



Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης
Πολυτεχνική Σχολή
Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών

Διπλωματική εργασία

με θέμα

**Υλοποίηση συστήματος
Παρακολούθησης κίνησης οφθαλμού
με χρήση συμβατικών συσκευών
οπτικής καταγραφής**

Ευθύμιος Χριστόφορος 8945

Θεσσαλονίκη, Φεβρουάριος 2022

Περίληψη

Η χρησιμότητα και οι εφαρμογές συστημάτων παρακολούθησης και ιχνηλάτησης οφθαλμών έχουν αυξηθεί σημαντικά τα τελευταία χρόνια. Οι περιπτώσεις χρήσης τους εκτείνονται σε πεδία όπως η ιατρική, η ψυχολογία, το μάρκετινγκ και η ψυχαγωγία. Παρά την αυξανόμενη ζήτησή τους, οι υπάρχουσες επαγγελματικές συσκευές έχουν ιδιαίτερα υψηλό κόστος, τόσο λόγω του εξειδικευμένου εξοπλισμού που χρησιμοποιούν, όσο λόγω του ολιγοπωλίου που επικρατεί στην συγκεκριμένη αγορά. Ταυτόχρονα, οι πιο απλές υλοποιήσεις, οι οποίες αξιοποιούν μόνο μια απλή συσκευή οπτικής καταγραφής, υποφέρουν από μειωμένη ακρίβεια και περιορισμένες επιτρεπτές συνθήκες λειτουργίας.

Στην παρούσα εργασία επιχειρήσαμε τη δημιουργία ενός συστήματος ιχνηλάτησης βλέμματος, το οποίο αποτελείται πλήρως από λογισμικό και αξιοποιεί ήδη υπάρχον υλικό καταγραφής όπως μία συμβατική διαδικτυακή κάμερα. Στην προσπάθεια αυτή, στόχος μας ήταν να πετύχουμε όσο το δυνατόν καλύτερη ακρίβεια, έχοντας, όμως, ως προτεραιότητα την ελαστικότητα, όσον αφορά την ελευθερία κινήσεων του χρήστη και τις συνθήκες λειτουργίας της συσκευής. Έγινε, λοιπόν, υλοποίηση αλγορίθμου ο οποίος χρησιμοποιεί γεωμετρικές μεθόδους για τον υπολογισμό της θέσης του προσώπου, των ματιών και του σημείου στο οποίο το βλέμμα του χρήστη συναντά την οθόνη. Στη συνέχεια, έγινε προσαρμογή του αλγορίθμου σε πρόγραμμα υπολογιστή, το οποίο επίσης υλοποιήσαμε, έτσι ώστε να έχουμε δυνατότητα παρατήρησης διαγνωστικών στοιχείων του αλγορίθμου (όπως τα υπολογισμένα σημεία ενδιαφέροντος), δυνατότητα βαθμονόμησης, καταγραφής της εξόδου του συστήματος και αξιολόγησης της απόδοσης του αλγορίθμου. Για την επίδειξη της χρησιμότητας του συστήματος που υλοποιήσαμε, αλλά και για την ελεύθερη χρήση του από το κοινό, αναπτύξαμε ιστότοπο, ο οποίος επιτρέπει στον κάθε χρήστη να έχει πρόσβαση στο σύστημά μας και να πραγματοποιήσει καταγραφές της εξόδου του σε παρατηρήσεις εικόνων που επιλέγει ο ίδιος.

Τα αποτελέσματα της αξιολόγησής του αλγορίθμου μας έδωσαν μια καλή ιδέα για την απόδοσή του σε διάφορες συνθήκες λειτουργίας. Σε σύγκριση με αντίστοιχες μετρικές επαγγελματικής συσκευής, πετύχαμε μειωμένη ακρίβεια στις περισσότερες των περιπτώσεων, γεγονός το οποίο ήταν αναμενόμενο, εφόσον η χρήση εξειδικευμένων αισθητήρων της επαγγελματικής συσκευής αποτελεί σημαντικό πλεονέκτημα. Παρά το γεγονός αυτό, η απόδοση του συστήματός μας προέκυψε αρκετά υψηλή σε όλες τις συνθήκες λειτουργίας, ώστε η έξοδός του να θεωρηθεί σεβαστά ορθή και χρήσιμη. Με την παρούσα εργασία, λοιπόν, έχουμε πετύχει την δημιουργία και την διανομή μιας επαρκώς αξιόπιστης λύσης ιχνηλάτησης οφθαλμών, η οποία εστιάζει στην ευκολία χρήσης, στη συμβατότητα με όσο το δυνατόν περισσότερα συστήματα και στην ελαστικότητα των συνθηκών λειτουργίας.

Abstract

Επιτομή της εργασίας στα αγγλικά (προαιρετικά)

Πίνακας Περιεχομένων

Περίληψη.....	2
Abstract	3
Πίνακας Περιεχομένων	4
Εισαγωγή.....	6
1. Θεώρηση του προβλήματος.....	7
1.1 Αναγκαιότητα - περιπτώσεις χρήσης.....	7
1.2 Κόστος εμπορικών συστημάτων	8
1.3 Περιορισμοί υλοποιήσεων με απλή κάμερα	8
2. Θεωρητικό υπόβαθρο (και προσέγγιση)	10
2.1 Βιβλιογραφική επισκόπηση	10
2.2 Η διαδικασία εντοπισμού και υπολογισμού βλέμματος	12
2.2.1 Εντοπισμός προσώπου	13
2.2.2 Εντοπισμός των οφθαλμών	14
2.2.3 Μέθοδοι υπολογισμού βλέμματος	15
3. Προτεινόμενη προσέγγιση - Μεθοδολογία	17
3.1 Η προτεινόμενη προσέγγιση - προδιαγραφές.....	17
3.2 Εξάλειψη διαφορών εισόδου	17
3.3 Μεθοδολογία	17
3.3.1 Εύρεση κέντρου των ματιών	19
3.3.1.1 Ομογραφία επιπέδου	19
3.3.1.2 Άμεσος γραμμικός μετασχηματισμός	21
3.3.2 Εύρεση σημείου βλέμματος στην οθόνη.....	22
4. Υλοποίηση συστήματος	23
4.1 Περιγραφή εφαρμογής συστήματος.....	23
4.2 Εργαλεία ανάπτυξης	23
4.3 Υλοποίηση αλγορίθμου – εμφάνιση υπολογισμών	26
4.4 Βαθμονόμηση	29
4.5 Διαδικασία καταγραφής	31
4.6 Δημιουργία οπτικοποιήσεων	31
4.7 Διαδικασία αξιολόγησης	33
4.7.1 Μετρικές	33
4.7.2 Συνθήκες-στάδια αξιολόγησης.....	34
4.7.2.1 Ιδανικές συνθήκες.....	34
4.7.2.2 Ακραία σημεία οθόνης	35

4.7.2.3 Μεταβαλλόμενες συνθήκες φωτισμού	36
4.7.2.4 Μεταβαλλόμενη θέση προσώπου	37
4.7.3 Αποτελέσματα αξιολόγησης	37
4.7.3.1 Ιδανικές συνθήκες	38
4.7.3.2 Ακραία σημεία οθόνης	38
4.7.3.3 Συνθήκες χαμηλού φωτισμού	38
4.7.3.4 Συνθήκες έντονου φωτισμού	39
4.7.3.5 Μεταβαλλόμενη θέση προσώπου (άξονας X)	39
4.7.3.6 Μεταβαλλόμενη θέση προσώπου (άξονας Y)	40
4.7.3.7 Μεταβαλλόμενη θέση προσώπου (άξονας Z)	40
4.7.3.8 Σύγκριση αποτελεσμάτων με επαγγελματική συσκευή	41
5. Αξιοποίηση του συστήματος σε διαδικτυακή σελίδα	51
5.1 Βασική λειτουργία της ιστοσελίδας	51
5.2 Αναλυτική περιγραφή	51
5.2.1 Αρχική σελίδα	52
5.2.2 Σελίδα βαθμονόμησης	52
5.2.3 Σελίδα καταγραφής	53
5.2.4 Σελίδα αποτελεσμάτων / οπτικοποιήσεων	55
6. Συμπεράσματα	59
6.1 Αποτελέσματα - Συμπεράσματα	59
6.2 Καινοτομία και συνεισφορά της εργασίας	59
6.3 Μελλοντικές εξελίξεις	60
Βιβλιογραφικές αναφορές	62
Λίστα Εικόνων	64
Λίστα Πινάκων	65
Λίστα Σχημάτων	66

Εισαγωγή

Η όραση αποτελεί την πρωταρχική αίσθηση του ανθρώπου η οποία του επιτρέπει να αντιλαμβάνεται και να καταλαβαίνει τον χώρο τριγύρω του. Βοηθάει στην πλοήγησή του στον κόσμο, στην αναγνώριση αντικειμένων και άλλων ανθρώπων, και στην αλληλεπίδραση με το περιβάλλον του. Η όραση είναι επίσης σημαντική στην επικοινωνία, καθώς μας επιτρέπει να διαβάζουμε και να ερμηνεύουμε τις εκφράσεις και τη στάση του σώματος άλλων ανθρώπων. Επιπροσθέτως, η όραση παίζει εξαιρετικά σημαντικό ρόλο σε διάφορες δραστηριότητες όπως η οδήγηση, η άθληση και η καλλιτεχνική αναζήτηση. Συνολικά, είναι ένα απαραίτητο εργαλείο στην καθημερινότητα του ανθρώπου και έχει σημαντικό αντίκτυπο στην ευημερία του.

Η ιχνηλάτηση των οφθαλμών αποτελεί μία τεχνολογία η οποία βρίσκει όλο και περισσότερες εφαρμογές τα τελευταία χρόνια. Επιτρέπει σε ερευνητές να μελετήσουν τον τρόπο με τον οποίο οι άνθρωποι παρατηρούν και αλληλεπιδρούν με διάφορων ειδών πολυμέσα, όπως ιστοσελίδες, διαφημίσεις και βιντεοπαιχνίδια. Οι συσκευές ιχνηλάτησης μπορούν να παρέχουν λεπτομερείς πληροφορίες για το που κοιτάει ο παρατηρητής κατά την παρακολούθηση των παραπάνω πολυμέσων. Τέτοιες πληροφορίες είναι χρήσιμες για την κατανόηση του πώς οι άνθρωποι επεξεργάζονται και αντιλαμβάνονται πληροφορίες και πώς παίρνουν αποφάσεις.

Συσκευές ιχνηλάτησης οφθαλμών χρησιμοποιούνται επίσης και σε πληθώρα άλλων εφαρμογών, όπως η έρευνα marketing, ο έλεγχος χρησιμότητας και διάφορα πειράματα ψυχολογίας. Μπορούν επίσης να βοηθήσουν στην κατανόηση του τρόπου με τον οποίο οι χρήστες αντιλαμβάνονται περιβάλλοντα διεπαφής και αλληλεπιδρούν με αυτά. Επίσης, τέτοιες συσκευές χρησιμοποιούνται και σε τομείς όπως η ιατρική, καθώς μπορούν να βοηθήσουν στη διάγνωση και στη θεραπεία προβλημάτων όρασης και άλλων προβλημάτων υγείας.

Η παρούσα διπλωματική εργασία, όπως δηλώνει και ο τίτλος, αφορά την ανάπτυξη λογισμικού το οποίο με μία απλή συμβατική συσκευή καταγραφής εικόνας, όπως μια διαδικτυακή κάμερα, θα μπορεί να ανιχνεύσει το βλέμμα του χρήστη και να αποτυπώσει στην οθόνη το σημείο στο οποίο κοιτάει με επαρκή ακρίβεια.

Στόχος μας είναι η ευκολία και η δυνατότητα χρήσης σε όσο το δυνατόν πιο αντίξοες συνθήκες λειτουργίας, ώστε να μπορεί να αξιοποιηθεί σε οποιαδήποτε από τις πιθανές περιπτώσεις χρήσης όπως αυτές που αναφέρθηκαν παραπάνω ή και άλλες οι οποίες θα αναλυθούν σε επόμενα κεφάλαια.

Κεφάλαιο 1^ο

1. Θεώρηση του προβλήματος

1.1 Αναγκαιότητα - περιπτώσεις χρήσης

Οι τεχνολογίες εντοπισμού και ιχνηλάτισης οφθαλμού έχουν τη δυνατότητα να επηρεάσει σε μεγάλο βαθμό ένα ευρύ φάσμα επιστημονικών πεδίων, συμπεριλαμβανομένης της διαφόρων ειδών έρευνας, της ιατρικής, της υποστηρικτικής τεχνολογίας, των βιντεοπαιχνιδιών και της αλληλεπίδρασης ανθρώπου-υπολογιστή. Συγκεκριμένα, η ιχνηλάτιση οφθαλμού μπορεί να παίξει σημαντικό ρόλο στους εξής τομείς:

- **Ακαδημαϊκή και επιστημονική έρευνα.** Δίχως αμφιβολία, ο τομέας ακαδημαϊκής και επιστημονικής έρευνας ωφελείται περισσότερο από τις τεχνολογίες eye tracking. Τέτοιες τεχνολογίες χρησιμοποιούνται σε περιβάλλοντα ηλεκτρονικής μάθησης (e-learning) και επιτρέπουν την καταγραφή της συμπεριφοράς του μαθητή σε πραγματικό χρόνο, γεγονός το οποίο είναι κρίσιμο για την ανάλυση της αποτελεσματικότητας τέτοιων περιβαλλόντων. Οι συσκευές eye tracking μας δίνουν πληροφορίες για τον βαθμό του ενδιαφέροντος, την ικανότητα προσοχής και εστίασης που επιδεικνύει ο χρήστης. Μπορούμε επίσης, μέσω της μελέτης των σημείων εστίασης, της διάρκειας και της θέσης των σημείων αυτών και του ρυθμού βλεφαρίσματος, να εξαγάγουμε συμπεράσματα για το επίπεδο προσοχής, άγχους, χαλάρωσης, αποδοτικότητας, κούρασης ή και διαφόρων συναισθημάτων του χρήστη. Γενικά, οι συσκευές ιχνηλάτισης βλέμματος αποτελεί εξαιρετικό εργαλείο για τη μελέτη της οπτικής προσοχής και κοινωνικής αλληλεπίδρασης σε όλα τα περιβάλλοντα μάθησης.
- **Ιατρική έρευνα.** Η ιχνηλάτιση βλέμματος σε συνδυασμό με συμβατικές μεθόδους έρευνας μπορεί να χρησιμοποιηθεί για τη διάγνωση ασθενειών όπως η Διαταραχή Ελλειμματικής Προσοχής και Υπερκινητικότητας (ΔΕΠΥ), η Διαταραχή Αυτιστικού Φάσματος (ΔΑΦ), η Ιδιοψυχαναγκαστική Διαταραχή (ΙΨΔ), η Σχιζοφρένεια, η νόσος του Parkinson και του Alzheimer. Σε διάφορα έργα πειραματικής ψυχολογίας και κλινικής νευροεπιστήμης έχει χρησιμοποιηθεί φορετός εξοπλισμός ιχνηλάτισης οφθαλμών για να οριστεί η σύνδεση μεταξύ της κίνησης των οφθαλμών και διαφόρων ψυχικών διαταραχών. Έχει διαπιστωθεί ότι είναι μεγάλη η χρησιμότητα της ιχνηλάτισης του βλέμματος για τη παρακολούθηση και αξιολόγηση των διαταραχών αυτών στους ασθενείς.
- **Υποστηρικτική τεχνολογία.** Η τεχνολογία eye tracking έχει τη δυνατότητα να βελτιώσει σε μεγάλο βαθμό την ανεξαρτησία και την ποιότητα ζωής ανθρώπων με αναπηρίες, προσφέροντάς τους ένα νέο τρόπο επικοινωνίας, αλληλεπίδρασης με την τεχνολογία και πρόσβασης πληροφοριών. Αυτό βασίζεται στο απλό γεγονός ότι η ιχνηλάτιση οφθαλμού επιτρέπει στα άτομα αυτά να αλληλεπιδρούν με τεχνολογικά μέσα χρησιμοποιώντας μόνο το βλέμμα τους.
- **Έρευνα βιντεοπαιχνιδιών.** Η ιχνηλάτιση οφθαλμού έχει αρχίσει να αξιοποιείται στη βιομηχανία των βιντεοπαιχνιδιών για την απόκτηση καλύτερων πληροφοριών από τους χρήστες σχετικά με την εμπειρία παιχνιδιού. Τέτοιες πληροφορίες είναι χρήσιμες είναι ανεκτίμητης αξίας για τους σχεδιαστές παιχνιδιών καθώς τους

επιτρέπει να τροποποιούν τα παιχνίδια για την καλύτερη αφοσίωση του χρήστη και τη δημιουργία χαρακτηριστικών που αυξάνουν την καθήλωση του.

- **Οδήγηση.** Οι ερευνητές στον τομέα της αυτοκινητοβιομηχανίας χρησιμοποιούν εδώ και καιρό τεχνολογίες ιχνηλάτησης βλέμματος για να μετρήσουν το επίπεδο προσοχής και συγκέντρωσης του οδηγού. Η γνώση αυτή είναι απαραίτητη για τον εντοπισμό λαθών και αδυναμιών του οδηγού καθώς και για τη δημιουργία πιο χρήσιμων και λιγότερο διασπαστικών διατάξεων ταμπλό αυτοκινήτου.
- **Εικονική πραγματικότητα.** Οι συσκευές εικονικής πραγματικότητας έχουν αρχίσει να κάνουν χρήση τεχνολογιών ιχνηλάτησης βλέμματος για την βελτίωση της απόδοσής τους. Αυτό γίνεται με την πιο στοχευμένη εμφάνιση λεπτομερειών μεγάλης ανάλυσης στα σημεία που κοιτάει ο χρήστης και τη μείωση ανάλυσης στα σημεία που βρίσκονται στην περιφερειακή όρασή του. Επίσης, η παρακολούθηση οφθαλμού στα συστήματα εικονικής πραγματικότητας θα μπορούσε να χρησιμοποιηθεί για την πιο ακριβή και ρεαλιστική εμφάνιση των avatar με εκφράσεις προσώπου, κίνηση οφθαλμών, βλεφάρισμα κ.λ.π.

Παρά τη μεγάλη χρησιμότητά και τις πολλές εφαρμογές τους, όμως, οι σημερινές συσκευές eye tracking, όπως θα δούμε παρακάτω, είτε είναι εξαιρετικά ακριβές για τον καθημερινό χρήστη, είτε υποφέρουν από θέματα χαμηλής ακρίβειας και αξιοπιστίας.

