

ΤΜΗΜΑ ΕΠΙΚΟΙΝΩΝΙΑΣ ΚΑΙ ΨΗΦΙΑΚΩΝ ΜΕΣΩΝ  
ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ  
ΑΝΑΠΤΥΞΗ ΨΗΦΙΑΚΩΝ ΠΑΙΧΝΙΔΙΩΝ ΚΑΙ ΠΟΛΥΜΕΣΙΚΩΝ  
ΕΦΑΡΜΟΓΩΝ

## **ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ**

**«ΑΝΑΠΤΥΞΗ ΕΝΟΣ ΔΙΑΔΡΑΣΤΙΚΟΥ ΠΑΙΧΝΙΔΙΟΥ ΕΙΚΟΝΙΚΗΣ  
ΠΡΑΓΜΑΤΙΚΟΤΗΤΑΣ ΒΑΣΙΣΜΕΝΟ ΣΤΗΝ ΑΝΑΓΝΩΡΙΣΗ  
ΚΙΝΗΣΕΩΝ ΤΟΥ ΧΡΗΣΤΗ»**

**του**

**Γκούτζιου Στέφανου**

**Επιβλέπων:**

Επίκουρος Καθηγητής Μιχαήλ Βρίγκας

**Μέλη της Τριμελούς:**

Καθηγητής Λάμπας Γεώργιος

Επίκουρος Καθηγητής Κλεφτοδήμος Αλέξανδρος

Μάρτιος 2024

## Περίληψη

Το αντικείμενο που πραγματεύεται η παρούσα διπλωματική εργασία είναι η δημιουργία ενός διαδραστικού πολυπρακτορικού παιχνιδιού με στόχο να αυξηθεί η αίσθηση εμπειρίας του παίχτη αλλά και η απλοποίηση λειτουργικότητας του παιχνιδιού. Η εργασία αυτή αποτελείται από δύο μέρη: από την ανίχνευση ανθρώπινων χειρονομιών και χρωματιστών αντικειμένων, για την αντικατάσταση των χειριστηρίων μιας παιχνιδομηχανής και από τη σχεδίαση ενός βιντεοπαιχνιδιού.

Στο πρώτο μέρος θα γίνει ανάλυση ορισμένων μεθόδων κατάτμησης της ψηφιακής εικόνας και θα επιλεγεί η αποδοτικότερη για να υλοποιηθούν εφαρμογές αναγνώρισης χειρονομιών και αντικειμένων με την βοήθεια της υπολογιστικής όρασης. Συγκεκριμένα με τη χρήση μίας κάμερας θα γίνεται η λήψη συνεχόμενων εικόνων, με κεντρικό θέμα το αντικείμενο που κρατά ο παίχτης στα χέρια του, στη συνέχεια μετά από κατάλληλη ψηφιακή επεξεργασία των εικόνων με βάση το χρώμα θα υπολογίζονται οι συντεταγμένες του έγχρωμου αντικειμένου που είναι συνδεδεμένες με την κίνηση του χαρακτήρα του παιχνιδιού ενώ με την ανίχνευση ανθρώπινων χειρονομιών του παίχτη θα επιτυγχάνεται η πλοήγηση στις υπόλοιπες λειτουργίες του παιχνιδιού.

Το δεύτερο μέρος αποτελείται από την δημιουργία ενός βιντεοπαιχνιδιού. Ο παίχτης θα συμμετέχει σε ένα αγώνα τένις με αντίπαλο τον υπολογιστή ή με έναν άλλον παίχτη ταυτόχρονα.

Για την υλοποίηση της εργασίας αξιοποιήθηκε η βιβλιοθήκη ανοιχτού κώδικα OpenCV η οποία περιέχει υλοποιήσεις αλγορίθμων που σχετίζονται με την υπολογιστική όραση. Η εργασία αναπτύχθηκε σε γλώσσα Python και έχει χρησιμοποιηθεί επίσης το εργαλείο Microsoft Visual Studio ενώ για την δημιουργία γραφικού περιβάλλοντος χρησιμοποιήθηκε το λογισμικό Unity.

Ο απώτερος στόχος αυτής της εργασίας είναι να δημιουργήσει μια πιο φυσική μορφή ελέγχου που επιτρέπει στους παίκτες να επικεντρωθούν στον ενθουσιασμό του αγώνα χωρίς να ανησυχούν για το πάτημα των κουμπιών ή τις κινήσεις των χειριστηρίων του παιχνιδιού.

## **Abstract**

The subject of the thesis is the creation of an interactive game with the aim of increasing the player's sense of experience but also simplifying the functionality of the game. This work consists of two parts, the detection of human gestures, objects to replace game controls and the design of a video game.

The first part consists of the recognition of gestures and colored objects with the help of computer vision. Specifically, a camera captures continuous images that, with the help of an algorithm, calculate the coordinates of the colored object that are connected to the movement of the game character, while by detecting human gestures, you can navigate the rest of the game's functions.

The second part consists of creating a 3D graphical video game environment. The player participates in a tennis match against the computer or simultaneously with another player.

For the implementation of the work, the OpenCV open-source library was used, which contains implementations of algorithms related to computer vision. The work was developed in Python language and the Microsoft Visual Studio tool was also used, while the Unity software was used to create the graphical environment. The ultimate goal of this thesis is to create a more natural form of control that allows players to focus on the excitement of the match without worrying about button presses or joystick movements.

# Πίνακας Περιεχομένων

<b>Πίνακας Εικόνων.....</b>	<b>v</b>
<b>Κεφάλαιο 1.....</b>	<b>1</b>
1.1 Εισαγωγή.....	1
1.2 Κίνητρο και επιδιωκόμενα αποτελέσματα .....	2
<b>Κεφάλαιο 2 Υπολογιστική όραση και επικοινωνία ανθρώπου - μηχανής.....</b>	<b>4</b>
2.1 Τι είναι η υπολογιστική όραση;.....	4
2.2 Κατανόηση της ανθρώπινης όρασης .....	5
2.3 Εξαγωγή πληροφοριών από εικόνες.....	6
2.3.1 Η όραση ως συσκευή μέτρησης .....	6
2.3.2 Πηγή σημασιολογικών πληροφοριών.....	6
2.4 Εφαρμογές υπολογιστικής όρασης.....	7
<b>Κεφάλαιο 3 Κατάτμηση εικόνας.....</b>	<b>9</b>
3.1 Κατάτμηση εικόνας .....	9
3.2 Βασικές τεχνικές τμηματοποίησης εικόνας.....	10
<b>Κεφάλαιο 4 Ανίχνευση αντικειμένου με βάση το χρώμα .....</b>	<b>12</b>
4.1 Εισαγωγή.....	12
4.2 Μοντέλο χρώματος RGB .....	12
4.2.1 Αριθμητικές παραστάσεις.....	14
4.3 Μοντέλο χρώματος CMY.....	15
4.4 HSV (απόχρωση, κορεσμός και τιμή) .....	16
4.5 Μέθοδος ανίχνευσης χρώματος.....	18
<b>Κεφάλαιο 5 Αναγνώριση χειρονομιών.....</b>	<b>20</b>
5.1 Εισαγωγή.....	20
5.2 Απόκτηση και προεπεξεργασία εικόνας .....	21
5.3 Τεχνικές αναγνώρισης χειρονομιών .....	23
5.3.1 Ανίχνευση ακμών.....	23
5.3.2 Αναγνώριση με την τεχνική τμηματοποίησης χρώματος.....	27
5.3.3 Αναγνώριση χεριού μέσω του MediaPipe .....	28

5.4 Χαρτογράφηση με χειρονομίες .....	30
5.5 Συμπεράσματα.....	31
<b>Κεφάλαιο 6 Υλοποίηση εφαρμογής.....</b>	<b>32</b>
6.1 Εισαγωγή.....	32
6.2 OpenCV .....	32
6.3 Αλγόριθμος ανίχνευσης αντικειμένων και παρακολούθησης με βάση το χρώμα .....	33
6.3.1 Είσοδος: Λήψη βίντεο μέσω κάμερας web.....	33
6.3.2 Ανάγνωση βίντεο .....	34
6.3.3 Ανάγνωση ροή βίντεο σε καρτέ εικόνων.....	36
6.3.4 Μετατροπή της εικόνας από RGB σε χρωματικό χώρο HSV .....	37
6.3.5 Λειτουργία του εύρους HSV στο OpenCV.....	37
6.3.6 Συγκάλυψη .....	39
6.3.7 Σχεδιασμός περιγραμμάτων .....	40
6.3.8 Συμπέρασμα .....	40
6.4 Αλγόριθμος ανίχνευσης χειρονομιών .....	41
6.4.1 Εισαγωγή.....	41
6.4.2 Αναγνώριση χειρονομίας με MediaPipe .....	41
6.4.6 Περιγραφή του συνόλου δεδομένων .....	44
6.5 Σχεδίαση παιχνιδιού .....	46
6.5.1 Εισαγωγή.....	46
6.5.2 Unity .....	46
6.5.3 Η διεπαφή χρήστη .....	46
6.5.4 Αντικείμενο παιχνιδιού .....	48
6.5.5 Ελεγκτής χαρακτήρων .....	49
6.5.6 Οι ήρωες του παιχνιδιού.....	49
6.5.7 Κανόνες – Συστήματα .....	52
6.5.7.1 Ανάλυση κανόνων.....	52
6.5.8 Φωτισμός.....	53
6.5.9 Διάγραμμα Ροής για την Έναρξη του Παιχνιδιού .....	54
6.6 Σύνδεση αλγορίθμων με το πρόγραμμα Unity μέσω πρωτόκολλου επικοινωνίας.....	56
6.6.1 Εισαγωγή.....	56
6.6.2 Πρωτόκολλα επικοινωνίας .....	56
6.6.3 Εκτέλεση του διακομιστή UDP σε Python.....	57
6.6.4 Λήψη δεδομένων μέσω UDP στη Unity.....	58
<b>Κεφάλαιο 7 Αξιολόγηση της εφαρμογής.....</b>	<b>61</b>
7.1 Αξιολόγηση .....	61
<b>Κεφάλαιο 8 Συμπεράσματα.....</b>	<b>64</b>
<b>Βιβλιογραφία .....</b>	<b>66</b>
Δήλωση Πνευματικών Δικαιωμάτων .....	69

## Πίνακας Εικόνων

Εικόνα 1: Η όραση υπολογιστή στη διασταύρωση πολλαπλών επιστημονικών πεδίων (Πηγή: Stanford CS 231n) .....	4
Εικόνα 2: Έγχρωμο μοντέλο RGB (Πηγή :Encyclopaedia Britannica .Inc).....	13
Εικόνα 3: Αναπαράσταση χρωματικού μοντέλου RGB σε κύβο (Πηγή: 2D Computer Vision: Principles, Algorithms and Applications Yu-Jin Zhang, Tsinghua University, China.).....	13
Εικόνα 4: Χρωματικό μοντέλο CMY (Πηγή :Encyclopaedia Britannica .Inc) .....	15
Εικόνα 5: Χρωματικό μοντέλο HSV (Πηγή : SharkD, CC BY-SA 3.0 via Wikimedia Commons <a href="https://en.wikipedia.org/wiki/HSL_and_HSV#/media/File:HSV_color_solid_cylinder_saturation_gray.png">https://en.wikipedia.org/wiki/HSL_and_HSV#/media/File:HSV_color_solid_cylinder_saturation_gray.png</a> ).....	17
Εικόνα 6: Μπλοκ διάγραμμα μεθόδου ανίχνευσης αντικειμένου με βάση το χρώμα .....	18
Εικόνα 7: Σύστημα καταγραφής απόσταση των σημείων στη σκηνή με αισθητήρες βάθους (Πηγή: <a href="http://www.ifixit.com/Teardown/Microsoft-Kinect-Teardown/4066/1">http://www.ifixit.com/Teardown/Microsoft-Kinect-Teardown/4066/1</a> )	22
Εικόνα 8: Εφαρμογή Gaussian Filter .....	23
Εικόνα 9: Εικονογράφηση του τρόπου με τον οποίο το φίλτρο Sobel δημιουργεί κλίσεις X και Y (Πηγή: <a href="https://medium.com/@rohit-krishna/coding-canny-edge-detection-algorithm-from-scratch-in-python-232e1fdceac7">https://medium.com/@rohit-krishna/coding-canny-edge-detection-algorithm-from-scratch-in-python-232e1fdceac7</a> ) .....	24
Εικόνα 10: Εφαρμογή φίλτρων Sobel ως προς τους άξονες X & Y .....	25
Εικόνα 11: Τελικό αποτέλεσμα του Canny Edge Detector .....	25
Εικόνα 12: Ανίχνευση άκρων με τη χρήση αλγορίθμου Canny edge detector .....	26
Εικόνα 13: Αφαίρεση φόντου, μάσκα εικόνας και τμηματοποίηση χρώματος .....	27
Εικόνα 14: Επισκόπηση αγωγού αντίληψης χεριού [32].....	28
Εικόνα 15: Ορόσημο χεριός στο MediaPipe [33] Αναγνωριστής χεριών.....	29
Εικόνα 16: Αναγνώριση χειρονομίας .....	30
Εικόνα 17: Εντολή δημιουργίας αντικειμένου Video Capture .....	35
Εικόνα 18: Βρόχος Video Capture .....	36
Εικόνα 19: Εύρος τιμών ροζ χρώματος .....	38
Εικόνα 20: Εύρος τιμών λαχανί χρώματος .....	39
Εικόνα 21: Περιγράμματα αντικειμένων .....	40
Εικόνα 22 : Ορόσημο χεριός στο MediaPipe [33].....	41
Εικόνα 23: Αναγνώριση χειρονομίας .....	44
Εικόνα 24: Απεικόνιση των συνολικών χειρονομιών .....	44
Εικόνα 25: Μενού διεπαφής παιχνιδιού .....	45

Εικόνα 26:Μενού διεπαφής παιχνιδιού εντός του παιχνιδιού .....	45
Εικόνα 27:Περιβάλλον χρήστη στη Unity .....	47
Εικόνα 28:Box collider παίχτη .....	49
Εικόνα 29: Ήρωες του παιχνιδιού pink player στα αριστερά & green player στα δεξιά	50
Εικόνα 30: Ήρωας mixamo (Πηγή : <a href="https://mixamo.com">https://mixamo.com</a> ).....	50
Εικόνα 31: Δεξί & αριστερό χτύπημα (Πηγή : <a href="https://mixamo.com">https://mixamo.com</a> ).....	51
Εικόνα 32: Γήπεδο τένις παιχνιδιού.....	51
Εικόνα 33: Οθόνη νικητή 1ου set παιχνιδιού .....	52
Εικόνα 34: Τελικοί πόντοι παιχνιδιού .....	53
Εικόνα 35: Σκηνή χωρίς ή με φωτισμό.....	53
Εικόνα 36: Διάγραμμα ροής παιχνιδιού.....	55
Εικόνα 37: Διάγραμμα κίνησης χαρακτήρα εξαρτώμενο από την κίνηση του αντικειμένου.....	56
Εικόνα 38: Δεδομένα αλγορίθμων .....	58
Εικόνα 39: Προσθήκη ιδιοτήτων μέσω script για επικοινωνία με UDP.....	59
Εικόνα 40: Play game .....	60
Εικόνα 41: Γράφημα απαντήσεων 1ης ερώτησης του ερωτηματολογίου .....	61
Εικόνα 42: Γράφημα απαντήσεων 2ης ερώτησης του ερωτηματολογίου .....	62
Εικόνα 43: Γράφημα απαντήσεων 3ης ερώτησης του ερωτηματολογίου .....	62
Εικόνα 44: Γράφημα απαντήσεων 4ης ερώτησης του ερωτηματολογίου .....	63

# Κεφάλαιο 1

## 1.1 Εισαγωγή

Η αναγνώριση χειρονομιών στα βιντεοπαιχνίδια είναι ένα θέμα αυξανόμενου ενδιαφέροντος στη βιομηχανία τυχερών παιχνιδιών την τελευταία δεκαετία [9]. Η τεχνολογία αναγνώρισης χειρονομιών είναι η δυνατότητα αναγνώρισης και ελέγχου κινήσεων των χεριών και των δακτύλων από ηλεκτρονικούς υπολογιστές. Είναι μια μορφή της υπολογιστικής όρασης που επιτρέπει στις μηχανές να ερμηνεύουν τις κινήσεις των ανθρώπινων χεριών και δακτύλων. Αναλύοντας και κατανοώντας τις χειρονομίες, οι υπολογιστές μπορούν να τις προσομοιώσουν σε ένα περιβάλλον παιχνιδιού. Έχει χρησιμοποιηθεί σε μια μεγάλη ποικιλία εφαρμογών, συμπεριλαμβανομένων των παιχνιδιών εικονικής πραγματικότητας (virtual reality - VR) [22] και επαυξημένης πραγματικότητας (augmented reality - AR), παιχνιδιών ελεγχόμενης κίνησης και σε κονσόλες βιντεοπαιχνιδιών ελεγχόμενες από χειρονομίες.

Η ανάπτυξη της τεχνολογίας αναγνώρισης χειρονομιών στα βιντεοπαιχνίδια έχει κάνει πολλά βήματα τα τελευταία χρόνια. Αρχικά, αυτή η τεχνολογία περιοριζόταν στους χρήστες που παρείχαν προ-προγραμματισμένες εντολές με τα δάχτυλα. Ωστόσο, με την ανάπτυξη αυτής της τεχνολογίας, κατέστη δυνατό να παρέχεται στους χρήστες πιο ακριβής αναγνώριση των χειρονομιών τους [5]. Για παράδειγμα, οι πρόσφατες εξελίξεις στους αλγόριθμους μηχανικής μάθησης και την χρήση έργων νευρωνικών δικτύων επέτρεψαν στους υπολογιστές να αναγνωρίζουν και να παρακολουθούν τις κινήσεις των χεριών και των δακτύλων [13].

Επιπλέον, η προσθήκη καμερών και ιχνηλατών έδωσε τη δυνατότητα στις κονσόλες βιντεοπαιχνιδιών και στους υπολογιστές να κατανοήσουν τις κινήσεις των χεριών των παικτών, επιτρέποντας πιο διαδραστικές εμπειρίες παιχνιδιού [8]. Για παράδειγμα, ορισμένα βιντεοπαιχνίδια έχουν εφαρμόσει τεχνολογία αναγνώρισης χειρονομιών για



προχωρημένους χρήστες που επιτρέπει στους παίκτες να εκτελούν κινήσεις με τα χέρια και τα δάχτυλά τους όπως θα έκαναν στον πραγματικό κόσμο [2].

Ο πιο συνηθισμένος τύπος αναγνώρισης χειρονομιών που χρησιμοποιείται στα βιντεοπαιχνίδια είναι γνωστός ως σκελετική παρακολούθηση [10]. Αυτή η μέθοδος χρησιμοποιεί κάμερες και αισθητήρες βάθους για να καταγράψει μια τρισδιάστατη εικόνα του σώματος του παίκτη, συμπεριλαμβανομένων των χεριών και των δακτύλων του. Στη συνέχεια, το λογισμικό αναλύει τις εικόνες για να ανιχνεύσει και να ερμηνεύσει τις χειρονομίες [6].

Η αναγνώριση χειρονομιών μπορεί να χρησιμοποιηθεί για τον έλεγχο μιας ποικιλίας στοιχείων του παιχνιδιού, όπως την κίνηση χαρακτήρων, το χειρισμό αντικειμένων και την πλοήγηση μενού του βιντεοπαιχνιδιού. Μπορεί επίσης να χρησιμοποιηθεί η ανίχνευση χειρονομιών για την εκτέλεση ειδικών ενεργειών, όπως ξόρκια σε παιχνίδια ρόλων.

Η παραπάνω τεχνολογία έχει τη δυναμική να φέρει επανάσταση στον τρόπο με τον οποίο οι άνθρωποι χειρίζονται τα βιντεοπαιχνίδια. Με την κίνηση ενός αντικειμένου ή ενός ανθρώπινου χεριού, σε βιντεοπαιχνίδια αγώνων, οι παίκτες μπορούν να ενσαρκώσουν περισσότερο το ρόλο του ήρωα και να αισθανθούν ότι πραγματικά βρίσκονται σε ένα γήπεδο αγώνων. Μπορεί να κάνει το παιχνίδι πιο καθηλωτικό και ελκυστικό [18]. Ωστόσο, είναι ακόμα στα αρχικά του στάδια και υπάρχουν πολλές προκλήσεις που πρέπει να αντιμετωπιστούν προτού μπορέσει να υιοθετηθεί ευρέως [23].

Συνοπτικά οι προσπάθειες αναγνώρισης χειρονομιών και αντικειμένων είναι πρώιμες στην ιστορία των βιντεοπαιχνιδιών, αλλά έχει δει μια ταχεία εξέλιξη με την έλευση των νέων τεχνολογιών. Οι πρόσφατες εξελίξεις στην εικονική πραγματικότητα και στην υπολογιστική όραση έχει οδηγήσει σε αναζωπύρωση του θέματος.

## **1.2 Κίνητρο και επιδιωκόμενα αποτελέσματα**

Στον συνεχώς εξελισσόμενο κόσμο των παιχνιδιών αγώνων, η είσοδος που βασίζεται σε ελεγκτές γίνεται γρήγορα απαρχαιωμένη. Οι καινοτομίες που στηρίζονται στην αναγνώριση χειρονομιών και αντικειμένων προσφέρουν στους παίκτες έναν νέο τρόπο ελέγχου του παιχνιδιού. Αυτή η εργασία στοχεύει να διερευνήσει τον τρόπο χειρονομίας απευθείας στις κινήσεις των χεριών του παίκτη ή την αναγνώριση αντικειμένων με βάση

το χρωματικό χαρακτηριστικό του, παρά το πάτημα ενός κουμπιού από ένα χειριστήριο μιας παιχνιδιομηχανής. Τα συστήματα αναγνώρισης μπορούν να εφαρμοστούν προκειμένου να αντικαταστήσουν τα παραδοσιακά συστήματα εισόδου που βασίζονται σε ελεγκτές που χρησιμοποιούνται σήμερα.

Μια από τις πιο συναρπαστικές εξελίξεις στην παιγνιοποίηση είναι η εμφάνιση της αναγνώρισης χειρονομιών ως μέσου ελέγχου των εμπειριών παιχνιδιού. Αυτή η τεχνολογία έχει τη δυνατότητα να επαναφέρει τον τρόπο που παίζουμε παιχνίδια, επιτρέποντας μια πιο διαισθητική και καθηλωτική εμπειρία.

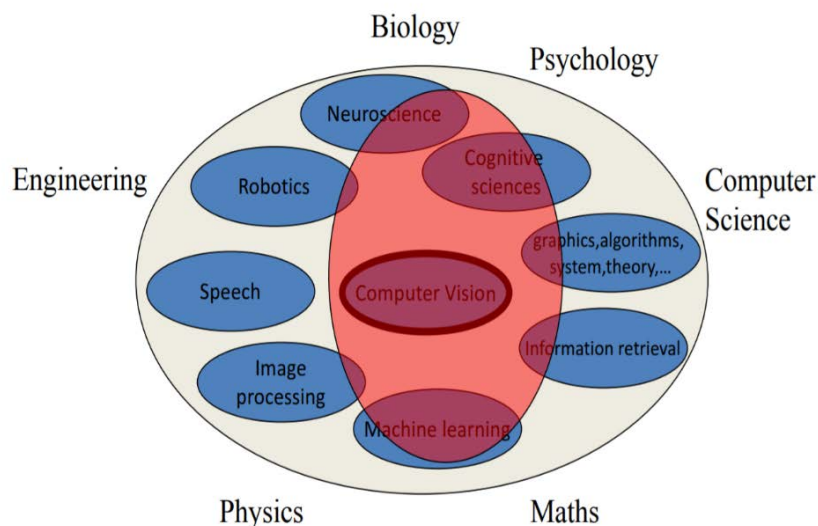
Ένα από τα κύρια κίνητρα για την αντικατάσταση των ελεγκτών με την αναγνώριση χειρονομιών ή την αναγνώριση αντικειμένου στα παιχνίδια αγώνων είναι η δυνατότητα για αυξημένο ρεαλισμό. Οι κινήσεις με το χέρι είναι πιο φυσικές και πιο διαισθητικός ο τρόπος για να αλληλοεπιδρούν οι άνθρωποι με το περιβάλλον τους ώστε να αισθάνονται περισσότερο σαν να βρίσκονται πραγματικά πίσω από το τιμόνι ενός αυτοκινήτου ή σε ένα αγώνα παιχνιδιού τένις κρατώντας πραγματικά μια ρακέτα. Αυτό συμβαίνει επειδή χαρτογραφούνται οι χειρονομίες ή οι κινήσεις ενός αντικειμένου και μεταφράζοντας τες κατά την διάρκεια του παιχνιδιού να εκτελούνται εντολές αντίστοιχες με το πάτημα ενός κουμπιού ή της κίνησης ενός μοχλού. Θα μπορούσε σε ένα αγώνα τένις το χτύπημα της μπάλας από μια ρακέτα να γίνεται απευθείας από τις σωματικές κινήσεις του παίκτη. Η αναγνώριση χειρονομιών ή η αναγνώριση αντικειμένου έχει ήδη εφαρμοστεί σε ορισμένες πλατφόρμες παιχνιδιών, όπως το Microsoft Kinect, σε αυτές τις πλατφόρμες, οι παίκτες χρησιμοποιούν τα χέρια τους για να ελέγχουν χαρακτήρες ή αντικείμενα στην οθόνη.

## Κεφάλαιο 2 Υπολογιστική όραση και επικοινωνία ανθρώπου - μηχανής

### 2.1 Τι είναι η υπολογιστική όραση;

Η υπολογιστική όραση μπορεί να οριστεί ως ένα επιστημονικό πεδίο που εξάγει πληροφορίες από ψηφιακές εικόνες. Ο τύπος των πληροφοριών που λαμβάνονται από μια εικόνα μπορεί να ποικίλλει από αναγνώριση, μετρήσεις χώρου για πλοήγηση ή εφαρμογές επαυξημένης πραγματικότητας. Ένας άλλος τρόπος για να ορίσετε την υπολογιστική όραση είναι μέσω των εφαρμογών του. Η υπολογιστική όραση δημιουργεί αλγόριθμους που μπορούν να κατανοήσουν το περιεχόμενο των εικόνων και τα συμπεράσματα να χρησιμοποιηθούν σε άλλες εφαρμογές.

Η υπολογιστική όραση συγκεντρώνει ένα μεγάλο σύνολο επιστημονικών κλάδων που μπορεί να εφαρμοστεί, όπως την Βιολογία, την Φυσιολογία, την Μηχανική, την Ρομποτική, την Φυσική, τα Μαθηματικά (βλ. εικόνα 1).



Εικόνα 1: Η όραση υπολογιστή στη διασταύρωση πολλαπλών επιστημονικών πεδίων (Πηγή: Stanford CS 231n)

Η νευροεπιστήμη μπορεί να βοηθήσει την υπολογιστική όραση καταλαβαίνοντας πρώτα την ανθρώπινη όραση, όπως θα δούμε στη συνέχεια. Η υπολογιστική όραση μπορεί να θεωρηθεί ως μέρος της επιστήμης των υπολογιστών και η θεωρία αλγορίθμων ή η μηχανική μάθηση είναι απαραίτητα για την ανάπτυξη αλγορίθμων υπολογιστική όρασης.

