

Όραση Υπολογιστών

2^η Εργαστηριακή Άσκηση

Θέμα: Εκτίμηση Οπτικής Ροής (Optical Flow) και

Εξαγωγή Χαρακτηριστικών σε Βίντεο

ΣΧΟΛΗ: ΣΗΜΜΥ

Ονοματεπώνυμο	Αριθμός Μητρώου
Γιάννης Πιτόσκας	03115077
Γιώργος Χιονάς	03115132



Μέρος 1: Παρακολούθηση Προσώπου με Χρήση του Πεδίου Οπτικής Ροής (Optical Flow) με τη μέθοδο Lucas-Kanade

Σκοπός: Θέλουμε να υλοποιήσουμε ένα σύστημα Face Tracking σε ένα βίντεο νοηματικής γλώσσας, ανιχνεύοντας στο πρώτο frame του βίντεο την περιοχή του προσώπου με πιθανοτικό ανιχνευτή ανθρώπινου δέρματος και ύστερα θα παρακολουθούμε το πλαίσιο αυτό μέσω του διανυσματικού πεδίου οπτικής ροής (αλγόριθμος Lucas-Kanade).

1.1 Ανίχνευση Δέρματος Προσώπου

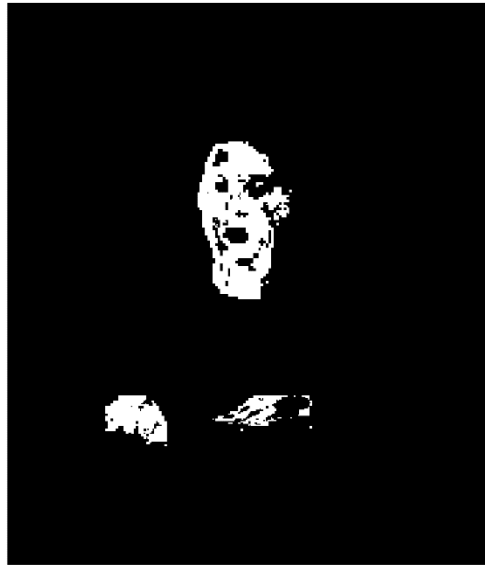
Θέλουμε να ανιχνεύσουμε την περιοχή του προσώπου στο πρώτο frame του βίντεο. Για να το επιτύχουμε αυτό αρχικά μεταφέρουμε την εικόνα στον YCbCr χρωματικό χώρο και αδιαφορούμε για την πληροφορία της φωτεινότητας που βρίσκεται στο Y κανάλι και επικεντρωνόμαστε στα κανάλια Cb και Cr που περιγράφουν την ταυτότητα του χρώματος. Το χρώμα του δέρματος μοντελοποιείται με δισδιάστατη Γκαουσιανή κατανομή ως εξής:

$$P(\mathbf{c} = \text{skin}) = \frac{1}{\sqrt{|\Sigma|} (2\pi)^2} e^{-\frac{1}{2}(\mathbf{c}-\boldsymbol{\mu})\Sigma^{-1}(\mathbf{c}-\boldsymbol{\mu})'}$$

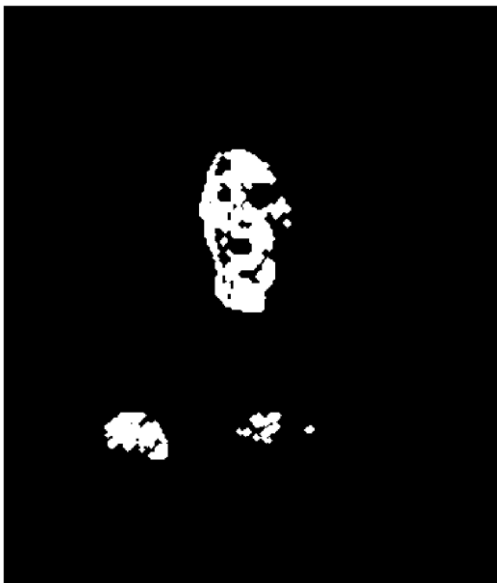
όπου $\boldsymbol{\mu} = [\mu_{Cb} \ \mu_{Cr}]$ και Σ ο 2×2 πίνακας συνδιακύμανσης.

Παίρνουμε από το αρχείο skinSamplesRGB.mat δείγματα δέρματος που δίνονται σε μορφή RGB τα μετατρέπουμε σε YcbCr και υπολογίζουμε τα $\boldsymbol{\mu}$, Σ της Γκαουσιανής μας. Έτσι, προκύπτει η εικόνα πιθανοτήτων $P(\mathbf{c}(x, y) = \text{skin})$ με κατωφλιοποίηση της οποίας προκύπτει η δυαδική εικόνα που καθορίζει αν ένα σημείο της εικόνας αντιστοιχεί σε δέρμα ή όχι.

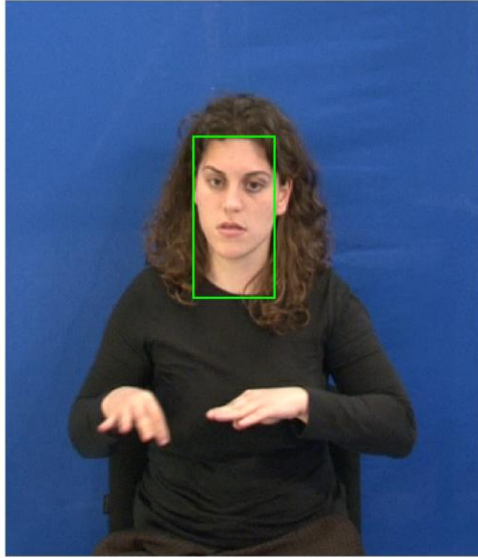
Με $Threshold = \frac{0.2}{\sqrt{|S|(2\pi)^2}}$ η δυαδική εικόνα ανίχνευσης δέρματος προκύπτει ως εξής:



Στη συνέχεια καλούμαστε να απομονώσουμε τα σημεία που αντιστοιχούν σε πρόσωπο κάνοντας discard τα υπόλοιπα. Θα επιλέξουμε λοιπόν την περιοχή με τον μεγαλύτερο εμβαδό. Ωστόσο, θα χρειαστεί πρώτα μια πρώτη μια μορφολογική επεξεργασία της δυαδικής εικόνας για να καλύψουμε τα μικρά κενά. Συγκεκριμένα θέλουμε να εξαλειφθούν οι μικρές περιοχές και να αποκτήσουν συνοχή οι περιοχές του προσώπου και των χεριών. Αυτό το επιτυγχάνουμε με opening με πολύ μικρό δομικό στοιχείο και στη συνέχεια closing με μεγάλο δομικό στοιχείο. Παρακάτω φαίνεται με τη σειρά ένα opening της παραπάνω εικόνας με δομικό στοιχείο δίσκο ακτίνας 2 και στη συνέχεια closing με δομικό στοιχείο δίσκο ακτίνας 30.



Και τώρα είμαστε σε θέση να επιλέξουμε την περιοχή με το μεγαλύτερο εμβαδό η οποία προφανώς αντιστοιχεί στην περιοχή του προσώπου και κατ'επέκταση να φτιάξουμε το πλαίσιο οριοθέτησης προσώπου (Bounding Box). Το αποτέλεσμα που λαμβάνουμε φαίνεται παρακάτω:



1.2 Υλοποίηση του αλγορίθμου των Lucas-Kanade

Σε αυτό το σημείο καλούμαστε να υλοποιήσουμε τον αλγόριθμο των Lucas-Kanade, ο οποίος δεδομένων δύο διαδοχικών εικόνων, υπολογίζει το πεδίο οπτικής ροής $\mathbf{d} = (d_x, d_y)$ φέρνοντας τις σε αντιστοιχία.

