

Τεχνολογία Και Ανάλυση Εικόνων και Βίντεο

Άσκηση 1

Επεξεργασία εικόνας, φίλτρα, ακμές και εκτίμηση κίνησης

*Γαλάνης Δημήτριος (03116088)
Χαρίτου Δανάη-Νικολέτα (03116045)*

Ερώτημα 2

Στο ερώτημα αυτό γίνεται χρήση αλγορίθμων εξαγωγής χαρακτηριστικών, καθώς και τεχνικών εκτίμησης κίνησης με σκοπό την παρακολούθηση κινούμενων στοιχείων σε surveillance videos.

1. Εισαγωγή video

Σύμφωνα με την δοθείσα λίστα, το video που αντιστοιχεί στην ομάδα μας είναι το εξής:

[VIRAT_S_010000_03_000442_000528.mp4](#)

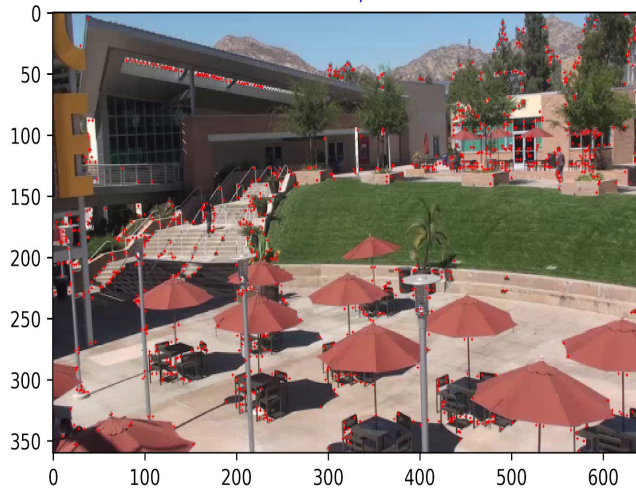
2. Επιλογή κατάλληλου χρωματικού χώρου

Για την επεξεργασία που πραγματοποιούμε στο εκάστοτε frame του προαναφερθέντος video, επιλέγουμε τον χρωματικό χώρο Grayscale, εφόσον κάθε εικόνα αναπαρίσταται με την βοήθεια ενός δισδιάστατου πίνακα, καθιστώντας όλη την διαδικασία αρκετά πιο εύκολη υπολογιστικά και πρακτικά, ενώ συνάμα η ‘εξάλειψη’ των χρωματικών συνιστωσών δεν επηρεάζει την επίδοση των εφαρμοζόμενων αλγορίθμων.

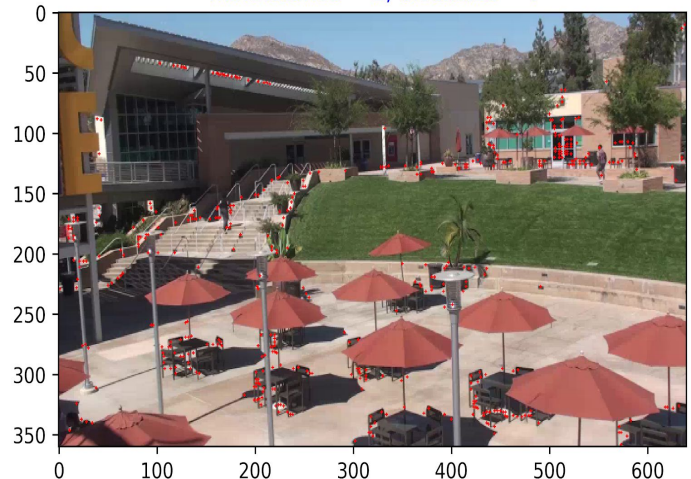
3. Εφαρμογή Shi-Tomasi και Harris corner Detectors

Στην συνέχεια παραθέτουμε ενδεικτικά αποτελέσματα που προέκυψαν κατά τον πειραματισμό μας με τις τιμές των παραμέτρων *maxCorners*, *qualityLevel*, *minDistance* των δύο ανιχνευτών με στόχο την καλύτερη δυνατή περιγραφή των αντικειμένων στο πρώτο frame του δοθέντος video.

