

Όραση Υπολογιστών  
Εργαστήριο 2

Αναστάσιος Στέφανος Αναγνώστου  
03119051

Σπυρίδων Παπαδόπουλος  
03119058

18 Ιουνίου 2023

# Περιεχόμενα 0

<b>1 Μέρος</b>	<b>3</b>
1.1 Ανίχνευση Δέρματος Προσώπου και Χεριών . . . . .	3
1.2 Παρακολούθηδη Προσώπου και Χεριών . . . . .	5
1.2.1 Υλοποίηση Αλγορίθμου Lucas-Kanade . . . . .	5
1.2.2 Υπολογισμός Μετατόπισης Παραθύρων και Οπτικής Ροής . . . . .	7
1.2.3 Πολυ-κλιμακωτός Υπολογισμός Οπτικής Ροής . . . . .	8
<b>2 Μέρος</b>	<b>10</b>
2.1 Χωροχρονικά Σημεία Ενδιαφέροντος . . . . .	10
2.1.1 Ανιχνευτής Harris . . . . .	10
2.1.2 Ανιχνευτής Gabor . . . . .	12
2.1.3 Πολυκλιμακωτές Εκδοχές . . . . .	16
2.2 Χωροχρονικοί Ιστογραφικοί Περιγραφητές . . . . .	19
2.3 Κατασκευή Bag of Visual Words και χρήση Support Vector Machines για ταξινόμηση δράσεων	19
<b>3 Μέρος</b>	<b>21</b>

Ο κάθισκας παρατίθεται σε ξεχωριστά αρχεία με επαρκή σχολιασμό και όχι στην παρούσα αναφορά.

## Μέρος 1

Στο πρώτο μέρος του εργαστηρίου υλοποιείται σύστημα παρακολούθησης προσώπου και χεριών σε βίντεο νοητοποίησης γλώσσας.

### 1.1 Ανίχνευση Δέρματος Προσώπου και Χεριών

Η ανίχνευση δέρματος προσώπου και χεριών επιτυγχάνεται χρήσει ενός πιθανοτικού ανιχνευτή ανθρωπίνου δέρματος. Συγκεκριμένα, θεωρείται ότι το χρώμα του δέρματος μοντελοποιείται με μία διδιάστατη Γκαουσιανή κατανομή.

$$P(\mathbf{c} = \text{skin}) = \frac{1}{\sqrt{|\Sigma|} (2\pi)^2} e^{-\frac{1}{2}(\mathbf{c}-\boldsymbol{\mu})\Sigma^{-1}(\mathbf{c}-\boldsymbol{\mu})^T} \quad (1.1)$$

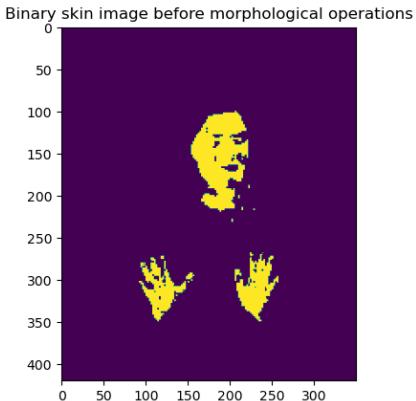
Το διάνυσμα τιμών  $\vec{c}$  της κατανομής είναι οι διαστάσεις  $C_b, C_r$  του χρωματικού χώρου YCbCr, δηλαδή:

$$\mathbf{c} = [C_b \quad C_r]^T \quad (1.2)$$

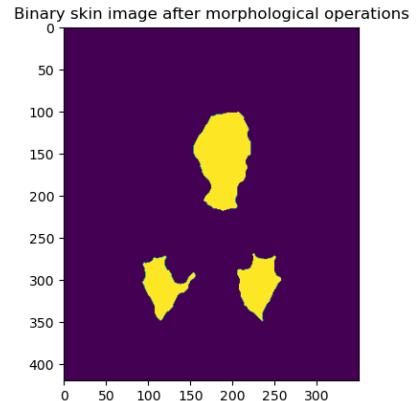
Η συνάρτηση πυκνότητας πιθανότητας εκπαιδεύεται χρήσει των παρεχόμενων δεδομένων skinSamplesRGB.mat. Φυσικά, προτού χρησιμοποιηθούν, πρέπει να μετασχηματιστούν από τον χώρο RGB στον χώρο YCbCr. Αφού εκπαιδευτεί η συνάρτηση πυκνότητας πιθανότητας, η δυαδική εικόνα ανίχνευσης δέρματος προκύπτει, με κατάλληλη κατωφλιοποίηση, ως:

$$I(x, y) = \begin{cases} 1, \arg \left\{ P(\mathbf{c}(x, y) = \text{skin}) \geq \theta \right\} \\ 0, \text{ειδικάλλως} \end{cases} \quad (1.3)$$

Σε αυτό το σημείο, η δυαδική εικόνα έχει μερικά κενά, τα οποία είναι ανεπιθύμητα, αφού είναι γνωστό ότι οι περιοχές δέρματος είναι συνεχείς. Αυτό επαληθύνεται από την ακόλουθη εικόνα 1α'.



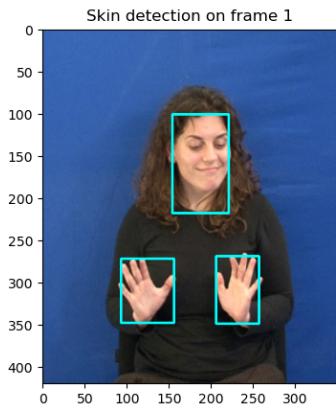
(α') Η ατελής δυαδική εικόνα δέρματος



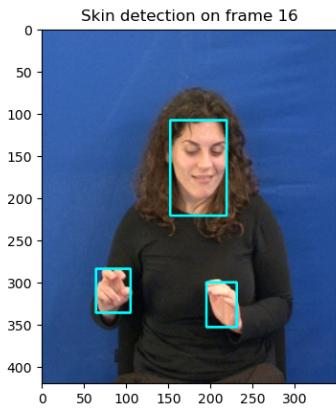
(β') Η βελτιωμένη δυαδική εικόνα δέρματος

Για την εξάλειψη των κενών, εφαρμόζεται opening με ένα πολύ μικρό δομικό στοιχείο και closing με ένα μεγαλύτερο ( $\times 6$ ) δομικό στοιχείο. Το αποτέλεσμα παρατίθεται στην εικόνα 1β'.

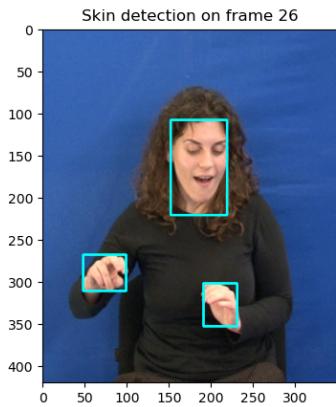
Παρατηρείται ότι οι περιοχές είναι τώρα συνεχείς και η απεικόνιση σαφώς βελτιωμένη, με τίμημα την ακρίβεια σε οριακές περιοχές όπως είναι τα δάχτυλα.



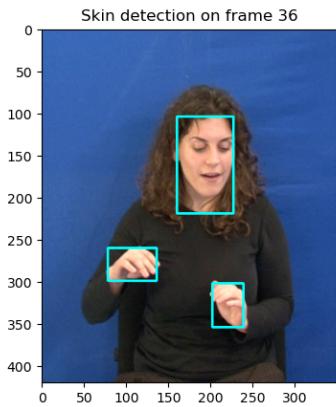
(α) Πλαίσιο 1



(β') Πλαίσιο 16



(γ') Πλαίσιο 26



(δ') Πλαίσιο 36

Σχήμα 2: Ενδεικτικός εντοπισμός δέρματος σε μερικά πλαίσια

Τέλος, όταν δημιουργηθούν ορθογώνια τα οποία περιβάλλουν τις περιοχές ενδιαφέροντος. Παρουσιάζονται ενδεικτικά μερικά πλαίσια από το βίντεο στην εικόνα 2.

Φαίνεται, ότι τα αποτελέσματα είναι αρκετά ικανοποιητικά. Συγκεκριμένα, οι περιοχές εντοπίζονται με αρκετή ακρίβεια σε όλες τις περιπτώσεις και τα χέρια παρακολουθούνται επιτυχώς, παρότι κινούνται και αλλάζουν το σχήματος. Μάλιστα, όπως αλλάζει το φαινόμενο μέγεθος των χεριών, αλλάζει μαζί και το περιβάλλον κουτί.

## 1.2 Παρακολούθηση Προσώπου και Χεριών

Σε αυτό το σημείο όταν επιχειρηθεί η παρακολούθηση του προσώπου και των χεριών, όχι όμως μέσω ανίχνευσης δέρματος, αλλά μέσω εκτίμησης οπτικής ροής. Αντί, δηλαδή, να βρίσκεται σε καθένα πλαίσιο εκ νέου η περιοχή γύρω από το δέρμα, γίνεται μία εκτίμηση της οπτικής ροής από πλαίσιο σε πλαίσιο.