1.2 Κόστος εμπορικών συστημάτων

Οι υπάρχουσες εμπορικές συσκευές οφείλουν την αξιοπιστία και ακρίβεια τους σε εξειδικευμένους αισθητήρες (όπως κάμερες υπέρυθρης ακτινοβολίας, στερεοσκοπικές κάμερες υψηλής ανάλυσης κ.α.). Τέτοιοι αισθητήρες, λόγω της πολυπλοκότητας και τις ακρίβειάς τους αυξάνουν σημαντικά το κόστος της όλης συσκευής. Επιπροσθέτως, η αγορά των συσκευών eye tracking είναι μικρή σε σχέση με αυτή άλλων συσκευών πράγμα το οποίο κάνει δύσκολη την μαζική παραγωγή τους και μπορεί να αυξήσει το κόστος τους. Επίσης, επικρατεί ολιγοπώλιο, το οποίο μειώνει τον ανταγωνισμό στην αγορά και δεν δημιουργεί ανάγκη στις εταιρίες να μειώσουν τις τιμές. Τέλος, η ανάπτυξη των αλγορίθμων και των συστημάτων ιχνηλάτησης βλέμματος αποτελεί μία εξαιρετικά χρονοβόρα και δαπανηρή διαδικασία, γεγονός το οποίο συμβάλλει με τη σειρά του στο αυξημένο κόστος τους. Δημιουργείται ανάγκη, λοιπόν, για ανάπτυξη συστημάτων τα οποία θα χρησιμοποιούν ήδη υπάρχον υλικό (όπως μια συμβατική διαδικτυακή κάμερα) και θα εστιάζει περισσότερο στο κομμάτι του λογισμικού έτσι ώστε να μειωθεί ή και να εξαλειφθεί το κόστος τους. Από ότι θα δούμε παρακάτω, στην περίπτωση της εστίασης στο λογισμικό αντιμετωπίζουμε άλλων ειδών προβλήματα.

1.3 Περιορισμοί υλοποιήσεων με απλή κάμερα

Στις υπάρχουσες υλοποιήσεις με απλή κάμερα παρατηρούμε το βασικό πρόβλημα της χαμηλής ακρίβειας και της περαιτέρω επιδείνωσής της σε μη ιδανικές συνθήκες λειτουργίας. Εφόσον απορρίπτουμε τη χρήση εξειδικευμένων αισθητήρων λογικό είναι να συναντούμε αυτή τη μείωση απόδοσης. Τα συστήματα λογισμικού αυτά είναι φτιαγμένα για να λειτουργούν σε ιδανικές συνθήκες φωτισμού με τον χρήστη να βρίσκεται στο κέντρο της κάμερας, χωρίς να κουνά ή να περιστρέφει το κεφάλι του. Οποιαδήποτε μη τήρηση των συνθηκών αυτών έχει ως αποτέλεσμα την περαιτέρω μείωση της απόδοσης. Επιπλέον, τα περισσότερα αυτού του είδους συστήματα αναπτύσσονται από την ακαδημαϊκή κοινότητα ή από απλούς προγραμματιστές και συνεπώς δεν διαθέτουν την υποστήριξη, την ασφάλεια ή την αξιοπιστία μιας εταιρίας τεχνολογίας. Παρόλο που η

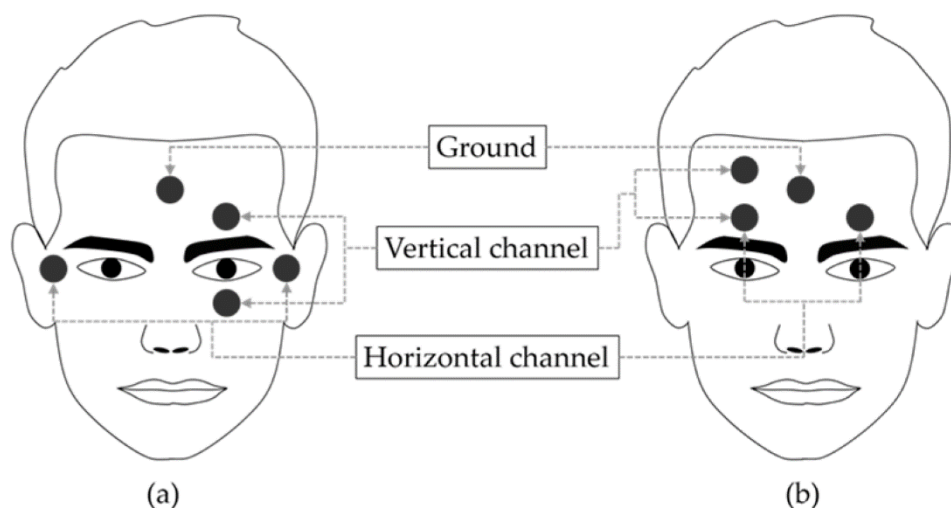
παρούσα εργασία δεν στοχεύει στη λύση του τελευταίου προβλήματος της υποστήριξης μιας εταιρίας τεχνολογίας, στοχεύει στην όσο το δυνατόν μεγαλύτερη βελτίωση της ακρίβειας του συστήματος που θα υλοποιηθεί και στην δυνατότητά του να λειτουργεί σε συνθήκες, ως ένα σεβαστό βαθμό, χειρότερες από τις ιδανικές.

Κεφάλαιο 2ο

2. Θεωρητικό υπόβαθρο (και προσέγγιση)

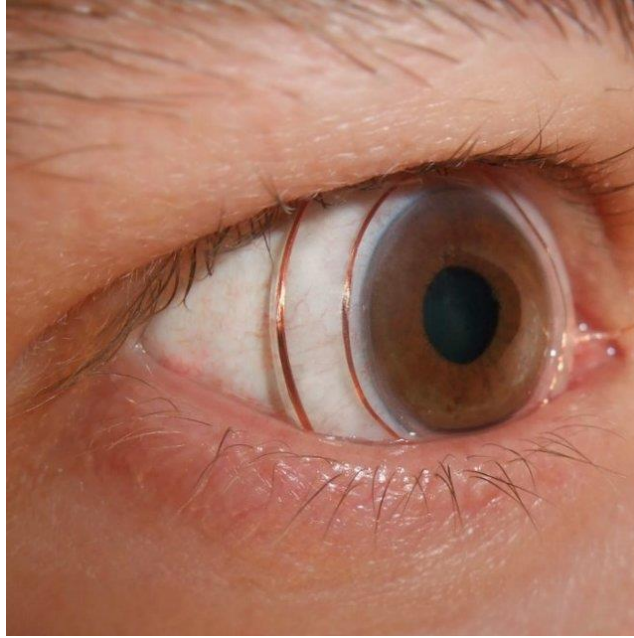
2.1 Βιβλιογραφική επισκόπηση

Σε αυτό το σημείο της εργασίας θα αναφέρουμε τις υπάρχουσες υλοποιήσεις συστημάτων ιχνηλάτησης οφθαλμού. Ξεκινώντας, έχουμε το Ηλεκτροοφθαλμογράφημα (ΗΟΓ). Η μέθοδος αυτή κάνει χρήση ηλεκτροδίων τα οποία τοποθετούνται γύρω από τον οφθαλμό και μετρούν τα ηλεκτρικά σήματα που παράγονται από τις κινήσεις των ματιών. Η βασική αρχή της μεθόδου αυτής είναι ότι το πίσω μέρος του ματιού έχει αρνητικό δυναμικό σε σχέση με το μπροστά κι έτσι όταν ο οφθαλμός αλλάζει θέση, αλλάζει και η κατανομή του ηλεκτρικού φορτίου. Το ΗΟΓ προσφέρει δεδομένα υψηλής ακρίβειας και χρησιμοποιείται κυρίως σε κλινικές περιπτώσεις για τη διάγνωση ασθενειών. Ένας περιορισμός του είναι ότι τα ηλεκτρόδια μπορούν να επηρεαστούν από ιδρώτα, τρίχες και κινήσεις άλλων κοντινών μυών (Janarthanan, 2013).



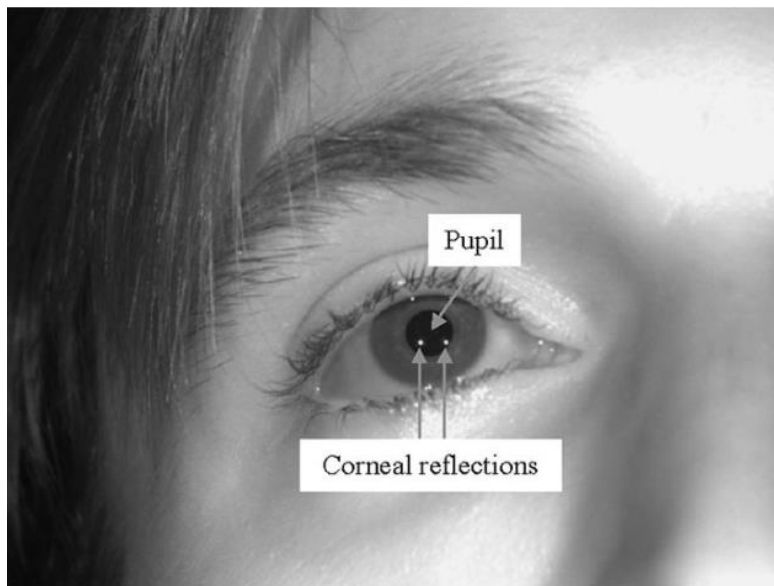
Εικόνα 1 Διαφορετικές τοποθετήσεις ηλεκτροδίων κατά το ηλεκτροοφθαλμογράφημα

Μια άλλη μέθοδος ιχνηλάτησης του οφθαλμού είναι η μέθοδος Πηνίων Αναζήτησης του Σκληρού Χιτώνα (Scleral search coils). Κατά τη μέθοδο αυτή τοποθετούνται πηνία στον οφθαλμό, είτε με μορφή φακών επαφής, είτε χειρουργικά, η κίνηση των οποίων στη συνέχεια ανιχνεύεται με τη χρήση ενός μαγνητικού πεδίου. Η μέθοδος των Πηνίων Αναζήτησης Σκληρού Χιτώνα θεωρείται η πιο ακριβής μέθοδος ανίχνευσης κινήσεων του οφθαλμού, καθώς προσφέρουν δεδομένα υψηλής ανάλυσης με ρυθμό έως και 1000 Hz. Για το λόγο αυτό, και λόγω της επεμβατικότητάς τους χρησιμοποιούνται μόνο σε ειδικές περιπτώσεις έρευνας, όπου η ύψιστη ακρίβεια παίζει σημαντικό ρόλο. Πέρα από την επεμβατικότητά της αυτή η μέθοδος μπορεί να προκαλέσει διαστρέβλωση των δεδομένων, καθώς τα πηνία είναι συνήθως ορατά από τον χρήστη κατά τη διάρκεια της καταγραφής και επομένως μπορεί να επηρεάσουν την κίνηση των οφθαλμών του (Robinson, 1963).



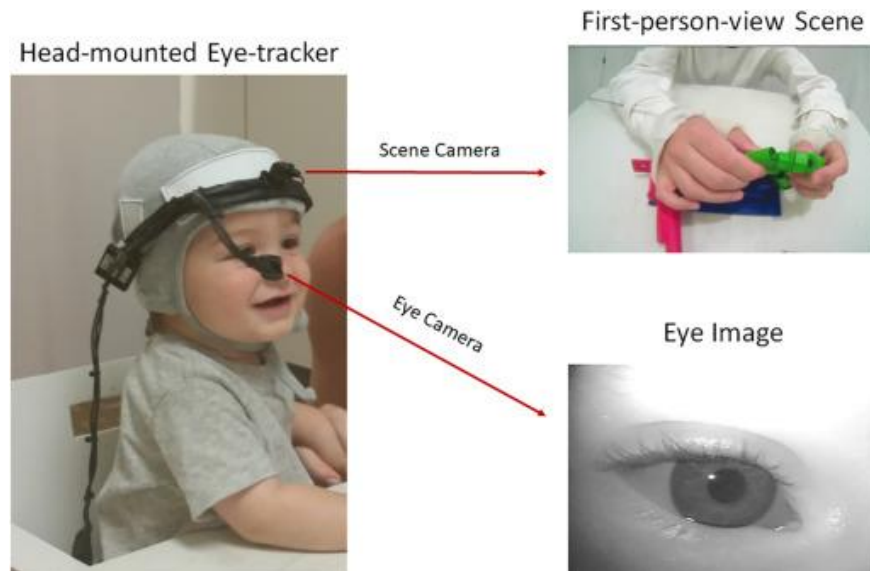
Εικόνα 2 Πηνία αναζήτησης σκληρού χιτώνα σε μορφή φακών επαφής

Στη συνέχεια έχουμε τις μεθόδους οι οποίες χρησιμοποιούν συνδυασμό μιας συσκευής καταγραφής (όπως η κάμερα) και μιας πηγής φωτός. Η πηγή φωτός μπορεί να είναι είτε πηγή υπέρυθρης ακτινοβολίας, είτε ένα απλό LED. Οι μέθοδοι αυτοί βασίζονται στην ανακλαστική ιδιότητα του κερατοειδούς χιτώνα του ματιού. Η αντανάκλαση της πηγής φωτός στα μάτια του χρήστη καθιστά εύκολο τον εντοπισμό τους από την συσκευή καταγραφής. Αποτελούν μια μη επεμβατική λύση η οποία προσφέρει αποτελέσματα υψηλής ανάλυσης, αλλά περιορίζεται από την επήρεια άλλων πηγών φωτός στον χώρο και τις αντανάκλασεις άλλων αντικειμένων (Kim, 2014).



Εικόνα 3 Εμφάνιση αντανάκλασης κερατοειδούς χιτώνα

Τέλος, έχουμε τις μεθόδους οι οποίες κάνουν χρήση μόνο συσκευών καταγραφής, όπως μια κάμερα. Τα συστήματα που χρησιμοποιούν αυτές τις μεθόδους χωρίζονται σε αυτά που είναι τοποθετημένα στο κεφάλι, και αυτά που λειτουργούν από απόσταση. Τα πρώτα συστήματα περιλαμβάνουν μία δομή στην οποία προσαρμόζεται η συσκευή καταγραφής και τοποθετείται στο κεφάλι του χρήστη, έτσι ώστε να παρακολουθεί την κίνηση των οφθαλμών από κοντά. Ταυτόχρονα, μία δεύτερη συσκευή καταγραφής, παρακολουθεί το περιβάλλον του χρήστη, έτσι ώστε να λαμβάνει πληροφορίες για αυτό, αλλά και για την κίνηση του κεφαλιού. Τα συστήματα αυτά είναι αποδοτικά στην ιχνηλάτηση του βλέμματος, αλλά χαρακτηρίζονται από περιορισμένη άνεση και ευκολία χρήσης (Cognolato, 2018).



Εικόνα 4 Συσκευή ιχνηλάτησης οφθαλμού τοποθετημένη στο κεφάλι του χρήστη. Φαίνονται οι όψεις της κάμερας για τους οφθαλμούς και της κάμερας για το περιβάλλον

Τα συστήματα που λειτουργούν από απόσταση βασίζονται στη λογική ότι η συσκευή καταγραφής είναι τοποθετημένη σε ένα σταθερό σημείο και καταγράφει τις κινήσεις των οφθαλμών και του κεφαλιού από απόσταση. Στη συνέχεια προσπαθούν, με τη χρήση διαφόρων αλγορίθμων, τους οποίους θα αναλύσουμε σε επόμενο κεφάλαιο, να εντοπίσουν τα επιθυμητά στοιχεία, όπως το πρόσωπο ή οι ίριδες των οφθαλμών. Η ακρίβεια τέτοιων μεθόδων είναι η χειρότερη από τις προαναφερθείσες, αλλά προσφέρει τη μεγαλύτερη δυνατή ευελιξία, αφού το μόνο που απαιτείται για την ανάπτυξη και την χρήση τους είναι μία συσκευή καταγραφής. Παρόλο που η ακρίβεια είναι ελλιπής σε σχέση με αυτή άλλων μεθόδων, εξαρτάται σε μεγάλο βαθμό από το λογισμικό και επομένως μπορεί να βελτιωθεί. Η εργασία αυτή διαπραγματεύεται την ανάπτυξη ενός τέτοιου συστήματος και αλγορίθμου (με χρήση μόνο συσκευής καταγραφής από απόσταση), με στόχο την αντιστάθμιση της έλλειψης υλικών συσκευών και εξειδικευμένων αισθητήρων και την επίτευξη της μεγαλύτερης δυνατής βελτίωσης της ακρίβειας.

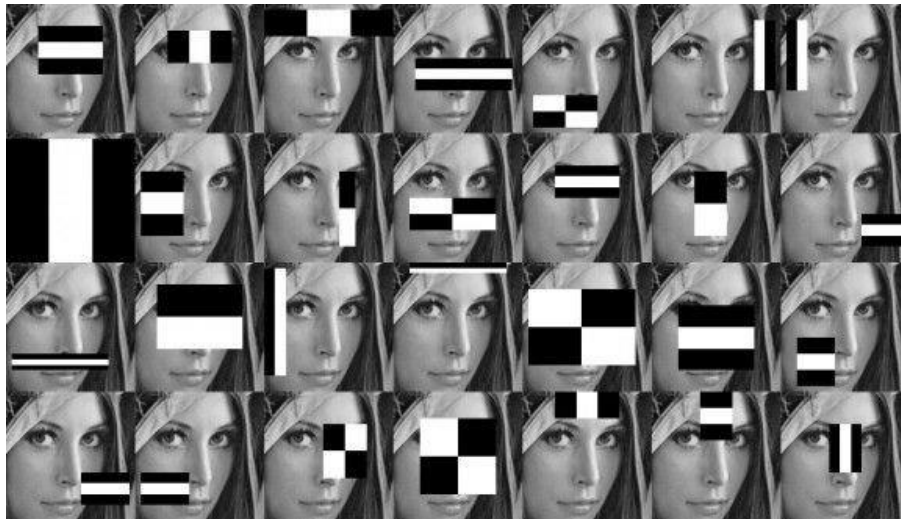
2.2 Η διαδικασία εντοπισμού και υπολογισμού βλέμματος

Η διαδικασία του υπολογισμού των συντεταγμένων του σημείου στην οθόνη στο οποίο κοιτάει ο χρήστης από βίντεο απλής συσκευής καταγραφής μπορεί να χωριστεί σε τρία επιμέρους κομμάτια. Αρχικά, έχουμε τον εντοπισμό προσώπου ο οποίος βοηθά στον υπολογισμό της θέσης, της απόστασης και της γωνίας του κεφαλιού σε σχέση με την

οθόνη. Στη συνέχεια απαιτείται ο εντοπισμός της κόρης ή ίριδας του ματιού ώστε να εξαχθεί η πληροφορία της γωνίας του βλέμματος και τέλος, από τις παραπάνω πληροφορίες, πρέπει να υπολογίσουμε το σημείο (x, y) στην οθόνη στο οποίο το βλέμμα του χρήστη συναντά την επιφάνειά της. Αξίζει να σημειωθεί ότι υπάρχουν μέθοδοι οι οποίες μπορούν με είσοδο την εικόνα να υπολογίσουν απευθείας τις προαναφερθείσες συντεταγμένες όπως θα δούμε παρακάτω.

