

Όραση Υπολογιστών
Εργαστήριο 1

Αναστάσιος Στέφανος Αναγνώστου
03119051

Σπυρίδων Παπαδόπουλος
03119058

7 Απριλίου 2023

Περιεχόμενα

1	Μέρος 1	3
1.1	Ανίχνευση Δέρματος Προσώπου και Χεριών	3
1.2	Παρακολούθηση 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
2	Μέρος 2	10
3	Μέρος 3	11

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

1 Μέρος 1

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

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

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

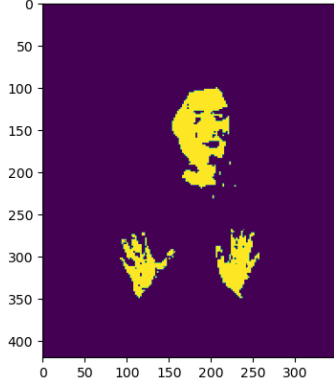
$$\mathbf{c} = [C_b \ C_r]^T \quad (1.2)$$

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:

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

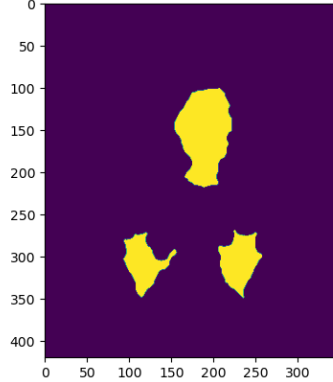
Σε αυτό το σημείο, η δυαδική εικόνα έχει μερικά κενά, τα οποία είναι ανεπιθύμητα, αφού είναι γνωστό ότι οι περιοχές δέρματος είναι συνεχείς. Αυτό επαληθεύεται από την ακόλουθη εικόνα 1α'.

Binary skin image before morphological operations



(α') Η ατελής δυαδική εικόνα δέρματος

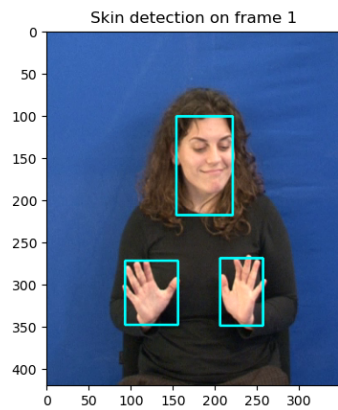
Binary skin image after morphological operations



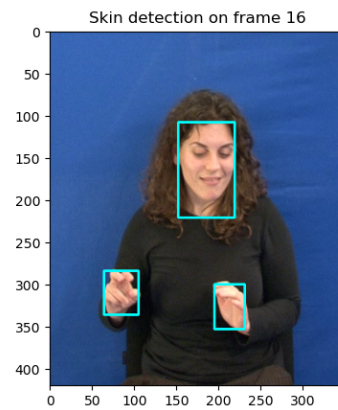
(β') Η βελτιωμένη δυαδική εικόνα δέρματος

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'.

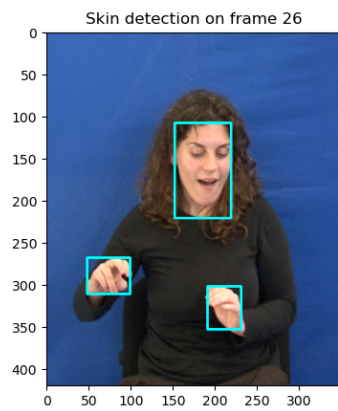
Παρατηρείται ότι οι περιοχές είναι τώρα συνεχείς και η απεικόνιση σαφώς βελτιωμένη, με τίμημα την ακρίβεια σε οριακές περιοχές όπως είναι τα δάχτυλα.



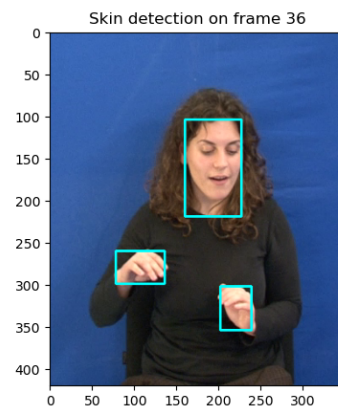
(α') Πλαίσιο 1



(β') Πλαίσιο 16



(γ') Πλαίσιο 26



(δ') Πλαίσιο 36

Σχήμα 2: Ενδεικτικός εντοπισμός δέρματος σε μερικά πλαίσια

Τέλος, θα 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 Παρακολούθηδη Προσώπου και Χεριών

