



Bonus Project

Nikolaos Mitsopoulos
TH20591

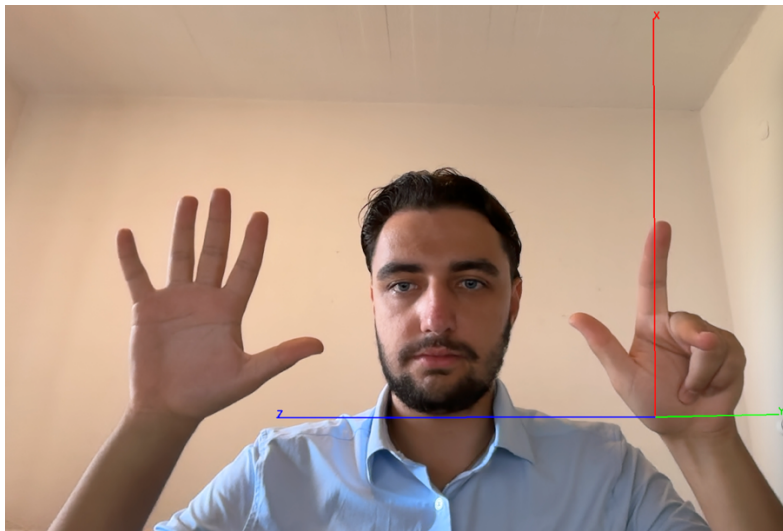
Περιγραφή Στόχου

Η εφαρμογή επιτυγχάνει **σε πραγματικό χρόνο (real-time)** ανίχνευση του δεξιού χεριού και την **οπτικοποίηση ενός τοπικού τρισδιάστατου συστήματος συντεταγμένων** που βασίζεται στις σχετικές θέσεις των δακτύλων.

Οι βασικές βιβλιοθήκες που χρησιμοποιούνται είναι:

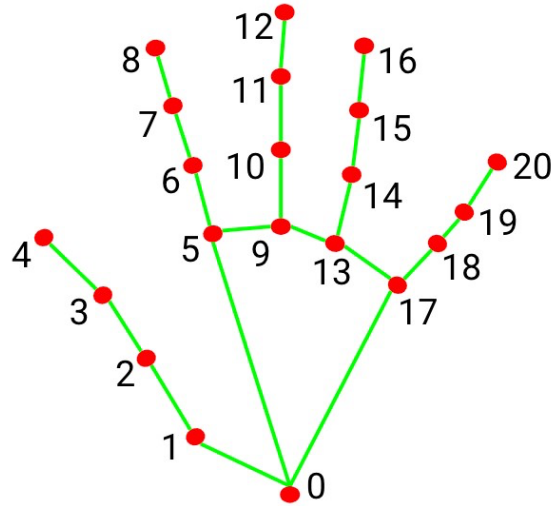
- **cv2** (Πρόσβαση στην κάμερα, προβολή και σχεδίαση)
- **mediapipe** (Ανίχνευση χαρακτηριστικών σημείων χεριού)
- **numpy** (Υπολογισμοί σε διανυσματικές ποσότητες)
- **matplotlib** (3D γραφική απεικόνιση)
- **sys** (Διαχείριση stdout για καλύτερη ανανέωση τερματικού)

Αναλυτικότερα με την χρήση του **mediapipe.solutions.hands.Hands()** εντοπίζει 21 landmark σημεία για κάθε χέρι. Στον κώδικα που έγραψα φρόντισα να εντοπίζει μόνο το δεξί χέρι όπως φαίνεται στην παρακάτω φωτογραφία.