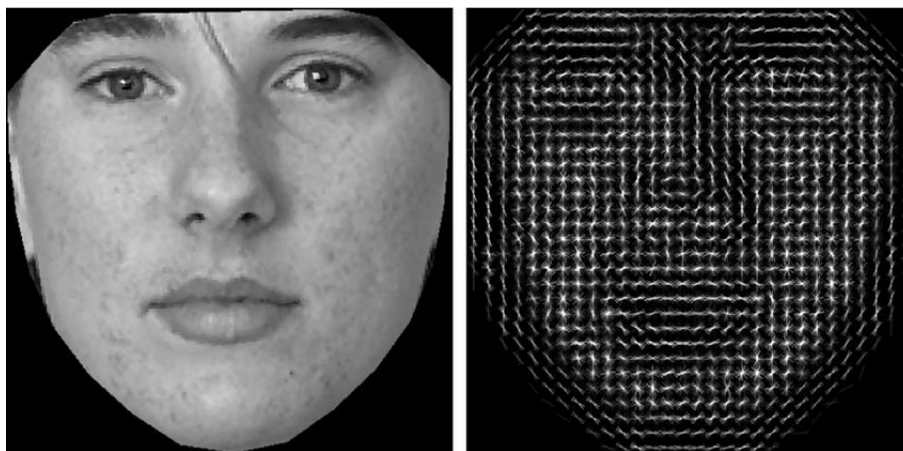
2.2.1 Εντοπισμός προσώπου

Για τον εντοπισμό προσώπου έχουν αναπτυχθεί πολυάριθμοι αλγόριθμοι, η περισσότεροι από τους οποίους χρησιμοποιούν μεθόδους αναγνώρισης προτύπων και μηχανικής μάθησης. Ο αλγόριθμος Viola-Jones είναι ευρέως χρησιμοποιούμενος για τον εντοπισμό προσώπων σε εικόνες και βίντεο (Viola, 2001). Αξιοποιεί ένα cascade από classifiers, όπου ο κάθε classifier είναι εκπαιδευμένος να εντοπίζει ένα συγκεκριμένο στοιχείο του προσώπου, όπως τα μάτια, η μύτη ή το στόμα. Ο αλγόριθμος ξεκινά με τον εντοπισμό των πιο απλών στοιχείων και όσο γίνεται πιο βέβαιος για την παρουσία του προσώπου προχωρά σε πιο περίπλοκα στοιχεία του προσώπου.



Εικόνα 5 Αντιστοίχιση Haar Features σε στοιχεία προσώπου από τον αλγόριθμο Viola - Jones

Η μέθοδος Histograms of Oriented Gradients (HOG) χρησιμοποιεί τις πληροφορίες gradient στην εικόνα για τον εντοπισμό του προσώπου. Κατασκευάζει ένα ιστόγραμμα κατευθύνσεων των gradient σε τοπικές περιοχές της εικόνας και χρησιμοποιεί τις πληροφορίες αυτές για να αναγνωρίσει πιθανά πρόσωπα (O. Déniz, 2011).



Εικόνα 6 Προσανατολισμός των gradient σε σημεία του προσώπου

Οι μέθοδοι βασισμένοι σε μοντέλα βαθιάς μάθησης χρησιμοποιούν συνελκτικά νευρωνικά δίκτυα (Convolutional Neural Networks or CNN) για την αναγνώριση προσώπων σε εικόνες και βίντεο. Το CNN εκπαιδεύεται με ένα dataset από επισημασμένες εικόνες και μετά από τη διαδικασία αυτή έχει την ικανότητα να εντοπίζει πρόσωπα σε νέες εικόνες, αναγνωρίζοντας τα χαρακτηριστικά προσώπου που έμαθε στη διάρκεια της εκπαίδευσης. Στην παρούσα διπλωματική θα εστιάσουμε στα κομμάτια υλοποίησης του εντοπισμού οφθαλμού και του υπολογισμού των συντεταγμένων, έχοντας διαθέσιμες τις απαραίτητες πληροφορίες εντοπισμού προσώπου από κάποιο έτοιμο εργαλείο.

2.2.2 Εντοπισμός των οφθαλμών

Οι μέθοδοι και οι αλγόριθμοι αναγνώρισης οφθαλμών βασίζονται σε παρόμοιες αρχές με αυτές των μεθόδων εντοπισμού προσώπου. Αρχικά έχουμε την μέθοδο μηχανικής μάθησης Haar cascades, η οποία χρησιμοποιεί ένα cascade από χαρακτηριστικά που πρέπει να εντοπιστούν μέσα στην εικόνα (Phase, 2019). Τα Haar cascades εκπαιδεύονται με τη χρήση θετικά και αρνητικά επισημασμένων εικόνων και είναι αποδοτικά για τον εντοπισμό οφθαλμών σε πραγματικό χρόνο. Θα μπορούσαμε να πούμε ότι η μέθοδος Haar cascades είναι η αντίστοιχη της μεθόδου Viola Jones της αναγνώρισης προσώπου, αλλά αυτή τη φορά σε αναγνώριση οφθαλμών. Η μέθοδος Histogram of Oriented Gradients (HOG) και Support Vector Machine (SVM) χρησιμοποιεί κατευθύνσεις από gradient για να ανιχνεύσει άκρες μέσα σε μία εικόνα και στη συνέχεια ένας SVM classifier καθορίζει εάν η περιοχή της άκρης περιέχει όντως οφθαλμό (J. K., 2017). Έχουμε, επίσης, παρόμοια με την ανίχνευση προσώπου, τις μεθόδους CNN οι οποίες χρησιμοποιούν τεχνικές βαθιάς μάθησης για τον εντοπισμό οφθαλμών σε εικόνες. Οι μέθοδοι CNN είναι πιο ακριβείς από τις Haar cascades και τις HOG + SVM, αλλά απαιτούν μεγαλύτερη υπολογιστική ισχύ και μεγαλύτερο όγκο δεδομένων για την εκπαίδευσή τους (Yu, 2018). Οι μέθοδοι Deep Neural Network (DNN), παρόμοια με τους CNN, βασίζονται σε μοντέλα βαθιάς μάθησης, αλλά είναι πιο δυνατοί και μπορούν να ανιχνεύουν οφθαλμούς με καλύτερη ακρίβεια. Τέλος, ο αλγόριθμος YOLO (You Only Look Once) μπορεί να εντοπίζει οφθαλμούς σε πραγματικό χρόνο χρησιμοποιώντας ένα μοναδικό convolutional neural network για να προβλέψει τις περιοχές και τις ενδείξεις αντικειμένων, όπως τα μάτια, στην εικόνα (Kumari, 2021). Είναι σημαντικό να σημειώσουμε ότι η χρησιμότητα της κάθε μεθόδου εξαρτάται από τη συγκεκριμένη περίπτωση χρήσης του συστήματος και τους διαθέσιμους πόρους (υπολογιστική δύναμη, μέγεθος dataset κλπ.).

2.2.3 Μέθοδοι υπολογισμού βλέμματος

Οι μέθοδοι βασισμένες στην εμφάνιση του οφθαλμού (appearance - based) χρησιμοποιούν τις φωτομετρικές ιδιότητες της περιοχής του στην εικόνα, σε συνδυασμό με μοντέλα μηχανικής μάθησης για να εντοπίσουν τα σημεία ενδιαφέροντός του. Αυτό σημαίνει ότι τα μοντέλα αυτά, με είσοδο τις τιμές φωτεινότητας της εικόνας στην περιοχή του ματιού, μπορούν να δώσουν σαν έξοδο την κατεύθυνση του βλέμματος ή, με κατάλληλη εκπαίδευση, και τις συντεταγμένες του σημείου στην οθόνη απευθείας. Για να επιτευχθούν τα παραπάνω, τα μοντέλα χρειάζονται εκτενή εκπαίδευση με μεγάλα και ποικίλα σετ δεδομένων έτσι ώστε να μπορούν να είναι αποδοτικά σε διάφορα είδη οφθαλμών, συνθηκών φωτισμού και ανάλυσης εικόνας. Οι πιο συνηθισμένοι αλγόριθμοι που χρησιμοποιούνται για το σκοπό αυτό είναι οι KNN (K-Nearest-Neighbor), παλινδρόμηση RF (Random Forest), παλινδρόμηση GP (Gaussian Process), SVM (Support Vector Machines), ANN (Artificial Neural Networks) και CNN (Convolutional Neural Networks) (Jiaqi Jiang, 2019). Παρά την απλότητα χρήσης τους και την ευελιξία τους όσον αφορά τις συνθήκες φωτισμού και ποιότητας κάμερας, τέτοιες μέθοδοι περιορίζονται από παράγοντες κλίμακας και περιστροφής του κεφαλιού και λόγω αυτού μπορούν να παρουσιάσουν μειωμένη ακρίβεια (Camilleri, 2018).

Οι μέθοδοι βασισμένες στα στοιχεία του ματιού (feature - based) προσπαθούν να εντοπίσουν συγκεκριμένα τοπικά στοιχεία ενδιαφέροντος του ματιού – όπως το κέντρο του, γωνίες ή περιγράμματα- ή και να ταιριάξει γεωμετρικά σχήματα – όπως κύκλοι ή ελλείψεις - στα στοιχεία αυτά. Στη δεύτερη περίπτωση, τα σχήματα τα οποία θα αντιστοιχηθούν στα διάφορα στοιχεία του ματιού μπορεί να είναι είτε απλά, όπως μία απλή έλλειψη στην κόρη ή την ίριδα του ματιού, είτε πιο σύνθετα που να περιλαμβάνουν τα βλέφαρα και τις γωνίες του ματιού (Camilleri, 2018). Οι μέθοδοι αυτές εφαρμόζονται σε μικρό κομμάτι της εικόνας το οποίο περιλαμβάνει αποκλειστικά τον οφθαλμό έτσι ώστε να αποφεύγονται αστοχίες στην αναγνώριση των σχημάτων και λανθασμένος εντοπισμός στοιχείων του ματιού εκεί που δεν υπάρχουν (Camilleri, 2018). Είναι προφανές ότι για να μπορέσουμε να εξάγουμε πληροφορίες από ένα μικρό κομμάτι της εικόνας που περιέχει τον οφθαλμό χρειαζόμαστε επαρκώς φωτισμένη εικόνα υψηλής ανάλυσης, γεγονός το οποίο περιορίζει αρκετά τις προδιαγραφές του συστήματος.

Για τον υπολογισμό της γωνίας όρασης και του στίγματος στην επιφάνεια της οθόνης μπορούν να χρησιμοποιηθούν μέθοδοι είτε γεωμετρικές, είτε παλινδρόμησης. Οι γεωμετρικές μέθοδοι μπορούν με τη σειρά τους να διαχωριστούν σε δυσδιάστατες και τρισδιάστατες. Οι δυσδιάστατες μέθοδοι χρησιμοποιούν μία συνάρτηση αντιστοίχισης της θέσης/γωνίας των ματιών και του κεφαλιού που εκλήφθηκαν από την εικόνα, στο σημείο στο οποίο κοιτάει ο χρήστης στην οθόνη. Το θετικό των μεθόδων αυτών είναι ότι έχουν αρκετά μικρή πολυπλοκότητα, η συνάρτηση αντιστοίχισης είναι σχετικά απλή και προσφέρει αξιόπιστα αποτελέσματα (Martínez, 2019). Οι τρισδιάστατες γεωμετρικές μέθοδοι έχουν μεγαλύτερη πολυπλοκότητα, αφού υπολογίζουν τα διανύσματα από τα μάτια προς το σημείο όρασης στην οθόνη με χρήση της γωνίας του κεφαλιού, της γωνίας όψης των ματιών και της θέσης και της απόστασης από την οθόνη. Πέρα από την αυξημένη πολυπλοκότητά τους, οι γεωμετρικές μέθοδοι απαιτούν εικόνα υψηλής ανάλυσης για να πετύχουν επαρκή ακρίβεια στους υπολογισμούς τους (Martínez, 2019). Τέλος, έχουμε τις μεθόδους παλινδρόμησης οι οποίες μπορούν να λαμβάνουν σαν είσοδο μέρη της εικόνας και να παράγουν απευθείας τις συντεταγμένες του στίγματος που μας ενδιαφέρει, ή να λαμβάνουν παραμέτρους σχετικές με τη θέση/γωνία των ματιών και του κεφαλιού από αυτές να παράγει το εν λόγω αποτέλεσμα (Fakotakis, 2015). Οι μέθοδοι

παλινδρόμησης όπως είδαμε είναι αρκετά ευέλικτες αλλά απαιτούν πολλή εκπαίδευση για να παρουσιάσουν ακριβή αποτελέσματα (Camilleri, 2018).

Κεφάλαιο 3ο

3. Προτεινόμενη προσέγγιση - Μεθοδολογία

3.1 Η προτεινόμενη προσέγγιση - προδιαγραφές

Όπως αναφέρθηκε παραπάνω, έχουν αναπτυχθεί αρκετά συστήματα με διαφορετικές προσεγγίσεις, αρχιτεκτονικές και αλγορίθμους για την επίτευξη του εντοπισμού του βλέμματος του χρήστη με χρήση απλής διαδικτυακής κάμερας, και του υπολογισμού της θέσης του αντίστοιχου στίγματος στην οθόνη. Ο τρόπος με τον οποίο επιχειρούμε να διαφοροποιηθούμε από τα συστήματα αυτά και να καινοτομήσουμε είναι η ευρεία συμβατότητα, οι ελαστικές συνθήκες λειτουργίας και η ταχύτητα/απλότητα του αλγορίθμου. Αν επιτευχθούν οι παραπάνω καινοτομίες, το σύστημά μας θα έχει μεγάλη ανοχή σε διάφορες αναλύσεις κάμερας και οθόνης, σε διάφορες συνθήκες φωτισμού, σε όλα τα είδη προσώπων και ματιών (με γυαλιά και μη), θα επιτρέπει μεγαλύτερη ελευθερία κινήσεων κεφαλιού και θα μπορεί να τρέχει σε σεβαστή συχνότητα ακόμα και στις πιο αδύναμες συσκευές. Η επίτευξη των παραπάνω συνεπάγεται την εύκολη προσαρμογή του συστήματός μας σε οποιαδήποτε από τις πιθανές περίπτωση χρήσης που μπορεί να προκύψουν.

3.2 Εξάλειψη διαφορών εισόδου

Η είσοδος του συστήματος που θα αναπτύξουμε αποτελείται από τη ροή βίντεο της κάμερας που διαθέτει η συσκευή του χρήστη. Η είσοδος αυτή μπορεί να διαφέρει δραματικά σε διαφορετικές περιπτώσεις χρήσης. Διαφορές θα μπορούσαμε να έχουμε αρχικά στις συνθήκες φωτισμού. Η ιδανική περίπτωση θα ήταν να έχουμε φωτεινή πηγή η οποία να βρίσκεται έξω από το οπτικό πεδίο της κάμερας (ώστε να μη χαλάει η ισορροπία φωτός στην εικόνα της κάμερας) και να φωτίζει επαρκώς το πρόσωπο του χρήστη. Σύμφωνα όμως με τις προδιαγραφές του συστήματός μας, θα μπορεί να λειτουργεί και σε περιπτώσεις χαμηλού ή κακού φωτισμού. Ένας άλλος τρόπος με τον οποίο θα μπορούσε να διαφέρει η είσοδος του συστήματος είναι το σχήμα και τα χαρακτηριστικά του προσώπου του κάθε χρήστη. Εδώ θέλουμε και πάλι να έχουμε σεβαστή λειτουργία του συστήματος ακόμα και στα πιο ιδιαίτερα χαρακτηριστικά προσώπου, ακόμα και γυαλιά, σκουλαρίκια, τατουάζ ή σημάδια. Τέλος, η είσοδος της κάμερας εξαρτάται από τα χαρακτηριστικά της ίδιας της κάμερας. Τέτοια χαρακτηριστικά είναι η ανάλυση, ο ρυθμός των καρέ, η γωνία θέασης και η ποιότητα του αισθητήρα.

Για να εξαλείψουμε τις διαφορές κάνουμε εφαρμογή του εργαλείου Mediapipe της Google. Το Mediapipe δέχεται σαν είσοδο τα καρέ της κάμερας και παρέχει, μεταξύ άλλων, αναγνώριση προσώπου και ίριδας. Το εργαλείο αυτό αποτελείται από μοντέλα τα οποία έχουν εκπαιδευτεί με χρήση τεράστιων dataset και λειτουργούν σε όλες τις διάφορες συνθήκες εισόδου που περιεγράφηκαν παραπάνω. Έτσι λαμβάνουμε όλες τις απαραίτητες πληροφορίες του προσώπου που μας ενδιαφέρουν ανεξαρτήτως των συνθηκών φωτισμού, της ποιότητας της κάμερας ή των ιδιοτήτων των χαρακτηριστικών του. Περισσότερες πληροφορίες για το σύστημα Mediapipe παρέχονται σε επόμενο κεφάλαιο.

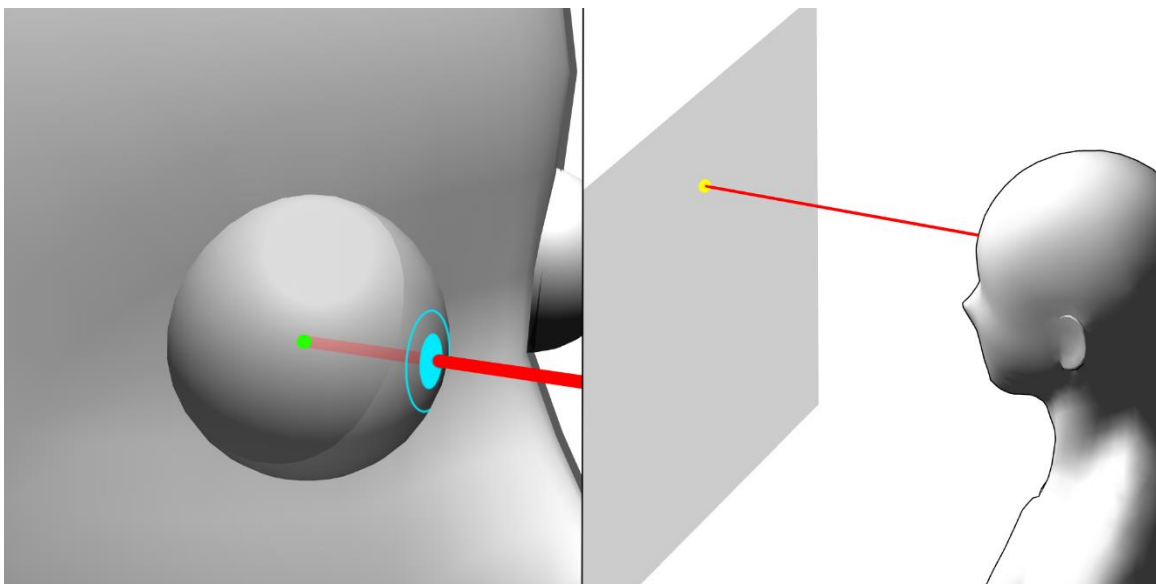
3.3 Μεθοδολογία

Έχοντας στη διάθεσή μας τα απαραίτητα σημεία του προσώπου και της ίριδας των ματιών, μπορούμε να ξεκινήσουμε να αναπτύσσουμε τον αλγόριθμο του κυρίως συστήματος εντοπισμού βλέμματος στην οθόνη. Επιλέγουμε να κάνουμε χρήση τρισδιάστατων γεωμετρικών μεθόδων για πληθώρα διαφόρων λόγων. Αρχικά, η μέθοδος αυτή ταιριάζει με την έξοδο του εργαλείου Mediapipe, αφού οι πληροφορίες της θέσης των στοιχείων του προσώπου παρέχονται σε τρισδιάστατες συντεταγμένες με τρόπο που θα περιγραφεί καλύτερα στο επόμενο κεφάλαιο. Έτσι, αξιοποιούμε πλήρως την πληροφορία που μας δίνεται και την χρησιμοποιούμε για να επιτύχουμε καλύτερη ακρίβεια. Επίσης, οι τρισδιάστατες γεωμετρικές μέθοδοι παρέχουν μεγαλύτερη δυνατότητα υπολογισμού του βλέμματος σε διάφορες θέσεις του κεφαλιού, αφού κάνουν δυνατό τον υπολογισμό της θέσης αυτής και του προσανατολισμού του. Ένας άλλος λόγος για τον οποίο επιλέγουμε τρισδιάστατες γεωμετρικές μεθόδους για τον υπολογισμό του σημείου στην οθόνη στο οποίο κοιτά ο χρήστης, είναι ότι έχουμε καταφέρει, με την εξάλειψη διαφορών εισόδου, να παρακάμψουμε ως ένα βαθμό το μοναδικό μειονέκτημά τους, το οποίο είναι η απαίτηση υψηλής ανάλυσης της κάμερας.

