



Όραση Υπολογιστών
2^η Εργαστηριακή Άσκηση

Ημερομηνία 03/06/2021

Ονοματεπώνυμο	A.M.
Σιαφαρίκας Χρήστος	03117123
Τσιγγέλης Κωνσταντίνος	03117149

**Θέμα: Εκτίμηση Οπτικής Ροής (Optical Flow) και Εξαγωγή
Χαρακτηριστικών σε Βίντεο για Αναγνώριση Δράσεων**

Περιεχόμενα

Μέρος 1. Παρακολούθηση Προσώπου και Χεριών με Χρήση της.....	3
Μεθόδου Οπτικής Ροής των Lucas-Kanade.....	3
1.1 Ανίχνευση Δέρματος Προσώπου και Χεριών	3
1.2 Παρακολούθηση Προσώπου και Χεριών	7
(1.2.1) Υλοποίηση του Αλγόριθμου των Lucas-Kanade	7
(1.2.2) Υπολογισμός της Μετατόπισης των Παραθύρων από τα Διανύσματα	8
Οπτικής Ροής	12
(1.2.3) Πολυ-Κλιμακωτός Υπολογισμός Οπτικής Ροής	19
Μέρος 2. Εντοπισμός Χωρο-χρονικών Σημείων Ενδιαφέροντος και Εξαγωγή Χαρακτηριστικών σε Βίντεο Ανθρωπίνων Δράσεων.....	21
2.1 Χωρο-χρονικά Σημεία Ενδιαφέροντος.....	21
(2.1.1)	21
(2.1.2)	23
(2.1.3)	24
2.2 Χωρο-χρονικοί Ιστογραφικοί Περιγραφητές.....	26
(2.2.1)	26
(2.2.2)	26
(2.2.3)	26
2.3 Κατασκευή Bag of Visual Words και χρήση Support Vector	27
Machine για την ταξινόμηση δράσεων	27
(2.3.1)	27
(2.3.2)	27
(2.3.3)	27
(2.3.4)	27
(2.3.5)	27
Βιβλιογραφία-Αναφορές	27

Μέρος 1. Παρακολούθηση Προσώπου και Χεριών με Χρήση της

Μεθόδου Οπτικής Ροής των Lucas-Kanade

Σε αυτό το μέρος, ο σκοπός του είναι να υλοποιηθεί ένα σύστημα παρακολούθησης προσώπου και χεριών (Face and Hands Tracking) σε μια ακολουθία βίντεο.

1.1 Ανίχνευση Δέρματος Προσώπου και Χεριών

Σε αυτό το ερώτημα υλοποιείται μια συνάρτηση για την ανίχνευση των περιοχών δέρματος του προσώπου και των χεριών. Η υλοποίησή της θα γίνει στο χρωματικό χώρο YCbCr, ο οποίος αποτελείται από τρεις συνιστώσες:

- Τη φωτεινότητα Y
- Τη χρωματική διαφορά ως προς το μπλε Cb
- Τη χρωματική διαφορά ως προς το κόκκινο Cr

Με μετατροπή των εικόνων από RGB σε YCbCr, οι εικόνες βρίσκονται στον επιθυμητό χώρο στον οποίο θα αξιοποιηθούν τα κανάλια Cb, Cr, τα οποία περιέχουν την πληροφορία για το δέρμα. Το χρώμα του δέρματος μοντελοποιείται με μια δισδιάστατη Γκαουσιανή κατανομή, όπως περιγράφεται από την ακόλουθη σχέση:

$$P(c=skin) = \frac{1}{\sqrt{|\Sigma|(2\pi)^2}} * e^{-\frac{1}{2}(c-\mu)\Sigma^{-1}(c-\mu)'} \quad (1.1)$$

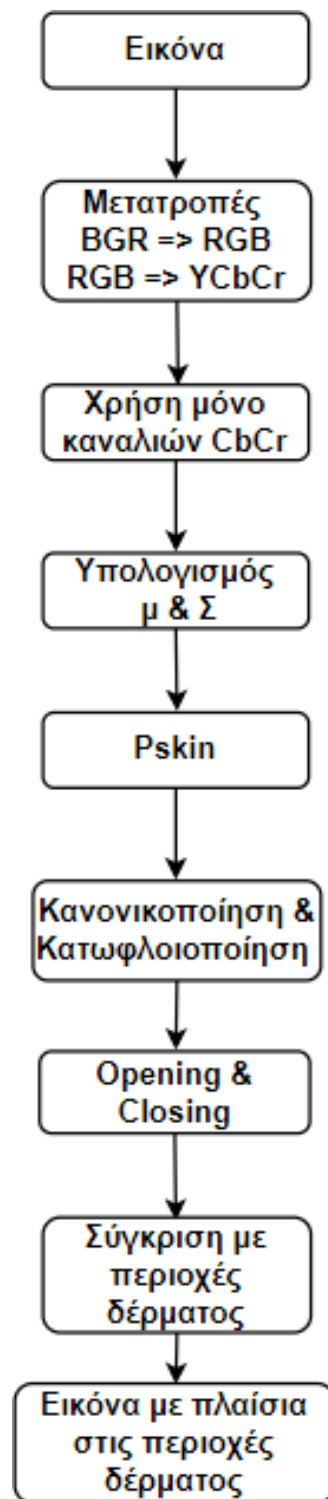
Όπου c : το διάνυσμα των τιμών Cb και Cr για κάθε σημείο της εικόνας

μ : το 2x1 διάνυσμα μέσης τιμής $\mu = [\mu_{cb}, \mu_{cr}]^T$ από τα δείγματα δέρματος

Σ : το 2x2 πίνακα συνδιακύμανσης από τα δείγματα δέρματος

Η παραπάνω σχέση μοντελοποιείται προγραμματιστικά με τη βοήθεια της συνάρτησης `multivariate_normal.pdf` της βιβλιοθήκης `scipy`.

Συνοπτικά ο αλγόριθμος, που ακολουθήθηκε για την ανίχνευση δέρματος προσώπου και χεριών, περιγράφεται από το ακόλουθο σχεδιάγραμμα.



Υλοποιώντας προγραμματιστικά τα στάδια του παραπάνω αλγόριθμου ανίχνευσης δέρματος προσώπου και χεριών προκύπτουν τα ακόλουθα αποτελέσματα.

Στην αρχή διαβάζονται οι εικόνες και εμφανίζονται σε RGB και YCbCr κανάλια αντιστοίχως.

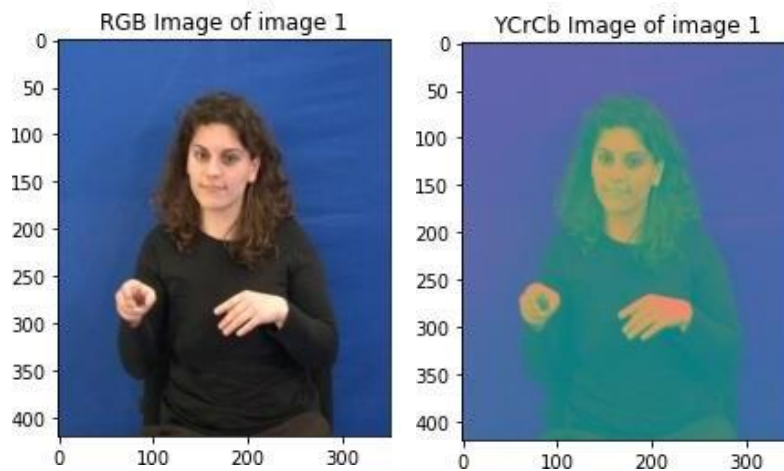


Figure 1. Η πρώτη εικόνα από το βίντεο της νοηματίστριας σε RGB (αριστερά) και σε YCbCr (δεξιά)

Ύστερα, κρατώντας μονό τα κανάλια, που περιέχουν χρήσιμη πληροφορία για το δέρμα, υπολογίζεται η πιθανότητα να υπάρχει δέρμα σε ορισμένο pixel στην εικόνα, όπως φαίνεται ακολούθως.