Προκύπτει ότι η λύση ελάχιστων τετραγώνων (με ανάπτυξη κατά Taylor) για την βελτίωση του πεδίου οπτικής ροής \mathbf{d} σε κάθε σημείο είναι:

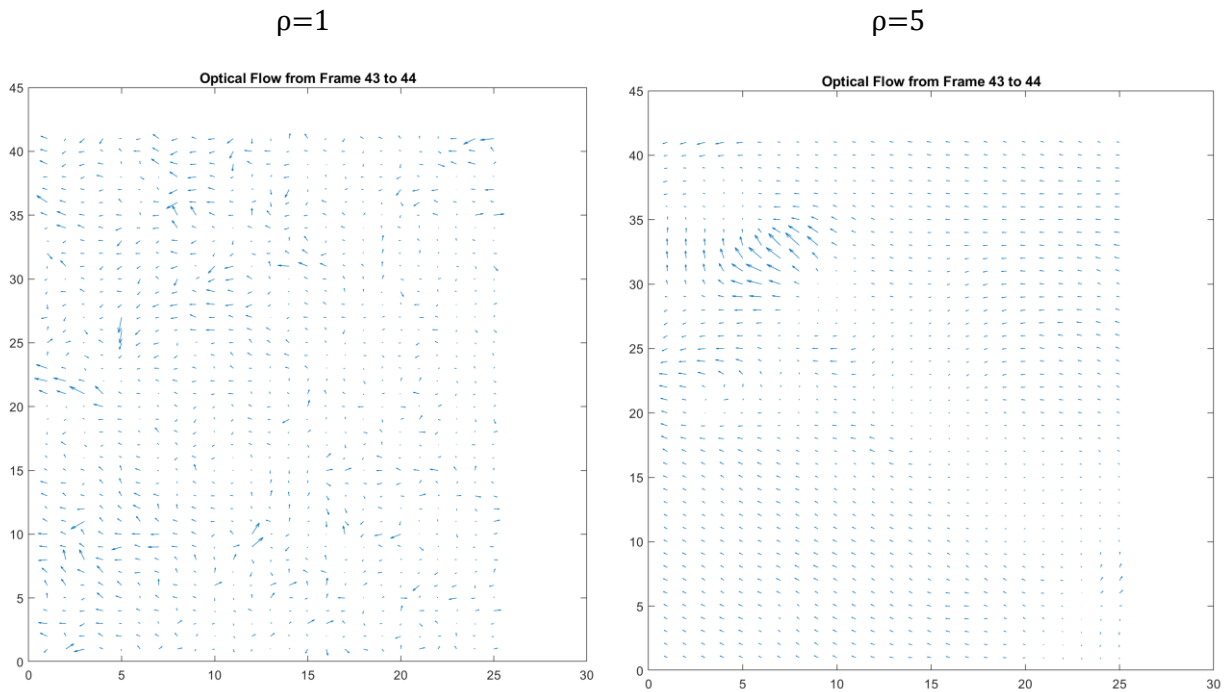
$$\mathbf{u}(\mathbf{x}) = \begin{bmatrix} (G_\rho * A_1^2)(\mathbf{x}) + \epsilon & (G_\rho * (A_1 A_2))(\mathbf{x}) \\ (G_\rho * (A_1 A_2))(\mathbf{x}) & (G_\rho * A_2^2)(\mathbf{x}) + \epsilon \end{bmatrix}^{-1} \cdot \begin{bmatrix} (G_\rho * (A_1 E))(\mathbf{x}) \\ (G_\rho * (A_2 E))(\mathbf{x}) \end{bmatrix}$$

όπου:

$$A(\mathbf{x}) = [A_1(\mathbf{x}) \quad A_2(\mathbf{x})] = \left[\frac{\partial I_{n-1}(\mathbf{x} + \mathbf{d}_i)}{\partial x} \quad \frac{\partial I_{n-1}(\mathbf{x} + \mathbf{d}_i)}{\partial y} \right]$$

$$E(\mathbf{x}) = I_n(\mathbf{x}) - I_{n-1}(\mathbf{x} + \mathbf{d}_i)$$

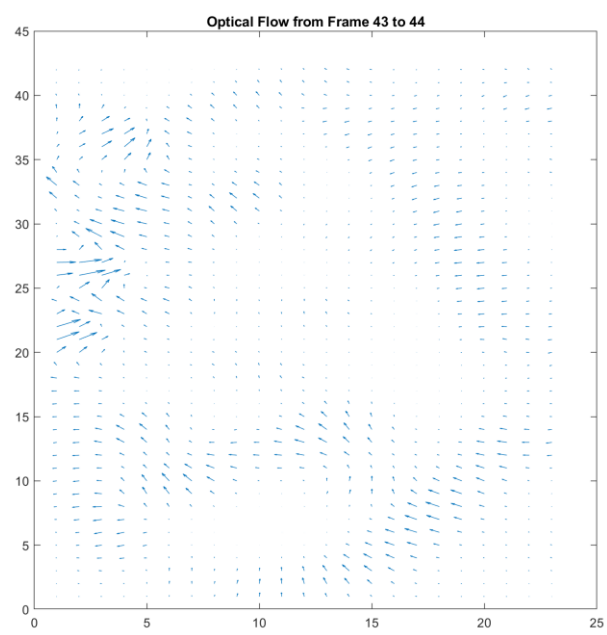
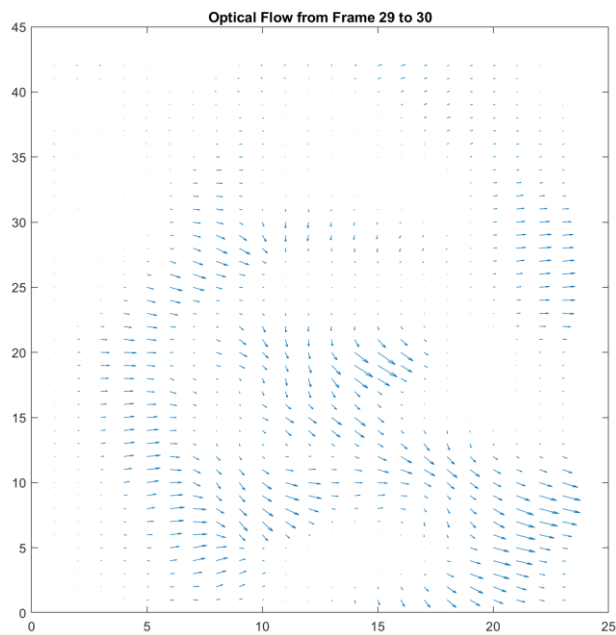
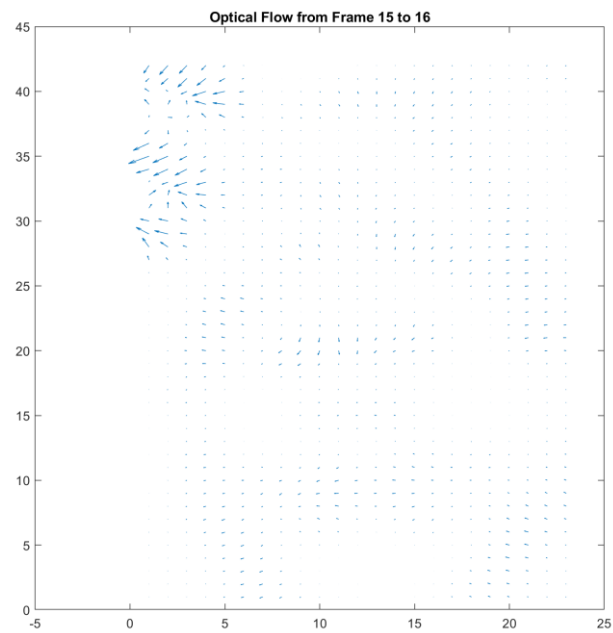
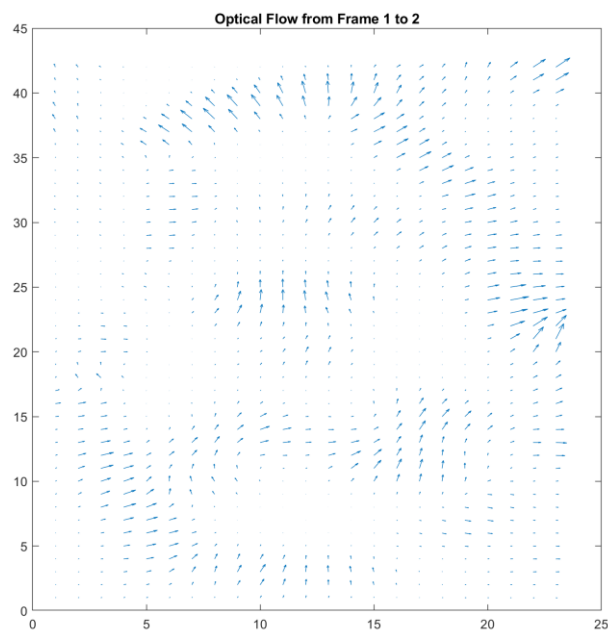
Αρχικά, να παρατηρήσουμε τι συμβαίνει καθώς μεταβάλλουμε την παράμετρο ρ :