### 1.2.1 Υλοποίηση Αλγορίθμου Lucas-Kanade

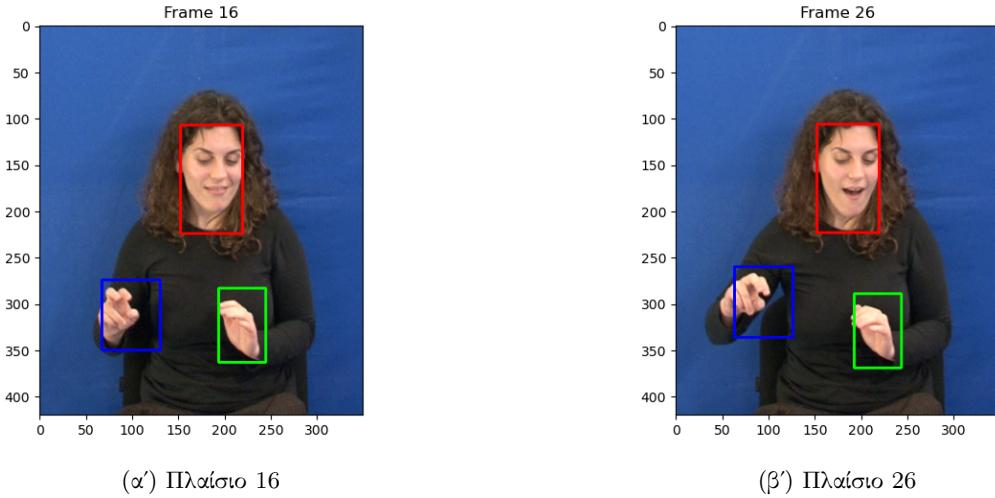
Η εκτίμηση της οπτικής ροής, εν προκειμένω, όταν γίνει χρήσει του αλγορίθμου Lucas-Kanade. Κατά τον αλγόριθμο αυτόν, υπολογίζεται η οπτική ροή καθενός σημείου της εικόνας, χρήσει της μεθόδου ελαχίστων τετραγώνων, υπό την παραδοχή ότι αυτή παραμένει σταθερή σε μία μικρή περιοχή γύρω από το σημείο.

$$\begin{aligned} I_n(\mathbf{x}) &\approx I_{n-1}(\mathbf{x} + \mathbf{d}) \xrightarrow{\text{Taylor}} \\ I_{n-1}(\mathbf{x} + \mathbf{d}) &\approx I_{n-1}(\mathbf{x} + \mathbf{d}_i) + \nabla I_{n-1}(\mathbf{x} + \mathbf{d}_i)^T \mathbf{u} \end{aligned} \quad (1.4)$$

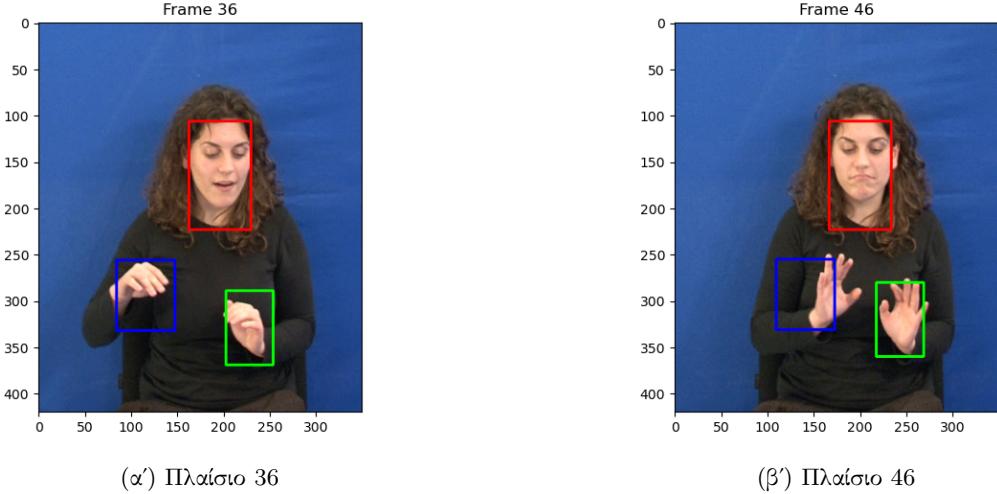
Ελαχιστοποιώντας το τετραγωνικό σφάλμα, προκύπτει ένα σύστημα το οποίο προσδιορίζει πλήρως την βελτίωση της εκτίμησης της οπτικής ροής.

Ο αλγόριθμος δεν εφαρμόζεται σε κάθε σημείο της εικόνας, αλλά αποδεικνύεται (απλώς κοιτάζοντας το προκύπτον σύστημα) ότι δίνει την βέλτιστη εκτίμηση όταν εφαρμόζεται σε γωνίες. Επομένως, πρώτα γίνεται εντοπισμός γωνιών στην εικόνα και έπειτα εφαρμόζεται ο αλγόριθμος. Η εφαρμογή γίνεται επαναληπτικά ανά σημείο, μέχρις ότου, είτε δύο διαδοχικές βελτιώσεις να διαφέρουν κατά αμελητέο ποσό, είτε να ξεπεραστεί κάποιο δριό επαναλήψεων.

Παρατίθενται ενδεικτικά πλαίσια για την αξιολόγηση της μεθόδου στην εικόνα 3. Σημειώνεται, επίσης, ότι στα αρχεία της αναφοράς δίνεται το αποτέλεσμα της συνένωσης όλων των πλαισίων σε μορφή gif.



Σχήμα 3: Ενδεικτικός εντοπισμός δέρματος χρήσει εκτίμησης οπτικής ροής

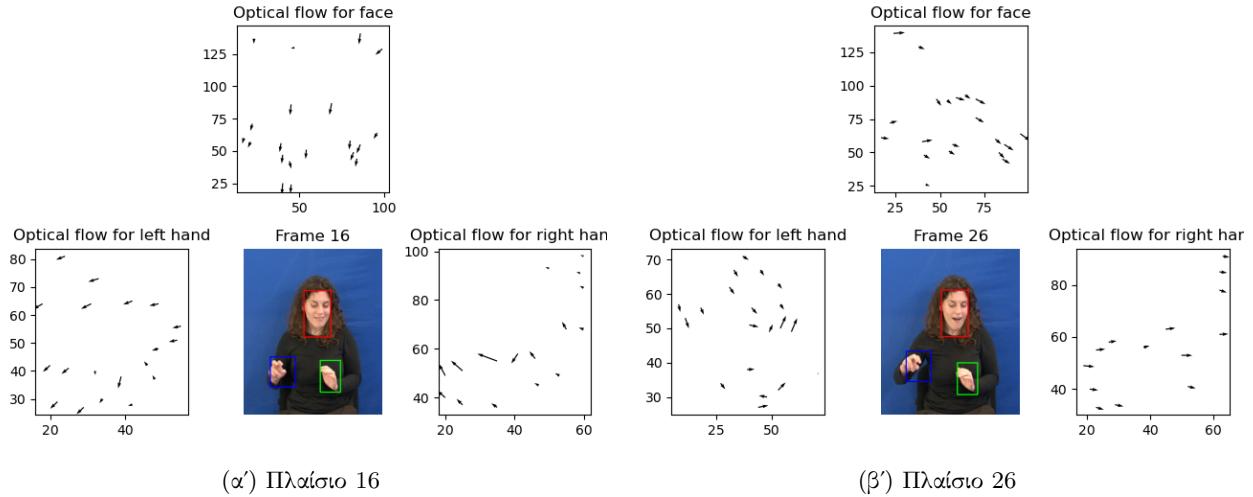


Σχήμα 4: Ενδεικτικός εντοπισμός δέρματος χρήσει εκτίμησης οπτικής ροής

Επισκοπώντας τα αποτελέσματα, συμπεραίνεται, ότι η εκτίμηση της οπτικής ροής λειτουργεί αρκετά καλά για μικρές, αργές κινήσεις. Αυτό επαληθεύεται από τα 3 πρώτα παρατιθέμενα πλάσια, στα οποία οι κινήσεις ήταν αργές και ομαλές. Ωστόσο, όταν οι κινήσεις γίνονται απότομες, η αποδοτικότητα του αλγορίθμου μειώνεται, με αποτέλεσμα να υστερούν τα κουτιά από τις περιοχές ενδιαφέροντος, όπως φαίνεται στο τελευταίο πλάσιο. Σημειώνεται, επίσης, ότι η κλίμακα του κουτιού δεν αλλάζει, όπως προηγουμένως, αλλά παραμένει σταθερή.