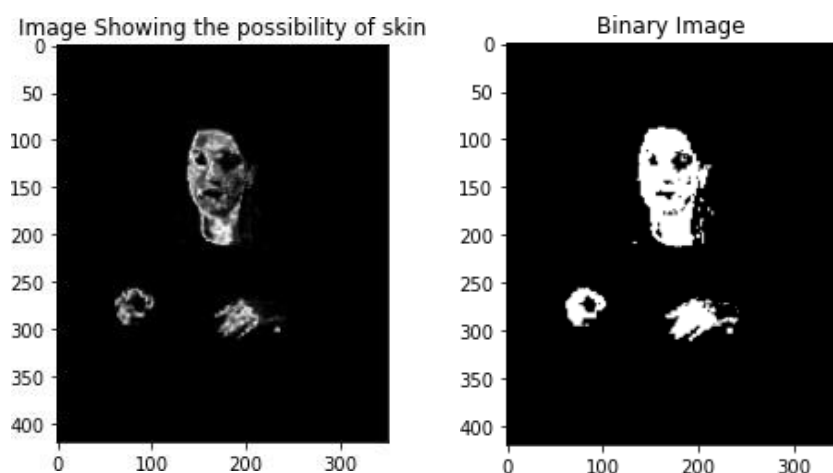


Figure 2. Η πιθανότητα να υπάρχει δέρμα στην πρώτη εικόνα από το βίντεο της νοηματίστριας σε grayscale (αριστερά) και ύστερα από την κανονικοποίηση και την κατωφλιοποίηση της (δεξιά)

Έπειτα, με κατάλληλη μορφολογική επεξεργασία, δηλαδή με χρήση μικρού πυρήνα για το opening για να καλυφθούν οι τρύπες, και με χρήση μεγάλου δομικού στοιχείου για το closing για να εξαλειφθούν μικρές περιοχές, προκύπτει το ακόλουθο αποτέλεσμα.

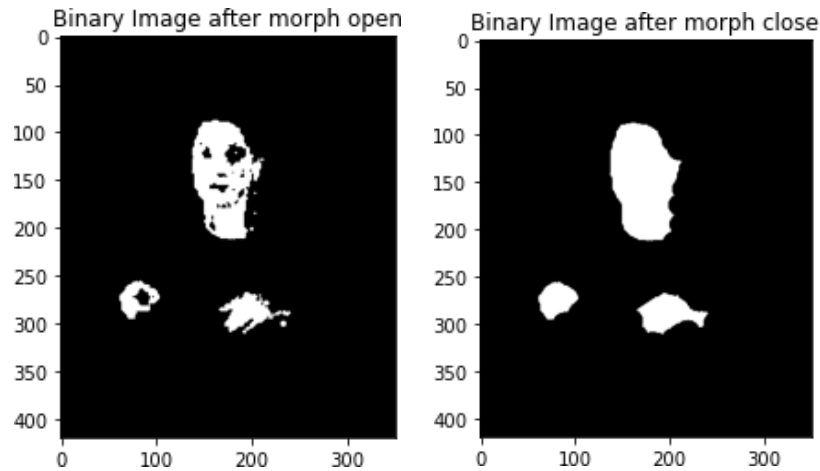


Figure 3. Η δυαδική εικόνα της νοσηματίστριας μετά την εφαρμογή μορφολογικής επεξεργασίας με opening (αριστερά) και με closing (δεξιά)

Τελικά, το αποτέλεσμα, που προκύπτει από την υλοποίηση του αλγορίθμου ανίχνευσης δέρματος φαίνεται στην παρακάτω εικόνα.

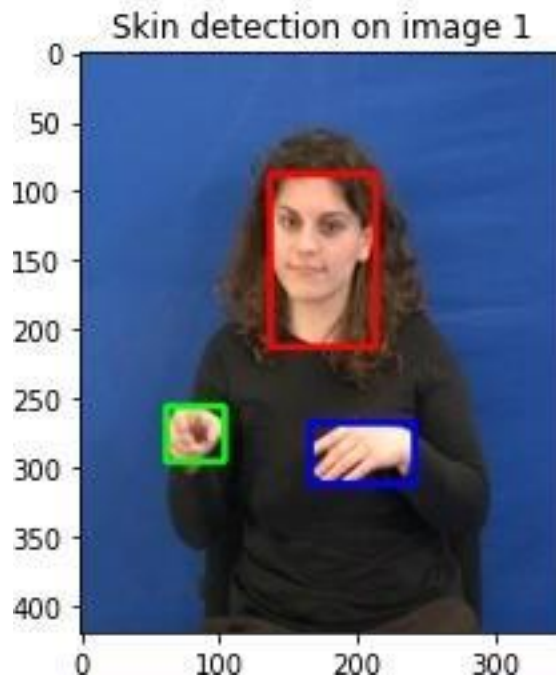


Figure 4. Η εικόνα της νοηματίστριας στην οποία έχουν ανιχνευθεί οι περιοχές δερματός της (πρόσωπο, δεξί και αριστερό χέρι)

1.2 Παρακολούθηση Προσώπου και Χεριών

Σε αυτό το στάδιο της εργαστηριακής άσκησης θα υλοποιηθούν οι ακόλουθοι αλγόριθμοι υπολογισμού της οπτικής ροής.

(1.2.1) Υλοποίηση του Αλγόριθμου των Lucas-Kanade

Σε αυτό το ερώτημα υλοποιείται ο αλγόριθμος Lucas – Kanade. Ο αλγόριθμος αυτός ανήκει στην κατηγορία αλγορίθμων, που εκτιμούν την οπτική ροή της εικόνας, επιλύοντας την εξίσωση διατήρησης της φωτεινότητας. Ο αλγόριθμος Lucas – Kanade στηρίζεται στην παραδοχή ότι το διάνυσμα οπτικής ροής εντός ενός παραθύρου γύρω από κάθε pixel.

Με αυτή την παραδοχή, ο αλγόριθμος Lucas – Kanade εξάγει μια εκτίμηση για το διάνυσμα οπτικής ροής d , την οποία συνεχώς βελτιώνει σύμφωνα με τις ακόλουθες σχέσεις.

$$d_{i+1} = d_i + u \quad (1.2)$$

Όπου u : διάνυσμα βελτίωσης το οποίο περιγράφεται από τη σχέση 1.3

$$u(x) = u_1(x) u_2(x) \quad (1.3)$$

Όπου u_1 : πίνακας που περιγράφεται από την πρώτη σχέση των (1.4)

u_2 : πίνακας που περιγράφεται από τη δεύτερη σχέση των (1.4)

$$u_1(x) = \begin{bmatrix} (G\rho * A1^2)(x) + \varepsilon & (G\rho * A1A2)(x) \\ (G\rho * A1A2)(x) & (G\rho * A2^2)(x) + \varepsilon \end{bmatrix}^{-1} \quad (1.4)$$

$$u_2(x) = \begin{bmatrix} (G\rho * A1E)(x) \\ (G\rho * A2E)(x) \end{bmatrix}$$

Όπου $A(x)$: πίνακας που περιγράφεται από την πρώτη σχέση των (1.5)

$E(x)$: πίνακας που περιγράφεται από τη δεύτερη σχέση των (1.5)

$$A(x) = \begin{bmatrix} A_1(x) & A_2(x) \end{bmatrix} = \begin{bmatrix} \frac{\partial \ln-1(x+di)}{\partial x} & \frac{\partial \ln-1(x+di)}{\partial y} \end{bmatrix} \quad (1.5)$$

$$u_2(x) = I_n(x) - I_{n-1}(x + d_i)$$

Συνεπώς, μοντελοποιώντας προγραμματιστικά τη σχέση για τη βελτίωση του διανύσματος οπτικής ροής προκύπτουν τα ακόλουθα αποτελέσματα για την εκτίμηση της οπτικής ροής κατά τη μετάβαση από το πρώτο frame του βίντεο στο δεύτερο.

Παρακάτω, φαίνεται η οπτική ροή των σημείων ενδιαφέροντος του προσώπου της νοηματίστριας. Από αυτές τις εικόνες φαίνεται ότι η νοηματίστρια πραγματοποιεί μια κίνηση με το κεφάλι προς τα πάνω και δεξιά. Αυτή η κίνηση, όπως φαίνεται, καταγράφεται επιτυχώς από τον αλγόριθμο αφού τα βελάκια στο διάγραμμα οπτικής ροής δείχνουν πάνω και δεξιά.