Shai-Tomasi detector using maxCorners = 2000 , qualityLevel = 0.1
minDistance = 2, blockSize = 7

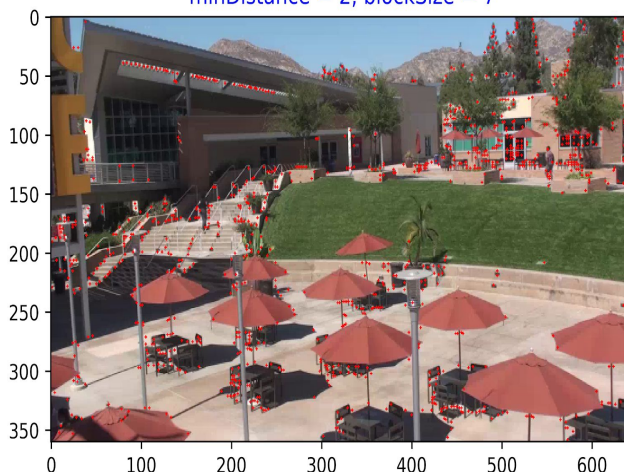


Harris detector using maxCorners = 2400 , qualityLevel = 0.08,
minDistance = 2, blockSize = 7

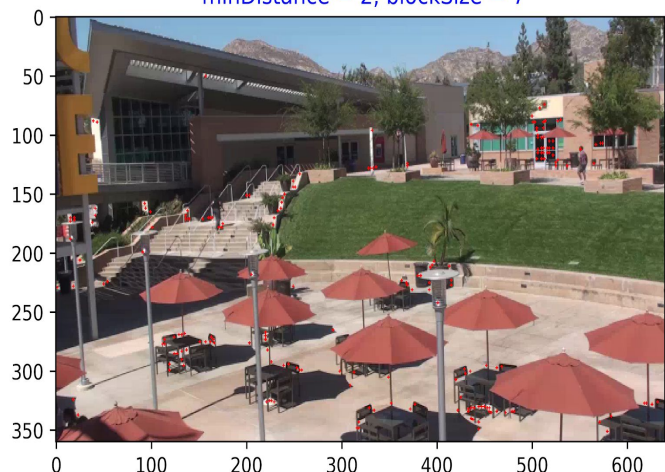


Παρατηρώντας τα αποτελέσματα που προέκυψαν με χρήση των ανιχνευτών Harris και Shai-Tomasi, διαπιστώνουμε ότι στην περίπτωση του δεύτερου εντοπίζονται περισσότερα σημεία ενδιαφέροντος, τα οποία αντιστοιχούν σε πραγματικές γωνίες του απεικονιζόμενου frame. Πιο συγκεκριμένα, χαρακτηριστικό παράδειγμα της προαναφερθείσας παρατήρησης αποτελούν τα ανιχνευθέντα σημεία στις ομπρέλες, τα κάγκελα της σκάλας αλλά και στα διάφορα δέντρα. Το γεγονός αυτό είναι αναμενόμενο, καθώς ο ανιχνευτής Harris παρουσιάζει ευαισθησία στις συνθήκες φωτισμού των προς ανίχνευση αντικειμένων στην εικόνα, οι οποίες επηρεάζονται άμεσα από τις δημιουργημένες σκιές. Αντίθετα, ο Shi Tomasi επιστρέφει σημεία που αναπαριστούν πιο ομοιόμορφα το σχήμα των αντικειμένων. Οι παράμετροι που αναγράφονται είναι οι βέλτιστες και για τους δύο ανιχνευτές, θεωρώντας ως κριτήριο το μέγιστο πλήθος των πραγματικών σημείων ενδιαφέροντος που κινούνται στα επόμενα frames του video, καθώς στα πλαίσια της εν λόγω ανάλυσης είναι και το επιθυμητό αποτέλεσμα.

Harris detector using maxCorners = 4400 , qualityLevel = 0.01,
minDistance = 2, blockSize = 7

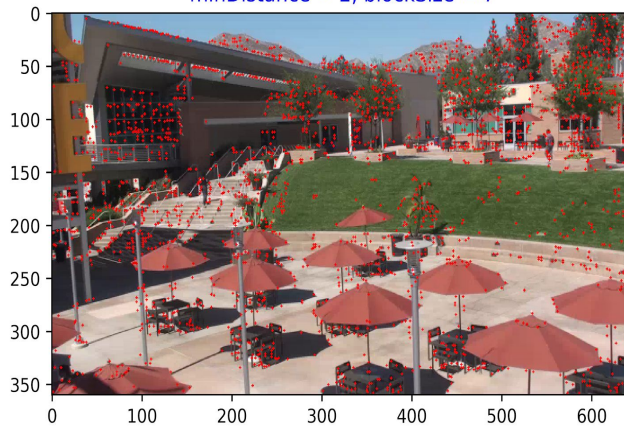


Harris detector using maxCorners = 200 , qualityLevel = 0.01,
minDistance = 2, blockSize = 7