Η υπολογιστική όραση δεν έχει λυθεί εδώ και 50 χρόνια και εξακολουθεί να είναι ένα πολύ δύσκολο πρόβλημα. Είναι κάτι που εμείς οι άνθρωποι κάνουμε ασυνείδητα, αλλά είναι πραγματικά δύσκολο για τους υπολογιστές.

Σήμερα είναι πιο εύκολο να δημιουργήσουμε ένα τρισδιάστατο μοντέλο ενός αντικειμένου με ακρίβεια χιλιοστού παρά να δημιουργήσουμε έναν αλγόριθμο που αναγνωρίζει καρέκλες. Η αναγνώριση αντικειμένων εξακολουθεί να είναι ένα πολύ δύσκολο πρόβλημα, αν και πλησιάζουμε την ανθρώπινη ακρίβεια.

Στην υπολογιστική όραση υπάρχει τεράστιο χάσμα μεταξύ των εικονοστοιχείων και του νοήματος μιας εικόνας που βλέπει ο άνθρωπος. Αυτό που βλέπει ο υπολογιστής σε μια ψηφιακή εικόνα με διάσταση  $200 \times 200$  εικονοστοιχεία είναι ένα σύνολο 120.000 τιμών. Ο δρόμος από αυτούς τους αριθμούς προς τις ουσιαστικές πληροφορίες είναι πολύ δύσκολος. Αναμφισβήτητα, ο οπτικός φλοιός του ανθρώπινου εγκεφάλου λύνει ένα πρόβλημα τόσο δύσκολο: την κατανόηση εικόνων που προβάλλονται στον αμφιβληστροειδή μας και μετατρέπονται σε σήματα νευρώνων. Η επόμενη ενότητα θα δείξει πώς η μελέτη του εγκεφάλου μπορεί να βοηθήσει στην όραση του υπολογιστή.

## **2.2 Κατανόηση της ανθρώπινης όρασης**

Μια πρώτη ιδέα για την επίλυση της υπολογιστική όρασης είναι η κατανόηση του τρόπου λειτουργίας της ανθρώπινης όρασης και η μεταφορά αυτής της γνώσης στους υπολογιστές. Είτε πρόκειται για υπολογιστή είτε για ζώο, η όραση αποτελείται από δύο στοιχεία. Πρώτον, μια συσκευή αντίληψης καταγράφει όσο το δυνατόν περισσότερες λεπτομέρειες από μια εικόνα. Το μάτι θα συλλάβει το φως που περνά μέσα από την ίριδα και θα το προβάλλει στον αμφιβληστροειδή, όπου εξειδικευμένα κύτταρα θα μεταδώσουν πληροφορίες στον εγκέφαλο μέσω των νευρώνων. Μια κάμερα καταγράφει εικόνες με παρόμοιο τρόπο και μεταδίδει εικονοστοιχεία στον υπολογιστή. Σε αυτό το κομμάτι, οι κάμερες είναι καλύτερες από τους ανθρώπους καθώς μπορούν να δουν πιο μακριά ή με μεγαλύτερη ακρίβεια. Δεύτερον, η συσκευή διερμηνείας πρέπει να επεξεργάζεται τις πληροφορίες και να εξάγει νόημα από αυτές. Ο άνθρωπος

εγκέφαλος το λύνει αυτό σε πολλαπλά βήματα σε διαφορετικές περιοχές του εγκεφάλου. Η υπολογιστική όραση εξακολουθεί να υστερεί σε σχέση με την ανθρώπινη απόδοση σε αυτόν τον τομέα.

## **2.3 Εξαγωγή πληροφοριών από εικόνες**

Μπορούμε να χωρίσουμε τις πληροφορίες που λαμβάνονται από εικόνες στην υπολογιστική όραση σε δύο κατηγορίες: μετρήσεις και σημασιολογικές πληροφορίες.

### **2.3.1 Η όραση ως συσκευή μέτρησης**

Τα ρομπότ που πλοηγούνται σε άγνωστη τοποθεσία πρέπει να μπορούν να σαρώνουν το περιβάλλον τους για να υπολογίσουν την καλύτερη διαδρομή. Χρησιμοποιώντας την υπολογιστική όραση, μπορούμε να μετρήσουμε τον χώρο γύρω από ένα ρομπότ και να δημιουργήσουμε έναν χάρτη του περιβάλλοντός του. Οι στερεοσκοπικές κάμερες δίνουν πληροφορίες βάθους, όπως τα δύο μάτια μας, μέσω του τριγωνισμού. Η στερεοσκοπική όραση είναι ένα μεγάλο πεδίο της υπολογιστικής όρασης και υπάρχει πολλή έρευνα που επιδιώκει να δημιουργήσει έναν ακριβή χάρτη βάθους δεδομένων στερεοσκοπικών εικόνων. Εάν αυξήσουμε τον αριθμό των σημείων προβολής για να καλύψουμε όλες τις πλευρές ενός αντικειμένου, μπορούμε να δημιουργήσουμε μια τρισδιάστατη επιφάνεια που αντιπροσωπεύει το αντικείμενο ακόμη πιο προκλητική ιδέα θα μπορούσε να είναι η ανακατασκευή του τρισδιάστατου μοντέλου ενός μνημείου μέσω όλων των αποτελεσμάτων μιας αναζήτησης εικόνων στο google για αυτό το μνημείο[24]. Υπάρχει επίσης έρευνα για τη σύλληψη, όπου η υπολογιστική όραση μπορεί να βοηθήσει στην κατανόηση της τρισδιάστατης γεωμετρίας ενός αντικειμένου για να βοηθήσει ένα ρομπότ να καταλάβει. Μέσω της κάμερας του ρομπότ, μπορούσαμε να αναγνωρίσουμε και να βρούμε τη λαβή του αντικειμένου και να συμπεράνουμε το σχήμα του, για να μπορέσει στη συνέχεια το ρομπότ να βρει μια καλή θέση λαβής [25] .

### **2.3.2 Πηγή σημασιολογικών πληροφοριών**

Πέρα από τις πληροφορίες μέτρησης, μια εικόνα περιέχει μια πολύ πυκνή ποσότητα σημασιολογικών πληροφοριών. Μπορούμε να βάλουμε ετικέτες σε αντικείμενα σε μια εικόνα, να βάλουμε ετικέτες σε ολόκληρη τη σκηνή, να αναγνωρίσουμε ανθρώπους, να αναγνωρίσουμε πράξεις, χειρονομίες, πρόσωπα. Οι ιατρικές εικόνες περιέχουν επίσης πολλές σημασιολογικές πληροφορίες. Η όραση υπολογιστή μπορεί να είναι χρήσιμη για

μια διάγνωση που βασίζεται σε εικόνες δερματικών κυττάρων, για παράδειγμα, για να αποφασίσει εάν είναι καρκινικά ή όχι.

## 2.4 Εφαρμογές υπολογιστικής όρασης

Οι κάμερες είναι παντού και ο αριθμός των εικόνων που ανεβαίνουν στο διαδίκτυο αυξάνεται με γεωμετρική πρόοδο. Έχουμε εικόνες στο Instagram, βίντεο στο YouTube, ροές από κάμερες ασφαλείας, ιατρικές και επιστημονικές εικόνες. Η υπολογιστική όραση είναι απαραίτητη γιατί πρέπει να ταξινομήσουμε αυτές τις εικόνες και να επιτρέψουμε στους υπολογιστές να κατανοήσουν το περιεχόμενό τους. Ακολουθεί μια μη εξαντλητική λίστα εφαρμογών της όρασης υπολογιστών.

- *Τρισδιάστατη αστική μοντελοποίηση*

Η λήψη φωτογραφιών με drone πάνω από μια πόλη μπορεί να χρησιμοποιηθεί για την απόδοση ενός τρισδιάστατου μοντέλου της πόλης. Η υπολογιστική όραση χρησιμοποιείται για να συνδυάσει όλες τις φωτογραφίες σε ένα ενιαίο τρισδιάστατο μοντέλο.

- *Αναγνώριση σκηνής*

Είναι δυνατή η αναγνώριση της τοποθεσίας όπου τραβήχτηκε μια φωτογραφία. Για παράδειγμα, μια φωτογραφία ενός ορόσημου μπορεί να συγκριθεί με δεσεκατομμύρια φωτογραφίες. Στη συνέχεια, μπορούμε να προσδιορίσουμε την καλύτερη αντιστοιχία και να συμπεράνουμε τη θέση της φωτογραφίας.

- *Ανίχνευση προσώπου*

Η ανίχνευση προσώπου χρησιμοποιείται εδώ και πολλά χρόνια σε κάμερες για τη λήψη καλύτερων φωτογραφιών και την εστίαση στα πρόσωπα. Η ανίχνευση χαμόγελου μπορεί να επιτρέψει σε μια κάμερα να τραβάει φωτογραφίες αυτόματα όταν το θέμα χαμογελάει. Η αναγνώριση προσώπου είναι πιο δύσκολη από την ανίχνευση προσώπου, αλλά με την κλίμακα των σημερινών δεδομένων, εταιρείες όπως το Facebook μπορούν να έχουν πολύ καλές επιδόσεις. Τέλος, μπορούμε επίσης να χρησιμοποιήσουμε την όραση υπολογιστή για βιομετρικά στοιχεία, χρησιμοποιώντας την αναγνώριση σχεδίου ίριδας ή δακτυλικά αποτυπώματα.

- *Οπτική αναγνώριση χαρακτήρων*

Μία από τις παλαιότερες επιτυχημένες εφαρμογές της υπολογιστικής όρασης είναι η αναγνώριση χαρακτήρων και αριθμών. Αυτό μπορεί να χρησιμοποιηθεί για την ανάγνωση ταχυδρομικών κωδίκων ή πινακίδων κυκλοφορίας.

- *Οπτική αναζήτηση για κινητά*

Με την υπολογιστική όραση, μπορούμε να κάνουμε μια αναζήτηση στη Google χρησιμοποιώντας μια εικόνα ως ερώτημα.

- *Αυτοοδηγούμενα αυτοκίνητα*

Η αυτόνομη οδήγηση είναι μια από τις πιο καυτές εφαρμογές της υπολογιστικής όρασης. Εταιρείες όπως η Tesla, η Google ή η General Motors ανταγωνίζονται για να είναι οι πρώτες που θα κατασκευάσουν ένα πλήρως αυτόνομο αυτοκίνητο. Αυτόματο ταμείο της Amazon Go είναι ένα νέο είδος καταστήματος που δεν έχει ταμείο. Με την υπολογιστική όραση, οι αλγόριθμοι εντοπίζουν ακριβώς ποια προϊόντα παίρνετε και σας χρεώνουν καθώς βγαίνετε από το κατάστημα.

- *Αλληλεπίδραση που βασίζεται στο Vision*

Το Kinect της Microsoft καταγράφει την κίνηση σε πραγματικό χρόνο και επιτρέπει στους παίκτες να αλληλεπιδρούν απευθείας με ένα παιχνίδι μέσω κινήσεων.

- *Η επαυξημένη πραγματικότητα AR*

Η επαυξημένη πραγματικότητα AR είναι επίσης ένα πολύ καυτό πεδίο αυτή τη στιγμή και πολλές εταιρείες ανταγωνίζονται για να παρέχουν την καλύτερη πλατφόρμα AR για κινητά.

- *Η εικονική πραγματικότητα VR*

Η εικονική πραγματικότητα (VR) χρησιμοποιεί παρόμοιες τεχνικές της υπολογιστικής όρασης με την επαυξημένη πραγματικότητα (AR). Ο αλγόριθμος πρέπει να γνωρίζει τη θέση ενός χρήστη και τις θέσεις όλων των αντικειμένων γύρω. Καθώς ο χρήστης μετακινείται, όλα πρέπει να ενημερώνονται με ρεαλιστικό και ομαλό τρόπο.

## Κεφάλαιο 3 Κατάτμηση εικόνας

### 3.1 Κατάτμηση εικόνας

Η κατάτμηση της εικόνας είναι η διαδικασία διαχωρισμού μιας εικόνας σε πολλαπλά τμήματα. Η κατάτμηση χρησιμοποιείται συνήθως για τον εντοπισμό αντικειμένων και ορίων σε εικόνες. Η τμηματοποίηση χωρίζει μια εικόνα σε διακριτές περιοχές που περιέχουν εικονοστοιχεία με παρόμοια γνωρίσματα. Για να είναι ουσιαστικές και χρήσιμες για ανάλυση και ερμηνεία εικόνων, οι περιοχές θα πρέπει να σχετίζονται με έντονα απ' εικονιζόμενα αντικείμενα ή χαρακτηριστικά ενδιαφέροντος. Η ουσιαστική τμηματοποίηση είναι το πρώτο βήμα από την επεξεργασία εικόνας χαμηλού επιπέδου που μετατρέπει μια εικόνα σε κλίμακα του γκρι ή μία έγχρωμη ή περισσότερες άλλες εικόνες σε περιγραφή εικόνας υψηλού επιπέδου όσον αφορά τα χαρακτηριστικά, τα αντικείμενα και τις σκηνές. Η επιτυχία της ανάλυσης εικόνας εξαρτάται από την αξιοπιστία της τμηματοποίησης.

Υπάρχουν δύο μορφές τμηματοποίησης εικόνας:

- *Τοπική τμηματοποίηση*: Αφορά μια συγκεκριμένη περιοχή της εικόνας.
- *Καθολική τμηματοποίηση* : Αφορά την τμηματοποίηση ολόκληρης της εικόνας.

Τρόποι προσεγγίσεων τμηματοποίησης εικόνας:

- *Ανίχνευση ασυνέχειας* : Αυτή είναι μια μέθοδος τμηματοποίησης μιας εικόνας σε περιοχές με βάση την ασυνέχεια. Εδώ μπαίνει η ανίχνευση ακμών. Η ασυνέχεια στις ακμές που δημιουργούνται λόγω της έντασης αναγνωρίζεται και χρησιμοποιείται για τον καθορισμό ορίων περιοχής. Παραδείγματα: Φιλτράρισμα ιστογράμματος και ανίχνευση περιγράμματος.
- *Ανίχνευση ομοιότητας* : Μια μέθοδος τμηματοποίησης μιας εικόνας σε ενότητες με βάση την ομοιότητα. Ο καθορισμός κατωφλίου, η επέκταση της περιοχής και



ο διαχωρισμός και η συγχώνευση περιοχών περιλαμβάνονται όλα σε αυτήν τη μεθοδολογία. Χωρίζοντας την εικόνα σε τμήματα με συγκρίσιμο αριθμό εικονοστοιχείου. Με βάση καθιερωμένα κριτήρια, χωρίζουν την εικόνα σε μια ομάδα συμπλεγμάτων με συγκρίσιμα χαρακτηριστικά. Παράδειγμα: ανίχνευση και ταξινόμηση χρωμάτων.

- *Προσέγγιση νευρωνικών δικτύων:* Για τον στόχο της λήψης αποφάσεων, οι αλγόριθμοι τμηματοποίησης που βασίζονται σε νευρωνικά δίκτυα αναπαράγουν τις τεχνικές εκμάθησης του ανθρώπινου εγκεφάλου. Αυτή η προσέγγιση χρησιμοποιείται ευρέως για την κατάτμηση ιατρικών εικόνων και τον διαχωρισμό τους από το φόντο. Ένα νευρωνικό δίκτυο αποτελείται από έναν τεράστιο αριθμό συνδεδεμένων κόμβων, ο καθένας με το δικό του βάρος.

### **3.2 Βασικές τεχνικές τμηματοποίησης εικόνας**

#### *Τμηματοποίηση εικόνας βάση κατωφλίου*

Η τμηματοποίηση κατωφλίου εικόνας είναι μια απλή μορφή τμηματοποίησης εικόνας. Είναι ένας τρόπος δημιουργίας μιας δυαδικής μονόχρωμης ή πολύχρωμης εικόνας που βασίζεται στον ορισμό μιας τιμής κατωφλίου βασιζόμενη στην ένταση των εικονοστοιχείων της αρχικής εικόνας.

Σε αυτή τη διαδικασία ορίου, θα εξετάσουμε το ιστόγραμμα έντασης όλων των εικονοστοιχείων στην εικόνα. Στη συνέχεια θα ορίσουμε ένα όριο για να χωρίσουμε την εικόνα σε τμήματα. Για παράδειγμα, θα εξετάσουμε όλα τα εικονοστοιχεία εικόνας που οι τιμές έντασης της φωτεινότητας τους κυμαίνονται από 0 έως 255, θα ορίσουμε ένα όριο 60. Έτσι, όλα τα εικονοστοιχεία με τιμές μικρότερες ή ίσες με 60 θα παρέχεται με τιμή 0 (μαύρο) και όλα τα εικονοστοιχεία με τιμή μεγαλύτερη από 60 θα παρέχεται τιμή 255 (λευκό). Σκεφτείτε μια εικόνα με φόντο και ένα αντικείμενο, μπορούμε να χωρίσουμε μια εικόνα σε περιοχές με βάση την ένταση του αντικειμένου και του φόντου. Αλλά αυτό το όριο πρέπει να ρυθμιστεί τέλεια για να τμηματοποιηθεί μια εικόνα σε ένα αντικείμενο και ένα φόντο.

#### *Τμηματοποίηση εικόνας βάση ακμών*

Η τμηματοποίηση βάση ακμών βασίζεται σε άκρες που βρίσκονται σε μια εικόνα χρησιμοποιώντας διάφορους τελεστές ανίχνευσης ακμών. Αυτές οι ακμές επισημαίνουν θέσεις ασυνέχειας στην εικόνα σε επίπεδα γκρι, χρώμα, υφή κ.λπ. Όταν μετακινούμαστε

από τη μια περιοχή στην άλλη, το επίπεδο του γκρι μπορεί να αλλάξει. Έτσι, αν μπορούμε να βρούμε αυτή την ασυνέχεια, μπορούμε να βρούμε και αυτή την ακμή. Είναι διαθέσιμοι διάφοροι τελεστές ανίχνευσης ακμών, αλλά η εικόνα που προκύπτει είναι ένα ενδιάμεσο αποτέλεσμα τμηματοποίησης και δεν πρέπει να συγχέεται με την τελική τμηματοποιημένη εικόνα. Πρέπει να εκτελέσουμε περαιτέρω επεξεργασία στην εικόνα για να την τμηματοποιήσουμε.

Τα πρόσθετα βήματα περιλαμβάνουν το συνδυασμό τμημάτων άκρων που λαμβάνονται σε ένα τμήμα προκειμένου να μειωθεί ο αριθμός των τμημάτων αντί για τα κομμάτια μικρών περιγραμμάτων που μπορεί να εμποδίσουν τη διαδικασία πλήρωσης περιοχής. Αυτό γίνεται για να ληφθεί ένα απρόσκοπτο περίγραμμα του αντικειμένου. Ο στόχος της τμηματοποίησης άκρων είναι να ληφθεί ένα ενδιάμεσο αποτέλεσμα τμηματοποίησης στο οποίο μπορούμε να εφαρμόσουμε τμηματοποίηση βάσει περιοχής ή οποιονδήποτε άλλο τύπο τμηματοποίησης για να λάβουμε την τελική τμηματοποιημένη εικόνα.

Τα άκρα συνήθως συνδέονται με το μέγεθος και την κατεύθυνση. Μερικοί ανιχνευτές ακμών δίνουν και κατευθύνσεις και μέγεθος. Μπορούμε να χρησιμοποιήσουμε διάφορους ανιχνευτές ακμών, όπως τον τελεστή ακμών Sobel, τον τελεστή ακμών Prewitt και τον τελεστή ακμών του Robert, κ.λπ.

#### *Τμηματοποίηση εικόνας βάση περιοχής*

Μια περιοχή μπορεί να ταξινομηθεί ως μια ομάδα συνδεδεμένων εικονοστοιχείων που εμφανίζουν παρόμοιες ιδιότητες. Η ομοιότητα των εικονοστοιχείων μπορεί να είναι ως προς την ένταση και το χρώμα, κ.λπ. Σε αυτόν τον τύπο τμηματοποίησης, υπάρχουν ορισμένοι προκαθορισμένοι κανόνες στους οποίους πρέπει να τηρούν τα εικονοστοιχεία προκειμένου να ταξινομηθούν σε παρόμοιες περιοχές εικονοστοιχείων. Οι μέθοδοι τμηματοποίησης βάση περιοχής προτιμώνται έναντι των μεθόδων τμηματοποίησης που βασίζονται σε άκρα, εάν πρόκειται για θορυβώδη εικόνα.

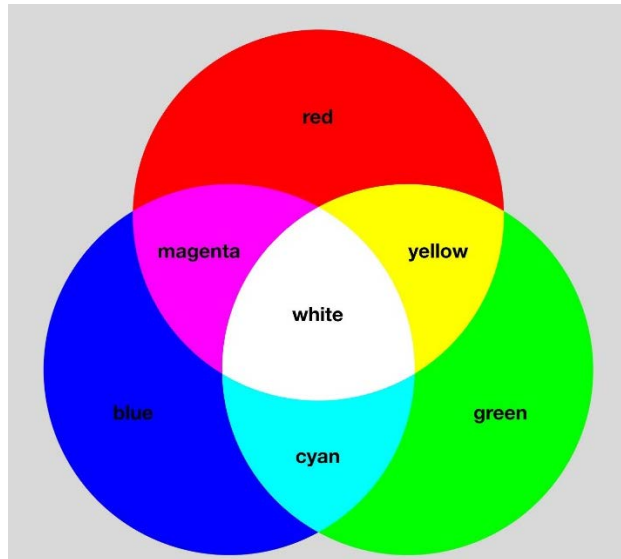
## Κεφάλαιο 4 Ανίχνευση αντικειμένου με βάση το χρώμα

### 4.1 Εισαγωγή

Τα χρώματα μπορούν να μετρηθούν με διαφορετικούς τρόπους. Αναμφίβολα, η εντύπωση ενός ανθρώπου για τα χρώματα είναι μια υποκειμενική διαδικασία κατά την οποία το μυαλό αντιδρά στα ερεθίσματα που δημιουργούνται όταν πλησιάζει το φως στο μάτι. Βασικά, διαφορετικά άτομα βλέπουν το ίδιο φωτισμένο αντικείμενο ή πηγή φωτός με διάφορους τρόπους.

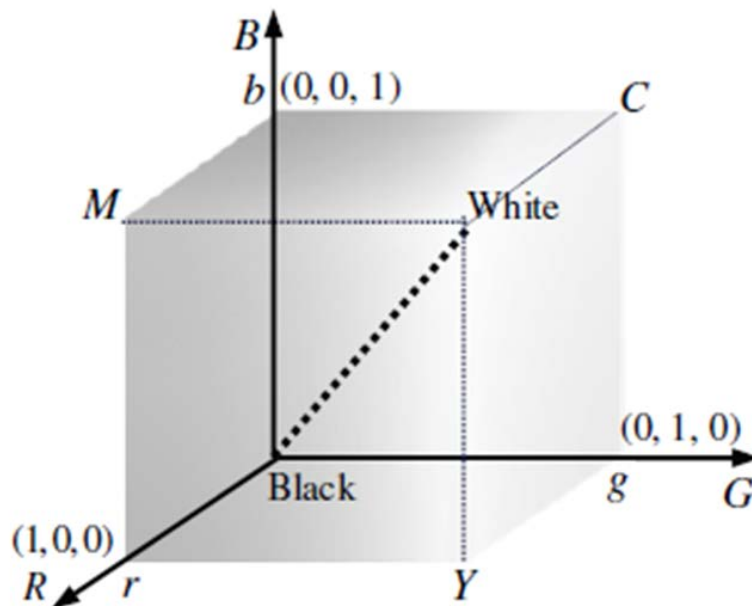
### 4.2 Μοντέλο χρώματος RGB

Ένα χρωματικό μοντέλο είναι ένα σύστημα για τη δημιουργία μιας πλήρους γκάμα χρωμάτων από ένα μικρό σύνολο βασικών χρωμάτων. Ένα χρωματικό μοντέλο ονομάζεται επιπλέον σύστημα χρωμάτων ή χρωματικός χώρος. Ο λόγος για το χρωματικό μοντέλο είναι να ενθαρρύνει τον προσδιορισμό του χρώματος σε κάποιο πρότυπο, σε μεγάλο βαθμό αναγνωρισμένο σε RGB (αναφέρεται σε μοντέλο χρώματος Κόκκινο, Πράσινο και Μπλε). Αυτό το χρωματικό μοντέλο παρουσιάστηκε για πρώτη φορά από τους Young *et al* και Helmholtz *et al* στη Θεωρία της Τριχρωματικής Χρωματικής όρασης και χρησιμοποιώντας το χρωματικό τρίγωνο του Maxwell *et al.* στις αρχές του 19ου αιώνα. Στο χρωματικό μοντέλο RGB, κάθε χρώμα εμφανίζεται στα κύρια φασματικά του τμήματα κόκκινο, πράσινο και μπλε. Το μοντέλο λαμβάνει υπόψη το καρτεσιανό σύστημα συντεταγμένων. Είναι ένα πρόσθετο χρωματικό μοντέλο, με την αιτιολογία ότι τα κόκκινα, πράσινα και μπλε φώτα περιλαμβάνονται σε διαφορετικό μείγμα για να αναδημιουργήσουν ένα ευρύ φάσμα χρωμάτων (βλ. εικόνα 2).



Εικόνα 2: Έγχρωμο μοντέλο RGB (Πηγή :Encyclopaedia Britannica .Inc)

Κατά την αναπαράσταση αυτού του χρωματικού μοντέλου, ορίζεται ένας κύβος στους άξονες των R, G και B (βλ. εικόνα 3).



Εικόνα 3: Αναπαράσταση χρωματικού μοντέλου RGB σε κύβο (Πηγή: 2D Computer Vision: Principles, Algorithms and Applications Yu-Jin Zhang, Tsinghua University, China.)

Το λευκό χρώμα εμφανίζεται όταν καθένα από τα κύρια χρώματα είναι παρόν στο M, όπου το M δείχνει τη μέγιστη ένταση του φωτός, έστω  $M = 255$ . Ο κύριος διαγώνιος άξονας που ενώνει τις ασπρόμαυρες γωνίες καθορίζει την ένταση.

$$I(R, G, B) = R + G + B$$

Όλα τα σημεία σε ένα επίπεδο κάθετο στον γκρι άξονα του κύβου χρώματος έχουν την ίδια ένταση. Το επίπεδο που διέρχεται από τον κύβο χρώματος στα σημεία  $R = G = B = M$  είναι ένα τέτοιο επίπεδο. Αυτό το επίπεδο σχηματίζει ένα ισόπλευρο τρίγωνο που είναι το τυπικό τρίγωνο χρωματικότητας RGB.

Σε αυτόν τον κύβο τα βασικά χρώματα (Κόκκινο, Πράσινο και Μπλε) βρίσκονται σε τρεις γωνίες, δηλαδή κόκκινο στον άξονα  $x$ , πράσινο στον άξονα  $y$  και μπλε στον άξονα  $z$ , τα δευτερεύοντα χρώματα κυανό, ματζέντα και κίτρινο είναι σε άλλες τρεις γωνίες, δηλαδή το κυανό είναι στο διαγώνια απέναντι γωνία του κόκκινου, το ματζέντα βρίσκεται στη διαγώνια απέναντι γωνία του πράσινου, το κίτρινο είναι στη διαγώνια απέναντι γωνία του μπλε. Το μαύρο είναι στην αρχή και το λευκό είναι στη γωνία πιο απομακρυσμένη από την προέλευση. Τα διακριτικά χρώματα σε αυτό το μοντέλο εστιάζονται στο σχήμα ή στο εσωτερικό του και χαρακτηρίζονται από διανύσματα που εκτείνονται από την αρχή.