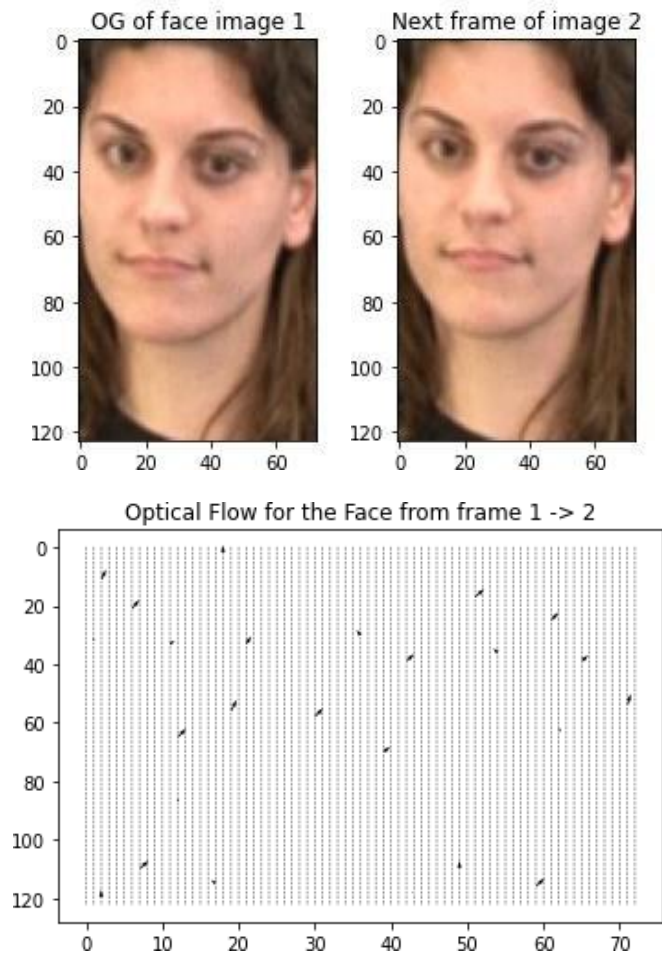


Figure 5. Το διάγραμμα οπτικής ροής του προσώπου της νοηματίστριας για τη μετάβαση από το 1^ο frame του βίντεο στο 2^ο

Παρακάτω, φαίνεται η οπτική ροή των σημείων ενδιαφέροντος του αριστερού χεριού της νοηματίστριας. Από τις παραπάνω εικόνες επαληθεύεται και σε αυτή την περίπτωση η ορθή λειτουργία του αλγορίθμου Lucas – Kanade .

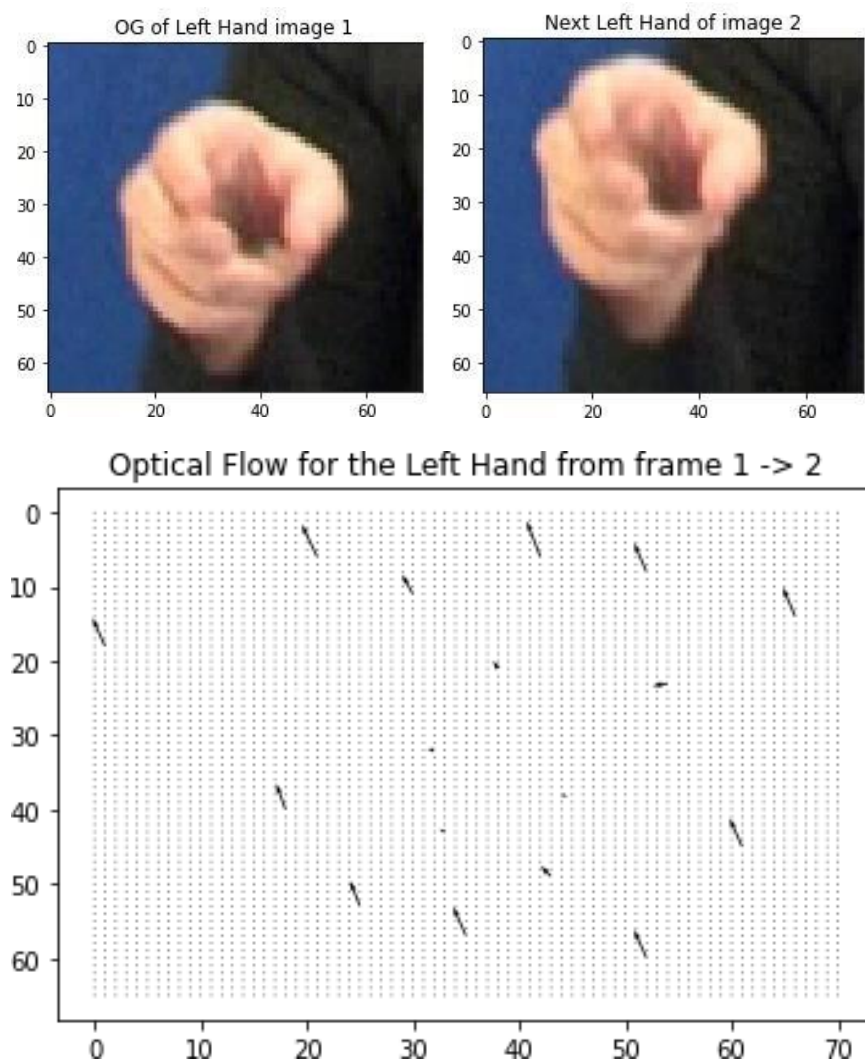


Figure 6.Το διάγραμμα οπτικής ροής του αριστερού χεριού της νοηματίστριας για τη μετάβαση από το 1^ο frame του βίντεο στο 2^ο

Παρακάτω, φαίνεται η οπτική ροή των σημείων ενδιαφέροντος του δεξιού χεριού της νοηματίστριας. Από τις παραπάνω εικόνες επαληθεύεται και σε αυτή την περίπτωση η ορθή λειτουργία του αλγορίθμου Lucas – Kanade .

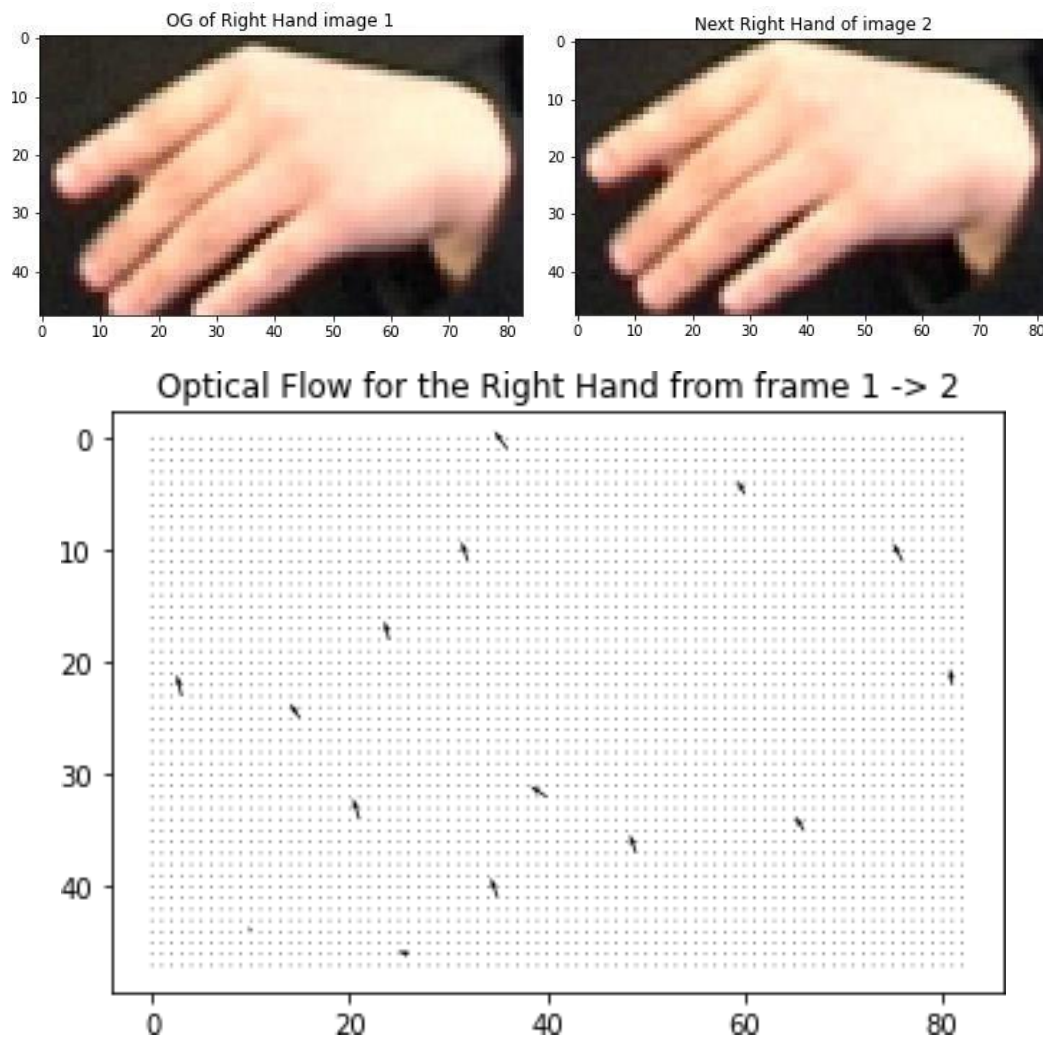
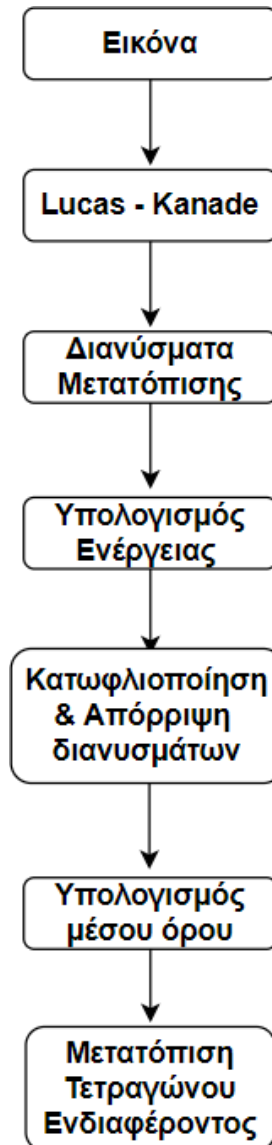