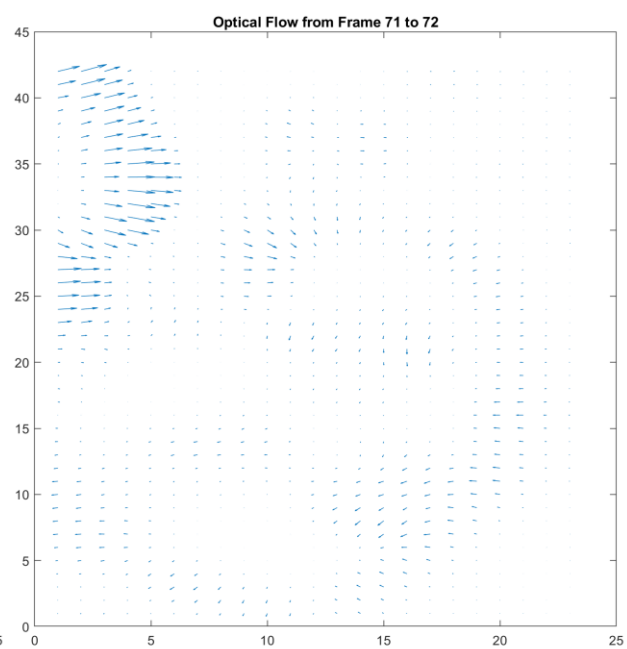
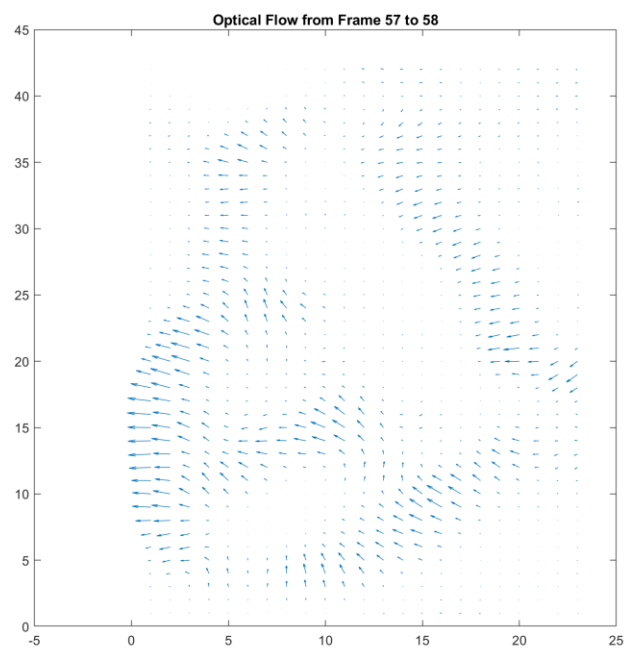


Αυτό που παρατηρούμε είναι ότι για μεγαλύτερες τιμές του ρ το πεδίο οπτικής έμφανίζει μεγαλύτερη ομοιομορφία και ομαλότητα και αυτό φαίνεται και από τις κατευθύνσεις των διανυσμάτων. Πειραματιστήκαμε και με μεγαλύτερες τιμές του ρ και παρατηρήσαμε ότι οι αλλαγές ύστερα είναι πολύ μικρές.

Όσο για την παράμετρο ϵ , δεν αξίζει κάποια οπτικοποίηση καθώς η διαφορές που φαίνονται στο διάστημα $[0.01, 0.1]$ είναι αμελητέες και αφορούν την περαιτέρω επέκταση των σχεδόν μηδενικών διανυσμάτων οπτικής ροής το οποίο φαίνεται μάλιστα να γίνεται λιγότερο αισθητό καθώς αυξάνουμε το ρ .

Για $\rho = 3$, $\varepsilon = 0.05$ το πεδίο οπτικής ροής για κάποιες ενδεικτικές περιπτώσεις φαίνεται παρακάτω:



