(t,z)/size(matrix,1);
end
t=t+1;

end
figure('Name','MinEigen Precision plot for 100 strongest points')
plot(strongest100_vima5,precision(:,z))
figure('Name','MinEigen Recall plot for 100 strongest points')
plot(strongest100_vima5,recall(:,z))
tempp=mean(recall);
mean_recall(z)=tempp(z);

tempp=mean(precision);
mean_precision(z)=tempp(z);

tempp=min(recall);
min_recall(z)=tempp(z);

tempp=min(precision);
min_precision(z)=tempp(z);

tempp=max(recall);
max_recall(z)=tempp(z);

tempp=max(precision);
max_precision(z)=tempp(z);
clearvars -except syntetagmenes_shmeiwn ar_shm_cracks precision recall
tunnel anot mean_recall mean_precision min_recall min_precision max_recall
max_precision
end
mean_precision_cracks=mean_precision(1:20);
mean_precision_spalling=mean_precision(21:37);

max_precision_cracks=max_precision(1:20);

```

```

max_precision_spalling=max_precision(21:37);

min_precision_cracks=min_precision(1:20);
min_precision_spalling=min_precision(21:37);

mean_recall_cracks=mean_recall(1:20);
mean_recall_spalling=mean_recall(21:37);

max_recall_cracks=max_recall(1:20);
max_recall_spalling=max_recall(21:37);

min_recall_cracks=min_recall(1:20);
min_recall_spalling=min_recall(21:37);

meanprecision_MinEigen=mean(mean_precision);
meanprecision_MinEigen_cracks=mean(mean_precision_cracks);
meanprecision_MinEigen_spalling=mean(mean_precision_spalling);
meanrecall_MinEigen=mean(mean_recall);
meanrecall_MinEigen_cracks=mean(mean_recall_cracks);
meanrecall_MinEigen_spalling=mean(mean_recall_spalling);

```

Αλγόριθμος με επιλογή τον FAST

```

tic;
tunnel=cell(1,37);
for z=1:37
    tunnel{z}=imread(sprintf('im%d.jpg',z));
    sgrtunnel=rgb2gray(tunnel{z});
    x=0;%arithos strongest shmeiwn
    sgrtunnel = histeq(sgrtunnel);
    sgrtunnel = medfilt2(sgrtunnel);
    figure, imshow(sgrtunnel);
    anot=cell(1,37);
    anot{z}=imread(sprintf('ima%d.jpg',z));
    anotation=rgb2gray(anot{z});
    row=1;
    [m,n]=size(anotation);
    for i=1:m %pinakas matrix pou emperiexei ta cracks/spalling ths
        annotation eikonas
            for j=1:n
                if anotation(i,j)==255;
                    matrix(row,1)=j;
                    matrix(row,2)=i;
                    row=row+1;
                end
            end
        end
    end
    t=1;
    strongest100_vima5=5:5:100;
    for x=strongest100_vima5

        %%
        points = detectFASTFeatures(sgrtunnel);
        figure('Name','fast detection'), imshow(sgrtunnel); hold on;
        plot(points.selectStrongest(x));

        %%
        a=1;
        w=1;
    end
end

```

```

rr=round(points.selectStrongest(x).Location);
myindex=ones(length(rr),1)*(-1);

for i=1:size(matrix,1) %syntetagmenes tw n shmeiwn pou vriskoun oi
detectors mesa sthn perioxh pou yparxoun ta cracks/spalling
    for w=1:size(rr,1)

        if ((rr(w,1) <matrix(i,1)+120) && (rr(w,1)> matrix(i,1)-120))
&& ((rr(w,2) <matrix(i,2)+120) && (rr(w,2)> matrix(i,2)-120))
&&(myindex(w)==-1)
            shmeia(a,1,t)=rr(w,1);
            syntetagmenes_shmeiwn(a,1,t,z)=shmeia(a,1,t);
            shmeia(a,2,t)=rr(w,2);
            syntetagmenes_shmeiwn(a,2,t,z)=shmeia(a,2,t);
            a=a+1;
            myindex(w)=1;

        end
    end

end

logiko = exist('shmeia','var');
if (logiko) == 0
    ar_shm_cracks(t,z)=0;
    precision(t,z)=0;
    recall(t,z)=0;
end
if (logiko) == 1
    ar_shm_cracks(t,z)=size(shmeia,1);
    precision(t,z)=ar_shm_cracks(t,z)/size(rr,1);
    recall(t,z)=ar_shm_cracks(t,z)/size(matrix,1);
end
t=t+1;

end
figure('Name','Fast Precision plot for 100 strongest points')
plot(strongest100_vima5,precision(:,z))
figure('Name','Fast Recall plot for 100 strongest points')
plot(strongest100_vima5,recall(:,z))
tempp=mean(recall);
mean_recall(z)=tempp(z);

tempp=mean(precision);
mean_precision(z)=tempp(z);

tempp=min(recall);
min_recall(z)=tempp(z);

tempp=min(precision);
min_precision(z)=tempp(z);

tempp=max(recall);
max_recall(z)=tempp(z);

tempp=max(precision);
max_precision(z)=tempp(z);
clearvars -except syntetagmenes_shmeiwn ar_shm_cracks precision recall
tunnel anot mean_recall mean_precision min_recall min_precision max_recall
max_precision
end
mean_precision_cracks=mean_precision(1:20);

```