Figure 7.Το διάγραμμα οπτικής ροής του δεξιού χεριού της νοηματίστριας για τη μετάβαση από το 1^ο frame του βίντεο στο 2^ο

(1.2.2) Υπολογισμός της Μετατόπισης των Παραθύρων από τα Διανύσματα Οπτικής Ροής

Σε αυτό το ερώτημα υλοποιείται η συνάρτηση που θα δέχεται σαν είσοδο τα διανύσματα της οπτικής ροής και θα υπολογίζει το τελικό διάνυσμα μετατόπισης του ορθογωνίου. Για τον υπολογισμό του τελικού διανύσματος μετατόπισης του ορθογωνίου ακολουθήθηκε η λογική, που περιγράφεται στο ακόλουθο σχεδιάγραμμα.



Μοντελοποιώντας προγραμματιστικά τη λογική, που περιγράφεται στο παραπάνω σχεδιάγραμμα, προκύπτουν τα ακόλουθα αποτελέσματα.

Παρακάτω, φαίνεται η μετατόπιση του τετραγώνου ενδιαφέροντος του προσώπου της νοηματίστριας από το πρώτο frame του βίντεο στο δεύτερο.

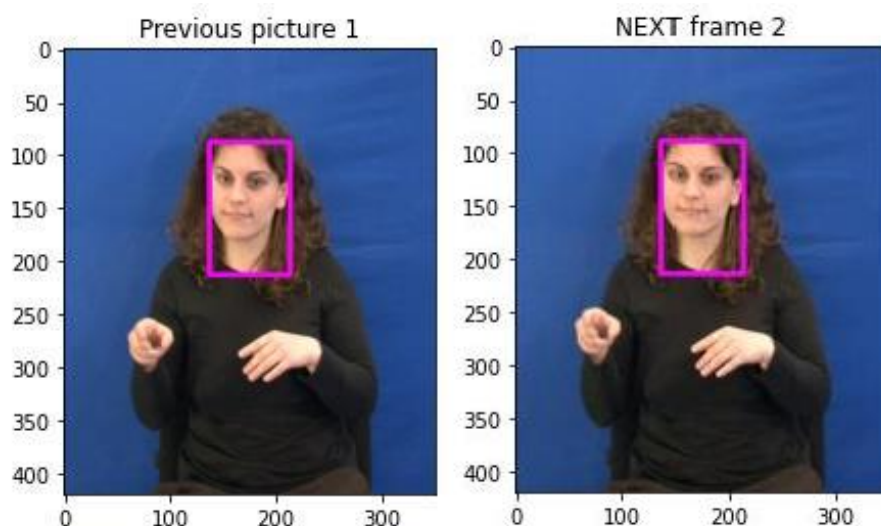


Figure 8.Οπτικοποίηση μετατόπισης τετραγώνου ενδιαφέροντος για το πρόσωπο με τη βοήθεια του αλγορίθμου υπολογισμού οπτικής ροής του Lucas - Kanade

Παρακάτω, φαίνεται η μετατόπιση του τετραγώνου ενδιαφέροντος του αριστερού χεριού της νοηματίστριας από το πρώτο frame του βίντεο στο δεύτερο.

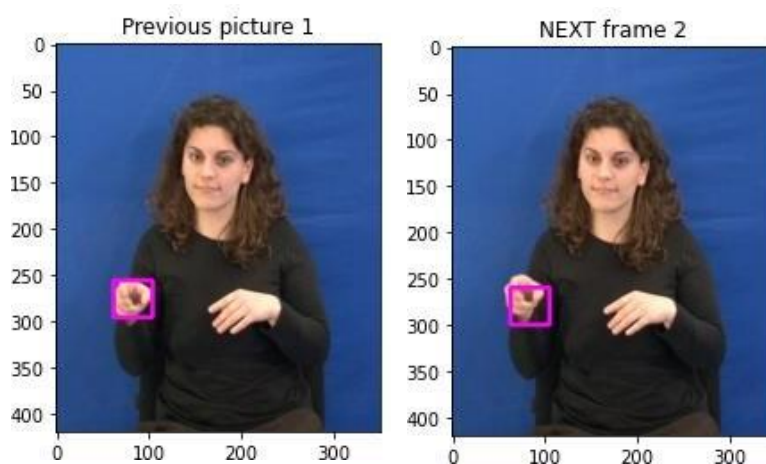


Figure 9.Οπτικοποίηση μετατόπισης τετραγώνου ενδιαφέροντος για το αριστερό χέρι με τη βοήθεια του αλγορίθμου υπολογισμού οπτικής ροής του Lucas - Kanade

Παρακάτω, φαίνεται η μετατόπιση του τετραγώνου ενδιαφέροντος του δεξιού χεριού της νοηματίστριας από το πρώτο frame του βίντεο στο δεύτερο.

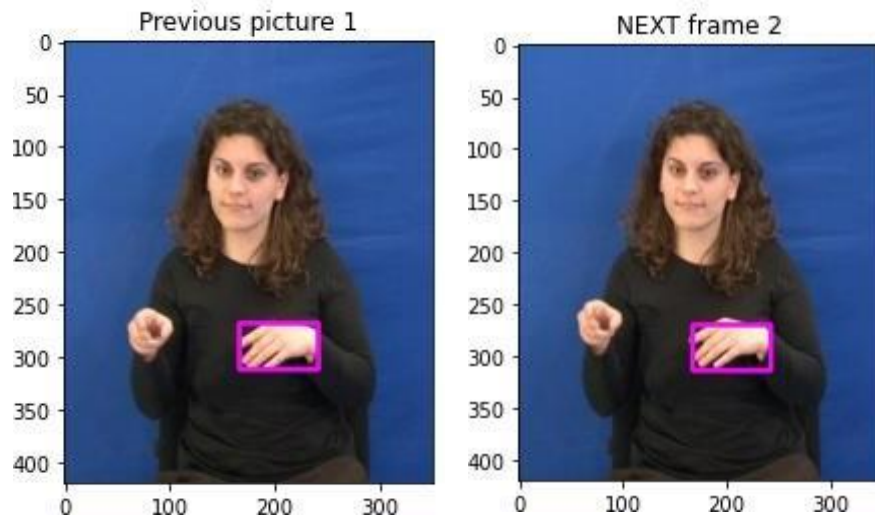


Figure 10.Οπτικοποίηση μετατόπισης τετραγώνου ενδιαφέροντος για το δεξί χέρι με τη βοήθεια του αλγορίθμου υπολογισμού οπτικής ροής του Lucas - Kanade

Οι παραπάνω εικόνες προέκυψαν για τιμές παραμέτρων $[\rho, \epsilon, rep] = [2, 0.1, 10]$ και αρχικές εκτιμήσεις $[dx_0, dy_0] = [0, 0]$. Κρατώντας, λοιπόν, σταθερές τις εκτιμήσεις dx_0, dy_0 και μεταβάλλοντας τις άλλες παραμέτρους θα αναζητηθεί ο καλύτερος συνδυασμός για τον αλγόριθμο Lucas - Kanade.