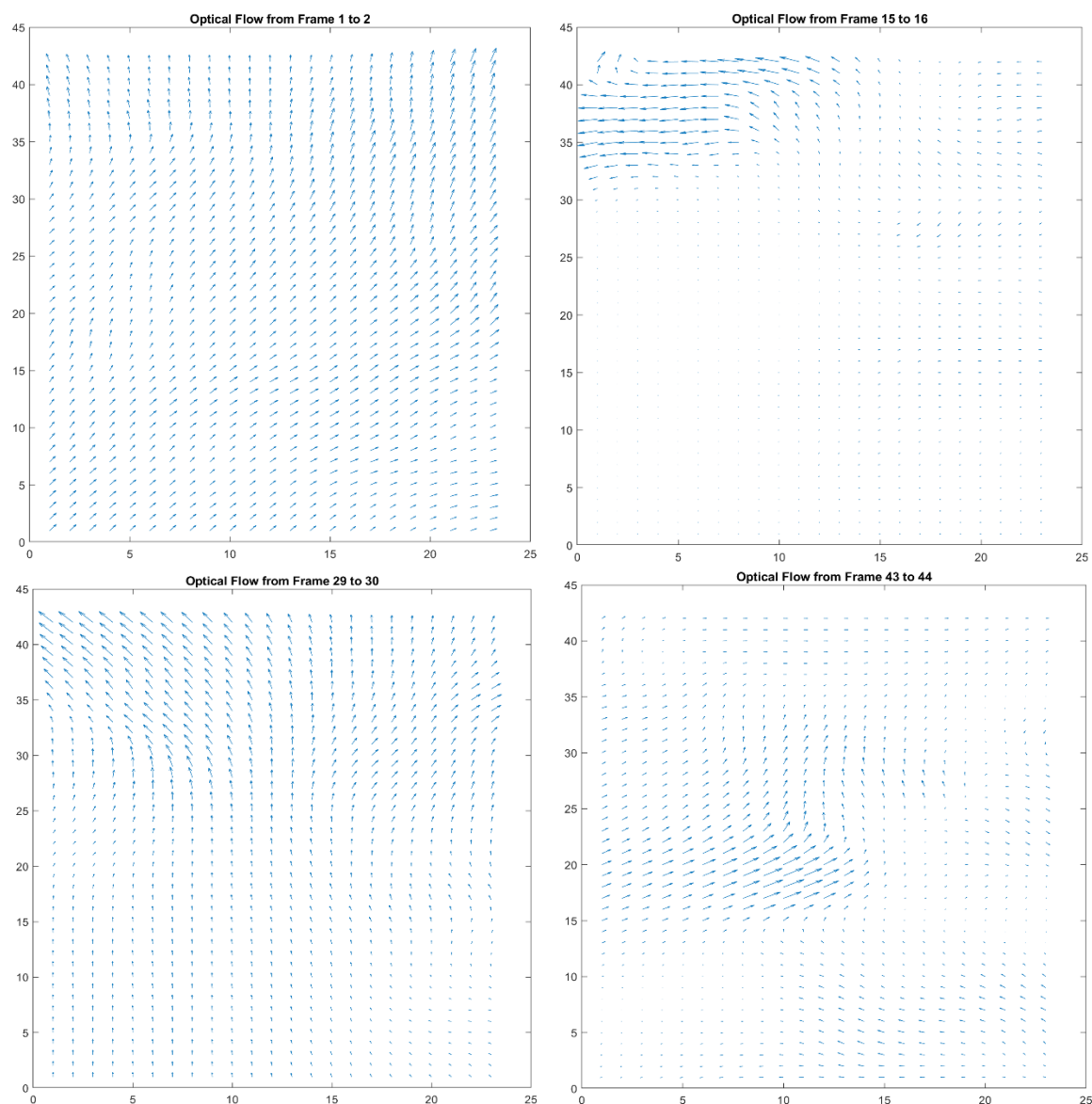


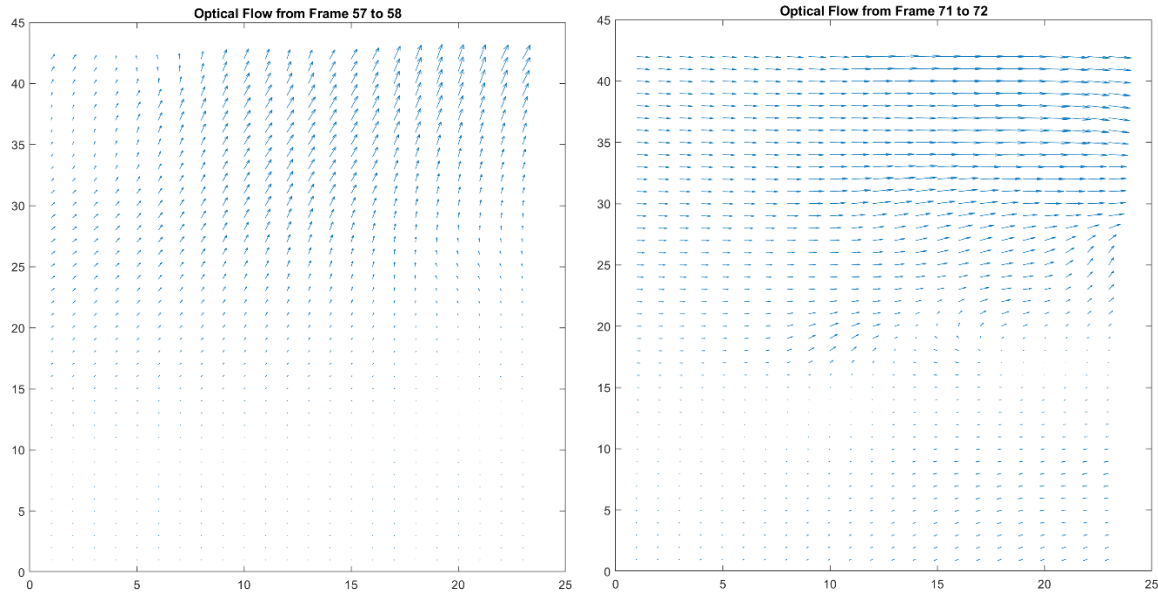
1.3 Πολυ-Κλιμακωτός Υπολογισμός Οπτικής Ροής

(Προαιρετικό για τους προπτυχιακούς με bonus 20%)

Στην πολυκλιμακωτή εκδοχή του αλγορίθμου των Lucas-Kanade αναλύουμε τις αρχικές εικόνες σε γκαουσιανές πυραμίδες (χρήση συνάρτησης `impyramid`) και θα γίνεται υπολογισμός του πεδίου οπτικής ροής από τα πιο μικρά στα πιο μεγάλα scales. Επαναλαμβάνουμε τον μονοκλιμακωτό αλγόριθμο των Lucas-Kanade για κάθε κλίμακα, τόσες φορές όσος είναι ο αριθμός των κλιμακών της πυραμίδας. Πρέπει να δοθεί ιδιαίτερη έμφαση στο γεγονός ότι κατά τη μετάβαση από μεγάλες σε μικρές κλίμακες πρέπει να πραγματοποιηθεί ένα βαθυπερατό φιλτράρισμα της εικόνας για την μείωση το aliasing κατά την υποδειγματοληψία. Γι'αυτόν τον σκοπό επιλέξαμε Gaussian με $\sigma = 3$ pixels ως βαθυπερατό φίλτρο, όπως υποδεικνύεται. Ακόμη προσέχουμε όταν μεταβαίνουμε από μικρές σε μεγάλες κλίμακες να διπλασιάζουμε το πεδίο οπτικής ροής.

Για $\rho = 5$, $\varepsilon = 0.05$ το πεδίο οπτικής ροής για τις ίδιες ενδεικτικές περιπτώσεις με πριν φαίνεται παρακάτω:





Συγκρίνοντας τα παραπάνω αποτελέσματα με εκείνα της μονοκλιμακωτής υλοποίησης, βλέπουμε ότι έχουμε καλύτερα αποτελέσματα αφού το πεδίο οπτικής ροής έχει σχεδόν μία συγκεκριμένη κατεύθυνση. Βέβαια, μπορούμε να παρατηρήσουμε πως αν και καλύτερος με αύξηση της παραμέτρου ρ στην μονοκλιμακωτή υλοποίηση μπορούμε να επιτύχουμε κοντινά αποτελέσματα όσον αφορά την συνολική ομοιομορφία της κατεύθυνσης του πεδίου οπτικής ροής.

Επίσης παρατηρούμε ότι στον πολυκλιμακωτό αλγόριθμο έχουμε πιο γρήγορη σύγκλιση. Το πλήθος των επαναλήψεων μέχρις ότου έχουμε σύγκλιση του αλγορίθμου είναι σημαντικά μικρότερο, πράγμα που παρατηρήσαμε δοκιμάζοντας να μειώσουμε τον «υποδειγμένο» αριθμό επαναλήψεων της μονοκλιμακωτής υλοποίησης.

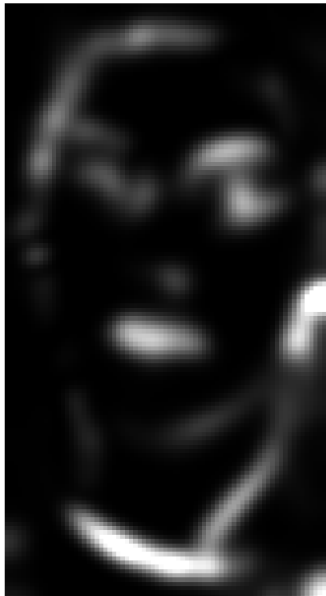
1.4 Υπολογισμός της Μετατόπισης του Προσώπου από το Πεδίο Οπτικής Ροής

Αυτό που μένει να πράξουμε πλέον, μιας και υπολογίσαμε το πεδίο οπτικής ροής, είναι να βρούμε κατά πόσο πρέπει να μετατοπίσουμε το Bounding Box από το i -οστό frame στο $(i+1)$ -οστό, ώστε να συνεχίσουμε να κάνουμε επιτυχώς το Face Tracking, δηλαδή σε κάθε επόμενο frame το Bounding Box να περικλείει το πρόσωπο με όσο το δυνατόν μεγαλύτερη ακρίβεια.