Η μεθοδολογία του συστήματος που επιλέξαμε να αναπτύξουμε βασίζεται στον υπολογισμό τριών στοιχείων:

- Το ζευγάρι των σημείων των κέντρων των ματιών στον τρισδιάστατο χώρο σε σχέση με την κάμερα
- Το αντίστοιχο ζευγάρι των ιρίδων των ματιών στον χώρο
- Τη θέση της οθόνης σε σχέση με την κάμερα

Έχοντας υπολογίσει αυτά τα τρία στοιχεία, το πρόβλημά μας μειώνεται στην εύρεση των σημείων στα οποία οι δύο ευθείες του βλέμματος των ματιών τέμνουν το επίπεδο της οθόνης.



Εικόνα 7 Ευθεία που ορίζεται από το κέντρο του ματιού (πράσινο σημείο) και την επιφάνειά του στο σημείο της ίριδας (μπλε σημείο), η οποία συναντά την επιφάνεια της οθόνης στο σημείο που ψάχνουμε (κίτρινο σημείο)

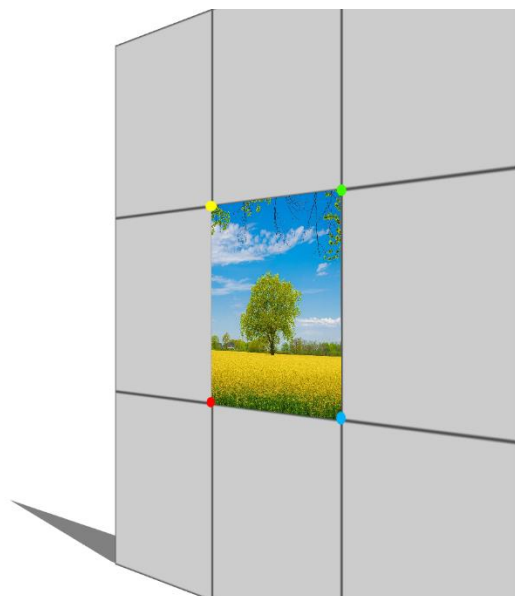
Τα σημεία της ίριδας του κάθε ματιού στον χώρο τα λαμβάνουμε για το κάθε καρέ μετά από μικρή επεξεργασία από την βιβλιοθήκη MediaPipe. Ο υπολογισμός των κέντρων των ματιών, όμως, αποτελεί πιο περίπλοκη διαδικασία. Πρέπει να βρούμε πρώτα την θέση του κέντρου του κάθε ματιού σε σχέση με τα υπόλοιπα σημεία του προσώπου και μετά να υπολογίσουμε την απόλυτη θέση των κέντρων αυτών με βάση την απόλυτη θέση και προσανατολισμό του προσώπου. Η σχετική με το υπόλοιπο πρόσωπο, θέση των κέντρων των ματιών θα υπολογιστεί μία φορά κατά τη διαδικασία της βαθμονόμησης και η απόλυτη θέση θα υπολογίζεται στο κάθε καρέ με βάση τα σημεία του προσώπου και την σχετική θέση που έχουμε ήδη υπολογίσει. Έχοντας για το κάθε μάτι το σημείο του κέντρου και της ίριδας του, μπορούμε να υπολογίσουμε το διάνυσμα και κατ' επέκταση την ευθεία του βλέμματος του. Ο υπολογισμός του επιπέδου της οθόνης θα γίνει κατά το στάδιο της βαθμονόμησης. Επιλέγουμε τρία συγκεκριμένα σημεία της οθόνης: το κέντρο, την δεξιά άκρη και την κάτω άκρη της. Κατά τη βαθμονόμηση λαμβάνουμε τις ευθείες του βλέμματος του κάθε ματιού όταν ο χρήστης κοιτάει αυτά τα σημεία της οθόνης που επιλέξαμε. Έχοντας τις ευθείες αυτές και γνωρίζοντας το μέγεθος της οθόνης, μπορούμε να υπολογίσουμε τη θέση της. Κατά την όλη διαδικασία που περιγράφηκε, υπολογίζονται κι άλλα αναγκαία μεγέθη, όπως η απόσταση του προσώπου, ο προσανατολισμός του, διάφορες μετρικές της κάμερας κ.λ.π., τα οποία συμβάλλουν στην διεξαγωγή των υπολογισμών των τριών βασικών στοιχείων της μεθόδου μας. Η αναλυτική διαδικασία του αλγορίθμου μας αναλύεται στο επόμενο κεφάλαιο.

3.3.1 Εύρεση κέντρου των ματιών

Όπως αναφέρθηκε παραπάνω η εύρεση των επιθυμητών συντεταγμένων στην οθόνη βασίζεται στη δημιουργία μίας ευθείας στον τρισδιάστατο χώρο για το κάθε μάτι, η οποία ορίζεται από τα εξής δύο σημεία: το κέντρο του ματιού και την ίριδα του ματιού. Τα σημεία της ίριδας του κάθε ματιού τα εξάγουμε από τα landmarks που μας δίνει η βιβλιοθήκη MediaPipe. Το κέντρα των ματιών θα πρέπει να υπολογίζονται στο κάθε καρέ από άλλα σημεία ενδιαφέροντος του προσώπου, τα οποία εξάγουμε από τα landmarks της βιβλιοθήκης, μέσω μιας διαδικασίας η οποία θα αναλυθεί στο κεφάλαιο αυτό.

3.3.1.1 Ομογραφία επιπέδου

Αρχικά, θεωρείται ότι τα μάτια είναι σφαίρες των οποίων τα κέντρα είναι κάθε στιγμή σταθερά σε σχέση με το πρόσωπο του χρήστη. Επίσης σε μία εικόνα στην οποία ο χρήστης κοιτάει (με τα μάτια αλλά και με την κατεύθυνση του προσώπου) απευθείας μέσα στην κάμερα, τα σημεία της κάθε ίριδας συμπίπτουν με τα σημεία των κέντρων των ματιών. Μία τέτοια εικόνα αποθηκεύεται στην αρχή της ροής του προγράμματος, στη φάση της ισοστάθμισης η οποία θα αναλυθεί σε επόμενο κεφάλαιο. Έτσι αποκτάμε τη θέση των κέντρων των ματιών σχετικά με τα σημεία του προσώπου. Λόγω της δυσδιάστατης εικόνας όμως, δεν μπορούμε να βγάλουμε συμπέρασμα για το βάθος των κέντρων σε σχέση με το πρόσωπο. Για την εύρεση και των τριών διαστάσεων, ξεκινάμε εφαρμόζοντας τη διαδικασία της ομογραφίας επιπέδου (planar homography). Η γεωμετρική αυτή διαδικασία χρησιμεύει στον μετασχηματισμό ενός επιπέδου και οποιωνδήποτε σημείων μέσα σε αυτό, έτσι ώστε να χωρέσει σε συγκεκριμένη θέση και κατεύθυνση του τρισδιάστατου χώρου ο οποίος αποτυπώνεται σε εικόνα με συγκεκριμένη προοπτική. Ένα παράδειγμα της χρήσης της ομογραφίας επιπέδου απεικονίζεται στην παρακάτω εικόνα.



Εικόνα 8 Εφαρμογή μεθόδου ομογραφίας επιπέδου σε δισδιάστατη εικόνα

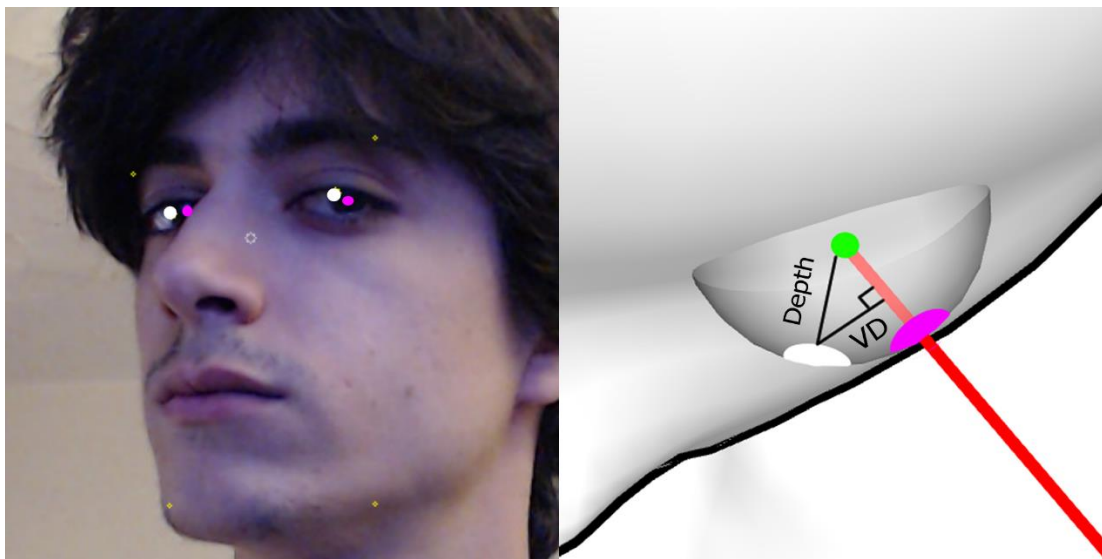
Στην περίπτωση μας, το επίπεδο το οποίο εισάγουμε στη διαδικασία είναι μέρος της αρχικής εικόνας που λήφθηκε κατά την ισοστάθμιση, όπου τα κέντρα συμπίπταν με τις ίριδες των ματιών. Χρησιμοποιούμε τέσσερα σημεία του προσώπου τα οποία επιλέγονται έτσι ώστε να κινούνται όσο το δυνατόν λιγότερο με τους διάφορους μορφασμούς του προσώπου, και τα δύο σημεία των κέντρων. Έτσι, μετά από τη φάση της ισοστάθμισης, πραγματοποιούμε τη διαδικασία της ομογραφίας του επιπέδου που ορίζεται από αυτά τα σημεία για κάθε καρέ του ζωντανού βίντεο της κάμερας. Με τον τρόπο αυτό αποκτάμε μία αρχική θέση των κέντρων των ματιών σε σχέση με το πρόσωπο, η οποία όμως έχει ακόμα λανθασμένο βάθος. Η διόρθωση του σφάλματος αυτού αναλύεται στο επόμενο κεφάλαιο.



Εικόνα 9 Εφαρμογή μεθόδου ομογραφίας επιπέδου σε τρισδιάστατα σημεία του προσώπου

3.3.1.2 Άμεσος γραμμικός μετασχηματισμός

Εφόσον ψάχνουμε το βάθος των κέντρων των ματιών σε σχέση με το πρόσωπο θα μετακινήσουμε τα σημεία που βρήκαμε από την ομογραφία επιπέδου στον άξονα z του προσώπου. Αυτό σημαίνει ότι πρέπει πρώτα να μπορούμε να βρούμε τον προσανατολισμό του προσώπου σε σχέση με το σύστημα συντεταγμένων της κάμερας σε κάθε καρέ. Για να τον σκοπό αυτό χρησιμοποιούμε τη μέθοδο του Άμεσου Γραμμικού Μετασχηματισμού (Direct Linear Transform). Σύμφωνα με αυτόν, έχοντας τις τρισδιάστατες συντεταγμένες τουλάχιστον 6 σημείων ενός αντικειμένου (τα οποία δεν βρίσκονται στο ίδιο επίπεδο) και τις αντίστοιχες δυσδιάστατες θέσεις τους σε μία εικόνα, μπορούμε να υπολογίσουμε τους πίνακες μετατόπισης, περιστροφής και αλλαγής κλίμακας, από τους οποίους μπορούμε να βγάλουμε συμπέρασμα για τον προσανατολισμό του αντικειμένου. Στην περίπτωση μας το αντικείμενο αυτό είναι το πρόσωπο του χρήστη, τα αρχικά τρισδιάστατα σημεία λαμβάνονται από τη βιβλιοθήκη MediaPipe στην πρώτη εικόνα που λαμβάνεται στην ισοστάθμιση (με τρόπο τέτοιο ώστε να μη βρίσκονται στο ίδιο επίπεδο), ενώ τα δυσδιάστατα σημεία λαμβάνονται και πάλι από τη βιβλιοθήκη MediaPipe στο κάθε καρέ του ζωντανού βίντεο. Έτσι, έχοντας την κατεύθυνση στην οποία πρέπει να κινηθούν τα αρχικά κέντρα των ματιών για να βρεθούν στην σωστή τους θέση, χρειάζεται να βρούμε το μέγεθος της μετατόπισης που πρέπει να εφαρμοστεί στην κατεύθυνση αυτή. Για γίνει αυτό λαμβάνουμε άλλη μία εικόνα στο στάδιο της ισοστάθμισης, στην οποία ζητείται στον χρήστη να γείρει το κεφάλι του προς τη μία μεριά κοιτώντας και πάλι την κάμερα. Στην εικόνα αυτή τα σημεία της ίριδας των ματιών συμπίπτουν και πάλι με τα κέντρα των ματιών και με έναν απλό τριγωνισμό βρίσκουμε το ακριβές βάθος των κέντρων των ματιών, όπως φαίνεται στην παρακάτω εικόνα.



Εικόνα 10 Εύρεση βάθους του κέντρου της σφαίρας του ματιού με αλλαγή του προσανατολισμού του προσώπου

3.3.2 Εύρεση σημείου βλέμματος στην οθόνη

Μέχρι στιγμής έχουμε καταφέρει να κατέχουμε σε κάθε καρτέλα τα σημεία της ίριδας και του κέντρου του κάθε ματιού. Το τελευταίο στοιχείο που χρειαζόμαστε για τον υπολογισμό των συντεταγμένων του σημείου στο οποίο κοιτάει ο χρήστης στην οθόνη είναι η απόσταση του προσώπου από την οθόνη. Θεωρούμε ότι η απόσταση αυτή είναι ίδια με την απόσταση του προσώπου από την κάμερα. Για να υπολογίσουμε την απόσταση προτρέπουμε τον χρήστη να βρίσκεται περίπου 60 εκατοστά μακριά από την κάμερα κατά την διαδικασία της βαθμονόμησης. Στην εικόνα που λαμβάνουμε κατά την βαθμονόμηση κρατάμε την τρισδιάστατη απόσταση μεταξύ δύο αυθαίρετων σημείων του προσώπου. Ξέροντας την απόσταση μεταξύ αυτών των δύο σημείων στα 60 εκατοστά μακριά από την οθόνη, την υπολογίζουμε ξανά στο κάθε καρτέλα και με απλή μέθοδο των τριών μπορούμε να βρούμε την απόσταση του χρήστη από την οθόνη.

Για τον υπολογισμό του τελικού σημείου της οθόνης που μας ενδιαφέρει βρίσκουμε την ευθεία του βλέμματος του κάθε ματιού με χρήση των σημείων κέντρου και ίριδας. Στη συνέχεια ξέροντας την απόσταση του προσώπου, βρίσκουμε την απόσταση του κάθε ματιού με δεδομένα που προσφέρονται από την βιβλιοθήκη Mediripe. Έτσι με χρήση του θεωρήματος όμοιων τριγώνων βρίσκουμε τις συντεταγμένες του σημείου όπου η ευθεία τη κάθε ματιού συναντά την οθόνη. Ο υπολογισμός του συνδυασμένου σημείου συμπίπτει με τον μέσο όρο των δύο αυτών σημείων.

Κεφάλαιο 4ο

4. Υλοποίηση συστήματος

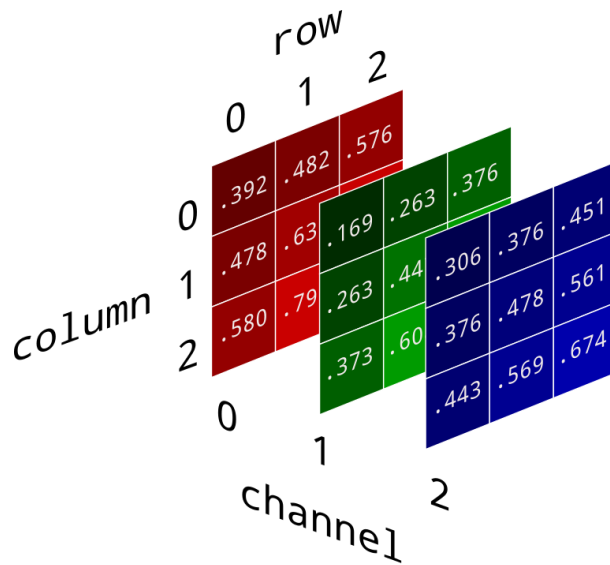
4.1 Περιγραφή εφαρμογής συστήματος

Για την εφαρμογή του αλγορίθμου που περιγράφηκε στο προηγούμενο κεφάλαιο, έγινε υλοποίηση προγράμματος το οποίο κάνει χρήση αυτού για τον υπολογισμό των επιθυμητών συντεταγμένων. Επιπροσθέτως, το πρόγραμμα αυτό πραγματοποιεί οπτικοποίηση διδαγνωστικών μεγεθών (όπως πχ το ίδιο το σημείο στην οθόνη ή τα εντοπισμένα σημεία της ίριδας των ματιών) σε ένα περιβάλλον διεπαφής, πραγματοποιεί αξιολόγηση των αποτελεσμάτων του αλγορίθμου και έχει τη δυνατότητα να καταγράφει σημεία στα οποία κοίταξε ο χρήστης σε ένα ορισμένο χρονικό διάστημα.