Σε αυτό το σημείο θα 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

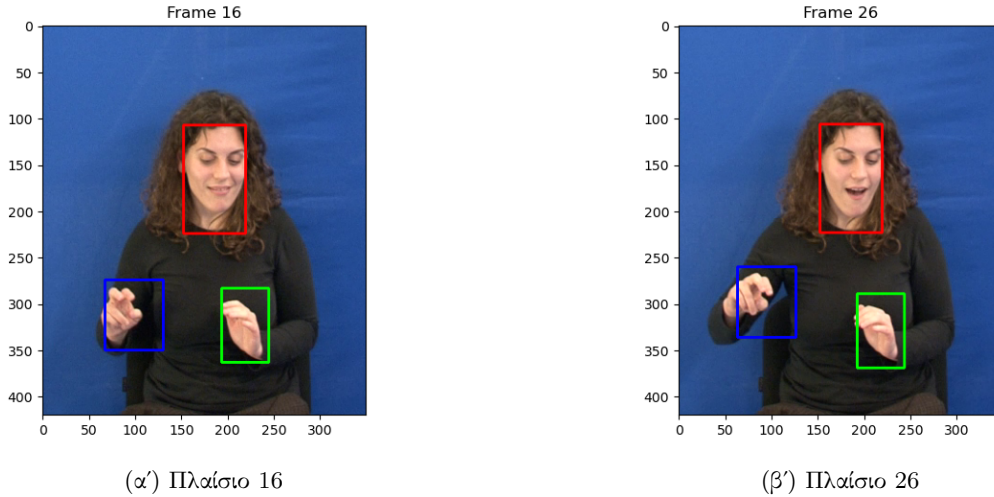
Η εκτίμηση της οπτικής ροής, εν προκειμένω, θα γίνει χρήση του αλγορίθμου Lucas-Kanade. Κατά τον αλγόριθμο αυτόν, υπολογίζεται η οπτική ροή καθενός σημείου της εικόνας, χρήση της μεθόδου ελαχίστων τετραγώνων, υπό την παραδοχή ότι αυτή παραμένει σταθερή σε μία μικρή περιοχή γύρω από το σημείο.

$$I_n(\mathbf{x}) \approx I_{n-1}(\mathbf{x} + \mathbf{d}) \stackrel{\text{Taylor}}{\Rightarrow} I_{n-1}(\mathbf{x} + \mathbf{d}) \approx I_{n-1}(\mathbf{x} + \mathbf{d}_i) + \nabla I_{n-1}(\mathbf{x} + \mathbf{d}_i)^T \mathbf{u} \quad (1.4)$$

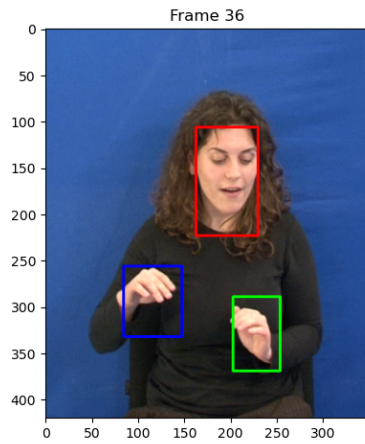
Ελαχιστοποιώντας το τετραγωνικό σφάλμα, προκύπτει ένα σύστημα το οποίο προσδιορίζει πλήρως την βελτίωση της εκτίμησης της οπτικής ροής.

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 όριο επαναλήψεων.

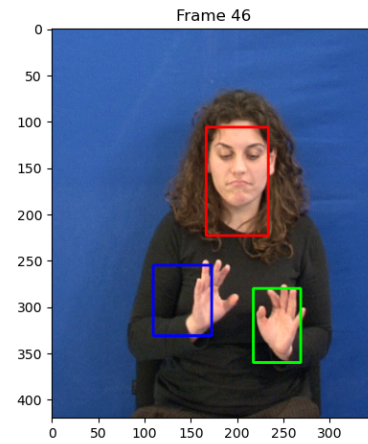
Παρατίθενται ενδεικτικά πλαίσια για την αξιολόγηση της μεθόδου στην εικόνα 3. Σημειώνεται, επίσης, ότι στα αρχεία της αναφοράς δίνεται το αποτέλεσμα της συνένωσης όλων των πλαισίων σε μορφή gif.