Η πρώτη σκέψη θα ήταν να υπολογίσουμε τη μέση τιμή όλων των διανυσμάτων οπτικής ροής και να μετατοπίσουμε το Bounding Box κατά αυτή τη μέση τιμή. Ωστόσο, αυτό θα μας δώσει ανακριβή αποτελέσματα, αφού τα διανύσματα του πεδίου οπτικής ροής έχουν μεγαλύτερο μήκος σε σημεία που ανήκουν σε περιοχές με έντονη πληροφορία και σχεδόν μηδενικό μήκος σε σημεία που ανήκουν σε περιοχές με ομοιόμορφη και επίπεδη υφή. Επομένως, μια λύση είναι να κάνουμε discard αυτά τα πολύ μικρού μήκους διανύσματα και να υπολογίσουμε τη μέση τιμή με τα εναπομείναντα. Πιο συγκεκριμένα, θα υπολογίσουμε την μέση τιμή των διανυσμάτων εκείνων των οποίων η ενέργεια $\|\mathbf{d}\|^2 = d_x^2 + d_y^2$ είναι μεγαλύτερη από κάποιο *Threshold*, και αυτή η μέση τιμή θα είναι ένα διάνυσμα $\mathbf{displ} = (displ_x, displ_y)$ κατά το οποίο θα μετατοπίσουμε, με πιο έγκυρο τρόπο, το Bounding Box.

Παρακάτω έχουμε την οπτικοποίηση της Ενέργειας των διανυσμάτων Οπτικής Ροής για τις προηγούμενες ενδεικτικές περιπτώσεις:

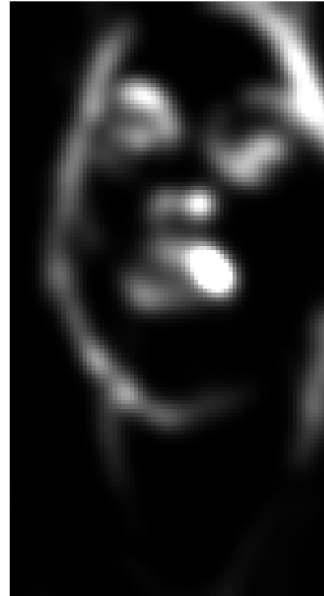
Optical Flow from Frame 1 to 2



Optical Flow from Frame 15 to 16



Optical Flow from Frame 29 to 30



Optical Flow from Frame 43 to 44



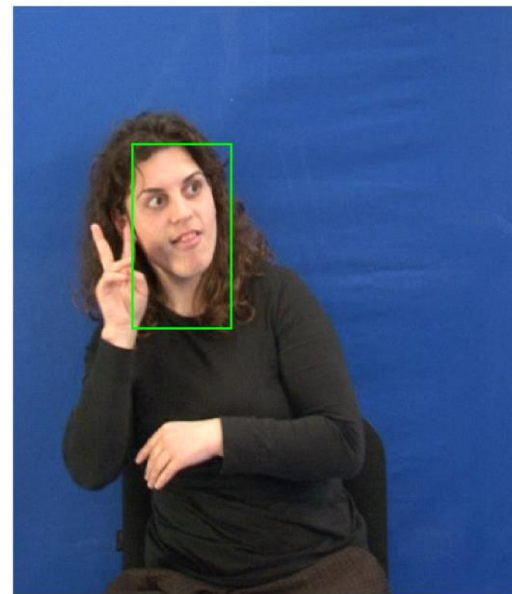
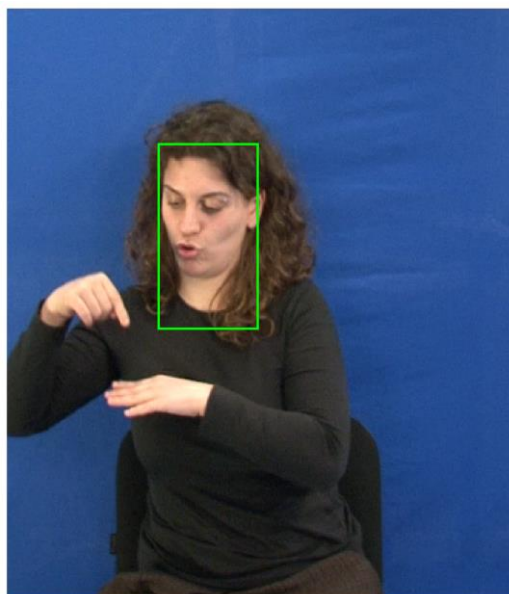
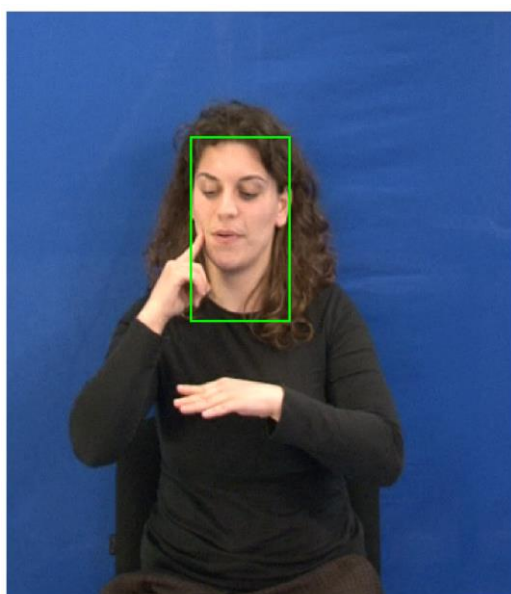
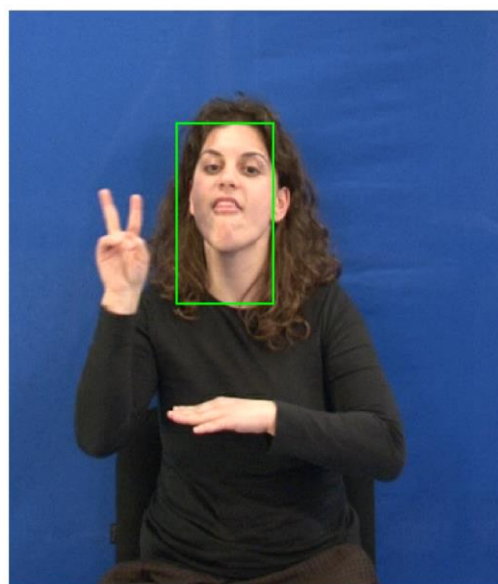
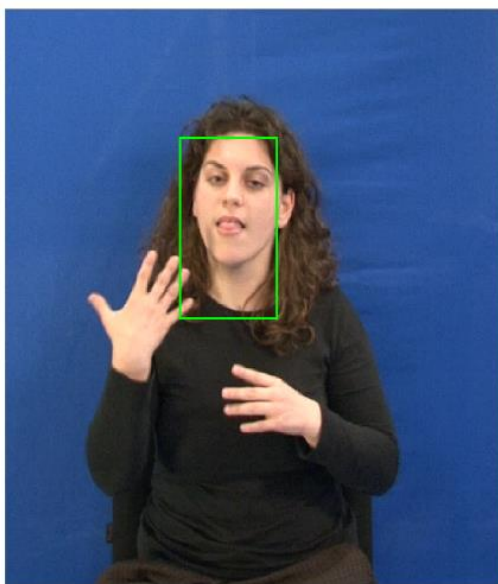
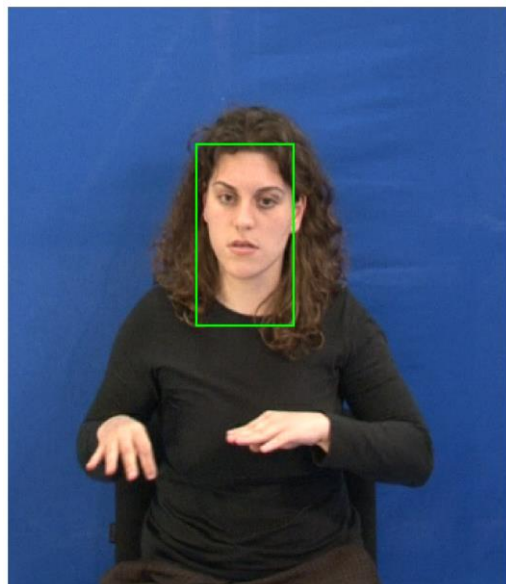
Optical Flow from Frame 57 to 58



Optical Flow from Frame 71 to 72



Τέλος, παραθέτουμε παρακάτω το τελικό αποτέλεσμα του Face Tracking με το Bounding Box να περικλείει επιτυχώς την περιοχή του προσώπου για κάποια ενδεικτικά frames:



Η πολυκλιμακωτή εκδοχή του αλγορίθμου δίνει τελείως όμοια αποτελέσματα, απλά σε μικρότερο χρόνο εκτέλεσης, καθώς, όπως προαναφέραμε, στην πολυκλιμακωτή εκδοχή έχουμε πολύ γρηγορότερη σύγκλιση.