Το κύριο κίνητρο πίσω από το χρωματικό μοντέλο RGB είναι η ανίχνευση, η αναπαράσταση, καθώς και η παρουσίαση εικόνων σε ηλεκτρονικές συσκευές, για παράδειγμα, τηλεοράσεις, υπολογιστές, κ.λπ. Ωστόσο, χρησιμοποιείται επιπλέον ως μέρος της παραδοσιακής φωτογραφίας.

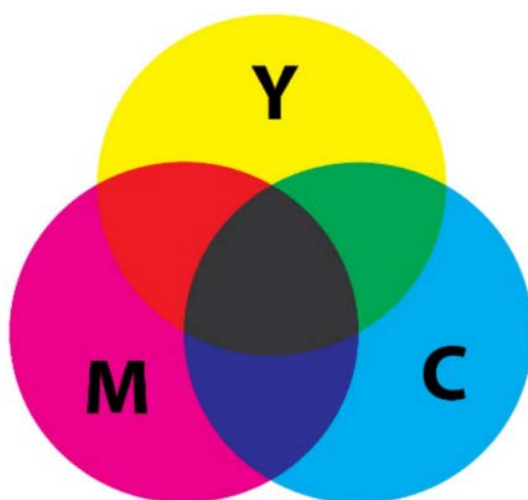
#### **4.2.1 Αριθμητικές παραστάσεις**

Αυτές οι εκτάσεις μπορούν να αξιολογηθούν με διάφορους τρόπους. Όπως, με οποιαδήποτε κλασματική τιμή μεταξύ 0 και 1. Αυτή η απεικόνιση χρησιμοποιείται στη θεωρητική ανάλυση και σε διατάξεις που χρησιμοποιούν απεικονίσεις κινητής υποδιαστολής. Επίσης, κάθε τιμή τμήματος χρώματος μπορεί επίσης να χρησιμοποιηθεί ως ποσοστιαία τιμή, που κυμαίνεται μεταξύ 0% και 100%. Ενώ οι τιμές των στοιχείων χρησιμοποιούνται συχνά ως ακέραιοι αριθμοί που κυμαίνονται από 0 έως 255, ειδικά σε υπολογιστές, το εύρος που μπορεί να προσφέρει ένα byte είναι «μόνο 8 bit». Δεκαδικό ή δεκαεξαδικό. Στο χρωματικό μοντέλο RGB, το χρώμα απεικονίζεται συχνά προσδιορίζοντας την ποσότητα καθενός από το κόκκινο, το πράσινο και το μπλε που ενσωματώνεται. Μια τριπλέτα RGB (R, G, B) χρησιμοποιείται για να εκφράσει το χρώμα, κάθε τμήμα του οποίου μπορεί να μετατοπιστεί από μια καθορισμένη ελάχιστη τιμή (συνήθως λαμβάνεται ως μηδέν) σε μια καθορισμένη μέγιστη τιμή. Στην

περίπτωση που όλα τα μέρη είναι στο μηδέν, το αποτέλεσμα είναι σκοτεινό, εάν όλα τα τμήματα έχουν την υψηλότερη τιμή, το λευκό έρχεται ως έξοδος.

### 4.3 Μοντέλο χρώματος CMY

Το CMY χρωματικό μοντέλο αναφέρεται στα χρώματα κυανό, ματζέντα και κίτρινο (βλ. εικόνα 4) .Αυτό εφευρέθηκε από τον Blou *et al.* που τελικά βοήθησε στην ανάπτυξη συστημάτων εκτύπωσης που περιλαμβάνουν συστήματα εκτύπωσης τριών χρωμάτων (CMY) και τεσσάρων χρωμάτων (CMYK).



Εικόνα 4: Χρωματικό μοντέλο CMY (Πηγή :Encyclopaedia Britannica .Inc)

Αυτό το χρωματικό μοντέλο χρησιμοποιείται σε διάφορες έγχρωμες εκτυπώσεις. Επί του παρόντος, τα πλαίσια εκτύπωσης χρησιμοποιούν ένα σύστημα που συνδυάζει τέσσερα χρώματα, συμπεριλαμβανομένου του μαύρου και κάποιου επιπλέον χρώματος, όπως κυανό, ματζέντα και κίτρινο. Αυτό το είδος μοντέλου ονομάζεται CMYK, όπου το K αντιπροσωπεύει το «κλειδί» που δείχνει μαύρο χρώμα. Οι αριθμοί χρησιμοποιούνται συχνά για να το αντιπροσωπεύσουν αυτό . Για παράδειγμα, διαφορετικοί συμβολισμοί RGB για το πιο φωτεινό κορεσμένο κόκκινο μπορούν να χαρακτηριστούν ως: Μοντέλο χρώματος CMYK .Το μαύρο χρησιμοποιείται ως μέρος της επέκτασης στο CMY για τους ακόλουθους λόγους: Το μαύρο μελάνι είναι λιγότερο ακριβό σε σύγκριση με αυτό με την ανάμειξη C, M και Y για να γίνει μαύρο. Το κείμενο που αποτυπώνεται σε μαύρο χρώμα έχει λεπτές λεπτομέρειες που θα θολώνονταν αν εκτυπωθεί με τρία διαφορετικά χρώματα. Μια ατελής απόχρωση του μαύρου παράγεται από το μείγμα των C, M και Y.

Το μοντέλο CMYK λειτουργεί καλύπτοντας εν μέρει ή συνολικά το χρώμα σε ένα πιο ανοιχτό, συνήθως λευκό φόντο. Το μελάνι μειώνει το φως που με κάποιο τρόπο θα αντανakλούσε το χρώμα. Ένα τέτοιο μοντέλο ονομάζεται αφαιρετικό επειδή τα μελάνια «αφαιρούν» τη φωτεινότητα από το λευκό.

Το χρωματικό μοντέλο CMY μπορεί να προκύψει από το χρωματικό μοντέλο RGB, με την παρακάτω μαθηματική πράξη:

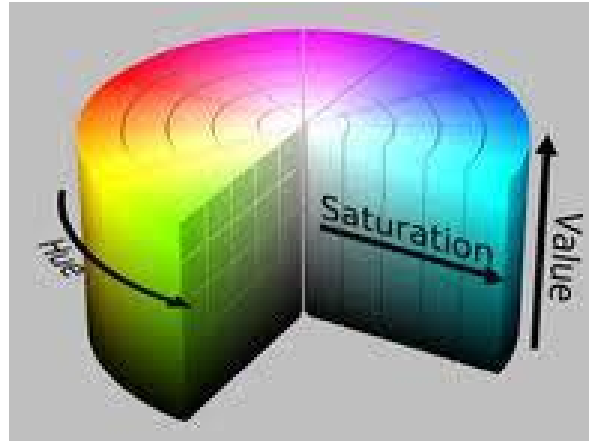
$$\begin{bmatrix} C \\ M \\ Y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

Αν αφαιρέσουμε από το λευκό της τιμές του κόκκινου, πράσινου, μπλε θα προκύψουν οι τιμές του χρωματικού μοντέλου CMY.

#### 4.4 HSV (απόχρωση, κορεσμός και τιμή)

Το HSV αναφέρεται στην απόχρωση, τον κορεσμό και την τιμή. Το χρωματικό μοντέλο HSV απεικονίστηκε από τον Ray *et al.*(1978). Το HSV είναι ένα κυλινδρικό χρωματικό μοντέλο που αντιστοιχίζει τα βασικά χρώματα RGB σε διαστάσεις που είναι ευκολότερο να κατανοήσουν οι άνθρωποι. Η απόχρωση, ο κορεσμός και η τιμή είναι οι κύριες ιδιότητες χρώματος που μας επιτρέπουν να διακρίνουμε μεταξύ διαφορετικών χρωμάτων. Η αποτελεσματική χρήση του χρώματος είναι ένα από τα πιο απαραίτητα στοιχεία στη φωτογραφία, καθώς το χρώμα μπορεί να τραβήξει το βλέμμα του θεατή στη σύνθεσή σας και να επηρεάσει τη διάθεση και τον συναισθηματικό αντίκτυπο της φωτογραφίας σας.

Είναι σημαντικό να σημειωθεί ότι οι τρεις διαστάσεις του χρωματικού μοντέλου HSV αλληλοεξαρτώνται. Εάν η διάσταση τιμής ενός χρώματος οριστεί στο 0%, η ποσότητα της απόχρωσης και του κορεσμού δεν έχει σημασία, καθώς το χρώμα θα είναι μαύρο. Ομοίως, εάν ο κορεσμός ενός χρώματος οριστεί στο 0%, η απόχρωση δεν έχει σημασία, καθώς δεν χρησιμοποιείται χρώμα. Επειδή η διάσταση απόχρωσης είναι κυκλική, το χρωματικό μοντέλο HSV απεικονίζεται καλύτερα ως κύλινδρος. Αυτό φαίνεται στην παρακάτω εικόνα 5, όπου όλα τα πιθανά μείγματα χρωμάτων αντιπροσωπεύονται εντός των ορίων του κυλίνδρου.



Εικόνα 5: Χρωματικό μοντέλο HSV (Πηγή : SharkD, CC BY-SA 3.0 via Wikimedia Commons

[https://en.wikipedia.org/wiki/HSL\\_and\\_HSV#/media/File:HSV\\_color\\_solid\\_cylinder\\_saturation\\_gray.png](https://en.wikipedia.org/wiki/HSL_and_HSV#/media/File:HSV_color_solid_cylinder_saturation_gray.png))

Το Hue καθορίζει τη γωνία του χρώματος στον κύκλο χρώματος RGB. Μια απόχρωση  $0^\circ$  έχει ως αποτέλεσμα κόκκινο,  $120^\circ$  σε πράσινο και  $240^\circ$  σε μπλε. Η απόχρωση είναι το χρωματικό τμήμα του μοντέλου, που εκφράζεται ως αριθμός από 0 έως 360 μοίρες:

- Το κόκκινο πέφτει μεταξύ 0 και 60 μοιρών.
- Το κίτρινο πέφτει μεταξύ 61 και 120 μοιρών.
- Το πράσινο πέφτει μεταξύ 121 και 180 μοιρών.
- Το κυανό πέφτει μεταξύ 181 και 240 μοιρών.
- Το μπλε πέφτει μεταξύ 241 και 300 μοιρών.
- Το ματζέντα πέφτει μεταξύ 301 και 360 μοιρών.

Ο κορεσμός ελέγχει την ποσότητα του χρώματος που χρησιμοποιείται. Ένα χρώμα με 100% κορεσμό θα είναι το πιο καθαρό δυνατό χρώμα, ενώ ο κορεσμός 0% αποδίδει κλίμακα του γκρι. Ο κορεσμός περιγράφει την ποσότητα του γκρι σε ένα συγκεκριμένο χρώμα, από 0 έως 100 %. Η μείωση αυτού του στοιχείου προς το μηδέν εισάγει περισσότερο γκρι και δημιουργεί ένα ξεθωριασμένο αποτέλεσμα. Μερικές φορές, ο κορεσμός εμφανίζεται ως ένα εύρος από 0 έως 1, όπου το 0 είναι γκρι και το 1 είναι ένα βασικό χρώμα.

Η τιμή ελέγχει τη φωτεινότητα του χρώματος. Ένα χρώμα με 0% φωτεινότητα είναι καθαρό μαύρο ενώ ένα χρώμα με 100% φωτεινότητα δεν έχει μαύρο .



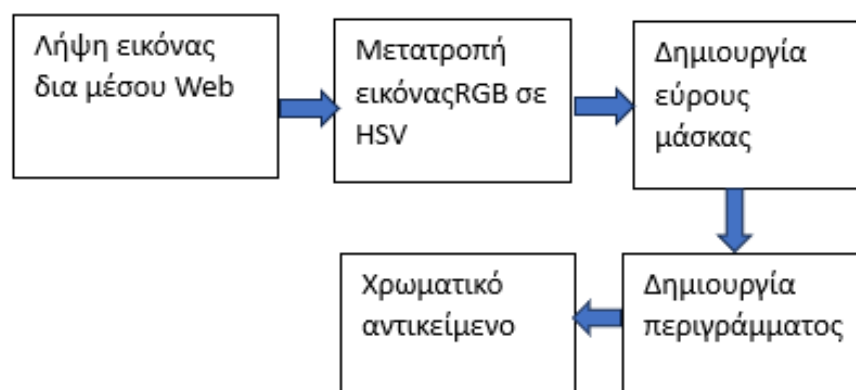
Ο χρωματικός τροχός HSV χρησιμοποιείται για την επιλογή του επιθυμητού χρώματος. Η απόχρωση αντιπροσωπεύεται από τον κύκλο στον τροχό. Ένα ξεχωριστό τρίγωνο χρησιμοποιείται για να αναπαραστήσει τον κορεσμό και την τιμή. Ο οριζόντιος άξονας του τριγώνου δείχνει τιμή και ο κάθετος άξονας αντιπροσωπεύει κορεσμό. Όταν χρειάζεστε ένα συγκεκριμένο χρώμα για την εικόνα σας, πρώτα πρέπει να επιλέξετε ένα χρώμα από την απόχρωση (την κυκλική περιοχή) και στη συνέχεια από την κατακόρυφη γωνία του τριγώνου μπορείτε να επιλέξετε τον επιθυμητό κορεσμό. Για φωτεινότητα, μπορείτε να επιλέξετε την επιθυμητή τιμή από την οριζόντια γωνία του τριγώνου.

#### 4.5 Μέθοδος ανίχνευσης χρώματος

Η παρακάτω αρχιτεκτονική κάνει τη διαδικασία πιο αποτελεσματική με βάση αρχές και ιδιότητες που σχετίζονται μεταξύ τους.

Η υπολογιστική όραση δεν έχει ακόμη φτάσει σε ένα επίπεδο στο οποίο να μπορεί να χρησιμοποιηθεί άμεσα για την επίλυση προβλημάτων της καθημερινής ζωής, καθώς βρίσκεται ακόμη στη φάση των εξελίξεων.

Η παρακάτω εικόνα 6 μας δείχνει την ικανότητα του έργου. Αποτελείται από καλά καθορισμένα στάδια ακολουθίας, προσαρμοσμένα από τον πηγαίο κώδικα. Αξιοποιεί τις πλούσιες δυνατότητες της τεχνολογίας όπως η βιβλιοθήκη OpenCV, στην Python.



Εικόνα 6:Μπλοκ διάγραμμα μεθόδου ανίχνευσης αντικειμένου με βάση το χρώμα

Η λειτουργία του μπλοκ διαγράμματος (βλ. εικόνα 6) :

- Η κάμερα web λειτουργεί ως αισθητήρας ανίχνευσης που θα κάνει λήψη φωτογραφιών ή βίντεο των αντικειμένων.
- Μετατροπή της εικόνας από το χρωματικό μοντέλο RGB σε HSV, με την τεχνική αυτή πετυχαίνουμε καλύτερα αποτελέσματα στην αναγνώριση χρωμάτων.
- Δημιουργία μάσκας που εστιάζει στην επιθυμητή χρωματική περιοχή ,ορίζοντας τις μέγιστες και ελάχιστες τιμές της.
- Δημιουργία περιγράμματος της επιθυμητής χρωματικής περιοχής.
- Απεικόνιση αποτελέσματος στην οθόνη.

Με τον προσδιορισμό του αντικειμένου με βάση το επιθυμητό χρώμα το σύστημα μπορεί τώρα να προχωρήσει στο τελευταίο βήμα, το οποίο είναι να παράγει μια δράση εντός του παιχνιδιού.

Αξίζει να σημειωθεί ότι το σύστημα που περιγράφεται δεν είναι ειδικά προσαρμοσμένο για να διερευνήσει τη χρήση αναγνώρισης αντικειμένων ως αντικατάσταση των χειριστηρίων σε παιχνίδια αγώνων. Ωστόσο, αυτή η εφαρμογή αναγνώρισης αντικειμένων με βάση το χρώμα έχει επιλεγεί καθώς παρέχει μια σαφή και συγκεκριμένη περίπτωση χρήσης της τεχνολογίας και επιτρέπει την εξέταση των πιθανών πλεονεκτημάτων της ως μέσου εισαγωγής σε παιχνίδια αγώνων. Το περιβάλλον ενός παιχνιδιού αγώνων είναι ένα τέλειο σενάριο για να δοκιμάσουμε την συγκεκριμένη τεχνική αναγνώριση αντικειμένων, καθώς οι είσοδοι σε αυτόν τον τύπο διεπαφής μας δίνουν μεγάλο βαθμό ελευθερίας.

## Κεφάλαιο 5 Αναγνώριση χειρονομιών

### 5.1 Εισαγωγή

Η αναγνώριση χειρονομιών είναι η διαδικασία αναγνώρισης και ερμηνείας των κινήσεων και των χειρονομιών των χεριών και των δακτύλων ενός ατόμου για τον έλεγχο ενός συστήματος υπολογιστή ή άλλης τεχνολογίας. Είναι μια μορφή μη λεκτικής επικοινωνίας που περιλαμβάνει τη χρήση των κινήσεων των χεριών και των δακτύλων για τον έλεγχο και την αλληλεπίδραση με μια συσκευή. Η τεχνολογία αναγνώρισης χειρονομιών γίνεται όλο και πιο δημοφιλής τα τελευταία χρόνια, καθώς παρέχει έναν βολικό και διαισθητικό τρόπο για τους χρήστες να αλληλεπιδρούν με υπολογιστές και άλλες συσκευές.

Οι εφαρμογές της αναγνώρισης χειρονομιών είναι πολλές. Έχει χρησιμοποιηθεί για τον έλεγχο βιντεοπαιχνιδιών, τον έλεγχο των ρομποτικών βραχιόνων, ακόμη και για την αλληλεπίδραση με περιβάλλοντα εικονικής πραγματικότητας. Χρησιμοποιείται επίσης όλο και περισσότερο για τη διερμηνεία της νοηματικής γλώσσας, επιτρέποντας στους χρήστες να επικοινωνούν με υπολογιστές χωρίς την ανάγκη λεκτικής επικοινωνίας. Τέλος, η αναγνώριση χειρονομιών διερευνάται ως ένας τρόπος αλληλεπίδρασης με οχήματα και άλλες μηχανές, επιτρέποντας πιο φυσικές, διαισθητικές αλληλεπιδράσεις όπως ερευνήθηκε από τον Bar *et al.* [15].

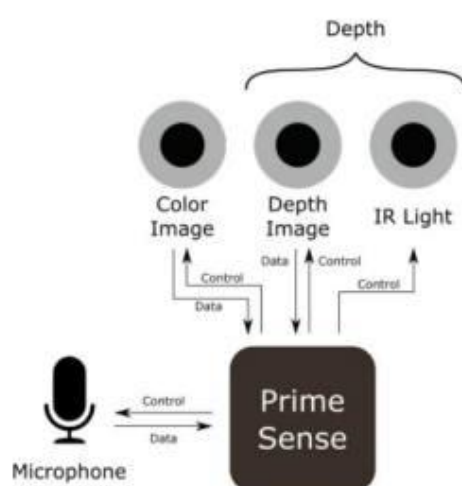
Η εφαρμογή αυτής της τεχνολογίας είναι μια διαδικασία πολλαπλών βημάτων που περιλαμβάνει μια σειρά από δραστηριότητες. Το αρχικό βήμα σε αυτή τη διαδικασία είναι η καταγραφή των χειρονομιών του χρήστη μέσω εξειδικευμένου υλικού, όπως κάμερες ή αισθητήρες. Μόλις καταγραφούν οι χειρονομίες, πρέπει να υποβληθούν σε επεξεργασία για να προσδιοριστούν με ακρίβεια οι συγκεκριμένες χειρονομίες που εκτελούνται. Αυτό το βήμα είναι κρίσιμο καθώς θέτει τα θεμέλια για τα επόμενα βήματα [20]. Μετά την αναγνώριση των χειρονομιών, το επόμενο βήμα είναι να τις αντιστοιχίσουν σε συγκεκριμένες ενέργειες που μπορεί να εκτελέσει ο υπολογιστής.

Αυτό μπορεί να είναι το άνοιγμα ενός αρχείου, το κλείσιμο ενός παραθύρου ή η προσομοίωση ενός χειριστηρίου παιχνιδιού .

## 5.2 Απόκτηση και προεπεξεργασία εικόνας

Αρχικά, είναι σημαντικό να αναγνωρίσουμε ότι η είσοδος για την ανίχνευση των χειρονομιών μπορεί να προέρχεται από διάφορες πηγές, όχι μόνο από μια κάμερα. Αυτή η ευελιξία στις πηγές εισόδου επιτρέπει μεγαλύτερη ευελιξία στους τύπους συστημάτων και εφαρμογών που μπορούν να χρησιμοποιήσουν την τεχνολογία αναγνώρισης χειρονομιών. Μια πρώιμη μέθοδος για τη λήψη δεδομένων για την ανίχνευση χειρονομιών ήταν η χρήση γαντιών. Αυτή η προσέγγιση απαιτούσε τη χρήση μιας εξειδικευμένης συσκευής υλικού που θα συλλέγει δεδομένα από τις αρθρώσεις στο χέρι. Ενώ αυτή η μέθοδος είχε τα πλεονεκτήματά της, όπως η ταχύτερη συλλογή δεδομένων και η ανάγκη για λιγότερα δεδομένα εισόδου λόγω της μεταφοράς των πληροφοριών απευθείας από τα δάχτυλα, είχε επίσης σημαντικά μειονεκτήματα [17]. Ένα από τα κύρια μειονεκτήματα ήταν το κόστος και το βάρος των γαντιών, καθιστώντας τα μη πρακτικά για φυσική αλληλεπίδραση και χρήση [3] σύμφωνα με την χειρονομία που εντοπίστηκε. Ο στόχος αυτού του βήματος είναι να μεταφράσει την αναγνωρισμένη χειρονομία σε μια συγκεκριμένη ενέργεια που μπορεί να εκτελεστεί μέσα στο παιχνίδι. Μια εφαρμογή μπορεί να είναι η προσομοίωση των παραδοσιακών εισόδων ελεγκτή. Αυτή η προσέγγιση είναι ιδιαίτερα χρήσιμη καθώς επιτρέπει τη συμβατότητα με ένα ευρύ φάσμα παιχνιδιών, καθώς τα περισσότερα παιχνίδια έχουν σχεδιαστεί για να ανταποκρίνονται σε εισόδους από παραδοσιακούς ελεγκτές. Μια εναλλακτική προσέγγιση λήψης δεδομένων από την ανίχνευση χειρονομιών, είναι η χρήση της υπολογιστικής όρασης. Αυτή η μέθοδος περιλαμβάνει τη λήψη εικόνων ή βίντεο από μια κάμερα ή παρόμοια συσκευή και την ανάλυση της εισόδου για τον εντοπισμό χειρονομιών. Ένα από τα κύρια πλεονεκτήματα αυτής της προσέγγισης είναι ότι επιτρέπει πιο φυσικές διεπαφές. Σε αντίθεση με την προσέγγιση που βασίζεται στα γάντια, η οποία απαιτούσε εξειδικευμένο υλικό, αυτή η μέθοδος μπορεί να εφαρμοστεί χρησιμοποιώντας κοινώς διαθέσιμες συσκευές όπως smartphone, tablet ή web κάμερες . Επιπλέον, αυτή η μέθοδος δεν περιορίζεται στις χειρονομίες και μπορεί να χρησιμοποιηθεί για την ανίχνευση άλλων κινήσεων του σώματος, όπως επίσης περιγράφεται από το άρθρο των Rui Ma *et al.* [12]. Επιπλέον, αυτή η προσέγγιση είναι πιο οικονομική και μπορεί να ενσωματωθεί σε ένα ευρύ φάσμα εφαρμογών, όπως παιχνίδια, εικονική πραγματικότητα και αλληλεπίδραση ανθρώπου-υπολογιστή.

Τέλος, υπάρχει η προσέγγιση με βάση το βάθος, χρησιμοποιώντας κάμερες βάθους [20]. Αυτές οι συσκευές λειτουργούν διαφορετικά από τις παραδοσιακές κάμερες, καθώς δεν καταγράφουν κανονικές εικόνες, αλλά καταγράφουν την απόσταση των σημείων στη σκηνή με αισθητήρες βάθους (βλ. εικόνα 7). Ένα πολύ γνωστό παράδειγμα αυτού του τύπου συσκευής είναι η Microsoft Kinect. Αρχικά αναπτύχθηκε για παιχνιοποίηση, τώρα έχει βρει άλλες χρήσεις πέρα από την παιχνιοποίηση, συμπεριλαμβανομένης της αναγνώρισης χειρονομιών.



Εικόνα 7: Σύστημα καταγραφής απόσταση των σημείων στη σκηνή με αισθητήρες βάθους (Πηγή: <http://www.ifixit.com/Teardown/Microsoft-Kinect-Teardown/4066/1> )

Μόλις καταγραφεί σήμα στην είσοδο της συσκευής, συχνά πραγματοποιείται προεπεξεργασία για την ανάλογη βελτίωση της ακρίβειας και της βέλτιστης απόδοσης του συστήματος [21]. Αυτό το βήμα, αν και δεν είναι απολύτως απαραίτητο, μπορεί να περιλαμβάνει εργασίες όπως αλλαγή μεγέθους και περικοπή εικόνων, κανονικοποίηση τιμών των εικονοστοιχείων και αφαίρεση θορύβου.

## 5.3 Τεχνικές αναγνώρισης χειρονομιών

### 5.3.1 Ανίχνευση ακμών

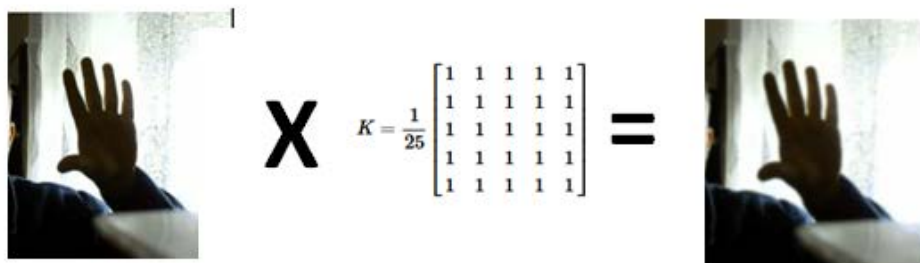
Ένα ιδανικό σύστημα παρακολούθησης χειρών μπορεί να ανιχνεύσει τα χέρια σε μια εικόνα χωρίς να χρειάζεται βαθμονόμηση.

Με τη χρήση του αλγορίθμου Canny edge detector [29], σε μία ψηφιακή εικόνα ή σε μία ροή εικόνων, μπορούμε να έχουμε ένα σύστημα παρακολούθησης των χειρών εντοπίζοντας τις περιοχές υψηλής αντίθεσης. Μέσω της χρήσης των διαβαθμίσεων εικόνας και του κατωφλίου υστέρησης, αυτός ο αλγόριθμος προσδιορίζει τις πιο εμφανείς γραμμές και τα περιγράμματα σε μια εικόνα. Ανάλογα με το σχήμα τους, συγκεκριμένα αντικείμενα μπορούν στη συνέχεια να αναγνωριστούν και να τμηματοποιηθούν.