Εικόνα 1 Ανίχνευση δεξιού χεριού με Mediapipe

Ο υπολογισμός βασίζεται σε **διανυσματικά cross products**. Για την κατασκευή του τοπικού συστήματος συντεταγμένων στο χέρι, ορίζεται ως origin (αφετηρία) το landmark 1, δηλαδή η βάση του δείκτη. Ο άξονας X ορίζεται ως το διάνυσμα από το landmark[1] προς το landmark[5], που είναι η βάση του δείκτη. Για την κατεύθυνση του αντίχειρα χρησιμοποιείται το διάνυσμα από το landmark[1] προς το landmark[2]. Ο άξονας Z υπολογίζεται ως το διάνυσμα που είναι όσο το δυνατόν πιο ευθυγραμμισμένο με την κατεύθυνση του αντίχειρα αλλά κάθετο στον άξονα X. Αυτό επιτυγχάνεται με το διάνυσμα $\text{cross}(X, \text{cross}(\text{Thumb}, X))$. Στη συνέχεια, ο άξονας Y προκύπτει από το cross product του Z με τον X, ώστε να διατηρείται δεξιόστροφο και ορθογώνιο το σύστημα.

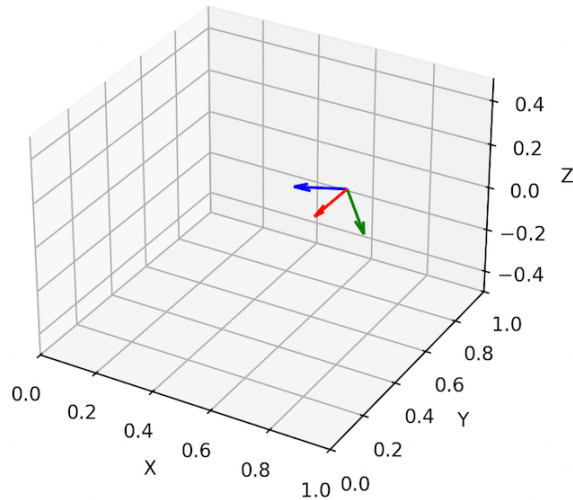


Εικόνα 2 Οπτική αναπαράσταση των σημείων landmark

Η οπτικοποίηση του συστήματος γίνεται με δύο τρόπους. Πρώτον, μέσω της βιβλιοθήκης matplotlib, χρησιμοποιείται η συνάρτηση `ax.quiver` για την απεικόνιση των αξόνων στο 3D διάγραμμα. Οι άξονες προβάλλονται με διαφορετικά χρώματα: ο X με κόκκινο, ο Y με πράσινο και ο Z με μπλε. Δεύτερον, στην 2D απεικόνιση μέσω OpenCV, οι τρισδιάστατες συντεταγμένες μετατρέπονται σε δισδιάστατες (pixel) και προβάλλονται απευθείας πάνω στο καρέ της κάμερας. Επιπλέον, με τη χρήση της `cv2.putText`, προστίθενται οι αντίστοιχες ετικέτες "X", "Y" και "Z" για να υποδεικνύονται οι κατευθύνσεις.

Η λειτουργία του κώδικα βασίζεται στην **πραγματικού χρόνου επεξεργασία εικόνας**, κάτι που επιτυγχάνεται με τη χρήση της εντολής `plt.pause(0.001)` σε κάθε επανάληψη του βρόχου. Αυτή η εντολή επιτρέπει στο matplotlib να ενημερώνει δυναμικά το 3D γράφημα χωρίς να διακόπτεται η εκτέλεση, εξασφαλίζοντας ομαλή και συνεχή απεικόνιση του συστήματος συντεταγμένων του χεριού.