### 1.2.2 Υπολογισμός Μετατόπισης Παραθύρων και Οπτικής Ροής

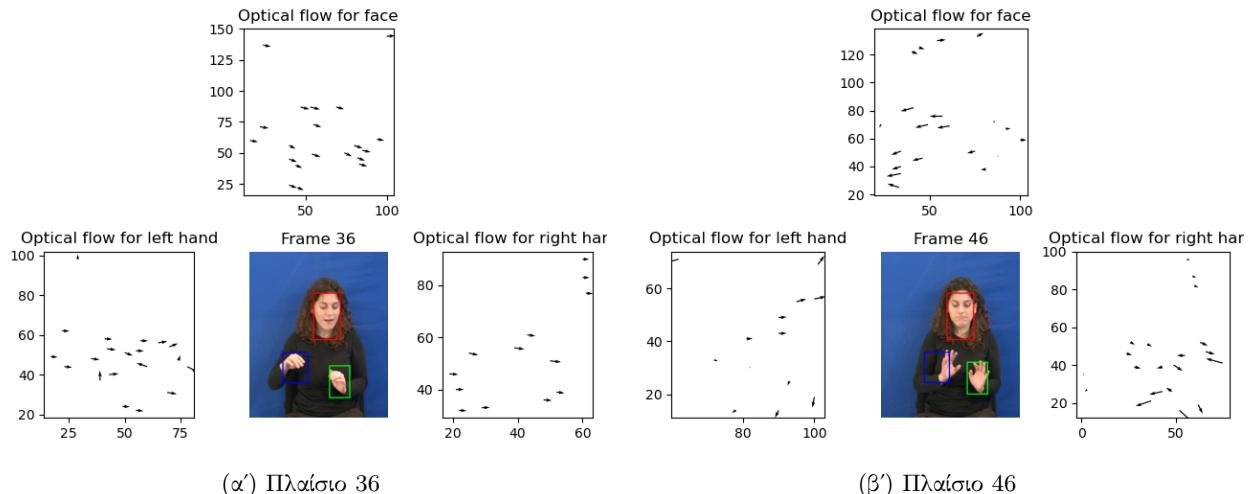
Το συνολικό διάνυσμα μετατόπισης του κουτιού βρίσκεται, ουσιαστικά, υπολογίζοντας έναν απλό μέσον όρο. Συγκεκριμένα, υπολογίζεται ο μέσος όρος των διανυσμάτων μετατόπισης αυτών, τα οποία έχουν ενέργεια  $\|\mathbf{d}^2\| = d_x^2 + dy^2$  μεγαλύτερη από ένα κατώφλι, ώστε να απορριφθούν outliers και να επιτευχθεί καλύτερη ακρίβεια.



(α') Πλαίσιο 16

(β') Πλαίσιο 26

Σχήμα 5: Εντοπισμός δέρματος και οπτική ροή σε μερικά πλαίσια



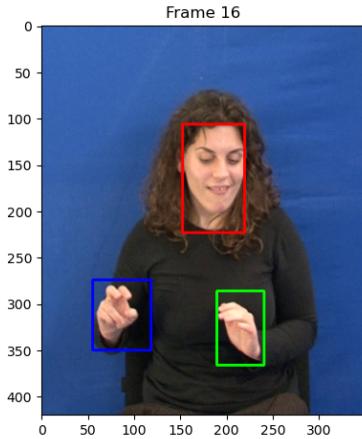
(α') Πλαίσιο 36

(β') Πλαίσιο 46

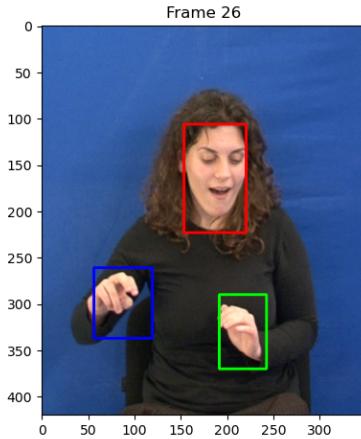
Σχήμα 6: Εντοπισμός δέρματος και οπτική ροή σε μερικά πλαίσια

### 1.2.3 Πολυ-κλιμακωτός Υπολογισμός Οπτικής Ροής

Η κακή επίδοση του αλγορίθμου στις απότομες δημιουργεί την ανάγκη επέκτασής του σε περισσότερες κλίμακες. Συγκεκριμένα, η πολυκλιμακωτή εκδοχή του, αναλύει την αρχική εικόνα σε μία Γκαουσιανή πυραμίδα εικόνων, στο καθένα επίπεδο της οποίας βρίσκεται μία υποδειγματοληπτημένη εκδοχή της αυθεντικής εικόνας. Σε καθεμία κλίμακα χρησιμοποιείται ο αυθεντικός αλγόριθμος και μεταφέρει την εκτίμησή του στην αμέσως ανώτερη κλίμακα, προσαρμόζοντάς την κατάλληλα. Σημειώνεται, ότι τα χαρακτηριστικά πρέπει τώρα να εξαχθούν για κάθε επίπεδο της κλίμακας, όχι μόνο για την αρχική εικόνα.

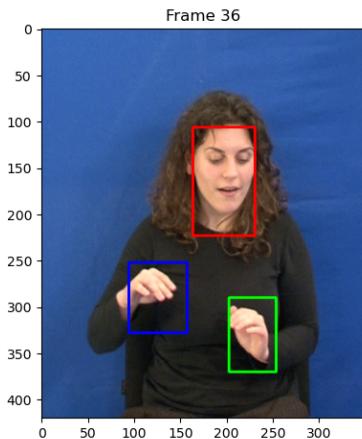


(α') Πλαίσιο 16

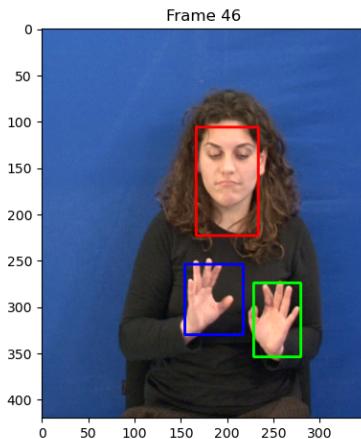


(β') Πλαίσιο 26

Σχήμα 7: Ενδεικτικός πολυκλιμακωτός εντοπισμός δέρματος χρήσει εκτίμησης οπτικής ροής



(α') Πλαίσιο 36



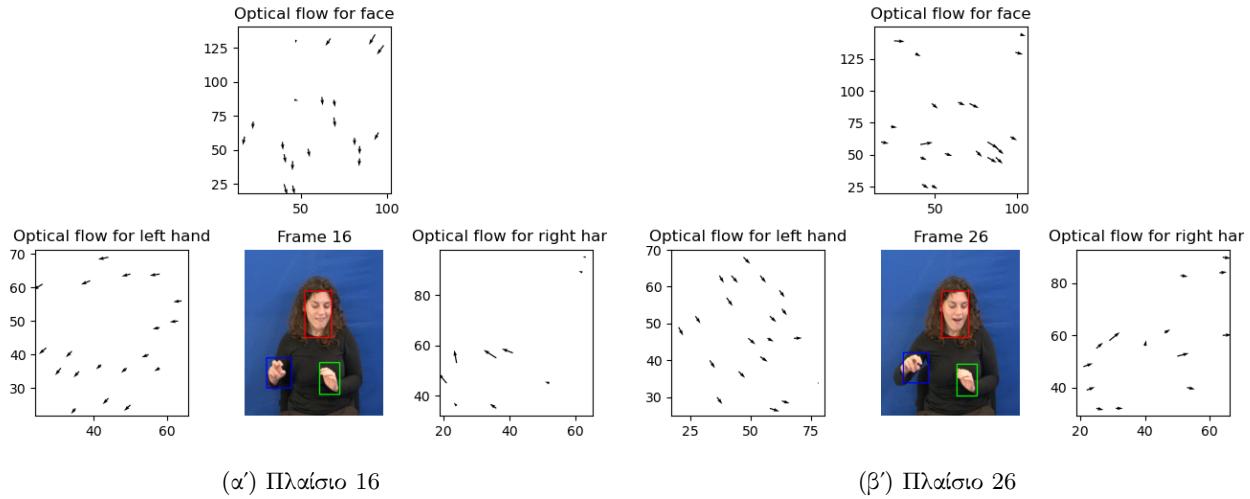
(β') Πλαίσιο 46

Σχήμα 8: Ενδεικτικός πολυκλιμακωτός εντοπισμός δέρματος χρήσει εκτίμησης οπτικής ροής

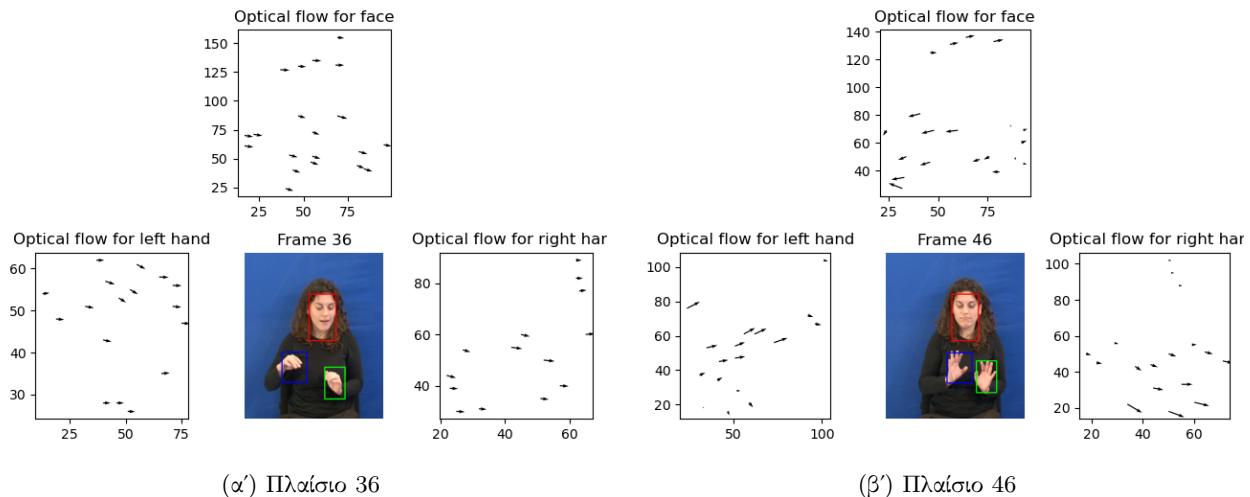
Παρατηρείται, ότι σε κάθε πλαίσιο ο εντοπισμός είναι ακριβέστερος. Η διαφορά είναι ιδιαίτερα αξιοσημειώτη στο τελευταίο παρατιθέμενο πλαίσιο, όπου το μπλε κουτί περικυκλώνει τέλεια το αριστερό χέρι, ενώ στην μονοκλιμακωτή εκδοχή το έχανε λόγω ταχύτητας.

