



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ
ΥΠΟΛΟΓΙΣΤΩΝ

Όραση Υπολογιστών

ΕΡΑΣΤΗΡΙΑΚΗ ΑΝΑΦΟΡΑ 3^{ης} ΑΣΚΗΣΗΣ

Εντοπισμός Χωρο-χρονικών Σημείων Ενδιαφέροντος και Εξαγωγή Χαρακτηριστικών
σε Βίντεο Ανθρωπίνων Δράσεων

ΡΟΛΑΝΔΟΣ-ΑΛΕΞΑΝΔΡΟΣ ΠΟΤΑΜΙΑΣ
A.M. 03114437

Μέρος 1

Στο πρώτο μέρος της εργαστηριακής άσκησης θα αναζητήσουμε χώρο-χρονικά σημεία ενδιαφέροντος με την χρήση του αλγορίθμου ανίχνευσης ακμών/γωνιών \$Harris\$ Detector και Gabor φίλτρων.

1.1 Υπολογισμός κριτηρίου γωνιότητας

Επεκτείνουμε τον αλγόριθμο του Harris στις τρεις διαστάσεις αφού πρώτα εξομαλύνουμε την εικόνα με Gauss φίλτρα. Η χρονική μεταβολή στις εικόνες μας επιβάλει να χρησιμοποιήσουμε ένα 3x3 πίνακα M (δομικό τανυστή) σε αντίθεση με την εργαστηριακή άσκηση 2, προσθέτοντας και τις μερικές παραγώγους ως προς τον χρόνο. Στην συνέχεια υπολογίζουμε τις μερικές παραγώγους της εικόνας με την βοήθεια Kernel πυρήνων διαφορών. Τελικά υπολογίζουμε τις περιοχές που μεγιστοποιούν το κριτήριο:

$$H(x,y,t) = \det(M(x,y,t)) - k \operatorname{trace}^3 M(x,y,t),$$

Με βάση τα τοπικά μέγιστα και καθορισμένα κατώφλια υπολογίζουμε τα σημεία ενδιαφέροντος που φαίνονται παρακάτω, με την χρήση της συνάρτησης `showDetection.m` :



Σχήμα 1: a) Box 1: Frame 100 b) Box 2: Frame 100 c) Box 3: Frame 120



Σχήμα 2 : a) Run 1: Frame 20 b) Run 2: Frame 140 c)



Σχήμα 3: a) Walk 1: Frame 80 b) Walk 2: Frame 100 c) Walk 3: Frame 120

1.2 Μεγιστοποίηση Κριτηρίου με Gabor Φίλτρα

Όπως και στην προηγούμενη μέθοδο το πρώτο βήμα για την εφαρμογή του κριτηρίου είναι η εξομάλυνσή της εικόνας με χρονικά, αυτή την φορά, Gauss φίλτρα. Στην συνέχεια εφαρμόζουμε το άρτιο/περιττό ζεύγος Gabor χρονικών φίλτρων στην εικόνα και υπολογίζουμε την τετραγωνική ενέργεια εξόδου του φίλτρου. Σαν σημεία ενδιαφέροντος επιλέγουμε πάλι, όπως και προηγουμένως, τα σημεία που μεγιστοποιούν το κριτήριο σε μια υποπεριοχή. Μερικά αποτελέσματα φαίνονται παρακάτω:



Σχήμα 4: a) Box 1: Frame 100 b) Box 2: Frame 170 c) Box 3: Frame 120



Σχήμα 5: a) Run 1: Frame 110 b) Run 2: Frame 130 c) Run 3: Frame 10



Σχήμα 6: a) Walk 1: Frame 30 b) Walk 2: Frame 120 c) Walk 3: Frame 60

Μεταβάλλοντας τις παραμέτρους σ , θ , k για κάθε φίλτρο καταλήγουμε στην επιλογή:

- Harris Detector: Χωρικό $\sigma = 3$, χρονικό $\sigma = 2$, $\theta = 0.001$, $k = 0.03$ για τα βίντεο Box και $k = 0.01$ για τα βίντεο Run, Walk.
- Gabor Filter Detector: Χωρικό $\sigma = 3$, χρονικό $\sigma = 2$, $\theta = 0.4$, για τα βίντεο Run και $\theta = 0.2$ για τα βίντεο Box, Walk.