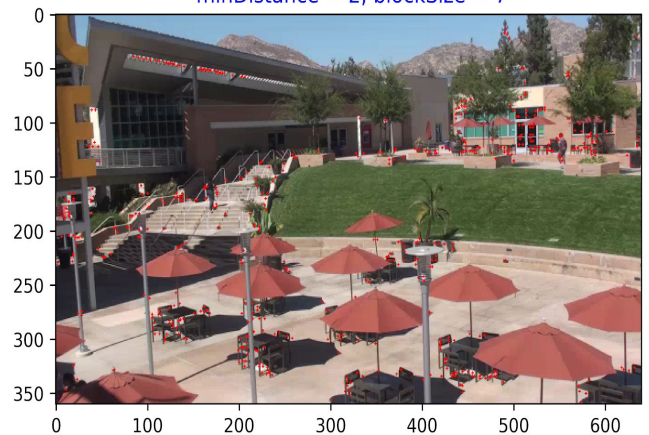


Μεταβάλλοντας το μέγιστο πλήθος γωνιών στην περίπτωση του Harris ανιχνευτή, παρατηρούμε ότι η επίδραση στον αριθμό των ανιχνευθέντων σημείων ενδιαφέροντος είναι και η ανάλογη σε σχέση με το είδος της (αύξηση-μείωση).

Shai-Tomasi detector using maxCorners = 4000 , qualityLevel = 0.01, minDistance = 2, blockSize = 7

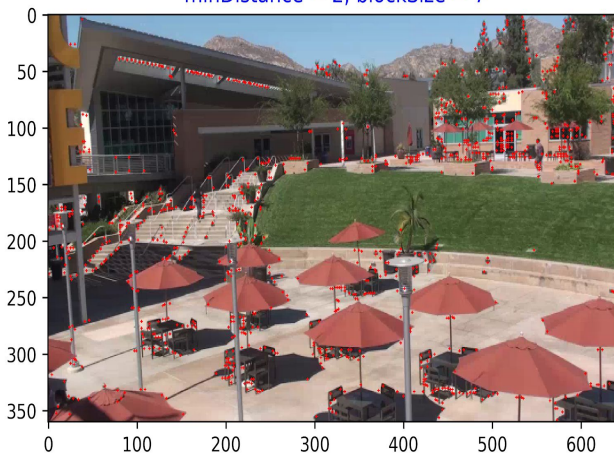


Shai-Tomasi detector using maxCorners = 500 , qualityLevel = 0.01, minDistance = 2, blockSize = 7

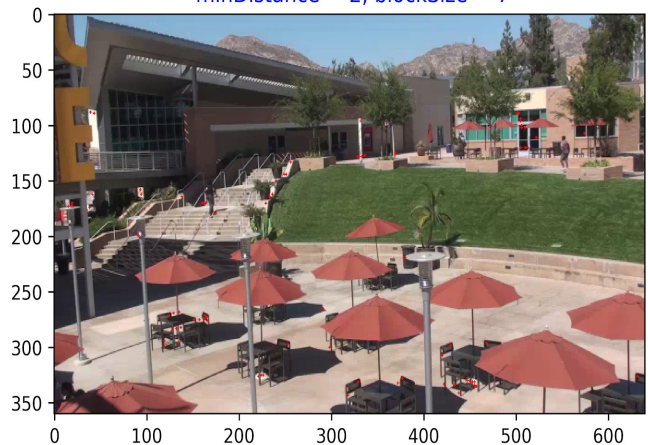


Αντίστοιχα αποτελέσματα παρατηρούμε και στην περίπτωση του Shi Tomasi detector. Αξίζει να σημειωθεί ότι στην περίπτωση αύξησης εντοπίζονται πάρα πολλά σημεία ενδιαφέροντος σε κάθε τμήμα της εικόνας.