Κρατώντας σταθερές τις μετρικές ρ , ε και μεταβάλλοντας των αριθμό των επαναλήψεων σε τιμές $\text{rep} = \{10, 20, 100\}$ προκύπτουν τα ακόλουθα αποτελέσματα.

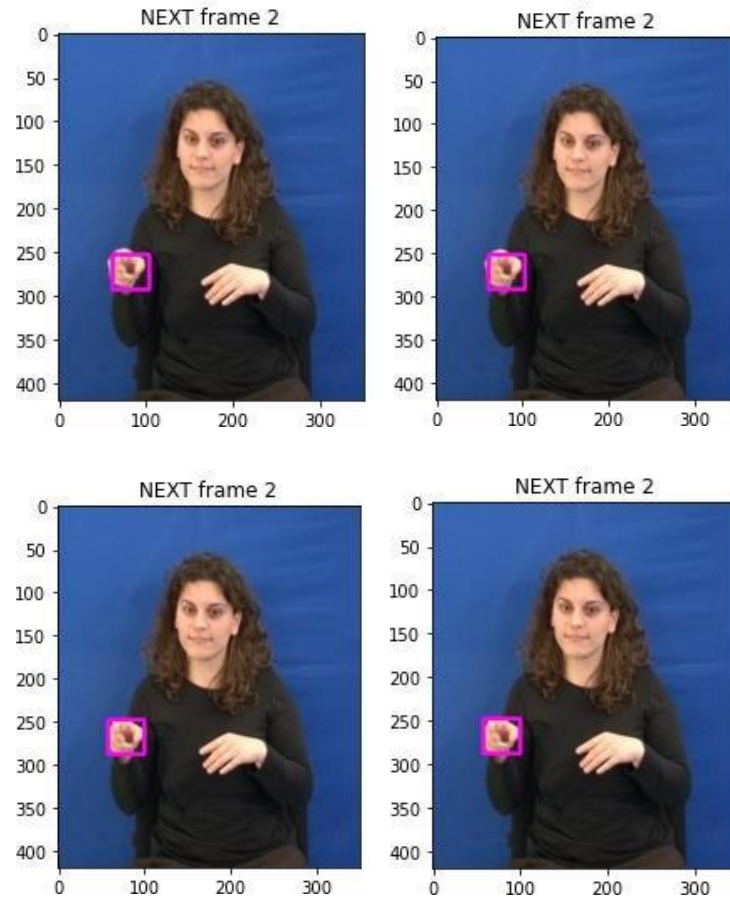


Figure 11.Οπτικοποίηση μετατόπισης τετραγώνου ενδιαφέροντος για επανάληψη βελτίωσης διανύσματος οπτικής ροής ίση με 10 (αριστερά) , με 20 (μεσαία) και με 100 (δεξιά)

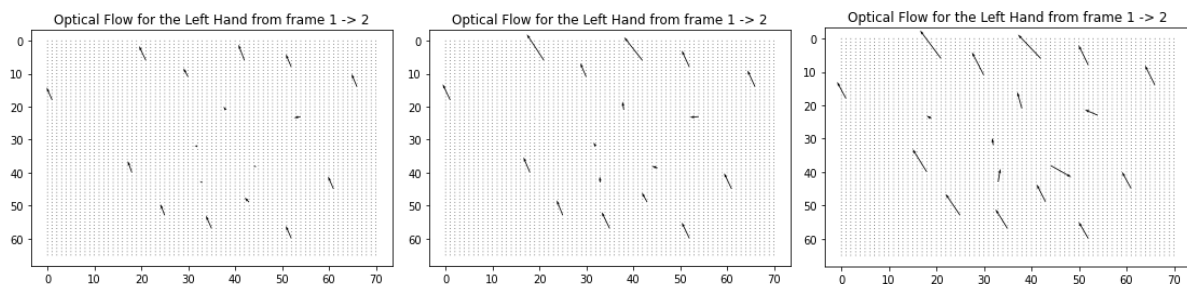


Figure 12.Οπτικοποίηση διανύσματος οπτικής ροής για επανάληψη βελτίωσης διανύσματος οπτικής ροής ίση με 10 (αριστερά) , με 20 (μεσαία) και με 100 (δεξιά)

Από τις παραπάνω εικόνες φαίνεται ότι η εκτίμηση του διανύσματος d , από ένα κατώφλι επανάληψης και μετά δεν βελτιώνει d , όμως για αύξηση αριθμού επαναλήψεων σε μικρές τιμές παρατηρείται σημαντική βελτίωση στον εντοπισμό της ενδιαφέρουσας περιοχής. Μιας και είναι μια χρονοβόρα διαδικασία είναι επιθυμητός ο αριθμός των επαναλήψεων να είναι όσο το δυνατόν πιο μικρός. Με διάφορες δοκιμές παρατηρήθηκε ότι το κατώφλι από το οποίο και μετά σταματάει να βελτιώνεται η εκτίμηση είναι περίπου στις 50 επαναλήψεις.

Άρα, η ανάλυση από εδώ και πέρα αν και θα έπρεπε να γίνεται για $rep = 50$, γίνεται για $rep=10$ για να είναι πιο εύκολα παρατηρήσιμες οι επιδράσεις των υπολοίπων παραμέτρων.

Κρατώντας σταθερές τις μετρικές rep , ϵ και μεταβάλλοντας των αριθμό των επαναλήψεων σε τιμές $\rho = \{1, 2, 3, 4, 5\}$ προκύπτουν τα ακόλουθα αποτελέσματα.

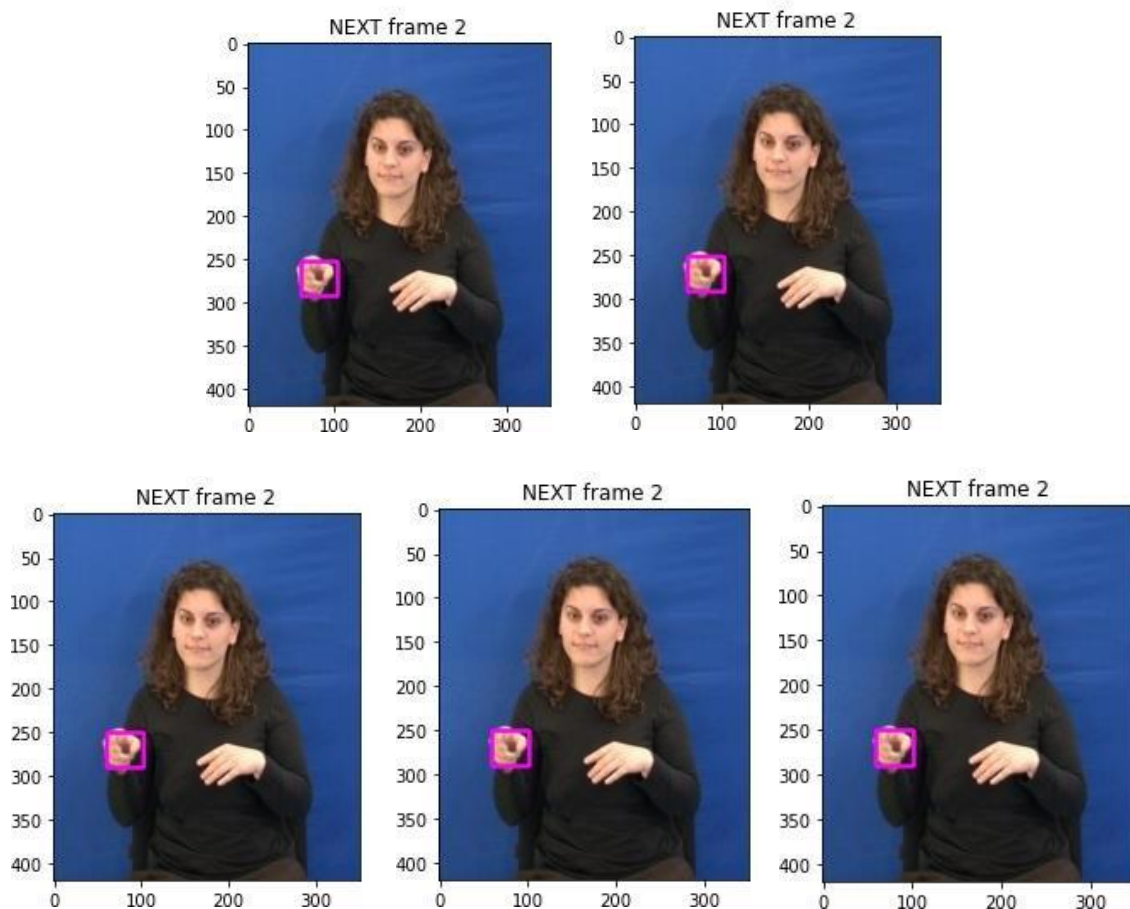
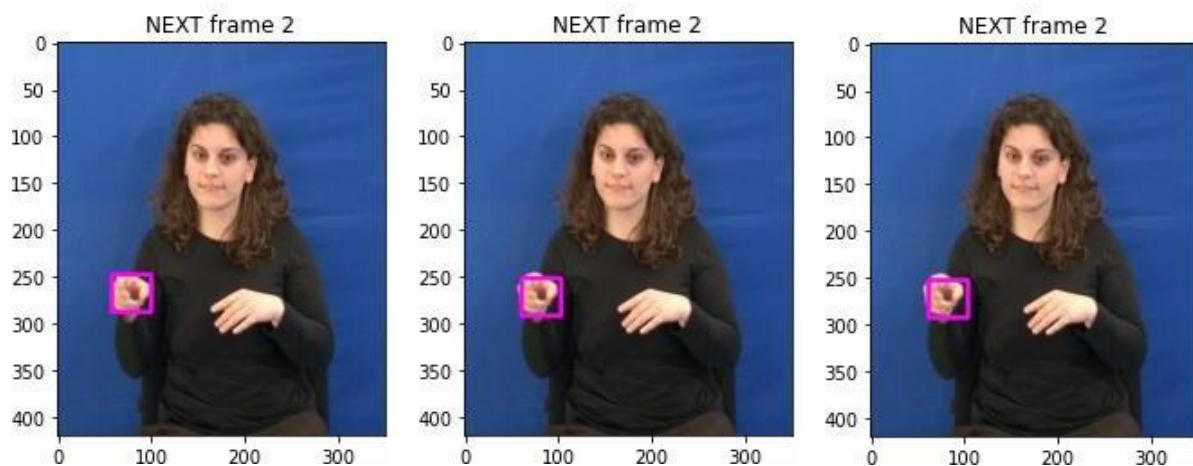