Τα στάδια του αλγορίθμου Canny edge detector είναι τα εξής:

- Περιορισμός θορύβου

Εφαρμόζουμε στην αρχική εικόνα ένα φίλτρο Gaussian Filter, αυτό βοηθά στην εξομάλυνση της εικόνας και στη μείωση του αντίκτυπου των μικρών τυχαίων διακυμάνσεων στις εντάσεις των εικονοστοιχείων (βλ. εικόνα 8).



Εικόνα 8: Εφαρμογή Gaussian Filter

- Gradient calculation

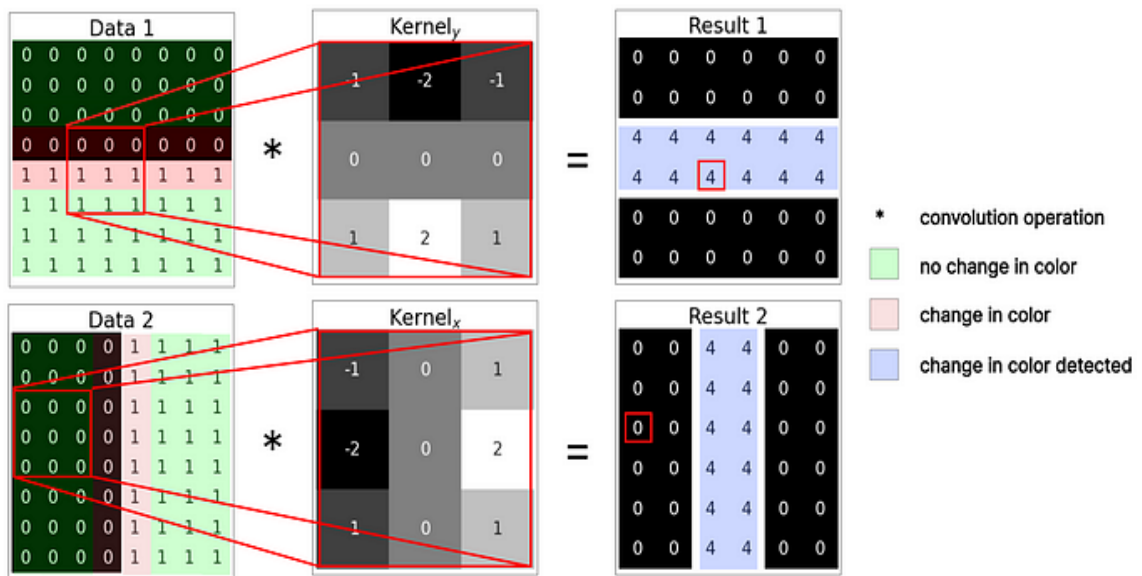
Σε αυτό το βήμα θα χρησιμοποιήσουμε ένα φίλτρο Sobel, το βήμα αυτό είναι στην πραγματικότητα η μέθοδος ανίχνευσης άκρων Sobel.

Η διαδικασία χρησιμοποιεί τα φίλτρα Sobel που είναι μια διακριτή προσέγγιση της παραγώγου, θα εκτελέσουμε μια λειτουργία συνέλιξης στην θολή εικόνα που δημιουργήσαμε παραπάνω (βλ. Εικόνα 8), έτσι ως αποτέλεσμα θα λάβουμε δύο διαβαθμίσεις ως προς τον άξονα X και Y, οι διαβαθμίσεις είναι διανύσματα που

δείχνουν προς την κατεύθυνση του μεγαλύτερου ρυθμού αλλαγής στην ένταση της εικόνας.

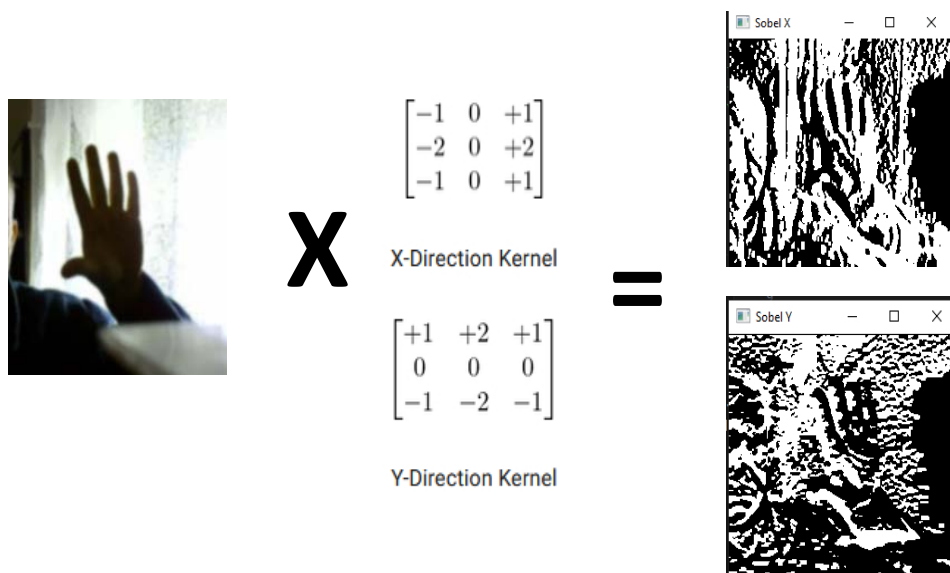
Χρησιμοποιώντας αυτές τις δύο διαβαθμίσεις, δημιουργούμε μια ενιαία συνολική κλίση, αν σχεδιάσουμε αυτή τη διαβάθμιση, παίρνουμε μια εικόνα που αποτελείται από τις άκρες της εικόνας.

Αν κάνουμε την πράξη συνέλιξης με το Sobel Filter X στην εικόνα (βλ. *Εικόνα 9*) θα πάρουμε έναν άλλο πίνακα που δείχνει την αλλαγή των χρωμάτων προς την κατεύθυνση x. Επίσης, η πράξη συνέλιξης του φίλτρου Sobel Y στην εικόνα (βλ. *Εικόνα 9*) έχει ως αποτέλεσμα έναν άλλο πίνακα που δείχνει την αλλαγή των χρωμάτων στην κατεύθυνση y.



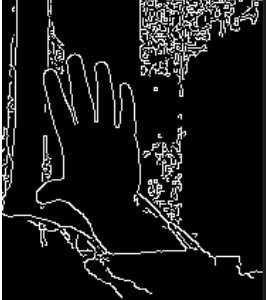
*Εικόνα 9: Εικονογράφηση του τρόπου με τον οποίο το φίλτρο Sobel δημιουργεί κλίσεις X και Y (Πηγή: <https://medium.com/@rohit-krishna/coding-canny-edge-detection-algorithm-from-scratch-in-python-232e1fdceac7>)*

Το οπτικό αποτέλεσμα μετά την πράξη συνέλιξης με το Sobel Filter X και Sobel Filter Y φαίνεται στην παρακάτω εικόνα 10.



Εικόνα 10: Εφαρμογή φίλτρων Sobel ως προς τους άξονες X & Y

Για να βρούμε τη συνολική αλλαγή, χρησιμοποιούμε το Πυθαγόρειο θεώρημα. Παίρνουμε τη Βάση ως κλίση X και το Υψός ως την κλίση Y, έτσι η Υποτείνουσα θα οδηγήσει στη συνολική αλλαγή (βλ. εικόνα 11).

$$G = \sqrt{G_x^2 + G_y^2} =$$


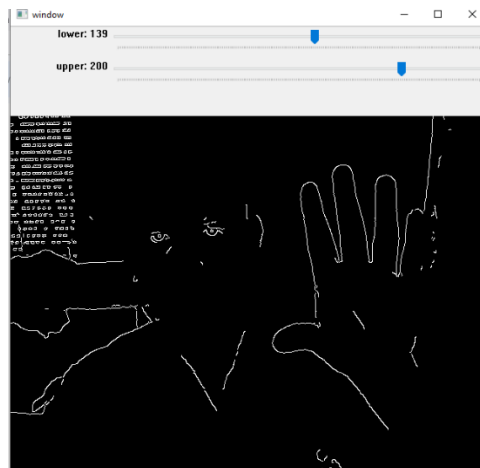
Εικόνα 11: Τελικό αποτέλεσμα του Canny Edge Detector

Δυστυχώς, αυτή η μέθοδος από μόνη της είναι ανεπαρκής για τον εντοπισμό των χεριών σε ένα πολύπλοκο φόντο. Με δεδομένη μια εικόνα, η ανίχνευση άκρων με τον αλγόριθμο Canny edge detector εξάγει μόνο τα εικονοστοιχεία με υψηλή τοπική



αντίθεση και δεν παρέχει καμία εικόνα για το περιβάλλον τους. Ως αποτέλεσμα, δεν υπάρχει καμία εγγύηση ότι κάθε περίγραμμα ανήκει μόνο σε ένα μόνο αντικείμενο. Ωστόσο, οι παραλλαγές των εικονοστοιχείων και κατ' επέκταση οι ακμές δημιουργούνται συχνά από τη διασταύρωση πολλών αντικειμένων. Ο προσδιορισμός των τμημάτων των περιγραμμάτων που σχετίζονται με τα χέρια είναι ένα δύσκολο πρόβλημα που δεν μπορεί να λυθεί με την ανίχνευση άκρων Canny edge detector.

Επιπλέον, σε εικόνες με πολλά αντικείμενα, το χέρι σπάνια αντιστοιχεί σε ένα μόνο συνδεδεμένο περίγραμμα. Αντίθετα, συχνά αποτελείται από πολλά μικρότερα τμήματα γραμμής. Παρά τις προσπάθειές μας να ενοποιήσουμε αυτά τα κομμάτια σε ένα συνεκτικό περίγραμμα, δεν μπορούσαμε να βρούμε μια μέθοδο για την αξιόπιστη εκτέλεση αυτής της εργασίας. Μορφολογικές επεμβάσεις [30] όπως η διαστολή μπορούν να κλείσουν μικρά κενά, αλλά συνδέουν επίσης το χέρι με άσχετα κοντινά περιγράμματα (βλ. εικόνα 12).



*Εικόνα 12: Ανίχνευση άκρων με τη χρήση αλγορίθμου Canny edge detector*

Ο αλγόριθμος Canny edge detector παρουσιάζει τους ίδιους περιορισμούς όταν ομαδοποιούνται ακμές. Ως αποτέλεσμα, καταλήξαμε τελικά στο συμπέρασμα ότι η ανίχνευση ακμών δεν ήταν μια αρκετά αξιόπιστη προσέγγιση για την παρακολούθηση χεριών.

### 5.3.2 Αναγνώριση με την τεχνική τμηματοποίησης χρώματος

Η τμηματοποίηση χρώματος είναι μια άλλη χρήσιμη τεχνική για την παρακολούθηση των χεριών. Με δεδομένη μια εικόνα, διατηρεί μόνο τα εικονοστοιχεία εντός ενός καθορισμένου εύρους χρωμάτων. Με άλλα λόγια, μπορούν να καθοριστούν ελάχιστες και μέγιστες τιμές χρώματος για την ανίχνευση του δέρματος και κατά συνέπεια των χεριών σε μια εικόνα.

Στην πράξη, η έρευνα [31] έδειξε ότι ο χρωματικός χώρος HSV είναι πιο κατάλληλος από τον χρωματικό χώρο RGB για αυτό το σκοπό.

Ωστόσο, το χρώμα του δέρματος μπορεί να ποικίλλει πολύ και η επιλογή στενού εύρους τιμών θα προκαλέσει την αποτυχία ανίχνευσης του δέρματος για ορισμένα άτομα. Από την άλλη πλευρά, άλλος μεγαλύτερος αριθμός ανεπιθύμητων αντικειμένων φόντου θα διατηρηθεί με ένα ευρύτερο φάσμα χρωμάτων.

Στην εικόνα 13 φαίνεται η εφαρμογή του αλγορίθμου τμηματοποίησης με βάση το χρώμα.



Εικόνα 13: Αφαίρεση φόντου, μάσκα εικόνας και τμηματοποίηση χρώματος

Στην συγκεκριμένη τεχνική πρώτα μετατρέπουμε την εικόνα στο χρωματικό μοντέλο HSV, ορίζουμε μια χρωματική περιοχή με ένα ανώτερο και ένα κατώτερο όριο με πρότυπο το ανιχνεύσιμο χρώμα του χεριού. Στο επόμενο βήμα δημιουργούμε μία μάσκα σύμφωνα με τα χρωματικά όρια, το αποτέλεσμα της μεθόδου φαίνεται στο πρώτο παράθυρο της εικόνας 13. Έτσι πετυχαίνουμε να διαχωρίσουμε την επιθυμητή χρωματική περιοχή και να εφαρμόσουμε ένα περίγραμμα πάνω της.

Παρατηρήθηκε ότι εάν οι συνθήκες φωτισμού αλλάξουν ή κάποιος άλλος χρήστης χειριστεί την οθόνη διεπαφής του παιχνιδιού τα αποτελέσματα εντοπισμού μπορεί να

μην είναι τα επιθυμητά. Έτσι καταλήξαμε ότι η παραπάνω τεχνική δεν ήταν μια αρκετά αξιόπιστη προσέγγιση για την παρακολούθηση χεριών.

### 5.3.3 Αναγνώριση χεριού μέσω του MediaPipe

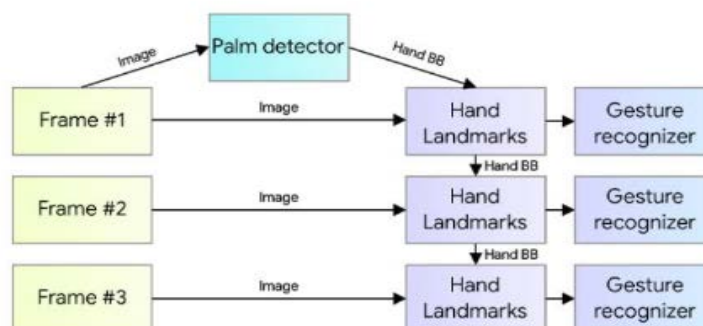
Το MediaPipe είναι ένα πλαίσιο λογισμικού ανοικτού κώδικα (open source framework) που προσφέρει καινοτόμες λύσεις σε προβλήματα Υπολογιστικής Όρασης και Μηχανικής Μάθησης. Ο σκοπός του MediaPipe είναι να καταστήσει σχετικά απλή τη χρήση σύγχρονων αλγορίθμων Βαθιάς Μάθησης με σύνθετες αρχιτεκτονικές.

Η χρήση του MediaPipe για την αναγνώριση χειρονομιών έχει ερευνηθεί από τον Zhang *et al.* [27] στο παρελθόν, χρησιμοποιώντας μια μοναδική κάμερα RGB για εφαρμογή AR/VR σε ένα σύστημα σε πραγματικό χρόνο που προβλέπει έναν σκελετό χεριού του ανθρώπου.

Η συγκεκριμένη τεχνική αποτελείται από τρία μοντέλα αναγνώρισης χειρονομιών ως εξής :

1. Ένα μοντέλο ανιχνευτή παλάμης επεξεργάζεται την εικόνα που τραβήχτηκε και γυρίζει την εικόνα με ένα προσανατολισμένο πλαίσιο οριοθέτησης του χεριού.
2. Ένα μοντέλο ορόσημο χειρός επεξεργάζεται την εικόνα περικομμένου πλαισίου οριοθέτησης και επιστρέφει τρισδιάστατα σημεία-κλειδιά με το χέρι.
3. Ένας αναγνωριστής χειρονομίας που ταξινομεί τα τρισδιάστατα σημεία-κλειδιά του χεριού και στη συνέχεια τα διαμορφώνει σε ένα διακριτό σύνολο χειρονομιών.

Στην παρακάτω εικόνα γίνεται αναπαράσταση της λειτουργίας του αναγνωριστή χεριού με τη βοήθεια του MediaPipe(βλ. εικόνα 14).



Εικόνα 14:Επισκόπηση αγωγού αντίληψης χεριού [32]

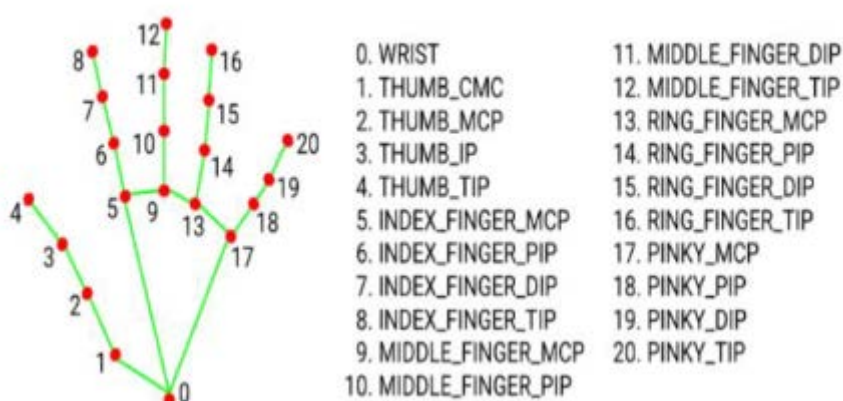
## 1. Μοντέλο ανιχνευτή παλάμης

Το πλαίσιο MediaPipe έχει ενσωματώσει τον αρχικό ανιχνευτή παλάμης εντοπισμού που ονομάζεται BlazePalm. Στη φάση αυτή του μοντέλου εφαρμόζουμε έναν αλγόριθμο SSD μέσω τετραγώνων όπου πετυχαίνετε η μοντελοποιησή της παλάμης . Στη συνέχεια, ένας κωδικοποιητής-αποκωδικοποιητής εξαγωγής χαρακτηριστικών παρόμοιο με το Feature Pyramid Networks (Lin et al.,2016) ,χρησιμοποιείται για μεγαλύτερη επίγνωση του πλαισίου σκηνής ακόμη και σε μικρότερα αντικείμενα. Τέλος, ελαχιστοποιούν την εστιακή απώλεια κατά τη διάρκεια της προπόνησης με υποστήριξη μεγάλου αριθμού αγκυρώσεων που προκύπτουν από τη διακύμανση υψηλής κλίμακας [28].

## 2. Ορόσημο χειρός

Επιτυγχάνει ακριβή εντοπισμό σημείων κλειδιού,21 βασικών σημείων με μια τρισδιάστατη συντεταγμένη χεριού-άρθρωσης που διεξάγεται εντός των ανιχνευόμενων περιοχών χεριού μέσω παλινδρόμησης που θα παράγει την πρόβλεψη συντεταγμένων απευθείας, η οποία είναι ένα μοντέλο του ορόσημου χεριού στο MediaPipe (βλ. εικόνα 15). Η συντεταγμένη αποτελείται από x, y και z όπου τα x και y κανονικοποιούνται σε [0,0, 1,0] κατά πλάτος και ύψος εικόνας, ενώ το z αναπαριστά το βάθος του ορόσημου. Το βάθος του ορόσημου συσχετίζεται με την θέση του καρπού. Όσο κλείνει το ορόσημο στην κάμερα, η τιμή γίνεται μικρότερη.

Τα 21 σημεία του χεριού απεικονίζονται στην παρακάτω εικόνα (βλ. εικόνα 15).



Εικόνα 15:Ορόσημο χειρός στο MediaPipe [33]Αναγνωριστής χεριών

### 3. Αναγνώριση χειρονομιών

Για την αναγνώριση των χειρονομιών του χεριού, γίνεται η εφαρμογή ενός απλού αλγορίθμου που είναι ο υπολογισμός των χειρονομιών με μια καθορισμένη συσσωρευμένη γωνία της κατάστασης της άρθρωσης ή των συνθήκων κάθε δακτύλου, όπως λυγισμένο δάχτυλο ή ίσιο δάχτυλο, και στη συνέχεια χαρτογραφείτε το σύνολο των καταστάσεων των δακτύλων που είχαμε πριν για να οριστεί ετικέτα προκαθορισμένων χειρονομιών όπως "Index Middle" (βλ. εικόνα 16).



Εικόνα 16: Αναγνώριση χειρονομίας

### 5.4 Χαρτογράφηση με χειρονομίες

Στην Αλληλεπίδραση ανθρώπου-υπολογιστή (HCI) μέσω της αναγνώρισης χειρονομιών, η επιλογή των χειρονομιών είναι μια κρίσιμη πτυχή για το σχεδιασμό ενός αποτελεσματικού λεξιλογίου χειρονομιών. Τα μοναδικά χαρακτηριστικά και το εύρος κίνησης των χειρονομιών πρέπει να λαμβάνονται υπόψη κατά την επιλογή αυτή, καθώς ο στόχος της αλληλεπίδρασης ανθρώπου-υπολογιστή (HCI) είναι να δημιουργήσει μια διεπαφή που επιτρέπει στους χρήστες να ελέγχουν εργασίες υπολογιστή μέσω ενός συνόλου εντολών με τη μορφή χειρονομιών.

Αυτό απαιτεί ενδελεχή κατανόηση των πιθανών διαθέσιμων χειρονομιών και τη δυνατότητα επιλογής εκείνων που είναι πιο εύχρηστες και κατανοητές.

Συνολικά, η επιλογή του λεξιλογίου (σε αυτήν την περίπτωση οι χειρονομίες) είναι μια ουσιαστική πτυχή της αλληλεπίδρασης ανθρώπου-υπολογιστή (HCI), καθώς παρέχει τη βάση για μια φυσική και διαισθητική διεπαφή [14]. Ένα καλά σχεδιασμένο λεξιλόγιο χειρονομιών μπορεί να οδηγήσει σε μεγαλύτερη αποτελεσματικότητα και ικανοποίηση για τον χρήστη και ως εκ τούτου, είναι ένα κρίσιμο βήμα στην πορεία για την ανάπτυξη ενός επιτυχημένου συστήματος αλληλεπίδρασης ανθρώπου-υπολογιστή (HCI).

## **5.5 Συμπεράσματα**

Συνοπτικά, αναγνωρίζοντας τις χειρονομίες στο πεδίο της αλληλεπίδρασης ανθρώπου-υπολογιστή τα μοναδικά χαρακτηριστικά και το εύρος κίνησης των χειρονομιών πρέπει να λαμβάνονται υπόψη κατά την επιλογή αυτή, καθώς ο στόχος του HCI είναι να δημιουργήσει μια διεπαφή που επιτρέπει στους χρήστες να ελέγχουν εργασίες στον υπολογιστή μέσω ενός συνόλου εντολών με τη μορφή χειρονομιών(HCI), ώστε το αποτέλεσμα να είναι μια πρόκληση.

## Κεφάλαιο 6 Υλοποίηση εφαρμογής

### 6.1 Εισαγωγή

Η εργασία αυτή όπως προαναφέρθηκε αποτελείται από δύο μέρη, την ανίχνευση των ανθρώπινων χειρονομιών, αντικειμένων και από την σχεδίαση ενός βιντεοπαιχνιδιού. Επίσης θα δημιουργηθεί και ένας διάλογος επικοινωνίας μεταξύ των αλγορίθμων και του βιντεοπαιχνιδιού.

Για να αναπτύξουμε το πρώτο μέρος της συγκεκριμένη ιδέας, επιλέξαμε την Python λόγω των πολλών πλεονεκτημάτων της. Ένα βασικό πλεονέκτημα είναι η μεγάλη και ενεργή κοινότητα προγραμματιστών, η οποία παρέχει πρόσβαση σε ένα ευρύ φάσμα βιβλιοθηκών όπως οι MediaPipe, OpenCV που χρησιμοποιούνται συνήθως για την ανάπτυξη σύνθετων εφαρμογών.

Επιπλέον, η Python είναι μια δημοφιλής επιλογή στον τομέα της μηχανικής μάθησης και της βαθιάς μάθησης, παρέχοντας άφθονους πόρους για μάθηση και ανάπτυξη. Ακόμη, είναι απλή στη σύνταξη και βοηθά στο να εστιάσουμε στην ανάπτυξη της απόδειξης της έννοιας χωρίς να κολλάμε κάτω από την πολυπλοκότητα της γλώσσας.

Παρακάτω θα παρουσιαστούν αλγοριθμικές τεχνικές που θα εφαρμοστούν στην εργασία μας, ακολουθώντας την βασική δομή που αναλύθηκε στις προηγούμενες ενότητες, πιο συγκεκριμένα:

- Αλγόριθμος ανίχνευσης χειρονομιών
- Αλγόριθμος ανίχνευσης αντικειμένων και παρακολούθησης με βάση το χρώμα

### 6.2 OpenCV

Μια από τις βασικές βιβλιοθήκες που θα χρησιμοποιήσουμε στην γλώσσα προγραμματισμού Python είναι το OpenCV. Το OpenCV (Open Source Computer Vision Library) στοχεύει στην επεξεργασία δυναμικών εικόνων σε πραγματικό χρόνο. Το

OpenCV είναι μια βιβλιοθήκη API (Application Programming Interface) που είναι πολύ εξοικειωμένη με την επεξεργασία εικόνας σχεδιασμένη για εφαρμογές σε πραγματικό χρόνο, έχει καλές λειτουργίες λήψης εικόνας / βίντεο (Syafriandi *et al.*, 2017). Το OpenCV έχει πολλές λειτουργίες που μπορούν να βοηθήσουν στο να γίνουν οι εφαρμογές επεξεργασίας εικόνας ευκολότερες και λιγότερο περίπλοκες (Edward *et al.*, 2020). Το OpenCV μπορεί να χρησιμοποιηθεί με διάφορες γλώσσες προγραμματισμού όπως Python, Java και C (Bradski *et al.*, 2008).

### **6.3 Αλγόριθμος ανίχνευσης αντικειμένων και παρακολούθησης με βάση το χρώμα**

Στην ενότητα αυτή θα αναλυθεί ο κώδικας που γράφτηκε για την ανίχνευση μηκών κύματος του οπτικού πεδίου χρησιμοποιώντας την μετατροπή του χώρου από RGB σε χώρο HSV όπως περιγράψαμε σε προηγούμενο κεφάλαιο.

Αυτή είναι μια εφαρμογή ανίχνευσης πολλαπλών χρωμάτων του λαχανί και του ροζ σε πραγματικό χρόνο χρησιμοποιώντας τη γλώσσα προγραμματισμού Python. Το χρώμα σε πραγματικό χρόνο είναι επίσης πολύ σημαντικό. Η ανίχνευση του στόχου πραγματοποιείται από το χρωματικό χαρακτηριστικό, χρησιμοποιώντας το χρωματικό μοντέλο HSV. Εδώ οριοθετούνται τα χρώματα του λαχανί και του ροζ. Επιλέχθηκαν τα χρώματα αυτά διότι θα είναι τα αντικείμενα που θα κρατούν οι παίχτες του παιχνιδιού και θα δίνουν με την αναγνώριση τους την κίνηση στον χαρακτήρα του παιχνιδιού.