Harris detector using maxCorners = 4400 , qualityLevel = 0.01, minDistance = 2, blockSize = 7

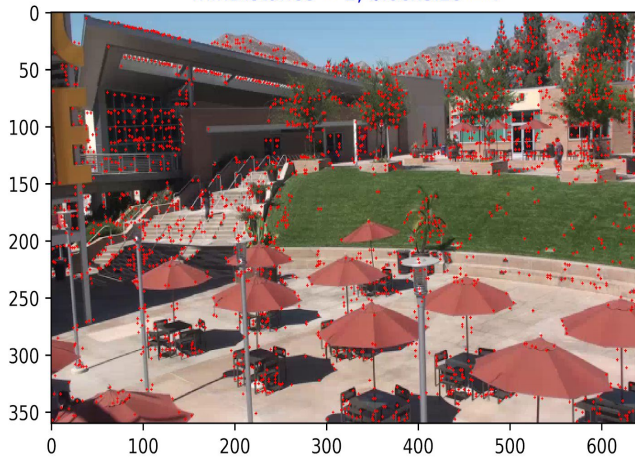


Harris detector using maxCorners = 4400 , qualityLevel = 0.3, minDistance = 2, blockSize = 7

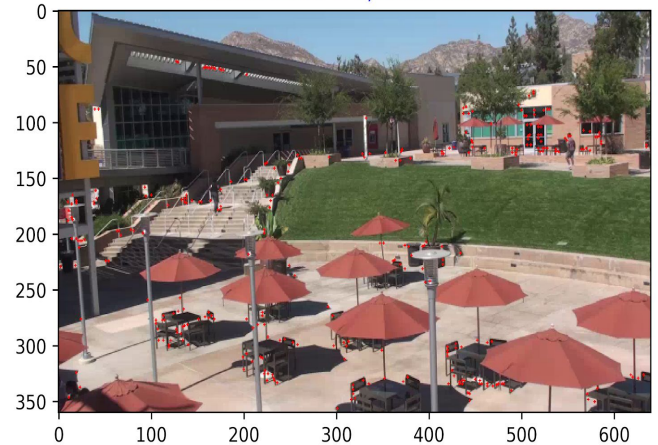


Όσον αφορά την παράμετρο qualityLevel, η οποία λειτουργεί ως κατώφλι στην διαδικασία επιλογής των τελικών σημείων ενδιαφέροντος που εντοπίζει ο αλγόριθμος του Harris detector, παρατηρούμε ότι η ελάττωσή της συνεπάγεται μεγαλύτερο αριθμό γωνιών, οι οποίες όμως ενδέχεται αντιστοιχούν σε μη-πραγματικά σημεία ενδιαφέροντος ή θόρυβο, ενώ αντίθετα η αύξησή της συνεπάγεται μικρότερο πλήθος γωνιών, οι οποίες ωστόσο αντιπροσωπεύουν πραγματικά σημεία ενδιαφέροντος που επιθυμούμε να ανιχνεύσουμε στην εικόνα.

Shai-Tomasi detector using maxCorners = 2400 , qualityLevel = 0.01,
minDistance = 2, blockSize = 7



Shai-Tomasi detector using maxCorners = 2400 , qualityLevel = 0.3,
minDistance = 2, blockSize = 7



Αντίστοιχα αποτελέσματα παρατηρούμε και στην περίπτωση του Shi Tomasi detector.

4. Εφαρμογή αλγορίθμου Lucas Kanade για τον υπολογισμό οπτικής ροής

Στην συνέχεια παραθέτουμε ενδεικτικά αποτελέσματα που προέκυψαν κατά τον πειραματισμό μας με τις τιμές των παραμέτρων *winSize*, *maxLevel*, *criteria* των δύο ανιχνευτών με στόχο την καλύτερη δυνατή παρακολούθηση των σημείων ενδιαφέροντος του δοθέντος video.