Figure 13.Οπτικοποίηση μετατόπισης τετραγώνου ενδιαφέροντος για διάφορες τιμές της τυπικής απόκλισης του Γκαουσιανού φίλτρου, για $\rho = \{1, 2, 3, 4, 5\}$ ξεκινώντας από πάνω αριστερά

Από τις παραπάνω εικόνες φαίνεται ότι όσο αυξάνεται η τυπική απόκλιση ρ του Γκαουσιανού φίλτρου τόσο αυξάνεται η ποιότητα της εκτίμησης του διανύσματος d , μιας και φαίνεται το τετράγωνο ενδιαφέροντος να μετατοπίζεται σε κατεύθυνση που έχει περισσότερο δέρμα. Αυτή η παρατήρηση φαίνεται να ισχύει μέχρι το ρ να πάρει την τιμή 4, καθώς μετά φαίνεται ελάχιστη μετατόπιση σε σημείο μη επιθυμητό. Άρα, ως ιδανική τιμή για την παράμετρο ρ επιλέγεται η τιμή 4.

Κρατώντας σταθερές τις μετρικές ρ , σ_{ep} και μεταβάλλοντας των αριθμό των επαναλήψεων σε τιμές $\varepsilon = \{0.01, 0.05, 0.1\}$ προκύπτουν τα ακόλουθα αποτελέσματα.



Από τις παραπάνω εικόνες φαίνεται ότι όσο αυξάνεται η σταθερά ε τόσο μειώνεται η εκτίμηση του διανύσματος d , μιας και η σταθερά ε , βελτιώνει το αποτέλεσμα σε επίπεδες περιοχές με μειωμένη υφή, δηλαδή με μειωμένη πληροφορία για τον υπολογισμό της οπτικής ροής. Άρα, ως ιδανική τιμή για την παράμετρο ε επιλέγεται η τιμή 0.01.

Συνεπώς, για τιμές παραμέτρων $[\rho, \varepsilon, \sigma_{\text{ep}}] = [4, 0.01, 50]$ και αρχικές εκτιμήσεις $[dx_0, dy_0] = [0, 0]$ προκύπτουν οι ακόλουθες εικόνες για τον αλγόριθμο Lucas – Kanade.



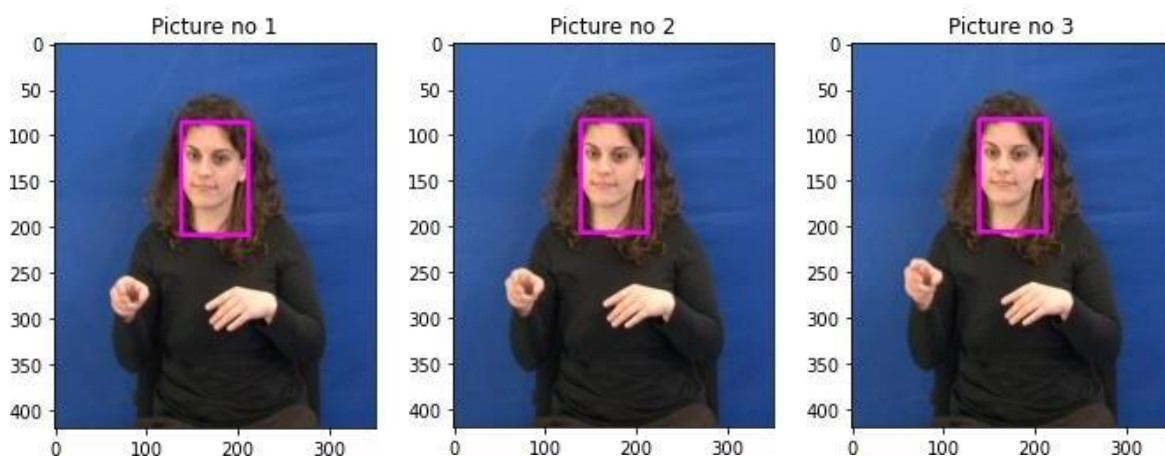
Από τις παραπάνω εικόνες φαίνεται ότι ο αλγόριθμος Lucas – Kanade μετατοπίζει με μεγάλη επιτυχία το τετράγωνο ενδιαφέροντος στο επόμενο frame, όταν η μετατόπιση