```

mean_precision_spalling=mean_precision(21:37);

max_precision_cracks=max_precision(1:20);
max_precision_spalling=max_precision(21:37);

min_precision_cracks=min_precision(1:20);
min_precision_spalling=min_precision(21:37);

mean_recall_cracks=mean_recall(1:20);
mean_recall_spalling=mean_recall(21:37);

max_recall_cracks=max_recall(1:20);
max_recall_spalling=max_recall(21:37);

min_recall_cracks=min_recall(1:20);
min_recall_spalling=min_recall(21:37);

meanprecision_fast=mean(mean_precision);
meanprecision_fast_cracks=mean(mean_precision_cracks);
meanprecision_fast_spalling=mean(mean_precision_spalling);
meanrecall_fast=mean(mean_recall);
meanrecall_fast_cracks=mean(mean_recall_cracks);
meanrecall_fast_spalling=mean(mean_recall_spalling);

```

Αλγόριθμος με επιλογή τον SURF

```

tunnel=cell(1,37);
for z=1:37
    tunnel{z}=imread(sprintf('im%d.jpg',z));
    sgrtunnel=rgb2gray(tunnel{z});
    x=0;%arithos strongest shmeiwn
    sgrtunnel = histeq(sgrtunnel);
    sgrtunnel = medfilt2(sgrtunnel);
    figure, imshow(sgrtunnel);
    anot=cell(1,17);
    anot{z}=imread(sprintf('ima%d.jpg',z));
    anotation=rgb2gray(anot{z});
    row=1;
    [m,n]=size(anotation);
    for i=1:m %pinakas matrix pou emperiexei ta cracks/spalling ths
        annotation eikonas
            for j=1:n
                if anotation(i,j)==255;
                    matrix(row,1)=j;
                    matrix(row,2)=i;
                    row=row+1;
                end
            end
        end
    end
    t=1;
    strongest100_vima5=5:5:100;
    for x=strongest100_vima5

        %%
        points = detectSURFFeatures(sgrtunnel) ;
        figure('Name','surf'),imshow(sgrtunnel); hold on;
        plot(points.selectStrongest(x));
    end
end

```



```

%%
a=1;
w=1;
rr=round(points.selectStrongest(x).Location);
myindex=ones(length(rr),1)*(-1);

for i=1:size(matrix,1) %syntetagmenes tw'n shmeiwn pou vriskoun oi
detectors mesa sthn perioxh pou yparxoun ta cracks/spalling
    for w=1:size(rr,1)

        if ((rr(w,1) <matrix(i,1)+120) && (rr(w,1)> matrix(i,1)-120))
&& ((rr(w,2) <matrix(i,2)+120) && (rr(w,2)> matrix(i,2)-120))
&& (myindex(w)==-1)
            shmeia(a,1,t)=rr(w,1);
            syntetagmenes_shmeiwn(a,1,t,z)=shmeia(a,1,t);
            shmeia(a,2,t)=rr(w,2);
            syntetagmenes_shmeiwn(a,2,t,z)=shmeia(a,2,t);
            a=a+1;
            myindex(w)=1;

        end
    end

end

logiko = exist('shmeia','var');
if (logiko) == 0
    ar_shm_cracks(t,z)=0;
    precision(t,z)=0;
    recall(t,z)=0;
end
if (logiko) == 1
    ar_shm_cracks(t,z)=size(shmeia,1);
    precision(t,z)=ar_shm_cracks(t,z)/size(rr,1);
    recall(t,z)=ar_shm_cracks(t,z)/size(matrix,1);
end
t=t+1;
end

figure('Name','SURF Precision plot for 100 strongest points')
plot(strongest100_vima5,precision(:,z))
figure('Name','SURF Recall plot for 100 strongest points')
plot(strongest100_vima5,recall(:,z))
tempp=mean(recall);
mean_recall(z)=tempp(z);

tempp=mean(precision);
mean_precision(z)=tempp(z);

tempp=min(recall);
min_recall(z)=tempp(z);

tempp=min(precision);
min_precision(z)=tempp(z);

tempp=max(recall);
max_recall(z)=tempp(z);

tempp=max(precision);
max_precision(z)=tempp(z);

```

```
clearvars -except syntetagmenes_shmeiwn ar_shm_cracks precision recall
tunnel anot mean_recall mean_precision min_recall min_precision max_recall
max_precision
end
mean_precision_cracks=mean_precision(1:20);
mean_precision_spalling=mean_precision(21:37);

max_precision_cracks=max_precision(1:20);
max_precision_spalling=max_precision(21:37);

min_precision_cracks=min_precision(1:20);
min_precision_spalling=min_precision(21:37);

mean_recall_cracks=mean_recall(1:20);
mean_recall_spalling=mean_recall(21:37);

max_recall_cracks=max_recall(1:20);
max_recall_spalling=max_recall(21:37);

min_recall_cracks=min_recall(1:20);
min_recall_spalling=min_recall(21:37);

meanprecision_SURF=mean(mean_precision);
meanprecision_SURF_cracks=mean(mean_precision_cracks);
meanprecision_SURF_spalling=mean(mean_precision_spalling);
meanrecall_SURF=mean(mean_recall);
meanrecall_SURF_cracks=mean(mean_recall_cracks);
meanrecall_SURF_spalling=mean(mean_recall_spalling);
```