Παρατηρούμε ότι ο αλγόριθμος Harris είναι ευαίσθητος σε μικρές μεταβολές και κινήσεις αφού βασίζεται στις κατευθυντηκές παραγώγους. Αντίθετα ο εντοπισμός σημείων ενδιαφέροντος με χρήση Gabor φίλτρων παρουσιάζει προβλήματα σε

μικρές και ομαλές κινήσεις, όπως στα Walking Videos, αλλά βελτιώνει τα αποτελέσματα του Harris σε βίντεο με απότομες κινήσεις όπως στα Box Videos.

Μέρος 2

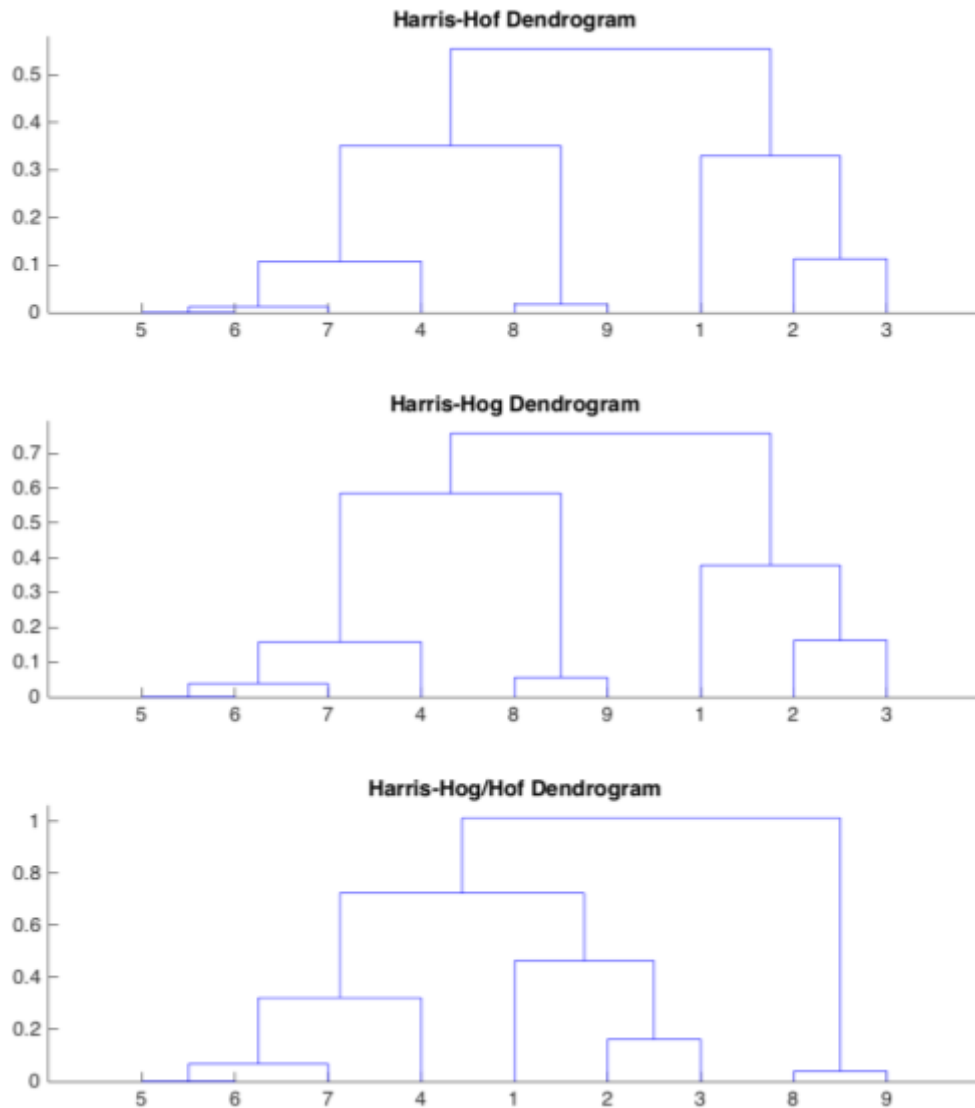
Στο δεύτερο μέρος της εργαστηριακής άσκησης σκοπός είναι να εξάγουμε τους χωρο-χρονικούς περιγραφητές Hog, Hof για μια περιοχή 4xScale γύρω από κάθε σημείο ενδιαφέροντος. Για την υλοποίηση της διαδικασίας υπολογίζουμε την οπτική ροή της εικόνας στις περιοχές που προαναφέραμε με την βοήθεια του αλγορίθμου Lukas-Kanade που δημιουργήσαμε στο δεύτερο εργαστήριο. Με την βοήθεια της συνάρτησης *OriantationHistogram.m* υπολογίζουμε τους περιγραφητές Hog και Hof χρησιμοποιώντας σαν είσοδο τις τετραγωνικές περιοχές των κατευθυντικών παραγόνων ή τις τετραγωνικές περιοχές της οπτικής ροής της εικόνας αντίστοιχα. Για να δημιουργήσουμε περιγραφητές Hof, Hog αλλά και τον συνδυασμό τους και να αξιολογήσουμε μετέπειτα την ικανότητα τους στην αναγνώριση ομοιοτήτων των κινήσεων καλούμε την συνάρτηση *Descriptors.m* που εκτελεί τις παραπάνω λειτουργίες για κάθε ένα βίντεο.