Βήματα για τον εντοπισμό και την παρακολούθηση:

Το πρώτο βήμα είναι να καταγράψουμε βίντεο μέσω μιας κάμερας web. Αυτό λαμβάνεται ως είσοδος. Τα καρέ της εικόνας είναι προσβάσιμα από τη ροή του βίντεο. Το χρωματικό μοντέλο HSV μπορεί να χειριστεί αλλαγές που προκαλούνται από το φωτισμό. Ως εκ τούτου, η εικόνα που είναι αποθηκευμένη σε μορφή RGB πρέπει να μετατραπεί σε HSV. Αυτό το χρωματικό μοντέλο περιγράφει το χρώμα ως προς την ποσότητα του γκρι και την τιμή φωτεινότητας του. Η τιμή απόχρωσης επεκτείνεται από 0-179, η τιμή κορεσμού είναι επεκτείνεται από 0-255 και η τιμή επεκτείνεται από 0-255 αντίστοιχα. Η αντίστοιχη μάσκα φτιάχνεται ορίζοντας το εύρος του επιθυμητού χρώματος (λαχανί, ροζ) προς ανίχνευση.

#### **6.3.1 Είσοδος: Λήψη βίντεο μέσω κάμερας web**

Ένα βίντεο είναι μια ακολουθία εικόνων που κινούνται γρήγορα. Το μέτρο του είναι πόσο γρήγορες είναι οι εικόνες. Η μετάβαση των εικόνων δίνεται από μια μέτρηση που



ονομάζεται καρτέ ανά δευτερόλεπτο (FPS). Εάν το βίντεο έχει FPS 40, σημαίνει ότι εμφανίζονται 40 εικόνες κάθε δευτερόλεπτο. Τα άλλα σημαντικά χαρακτηριστικά είναι το πλάτος και το ύψος, του πλαίσιο της εικόνας.

Σε αυτή τη μελέτη, η προσέγγιση που βασίζεται στην υπολογιστική όραση επιλέχθηκε μια απλή web κάμερα για την αναγνώριση των χειρονομιών και των χρωμάτων καθώς προσφέρει μια οικονομικά αποδοτική και πρακτική λύση.

Αυτή ήταν μια σκόπιμη επιλογή, καθώς δεν είναι μια υψηλής ποιότητας και όχι ακριβή κάμερα. Χρησιμοποιώντας μια λιγότερο δαπανηρή κάμερα, επισημαίνεται ότι η προσέγγιση που βασίζεται στην υπολογιστική όραση για την αναγνώριση χειρονομιών μπορεί να χρησιμοποιηθεί χωρίς να απαιτείται υψηλό κόστος. Χρησιμοποιεί επίσης ως απόδειξη ότι το σύστημα μπορεί να λειτουργήσει αποτελεσματικά με μια κάμερα σχετικά χαμηλού κόστους, καθιστώντας το πιο προσιτό σε ένα ευρύτερο κοινό.

Εν τω μεταξύ, στο βήμα προ επεξεργασίας, θα πάρουμε την ακατέργαστη εικόνα εισόδου και θα την προετοιμάσουμε για το επόμενο στάδιο εκτελώντας μια σειρά λειτουργιών σε αυτήν. Μία από τις βασικές λειτουργίες είναι η αλλαγή μεγέθους της εικόνας.

Μια άλλη ενέργεια που μπορεί να χρειαστεί να κάνουμε είναι να μετατρέψουμε τον χρωματικό χώρο της εικόνας. Ο χρωματικός χώρος μιας εικόνας αναφέρεται στον τρόπο με τον οποίο αναπαρίστανται τα χρώματα στην εικόνα και διαφορετικά μοντέλα ενδέχεται να περιμένουν διαφορετικούς χρωματικούς χώρους. Ίσως χρειαστεί να μετατρέψουμε την εικόνα σε συγκεκριμένο χρωματικό χώρο, ώστε να πληροί τις απαιτήσεις του μοντέλου. Αυτό θα διασφαλίσει ότι το μοντέλο θα είναι σε θέση να επεξεργαστεί σωστά την εικόνα και να εξαγάγει σχετικές λειτουργίες από αυτήν.

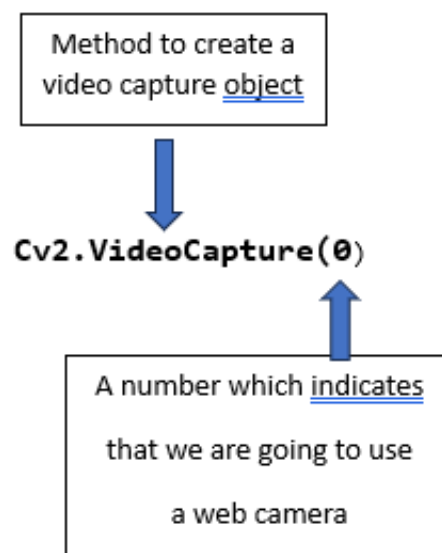
Όλες αυτές οι λειτουργίες είναι σημαντικές καθώς βοηθούν το μοντέλο να εξαγάγει σχετικά χαρακτηριστικά από την εικόνα που βοηθούν στη βελτίωση της απόδοσης του μοντέλου.

### **6.3.2 Ανάγνωση βίντεο**

Στο OpenCV, ένα βίντεο μπορεί να διαβαστεί είτε χρησιμοποιώντας τη ροή από μια κάμερα συνδεδεμένη σε ένα υπολογιστή ή διαβάζοντας ένα αρχείο βίντεο. Το πρώτο βήμα για την ανάγνωση ενός αρχείου βίντεο είναι η δημιουργία ενός αντικείμενου *VideoCapture*. Το όρισμά του μπορεί να είναι είτε το ευρετήριο συσκευής είτε το όνομα του αρχείου βίντεο προς ανάγνωση. Στις περισσότερες περιπτώσεις, μόνο μία κάμερα

είναι συνδεδεμένη στο σύστημα, αν έχουμε μια κάμερα. Όλα λοιπόν τότε περνάνε από το '0' και το OpenCV χρησιμοποιεί τη μοναδική κάμερα που είναι συνδεδεμένη στον υπολογιστή. Όταν περισσότερες από μία κάμερες είναι συνδεδεμένες στον υπολογιστή, μπορούμε να επιλέξουμε τη δεύτερη κάμερα περνώντας το '1', την τρίτη κάμερα περνώντας το '2' .

Θα εφαρμόσουμε τη συνάρτηση `cv2.VideoCapture` και θα δημιουργήσουμε μια παρουσία κλάσης. Προκειμένου να έχουμε πρόσβαση για ροή βίντεο, θα βάλουμε τις παραμέτρους της κάμερας. Μπορούμε να παρέχουμε ένα ευρετήριο του, την κάμερά μας από την οποία θέλουμε να διαβάσουμε δεδομένα. Σε περίπτωση που έχουμε μόνο μία κάμερα, από προεπιλογή, θα ευρετηριαστεί με 0 (βλ. εικόνα 17). Εάν έχουμε περισσότερες από μία κάμερες, η δεύτερη θα είναι με το όνομα 1, τρίτο με 2 και ούτω καθεξής.

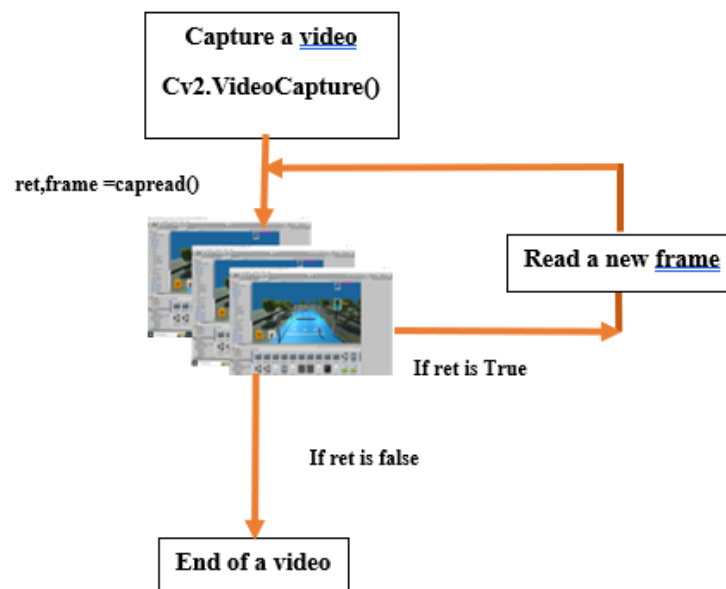


Εικόνα 17: Εντολή δημιουργίας αντικειμένου Video Capture

Αφού διαβάσουμε ένα αρχείο βίντεο, μπορούμε να εμφανίσουμε το βίντεο καρέ καρέ. Ένα καρέ ενός βίντεο είναι απλά μια εικόνα και εμφανίζουμε κάθε καρέ με τον ίδιο τρόπο που εμφανίζουμε τις εικόνες.

### 6.3.3 Ανάγνωση ροή βίντεο σε καρτέ εικόνων

Στην αρχή αυτής της διαδικασίας ο δείκτης μας βρίσκεται στο πρώτο καρτέ. Όταν κάνουμε αίτηση με την εντολή `cap.read()` θα φορτωθεί το πρώτο καρτέ από το αρχείο βίντεο. Θα αποθηκευτεί σε ένα μεταβλητό πλαίσιο. Εάν καλέσουμε ξανά αυτήν την εντολή, θα φορτωθεί το δεύτερο πλαίσιο και ούτω καθεξής. Η μεταβλητή `ret` είναι ένας τύπος δεδομένων `boolean` που επιστρέφει `True`, τότε μόνο μπορούμε να εκτελέσουμε τη λειτουργία ανάγνωσης με επιτυχία. Το καρτέ μας μπορεί να φορτωθεί ως έγχρωμη εικόνα (θα έχει 3 κανάλια) ή εικόνα σε κλίμακα του γκρι (θα έχει 1 κανάλι). Στην παρακάτω εικόνα 18 φαίνεται η διαδικασία φόρτωσης και ανάγνωσης από μια συνεχόμενη ροή εικόνων.



Εικόνα 18:Βρόχος Video Capture

Αφού ολοκληρώσουμε αυτή τη διαδικασία, μπορούμε να απελευθερώσουμε την έξοδο μας με το χρωματικό μοντέλο HSV ή Hue, Saturation και Value ενός δεδομένου αντικειμένου, όπου είναι ο χρωματικός χώρος που σχετίζεται μια μεταβλητή πλαισίου και το `ret` που είναι τύπος `boolean` επιστρέφει `True` εάν η Python μπορεί να διαβάσει έξοδο βίντεο. Επίσης θα δημιουργήσουμε κώδικα βρόχο με μια εντολή `while` που διαβάζει τα καρτέ από το βίντεο μας συνεχώς. Μπορούμε να το κάνουμε αυτό με την εντολή `cap.read()`. Αφού διαβάσουμε ένα αρχείο βίντεο ή καταγράψουμε μια ζωντανή ροή, θέλουμε να επεξεργαστούμε και να εμφανίσουμε το δικό μας βίντεο και αυτό το πετυχαίνουμε με την εντολή `cap.release()`.

#### 6.3.4 Μετατροπή της εικόνας από RGB σε χρωματικό χώρο HSV

Όπως εξηγήσαμε στην ενότητα 4.4, μια εικόνα μπορεί να αναπαρασταθεί όχι μόνο σε RGB αλλά και στο χρωματικό χώρο HSV, το OpenCV μπορεί εύκολα να μετατρέψει την εικόνα από το χρωματικό μοντέλο RGB σε HSV. Για την μετατροπή της εικόνας από RGB σε HSV χρησιμοποιείται η συνάρτηση:

```
hsvcolorspace = cv.cvtColor (εικόνα, cv.COLOR_BGR2HSV)
```

Η παράμετρος `cv.COLOR_BGR2HSV` δηλώνει την μετατροπή της εικόνας από το χρωματικό μοντέλο BGR σε HSV.

Η εικόνα HSV με την εντολή `return hsv` μπορεί επίσης να χωριστεί σε τρία κανάλια, με βάση την απόχρωση, τον κορεσμό και την τιμή.

#### 6.3.5 Λειτουργία του εύρους HSV στο OpenCV

Όποτε θέλουμε να λύσουμε προβλήματα που σχετίζονται με την ανίχνευση αντικειμένων, είναι απαραίτητο να χρησιμοποιούμε το HSV και να βρίσκουμε το χρωματικό εύρος του αντικειμένου. Η απόχρωση, ο κορεσμός και η τιμή στο HSV έχουν το δικό τους εύρος τιμών για κάθε χρώμα.

Η σύνταξη για τον ορισμό του εύρους HSV στο OpenCV είναι η εξής:

```
hsvcolorspace = cv.cvtColor (εικόνα, cv.COLOR_BGR2HSV)
```

```
low_hsvcolorspace = np.array ([Εύρος απόχρωσης, εύρος  
κορεσμού, εύρος τιμών])
```

```
upper_hsvcolorspace = np.array ([Εύρος απόχρωσης, εύρος  
κορεσμού, εύρος τιμών])
```

Όπου `hsvcolorspace` είναι η μετατροπή της δεδομένης εικόνας από την μορφή RGB σε HSV μορφή, `low_hsvcolorspace` είναι το κατώτερο όριο για μια σειρά χρωμάτων, το `upper_hsvcolorspace` είναι το ανώτερο όριο για μια σειρά χρωμάτων.

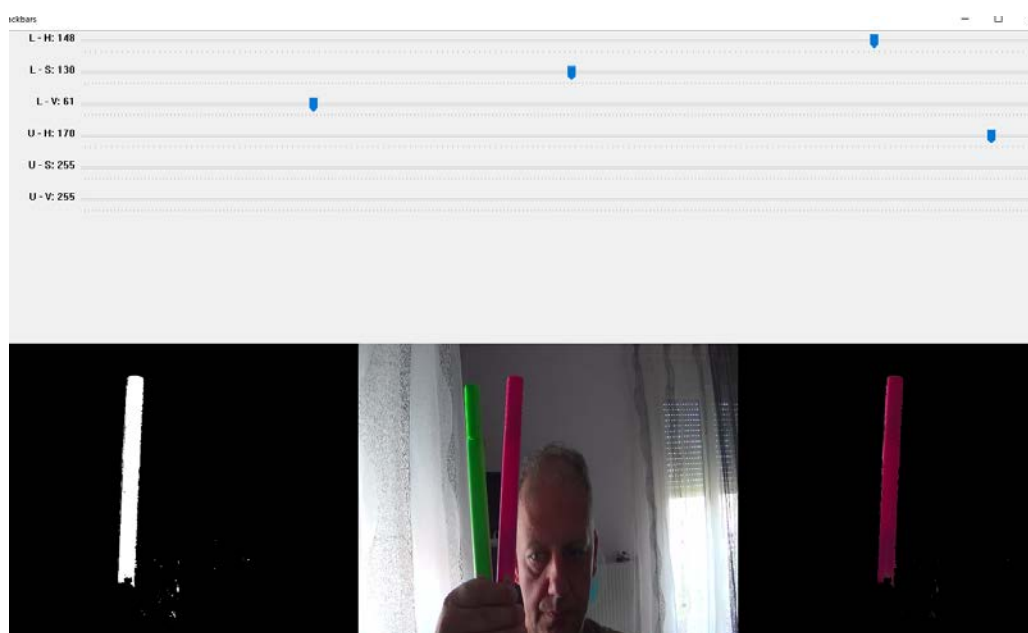
Τα χρώματα που θα πρέπει ανιχνευτούν από τον αλγόριθμο είναι το ροζ και το λαχανί διότι αυτά τα χρώματα έχουν τα αντικείμενα που επιλέχτηκαν να χρησιμοποιεί ο παίχτης του παιχνιδιού.

Τα όρια των τιμών για το ροζ χρώμα είναι:

'hmin': 148, 'smin': 130, 'vmin': 61

'hmax': 170, 'smax': 255, 'vmax': 255

Για μεγαλύτερη ακρίβεια μπορούμε να εντοπίσουμε τα κατώτερα και τα ανώτερα όρια του επιθυμητού χρώματος μέσω προγράμματος της Python και της βιβλιοθήκης OpenCV (βλ. εικόνα 19).



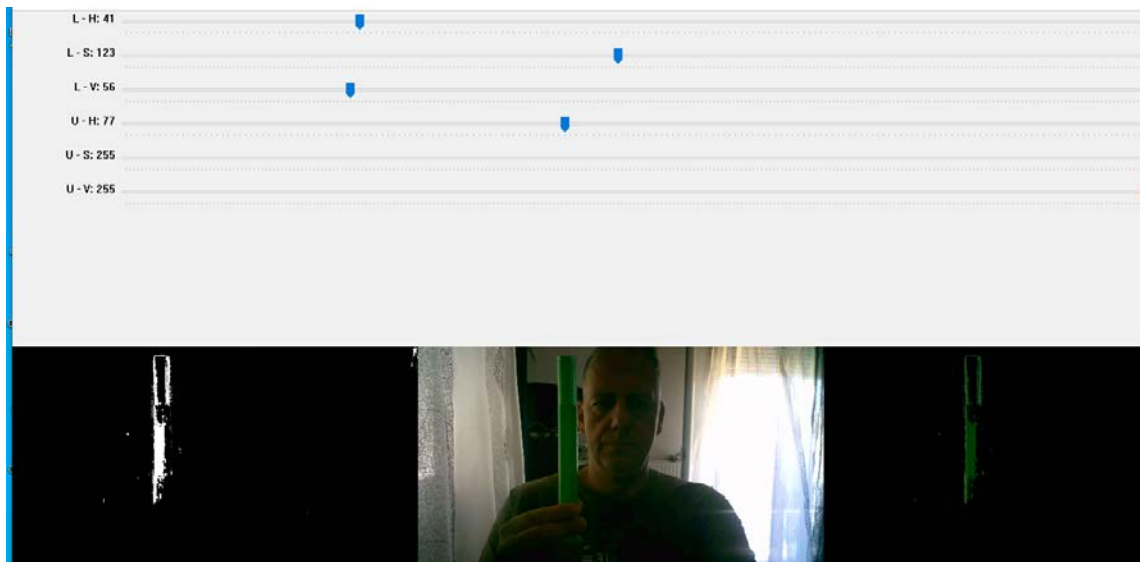
Εικόνα 19:Εύρος τιμών ροζ χρώματος

Ενώ για το λαχανί είναι :

'hmin': 41, 'smin': 123, 'vmin': 56

'hmax': 77, 'smax': 255, 'vmax': 255

Αντίστοιχα εκτελώντας το πρόγραμμα ανίχνευσης χρωματικών ορίων του δεύτερου αντικειμένου βρίσκουμε και τα όρια του λαχανί χρώματος (βλ. εικόνα 20)



Εικόνα 20:Εύρος τιμών λαχανί χρώματος

### 6.3.6 Συγκάλυψη

Η τμηματοποίηση χρωμάτων ή το φιλτράρισμα χρώματος χρησιμοποιείται ευρέως στο OpenCV για τον προσδιορισμό αντικείμενων/περιοχών που έχουν συγκεκριμένο χρώμα. Ο πιο ευρέως χρησιμοποιούμενος χρωματικός χώρος είναι το RGB, ονομάζεται πρόσθετος χρωματικός χώρος καθώς οι τρεις χρωματικές αποχρώσεις αθροίζονται για να δώσουν χρώμα στην εικόνα. Για να προσδιορίσετε μια περιοχή συγκεκριμένου χρώματος, ορίζουμε ένα κατώφλι και δημιουργούμε μια μάσκα, για να διαχωριστούν τα διαφορετικά χρώματα. Ο χρωματικός χώρος HSV είναι πολύ πιο χρήσιμος καθώς τα χρώματα στο χώρο HSV είναι πολύ πιο εντοπισμένα. Το Color Filtering έχει πολλές εφαρμογές και χρησιμοποιεί περιπτώσεις όπως κρυπτογραφία, υπέρυθη ανάλυση, συντήρηση τροφίμων ευπαθών τροφίμων κ.λπ. Σε τέτοιες περιπτώσεις, οι έννοιες της επεξεργασίας εικόνας μπορούν να χρησιμοποιηθούν για την εύρεση ή εξαγωγή περιοχών του συγκεκριμένου χρώματος. Για την τμηματοποίηση χρώματος, το μόνο που χρειαζόμαστε είναι οι τιμές κατωφλίου ή η γνώση του κάτω ορίου και του άνω ορίου του φάσματος χρωμάτων. Αφού καθοριστεί το εύρος του χρώματος που θα τμηματοποιηθεί, χρειάζεται να δημιουργηθεί μια μάσκα, κατά συνέπεια και χρησιμοποιώντας το, μπορεί να διαχωριστεί μια συγκεκριμένη περιοχή ενδιαφέροντος. Η κάλυψη περιλαμβάνει τη ρύθμιση ορισμένων από τις τιμές των εικονοστοιχείων σε μια εικόνα στο μηδέν ή σε κάποια άλλη τιμή. Η κάλυψη μπορεί να γίνει με τη χρήση μιας εικόνας ως μάσκα.

Μια εικόνα μάσκας είναι απλώς μια εικόνα όπου μερικά από τα εικονοστοιχεία οι τιμές έντασης των εικονοστοιχείων είναι μηδέν και άλλες είναι μη μηδενικές. Αν ένα εικονοστοιχείο βρίσκεται εκτός του εύρους των χρωματικών ορίων που έχουμε ορίσει τότε οι τιμές έντασης των εικονοστοιχείων είναι 0 ενώ αν είναι εντός των χρωματικών ορίων οι τιμές των εικονοστοιχείων είναι 1.

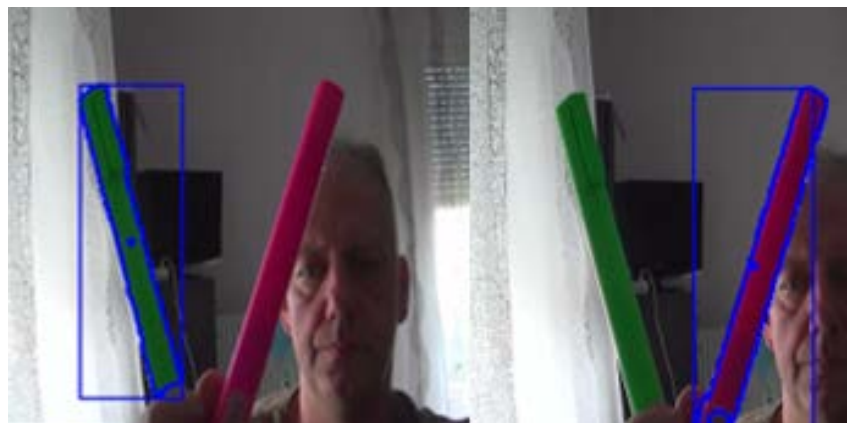
### 6.3.7 Σχεδιασμός περιγραμμάτων

Για να σχεδιάσουμε τα περιγράμματα, χρησιμοποιείται η λειτουργία `cv.drawContours`. Το πρώτο του όρισμα είναι η εικόνα πηγής, δεύτερον όρισμα είναι τα περιγράμματα που πρέπει να περάσουν ως λίστα Python, το τρίτο όρισμα είναι το ευρετήριο των περιγραμμάτων, για να σχεδιαστούν όλα τα περιγράμματα, ορίζουμε τη τιμή -1 και τα υπόλοιπα ορίσματα είναι το χρώμα, το πάχος κ.λπ.

- Για να σχεδιαστούν τα περιγράμματα σε μια εικόνα γράφουμε την εντολή:

```
cv.drawContours(img, περιγράμματα, -1, (0,255,0), 3)
```

το αποτέλεσμα της παραπάνω εντολής φαίνεται στην εικόνα 21.



Εικόνα 21:Περιγράμματα αντικειμένων

### 6.3.8 Συμπέρασμα

Εύκολα μπορούμε να αντιληφθούμε ότι με την βοήθεια μιας κάμερας που θα την χρησιμοποιήσουμε ως πηγή εισόδου και εφαρμόζοντας τον αλγόριθμο ανίχνευσης αντικειμένων και παρακολούθησης με χρήση HSV μπορούμε να εντοπίσουμε συγκεκριμένα αντικείμενα. Επίσης με την βοήθεια του αλγορίθμου κάθε φορά που ανιχνεύεται το επιθυμητό χρώμα λαμβάνουμε πληροφορίες για την θέση του αντικειμένου πάνω στην οθόνη. Οι τιμές αυτές θα μας χρειαστούν ως πηγή εισόδου στο παιχνίδι που θα ορίζουνε κάθε χρονική στιγμή την νέα θέση του χαρακτήρα.

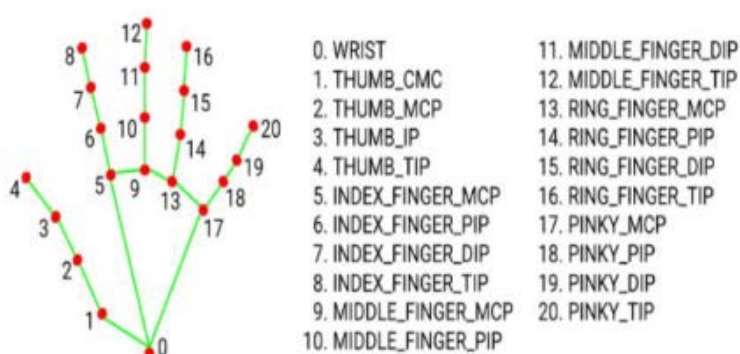
## 6.4 Αλγόριθμος ανίχνευσης χειρονομιών

### 6.4.1 Εισαγωγή

Στην ενότητα αυτή θα ασχοληθούμε με τον αλγόριθμο ανίχνευσης χειρονομιών που θα χρησιμοποιηθεί στο παιχνίδι. Υπάρχουν πολλές βιβλιοθήκες και τεχνικές για την ανίχνευση των χειρονομιών. Στην συγκεκριμένη εργασία μετά από δοκιμές κατέληξα να χρησιμοποιήσω την βιβλιοθήκη MediaPipe της Google. Η απόδοση του συγκεκριμένου αλγορίθμου για τα δεδομένα της εργασίας ήταν πολύ θετικά.

### 6.4.2 Αναγνώριση χειρονομίας με Mediapipe

Για να προσδιορίσουμε τις διαφορετικές στάσεις του χεριού, θα υπολογίσουμε τα 21 βασικά σημεία στο ορόσημο χεριός, εξηγείται στην ενότητα 5.3.3. Η ταυτοποίηση της χειρονομίας θα μπορούσε να γίνει με το να προσδιορίσουμε αν το κάθε δάχτυλο του χεριού είναι ανοιχτό ή κλειστό.



Εικόνα 22 :Ορόσημο χεριός στο MediaPipe [33]

Στην εικόνα 22, μπορούμε να δούμε ότι οι συντεταγμένες των σημείων [4,8,12,16,20] είναι συντεταγμένες στις άκρες των δακτύλων και δηλώνονται ως ακροδάχτυλα. Η δήλωση του χεριού με συντεταγμένες 0 έως 20 λαμβάνονται από 21 βασικά σημεία με μια συντεταγμένη σε κάθε άρθρωση, στο ορόσημο χεριός. Θα συγκρίνουμε τις συντεταγμένες ενός ακροδαχτύλου με τις συντεταγμένες ενός μεσαίου σημείου του ίδιου δαχτύλου. Εάν η συντεταγμένη ενός ακροδαχτύλου έχει τιμή μεγαλύτερη από τη συντεταγμένη ενός μεσαίου σημείου, τότε ορίζουμε το δάχτυλο με την τιμή 1, τότε το



δάκτυλο δείκτης είναι σε κατάσταση ανοιχτό και αντίστροφα. Η σύγκριση μπορεί να γίνει ως προς τον άξονα y.

```
if hand_landmarks.landmark[mp_hands.HandLandmark.INDEX_FINGER_TIP].y  
> hand_landmarks.landmark[mp_hands.HandLandmark.INDEX_FINGER_DIP].y:
```

Οι παραπάνω γραμμές κώδικα χρησιμοποιούνται για την αναγνώριση της κατάστασης των δακτύλων του ενός χεριού που σχετίζεται με την εικόνα 22. Συγκεκριμένα στον κώδικα αυτό ελέγχουμε την συντεταγμένη ως προς τον άξονα y του σημείου[4] με την συντεταγμένη του σημείου [2].