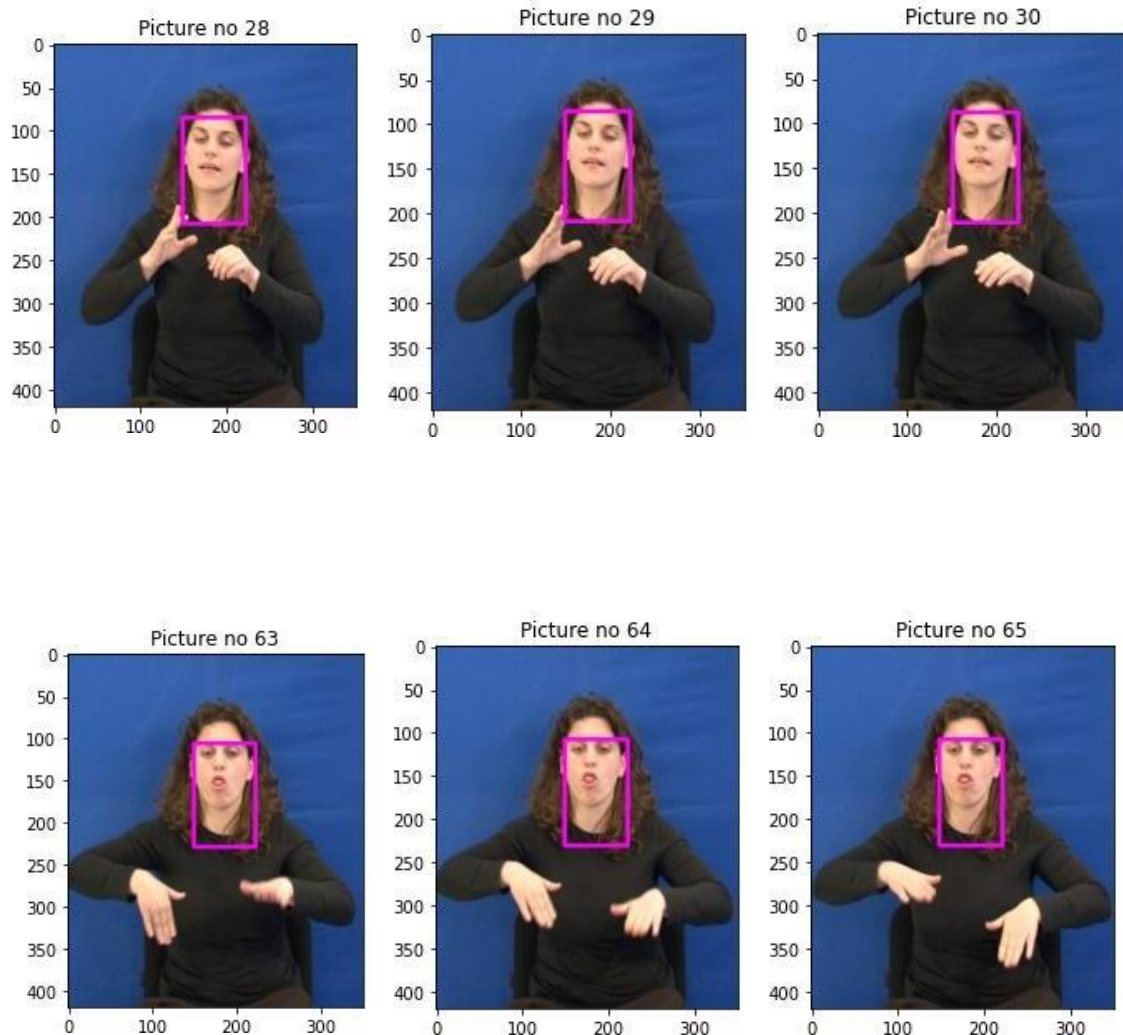
στην εικόνα είναι μικρή. Αντιθέτως, όταν το διάνυσμα της οπτικής ροής παίρνει μεγάλες τιμές ο Lucas – Kanade δυσκολεύεται να μετατοπίσει επακριβώς το τετράγωνο ενδιαφέροντος.

(1.2.3) Πολυ-Κλιμακωτός Υπολογισμός Οπτικής Ροής

Σε αυτό το ερώτημα υλοποιείται η πολύ-κλιμακωτή εκδοχή του αλγόριθμου των Lucas - Kanade . Ο αλγόριθμος αυτός αναλύει τις αρχικές εικόνες σε Γκαουσιανές πυραμίδες και υπολογίζει την οπτική ροή από τις πιο μικρές στις πιο μεγάλες κλίμακες, χρησιμοποιώντας τη λύση της μικρής κλίμακας ως αρχική συνθήκη για τη μεγάλη κλίμακα. Οι Γκαουσιανές πυραμίδες προκύπτουν από την υποδειγματοληψία της εικόνας μειώνοντας την ανάλυση κάθε φορά στο μισό.

Συνεπώς, η πολύ-κλιμακωτή εκδοχή του αλγόριθμου των Lucas -Kanade δίνει τα ακόλουθα αποτελέσματα.





Οι παραπάνω εικόνες έχουν προκύψει με κατάλληλη επιλογή των παραμέτρων $[ρ, ε, ger]$, όπως έγινε η ανάλυση και για τον Lucas – Kanade σε μια κλίμακα. Συγκεκριμένα, μετά από δοκιμές επιλέχθηκε ότι ο καλύτερος συνδυασμός είναι $[ρ, ε, ger] = [4, 0.1, 50]$, δηλαδή παρατηρείται παρόμοια συμπεριφορά για τις μετρικές $ρ$, ger ενώ παρατηρήθηκε αντίθετη συμπεριφορά για την παράμετρο $ε$.

Συγκρίνοντας τους δύο παραπάνω αλγόριθμους, προκύπτει ότι ο αλγόριθμος Lucas – Kanade σε μονή κλίμακα λειτουργεί καλύτερα σε μικρές μετατοπίσεις ενώ η πολυκλιμακωτή εκδοχή του Lucas – Kanade δίνει τη λύση σε προβλήματα, που αντιμετωπίζει ο μονο-κλιμακωτός μιας και υπολογίζεται η οπτική ροή σε μεγαλύτερες κινήσεις, που υπερβαίνουν αρκετά το ένα pixel και οι παραδοχή για το ίδιο διάνυσμα σε ένα τετράγωνο δεν ισχύει πια.

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

Σε αυτό το μέρος, ο σκοπός του είναι η εξαγωγή χωρο-χρονικών χαρακτηριστικών, για την αξιοποίηση τους στο πρόβλημα κατηγοριοποίησης βίντεο, που περιέχουν ανθρώπινες δράσεις.

2.1 Χωρο-χρονικά Σημεία Ενδιαφέροντος

Οι ανιχνευτές τοπικών χαρακτηριστικών αναζητούν χωρο-χρονικά σημεία και κλίμακες ενδιαφέροντος, τα οποία αντιστοιχούν σε περιοχές που χαρακτηρίζονται από σύνθετη κίνηση ή από απότομες μεταβολές στην εμφάνιση του video εισόδου. Αυτό επιτυγχάνεται μεγιστοποιώντας μια συνάρτηση οπτικής σημαντικότητας.

(2.1.1)

Σε αυτό το βήμα υλοποιείται μια επέκταση σε 3 διαστάσεις του ανιχνευτή γωνιών Harris - Stephens. Για κάθε voxel του βίντεο εφαρμόζεται το ακόλουθο κριτήριο γωνιότητας :

$$H(x, y, t) = \det(M(x, y, t)) - k \text{trace}^3(M(x, y, t)) \quad (2.1)$$

Όπου $M(x, y, t)$: πίνακας 3×3 που περιγράφεται από τη σχέση (2.2)

$$M(x, y, t; \sigma, \tau) = g(x, y, t; \sigma, \tau) * (\nabla L(x, y, t; \sigma, \tau) (\nabla L(x, y, t; \sigma, \tau)^T)) \quad (2.2)$$

Όπου τα δομικά στοιχεία περιγράφονται από τις ακόλουθες σχέσεις

$$L(x, y, t; \sigma, \tau) = G(x, y, t; \sigma, \tau) * I(x, y, t) \quad (2.3)$$

$$G(x, y, t; \sigma, \tau) = \frac{1}{\sqrt{(2\pi)^3 \sigma^4 c^2}} e^{\left(-\frac{x^2+y^2}{2\sigma^2} - \frac{t^2}{2\tau^2}\right)} \quad (2.4)$$

Άρα, ο πίνακας M 3×3 ισοδύναμα περιγράφεται από τη σχέση την ακόλουθη σχέση.

$$M(x, y, t; \sigma, \tau) = \begin{bmatrix} J_{11} & J_{12} & J_{13} \\ J_{21} & J_{22} & J_{23} \\ J_{31} & J_{32} & J_{33} \end{bmatrix} \quad (2.5)$$

Όπου τα στοιχεία του πίνακα M περιγράφονται από τις σχέσεις στη (2.6) .

$$\begin{aligned} J_{11}(x, y, t) &= g(x, y, t; \sigma, \tau) * \left(\frac{\partial L}{\partial x} \frac{\partial L}{\partial x} \right) \\ J_{12}(x, y, t) &= g(x, y, t; \sigma, \tau) * \left(\frac{\partial L}{\partial x} \frac{\partial L}{\partial y} \right) \\ J_{13}(x, y, t) &= g(x, y, t; \sigma, \tau) * \left(\frac{\partial L}{\partial x} \frac{\partial L}{\partial t} \right) \\ J_{21}(x, y, t) &= J_{12}(x, y, t) \\ J_{22}(x, y, t) &= g(x, y, t; \sigma, \tau) * \left(\frac{\partial L}{\partial y} \frac{\partial L}{\partial y} \right) \end{aligned} \quad (2.6)$$

$$J_{23}(x,y,t) = g(x, y, t; s\sigma, s\tau) * \left(\frac{\partial L}{\partial y} \frac{\partial L}{\partial t} \right)$$

$$J_{31}(x,y,t) = J_{13}(x,y,t)$$

$$J_{32}(x,y,t) = J_{23}(x,y,t)$$

$$J_{33}(x,y,t) = g(x, y, t; s\sigma, s\tau) * \left(\frac{\partial L}{\partial t} \frac{\partial L}{\partial t} \right)$$

Μοντελοποιώντας προγραμματιστικά το 3Δ κριτήριο γωνιότητας προκύπτουν τα ακόλουθα αποτελέσματα για 20 πρώτα frames του βίντεο :

“person01_boxing_d2_uncomp.avi”

Τα αποτελέσματα φαίνονται στο ερώτημα (2.1.3) μαζί με την υλοποίηση του άλλου ανισχευτή.