Live 3D Coordinate Frame



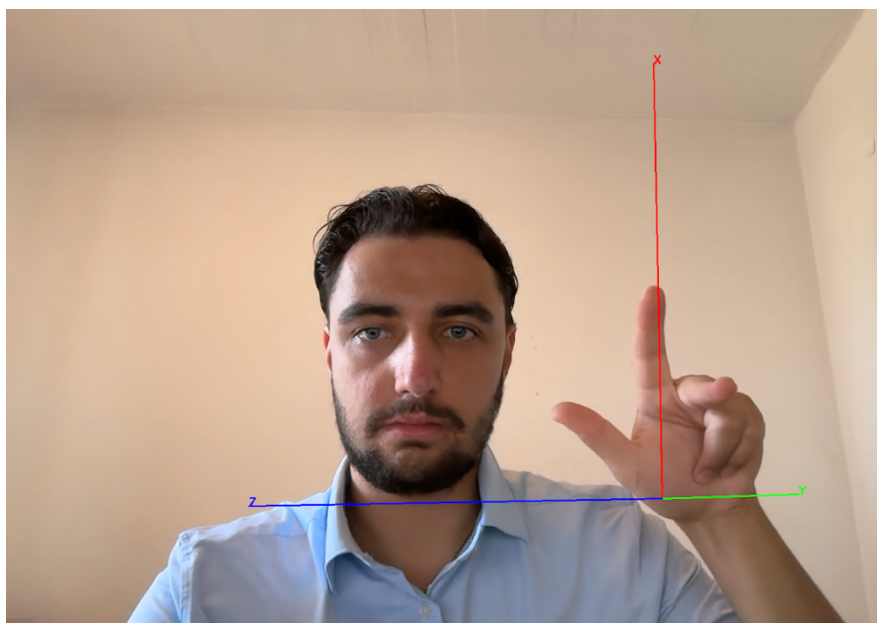
Εικόνα 3 Τρισδιάστατη απεικόνιση τοπικού συστήματος συντεταγμένων χεριού

Στην παραπάνω εικόνα φαίνεται η ζωντανή απεικόνιση του τοπικού συστήματος συντεταγμένων του χεριού, όπως αυτό προκύπτει από την επεξεργασία του Mediapipe. Ο άξονας X (κόκκινος) δείχνει προς τον δείκτη, ο Z (μπλε) είναι κάθετος στον X και ευθυγραμμίζεται όσο το δυνατόν περισσότερο με τον αντίχειρα, ενώ ο Y (πράσινος) προκύπτει από το δεξιόστροφο γινόμενο $Z \times X$. Παρατηρείται ότι το σύστημα είναι κανονικοποιημένο και τοποθετημένο στη βάση του αντίχειρα, παρέχοντας σαφή γεωμετρική ερμηνεία της στάσης του χεριού.

Για την καλύτερη κατανόηση της ροής της εφαρμογής, παρακάτω συνοψίζεται η βασική λογική της λειτουργίας της με μορφή διαγράμματος:

1. Εκκίνηση κάμερας
2. Ανάγνωση καρέ
3. Αντιστροφή και μετατροπή σε RGB
4. Εντοπισμός χεριού μέσω Mediapipe
5. Έλεγχος αν είναι δεξί χέρι
6. Υπολογισμός αξόνων (X, Y, Z)
7. Οπτικοποίηση σε 3D και 2D

8. Εμφάνιση τελικής εικόνας



Εικόνα 4 Τελική απεικόνιση του X , Y , Z άξονα πάνω στην εικόνα του χεριού με χρήση *OpenCV* – κόκκινος (X), πράσινος (Y), μπλε (Z).

Συμπεράσματα

Η παρούσα εργασία πέτυχε την οπτικοποίηση ενός τοπικού τρισδιάστατου συστήματος συντεταγμένων στο ανθρώπινο χέρι, σε πραγματικό χρόνο, με χρήση της βιβλιοθήκης Mediarpipe για τον εντοπισμό σημείων landmark. Μέσω γεωμετρικών υπολογισμών (διανυσματικές πράξεις και κανονικοποίηση), το σύστημα αποδίδει με ακρίβεια τους άξονες X, Y και Z του χεριού, ενώ η οπτικοποίηση υλοποιείται τόσο σε 3D διάγραμμα όσο και πάνω στο καρέ της κάμερας.