Στην συνέχεια δημιουργούμε ένα καθολικό ιστόγραμμα για όλα τα βίντεο για κάθε περιγραφητή και τα διαχωρίζουμε σε Clusters με την βοήθεια του αλγορίθμου *k-means*. Ο διαχωρισμός γίνεται με βάση την ευκλείδεια απόσταση από τα 30 κέντρα των κλάσεων όπως ορίζει η τεχνική ταξινόμησης Bag of Visual Words. Η συνάρτηση που υλοποιεί τα παραπάνω ονομάζεται *BagOfWords.m*

Μέρος 3

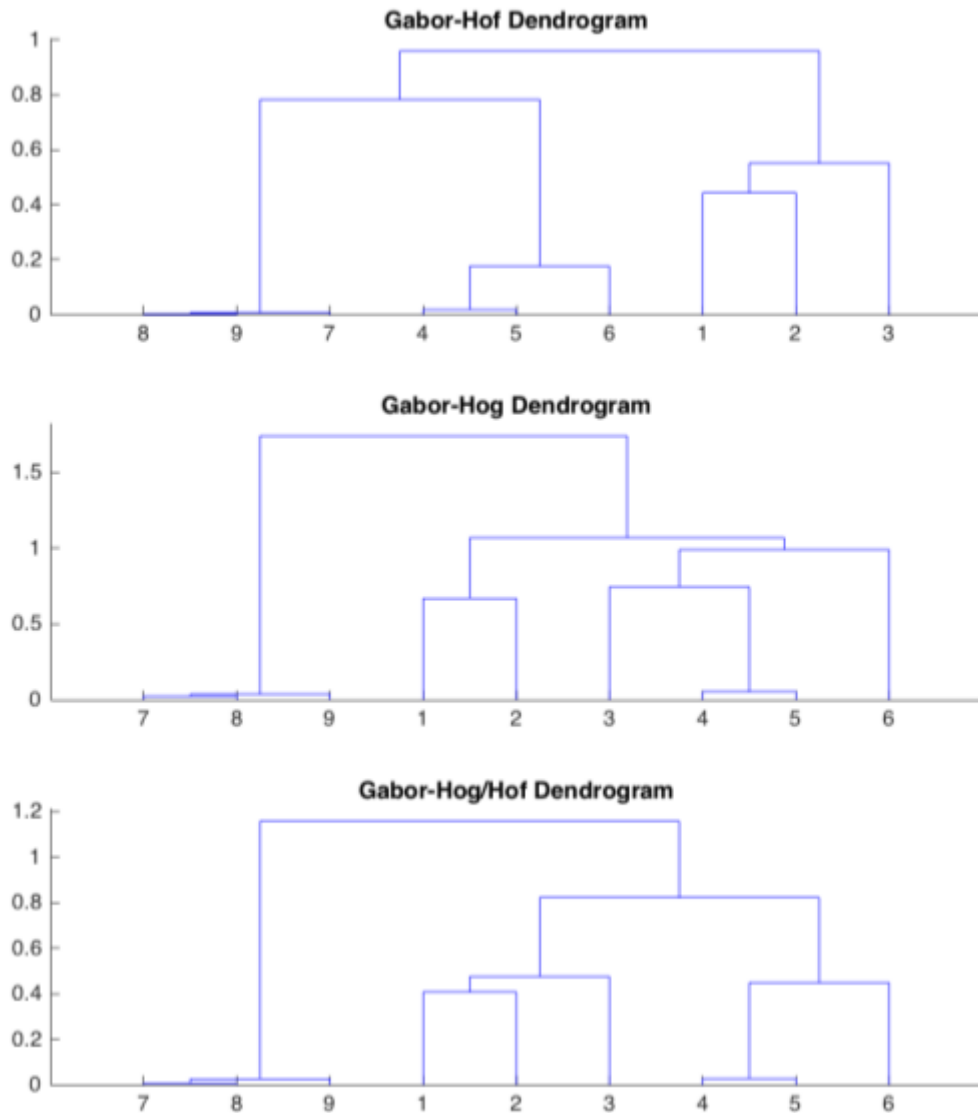
Στο τρίτο μέρος της εργαστηριακής άσκησης δημιουργούμε ένα δενδρόγραμμα για να παρατηρήσουμε τα αποτελέσματα της ταξινόμησης Bag Of Words του προηγούμενου μέρους. Τα δενδρογράμματα προκύπτουν από τα καθολικά ιστογράμματα της συνάρτησης *BagOfWords.m* για κάθε έναν από τους ανιχνευτές και περιγραφητές με την βοήθεια των συναρτήσεων *linkage* και *distChisq*.

Παρακάτω μπορούμε να παρατηρήσουμε τα αποτελέσματα που προέκυψαν:



Σχήμα 7: Δενδρογράμματα ανιχνευτή Harris για τους τρεις περιγραφητές