Σχήμα 3: Ενδεικτικός εντοπισμός δέρματος χρήση εκτίμησης οπτικής ροής



(α') Πλαίσιο 36



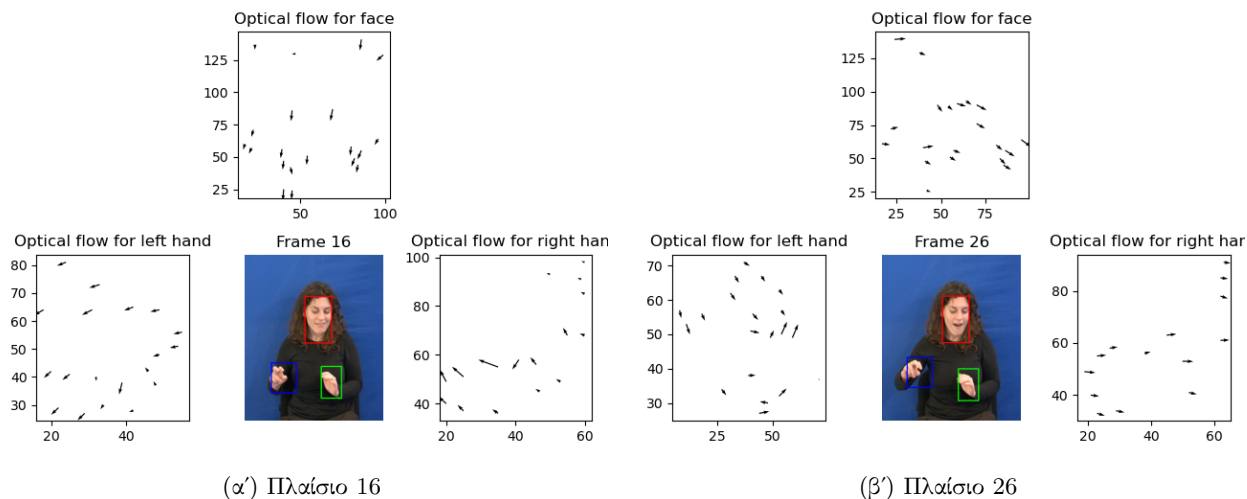
(β') Πλαίσιο 46

Σχήμα 4: Ενδεικτικός εντοπισμός δέρματος χρήσει εκτίμησης οπτικής ροής

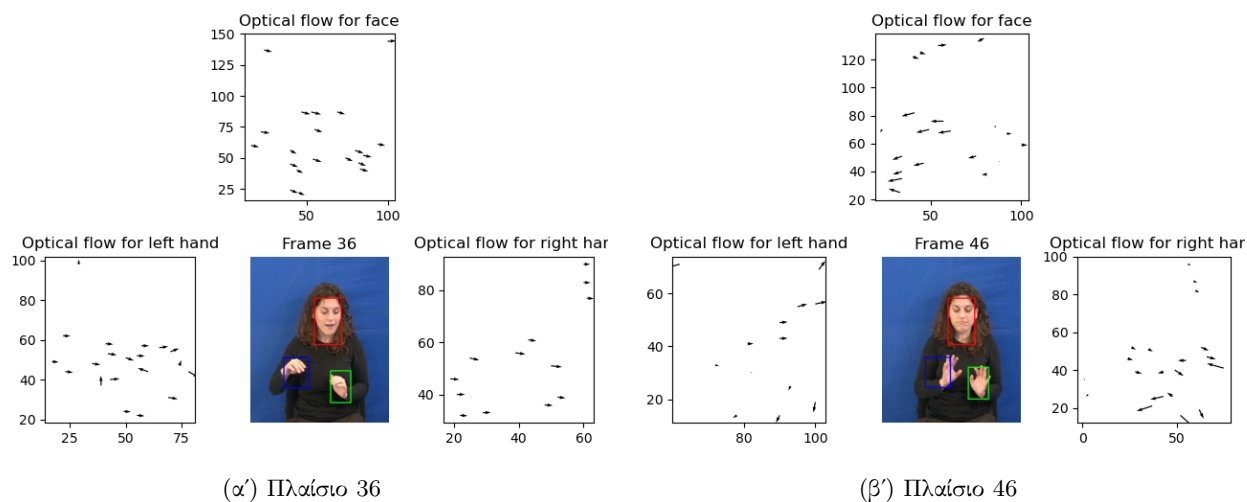
Επισκοπώντας τα αποτελέσματα, συμπεραίνεται, ότι η εκτίμηση της οπτικής ροής λειτουργεί αρκετά καλά για μικρές, αργές κινήσεις. Αυτό επαληθεύεται από τα 3 πρώτα παρατιθέμενα πλαίσια, στα οποία οι κινήσεις ήταν αργές και ομαλές. Ωστόσο, όταν οι κινήσεις γίνονται απότομες, η αποδοτικότητα του αλγορίθμου μειώνεται, με αποτέλεσμα να υστερούν τα κουτιά από τις περιοχές ενδιαφέροντος, όπως φαίνεται στο τελευταίο πλαίσιο. 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 $\|d^2\| = d_x^2 + d_y^2$ μεγαλύτερη από ένα κατώφλι, ώστε να απορριφθούν outliers και να επιτευχθεί καλύτερη ακρίβεια.