Παρακάτω δίνονται τα αντίστοιχα αποτελέσματα για την απεικόνιση της οπτικής ροής. Η βελτίωση της ποιότητας είναι ορατή και εκεί, αφού τα διανύσματα οπτικής ροής είναι πλέον πολύ ομοιόμορφα και όλα σαφώς

προσανατολισμένα προς κάποια κατεύθυνση, σε αντίθεση με την προηγούμενη μέθοδο όπου υπήρχαν κάποιες διακυμάνσεις.



Σχήμα 9: Πολυκλιμακωτός εντοπισμός δέρματος και οπτική ροή σε μερικά πλαίσια



Σχήμα 10: Πολυκλιμακωτός εντοπισμός δέρματος και οπτική ροή σε μερικά πλαίσια

## Μέρος 2

Στόχος αυτού του μέρους είναι η εξαγωγή χωροχρονικών χαρακτηριστικών με στόχο την εφαρμογή τους στο πρόβλημα κατηγοριοποίησης βίντεο ανθρωπίνων δράσεων.

### 2.1 Χωροχρονικά Σημεία Ενδιαφέροντος

#### 2.1.1 Ανιχνευτής Harris

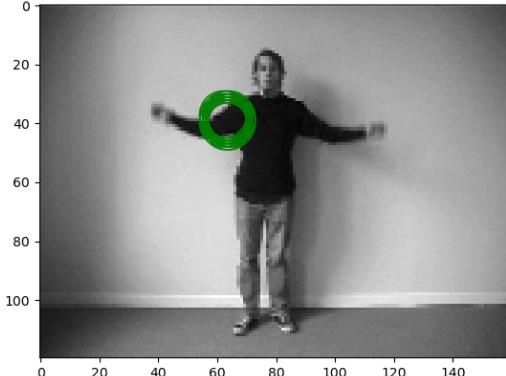
Αρχικά, υλοποιείται ο ανιχνευτής Harris, καταλλήλως προσαρμοσμένος για βίντεο. Συγκεκριμένα, υπολογίζεται ο τανυστής:

$$M(x, y, t|\sigma, \tau) = g(x, y, t|\sigma, \tau) * \begin{bmatrix} L_x^2 & L_x L_y & L_x L_t \\ L_x L_y & L_y^2 & L_y L_t \\ L_x L_t & L_y L_t & L_t^2 \end{bmatrix} \quad (2.1)$$

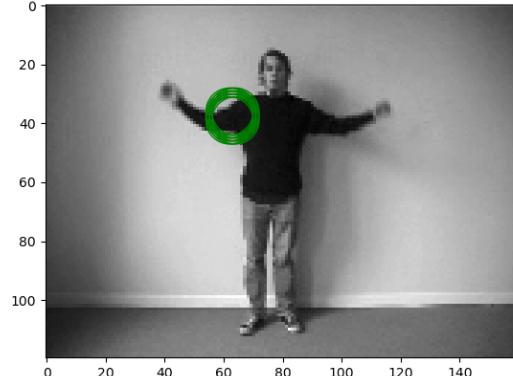
Οι μερικές παράγωγοι της εικόνας υπολογίζονται στην εξομαλυμένη εκδοχή της και κατόπιν εξομαλύνονται εκ νέου από τον Γκαουσσιανό πυρήνα. Το κριτήριο γωνιότητας είναι:

$$H(x, y, t) = \det(M(x, y, t)) - k \cdot \text{trace}^3(M(x, y, t)) \quad (2.2)$$

Παρατίθενται τα αποτελέσματα της ανίχνευσης. Σημειώνεται, ότι τόσο στον παρόν ανιχνευτή όσο και στους άλλους υλοποιηθέντες, απεικονίζονται τα σημεία ενδιαφέροντος ως τοπικά μέγιστα του κριτηρίου σημαντικότητας. Συγκεκριμένα, απεικονίζονται τα 500 μέγιστα.



(α') Πλαίσιο 8



(β') Πλαίσιο 9



( $\alpha'$ ) Πλαίσιο 8



( $\beta'$ ) Πλαίσιο 9



( $\gamma'$ ) Πλαίσιο 69



( $\delta'$ ) Πλαίσιο 70

Σχήμα 12: Εντοπισμός σημείων ενδιαφέροντος από τον ανιχνευτή harris

Ο ανιχνευτής Harris φαίνεται να μην ανιχνεύει μεγάλη ποικιλία σημείων. Συγκεκριμένα, φαίνεται να ανιχνεύει κυρίως γωνίες, όπως φαίνεται στο τρέξιμο και στο περπάτημα όπου έχουν ανιχνευτεί οι φτέρνες και στον χαιρετισμό όπου έχει ανιχνευτεί ο βραχίωνας του ενός χεριού.

### 2.1.2 Ανιχνευτής Gabor

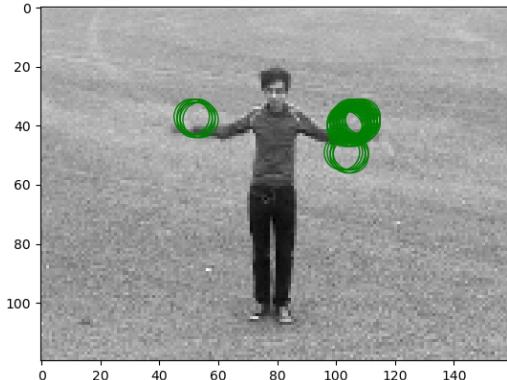
Ο ανιχνευτής Gabor βασίζεται στο χρονικό φιλτράρισμα της εξομαλυμένης από Γκαουσσιανό πυρήνα εικόνας με τα παρακάτω φίλτρα:

$$\begin{aligned} h_e v(t|\tau, \omega) &= \cos(2\pi t\omega) e^{-\frac{t^2}{r\tau^2}} \\ h_o d(t|\tau, \omega) &= \sin(2\pi t\omega) e^{-\frac{t^2}{r\tau^2}} \end{aligned} \quad (2.3)$$

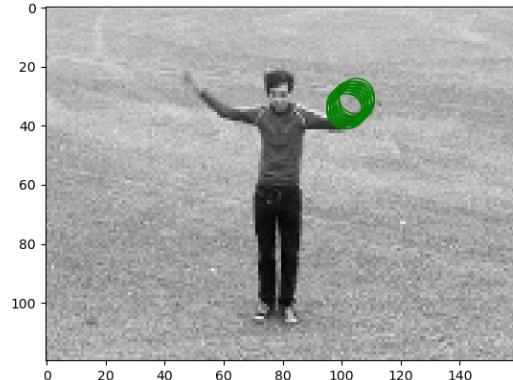
Το κριτήριο σημαντικότητας είναι:

$$H(x, y, t) = (I(x, y, t) * g * h_e v)^2 + (I(x, y, t) * g * h_o d)^2 \quad (2.4)$$

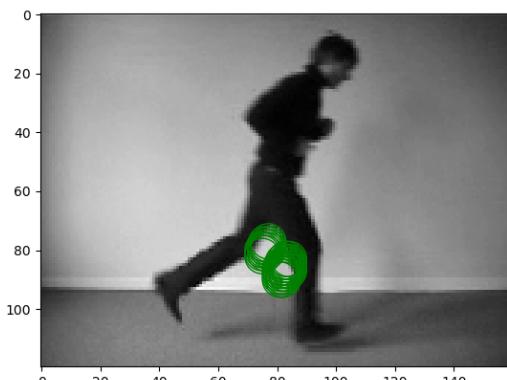
Παρατίθενται ενδεικτικά σημεία ενδιαφέροντος για καθεμία από τις δράσεις χαιρετισμού, τρεξίματος και περπατήματος.