4.2 Εργαλεία ανάπτυξης

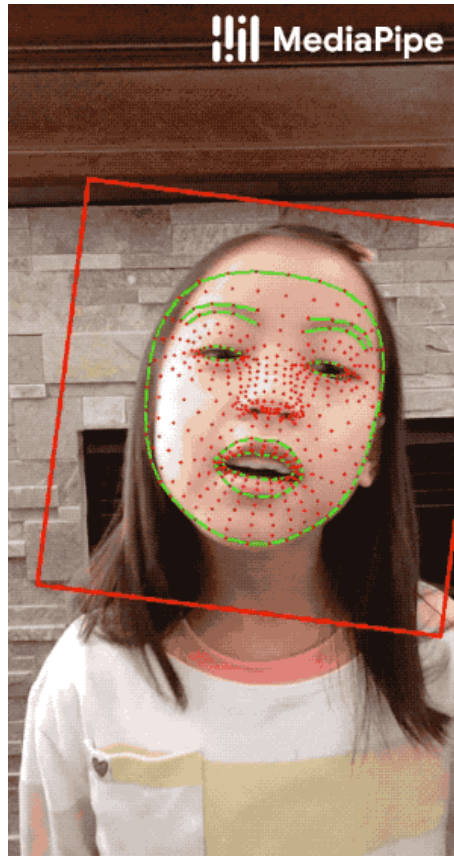
Το πρόγραμμα είναι γραμμένο στη γλώσσα Python στο περιβάλλον ανάπτυξης Pycharm. Η επιλογή της γλώσσας Python έγινε λόγω του υψηλού επιπέδου της, της ευκολίας χρήσης της και της υποστήριξής της σε πολλές χρήσιμες βιβλιοθήκες (όπως OpenCV, MediaPipe, NumPy μεταξύ άλλων) οι οποίες βοήθησαν σημαντικά στην υλοποίηση του αλγορίθμου και του όλου συστήματος.

Η βιβλιοθήκη OpenCV (Open Source Computer Vision Library) είναι μια βιβλιοθήκη η οποία παρέχει μεθόδους υπολογιστικής όρασης σε πραγματικό χρόνο. Είχε αρχικά υλοποιηθεί από την Intel και τώρα συντηρείται από την Willow Garage. Η OpenCV μπορεί να χρησιμοποιηθεί για την επεξεργασία εικόνων και βίντεο, τον εντοπισμό χαρακτηριστικών και την εκτέλεση άλλων διαδικασιών υπολογιστικής όρασης. Στην περίπτωσή μας, χρησιμοποιώντας εντολές της βιβλιοθήκης OpenCV, αρχικά λαμβάνουμε τη ζωντανή μετάδοση βίντεο της συνδεδεμένης κάμερας. Συγκεκριμένα λαμβάνουμε κάθε καρέ σε ένα πίνακα τριών διαστάσεων όπου οι δύο διαστάσεις αντιπροσωπεύουν το πλάτος και ύψος της εικόνας σε pixel, ενώ η τρίτη χωρίζει την τιμή του κάθε pixel σε κανάλια του μπλε, πράσινου και κόκκινου όπως φαίνεται στο παρακάτω σχήμα.



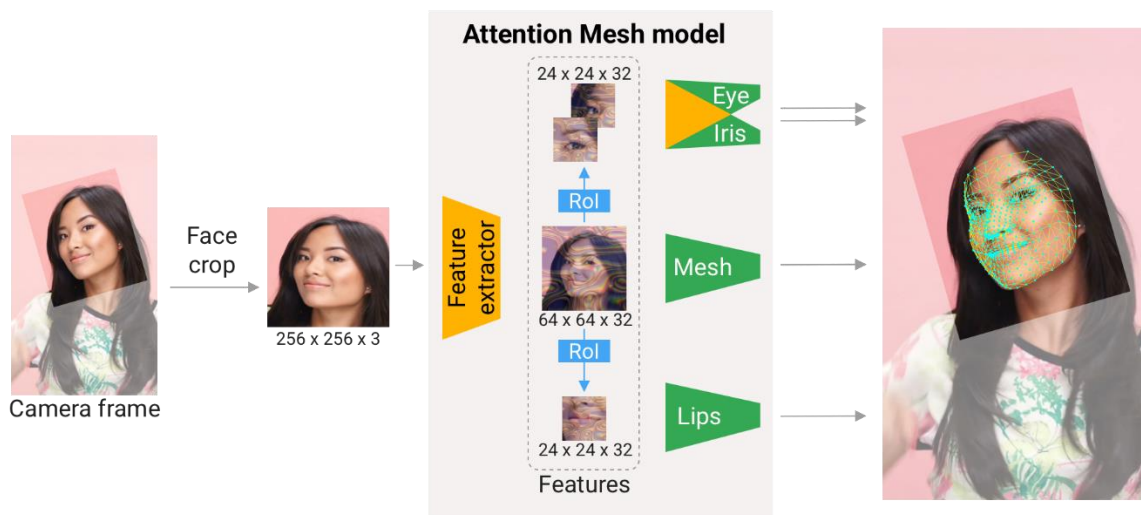
Εικόνα 11 Διαχωρισμός εικόνας σε RGB πίνακα

Η αναγνώριση προσώπου και η ανίχνευση των σημείων ενδιαφέροντός του γίνεται με χρήση της λειτουργίας face mesh της βιβλιοθήκης Mediapipe της Google. Η βιβλιοθήκη λαμβάνει ως είσοδο τον τρισδιάστατο πίνακα που αντιπροσωπεύει το κάθε καρέ του ζωντανού βίντεο και υπολογίζει 468 σημεία του προσώπου σε τρισδιάστατες συντεταγμένες. Ο υπολογισμός αυτός γίνεται με χρήση δύο μοντέλων βαθιών νευρωνικών δικτύων πραγματικού χρόνου τα οποία λειτουργούν συνεργατικά. Το πρώτο είναι υπεύθυνο για τον εντοπισμό του προσώπου και την περικοπή της εικόνας έτσι ώστε σε κάθε καρέ να παράγεται μία ευθυγραμμισμένη, απομονωμένη εικόνα του προσώπου. Η εικόνα αυτή εισάγεται στο δεύτερο μοντέλο το οποίο προβλέπει κατά προσέγγιση τις 3D θέσεις των σημείων του προσώπου. Η διαδικασία αυτή βελτιώνει δραματικά την απόδοση του συστήματος αφού εξαλείφει την ανάγκη για πράξεις όπως περιστροφή, μετατόπιση και αλλαγή κλίμακας. Επιπλέον, ο υπολογισμός της περικομμένης εικόνας από το πρώτο μοντέλο βασίζεται στα σημεία τα οποία εντοπίστηκαν στο προηγούμενο καρέ και μόνο όταν το μοντέλο σταματήσει να εντοπίζει πρόσωπο πραγματοποιεί τους πλήρεις υπολογισμούς από την αρχή.



Εικόνα 12 Πλέγμα προσώπου που προσφέρει η βιβλιοθήκη Mediapipe

Τα σημεία που δίνει το σύστημα της βιβλιοθήκης στην έξοδό του είναι σημεία ενός πλέγματος προσώπου το οποίο μορφοποιείται κατάλληλα και αντιστοιχίζεται στο πρόσωπο σε κάθε καρέ. Κάθε σημείο περιέχει τρεις συντεταγμένες όπου x και y είναι κανονικοποιημένες συντεταγμένες της εικόνας ενώ z είναι σχετικό μέγεθος όμοιας κλίμακας με το x .



Εικόνα 13 Αρχιτεκτονική βιβλιοθήκης Mediapipe

Η λειτουργία face mesh, εκτός από τα μοντέλα που περιεγράφηκαν παραπάνω, περιλαμβάνει άλλο ένα μοντέλο το οποίο εστιάζει σε συγκεκριμένες περιοχές του προσώπου και βελτιώνει την ακρίβεια του συστήματος. Οι περιοχές αυτές περιλαμβάνουν, εκτός από τα χείλη, τα μάτια και τις ίριδες των ματιών τα οποία μας ενδιαφέρουν ιδιαίτερα στην εργασία αυτή.

Αξίζει επίσης να αναφερθεί η χρήση που κάναμε στη βιβλιοθήκη Numpy. Η Numpy είναι βιβλιοθήκη η οποία βοηθά στη δημιουργία και διαχείριση πολύπλοκων δομών δεδομένων, όπως διανυσμάτων και πολυδιάστατων πινάκων, στην εκτέλεση πράξεων μεταξύ των δομών αυτών και παρέχει πληθώρα μαθηματικών συναρτήσεων και μεθόδων υψηλού επιπέδου. Κάποιες από τις πράξεις που προσφέρει είναι μαθηματικές, λογικές, διαχείρισης διαστάσεων και μεγέθους, ταξινόμησης, διακριτούς μετασχηματισμούς Fourier, γραμμικής άλγεβρας κλπ.

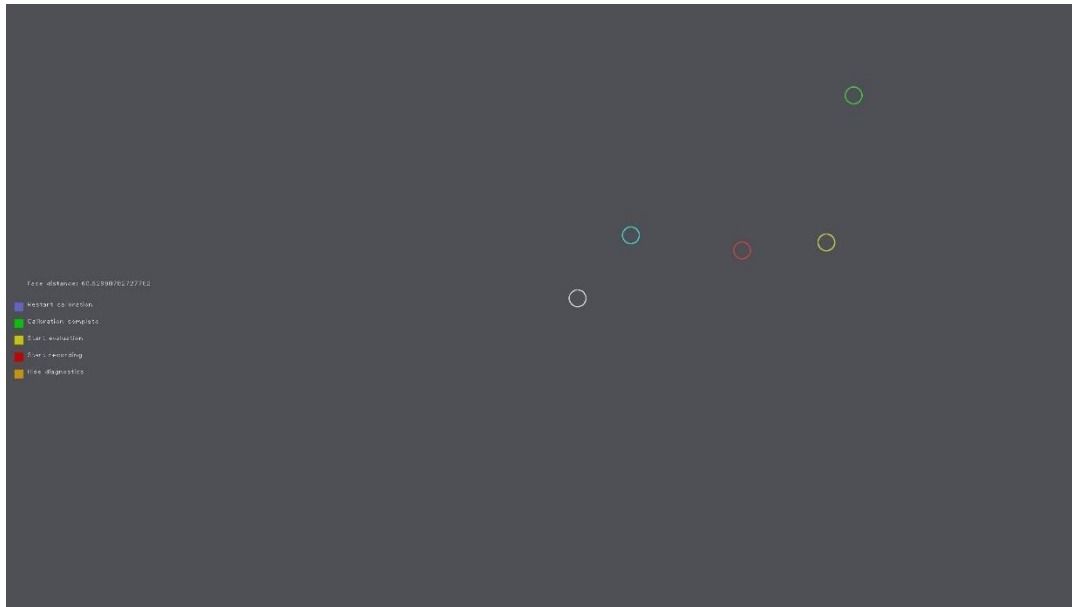
4.3 Υλοποίηση αλγορίθμου – εμφάνιση υπολογισμών

Έχοντας στη διάθεσή μας, από τα παραπάνω εργαλεία, τις θέσεις της ίριδας του κάθε ματιού στο κάθε καρέ που μας προσφέρει η κάμερα, καθώς και τα απαραίτητα landmarks του προσώπου, μπορούμε να υλοποιήσουμε τον αλγόριθμο εύρεσης του σημείου βλέμματος στην οθόνη που περιγράφηκε στο 3^ο κεφάλαιο. Για την εμφάνιση, αρχικά, των εντοπισμένων σημείων του προσώπου και της κάθε ίριδας δημιουργήθηκε ένα παράθυρο στο οποίο φαίνεται η είσοδος της κάμερας και πάνω σε αυτή σχεδιάζονται με εντολές της βιβλιοθήκης OpenCV τα σημεία αυτά, όπως φαίνεται στην παρακάτω εικόνα.



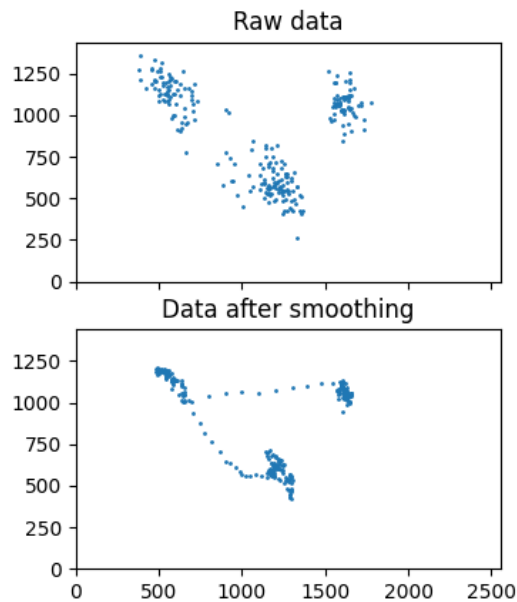
Εικόνα 14 Διαγνωστικά στοιχεία παραθύρου κάμερας. Φαίνονται τα σημεία των κέντρων των ματιών (γαλάζιο), της κάθε ίριδας (κόκκινο και πράσινο) και τα σημεία – άγκυρες του προσώπου (κίτρινο)

Για την εμφάνιση των πραγματοποιημένων υπολογισμών δημιουργήθηκε ένα παράθυρο πλήρους οθόνης στο οποίο σχεδιάζονται, με τη μορφή χρωματιστών κύκλων, τα σημεία στα οποία κοιτά το κάθε μάτι στην οθόνη, ο μέσος όρος των σημείων αυτών, αλλά και το σημείο μετατόπισης του προσώπου και το σημείο στο οποίο η ευθεία προσανατολισμού του προσώπου συναντά την οθόνη. Φαίνεται επίσης η υπολογισμένη απόσταση του προσώπου από την κάμερα με τη μορφή απλού κειμένου. Σημειώνεται ότι τα διαγνωστικά σημεία αυτά μπορούν να κρυφτούν και να ξαναεμφανιστούν με το πάτημα του πορτοκαλί κουμπιού στα αριστερά του παραθύρου (Show/Hide diagnostics).



Εικόνα 15 Διαγνωστικά στοιχεία παραθύρου οθόνης. Φαίνονται το σημείο βλέμματος του δεξιού και αριστερού ματιού (κόκκινο και πράσινο σημείο αντίστοιχα), το τελικό συνολικό σημείο του βλέμματος (κίτρινο σημείο) και το σημείο προσανατολισμού και θέσης του προσώπου (μπλε και λευκό σημείο αντίστοιχα)

Όπως φαίνεται στην παραπάνω εικόνα, ο συνδυασμός των δύο σημείων εξόδου στην οθόνη, παρόλο που αποτελεί τον μέσο όρο των δύο σημείων, δεν βρίσκεται ανάμεσά τους. Αυτό συμβαίνει γιατί το κίτρινο σημείο αναπαριστά την έξοδο του συστήματος μετά την εφαρμογή εξομάλυνσης. Αυτή προστέθηκε διότι παρατηρήθηκε, ειδικότερα σε περιπτώσεις που το πρόσωπο βρισκόταν μακριά από την κάμερα, μια έλλειψη σταθερότητας στην ένδειξη του τελικού σημείου του βλέμματος του χρήστη στην οθόνη. Έτσι εφαρμόστηκε μια τεχνική εξομάλυνσης, σύμφωνα με την οποία κάθε σημείο εξομαλυσμένης εξόδου αποτελεί μέσο όρο των 10 τελευταίων σημείων απλής εξόδου. Με τον τρόπο αυτό προστίθεται μεν μια καθυστέρηση και, θεωρητικά, κάποιο σφάλμα στο σύστημα, αλλά η έξοδος είναι πολύ πιο σταθερή και οπτικά κατανοητή από τον χρήστη. Παρακάτω παρουσιάζεται η απλή και η εξομαλυσμένη έξοδος του συστήματος για μια τυχαία καταγραφή του βλέμματος ενός χρήστη.

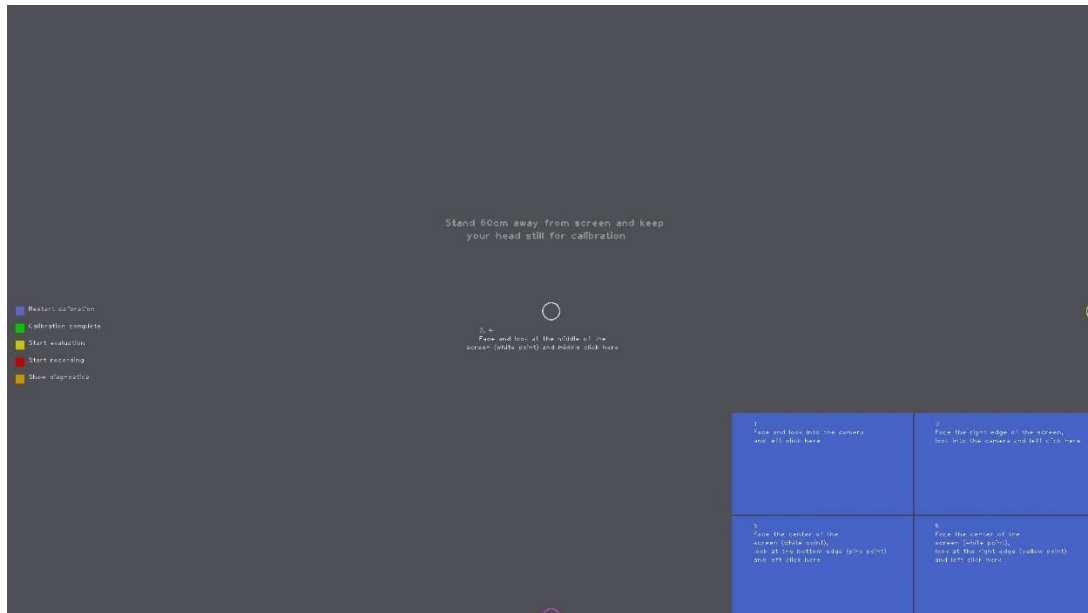


Σχήμα 1 Δεδομένα εξόδου του συστήματος πριν την ομαλοποίηση παρελθοντικών τιμών (επάνω σχήμα) και μετά την ομαλοποίηση (κάτω σχήμα)

Παρατηρούμε ότι μετά την εξομάλυνση, φαίνονται πιο ξεκάθαρα όχι μόνο τα σημεία εστίασης της παρατήρησης, αλλά και η διαδρομή που ακολουθήθηκε από το βλέμμα του χρήστη.

4.4 Βαθμονόμηση

Το στάδιο της βαθμονόμησης πραγματοποιείται στην αρχή της ροής του προγράμματος, καθώς παρέχει απαραίτητες παραμέτρους για τη λειτουργία του αλγορίθμου. Στο στάδιο αυτό ζητείται από τον χρήστη να κοιτάξει σε συγκεκριμένα σημεία της οθόνης ή να γείρει το πρόσωπό του προς αυτά, και στη συνέχεια να κάνει κλικ σε συγκεκριμένα κουμπιά στο παράθυρο της πλήρους οθόνης. Για την πιο εύκολη εύρεση των σημείων στα οποία καλείται να εστιάσει ο χρήστης την προσοχή του, έχουμε σχεδιάσει στο παράθυρο τα σημεία αυτά με μορφή κύκλων με ξεχωριστό χρώμα. Στην παρακάτω εικόνα φαίνεται το παράθυρο πλήρους οθόνης κατά το στάδιο της βαθμονόμησης.