Αν  $[y4] > [y2]$  τότε το δάκτυλο είναι ανοιχτό .

Ενώ αν  $[y4] < [y2]$  τότε το δάκτυλο είναι κλειστό.

Στο επόμενο βήμα, αφού προσδιορίσαμε την κατάσταση του κάθε δακτύλου δηλαδή αν είναι ανοιχτό ή κλειστό, όπως αναφέρθηκε προηγουμένως, συγκεντρώνουμε όλες τις τιμές των δακτύλων του χεριού και περνάμε στο στάδιο ελέγχου της συνθήκης.

Αν ένα δάκτυλο βρίσκεται σε κατάσταση ανοιχτό τότε το αποτέλεσμα της συνθήκης μας επιστρέφει την μεταβλητή  $val_i=1$  αλλιώς  $val_i=0$  όπου i παίρνει τιμές από 1 έως 5 όσα και τα δάκτυλα του ενός χεριού. Επίσης κάθε φορά που η μεταβλητή  $val_i=1$  το πρόγραμμα μας επιστρέφει και μια ετικέτα  $fin = \text{“ Το όνομα του δακτύλου”}$ . Όπως βλέπουμε στον παρακάτω κώδικα όταν η μεταβλητή  $val_i=1$  τότε η ετικέτα είναι  $fin = \text{“ Index”}$ . Στο τέλος του κώδικα αθροίζονται όλες οι μεταβλητές  $val_i$  σε μια μεταβλητή  $val = val_1 + val_2 + val_3 + val_4 + val_5$ , το αποτέλεσμα της μεταβλητής  $val$  μας δίνει ένα αριθμό που ισούται με τα δάκτυλα που είναι στην κατάσταση ανοιχτό .

Ακολουθεί παρακάτω ο κώδικας αναγνώρισης χειρονομίας:

```
fin=""  
if hand_landmarks.landmark[mp_hands.HandLandmark.INDEX_FINGER_TIP].y  
> hand_landmarks.landmark[mp_hands.HandLandmark.INDEX_FINGER_DIP].y:  
    val1 = 0  
else:  
    val1 = 1  
    fin = 'Index '
```

```

if
hand_landmarks.landmark[mp_hands.HandLandmark.MIDDLE_FINGER_TIP].y
>
hand_landmarks.landmark[mp_hands.HandLandmark.MIDDLE_FINGER_DIP].y:
    val2 = 0
else:
    val2 = 1
    fin += 'Middle '

if hand_landmarks.landmark[mp_hands.HandLandmark.RING_FINGER_TIP].y >
hand_landmarks.landmark[mp_hands.HandLandmark.RING_FINGER_DIP].y:
    val3 = 0
else:
    val3 = 1
    fin += 'Ring '

if hand_landmarks.landmark[mp_hands.HandLandmark.PINKY_TIP].y >
hand_landmarks.landmark[mp_hands.HandLandmark.PINKY_DIP].y:
    val4 = 0
else:
    val4 = 1
    fin += 'PINKY '

if hand_landmarks.landmark[mp_hands.HandLandmark.THUMB_TIP].x <
hand_landmarks.landmark[mp_hands.HandLandmark.THUMB_IP].x:
    val5 = 0
else:
    val5 = 1
    fin += 'Tump '

val=val1 +val2+val3+val4+val5

print(val)
fps= str(val)+' fingers'

```

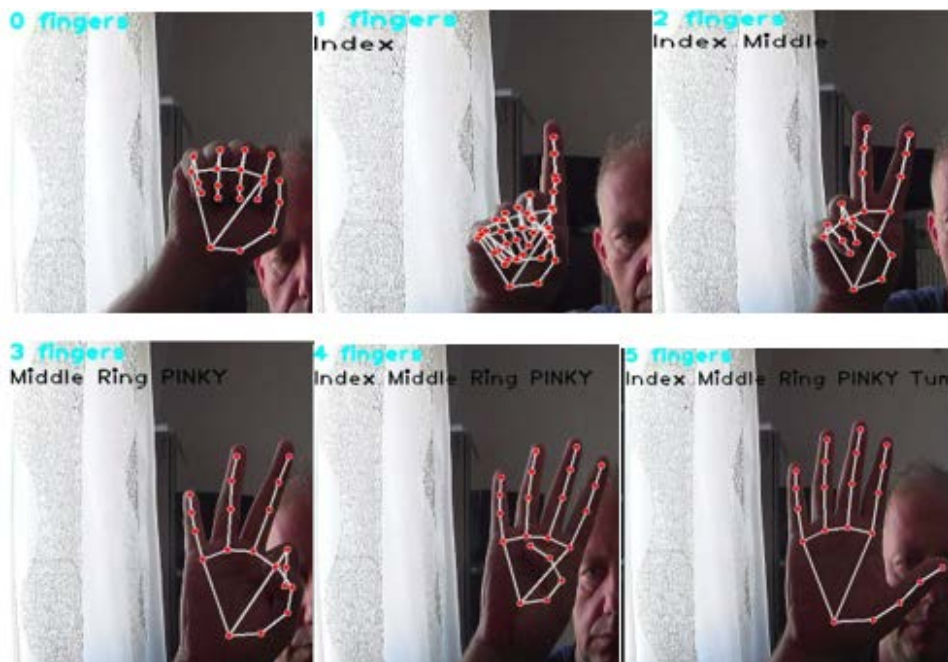
Στην εικόνα 23 που ακολουθεί βλέπουμε την ανίχνευση της παλάμης του χεριού, την σκελετική άρθρωση του ορόσημου χειρός, επίσης λαμβάνουμε την πληροφορία 5 *figure* που μας δηλώνει ότι και τα πέντε δάκτυλα του χεριού είναι στην κατάσταση *ανοιχτό*. Η τελευταία αριθμητική πληροφορία είναι σημαντική και θα μας χρειαστεί για την λειτουργία της οθόνης διεπαφής του παιχνιδιού.



Εικόνα 23:Αναγνώριση χειρονομίας

#### 6.4.6 Περιγραφή του συνόλου δεδομένων

Ο κύριος στόχος είναι να αναγνωριστούν οι χειρονομίες του παίχτη από την οθόνη διεπαφής του παιχνιδιού και να εμφανιστεί ένα μενού λειτουργιών του βιντεοπαιχνιδιού. Χρησιμοποιήθηκαν στο παιχνίδι πέντε χειρονομίες που αποτυπώνονται παρακάτω στην εικόνα 24.



Εικόνα 24:Απεικόνιση των συνολικών χειρονομιών

Ξεκινώντας το παιχνίδι που δημιουργήσαμε ο χρήστης στην πρώτη οθόνη του παιχνιδιού μπορεί να επιλέξει τις επιλογές που εμφανίζονται κάνοντας κάποια χειρονομία.

Ο χρήστης έχοντας όλα τα δάκτυλα κλειστά δηλαδή την επιλογή «0 figure» όπως βλέπουμε στην εικόνα 24 μπορεί να επιλέξει να παίξει το παιχνίδι με έναν παίχτη, ενώ με την επιλογή «3 figure» μπορεί να επιλέξει να παίξει με δύο παίχτες (βλ. εικόνα 25). Τέλος έχει οριστεί στο μενού του παιχνιδιού ότι με την επιλογή «4 figure» μπορεί ο παίχτης να τερματίσει το παιχνίδι, παρακάτω στην εικόνα 25 φαίνεται η οθόνη διεπαφής του παιχνιδιού.



*Εικόνα 25:Μενού διεπαφής παιχνιδιού*

Κατά την διάρκεια του παιχνιδιού υπάρχει δυνατότητα από τον παίχτη να παγώνει το παιχνίδι για όσο χρόνο θέλει επιλέγοντας την χειρονομία «2 figure» παράδειγμα φαίνεται στην παρακάτω εικόνα 26. Στην κάτω αριστερή γωνία της οθόνης του παιχνιδιού υπάρχει ένα εικονίδιο που δείχνει στον χρήστη την χειρονομία που θα πρέπει να κάνει μπροστά στην κάμερα και με την βοήθεια της υπολογιστικής όρασης να σταματήσει ο χρόνος του παιχνιδιού.



*Εικόνα 26:Μενού διεπαφής παιχνιδιού εντός του παιχνιδιού*

## **6.5 Σχεδίαση παιχνιδιού**

### **6.5.1 Εισαγωγή**

Όπως αναφέρθηκε στην περίληψη το δεύτερο στάδιο αυτής της εργασίας αποτελείται από τη δημιουργία ενός τρισδιάστατου παιχνιδιού με θέμα έναν «αγώνα τένις» .Ο παίχτης έχει την δυνατότητα να βρίσκεται σε ένα γήπεδο τένις και να παίζει το παιχνίδι είτε με τον υπολογιστή σαν αντίπαλο είτε με έναν δεύτερο παίχτη διαμοιράζοντας την οθόνη προβολής του παιχνιδιού.

### **6.5.2 Unity**

Ένα βασικό βήμα του σχεδιαστή παιχνιδιών είναι να διαλέξει το κατάλληλο πρόγραμμα ανάπτυξης και σχεδίασης βιντεοπαιχνιδιού. Οι πιο δημοφιλείς μηχανές ανάπτυξης είναι το Godot, η Unreal (UDK) και φυσικά το Unity3d .

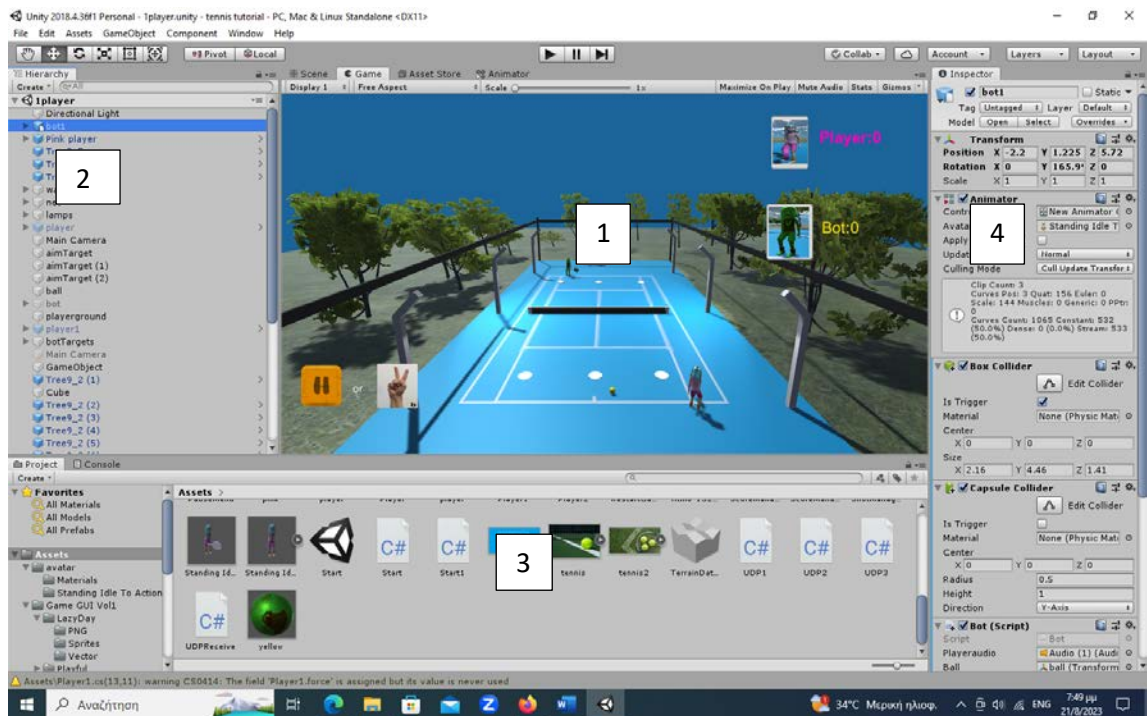
Το Unity είναι μια μηχανή παιχνιδιών που κυκλοφόρησε το 2005 και υποστηρίζει την ανάπτυξη για περισσότερες από 25 πλατφόρμες, περισσότερες από οποιαδήποτε άλλη μηχανή δημιουργίας. Υποστηρίζει την ανάπτυξη 2D και 3D παιχνιδιών.

Η κύρια γλώσσα δέσμης ενεργειών του Unity είναι η C#, την οποία υποστηρίζει εγγενώς. Άλλα αξιοσημείωτα χαρακτηριστικά του Unity είναι οι δύο μηχανές φυσικής του (3D και 2D), εργαλεία κινούμενων εικόνων, γραφική παράσταση shader (προγραμματισμός shader βάσει κόμβου) και assetstore (Unity 2019).

Στη συγκεκριμένη εργασία επιλέχθηκε το Unity3d επειδή τη διδαχθήκαμε στα πλαίσια του Μεταπτυχιακού, καθώς επίσης για την προτίμηση που έχουμε για την γλώσσα προγραμματισμού η οποία είναι η C#.

### **6.5.3 Η διεπαφή χρήστη**

Το πρώτο βήμα για την εκμάθηση οποιουδήποτε νέου λογισμικού είναι να εξοικειωθείτε με το περιβάλλον διεπαφής. Η προεπιλεγμένη διεπαφή χρήστη του Unity που φαίνεται στην εικόνα 27 αποτελείται από τέσσερα παράθυρα που εξυπηρετούν τους δικούς τους σκοπούς το καθένα.



Εικόνα 27:Περιβάλλον χρήστη στη Unity

Το παράθυρο ιεραρχίας (παράθυρο 2 στην εικόνα 27) βρίσκεται στην αριστερή πλευρά της διεπαφής χρήστη. Περιέχει όλα τα αντικείμενα του παιχνιδιού που βρίσκονται στη σκηνή, όταν είναι ανοιχτή. Χρησιμοποιώντας το παράθυρο ιεραρχίας, ο προγραμματιστής μπορεί να επιλέξει ένα αντικείμενο για προβολή ή για επεξεργασία των ιδιοτήτων του σε άλλα παράθυρα. Όπως πολλά άλλα επαγγελματικά προγράμματα, το Unity υποστηρίζει την ανατροφή αντικειμένων. Αυτό σημαίνει ομαδοποίηση των «παιδικών αντικειμένων» κάτω από το πρώτο «γονικό αντικείμενο». Η ανατροφή των παιδιών διευκολύνει τη διαχείριση ενός μεγάλου αριθμού αντικειμένων κάνοντας το καθαριστικό στο παράθυρο της ιεραρχίας (παράθυρο 2 στην εικόνα 27) . Λειτουργίες μετασχηματισμού όπως η μετακίνηση και η κλιμάκωση που εφαρμόζονται στο γονικό αντικείμενο εφαρμόζονται επίσης και στα θυγατρικά του αντικείμενα (Unity Technologies 2019g.).Το παράθυρο σκηνής(παράθυρο 1 στην εικόνα 27) χρησιμοποιείται για την προβολή και την επεξεργασία του περιβάλλοντος της ανοιχτής σκηνής.

Τα επιλεγμένα αντικείμενα επισημαίνονται με πορτοκαλί χρώμα και μπορούν να μετακινηθούν, να κλιμακωθούν και να περιστραφούν χρησιμοποιώντας το εργαλείο gizmo. Η λειτουργία προβολής του παραθύρου σκηνής μπορεί να αλλάξει από το αναπτυσσόμενο μενού πάνω αριστερά. Αυτές οι λειτουργίες προβολής μπορούν να

χρησιμοποιηθούν για την προβολή σχετικών δεδομένων. Αυτό το παράθυρο έχει επίσης την προβολή παιχνιδιού και το κατάστημα Asset store ως καρτέλες του. Η προβολή παιχνιδιού είναι ενεργοποιημένη κατά τη δοκιμή του παιχνιδιού ενώ αναπτύσσεται και το κατάστημα Asset store που χρησιμοποιείται για την περιήγηση και την απόκτηση αντικειμένων από το Unity Asset store. (Unity Technologies 2019h.) Το παράθυρο επιθεωρητή (παράθυρο 4 στην εικόνα 27) χρησιμοποιείται για την επεξεργασία των ιδιοτήτων των επιλεγμένων στοιχείων του παιχνιδιού, υλικά, μοντέλα κα. Εδώ τα στοιχεία του GameObjects μπορούν επίσης να προσπεραστούν και να δημιουργηθούν. Το παράθυρο του έργου (παράθυρο 3 στην εικόνα 27) είναι ο διαχειριστής αρχείων του Unity. Μπορεί να χρησιμοποιηθεί για την προβολή των στοιχείων που εισάγονται και δημιουργήθηκαν στο φάκελο του έργου. Η αριστερή πλευρά δείχνει τη δομή του φακέλου και η δεξιά εμφανίζει όλα τα στοιχεία μέσα στον επιλεγμένο φάκελο. Διαθέτει επίσης λίστα αγαπημένων και γραμμή αναζήτησης που μπορεί να χρησιμοποιηθεί για την αναζήτηση στοιχείων με λέξεις-κλειδιά.

#### **6.5.4 Αντικείμενο παιχνιδιού**

Κάθε στοιχείο σε ένα παιχνίδι Unity είναι ένα αντικείμενο παιχνιδιού που περιλαμβάνει τρισδιάστατα στοιχεία, χαρακτήρες και φώτα. Ένα αντικείμενο παιχνιδιού από μόνο του δεν έχει καμία λειτουργικότητα, επομένως πρέπει να έχει εξαρτήματα συνδεδεμένα σε αυτό. Κάθε αντικείμενο παιχνιδιού περιέχει το στοιχείο Transform, το οποίο περιέχει δεδομένα θέσης, περιστροφής και κλίμακας του αντικειμένου. Άλλα στοιχεία περιλαμβάνουν σενάρια, πηγές ήχου, επιλογές διεπαφής χρήστη.

Το μενού περιέχει πολλά προκαθορισμένα αντικείμενα παιχνιδιού που έχουν κάποια σύνθεση και πολλά άλλα. Τα αντικείμενα παιχνιδιού μπορούν να δημιουργηθούν από το επάνω αριστερά αναπτυσσόμενου μενού ή από το Hierarchy παράθυρο.

Τα αντικείμενα παιχνιδιού περιέχουν επίσης ιδιότητες που ονομάζονται ετικέτες και επίπεδα. Οι ετικέτες χρησιμοποιούνται για την κατηγοριοποίηση των αντικειμένων σε διαφορετικές ομάδες, για ευκολότερο χειρισμό με σενάρια. Αυτά είναι ιδιαίτερα χρήσιμα σε σενάρια ελέγχου collider. Για παράδειγμα, μια προεπιλεγμένη ετικέτα "Παίκτης" είναι χρήσιμη για τον εντοπισμό εάν ο παίκτης έχει έρθει σε επαφή με την μπάλα του τένις, αυτό γίνεται με τον έλεγχο collider. Παρακάτω στην εικόνα 28 φαίνεται ένα παράδειγμα box collider από το παιχνίδι.



Εικόνα 28:Box collider παίχτη

### 6.5.5 Ελεγκτής χαρακτήρων

Ο ελεγκτής χαρακτήρων είναι ένα αντικείμενο παιχνιδιού που χρησιμοποιείται για τον έλεγχο ενός χαρακτήρα πρώτου ή τρίτου προσώπου μέσα στο πρόγραμμα Unity. Περιέχει όλα τα απαραίτητα στοιχεία που συνθέτουν τον χαρακτήρα του παιχνιδιού όπως σενάρια απόδοσης, συγκρουόμενων και κινήσεων. Τα παιχνίδια τένις της εργασίας είναι τρίτου προσώπου, έχει σχεδιαστεί για να κάνει τον χρήστη να αισθάνεται σαν να είναι πλάι στον κόσμο του παιχνιδιού, δηλαδή μέσα σε ένα γήπεδο.

Για να μετακινήσετε έναν χαρακτήρα σε μια μηχανή παιχνιδιών, ο χαρακτήρας πρέπει να έχει δέσμη ενεργειών και να αντιδρά στην εισαγωγή δεδομένων του χρήστη από μια συσκευή εισόδου. Στην περίπτωση μας η πηγή εισόδου θα προέρχεται από τα αντικείμενα που κρατά ο παίχτης στα χέρια, αυτά θα μας δώσουν την κίνηση αριστερά ή δεξιά.

### 6.5.6 Οι ήρωες του παιχνιδιού

Ο παίχτης ενσαρκώνεται με έναν 3D χαρακτήρα που απεικονίζεται στην παρακάτω εικόνα 29.

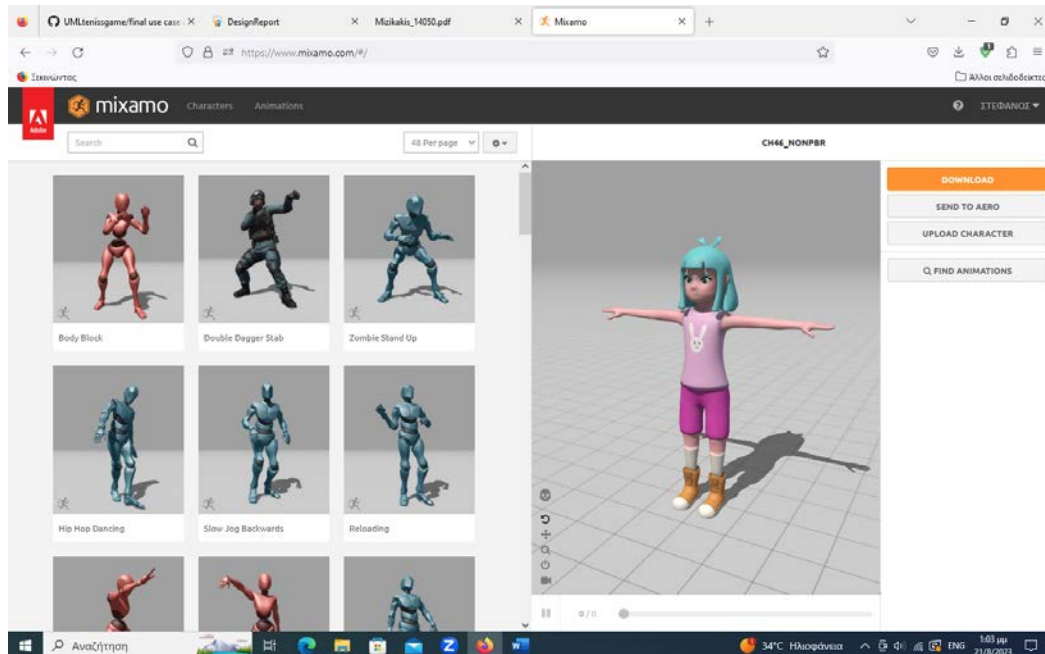




Εικόνα 29: Ήρωες του παιχνιδιού pink player στα αριστερά & green player στα δεξιά

Ο βασικός ήρωας είναι ο pink player ενώ όταν υπάρχει συμμετοχή δύο παιχτών, ο δεύτερος παίχτης είναι ο green player. Επιλέχθηκαν τα χρώματα ροζ και λαχανί για να υπάρχει μια αντιστοιχία με το χρωματικό χαρακτηριστικό των αντικειμένων που θα κρατά ο παίχτης ή οι παίχτες στα χέρια τους και θα δίνουν κίνηση στους χαρακτήρες με απώτερο σκοπό το χτύπημα της μπάλας με τη ρακέτα που έχει στα χέρια του ο ήρωας.

Ο βασικός χαρακτήρας ανακτήθηκε από την σελίδα της mixamo (βλ. εικόνα 30).



Εικόνα 30: Ήρωας mixamo (Πηγή :<https://mixamo.com>)

Ενώ για να αποδώσουμε περισσότερη ρεαλιστικότητα στον χαρακτήρα μας προστέθηκαν και χαρακτηριστικά κινήσεων χτυπήματος της μπάλας (βλ. εικόνα 31) με την βοήθεια της ρακέτας που θα κρατά ο ήρωας στα χέρια του.

Εικόνα 31: Δεξί & αριστερό χτύπημα (Πηγή :<https://mixamo.com>)

Το παιχνίδι περιλαμβάνει τον κύριο χώρο εξέλιξης του παιχνιδιού που είναι ένα γήπεδο τένις (βλ. εικόνα 32) .



Εικόνα 32: Γήπεδο τένις παιχνιδιού

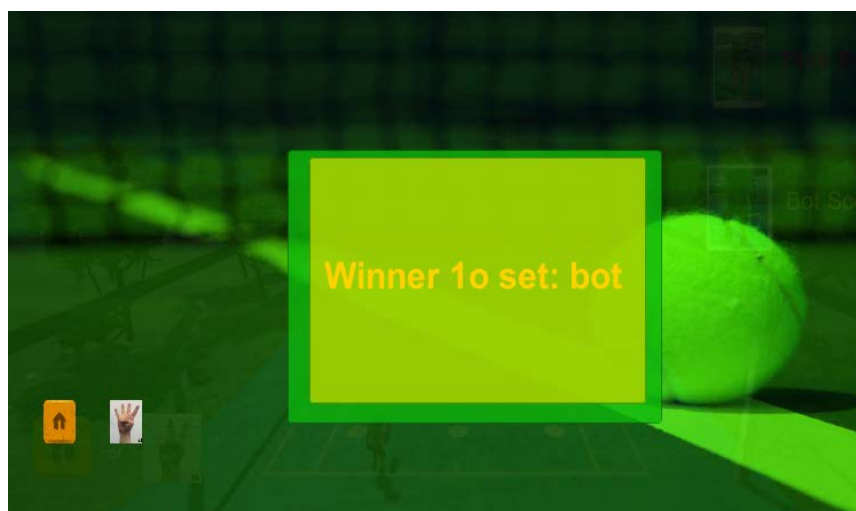
### 6.5.7 Κανόνες – Συστήματα

Τέλος θα αναφέρουμε ενδεικτικά τους μηχανισμούς εκείνους, που χρειάζεται το παιχνίδι μας να φαίνεται ολοκληρωμένο. Θα χρειαστούμε να κρατάμε τις εξής τιμές όταν θα παίζουμε:

- Πόντοι (Pink Player) ,οι πόντοι του πρώτου παίχτη
- Πόντοι (Bot Player), οι πόντοι του αντίπαλου παίχτη που είναι ο υπολογιστής
- Πόντοι (Green Player), οι πόντοι του δεύτερου παίχτη
- Final Score, το τελικό score του νικητή

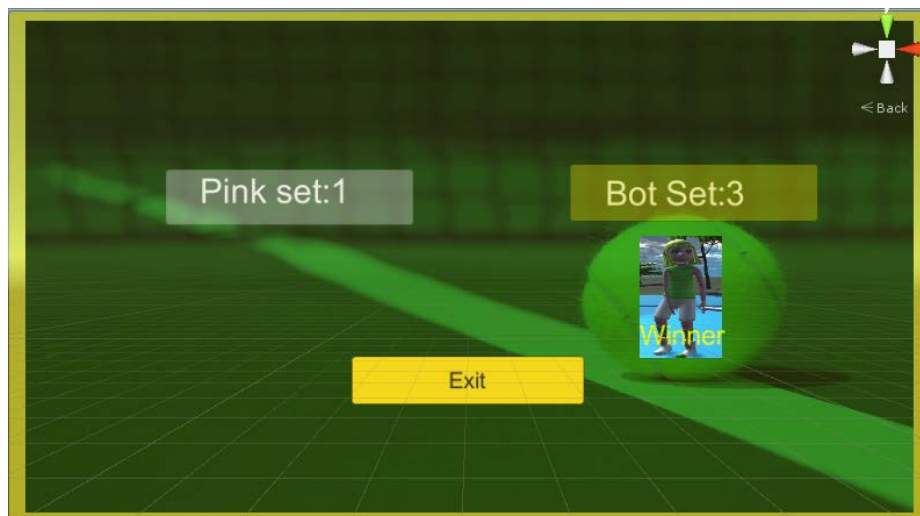
#### 6.5.7.1 Ανάλυση κανόνων

Με την έναρξη του παιχνιδιού και οι δύο παίχτες έχουν 0 πόντους , για να αυξηθούν οι πόντοι θα πρέπει το σέρβις που κάνει ο παίχτης ή ο υπολογιστής να είναι επιτυχημένο δηλαδή να χτυπήσει η μπάλα εντός των ορίων της περιοχής του αντιπάλου και ο αντίπαλος να μην μπορεί να αποκρούσει την μπάλα και να βγει εκτός των ορίων περιοχής του γηπέδου τένις, τότε σε αυτήν την περίπτωση ο παίχτης που έκανε ένα επιτυχημένο σέρβις κερδίζει 15 πόντους. Όταν ένας από τους δύο παίχτες φτάσει τους 60 πόντους κερδίζει τον πρώτο γύρο του παιχνιδιού (βλ. εικόνα 33).