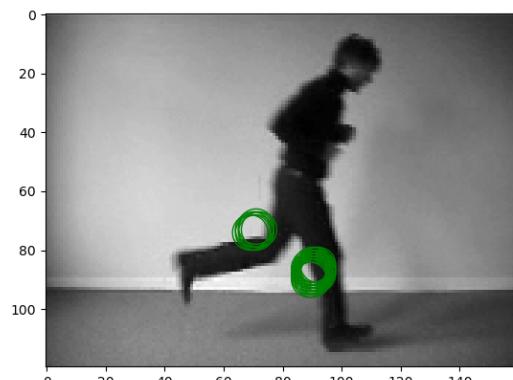
(α') Πλαίσιο 7



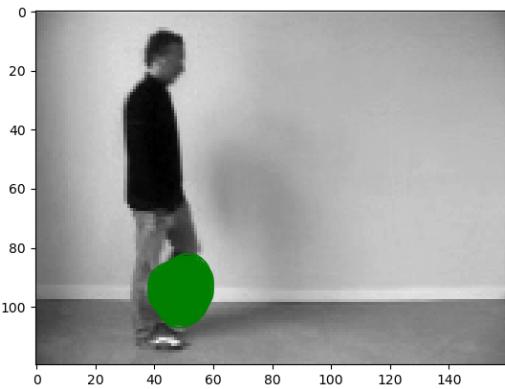
(β') Πλαίσιο 30



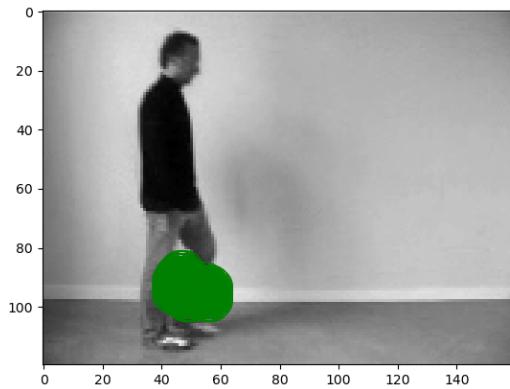
(γ') Πλαίσιο 29



(δ') Πλαίσιο 30



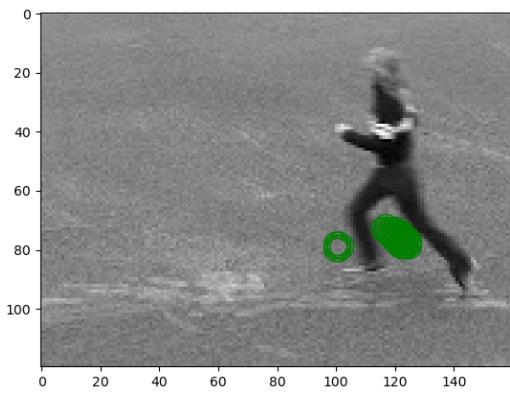
( $\alpha'$ ) Πλάσιο 16



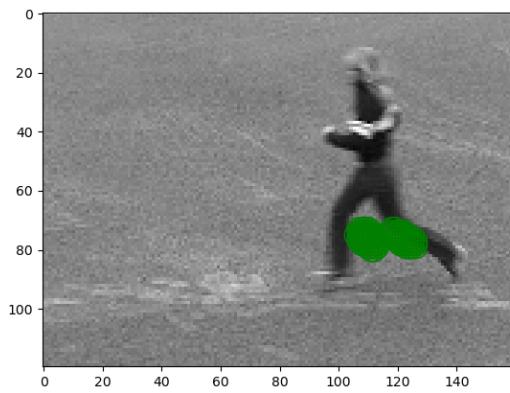
( $\beta'$ ) Πλάσιο 17

Σχήμα 14: Εντοπισμός σημείων ενδιαφέροντος από τον ανιχνευτή Gabor  
 $\sigma = 3, \tau = 1.5$

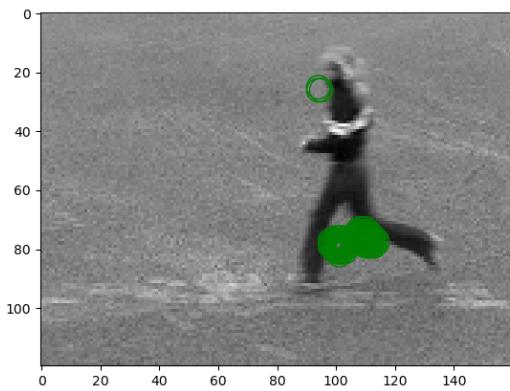
Δοκιμάζοντας διάφορες τιμές των παραμέτρων, παρατηρήθηκε ότι η αύξηση της παραμέτρου το οδηγούσε στην ανίχνευση πολλών διαδοχικών χρονικά σημείων, ενώ η μικρή τιμή της παραμέτρου οδηγούσε σε πιο ‘σύντομες’ ανίχνεύσεις. Αντιστοίχως, μεγάλες τιμές της παραμέτρου σ οδηγούσαν στην ανίχνευση μεγαλύτερων χωρικά σημείων, ενώ μικρές τιμές στην ανίχνευση μικρών χωρικά σημείων. Παρουσιάζονται ενδεικτικά οι παρατηρήσεις στις περιπτώσεις του τρεξίματος και του χαιρετισμού.



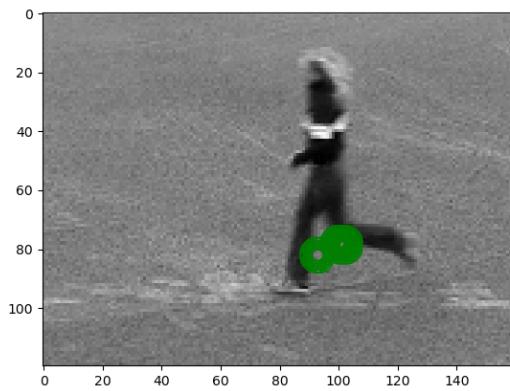
( $\alpha'$ ) Πλαίσιο 13



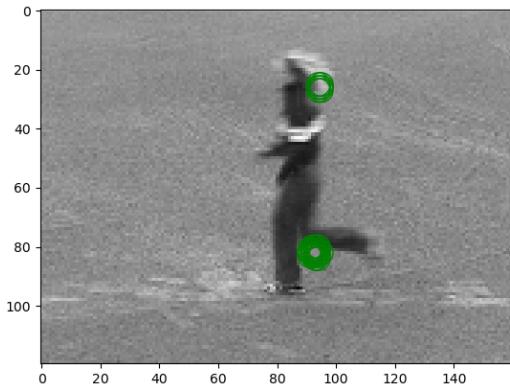
( $\beta'$ ) Πλαίσιο 14



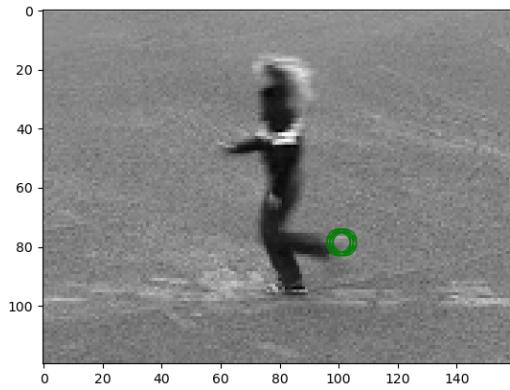
( $\gamma'$ ) Πλαίσιο 15



( $\delta'$ ) Πλαίσιο 16

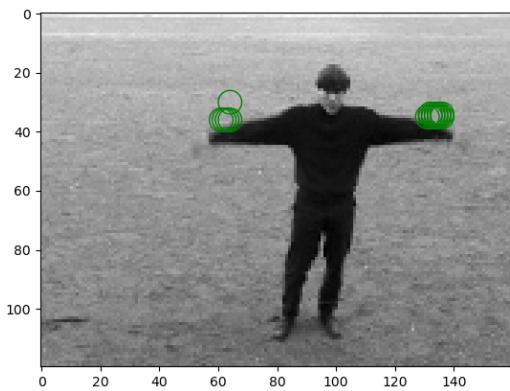


( $\alpha'$ ) Πλαίσιο 17

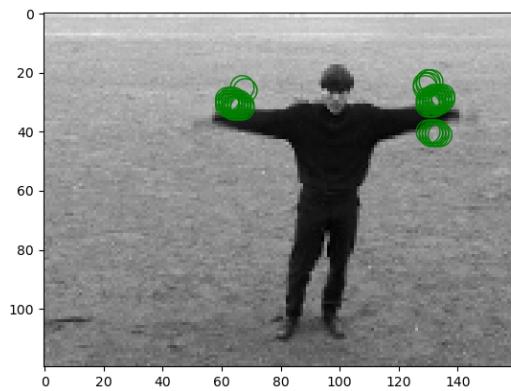


( $\beta'$ ) Πλαίσιο 18

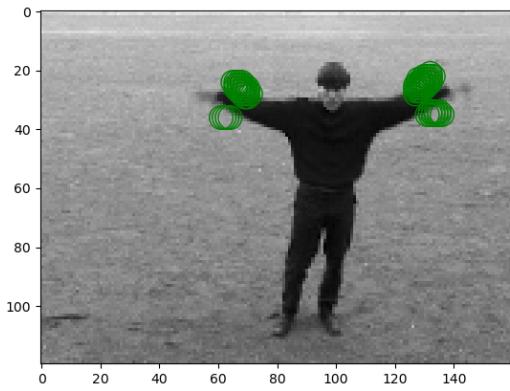
**Σχήμα 16:** Εντοπισμός σημείων ενδιαφέροντος από τον ανιχνευτή Gabor  
 $\sigma = 2, \tau = 3$