Εικόνα 16 Διάταξη παραθύρου οθόνης στο στάδιο της βαθμονόμησης. Φαίνονται κάτω δεξιά τα τέσσερα μπλε κουμπιά, και τα βοηθητικά σημεία στο κέντρο και τις άκρες της οθόνης.

Οι οδηγίες του σταδίου βαθμονόμησης, έχουν επιλεγθεί έτσι ώστε να συλλέγονται όλες οι απαραίτητες πληροφορίες, με τα ελάχιστα δυνατά βήματα. Οι οδηγίες αυτές περιλαμβάνουν τα εξής:

- Αρχικά, ζητείται από τον χρήστη να φέρει το πρόσωπό του σε απόσταση περίπου 60 εκατοστών από την οθόνη και να διατηρήσει το κεφάλι του όσο πιο σταθερό γίνεται. Ο λόγος για τον οποίο ζητάμε από τον χρήστη να σταθεί σε 60 εκατοστά από την κάμερα είναι έτσι ώστε να λάβουμε την απόσταση μεταξύ δύο άκρων του προσώπου. Ξέροντας το μέγεθος αυτό σε απόσταση 60 εκατοστών, μπορούμε να υπολογίσουμε την απόσταση του προσώπου από την οθόνη ανάλογα με την αλλαγή του στο κάθε καρέ.
- Το πρώτο μπλε κουμπί ζητά από τον χρήστη να κοιτάξει στην κάμερα και στη συνέχεια να πατηθεί. Έτσι βρίσκουμε τις ευθείες στις οποίες βρίσκεται το κέντρο του κάθε ματιού.
- Στη συνέχεια, ζητείται από τον χρήστη να κατευθύνει το πρόσωπο και το βλέμμα του στο κέντρο της οθόνης, το οποίο παρουσιάζεται με έναν λευκό κύκλο, και να πατήσει το μεσαίο κουμπί του ποντικιού οπουδήποτε στην οθόνη. Εδώ χρησιμοποιώντας την θέση των ματιών παίρνουμε πληροφορίες για τη σχετική θέση του κέντρου της οθόνης.
- Το δεύτερο κουμπί ζητά από τον χρήστη να κατευθύνει το πρόσωπό του στο δεξιό άκρο της οθόνης, το οποίο παρουσιάζεται με έναν κίτρινο κύκλο, και να κατευθύνει το βλέμμα του στο κέντρο της οθόνης πριν το πατήσει. Με τον τρόπο αυτό λαμβάνουμε πληροφορίες για το βάθος του κέντρου των ματιών στο πρόσωπο του χρήστη όπως αναφέραμε και στο 3^ο κεφάλαιο. Έχοντας τις ευθείες που βρέθηκαν στο πάτημα του πρώτου κουμπιού και τις νέες ευθείες με τον διαφορετικό προσανατολισμό προσώπου, τα σημεία των κέντρων των ματιών βρίσκονται στα σημεία τομής των αντίστοιχων ευθειών.

- Στο επόμενο στάδιο ζητείται και πάλι από τον χρήστη να κατευθύνει το πρόσωπο και το βλέμμα του στο κέντρο της οθόνης και να πατήσει το μεσαίο κουμπί του ποντικιού οπουδήποτε στην οθόνη. Το βήμα αυτό επαναλαμβάνεται για εξάλειψη πιθανών σφαλμάτων.
- Στη συνέχεια, ζητείται από τον χρήστη να κατευθύνει το πρόσωπό του στο κέντρο της οθόνης και το βλέμμα του στο κάτω άκρο της, το οποίο σηματοδοτείται με έναν ροζ κύκλο, και στη συνέχεια να πατήσει το τρίτο μπλε κουμπί. Εδώ λαμβάνουμε πληροφορίες για τον προσανατολισμό της οθόνης, το κάθετο μέγεθός της και την απόστασή της.
- Τέλος, το τέταρτο μπλε κουμπί ζητά από τον χρήστη να κατευθύνει το πρόσωπό του στο κέντρο της οθόνης και το βλέμμα του στο δεξιό άκρο της πριν πατηθεί. Έτσι λαμβάνουμε τις αντίστοιχες πληροφορίες με το προηγούμενο βήμα, αλλά για το οριζόντιο άξονα.

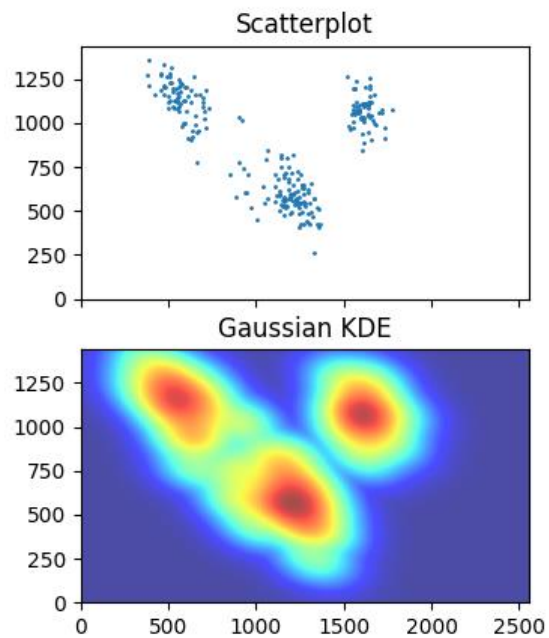
Έτσι λαμβάνονται όλες οι απαραίτητες πληροφορίες για τη λειτουργία του αλγορίθμου του συστήματος που υλοποιήθηκε. Με το πάτημα του πράσινου κουμπιού στα αριστερά του παραθύρου (Calibration complete) τελειώνει το στάδιο της βαθμονόμησης και εξαφανίζονται τα βοηθητικά σημεία, οι οδηγίες και τα κουμπιά που το συνόδευσαν. Σε περίπτωση που ο χρήστης διαπιστώσει ότι υπήρξε σφάλμα στη βαθμονόμηση μπορεί να την επαναλάβει πατώντας το μωβ κουμπί στα αριστερά του παραθύρου (Restart calibration).

4.5 Διαδικασία καταγραφής

Η διαδικασία καταγραφής ξεκινά με το πάτημα του κόκκινου κουμπιού στα αριστερά της οθόνης (Start recording). Κατά την εκκίνηση, προβάλλεται μία τυχαία εικόνα από ένα σύνολο επιλεγμένων εικόνων σε πλήρη οθόνη. Ταυτόχρονα ξεκινάει η καταγραφή των διαδοχικών σημείων του βλέμματος του χρήστη στην οθόνη. Επομένως, για κάθε καρτέ γίνεται ο υπολογισμός των συντεταγμένων στην οθόνη από το υλοποιημένο σύστημα και αποθηκεύεται σε ένα αρχείο csv. Μόλις ο χρήστης κάνει κλικ σε οποιοδήποτε σημείο της εικόνας η καταγραφή σταματά και ξεκινά η δημιουργία των οπτικοποιήσεων.

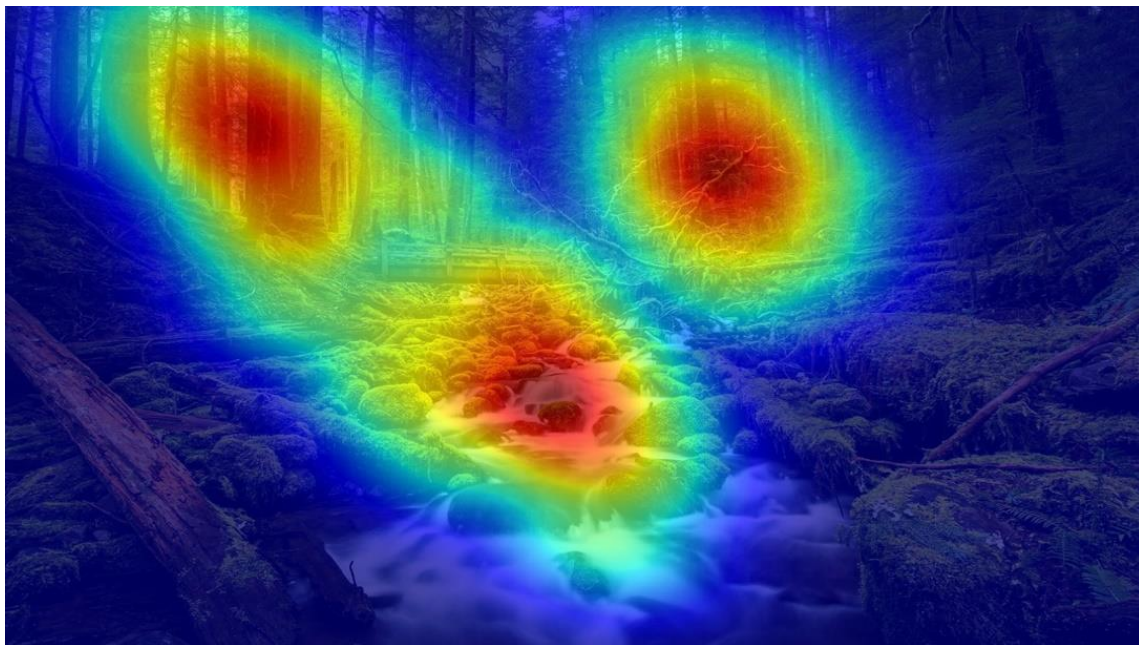
4.6 Δημιουργία οπτικοποιήσεων

Μετά την καταγραφή καταλήγουμε με μία λίστα των σημείων της οθόνης στα οποία κοίταξε ο χρήστης στο χρονικό διάστημα της καταγραφής. Στη λίστα αυτή εφαρμόζουμε την εκτίμηση πυκνότητας πυρήνων με χρήση Γκαουσιανών πυρήνων (Gaussian Kernel Density Estimator). Η διαδικασία αυτή χρησιμοποιείται για την εκτίμηση της συνάρτησης πυκνότητας πιθανότητας μίας τυχαίας μεταβλητής, χρησιμοποιώντας ως βάρη Γκαουσιανούς πυρήνες. Στην περίπτωσή μας, μετατρέπει το σύνολο των σημείων της οθόνης στα οποία κοίταξε ο χρήστης σε μία συνάρτηση δύο μεταβλητών, η οποία έχει πιο υψηλές τιμές στις περιοχές της οθόνης που βρίσκονται πολλά συνοσπισμένα σημεία και πιο χαμηλές τιμές σε περιοχές με λιγότερα σημεία. Αντιστοιχώντας το πεδίο τιμών της συνάρτησης αυτής σε ένα χάρτη χρωμάτων παίρνουμε τον επιθυμητό θερμοχάρτη, όπως φαίνεται παρακάτω.



Σχήμα 2 Φαίνονται (πάνω) τα δεδομένα εξόδου του συστήματος και (κάτω) ο θερμοχάρτης που παράγεται από το Gaussian KDE

Ο θερμοχάρτης αυτός τοποθετείται πάνω από την αρχική εικόνα που παρατηρήθηκε από τον χρήστη με μειωμένη ορατότητα, έτσι ώστε να φανερωθούν τα σημεία ενδιαφέροντός της.



Εικόνα 17 Ο θερμοχάρτης που παράγεται από την καταγραφή τοποθετημένος πάνω από την εικόνα που παρατηρήθηκε

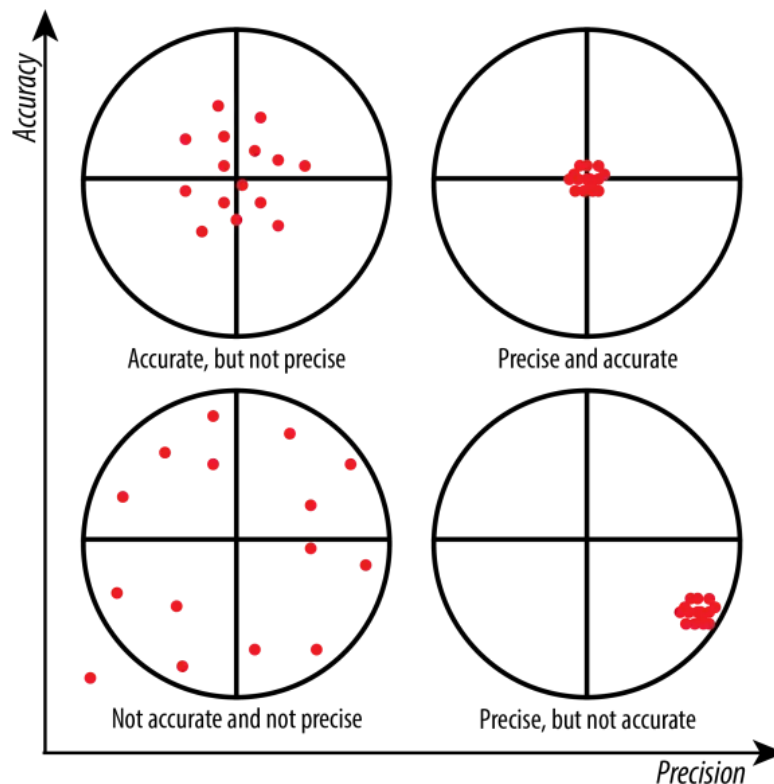
Περισσότερες οπτικοποιήσεις προσφέρονται στην διαδικτυακή υλοποίηση της εφαρμογής σε ιστοσελίδα που παρουσιάζεται σε επόμενο κεφάλαιο.

4.7 Διαδικασία αξιολόγησης

Η διαδικασία της αξιολόγησης του υλοποιημένου συστήματος, η οποία ξεκινά με το πάτημα του κίτρινου κουμπιού στα αριστερά του παραθύρου (Start evaluation), πραγματοποιήθηκε με βάση τη διαδικασία αξιολόγησης επαγγελματικών συσκευών εντοπισμού βλέμματος. Αυτό γίνεται για να έχουμε ένα αξιόπιστο σημείο αναφοράς μετρικών έτσι ώστε να μπορούμε να συγκρίνουμε τα αποτελέσματά μας με ένα σύστημα το οποίο έχει θεωρητικά τις καλύτερες δυνατές μετρικές.

4.7.1 Μετρικές

Για την σωστή αξιολόγηση του συστήματος χρησιμοποιήθηκαν μετρικές ακρίβειας (precision) αλλά και ορθότητας (accuracy) της εξόδου του. Οι μετρικές αυτές υπολογίζονται τόσο μονοφθαλμικά, για το κυρίαρχο μάτι, όσο και διοφθαλμικά. Το κυρίαρχο μάτι του χρήστη βρίσκεται ζητώντας από αυτόν να κρατήσει μία κόλλα χαρτί με μία τρύπα διαμέτρου τριών εκατοστών με τεντωμένα χέρια, και να κοιτάξει μέσα από αυτή σε ένα μακρινό αντικείμενο. Μετά ζητείται από τον χρήστη να κλείσει το κάθε μάτι με τη σειρά. Το μάτι με το οποίο βλέπει όντως το αντικείμενο μέσα από την τρύπα είναι το κυρίαρχο μάτι.



Σχήμα 3 Διαφορά ακρίβειας (precision) και ορθότητας (accuracy)

Για τον υπολογισμό της ορθότητας αλλά και της ακρίβειας λαμβάνονται 30 μετρήσεις σε pixel για το κάθε σημείο που ελέγχεται. Η ορθότητα είναι μετρική η οποία μετρά το πόσο

κοντά μία μέτρηση της εξόδου είναι στην πραγματική τιμή. Επομένως από αυτές τις 30 μετρήσεις βρίσκουμε το σημείο που αντιπροσωπεύει τον μέσο όρο τους και στη συνέχεια υπολογίζουμε την απόσταση σε ρίxel αυτού του σημείου από το πραγματικό σημείο που ελέγχεται κι έτσι βρίσκουμε την μονοφθαλμική ορθότητα. Για να βρούμε την διοφθαλμική ορθότητα σε ρίxel παίρνουμε τον μέσο όρο του ζεύγους της μέτρησης δεξιού και αριστερού ματιού και στη συνέχεια επαναλαμβάνουμε την παραπάνω διαδικασία. Για να βρούμε την ορθότητα σε χιλιοστά στην οθόνη προσαρμόζουμε την τιμή που βρήκαμε σε ρίxel ανάλογα με την πυκνότητα ρίxel της οθόνης. Τέλος για να βρούμε την ορθότητα σε γωνία του οφθαλμού χρησιμοποιούμε την απόστασή του από την οθόνη.

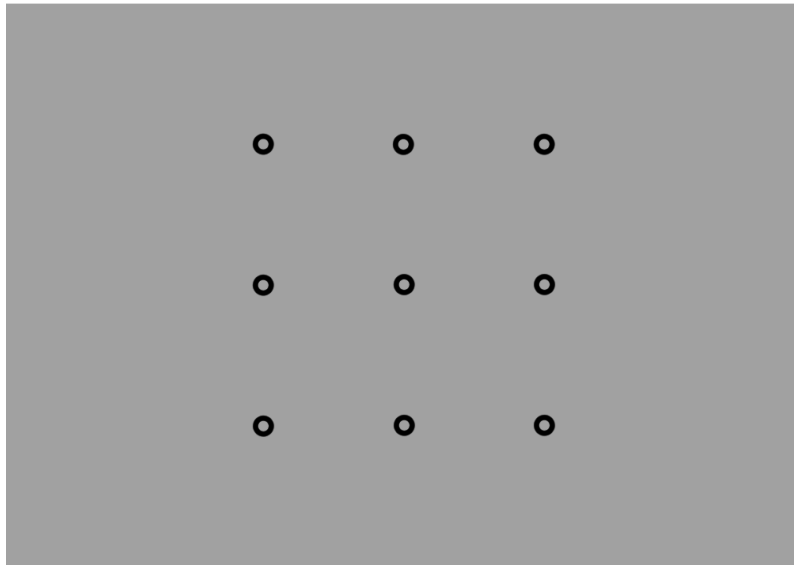
Η ακρίβεια είναι μετρική η οποία μετρά το πόσο κοντά είναι ένας αριθμός μετρήσεων της εξόδου μεταξύ τους και επομένως την ικανότητα του συστήματος να αναπαράγει αξιόπιστα τα ίδια αποτελέσματα με την ίδια είσοδο. Η τιμή της ακρίβειας σε ρίxel συμπίπτει με τη ρίζα του μέσου όρου των τετραγώνων των διαφορών συνεχόμενων μετρήσεων. Αντίστοιχα η τιμή σε γωνία του οφθαλμού υπολογίζεται ως ο μέσος όρος των τετραγώνων των διαφορών σε γωνία συνεχόμενων μετρήσεων. Τέλος υπολογίζεται η ακρίβεια τυπικής απόκλισης η οποία συμπίπτει, με την τυπική απόκλιση των σημείων από τον μέσο όρο τους.