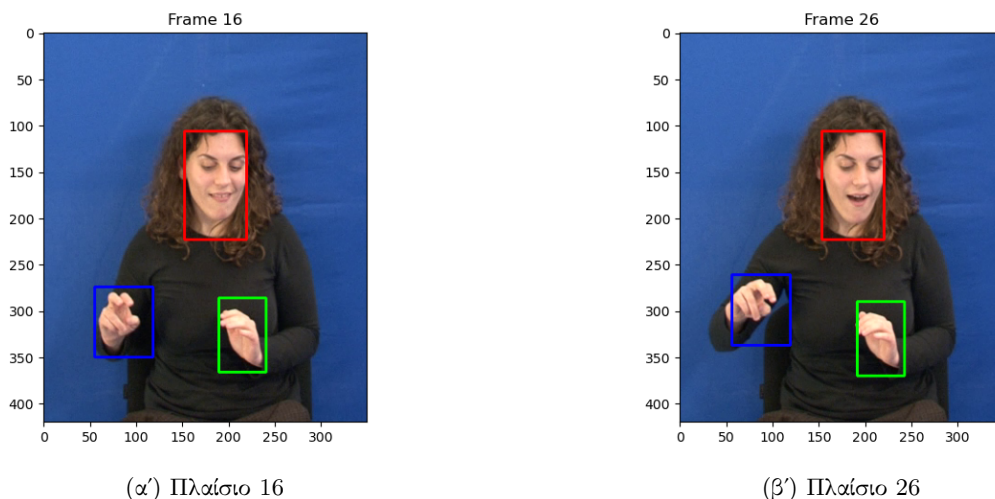
Σχήμα 5: Εντοπισμός δέρματος και οπτική ροή σε μερικά πλαίσια



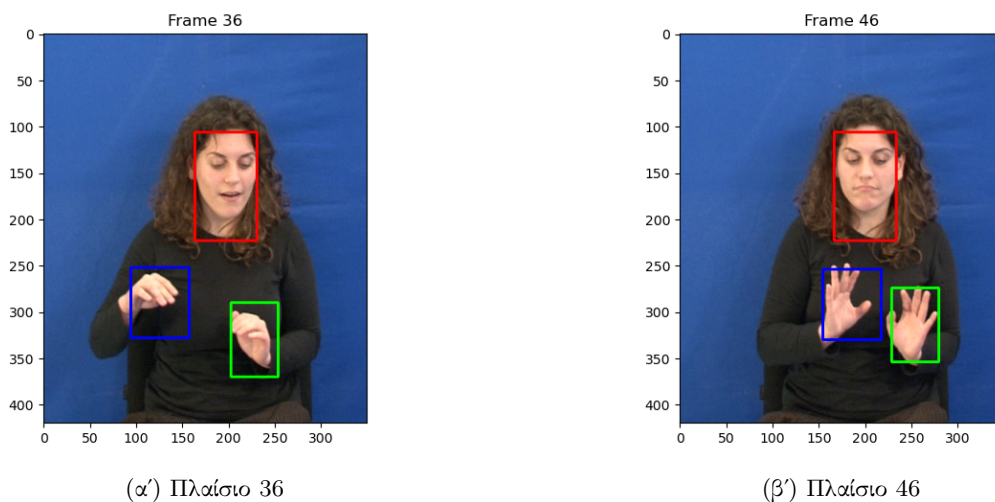
Σχήμα 6: Εντοπισμός δέρματος και οπτική ροή σε μερικά πλαίσια

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 υποδειγματοληπτική 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.