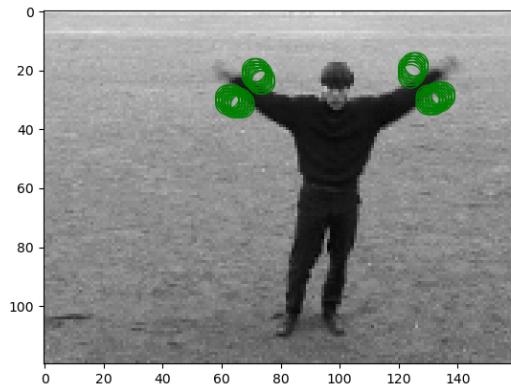
(α') Πλαίσιο 11



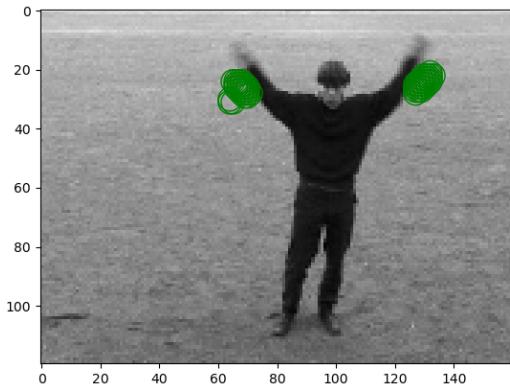
(β') Πλαίσιο 12



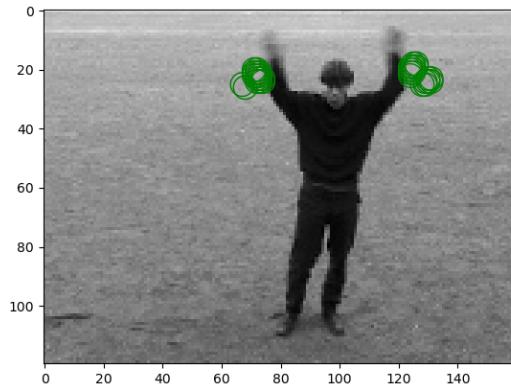
(γ') Πλαίσιο 13



(δ') Πλαίσιο 14



(ε') Πλαίσιο 15



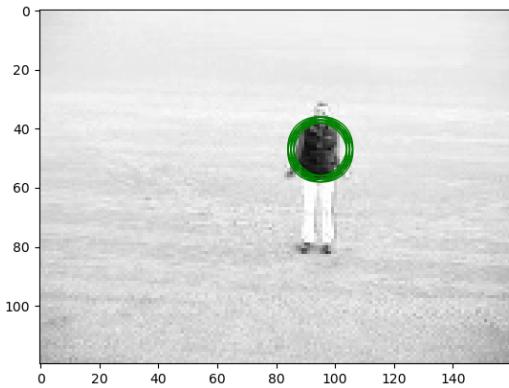
(τ') Πλαίσιο 16

Σχήμα 17: Εντοπισμός σημείων ενδιαφέροντος από τον ανιχνευτή Gabor  
 $\sigma = 2, \tau = 3$

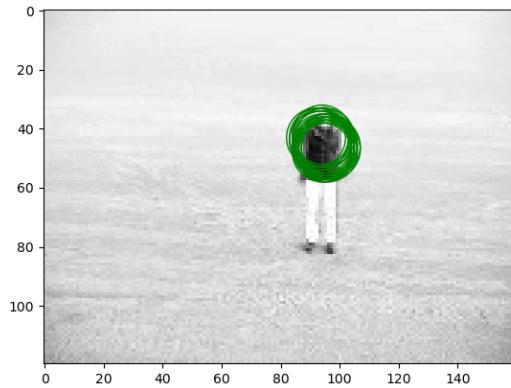
Γενικά, ο ανιχνευτής Gabor φαίνεται να εντοπίζει πιο ποικιλόμορφα χαρακτηριστικά. Συγκεκριμένα, στον κυματισμό χεριών εντοπίζει χαρακτηριστικά και στα δύο χέρια, σε αντίθεση με τον ανιχνευτή Harris ο οποίος εντοπίζει ουσιαστικά γωνίες και καταλήγει να βρίσκει χαρακτηριστικά μόνο στο ένα χέρι. Το ίδιο παρατηρείται και στο τρέξιμο και στο περπάτημα. Μάλιστα, παρότι ο διασκελισμός στο περπάτημα είναι μικρός και, εκ φύσεως, δεν μπορούν να διακριθούν με ευκολία τα χαρακτηριστικά στα πόδια, ο ανιχνευτής Gabor καταφέρνει και εκεί να τα διαφοροποιήσει.

### 2.1.3 Πολυκλιμακωτές Εκδοχές

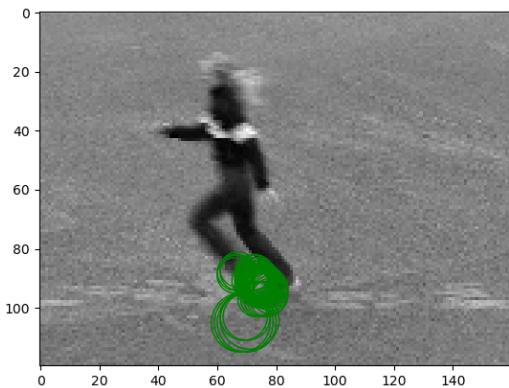
Για αμφότερους τους ανιχνευτές υλοποιήθηκαν επίσης πολυκλιμακωτές εκδοχές, αναζητώντας σημεία ενδιαφέροντος σε πολλαπλές χωρικές (όχι χρονικές) κλίμακες. Συγκεκριμένα, ακολουθήθηκε ανάλογη διαδικασία με αυτήν του πρώτου εργαστηρίου. Δεδομένης μίας αρχικής χωρικής κλίμακας, ενός πολλαπλασιαστικού όρου και ενός πλήθους κλιμάκων, υπολογίστηκαν όλες οι χωρικές κλίμακες και εκτέλεστηκε ο αντίστοιχος αλγόριθμος ανίχνευσης για καθεμία από αυτήν. Στην συνέχεια, επιλέχθηκαν τα σημεία αυτά τα οποία μεγιστοποιούσαν την LoG μετρική σε μία γειτονιά τριών κλιμάκων.



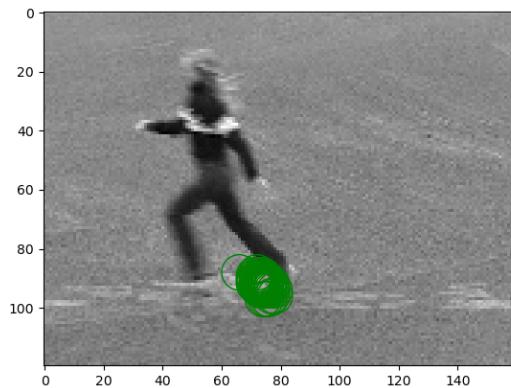
(α') Πλαίσιο 37



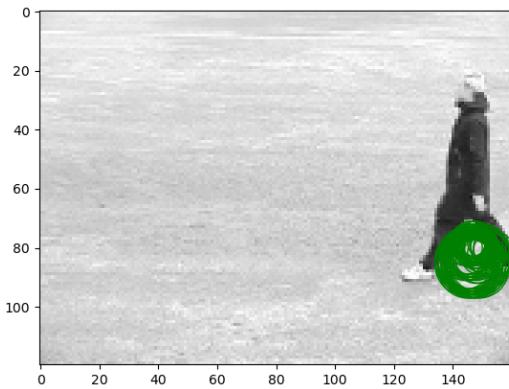
(β') Πλαίσιο 38



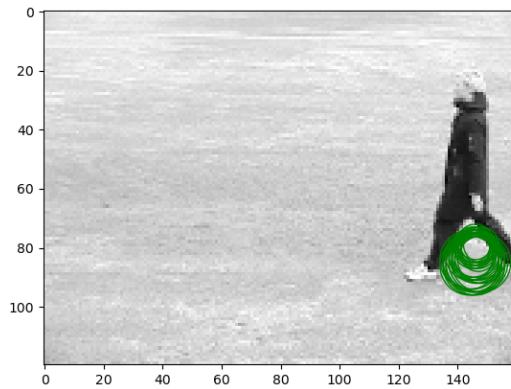
(α') Πλαίσιο 20



(β') Πλαίσιο 21



(α') Πλαίσιο 18



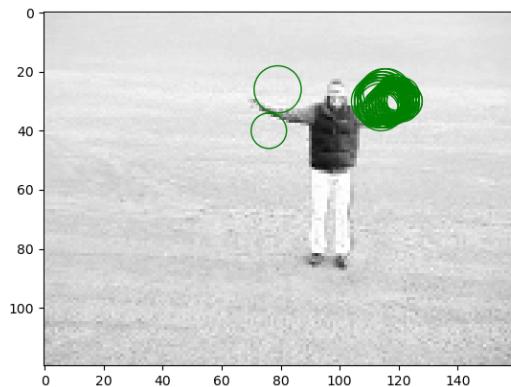
(β') Πλαίσιο 19

Σχήμα 20: Εντοπισμός σημείων ενδιαφέροντος από τον πολυκλιμακωτό ανιχνευτή harris

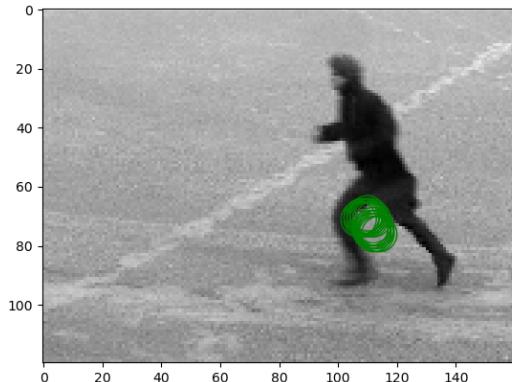
Τα αποτελέσματα για τον ανιχνευτή Gabor:



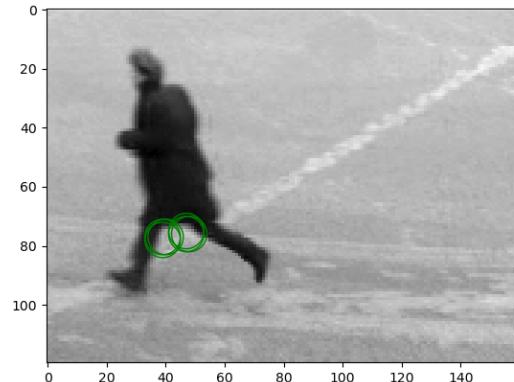
(α') Πλαίσιο 31



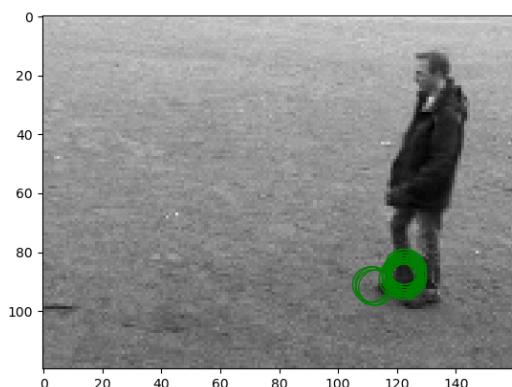
(β') Πλαίσιο 32



(α') Πλαίσιο 14



(β') Πλαίσιο 25



(γ') Πλαίσιο 24



(δ') Πλαίσιο 63

Σχήμα 22: Εντοπισμός σημείων ενδιαφέροντος από τον πολυκλιμακωτό ανιχνευτή Gabor

## 2.2 Χωροχρονικοί Ιστογραφικοί Περιγραφητές

Εξαχθέντων των σημείων ενδιαφέροντος, υπολογίζονται ιστογραφικοί περιγραφητές σε μία περιοχή γύρω από το καθένα. Συγχεκριμένα, υπολογίζεται αφενός η κατευθυντική παράγωγος καθενός πλαισίου της εικόνας, αφετέρου η οπική ροή, και εξάγονται τα διανυσματικά πεδία για μία περιοχή γύρω από το εκάστοτε σημείο ενδιαφέροντος.

## 2.3 Κατασκευή Bag of Visual Words και χρήση Support Vector Machines για ταξινόμηση δράσεων

Στο ερώτημα αυτό γίνεται η πραγματική κατηγοριοποίηση των βίντεο. Ως χαρακτηριστικά χρησιμοποιούνται Bag of Visual Words αναπαραστάσεις οι οποίες προκύπτουν είτε από HOG περιγραφητές, είτε από HOF περιγραφητές, είτε από την παράθεσή τους HOG / HOF.

Accuracy: 1.0

Predictions: [2 1 0 2 0 0 1 0 0 1 2 0 2 0 1]

Test labels: [2, 1, 0, 2, 0, 0, 1, 0, 0, 1, 2, 0, 2, 0, 1]

Accuracy: 0.8

Predictions: [2 1 0 2 0 0 2 0 0 0 1 0 2 0 1]

Test labels: [2, 1, 0, 2, 0, 0, 1, 0, 0, 1, 2, 0, 2, 0, 1]

Accuracy: 0.8

Predictions: [2 1 0 0 0 0 1 0 0 0 1 0 2 0 1]

Test labels: [2, 1, 0, 2, 0, 0, 1, 0, 0, 1, 2, 0, 2, 0, 1]

Σχήμα 23: Απόδοση HOG, HOF και HOG-HOF αντιστοίχως, για 500 σημεία ενδιαφέροντος από τον ανιχνευτή Gabor

Accuracy: 0.6666666666666666

Predictions: [2 1 0 1 0 1 1 0 1 1 0 1 2 0 1]

Test labels: [2, 1, 0, 2, 0, 0, 1, 0, 0, 1, 2, 0, 2, 0, 1]

Accuracy: 0.6

Predictions: [2 2 0 1 1 1 1 0 0 1 1 1 2 0 1]

Test labels: [2, 1, 0, 2, 0, 0, 1, 0, 0, 1, 2, 0, 2, 0, 1]

Accuracy: 0.8666666666666667

Predictions: [2 2 0 1 0 0 1 0 0 1 2 0 2 0 1]

Test labels: [2, 1, 0, 2, 0, 0, 1, 0, 0, 1, 2, 0, 2, 0, 1]

Σχήμα 24: Απόδοση HOG, HOF και HOG-HOF αντιστοίχως, για 500 σημεία ενδιαφέροντος από τον ανιχνευτή Harris

Όπως αναμενόταν από τον σχολιασμό των σημείων ενδιαφέροντος, ο ανιχνευτής Gabor αποδίδει καλύτερα από τον ανιχνευτή Harris, μάλλον χάρις στην ποικιλομορφία τους.

Εν πάση περιπτώσει, τα αποτελέσματα είναι αρκετά ικανοποιητικά. Φυσικά, δεν είναι πλήρως αντιπροσωπευτικά της πραγματικότητας, αφού το σύνολο δεδομένων είναι μικρό. Παραδείγματος χάριν, προφανώς δεν είναι δυνατόν να πετύχει 100 % ακρίβεια ο ανιχνευτής Gabor σε πλούσιότερο σύνολο δεδομένων.

Ο ανιχνευτής Gabor φαίνεται να μην οφελείται από τον συνδυασμό των περιγραφητών, σε αντίθετη με τον ανιχνευτή Harris, του οποίου η επίδοση βελτιώνεται.

Επιχειρήθηκε η αξιολόγηση και των πολυκλιμακωτών ανιχνευτών αλλά δεν χατέστη εφικτή λόγω τεχνικών προβλημάτων.

## Μέρος 3

Σε αυτό το μέρος του εργαστηρίου, στόχος είναι η δημιουργία μίας εφαρμογής συνένωσης διαδοχικών εικόνων, ληφθεισών από ένα σταθερό σημείο, περιστρέφοντας την κάμερα. Αυτό επιτυγχάνεται υπολογίζοντας την ομογραφία μεταξύ διαδοχικών εικόνων, ώστε να τοποθετηθούν στο ίδιο πλαίσιο αναφοράς.

Η διαδικασία που ακολουθείται είναι η εξής:

1. Διαβάζονται οι εικόνες.
2. Εντοπίζονται χαρακτηριστικά ενδιαφέροντος και τοπικοί περιγραφητές.
3. Αντιστοιχίζεται καθένα χαρακτηριστικό της πρώτης εικόνας με τα δύο κοντινότερά του στην δεύτερη.
4. Από όλες τις αντιστοιχίσεις, επιλέγονται όσες πληρούν το κριτήριο Lowe. Δηλαδή, αυτές στις οποίες ο λόγος απόστασης των ζευγών περιγραφητών είναι κοντά στην μονάδα. Έτσι, απορρίπτονται πολλά λάθος ταιριάσματα, χάνοντας ελάχιστα σωστά.
5. Υπολογίζεται η ομογραφία, χρήσει των καλών αντιστοιχίσεων, με την μέθοδο RANSAC.
6. Μετασχηματίζεται η πρώτη εικόνα μέσω της ομογραφίας και εκφράζονται οι συντεταγμένες της πάνω αριστερά γωνίας της ως προς το πλαίσιο αναφοράς της δεύτερης εικόνας.
7. Συνενώνονται οι εικόνες ώστε η πάνω αριστερά γωνίας της μετασχηματισμένης πρώτης να βρίσκεται στις προσδιορισμένες συντεταγμένες ως προς το πλαίσιο της δεύτερης.

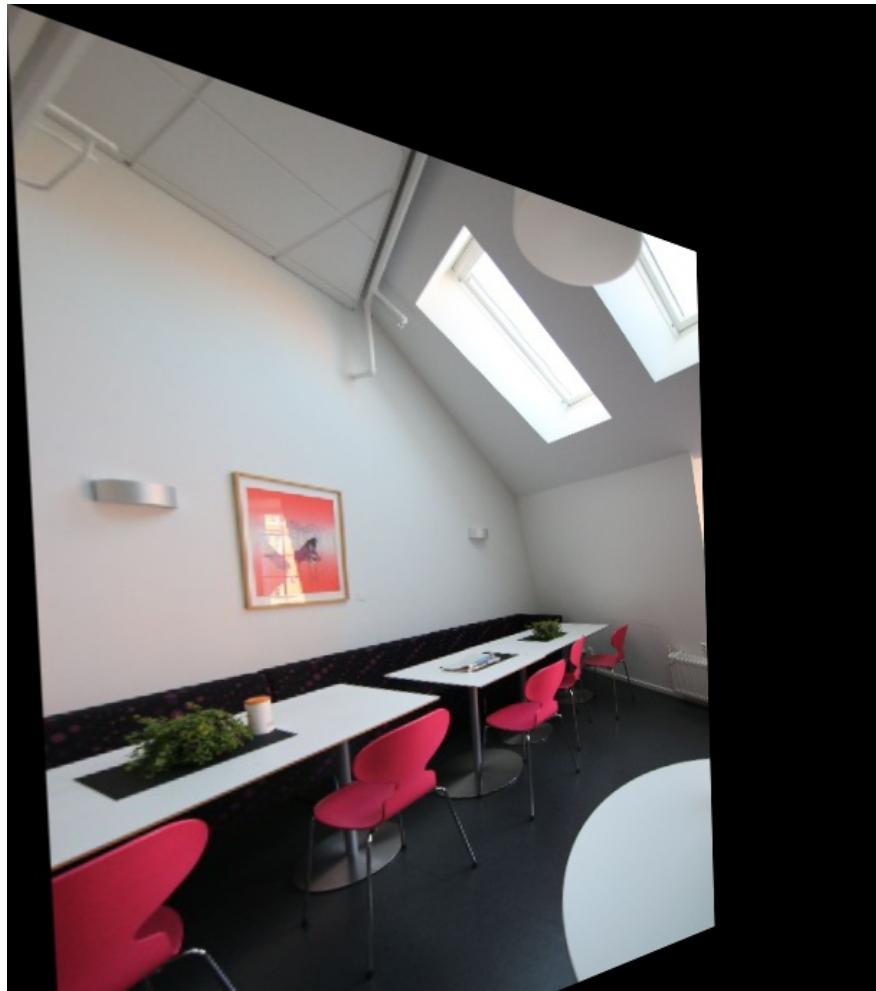
Στην παρούσα εργασία επιλέχθηκε να υλοποιηθούν τα βήματα 6 και 7 μαζί, σε μία ενιαία συνάρτηση `combine_images`.