Εικόνα 33: Οθόνη νικητή 1ου set παιχνιδιού

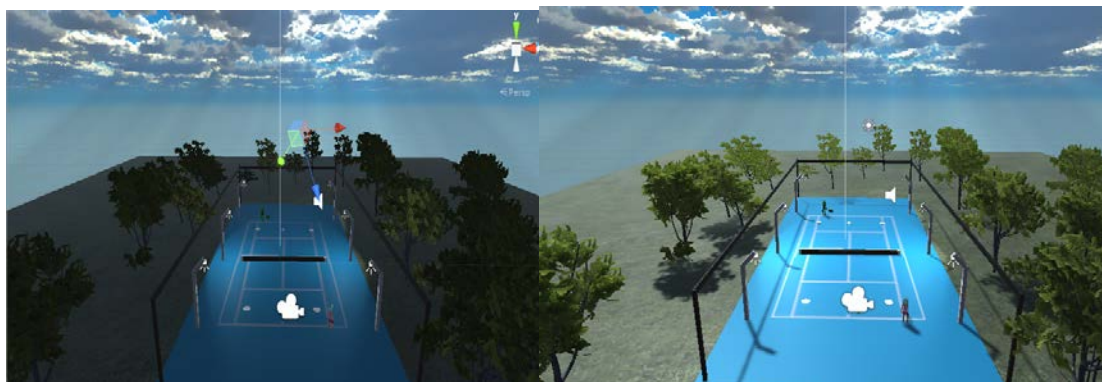
Συνεχίζοντας το παιχνίδι όποιος τελικά κερδίσει 3 γύρους παιχνιδιού θα είναι και ο τελικός νικητής (βλ. εικόνα 34).



Εικόνα 34: Τελικοί πόντοι παιχνιδιού

#### 6.5.8 Φωτισμός

Όπως στον πραγματικό κόσμο χρειάζεται φως για να δούμε αντικείμενα, έτσι και σε ένα παιχνίδι είναι αναγκαίο. Ο φωτισμός επηρεάζει δραματικά την εμφάνιση ενός παιχνιδιού και είναι ένα από τα πιο βασικά μέρη στη δημιουργία ρεαλιστικών γραφικών.



Εικόνα 35: Σκηνή χωρίς ή με φωτισμό

Στην εικόνα 35 βλέπουμε μια σκηνή σκοτεινή και την ίδια σκηνή πλήρως φωτισμένη δίπλα δίπλα. Με την τοποθέτηση κατάλληλου φωτισμού στην σκηνή επιτυγχάνουμε μια πιο δελεαστική σκηνή που με την τοποθέτηση κατάλληλων αντικειμένων να έχουμε ένα αποτέλεσμα πιο ρεαλιστικό για τον παίχτη. Ο φωτισμός στη Unity μπορεί να χωριστεί σε δύο κατηγορίες: σε πραγματικού χρόνου φωτισμός και προϋπολογισμένος φωτισμός. Συνήθως, αυτές οι τεχνικές συνδυάζονται για να γίνουν εύχρηστες οι ρυθμίσεις

φωτισμού. Ο φωτισμός σε πραγματικό χρόνο χρησιμοποιείται κυρίως για να φωτίζει κινητά αντικείμενα όπως χαρακτήρες ή οχήματα. Επειδή ο φωτισμός σε πραγματικό χρόνο υπολογίζεται κατά την εκτέλεση του παιχνιδιού, μπορεί να απαιτεί μεγάλη ένταση πόρων. Αυτός είναι επίσης ο λόγος που οι ακτίνες φωτός σε πραγματικό χρόνο δεν παράγουν φως που αναπηδά.

Προκειμένου να δημιουργηθεί πιο ρεαλιστικός φωτισμός, απαιτείται ο παγκόσμιος φωτισμός (GI). Παγκόσμιο illumination είναι ένας όρος που χρησιμοποιείται για να περιγράψει το φως που ανακλάται ή αναπηδά από άλλα αντικείμενα. Η συλλογή του παγκόσμιου φωτισμού μπορεί να είναι πολύ αργή για να εκτελεστεί σε πραγματικό χρόνο, επομένως απαιτείται λύση. Το burning Lightmap είναι μια μέθοδος που υπολογίζει τα παγκόσμια δεδομένα φωτισμού και τα αποθηκεύει σε χάρτες που ονομάζονται lightmaps. Αυτοί οι υπολογισμοί γίνονται κατά την ανάπτυξη του παιχνιδιού. Ένα μειονέκτημα αυτής της μεθόδου είναι ότι μπορεί να χρησιμοποιηθεί μόνο για στατικά αντικείμενα, λόγω του lightmap οι υφές δεν μπορούν να ενημερωθούν σε πραγματικό χρόνο. Αυτός είναι ο λόγος για τον οποίο τα περισσότερα παιχνίδια χρησιμοποιούν συνδυασμό τόσο burn φωτισμό όσο και φωτισμό σε πραγματικό χρόνο. Οι ρυθμίσεις φωτισμού βρίσκονται στο οριζόντιο μενού της Unity συγκεκριμένα πηγαίνοντας στο Παράθυρο > Απόδοση > Ρυθμίσεις φωτισμού. Το μενού έχει μια σειρά επιλογών για τον έλεγχο του φωτισμού περιβάλλοντος και των ρυθμίσεων GI. Αυτό το παράθυρο χρησιμοποιείται επίσης για το burn των φωτεινών χαρτών. Πριν από το burn, πρέπει να καθοριστούν ιδιότητες όπως ο αριθμός των αναπηδήσεων φωτός που υπολογίζεται για τον παγκόσμιο φωτισμό και η ανάλυση του χάρτη φωτός. Αυτές οι ρυθμίσεις επηρεάζουν τη διάρκεια του burning.

### **6.5.9 Διάγραμμα Ροής για την Έναρξη του Παιχνιδιού**

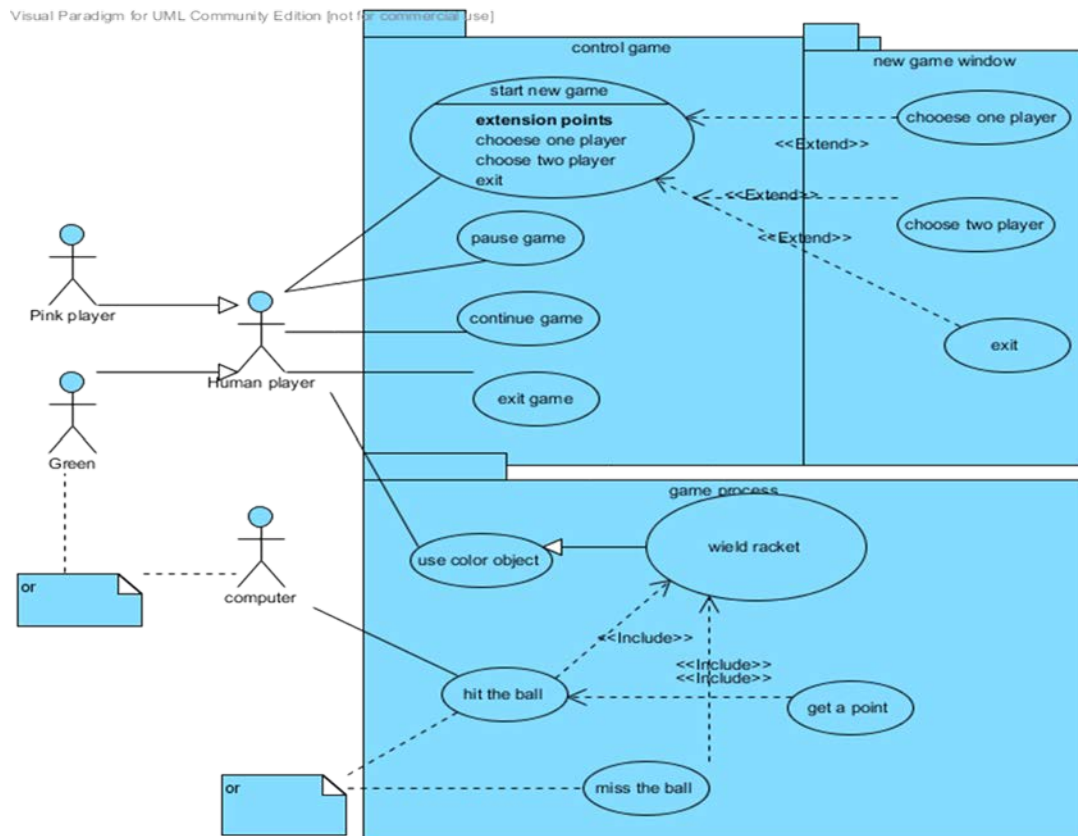
Στην παράγραφο αυτή θα δείξουμε το διάγραμμα ροής του παιχνιδιού πάνω στο οποίο στηρίχθηκε ο αλγόριθμος υλοποίησης του παιχνιδιού (βλ. εικόνα 36). Ο χρήστης ξεκινώντας το παιχνίδι στην 1<sup>η</sup> οθόνη συναντά τις παρακάτω επιλογές:

- 1 player (εκκίνηση του παιχνιδιού με έναν παίχτη)
- 2 player (εκκίνηση του παιχνιδιού με δύο παίχτες)
- Έξοδος από το παιχνίδι

Εφόσον επιλέξει ο χρήστης να παίξει το παιχνίδι, τότε υπάρχουνε τρεις διαδρομές :

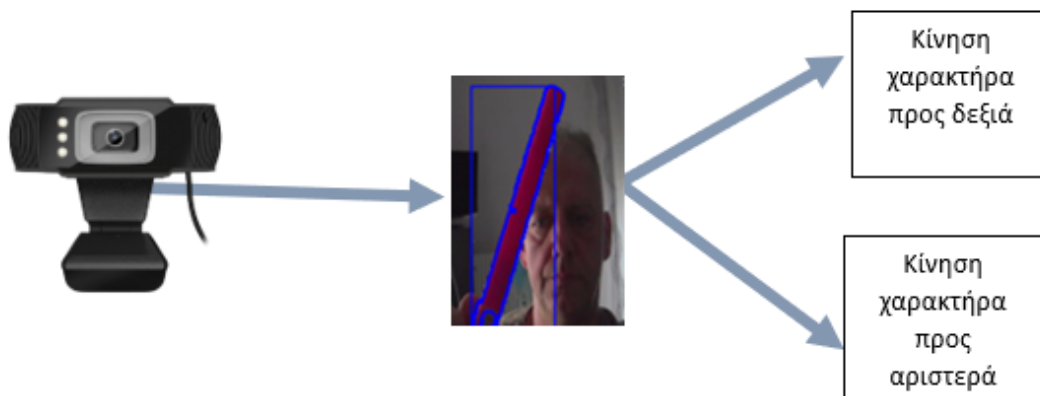
- Να παίξει και να κερδίσει το παιχνίδι

- Να παίξει και να χάσει
- Να δοκιμάσει ξανά το παιχνίδι



Εικόνα 36: Διάγραμμα ροής παιχνιδιού

Στην επόμενη εικόνα 37 βλέπουμε ότι ο χαρακτήρας του παιχνιδιού έχει την δυνατότητα να μετακινείται αριστερά ή δεξιά αρκεί ο παίχτης του παιχνιδιού μπροστά στην κάμερα να μετακινεί το χρωματιστό αντικείμενο προς την αντίστοιχη κατεύθυνση.



Εικόνα 37: Διάγραμμα κίνησης χαρακτήρα εξαρτώμενο από την κίνηση του αντικειμένου

## 6.6 Σύνδεση αλγορίθμων με το πρόγραμμα Unity μέσω πρωτόκολλου επικοινωνίας

### 6.6.1 Εισαγωγή

Σε αυτή την ενότητα θα αναλυθεί ο διάυλος επικοινωνίας μεταξύ των αλγορίθμων ανίχνευσης χειρονομιών και χρωματιστών αντικειμένων και του παιχνιδιού που δημιουργήσαμε στο πρόγραμμα Unity. Όπως προαναφέρθηκε η διαχείριση του μενού διεπαφής του παιχνιδιού αλλά και οι κινήσεις του ήρωα του παιχνιδιού γίνονται αποκλειστικά από την αναγνώριση των χρωματιστών αντικειμένων και των χειρονομιών, για να επιτευχθεί αυτό θα πρέπει να χρησιμοποιηθεί ένα πρωτόκολλο επικοινωνίας.

### 6.6.2 Πρωτόκολλα επικοινωνίας

Για να κατανοήσουμε πώς το σενάριο Python-OpenCV και η εφαρμογή στη Unity θα επικοινωνούν μεταξύ τους, είναι σημαντικό να γνωρίζουμε τις βασικές αρχές του τρόπου με τον οποίο τα δεδομένα μοιράζονται μεταξύ δύο ή περισσότερων εφαρμογών σε συστήματα υπολογιστών.

Τα πρωτόκολλα επικοινωνίας είναι επίσημες περιγραφές μορφών και κανόνων ψηφιακών μηνυμάτων. Απαιτούνται για να ανταλλάσσονται μηνύματα μεταξύ υπολογιστικών συστημάτων που είναι απαραίτητα στις τηλεπικοινωνίες.

Δύο κοινά χρησιμοποιούμενα πρωτόκολλα είναι το *TCP* (Transmission Control Protocol) και το *UDP* (User Datagram Protocol).

Το *TCP* είναι το κυρίαρχο πρωτόκολλο που χρησιμοποιείται για το μεγαλύτερο μέρος της συνδεσιμότητας στο Διαδίκτυο λόγω των υπηρεσιών που παρέχει για τη διάσπαση μεγάλων συνόλων δεδομένων σε μεμονωμένα πακέτα, τον έλεγχο και την εκ νέου αποστολή χαμένων πακέτων και τη συναρμολόγηση πακέτων στη σωστή σειρά - αποστολή και λήψη email (SMTP) και περιήγηση στο διαδίκτυο, για παράδειγμα.

Αντίθετα, οι εφαρμογές που επικοινωνούν χρησιμοποιώντας το πρωτόκολλο *UDP* απλώς στέλνουν τα πακέτα και δεν ελέγχουν ή περιμένουν μια επιβεβαίωση πριν στείλουν το επόμενο πακέτο, πράγμα που σημαίνει ότι έχουν πολύ χαμηλότερο εύρος ζώνης και δεν έχουν όμως χρονική καθυστέρηση.

Μερικές κοινές εφαρμογές του *UDP* στον πραγματικό κόσμο: Tunneling/VPN, υπηρεσίες ροής ζωντανών πολυμέσων και συνηθέστερα σε διαδικτυακά παιχνίδια για πολλούς παίκτες όπως το *Counter Strike:Global Offensive* ή το *DOTA 2* .

Στην περίπτωση μας επιλέχθηκε η επικοινωνία μέσω *UDP* μεταξύ της Python και του προγράμματος Unity .

### **6.6.3 Εκτέλεση του διακομιστή UDP σε Python**

Στο βασικό πρόγραμμα της Python κάνουμε *import* την βιβλιοθήκη *socket* ώστε όταν θα τρέξουμε τους αλγορίθμους αναγνώρισης χειρονομιών και χρωματιστών αντικειμένων να μπορέσουμε να κάνουμε αποστολή των δεδομένων του αλγορίθμου.

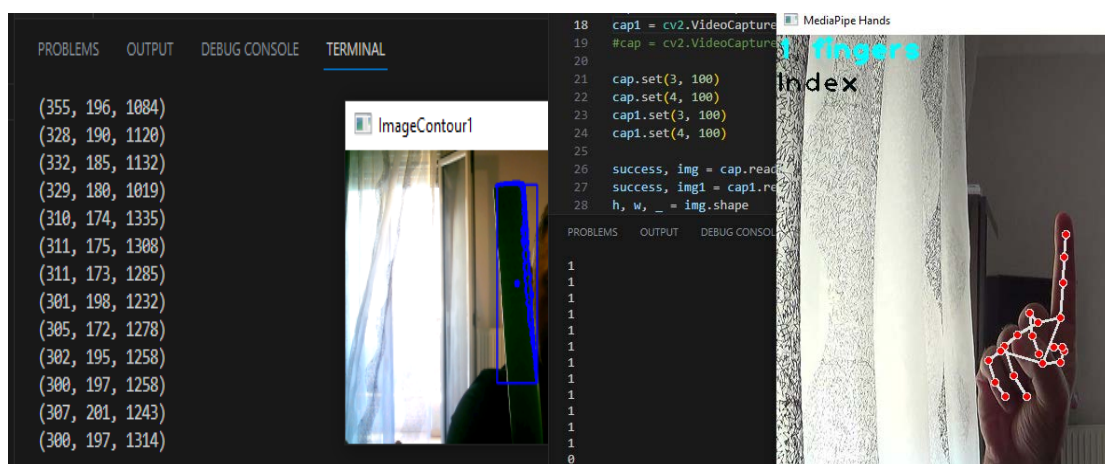
Τρέχουμε μέσω terminal το πρόγραμμα και τότε η webcam θα πρέπει να ενεργοποιηθεί και να ανοίξει τρία νέα παράθυρα: “hand recognition”, “pink color recognition”, “green color recognition”.

Μόλις αναγνωριστεί ο αριθμός των δακτύλων και τα αντίστοιχα χρωματιστά αντικείμενα η βιβλιοθήκη *socket* θα χρησιμοποιηθεί για την αποστολή των σχετικών δεδομένων μέσω *UDP* στις θύρες με αριθμούς:



- με την port 5052 θα αποστέλλονται τα δεδομένα αναγνώρισης του ροζ χρώματος.
- με την port 5053 θα αποστέλλονται τα δεδομένα αναγνώρισης του λαχανί χρώματος.
- με την port,5054 θα αποστέλλονται τα δεδομένα αναγνώρισης χειρονομιών.

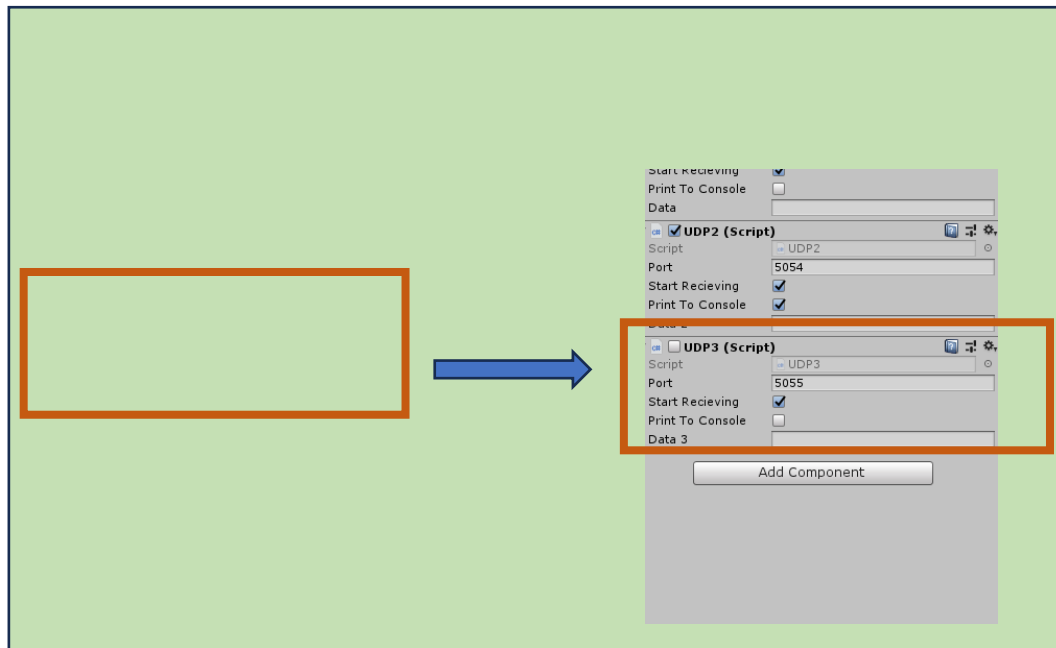
στην εικόνα 38 που ακολουθεί φαίνεται δείγμα δεδομένων που αποστέλλονται στην εφαρμογή Unity από την αναγνώριση αντικειμένων μέσω χρώματος αλλά και την αναγνώριση χειρονομιών .



Εικόνα 38: Δεδομένα αλγορίθμων

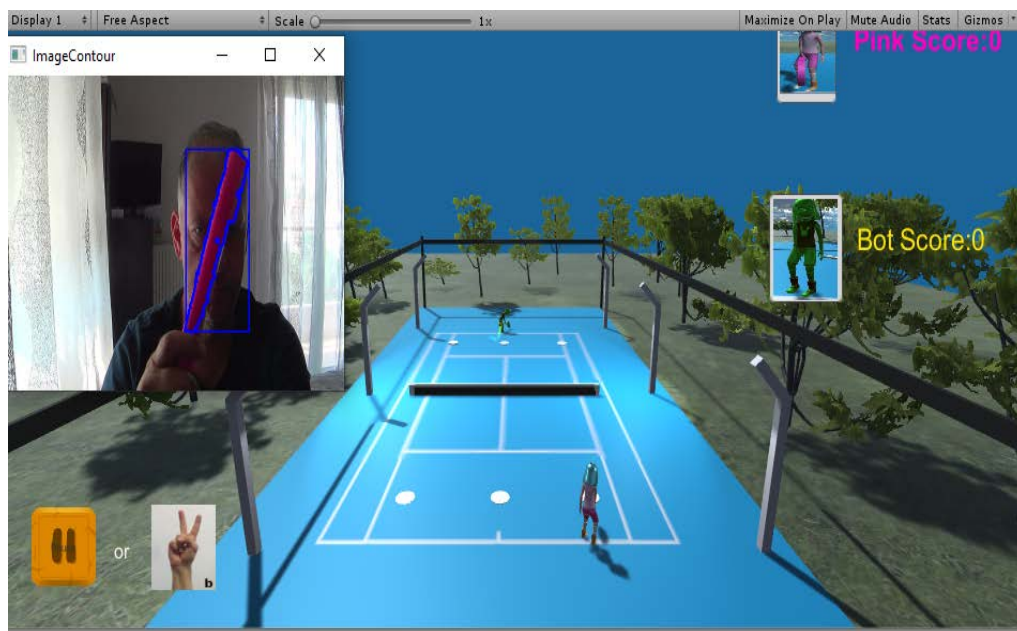
#### 6.6.4 Λήψη δεδομένων μέσω UDP στη Unity

Ανοίγοντας το έργο στην Unity δημιουργούμε ένα αντικείμενο όπου του προσθέτουμε ιδιότητες λήψης δεδομένων μέσω του πρωτοκόλλου UDP (βλ. εικόνα 39).Ο κώδικας για την λήψη δεδομένων στην Unity γράφτηκε σε γλώσσα προγραμματισμού C# που μας δίνει την δυνατότητα να λαμβάνει το συγκεκριμένο αντικείμενο τα αποτελέσματα των αλγορίθμων από την Python, μέσω του πρωτοκόλλου επικοινωνίας UDP.Τα δεδομένα που καταλήγουν στο αντικείμενο της Unity είναι αυτά που εμφανίζονται στην εικόνα 38 ,αυτά είναι δεδομένα από την αναγνώριση χειρονομιών αλλά και αντικειμένων.



Εικόνα 39: Προσθήκη ιδιοτήτων μέσω script για επικοινωνία με UDP

Προσθέτοντας τα κατάλληλα script στα διάφορα αντικείμενα που βρίσκονται στη σκηνή της Unity, όπως τα κουμπιά στην οθόνη διεπαφής και τον κεντρικό ήρωα του παιχνιδιού πετυχαίνουμε να παίξουμε το παιχνίδι μας μέσω της αναγνώρισης των αντικειμένων και των χειρονομιών (βλ. εικόνα 40). Το πρόγραμμα αναγνώρισης στην Python εκτελείται συνεχώς και ελέγχει αν υπάρχει στην πηγή εισόδου του κάποια χειρονομία ή κάποιο χρωματιστό αντικείμενο, αν ανιχνευτούν τα προηγούμενα τότε τα δεδομένα από την έξοδο του αλγορίθμου ανίχνευσης αποστέλλονται απευθείας μέσω του πρωτοκόλλου *UDP* στα αντίστοιχα αντικείμενα της Unity που έχουν οριστεί να κάνουν λήψη δεδομένων. Τα δεδομένα που λαμβάνουν τα αντικείμενα αποκωδικοποιούνται και προβάλλονται τελικά στην σκηνή της Unity.