Σχήμα 7: Ενδεικτικός πολυκλιμακωτός εντοπισμός δέρματος χρήσει εκτίμησης οπτικής ροής

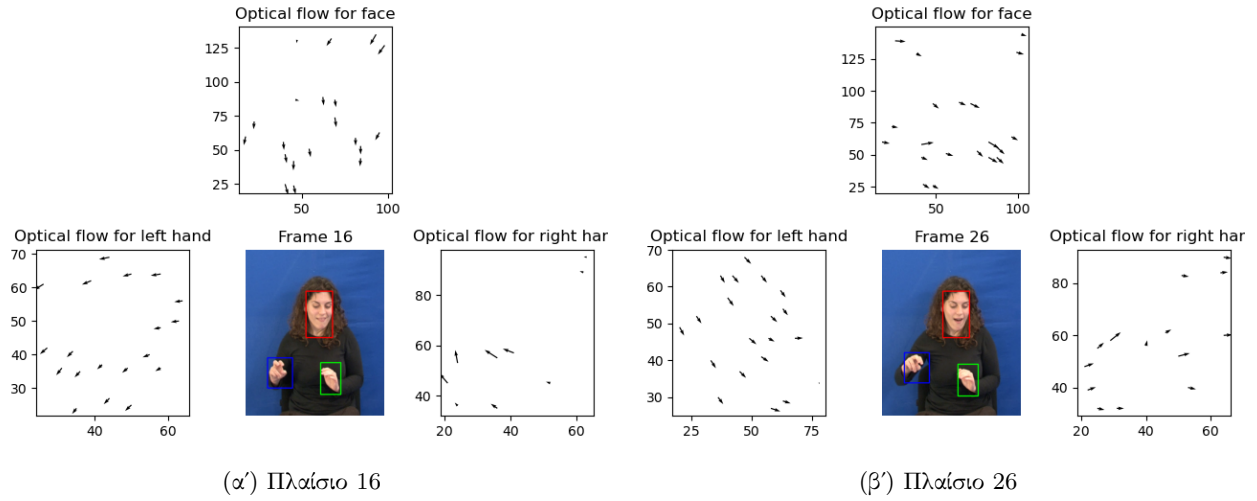


Σχήμα 8: Ενδεικτικός πολυκλιμακωτός εντοπισμός δέρματος χρήσει εκτίμησης οπτικής ροής

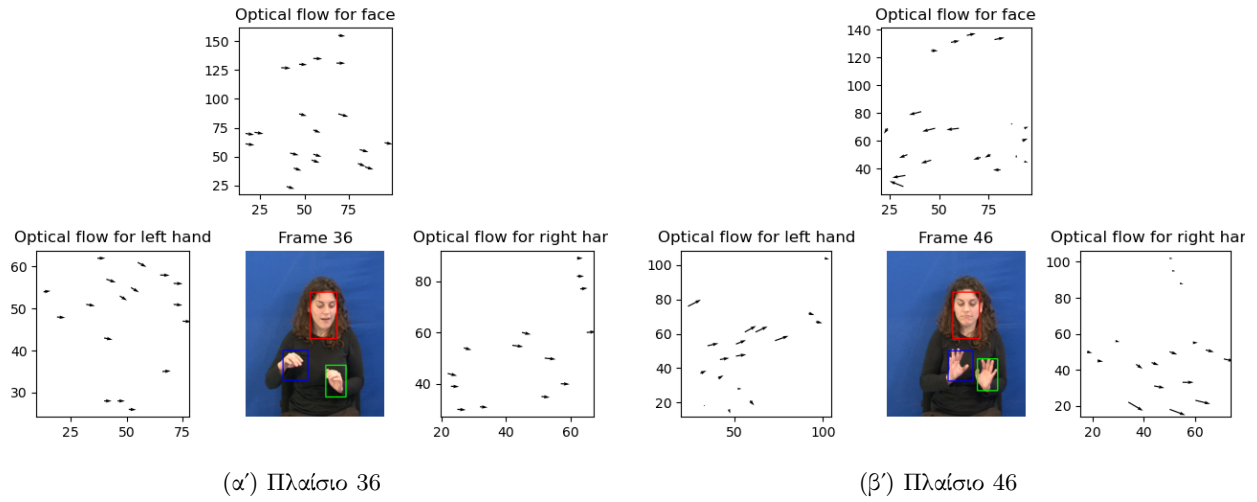
Παρατηρείται, ότι σε κάθε πλαίσιο ο εντοπισμός είναι ακριβέστερος. Η διαφορά είναι ιδιαίτερα αξιοσημείωτη στο τελευταίο παρατηθέμενο πλαίσιο, όπου το μπλε κουτί περικυκλώνει τέλεια το αριστερό χέρι, ενώ στην μονοκλιμακωτή εκδοχή το έχανε λόγω ταχύτητας.

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 διαχυμάνσεις.



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



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

2 Μέρος 2

3 Μέρος 3