(2.1.2)

Σε αυτό το βήμα υλοποιείται ένας ανιχνευτή Gabor ο οποίος βασίζεται στο χρονικό φιλτράρισμα του βίντεο με ένα ζεύγος Gabor φίλτρων, όπως περιγράφονται από τις ακόλουθες σχέσεις.

$$h_{\text{even}}(t; \tau, \omega) = \cos(2\pi t\omega) e^{-t^2 / 2\tau^2} \quad (2.7)$$

$$h_{\text{odd}}(t; \tau, \omega) = \sin(2\pi t\omega) e^{-t^2 / 2\tau^2}$$

Όπου η συχνότητα ω του Gabor φίλτρου συνδέεται με την χρονική κλίμακα τ μέσω της σχέσης: $\omega = 4/\tau$. Τελικά, το κριτήριο σημαντικότητας προκύπτει παίρνοντας την τετραγωνική ενέργεια της εξόδου για το ζεύγος Gabor φίλτρων, όπως περιγράφεται από

την ακόλουθη σχέση.

$$H(x, y, t) = (I(x, y, t) * g * h_{\text{even}})^2 + (I(x, y, t) * g * h_{\text{odd}})^2 \quad (2.8)$$

Μοντελοποιώντας προγραμματιστικά το 3Δ κριτήριο γωνιότητας προκύπτουν τα ακόλουθα αποτελέσματα για 20 πρώτα frames του βίντεο :

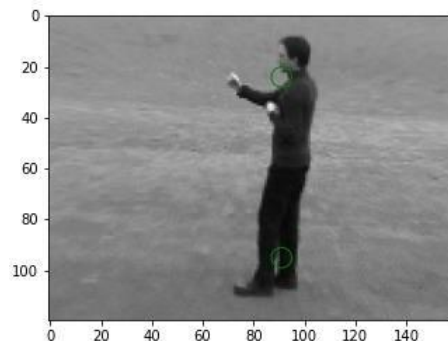
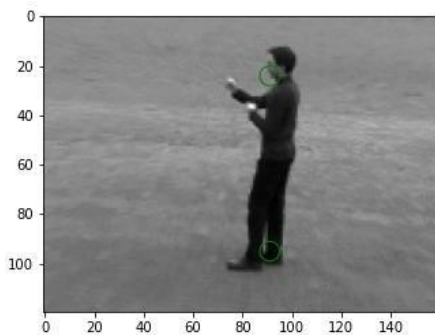
“person01_boxing_d2_uncomp.avi”

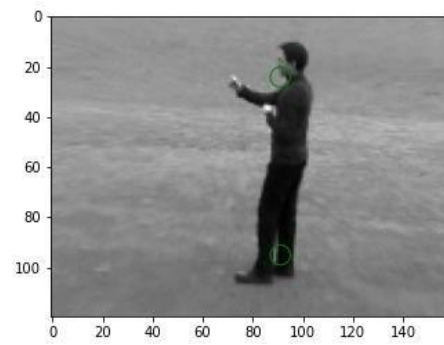
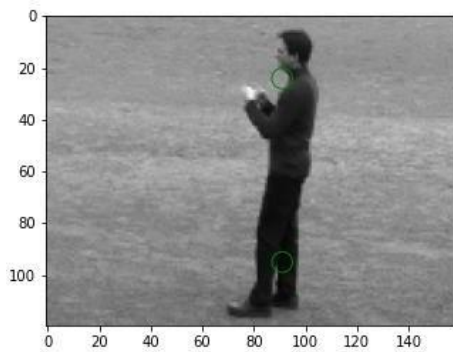
Τα αποτελέσματα φαίνονται στο ερώτημα (2.1.3) μαζί με την υλοποίηση του άλλου ανιχνευτή.

(2.1.3)

Σε αυτό το βήμα απεικονίζονται για επιλεγμένα frames τα σημεία που προκύπτουν χρησιμοποιώντας τη δοθείσα συνάρτηση show detection, δίνοντας τα ακόλουθα αποτελέσματα.

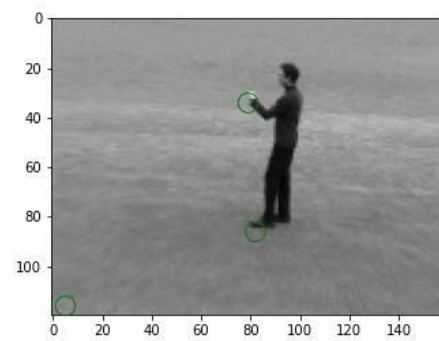
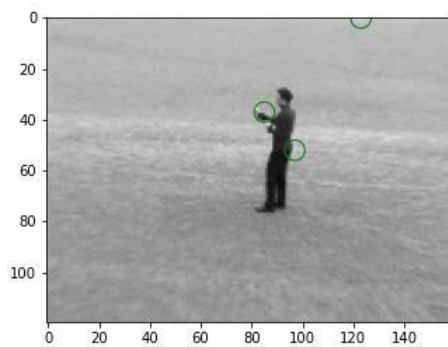
Για τον Harris – Stephens :

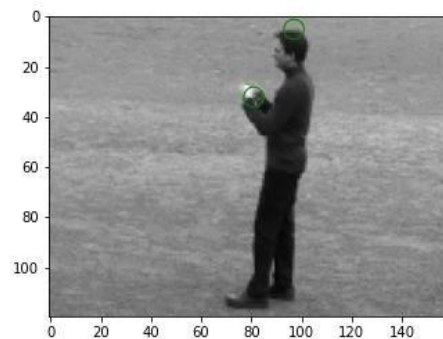
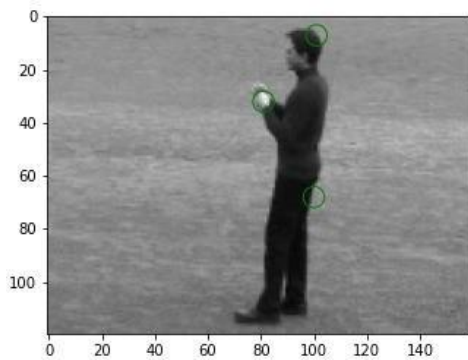




Από τις παραπάνω εικόνες φαίνεται ότι υπήρξε κάποιο πρόβλημα στο κριτήριο γωνιότητας που εφαρμόστηκε μιας και δεν εντοπίζει τις μεταβολές στις κινήσεις του ανθρώπου.

Για τον Gabor :





Από τις παραπάνω εικόνες φαίνεται ότι ο ανιχνευτής Gabor εντόπισε με επιτυχία τις γωνίες που σχηματίζονται κατά την κίνηση του ανθρώπου. Μάλιστα εντόπισε και κάποια άλλα σημεία τα οποία πιθανόν να οφείλονται στην κίνηση του γρασιδιού.

2.2 Χωρο-χρονικοί Ιστογραφικοί Περιγραφητές

Σε αυτό το μέρος του εργαστηρίου θα χρησιμοποιηθούν χωρο-χρονικοί περιγραφητές, οι οποίοι βασίζονται στον υπολογισμό ιστογραμμάτων της κατευθυντικής παραγώγου (HOG) και της οπτικής ροής (HOF - Histogram of Oriented Flow) γύρω από τα σημεία ενδιαφέροντος, που βρέθηκαν από τους ανιχνευτές του μέρους (2.1).

(2.2.1)

Σε αυτό το βήμα υπολογίζεται το διάνυσμα κλίσης για κάθε frame του βίντεο σύμφωνα με συναρτήσεις που υλοποιήθηκαν στο Μέρος 1 του εργαστηρίου.

(2.2.2)

(2.2.3)

2.3 Κατασκευή Bag of Visual Words και χρήση Support Vector Machine για την ταξινόμηση δράσεων

(2.3.1)

(2.3.2)

(2.3.3)

(2.3.4)

(2.3.5)

Βιβλιογραφία-Αναφορές:

[1] [splitting-and-merging-channels-with-opencv](#)