Παρατηρούμε ότι για τον ανιχνευτή Harris έχουμε πάντοτε σωστή ομαδοποίηση των Boxing βίντεο που περιέχουμε αργές και σταθερές κινήσεις. Αντίθετα ο διαχωρισμός των running videos δεν γίνεται σωστά, με το πρώτο running βίντεο να απέχει πάντοτε από τα άλλα δύο που εμφανίζουν περισσότερα σημεία ενδιαφέροντος. Τέλος παρόμοια αποτελέσματα έχουμε και στα Walking βίντεο και συγκεκριμένα το πρώτο Walk video ταυτίζεται με τα Running Videos.



Σχήμα 7: Δενδρογράμματα ανιχνευτή με χρήση Gabor φίλτρων για τους τρεις περιγραφητές

Το αντίστοιχο δενδρόγραμμα για Gabor φίλτρα παρατηρούμε ότι έχει πολύ καλύτερα αποτελέσματα από τον Harris Detector. Συγκεκριμένα και στις τρεις περιπτώσεις περιγραφητών τα Walking και Running videos έχουν ταξινομηθεί σωστά, και ειδικότερα ο Hof ομαδοποιεί τα αποτελέσματα ιδανικά. Οι περιγραφητές Hog και ο συνδυασμός Hog/Hof ομαδοποιούν και αυτοί αρκετά σωστά τα αποτελέσματα και συνεπώς μπορούμε να βγάλουμε ασφαλές συμπέρασμα πως ο ανιχνευτής με την χρήση Gabor φίλτρων δίνει πολύ καλύτερα αποτελέσματα.