4.7.2 Συνθήκες-στάδια αξιολόγησης

Η ακρίβεια και η ορθότητα του συστήματος ελέγχονται σε διάφορες συνθήκες λειτουργίας έτσι ώστε να αποκτήσουμε μία πιο σφαιρική εικόνα για την απόδοσή του. Για να γίνει αυτό δημιουργήσαμε τέσσερα στάδια αξιολόγησης καθένα από τα οποία ελέγχουν την απόδοση του συστήματος σε ειδικά επιλεγμένες συνθήκες. Οι συνθήκες αυτές επιλέχθηκαν έτσι ώστε να είναι όσο το δυνατό πιο όμοιες με αυτές που χρησιμοποιούνται στον έλεγχο επαγγελματικών συσκευών. Κάθε ένα από τα επόμενα υποκεφάλαια αναλύει τα στάδια αυτά.

4.7.2.1 Ιδανικές συνθήκες

Στο στάδιο αυτό αξιολογούμε το σύστημα σε ιδανικές συνθήκες γωνίας θέασης, φωτισμού και θέσης προσώπου του χρήστη. Συγκεκριμένα, επιλέγουμε να ελεγχθούν σημεία της οθόνης στα οποία η γωνία θέασης δεν ξεπερνά τις είκοσι μοίρες, δηλαδή βρίσκονται κοντά στο κέντρο της οθόνης. Η διάταξη των σημείων αυτών φαίνεται στην παρακάτω εικόνα.



Εικόνα 18 Ιδανικά σημεία παρατήρησης για το στάδιο της αξιολόγησης

Φροντίζουμε το δωμάτιο στο οποίο γίνεται η αξιολόγηση να είναι καλά φωτισμένο και ο χρήστης να διατηρεί την θέση του προσώπου του σταθερή και ευθυγραμμισμένη με την οθόνη κατά τη διάρκεια της λήψης των μετρήσεων. Στο στάδιο αυτό αναμένουμε να έχουμε τα καλύτερα δυνατά αποτελέσματα ορθότητας, ακρίβειας και ακρίβειας τυπικής απόκλισης σε σχέση με όλα τα άλλα.

4.7.2.2 Ακραία σημεία οθόνης

Για να εξετάσουμε την αποδοτικότητα του συστήματος στην περίπτωση που ο χρήστης κοιτάει σε ακραία μέρη της οθόνης, αλλάζουμε τα σημεία τα οποία ελέγχονται έτσι ώστε να βρίσκονται στα μέρη αυτά. Η νέα διάταξη των σημείων φαίνεται παρακάτω.



Εικόνα 19 Ακραία σημεία παρατήρησης για το στάδιο της αξιολόγησης

Η θέση του προσώπου του χρήστη πρέπει να παραμείνει σταθερή όπως και στο στάδιο ιδανικών συνθηκών. Εδώ αναμένουμε να πάρουμε χειρότερα αποτελέσματα ορθότητας και ακρίβειας σε σχέση με το στάδιο των ιδανικών συνθηκών.

4.7.2.3 Μεταβαλλόμενες συνθήκες φωτισμού

Σε αυτό το στάδιο εστιάζουμε στις κακές συνθήκες φωτισμού, οπότε αρχικά φροντίζουμε στο περιβάλλον του χρήστη να μην υπάρχουν σημαντικές πηγές φωτός πέρα από την ίδια την οθόνη. Η διάταξη των σημείων που ελέγχουμε είναι ίδια με αυτή στο στάδιο ιδανικών συνθηκών και ο χρήστης διατηρεί το κεφάλι του σε σταθερή ευθυγραμμισμένη θέση. Στο στάδιο αυτό αναμένουμε αποτελέσματα κοντά σε αυτά των ιδανικών συνθηκών, αφού το σύστημα εντοπισμού προσώπου και ματιών της βιβλιοθήκης *mediarpipe* λειτουργεί εξίσου καλά σε συνθήκες χαμηλού αλλά και υψηλού φωτισμού και επομένως μας εξαλείφει θεωρητικά τις συνθήκες φωτισμού σαν παράγοντα.

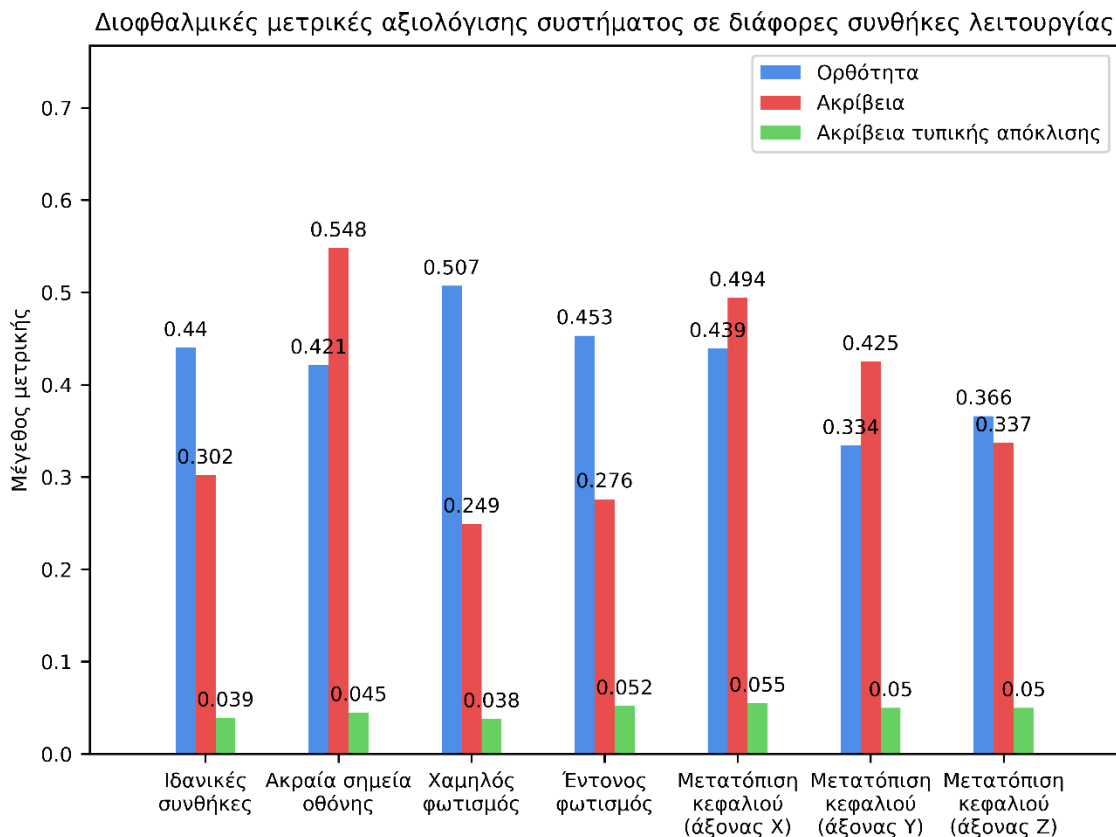
Στη συνέχεια επαναλαμβάνεται ο έλεγχος αλλά με φωτισμό πολύ μεγαλύτερο από αυτό του κοινού δωματίου. Πιο συγκεκριμένα τοποθετούμε μια πηγή φωτός πίσω από το κεφάλι του χρήστη έτσι ώστε να δυσκολευτεί η κάμερα να αποτυπώσει ξεκάθαρα τα στοιχεία του προσώπου του, εξετάζοντας έτσι τη χειρότερη δυνατή περίπτωση. Εδώ αναμένουμε πιθανή πτώση της απόδοσης του συστήματος, καθώς άμα δεν αποτυπώνονται ξεκάθαρα τα στοιχεία του προσώπου στην εικόνα που δίνει η κάμερα, ακόμα και η βιβλιοθήκη *mediarpipe* αδυνατεί να δώσει αξιόπιστα αποτελέσματα.

4.7.2.4 Μεταβαλλόμενη θέση προσώπου

Σε σύγκριση με το στάδιο ιδανικών συνθηκών, στο στάδιο μεταβαλλόμενης θέσης προσώπου, όπως προτείνει και το όνομα, κρατάμε όλες τις συνθήκες ίδιες εκτός από τη θέση του κεφαλιού του χρήστη. Ο έλεγχος πραγματοποιείται ξεχωριστά για την κάθε διάσταση του χώρου. Σε κάθε έλεγχο μετατοπίζουμε το κεφάλι του χρήστη σε διαστήματα των πέντε εκατοστών στον άξονα της αντίστοιχης διάστασης. Εδώ αναμένουμε τα αποτελέσματα να είναι τα χειρότερα από όλα τα άλλα στάδια, αφού εμπειρικά το σύστημα χάνει αρκετή από την απόδοσή του με την κίνηση του κεφαλιού. Παρόλα αυτά αναμένουμε να είναι και πάλι αρκετά ικανοποιητικά.

4.7.3 Αποτελέσματα αξιολόγησης

Αρχικά παραθέτουμε μία σύνοψη των αποτελεσμάτων για να γίνει πιο οπτικά ξεκάθαρη η απόδοση του συστήματος στις διάφορες συνθήκες λειτουργίας. Αυτό γίνεται με ένα ομαδοποιημένο ραβδόγραμμα το οποίο περιλαμβάνει τα αποτελέσματα της κάθε μίας μετρικής στις αντίστοιχες συνθήκες.



Σχήμα 4 Συνολικό γράφημα αποτελεσμάτων αξιολόγησης

Παρακάτω παρουσιάζονται αναλυτικά τα αποτελέσματα των μετρικών που περιεγράφηκαν στο κάθε αντίστοιχο στάδιο αξιολόγησης.

4.7.3.1 Ιδανικές συνθήκες

Αρχικά παρατηρούμε τα αποτελέσματα της αξιολόγησης σε ιδανικές συνθήκες λειτουργίας. Τα αποτελέσματα αυτά θα συγκρίνουμε με αυτά των άλλων σταδίων αξιολόγησης για να βγάλουμε κάποια αρχικά συμπεράσματα για την αξιοπιστία του συστήματος. Στο σημείο αυτό σημειώνουμε ότι μικρότερες τιμές μετρικών ορθότητας, ακρίβειας και τυπικής απόκλισης, σημαίνουν μεγαλύτερη αξιοπιστία του συστήματος.

Πίνακας 1 Μονοφθαλμικά και διοφθαλμικά αποτελέσματα αξιολόγησης σε ιδανικές συνθήκες

	Ορθότητα	Ακρίβεια	Ακρίβεια τυπικής απόκλισης
Μονοφθαλμική	0.401	0.349	-
Διοφθαλμική	0.44	0.302	0.039

4.7.3.2 Ακραία σημεία οθόνης

Παρατηρούμε ότι στην αξιολόγηση ακραίων θέσεων της οθόνης η ορθότητα παραμένει παρόμοια με την αξιολόγηση ιδανικών συνθηκών, αλλά τόσο η ακρίβεια, όσο και η ακρίβεια τυπικής απόκλισης έχουν αυξηθεί σημαντικά. Το γεγονός αυτό είναι αναμενόμενο, καθώς το σύστημα δυσκολεύεται στον εντοπισμό του βλέμματος στις θέσεις αυτές.

Πίνακας 2 Μονοφθαλμικά και διοφθαλμικά αποτελέσματα αξιολόγησης σε ακραία σημεία της οθόνης

	Ορθότητα	Ακρίβεια	Ακρίβεια τυπικής απόκλισης
Μονοφθαλμική	0.383	0.571	-
Διοφθαλμική	0.421	0.548	0.045

4.7.3.3 Συνθήκες χαμηλού φωτισμού

Στο στάδιο αυτό παρατηρούμε αυξημένη ορθότητα, τόσο μονοφθαλμική όσο και διοφθαλμική, και μειωμένη ακρίβεια.

Πίνακας 3 Μονοφθαλμικά και διοφθαλμικά αποτελέσματα αξιολόγησης σε συνθήκες χαμηλού φωτισμού

	Ορθότητα	Ακρίβεια	Ακρίβεια τυπικής απόκλισης
Μονοφθαλμική	0.492	0.266	-

Διοφθαλμική	0.507	0.249	0.038
-------------	-------	-------	-------

4.7.3.4 Συνθήκες έντονου φωτισμού

Στις συνθήκες έντονου φωτισμού παρατηρούμε μειωμένη ακρίβεια αλλά αυξημένη ακρίβεια τυπικής απόκλισης.

Πίνακας 4 Μονοφθαλμικά και διοφθαλμικά αποτελέσματα αξιολόγησης σε συνθήκες έντονου φωτισμού

	Ορθότητα	Ακρίβεια	Ακρίβεια τυπικής απόκλισης
Μονοφθαλμική	0.407	0.292	-
Διοφθαλμική	0.453	0.276	0.052

4.7.3.5 Μεταβαλλόμενη θέση προσώπου (άξονας Χ)

Στο στάδιο αυτό, όπως και σε όλα τα στάδια μεταβαλλόμενης θέσης προσώπου, τα αποτελέσματα είναι εμφανώς χειρότερα από αυτά των ιδανικών συνθηκών. Παρόλο που παρατηρείται μία βελτίωση σε κάποιες μετρικές, τα συνολικά αποτελέσματα δείχνουν μείωση της απόδοσης του συστήματος. Η μείωση αυτή είναι αναμενόμενη όπως αναφέρθηκε προηγουμένως στο κεφάλαιο αυτό.

Πίνακας 5 Μονοφθαλμικά και διοφθαλμικά αποτελέσματα αξιολόγησης με το κεφάλι του χρήστη μετατοπισμένο προς τα αριστερά

Αριστερά του κέντρου	Ορθότητα	Ακρίβεια	Ακρίβεια τυπικής απόκλισης
Μονοφθαλμική	0.519	0.867	-
Διοφθαλμική	0.524	0.685	0.049

Πίνακας 6 Μονοφθαλμικά και διοφθαλμικά αποτελέσματα αξιολόγησης με το κεφάλι του χρήστη μετατοπισμένο προς τα δεξιά

Δεξιά του κέντρου	Ορθότητα	Ακρίβεια	Ακρίβεια τυπικής απόκλισης
Μονοφθαλμική	0.289	0.578	-
Διοφθαλμική	0.354	0.303	0.061

4.7.3.6 Μεταβαλλόμενη θέση προσώπου (άξονας Υ)

Πίνακας 7 Μονοφθαλμικά και διοφθαλμικά αποτελέσματα αξιολόγησης με το κεφάλι του χρήστη μετατοπισμένο προς τα επάνω

Πάνω από το κέντρο	Ορθότητα	Ακρίβεια	Ακρίβεια τυπικής απόκλισης
Μονοφθαλμική	0.489	0.166	-
Διοφθαλμική	0.496	0.189	0.039

Πίνακας 8 Μονοφθαλμικά και διοφθαλμικά αποτελέσματα αξιολόγησης με το κεφάλι του χρήστη μετατοπισμένο προς τα κάτω

Κάτω από το κέντρο	Ορθότητα	Ακρίβεια	Ακρίβεια τυπικής απόκλισης
Μονοφθαλμική	0.165	0.632	-
Διοφθαλμική	0.173	0.668	0.062

4.7.3.7 Μεταβαλλόμενη θέση προσώπου (άξονας Ζ)

Πίνακας 9 Μονοφθαλμικά και διοφθαλμικά αποτελέσματα αξιολόγησης με το κεφάλι του χρήστη μετατοπισμένο πιο κοντά στην οθόνη

Απόσταση <60cm	Ορθότητα	Ακρίβεια	Ακρίβεια τυπικής απόκλισης
Μονοφθαλμική	0.477	0.194	-
Διοφθαλμική	0.47	0.181	0.021

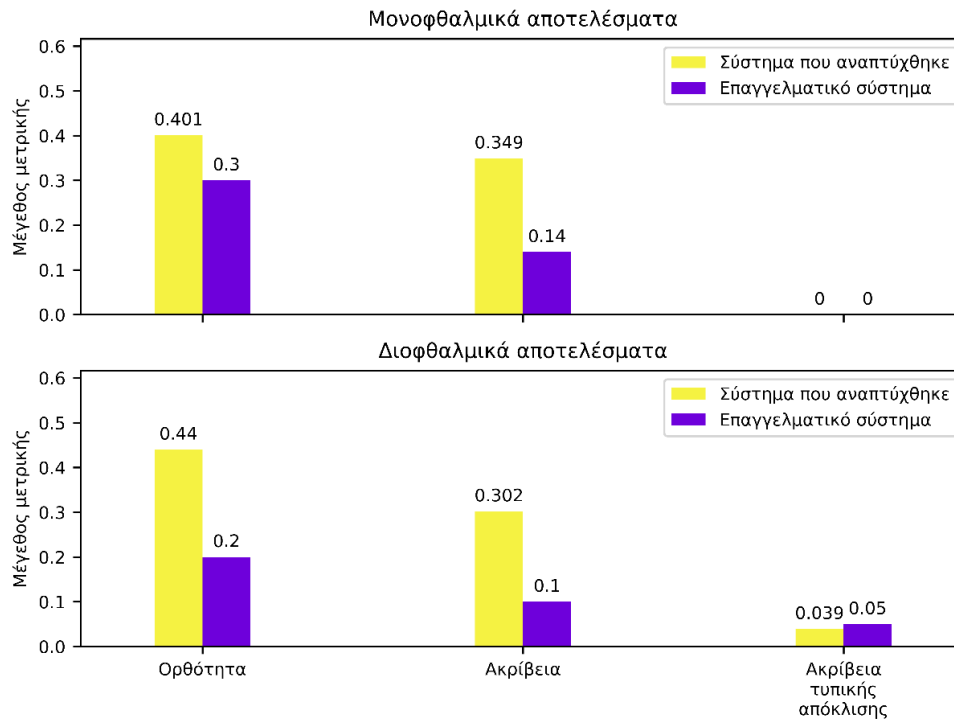
Πίνακας 10 Μονοφθαλμικά και διοφθαλμικά αποτελέσματα αξιολόγησης με το κεφάλι του χρήστη μετατοπισμένο πιο μακριά από την οθόνη

Απόσταση >60cm	Ορθότητα	Ακρίβεια	Ακρίβεια τυπικής απόκλισης
Μονοφθαλμική	0.212	0.664	-
Διοφθαλμική	0.292	0.493	0.08

4.7.3.8 Σύγκριση αποτελεσμάτων με επαγγελματική συσκευή

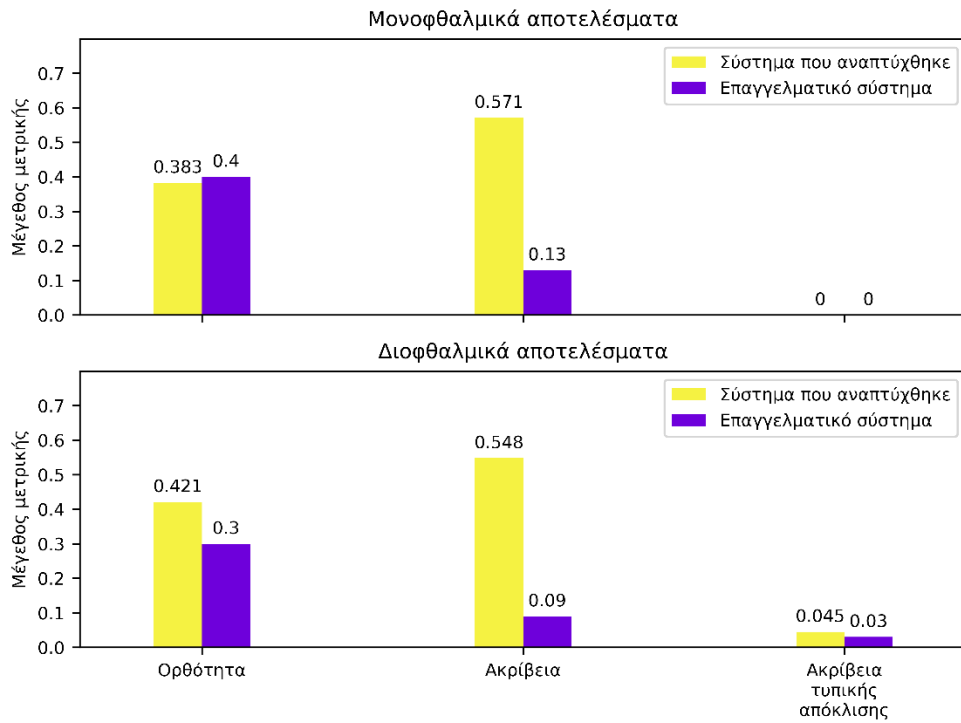
Στο υπό κεφάλαιο αυτό συγκρίνουμε τα αποτελέσματα των μετρικών που βρήκαμε για το σύστημά μας με αυτά μιας επαγγελματικής συσκευής εντοπισμού βλέμματος. Συγκεκριμένα, η συσκευή είναι η Tobii pro nano και αποτελεί μοντέλο μίας από τις πιο έμπιστες και ευρέως χρησιμοποιούμενες μάρκες στην αγορά. Η απόδοσή της αποτελεί άξιο σημείο σύγκρισης για το σύστημά μας. Σημειώνουμε και πάλι ότι μικρότερες τιμές της κάθε μετρικής σημαίνουν καλύτερη απόδοση του συστήματος.