Harris Detector



window size = (5,5)



window size = (10,10)



window size = (15,15)

Shi Tomasi Detector



window size = (5,5)

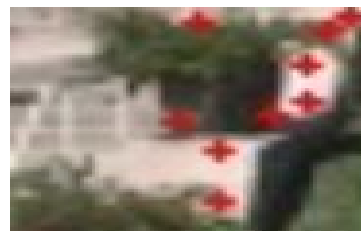
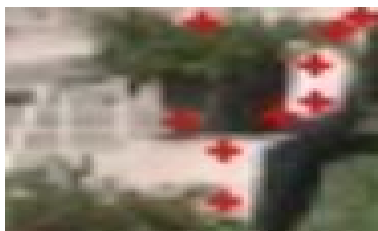


window size = (10,10)



window size = (25,25)

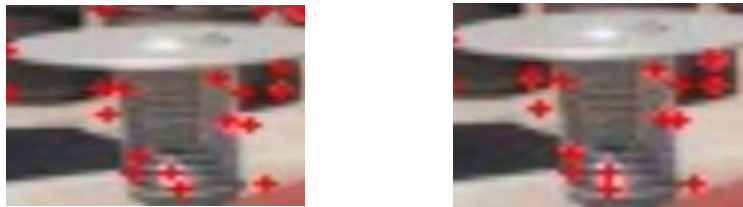
Πειραματιζόμενοι με το μέγεθος του παραθύρου που περιβάλλει το εκάστοτε σημείο ενδιαφέροντος σε ένα συγκεκριμένο frame του βίντεο, διατηρώντας σταθερές τις υπόλοιπες παραμέτρους του Lucas-Kanade και πιο συγκεκριμένα για $maxLevel = 4$, $criteria = (cv2.TERM_CRITERIA_EPS | cv2.TERM_CRITERIA_COUNT, 10, 0.03)$, παρατηρούμε το θεωρητικά αναμενόμενο, δηλαδή ότι η αύξησή του συνεπάγεται μεγαλύτερο αριθμό ανιχνευθέντων γωνιών, οι οποίες όμως ενδέχεται να μην αντιστοιχούν σε πραγματικά επιθυμητά σημεία.



Μεταβάλλοντας το πλήθος των επιπέδων της πυραμίδας κλιμάκων που χρησιμοποιεί ο αλγόριθμος Lucas Kanade, δεν παρατηρούμε ιδιαίτερη διαφοροποίηση στα οπτικά αποτελέσματα, διότι τα περισσότερα επίπεδα προσφέρουν καλύτερη ανίχνευση κίνησης όταν έχουμε γρήγορες κινήσεις, κάτι που λείπει από το δοθέν βίντεο. Όμως, τα περισσότερα επίπεδα στον αλγόριθμο Lucas-Kanade, επιφέρουν και περισσότερο flickering, δηλαδή αστάθεια μετατοπίσεων στα σημεία ενδιαφέροντος, όσον αφορά τις αργές κινήσεις. Επομένως προτιμούμε ένα σχετικά χαμηλό πλήθος επιπέδων (**4**) ώστε στο βήμα 5 τα σημεία ενδιαφέροντος που κινούνται λόγω θορύβου να μην εμφανίζονται στην μάσκα σημείων ενδιαφέροντος. Οι εν λόγω παρατηρήσεις ισχύουν και για τους δύο ανιχνευτές.

Ορίζουμε τα κριτήρια τερματισμού δηλώνοντας το μέγιστο αριθμό επαναλήψεων ύστερα από τις οποίες ο αλγόριθμος συγκλίνει με χρήση της μεταβλητής `cv2.TERM_CRITERIA_COUNT` αλλά και εναλλακτικά το κατώφλι μετατόπισης του παραθύρου που υπολογίζει ο αλγόριθμος, κάτω από το οποίο σταματά την εκτέλεσή του, με χρήση της μεταβλητής `cv2.TERM_CRITERIA_EPS`.

Shi Tomasi Detector for `cv2.TERM_CRITERIA_EPS = 9`



Πειραματιστήκαμε, αρχικά διατηρώντας σταθερό τον αριθμό των μέγιστων επαναλήψεων και συγκεκριμένα ίσο με **40** και δίνοντας στην μεταβλητή κατώτατου ορίου διαδοχικά τις τιμές **0,0001**, **0.03**, **9**. Παρατηρήσαμε ότι για μικρότερες τιμές του εν λόγω μεγέθους, οι υπολογιζόμενες μετατοπίσεις των σημείων ενδιαφέροντος λαμβάνουν και πιο μικρές τιμές, οπότε προκύπτει λιγότερο flickering στο τελικό video.