Εικόνα 40: Play game

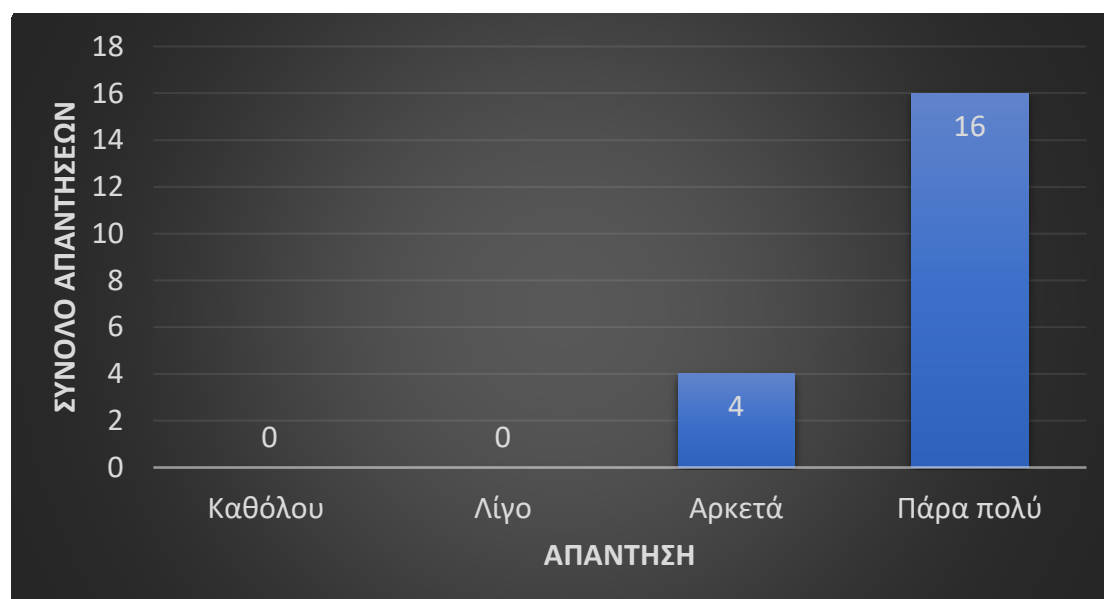
## Κεφάλαιο 7 Αξιολόγηση της εφαρμογής

### 7.1 Αξιολόγηση

Προκειμένου να κριθεί η αποτελεσματικότητα της εφαρμογής, διενεργήθηκε πιλοτική έρευνα που περιλάμβανε σε πρώτη φάση την δοκιμή του παιχνιδιού και σε δεύτερη φάση την απάντηση ερωτηματολογίου. Οι συμμετέχοντες της αξιολόγησης ήταν προπτυχιακοί φοιτητές του τμήματος Επικοινωνίας και Ψηφιακών Μέσων του Πανεπιστημίου Δυτικής Μακεδονίας. Στην αξιολόγηση της πιλοτικής έρευνας συμμετείχαν 20 άτομα, η ηλικία των συμμετεχόντων κυμαινόταν από 18 έως 24 έτη. Το φύλο των ατόμων ήταν Γυναίκες(6 άτομα) και Άντρες(14 άτομα) όλοι τους ασχολούταν με τα βιντεοπαιχνίδια. Η δοκιμή του παιχνιδιού έγινε σε ηλεκτρονικό υπολογιστή σε εργαστήριο του τμήματος.

Παρακάτω παρουσιάζονται τα αποτελέσματα που προέκυψαν από τις απαντήσεις των συμμετεχόντων στο ερωτηματολόγιο.

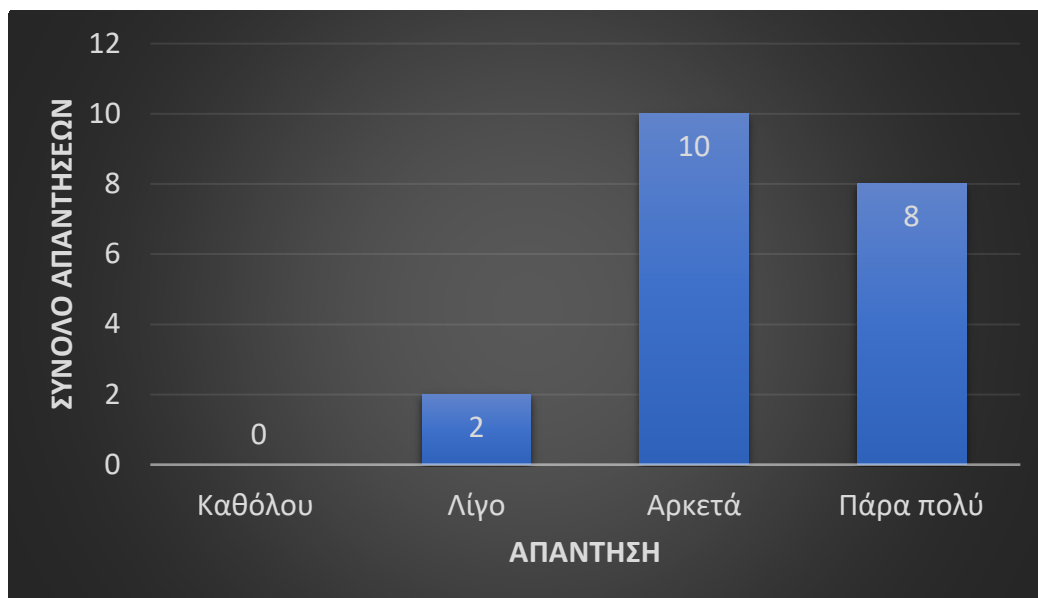
Ερώτηση 1<sup>η</sup>: *Το παιχνίδι ήταν ενδιαφέρον με το να κρατάς μια ρακέτα και να έχεις τον έλεγχο του παιχνιδιού;*



Εικόνα 41: Γράφημα απαντήσεων 1ης ερώτησης του ερωτηματολογίου

Από την παραπάνω εικόνα 41, φαίνεται πως οι συμμετέχοντες θεωρούν ότι το παιχνίδι ήταν ενδιαφέρον να κρατούν μια ρακέτα και να έχουν τον έλεγχο του παιχνιδιού (16 άτομα απάντησαν *Πάρα πολύ* και 4 άτομα *Αρκετά*).

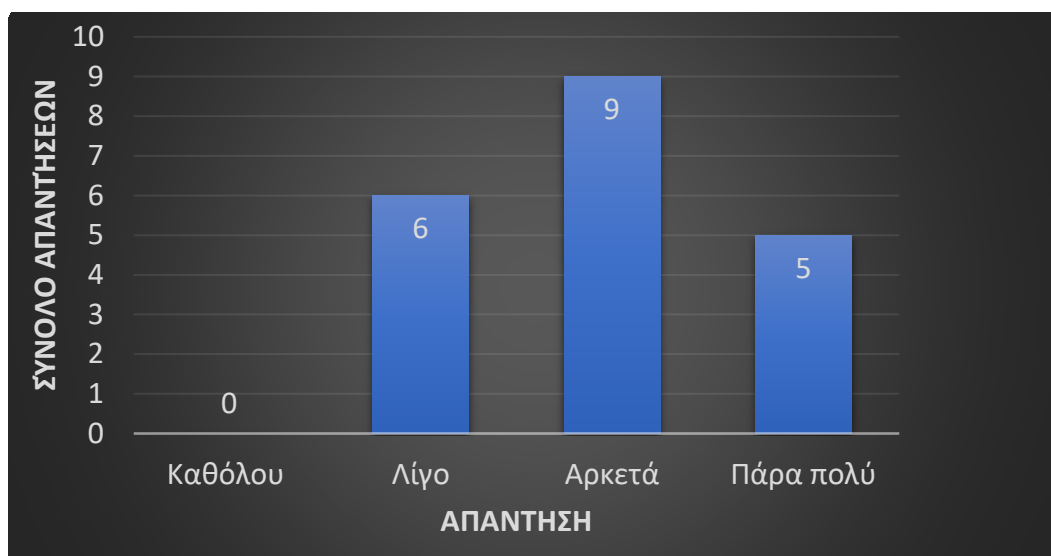
Ερώτηση 2<sup>η</sup>: *Το παιχνίδι ήταν εύχρηστο;*



Εικόνα 42: Γράφημα απαντήσεων 2ης ερώτησης του ερωτηματολογίου

Ενώ από την εικόνα 42 συμπεραίνουμε, ότι το παιχνίδι ήταν εύχρηστο για τους συμμετέχοντες (8 άτομα απάντησαν *Πάρα πολύ* και 10 άτομα *Αρκετά* ενώ 2 άτομα απάντησαν *Λίγο*).

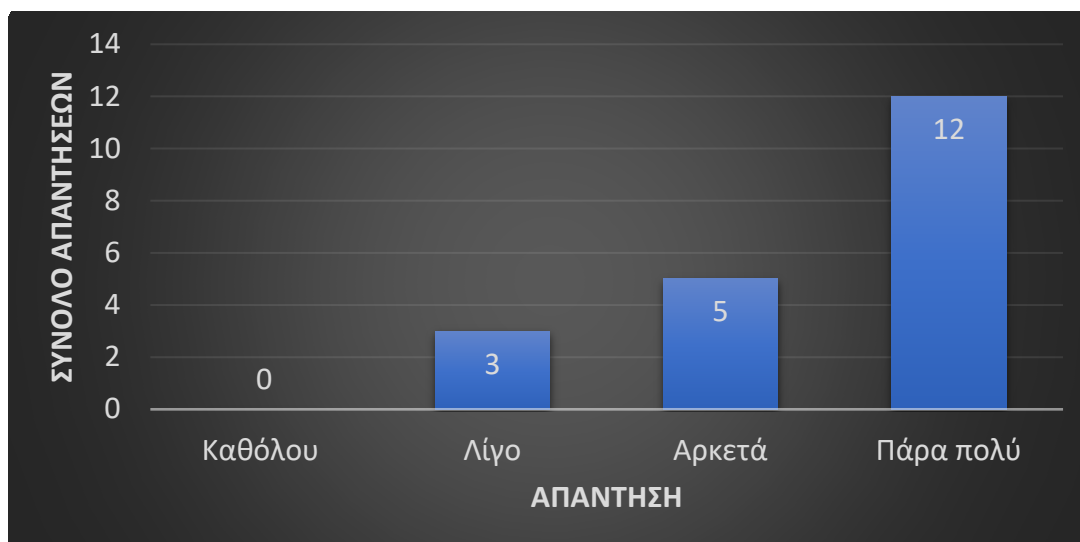
Ερώτηση 3<sup>η</sup>: *Υπήρχε ελευθερία κινήσεων κατά την διάρκεια του παιχνιδιού;*



Εικόνα 43: Γράφημα απαντήσεων 3ης ερώτησης του ερωτηματολογίου

Όσον αφορά, την ελευθερία κινήσεων κατά την διάρκεια του παιχνιδιού φαίνεται από την παραπάνω εικόνα (βλ. *Εικόνα 43*) ότι αισθάνθηκαν ελευθερία στις κινήσεις κατά την εξέλιξη του παιχνιδιού ( 9 άτομα απάντησαν *Αρκετά* και 5 άτομα *Πάρα πολύ* ενώ 6 άτομα *Λίγο*).

Ερώτηση 4<sup>η</sup>: Πιστεύετε ότι θα ήταν ενδιαφέρον να εφαρμοστεί ο συγκεκριμένος τρόπος χειρισμού σε tablet και mobile συσκευές;



*Εικόνα 44: Γράφημα απαντήσεων 4ης ερώτησης του ερωτηματολογίου*

Στην τελευταία ερώτηση, οι συμμετέχοντες απάντησαν (βλ. *Εικόνα 44*) ότι ο συγκεκριμένος τρόπος χειρισμού παιχνιδιού θα τους ενδιέφερε να εφαρμοστεί και σε άλλες συσκευές είτε mobile είτε tablet (12 άτομα απάντησαν *Πάρα πολύ* ,5 *Αρκετά* ενώ 3 άτομα απάντησαν *Λίγο*).

Τέλος 6 από τους συμμετέχοντες άφησαν σχόλια σχετικά με την ελευθερία κινήσεων ότι θα περιμέναν ο χαρακτήρας του παιχνιδιού πέρα από την κίνηση αριστερά-δεξιά να μπορεί να έχει κίνηση και εμπρός-πίσω.

## Κεφάλαιο 8 Συμπεράσματα

Αντικείμενο της διπλωματικής εργασίας αποτέλεσε η ανάπτυξη κατάλληλου λογισμικού για τον εντοπισμό ανθρώπινων χειρονομιών αλλά και αντικειμένων με βάση το χρωματικό χαρακτηριστικό τους και η σύνδεση των ανιχνεύσιμων δεδομένων με ένα παιχνίδι αγώνα τένις. Μελετήθηκαν διάφορες τεχνικές κατάτμησης εικόνας και επιλέχθηκε η τεχνική κατάτμησης με βάση το χρώμα για την ανίχνευση αντικειμένων, ενώ για την ανίχνευση ανθρώπινων χειρονομιών επιλέχθηκε το λογισμικό της MediaPipe. Η ανίχνευση αντικειμένων από μια συνεχόμενη ροή εικόνων ήταν επιτυχής, χωρίς ιδιαίτερες απαιτήσεις φωτισμού .

Επίσης η επιλογή του πρωτοκόλλου επικοινωνίας *UDP*, για την αποστολή δεδομένων προς το παιχνίδι, στην πράξη αποδείχθηκε επαρκής διότι δεν υπήρχαν χρονικές καθυστερήσεις. Η ανταπόκριση του χαρακτήρα του παιχνιδιού στις ανιχνεύσιμες κινήσεις του χρωματιστού αντικειμένου ήταν άμεσες και χωρίς καθυστερήσεις. Ενώ η γρήγορη ανίχνευση των ανθρώπινων χειρονομιών είχε το ίδιο θετικά αποτελέσματα στο να διαχειριστούμε την οθόνη διεπαφής του παιχνιδιού .

Η διαχείριση του παιχνιδιού συμπερασματικά σύμφωνα και με τις απαντήσεις του ερωτηματολογίου που είδαμε στην *ενότητα 7.1* μπορεί να υλοποιείται εξ' ολοκλήρου μέσω της υπολογιστικής όρασης χωρίς χρονικές καθυστερήσεις, δίνοντας στον παίχτη του βιντεοπαιχνιδιού ελευθερία σκελετικών κινήσεων αυξάνοντας την αίσθηση εμπειρίας του. Η εφαρμογή της υπολογιστικής όρασης στο παιχνίδι κατάφερε να δώσει περισσότερη ρεαλιστικότητα και μεγαλύτερη ελευθερία κινήσεων.

Το παιχνίδι δοκιμάστηκε σε μια οθόνη ηλεκτρονικού υπολογιστή, ίσως η αίσθηση να ήταν πιο πειστική σε μια μεγαλύτερη οθόνη προβολής. Επίσης όπως αναφέρθηκε χρησιμοποιήθηκε μια απλή κάμερα υπολογιστή με πολύ καλά αποτελέσματα με μόνο περιορισμό ίσως την απόσταση του χρήστη από την κάμερα για το λόγο αυτό έγινε δοκιμή της εφαρμογής και με την χρήση μίας βιντεοκάμερας με απόρροια να αυξηθεί ο

χώρος ελευθερίας κινήσεων του παίχτη. Επίσης σε επόμενη έκδοση της εφαρμογής θα μπορούσε να προστεθεί και μια επιπλέον κίνηση του χαρακτήρα κατά τον άξονα Z.

Συμπερασματικά ο απώτερος στόχος αυτής της εργασίας έχει επιτευχθεί, δημιούργησε μια πιο φυσική μορφή ελέγχου που επιτρέπει στους παίκτες να επικεντρωθούν στον ενθουσιασμό του αγώνα χωρίς να ανησυχούν για το πάτημα των κουμπιών ή τις κινήσεις των χειριστηρίων του παιχνιδιού. Η χρήση της υπολογιστικής όρασης μπορεί να χρησιμοποιηθεί στην αναπτυσσόμενη βιομηχανία των παιχνιδιών με πολλά προσδοκώμενα οφέλη για το άμεσο μέλλον .



## Βιβλιογραφία

- [1] S. Albawi, T. A. Mohammed, και S. Al-Zawi, ‘Understanding of a convolutional neural network’, *2017 International Conference on Engineering and Technology (ICET)*, December 2017, pp. 1–6. doi: [10.1109/ICEngTechnol.2017.8308186](https://doi.org/10.1109/ICEngTechnol.2017.8308186).
- [2] M. H. Amir, A. Quek, N. R. B. Sulaiman, και J. See, ‘DUKE: Enhancing Virtual Reality based FPS Game with Full-body Interactions’, στο *Proceedings of the 13th International Conference on Advances in Computer Entertainment Technology*, στο ACE ’16. New York, NY, USA: Association for Computing Machinery, August 2016, pp. 1–6. doi: [10.1145/3001773.3001804](https://doi.org/10.1145/3001773.3001804).
- [3] L. Chen, F. Wang, H. Deng, και K. Ji, ‘A Survey on Hand Gesture Recognition’, στο *2013 International Conference on Computer Sciences and Applications*, September 2013, pp. 313–316. doi: [10.1109/CSA.2013.79](https://doi.org/10.1109/CSA.2013.79).
- [4] H.-Y. Chung, Y.-L. Chung, και W.-F. Tsai, ‘An Efficient Hand Gesture Recognition System Based on Deep CNN’, στο *2019 IEEE International Conference on Industrial Technology (ICIT)*, October 2019, pp. 853–858. doi: [10.1109/ICIT.2019.8755038](https://doi.org/10.1109/ICIT.2019.8755038).
- [5] J. Galván-Ruiz, C. M. Travieso-González, A. Tejera-Fetmilch, A. Pinan-Roescher, L. Esteban-Hernández, και L. Domínguez-Quintana, ‘Perspective and Evolution of Gesture Recognition for Sign Language: A Review’, *Sensors*, Art. no. 12, January 2020, doi: [10.3390/s20123571](https://doi.org/10.3390/s20123571).
- [6] J. Han, L. Shao, D. Xu, και J. Shotton, ‘Enhanced Computer Vision With Microsoft Kinect Sensor: A Review’, *IEEE Transactions on Cybernetics*, vol. 43, no. 5, pp. 1318–1334, July 2013, doi: [10.1109/TCYB.2013.2265378](https://doi.org/10.1109/TCYB.2013.2265378).
- [7] A. Kapitanov, A. Makhlyarchuk, και K. Kvanchiani, ‘HaGRID - HAnd Gesture Recognition Image Dataset’. arXiv, 16 June 2022. doi: [10.48550/arXiv.2206.08219](https://doi.org/10.48550/arXiv.2206.08219).
- [8] N. V. Le, M. Qarmout, Y. Zhang, H. Zhou, και C. Yang, ‘Hand Gesture Recognition System for Games’, στο *2021 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE)*, September 2021, pp.1–6. doi: [10.1109/CSDE53843.2021.9718421](https://doi.org/10.1109/CSDE53843.2021.9718421).
- [9] J. Liu και M. Kavakli, ‘A survey of speech-hand gesture recognition for the development of multimodal interfaces in computer games’, στο *2010 IEEE International Conference on Multimedia and Expo*, July 2010, pp. 1564–1569. doi: [10.1109/ICME.2010.5583252](https://doi.org/10.1109/ICME.2010.5583252).

- [10] M. A. Livingston, J. Sebastian, Z. Ai, και J. W. Decker, 'Performance measurements for the Microsoft Kinect skeleton', στο *2012 IEEE Virtual Reality Workshops (VRW)*, March 2012, pp. 119–120. doi: [10.1109/VR.2012.6180911](https://doi.org/10.1109/VR.2012.6180911).
- [11] B. Ma, W. Xu, και S. Wang, 'A Robot Control System Based on Gesture Recognition Using Kinect', *TELKOMNIKA Indonesian Journal of Electrical Engineering*, vol 11, no. 5, Art. no. 5, May 2013.
- [12] R. Ma, Z. Zhang, και E. Chen, 'Human Motion Gesture Recognition Based on Computer Vision', *Complexity*, vol. 2021, February 2021, doi: [10.1155/2021/6679746](https://doi.org/10.1155/2021/6679746).
- [13] 'Applied Sciences | Free Full-Text | Real-Time Hand Gesture Recognition Based on Deep Learning YOLOv3 Model'. <https://www.mdpi.com/2076-3417/11/9/4164> (access date 1 September 2023).
- [14] K. Norman και J. Kirakowski, *The Wiley Handbook of Human Computer Interaction Set*. John Wiley & Sons, 2017.
- [15] E. Ohn-Bar και M. M. Trivedi, 'Hand Gesture Recognition in Real Time for Automotive Interfaces: A Multimodal Vision-Based Approach and Evaluations', *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 6, pp. 2368–2377, September 2014, doi: [10.1109/TITS.2014.2337331](https://doi.org/10.1109/TITS.2014.2337331).
- [16] K. O'Shea και R. Nash, 'An Introduction to Convolutional Neural Networks'. arXiv, 2 December 2015. doi: [10.48550/arXiv.1511.08458](https://doi.org/10.48550/arXiv.1511.08458).
- [17] M. Oudah, A. Al-Naji, και J. Chahl, 'Hand Gesture Recognition Based on Computer Vision: A Review of Techniques', *Journal of Imaging*, vol. 6, no. 8, Art. no. 8, August 2020, doi: [10.3390/jimaging6080073](https://doi.org/10.3390/jimaging6080073).
- [18] S. S. Rautaray και A. Agrawal, 'Vision based hand gesture recognition for human computer interaction: a survey', *Artif Intell Rev*, vol 43, no 1, pp. 1–54, January 2015, doi: [10.1007/s10462-012-9356-9](https://doi.org/10.1007/s10462-012-9356-9).
- [19] C. Saha, D. Goswami, S. Saha, A. Konar, A. Lekova, και A. K. Nagar, 'A novel gesture driven fuzzy interface system for car racing game', στο *2015 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, December 2015, pp. 1–8. doi: [10.1109/FUZZ-IEEE.2015.7337954](https://doi.org/10.1109/FUZZ-IEEE.2015.7337954).
- [20] J. Suarez και R. R. Murphy, 'Hand gesture recognition with depth images: A review', στο *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication*, September 2012, pp. 411–417. doi: [10.1109/ROMAN.2012.6343787](https://doi.org/10.1109/ROMAN.2012.6343787).
- [21] X. Y. Wu, 'A hand gesture recognition algorithm based on DC-CNN', *Multimed Tools Appl*, vol. 79, no. 13, pp. 9193–9205, April 2020, doi: [10.1007/s11042-019-7193-4](https://doi.org/10.1007/s11042-019-7193-4).

- [22] D. Xu, ‘A Neural Network Approach for Hand Gesture Recognition in Virtual Reality Driving Training System of SPG’, στο *18th International Conference on Pattern Recognition (ICPR’06)*, December 2006, pp. 519–522. doi: [10.1109/ICPR.2006.109](https://doi.org/10.1109/ICPR.2006.109).
- [23] M. Yasen και S. Jusoh, ‘A systematic review on hand gesture recognition techniques, challenges and applications’, *PeerJ Comput. Sci.*, vol. 5 September 2019, doi: [10.7717/peerj-cs.218](https://doi.org/10.7717/peerj-cs.218).
- [24] M. Goesele, N. Snavely, B. Curless, H. Hoppe, και S. M. Seitz, ‘Multi-View Stereo for Community Photo Collections’, *2007 IEEE 11th International Conference on Computer Vision*, July 2007, pp. 1–8. doi: [10.1109/ICCV.2007.4408933](https://doi.org/10.1109/ICCV.2007.4408933).
- [25] A. Saxena, J. Driemeyer, και A. Y. Ng, ‘Robotic Grasping of Novel Objects using Vision’, *The International Journal of Robotics Research*, vol. 27, no. 2, pp. 157–173, February 2008, doi: [10.1177/0278364907087172](https://doi.org/10.1177/0278364907087172).
- [26] C. Lugaresi, ‘MediaPipe: A Framework for Perceiving and Processing Reality’.
- [27] F. Zhang κ.ά., ‘MediaPipe Hands: On-device Real-time Hand Tracking’. arXiv, 17 Juny 2020. doi: [10.48550/arXiv.2006.10214](https://doi.org/10.48550/arXiv.2006.10214).
- [28] ‘On-Device, Real-Time Hand Tracking with MediaPipe – Google Research Blog’. <https://blog.research.google/2019/08/on-device-real-time-hand-tracking-with.html> (access date 1 September 2023).
- [29] J. Canny, ‘A Computational Approach to Edge Detection’, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-8, no. 6, pp. 679–698, August 1986, doi: [10.1109/TPAMI.1986.4767851](https://doi.org/10.1109/TPAMI.1986.4767851).
- [30] R. M. Haralick, S. R. Sternberg, και X. Zhuang, ‘Image Analysis Using Mathematical Morphology’, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-9, no. 4, pp. 532–550, July 1987, doi: [10.1109/TPAMI.1987.4767941](https://doi.org/10.1109/TPAMI.1987.4767941).
- [31] Q. Liu και G. Peng, ‘A robust skin color based face detection algorithm’, at *2010 2nd International Asia Conference on Informatics in Control, Automation and Robotics (CAR 2010)*, March 2010, pp. 525–528. doi: [10.1109/CAR.2010.5456614](https://doi.org/10.1109/CAR.2010.5456614).
- [32] MediaPipe: On-Device, Real Time Hand Tracking, In <https://ai.googleblog.com/2019/08/on-device-real-time-hand-tracking-with.html>. 2019. Access 2021.
- [33] MediaPipe Github: <https://google.github.io/mediapipe/solutions/hands>. Access 2021.

## Δήλωση Πνευματικών Δικαιωμάτων

Δηλώνω ρητά ότι, σύμφωνα με το άρθρο 8 του Ν. 1599/1986 και τα άρθρα 2,4,6 παρ. 3 του Ν. 1256/1982, η παρούσα Μεταπτυχιακή Διπλωματική Εργασία με τίτλο:

«Ανάπτυξη ενός διαδραστικού παιχνιδιού εικονικής πραγματικότητας βασισμένο στην αναγνώριση κινήσεων του χρήστη»

καθώς και τα ηλεκτρονικά αρχεία και πηγαίοι κώδικες που αναπτύχθηκαν ή τροποποιήθηκαν στα πλαίσια αυτής της εργασίας και αναφέρονται ρητώς μέσα στο κείμενο που συνοδεύουν, και η οποία εκπονηθεί στο Πρόγραμμα Μεταπτυχιακών Σπουδών «Ανάπτυξη Ψηφιακών Παιχνιδιών και Πολυμεσικών Εφαρμογών» του Τμήματος Επικοινωνίας & Ψηφιακών Μέσων του Πανεπιστημίου Δυτικής Μακεδονίας, υπό την επίβλεψη του Επίκουρου Καθηγητή Βρίγκα Μιχάλη του Τμήματος Επικοινωνίας και Ψηφιακών Μέσων του Πανεπιστημίου Δυτικής Μακεδονίας αποτελεί αποκλειστικά προϊόν προσωπικής εργασίας και δεν προσβάλλει κάθε μορφής πνευματικά δικαιώματα τρίτων και δεν είναι προϊόν μερικής ή αντιγραφής, οι πηγές δε που χρησιμοποιήθηκαν περιορίζονται στις βιβλιογραφικές αναφορές και μόνον. Τα σημεία όπου έχω χρησιμοποιήσει ιδέες, κείμενο, αρχεία ή / και πηγές άλλων συγγραφέων, αναφέρονται ευδιάκριτα στο κείμενο με την κατάλληλη παραπομπή και ή σχετική αναφορά περιλαμβάνεται στο τμήμα των βιβλιογραφικών αναφορών με πλήρη περιγραφή.

Απαγορεύεται ή αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται ή ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα. Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και μόνο.

Copyright (C) Ονοματεπώνυμο Φοιτητή & Επιβλέποντα, Έτος, Πόλη

Copyright (C) ΓΚΟΥΖΙΟΣ ΣΤΕΦΑΝΟΣ, ΒΡΙΓΚΑΣ ΜΙΧΑΛΗΣ, 2024, ΚΑΣΤΟΡΙΑ

**Υπογραφή Φοιτητή**