Μέρος 2: Εντοπισμός Χωρο-χρονικών Σημείων Ενδιαφέροντος και Εξαγωγή Χαρακτηριστικών σε Βίντεο Ανθρωπίνων Δράσεων

Σκοπός: Σε αυτό το μέρος της εργαστηριακής άσκησης θέλουμε να εξάγουμε χωρο-χρονικά χαρακτηριστικά σε βίντεο με ανθρώπινες δράσεις προκειμένου να το κατηγοριοποιήσουμε αντιστοίχως. Ουσιαστικά καλούμαστε να βρούμε σημεία ενδιαφέροντος στο εκάστοτε βίντεο, ύστερα να χρησιμοποιήσουμε τοπικούς περιγραφητές πάνω σε αυτά για να συλλέξουμε τοπικά χαρακτηριστικά και τελικά να τα ενσωματώσουμε σε μια τελική αναπαράσταση global representation (συγκεκριμένα Bag of Visual Words στην περίπτωση μας) η οποία θα αναπαριστά τη στατιστική κατανομή τους.

Στην περίπτωση μας έχουμε ένα σύνολο βίντεο από 3 κλάσεις δράσεων (boxing, walking, running), 3 βίντεο για κάθε κατηγορία.

2.1 Χωρο-χρονικά Σημεία Ενδιαφέροντος

Αρχικά, όπως προαναφέρθηκε, θέλουμε να ανιχνεύσουμε χωρο-χρονικά σημεία και κλίμακες ενδιαφέροντος που αντιστοιχούν σε περιοχές που εμφανίζουν σύνθετη κίνηση ή απότομες αλλαγές.

2.1.1 Χρήση Harris Detector

Καλούμαστε να υλοποιήσουμε τον ανιχνευτή Harris σε 3 πλέον διαστάσεις πράγμα που αποτελεί απλά μια επέκταση του Harris-Stephens Corner Detector σε 3 διαστάσεις. Σε πρώτη φάση πρέπει να προσθέσουμε στον 2D δομικό τανυστή τη χρονική παράγωγο, οπότε παίρνουμε τον αντίστοιχο 3D δομικό τανυστή:

$$M(x, y, t; \sigma, \tau) = g(x, y, t; \sigma\sigma, \sigma\tau) * (\nabla L(x, y, t; \sigma, \tau))(\nabla L(x, y, t; \sigma, \tau))^T \Rightarrow$$
$$M(x, y, t; \sigma, \tau) = \begin{bmatrix} g * L_x^2 & g * (L_x L_y) & g * (L_x L_t) \\ g * (L_x L_y) & g * L_y^2 & g * (L_y L_t) \\ g * (L_x L_t) & g * (L_y L_t) & g * L_t^2 \end{bmatrix}$$

Όπου υπολογίσαμε τις μερικές παραγώγους (χωρικές και χρονικές) εφαρμόζοντας συνέλιξη με τον πυρήνα κεντρικών διαφορών $[-1 \ 0 \ 1]^T$ προσαρμοσμένο στην κατάλληλη διάσταση.

Το 3D κριτήριο γωνιότητας είναι το εξής:

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

2.1.2 Χρήση Gabor Detector

Καλούμαστε να υλοποιήσουμε τον ανιχνευτή Gabor ο οποίος βασίζεται στο χρονικό φίλτράρισμα του βίντεο με ένα ζεύγος Gabor φίλτρων:

$$h_{ev}(t; \tau, \omega) = -\cos(2\pi t\omega) \exp(-t^2/2\tau^2) \text{ και } h_{od}(t; \tau, \omega) = -\sin(2\pi t\omega) \exp(-t^2/2\tau^2)$$

Πρώτα, ωστόσο, πρέπει το βίντεο να υποστεί εξομάλυνση στις χωρικές διαστάσεις μέσω ενός 2D γκαουσιανού πυρήνα $g(x, y; \sigma)$. Ύστερα, θεωρούμε μέγεθος παραθύρου $[-2\tau, 2\tau]$ για να

υπολογίσουμε την κρουστική απόκριση των Gabor και εν συνεχεία κανονικοποιούμε με την $L1$ νόρμα. Όσον αφορά την συχνότητα ω του Gabor φίλτρου είναι $\omega = 4/\tau$.

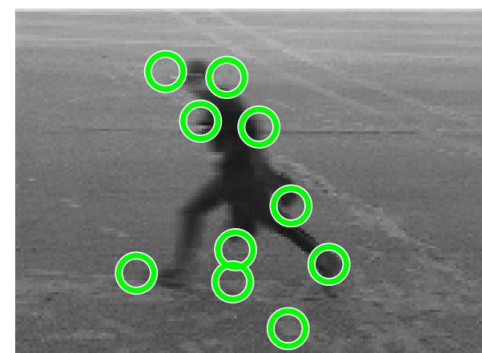
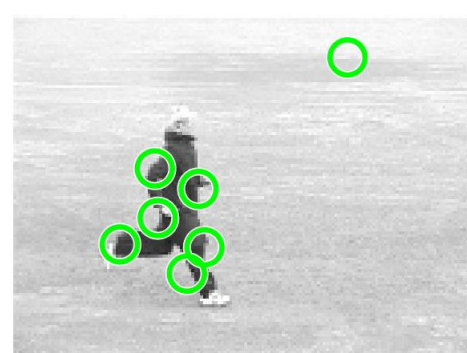
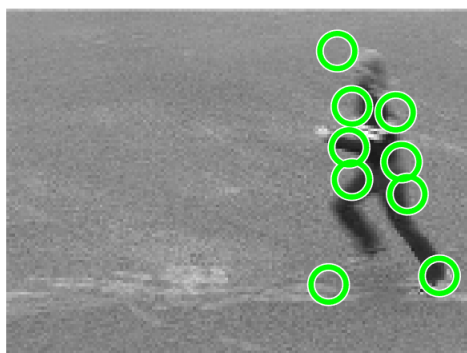
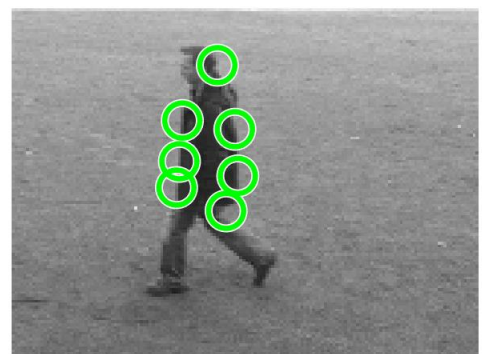
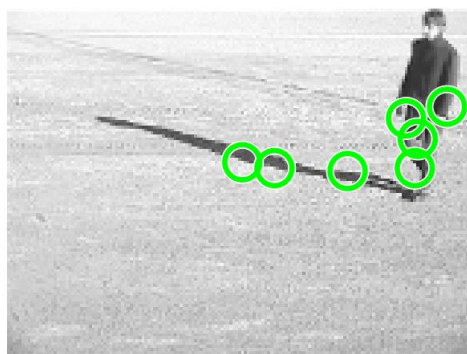
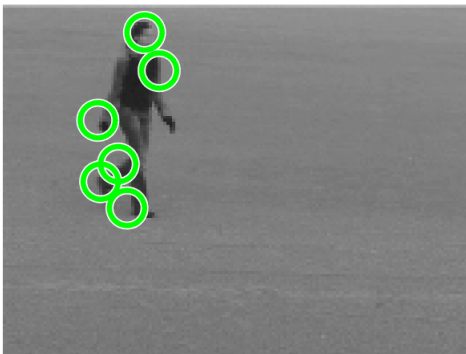
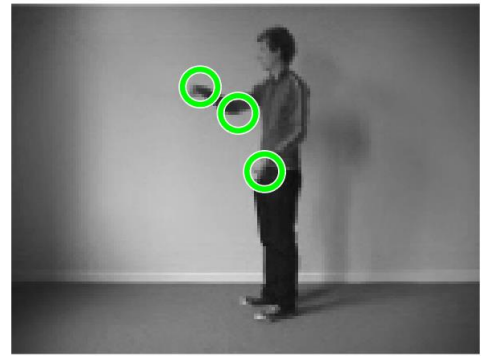
Το κριτήριο σημαντικότητας εδώ προκύπτει παίρνοντας την τετραγωνική ενέργεια της εξόδου για το ζεύγος Gabor φίλτρων:

$$H(x, y, t) = (I(x, y, t) * g * h_{ev})^2 + (I(x, y, t) * g * h_{od})^2$$

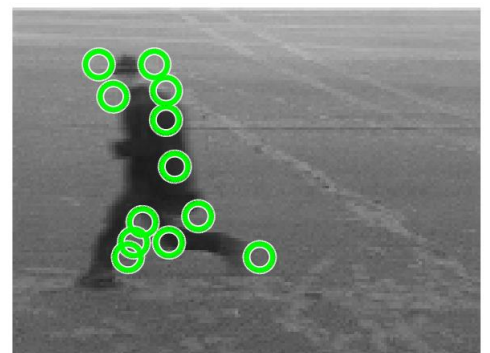
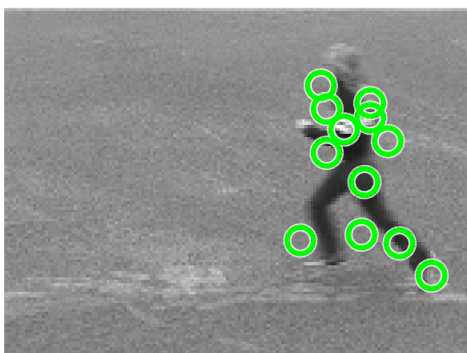
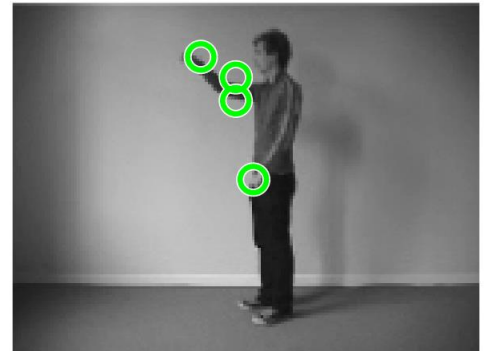
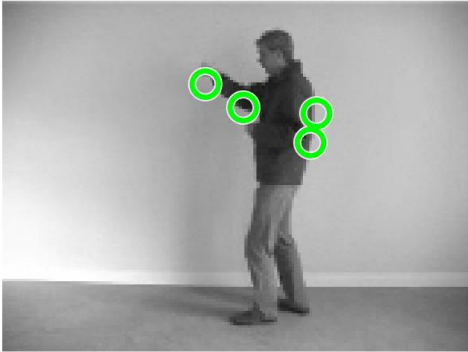
2.1.3 Τελική επιλογή σημείων ενδιαφέροντος

Υπολογίζουμε τα σημεία ενδιαφέροντος ως τα τοπικά μέγιστα του κριτηρίου σημαντικότητας. Για το εκάστοτε βίντεο, για λόγους απλότητας, θα πάρουμε από κάθε ανιχνευτή τα 500 σημεία με τις μεγαλύτερες τιμές του κριτηρίου σημαντικότητας όπως υποδεικνύεται από την εκφώνηση. Τα λαμβανόμενα αποτελέσματα για τον κάθε ανιχνευτή είναι τα εξής:

- Harris Detector sigma = 2, tau = 2.5:



➤ Gabor Detector $\sigma = 1.5$, $\tau = 3$:



2.2 Χωρο-χρονικοί Ιστογραφικοί Περιγραφητές

Θα χρησιμοποιήσουμε δύο ειδών χωρο-χρονικούς ιστογραφικούς περιγραφητές:

- HOG (Histogram of Oriented Gradients), που βασίζεται στον υπολογισμό ιστογραμμάτων της κατευθυντικής παραγώγου
- HOF (Histogram of Optical Flow), που βασίζεται στον υπολογισμό ιστογραμμάτων της οπτικής ροής
- HOG/HOF, δηλαδή συνδυασμό των δύο ανωτέρω περιγραφητών μέσω απλής συνένωσης τους

Ουσιαστικά, για κάθε βίντεο παίρνουμε τα σημεία ενδιαφέροντος που έχουμε βρει με βάση τον χρησιμοποιούμενο ανιχνευτή. Για κάθε ένα από αυτά τα χωροχρονικά σημεία ενδιαφέροντος υπολογίζουμε έναν χωροχρονικό ιστογραφικό περιγραφητή, δηλαδή για κάθε σημείο ενδιαφέροντος εξάγουμε ένα διάνυσμα που το περιγράφει (με τρόπο τέτοιο που αντιστοιχεί στον ιστογραφικό περιγραφητή).

Στη συνέχεια, αφού έχουμε ένα τέτοιο διάνυσμα για κάθε σημείο ενδιαφέροντος και για κάθε βίντεο, καλούμαστε να υπολογίσουμε την τελική αναπαράσταση global representation για κάθε βίντεο υλοποιώντας την Bag of Visual Words τεχνική. Πρέπει να δοθεί ωστόσο προσοχή στο γεγονός ότι εδώ δεν έχουμε train και test data, παραμόνο 9 βίντεο για το οποίο αρχικά υπολογίζουμε τις οπτικές λέξεις του λεξικού. Για αυτόν τον σκοπό θα χρησιμοποιήσουμε τον αλγόριθμο kmeans. Ο αλγόριθμός kmeans θα μας επιστρέψει έναν αριθμό από κέντρα (όσα εμείς απαιτήσουμε ως είσοδο). Στη συνέχεια, για κάθε βίντεο, αντιστοιχίζουμε το διάνυσμα του χρησιμοποιούμενου ιστογραφικού περιγραφητή του κάθε σημείου ενδιαφέροντος στο κέντρο από το οποίο απέχει την μικρότερη ευκλείδεια απόσταση, δηλαδή πρακτικά αντιστοιχίζουμε κάθε σημείο ενδιαφέροντος σε μια οπτική λέξη του λεξικού, και με αυτόν τον τρόπο λαμβάνουμε για κάθε βίντεο το ιστογράμμα εμφάνισης ή αλλιώς συχνότητας εμφάνισης της κάθε οπτικής λέξης του λεξικού μας. Τέλος, κανονικοποιούμε και με την $L2$ νόρμα, και έχουμε πλέον για κάθε βίντεο την τελική αναπαράσταση global representation με BoVW.

2.3 Κατασκευή Δενδρογράμματος για τον Διαχωρισμό των Δράσεων

Για την οπτικοποίηση των αποτελεσμάτων μας θα προσπαθήσουμε να κατηγοριοποιήσουμε τα δοθέντα βίντεο στις κλάσεις τους. Αυτό θα φανεί με την οπτικοποίηση της απόστασης των διανυσμάτων χαρακτηριστικών μέσω της κατασκευής ενός δενδρογράμματος αποστάσεων στο οποίο θα φαίνεται ο διαχωρισμός μεταξύ των 3 κλάσεων.