Στην συνέχεια, διατηρώντας σταθερή την τιμή του κατώτατου ορίου και ίση με **0.03**, μεταβάλλουμε τον αριθμό των επαναλήψεων εφαρμόζοντας τις τιμές **2**, **10**, **50**. Διαπιστώσαμε ότι μεγαλύτερος αριθμός επαναλήψεων επιφέρει καλύτερη σύγκλιση στις υπολογιζόμενες μετατοπίσεις.

Συνολικά, οι βέλτιστες παράμετροι του Lucas-Kanade είναι ***winSize = (10,10)***, ***maxLevel = 4***, ***criteria = (cv2.TERM_CRITERIA_EPS | cv2.TERM_CRITERIA_COUNT, 10, 0.03)***.

5. Παρακολούθηση γωνιών

Στο βήμα αυτό, τροποποιούμε τον αλγόριθμό μας, ώστε να ενημερώνει τα προς ακολούθηση σημεία ενδιαφέροντος, ανά τακτά χρονικά διαστήματα, κατά την διάρκεια του βίντεο. Επίσης φροντίζουμε ώστε να απεικονίζουμε τις κινήσεις των σημείων ενδιαφέροντος που είναι μεγάλες, δηλαδή αντιστοιχούν σε μετατόπιση άνω των 2 pixels αθροιστικά ως προς τους δύο άξονες, και να τις διατηρούμε στην εικόνα για ορισμένο χρονικό διάστημα. Έτσι χρησιμοποιώντας τις βέλτιστες παραμέτρους που βρήκαμε στα βήματα 3, 4 και ενημερώνοντας τα σημεία ενδιαφέροντος ανά **[20]** frames έχουμε τα παρακάτω αποτελέσματα για τον κάθε ανιχνευτή:

Harris Detector



Shi Tomasi Detector



Παρατηρούμε παρόμοια συμπεριφορά των ανιχνευτών Harris και Shi-Tomasi. Λόγω των μικρών αντικειμένων στην εικόνα δεν μπορούμε εύκολα να διακρίνουμε τις διαφορές στην κίνηση των σημείων ενδιαφέροντος τα οποία έχουν εντοπιστεί από διαφορετικές μεθόδους, όπως περιγράψαμε στο βήμα (3). Όμως παρατηρούμε γενικά πιο σταθερή συμπεριφορά στη ανίχνευση κίνησης με Shi-Tomasi ανιχνευτή σημείων ενδιαφέροντος, χωρίς όμως να απέχουν πολύ τα αποτελέσματά του από τα αντίστοιχα του ανιχνευτή Harris.

Και στις δύο περιπτώσεις ανιχνευτών, η νέα μέθοδος οπτικοποίησης αποκλείει ικανοποιητικά τον θόρυβο και εμφανίζει κατά πλειονότητα, σημεία ενδιαφέροντος που αντιστοιχούν σε κινούμενους ανθρώπους.

Shi Tomasi Detector



Harris Detector



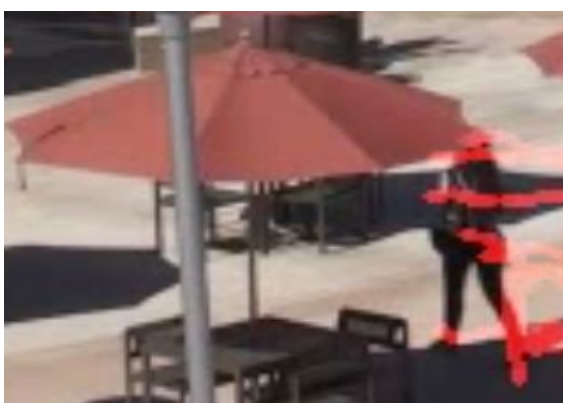
6. Εισαγωγή Salt and Pepper θορύβου

Σε αυτό το μέρος μελετούμε τον αντίκτυπο της ύπαρξης *salt & pepper* θορύβου στην αναγνώριση κίνησης μέσω του αλγορίθμου Lucas-Kanade και διαφορετικών ανιχνευτών σημείων ενδιαφέροντος. Παρατηρούμε, πως και με τους δύο ανιχνευτές, η αναγνώριση κίνησης δυσχεραίνεται από την ύπαρξη *salt & pepper* θορύβου, αφού λόγω των έντονων αντιθέσεων φωτεινότητας που δημιουργεί σε pixels του κάθε frame, έχουμε πολλά περισσότερες λανθασμένες ανιχνεύσεις σημείων ενδιαφέροντος και επομένως παραλείπονται από την παρακολούθηση πολλά πραγματικά σημεία ενδιαφέροντος. Παρακάτω παραθέτουμε παραδείγματα αναγνώρισης κίνησης με και χωρίς θόρυβο για το ίδιο frame:

Harris Detector



Shi Tomasi Detector



Παρατηρώντας τις προκύπτουσες εικόνες, διαπιστώνουμε ότι οι δύο ανιχνευτές παρουσιάζουν παρόμοια συμπεριφορά ως προς τον θόρυβο. Αυτό είναι κάτι που περιμέναμε, αφού και οι δύο μέθοδοι, για την εύρεση σημείων ενδιαφέροντος, χρησιμοποιούν Gaussian εξομάλυνση, που όπως έχουμε δει και στο Ερώτημα 1 της συγκεκριμένης Εργαστηριακής Άσκησης, δεν είναι αποτελεσματική στο να αφαιρέσει τον Salt & Pepper θόρυβο.

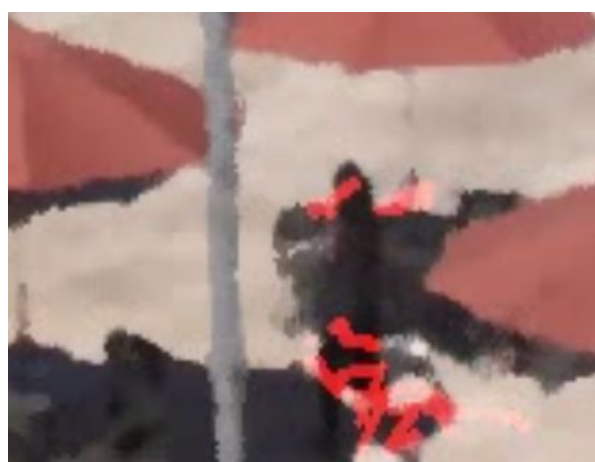
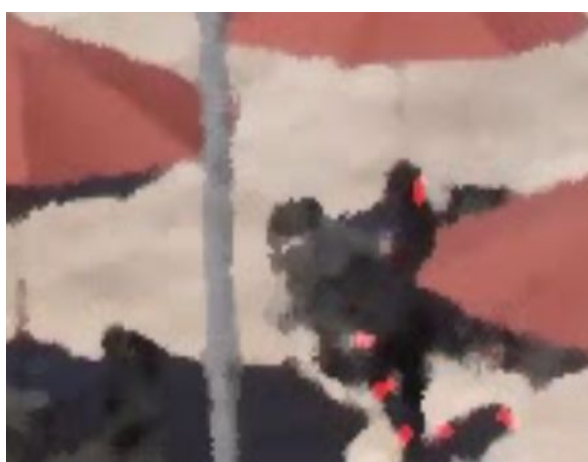
7. Αποθορυβοποίηση με χρήση Median φίλτρου

Στο βήμα αυτό, αποθορυβοποιούμε τον Salt & Pepper θόρυβο που παράγαμε στο προηγούμενο βήμα μέσω του βέλτιστου Median φίλτρου του Ερωτήματος 1 με σκοπό την καλύτερη παρακολούθηση κίνησης. Το δομικό στοιχείο του Median φίλτρου είναι δίσκος με ακτίνα 2 pixels. Τα αποτελέσματα που παίρνουμε παρατίθενται στην συνέχεια:

Shi Tomasi Detector



Harris Detector



Σε γενικές γραμμές παρατηρούμε, όπως είναι και θεωρητικά αναμενόμενο, ότι το median φιλτράρισμα που εφαρμόζουμε με στόχο την εξάλειψη του θορύβου salt and pepper, λειαίνει τα αιχμηρά χαρακτηριστικά του εκάστοτε frame, όπως για παράδειγμα τις ακμές ορισμένων απεικονιζόμενων αντικειμένων και συγχρόνως δημιουργεί ένα αισθητό θόλωμα. Επιπλέον, διαπιστώνουμε ότι η παρακολούθηση της κίνησης των ανιχνευθέντων σημείων ενδιαφέροντος είναι ικανοποιητική, αλλά σαφώς χειρότερη συγκριτικά με την περίπτωση απουσίας θορύβου.