Παρατίθεται, ενδεικτική εικόνα της διαδικασίας ταιριάσματος:



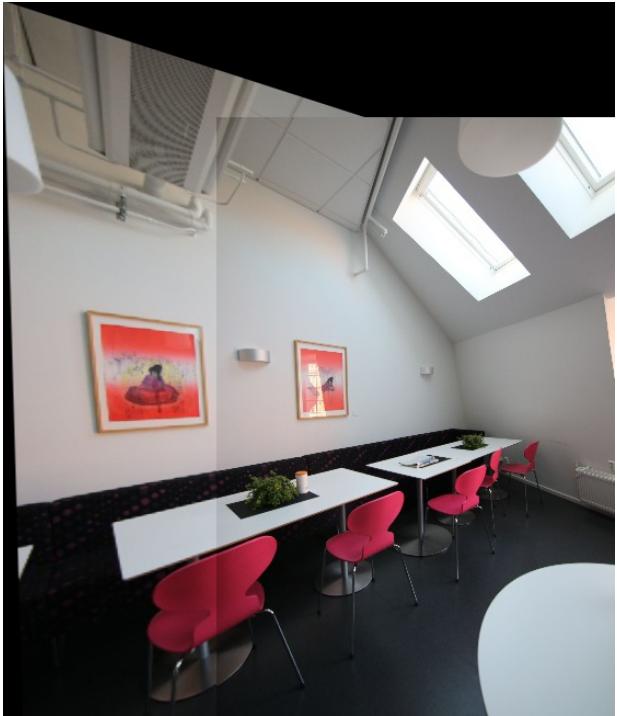
Σχήμα 25: Ταιριάσμα σημείων ενδιαφέροντος

Εντοπίζονται μερικά σφράλματα η συντριπτική πλειονότητα των ταιριασμάτων είναι ικανοποιητική.  
Παρατίθεται μετασχηματισμένη εικόνα μέσω ομογραφίας:

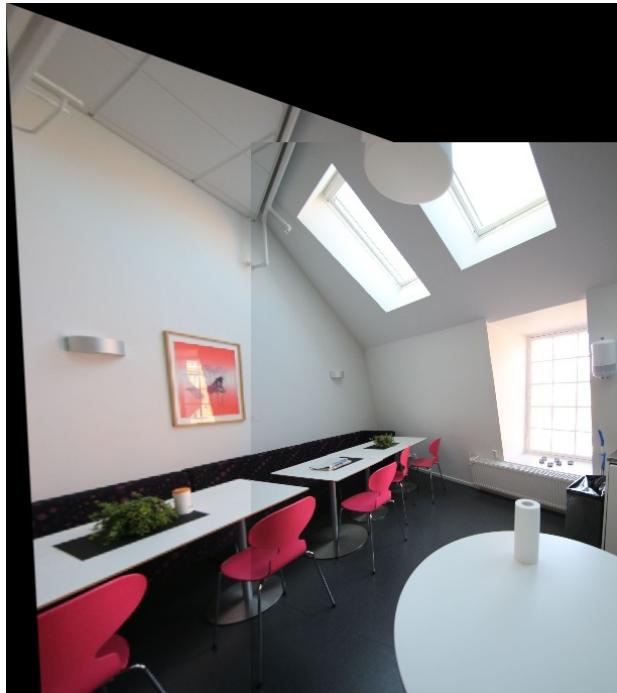


Σχήμα 26: Μετασχηματισμένη εικόνα 2 μέσω ομογραφίας

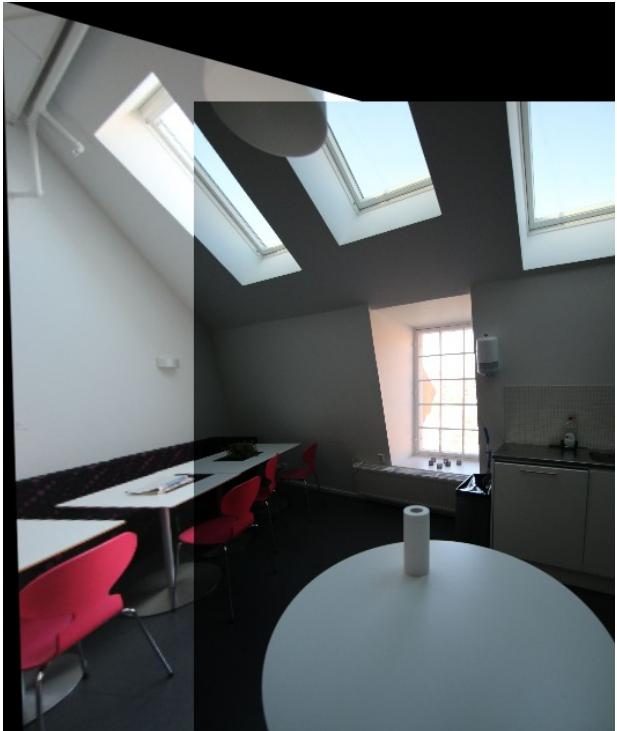
Παρατίθενται ζεύγη συνενώσεων:



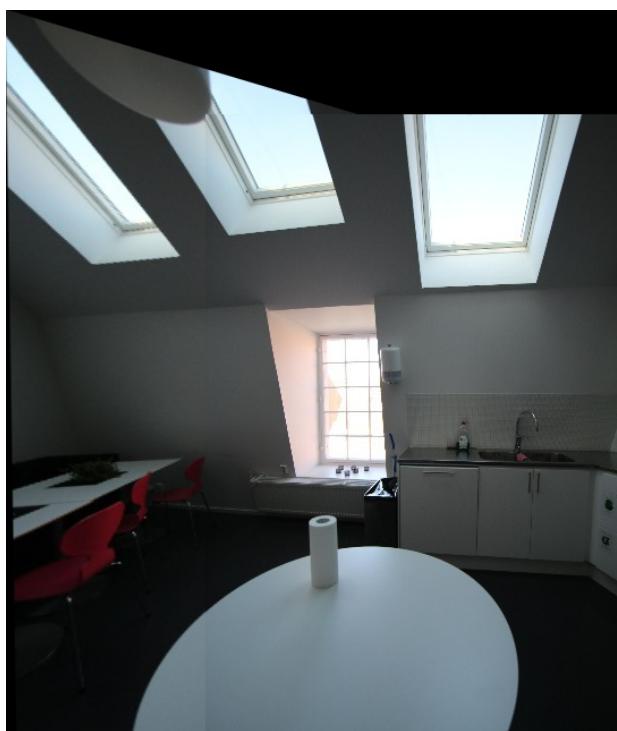
(α') Εικόνες 1 και 2



(β') Εικόνες 2 και 3



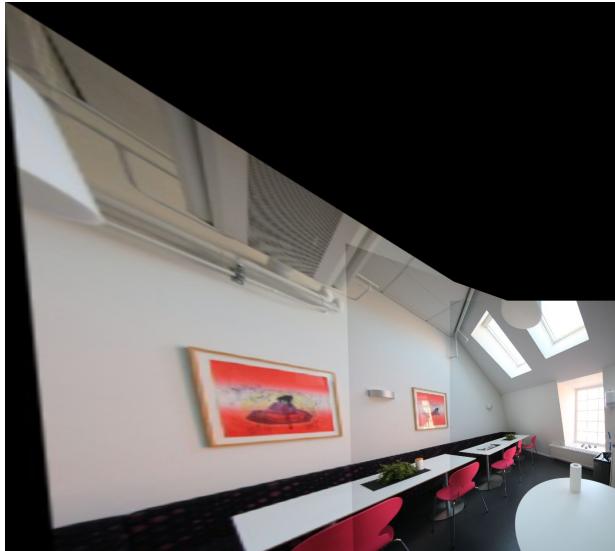
(γ') Εικόνες 3 και 4



(δ') Εικόνες 4 και 5

Σχήμα 27: Συνένωση εικόνων

Έγινε προσπάθεια, επίσης, για την συνένωση τριών συνολικά φωτογραφιών:



(α') Εικόνες 1, 2 και 3



(β') Εικόνες 4, 5 και 6

Σχήμα 28: Συνένωση εικόνων