Θα χρησιμοποιήσουμε την απόσταση χ^2 η οποία είναι κατάλληλη για ιστογράμματα και ορίζεται ως εξής:

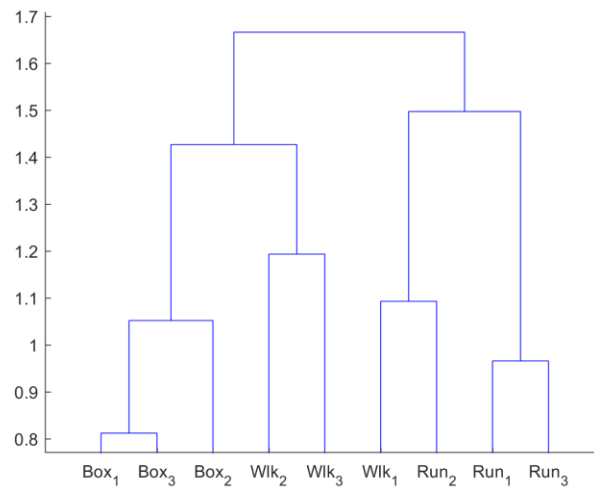
$$D(H_i, H_j) = \sum_{n=1}^K \frac{(h_{in} - h_{jn})^2}{h_{in} + h_{jn}}$$

όπου H_i, H_j δύο ιστογράμματα με $H_i = \{h_{i1}, h_{i2}, \dots, h_{iK}\}$, $H_j = \{h_{j1}, h_{j2}, \dots, h_{jK}\}$

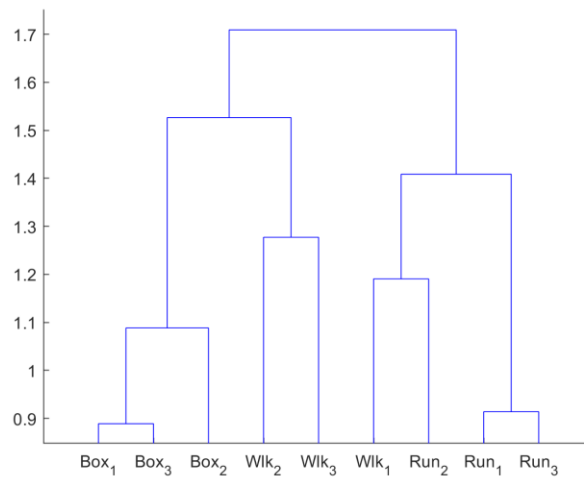
και K το πλήθος των κέντρων.

Εφαρμόζουμε τώρα όλα τα παραπάνω για 6 διαφορετικούς συνδυασμούς χωροχρονικού ανιχνευτή με χωροχρονικό ιστογραφικό περιγραφητή και λαμβάνουμε τα παρακάτω αποτελέσματα:

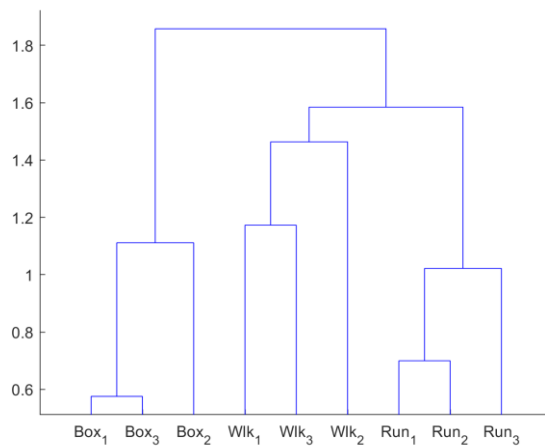
➤ Harris Detector με HOG



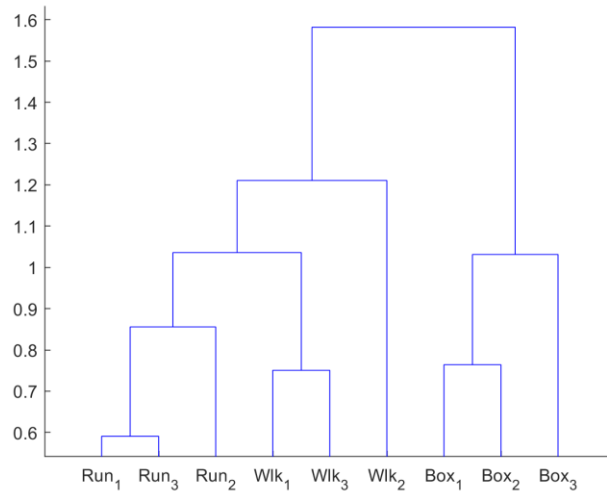
➤ Harris Detector με HOF



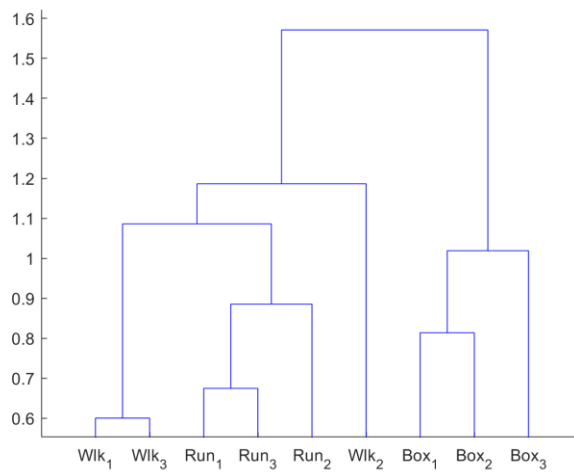
➤ Harris Detector με συνδυασμό HOG/HOF



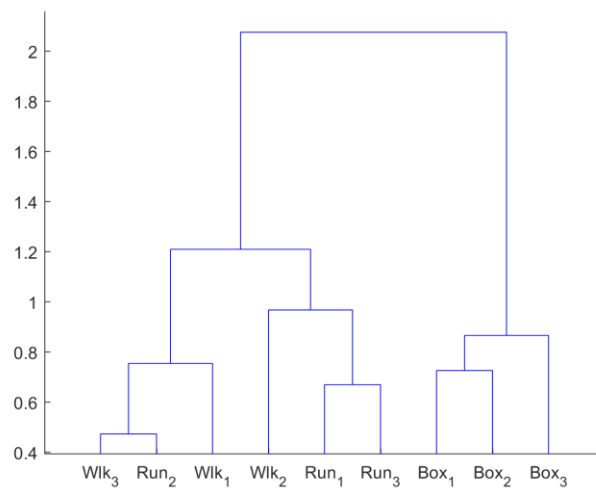
➤ Gabor Detector με HOG



➤ Gabor Detector με HOF



○ Gabor Detector με συνδυασμό HOG/HOF



Παρατηρώντας τα παραπάνω δενδρογράμματα βλέπουμε ότι το clustering με χρήση του Harris Detector και συνδυασμό HOG/HOF δίνει τα καλύτερα αποτελέσματα, καθώς γίνεται καλύτερος διαχωρισμός των 3 κατηγοριών.

Σε όλες τις υπόλοιπες περιπτώσεις βλέπουμε ότι γίνεται καλός διαχωρισμός μεταξύ του boxing και των walking-running, μιας και διαισθητικά καταλαβαίνει κανείς ότι είναι προφανείς οι διαφοροποιήσεις. Ωστόσο, αποτυγχάνουν στον διαχωρισμό του walking από το running. Αυτό βέβαια φαίνεται να είναι σε έναν βαθμό αναμενόμενο αφού το τρέξιμο και το περπάτημα σε ένα βίντεο μοιάζουν περισσότερο από το box.