Σε ιδανικές συνθήκες βλέπουμε ότι η επαγγελματική συσκευή αποδίδει καλύτερα από το σύστημά μας με αρκετή διαφορά τόσο στα μονοφθαλμικά όσο και στα διοφθαλμικά αποτελέσματα.



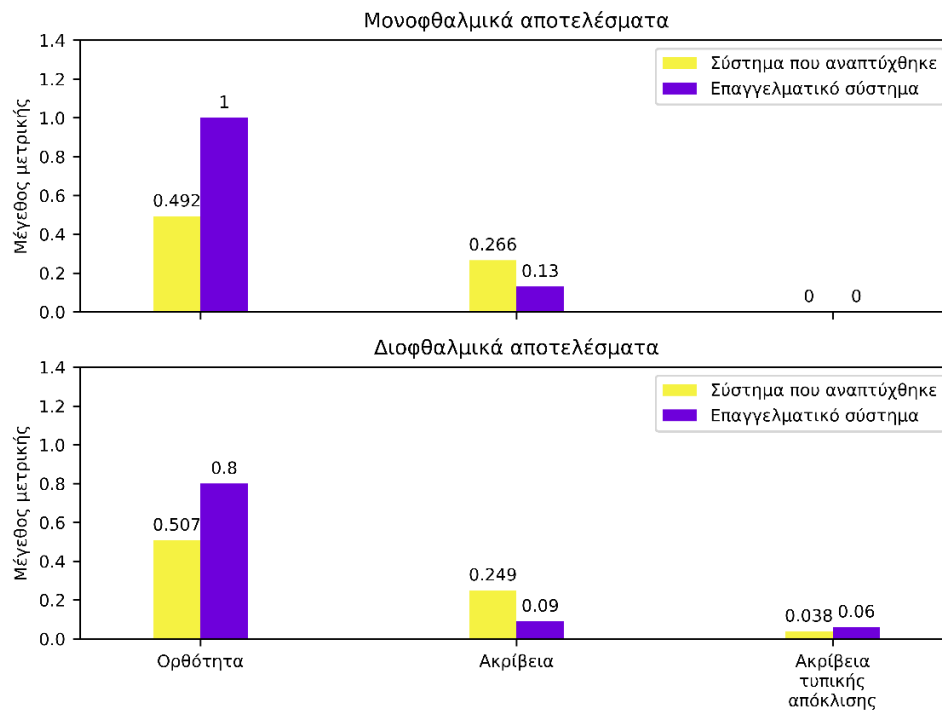
Σχήμα 5 Μονοφθαλμικά (επάνω) και διοφθαλμικά (κάτω) αποτελέσματα αξιολόγησης του συστήματός μας σε ιδανικές συνθήκες, συγκριτικά με επαγγελματικό σύστημα

Στα ακραία σημεία της οθόνης τα δύο συστήματα φαίνονται να αποδίδουν παρόμοια όσον αφορά την ορθότητα, αλλά το σύστημά μας υστερεί σημαντικά στην ακρίβειά του.



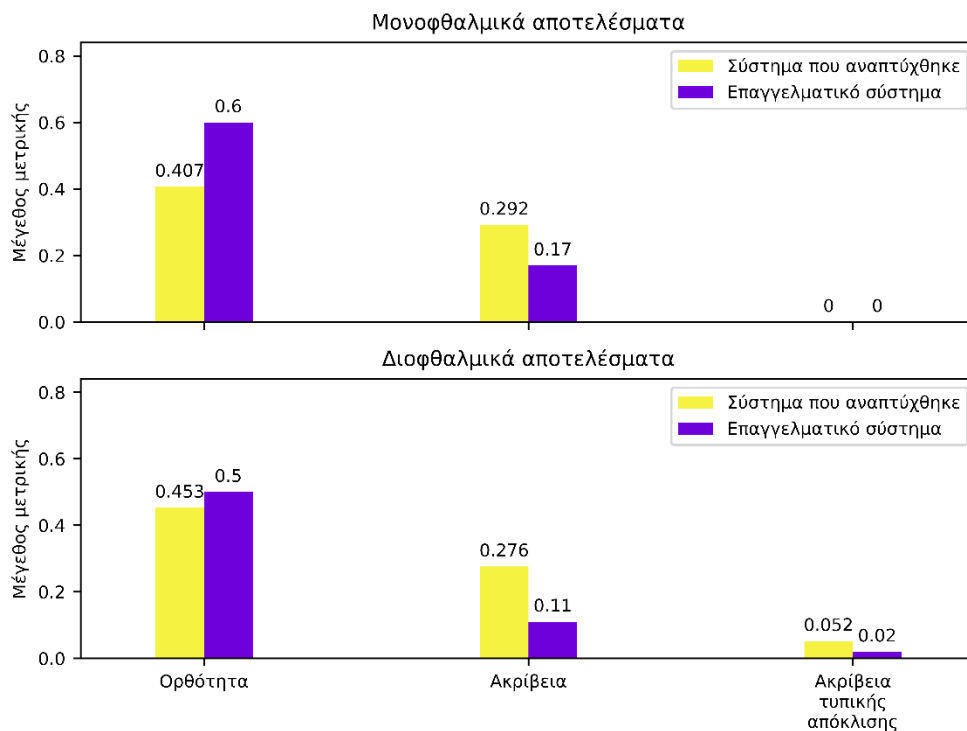
Σχήμα 6 Μονοφθαλμικά (επάνω) και διοφθαλμικά (κάτω) αποτελέσματα αξιολόγησης του συστήματός μας σε ακραία σημεία της οθόνης, συγκριτικά με επαγγελματικό σύστημα

Σε σκοτεινές συνθήκες φωτισμού βλέπουμε ότι το σύστημά μας προσφέρει καλύτερες τιμές ορθότητας και χειρότερες, αλλά συγκρίσιμες, τιμές ακρίβειας.



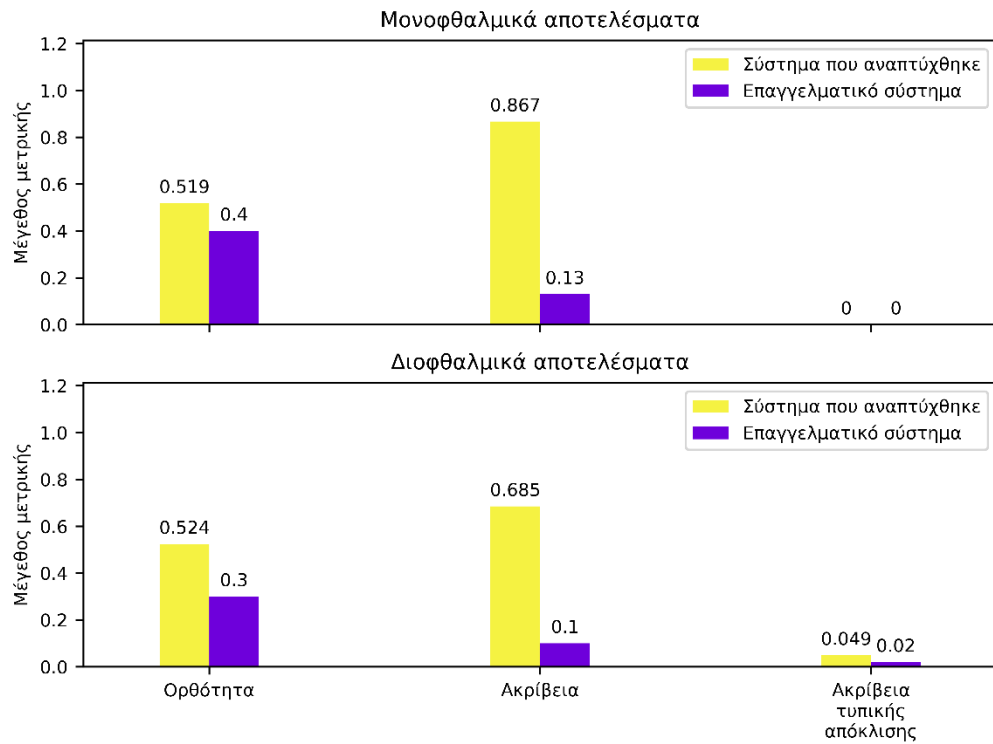
Σχήμα 7 Μονοφθαλμικά (επάνω) και διοφθαλμικά (κάτω) αποτελέσματα αξιολόγησης του συστήματός μας σε συνθήκες χαμηλού φωτισμού, συγκριτικά με επαγγελματικό σύστημα

Στις συνθήκες έντονου φωτισμού το σύστημά μας παρουσιάζει ελαφρώς καλύτερες τιμές ορθότητας, αλλά υστερεί στην ακρίβειά του.

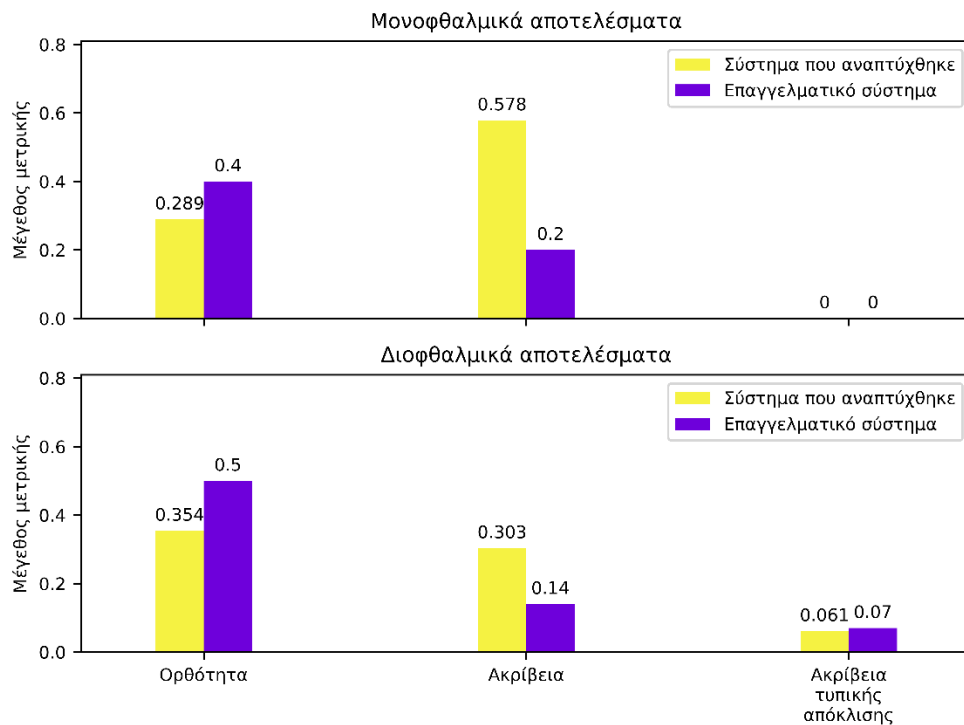


Σχήμα 8 Μονοφθαλμικά (επάνω) και διοφθαλμικά (κάτω) αποτελέσματα αξιολόγησης του συστήματός μας σε συνθήκες υψηλού φωτισμού, συγκριτικά με επαγγελματικό σύστημα

Στο στάδιο μετατόπισης του κεφαλιού του χρήστη κατά τον άξονα X παρατηρούμε παρόμοιες τιμές ορθότητας, αλλά η ακρίβεια είναι εμφανώς χειρότερη στο σύστημά μας, όπως αναμέναμε.

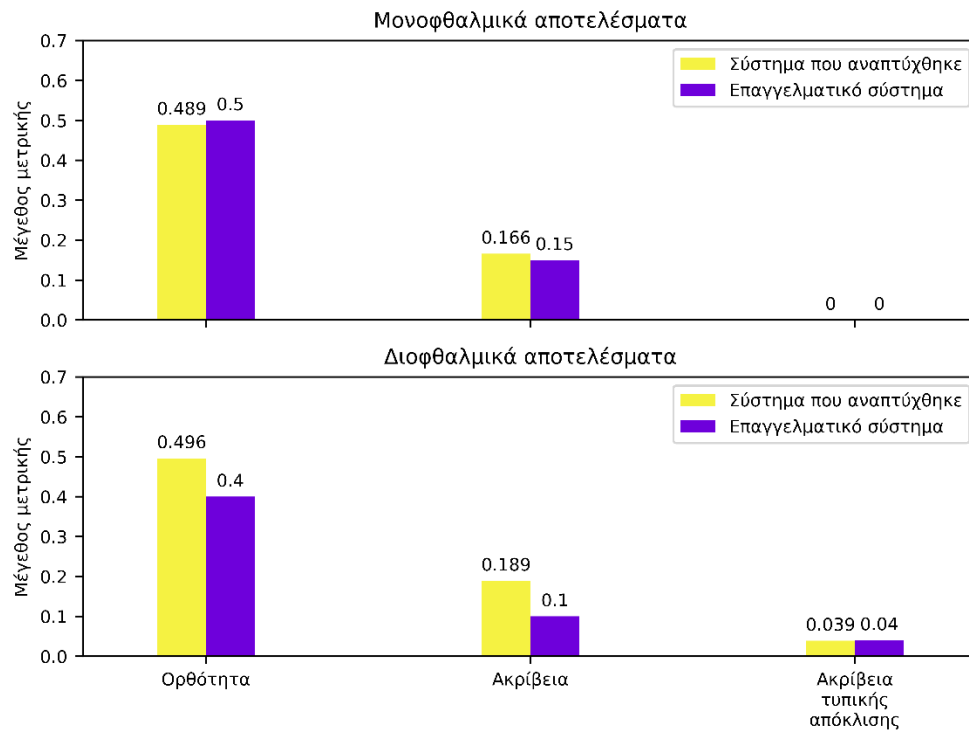


Σχήμα 9 Μονοφθαλμικά (επάνω) και διοφθαλμικά (κάτω) αποτελέσματα αξιολόγησης του συστήματός μας με το κεφάλι του χρήστη μετατοπισμένο προς τα αριστερά, συγκριτικά με επαγγελματικό σύστημα

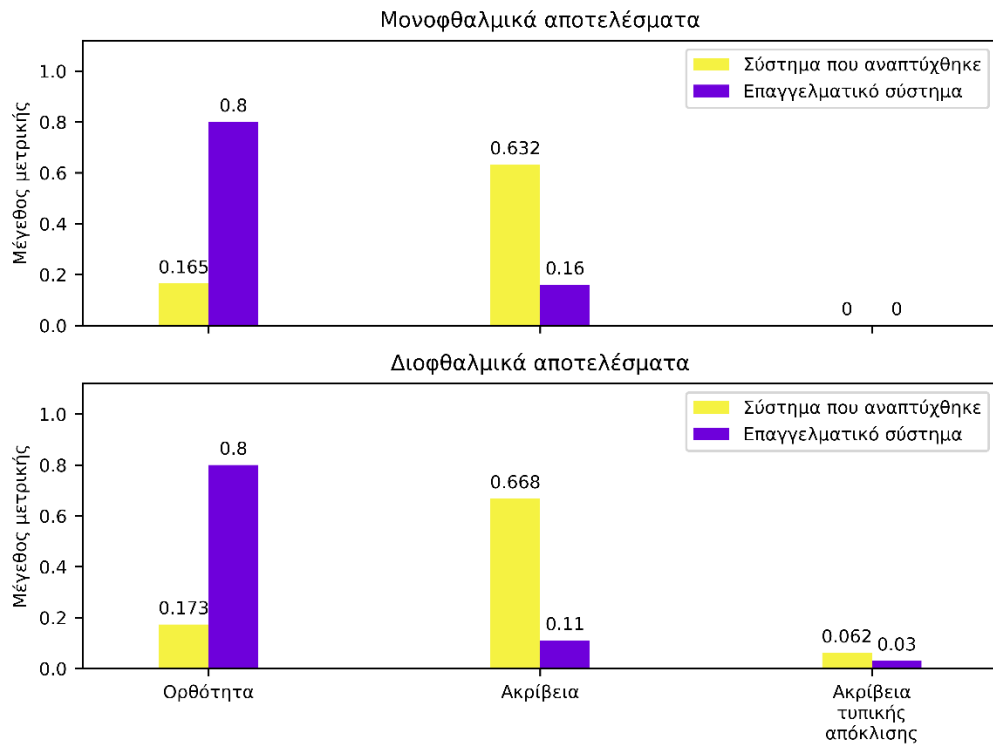


Σχήμα 10 Μονοφθαλμικά (επάνω) και διοφθαλμικά (κάτω) αποτελέσματα αξιολόγησης του συστήματός μας με το κεφάλι του χρήστη μετατοπισμένο προς τα δεξιά, συγκριτικά με επαγγελματικό σύστημα

Όσον αφορά τη μετατόπιση του προσώπου στον άξονα Υ παρατηρούμε μία ιδιαιτερότητα. Στις μετρήσεις όπου το κεφάλι του χρήστη ήταν πάνω από το κέντρο της οθόνης τα δύο συστήματα δίνουν παρόμοιες μετρικές, με αυτές του εμπορικού συστήματος να είναι ελαφρώς καλύτερες. Στις μετρήσεις όμως που το κεφάλι βρισκόταν κάτω από το κέντρο της οθόνης τα δύο συστήματα διαφέρουν έντονα. Το σύστημά μας φαίνεται να έχει αρκετά καλύτερη ορθότητα, αλλά υστερεί στην ακρίβεια.

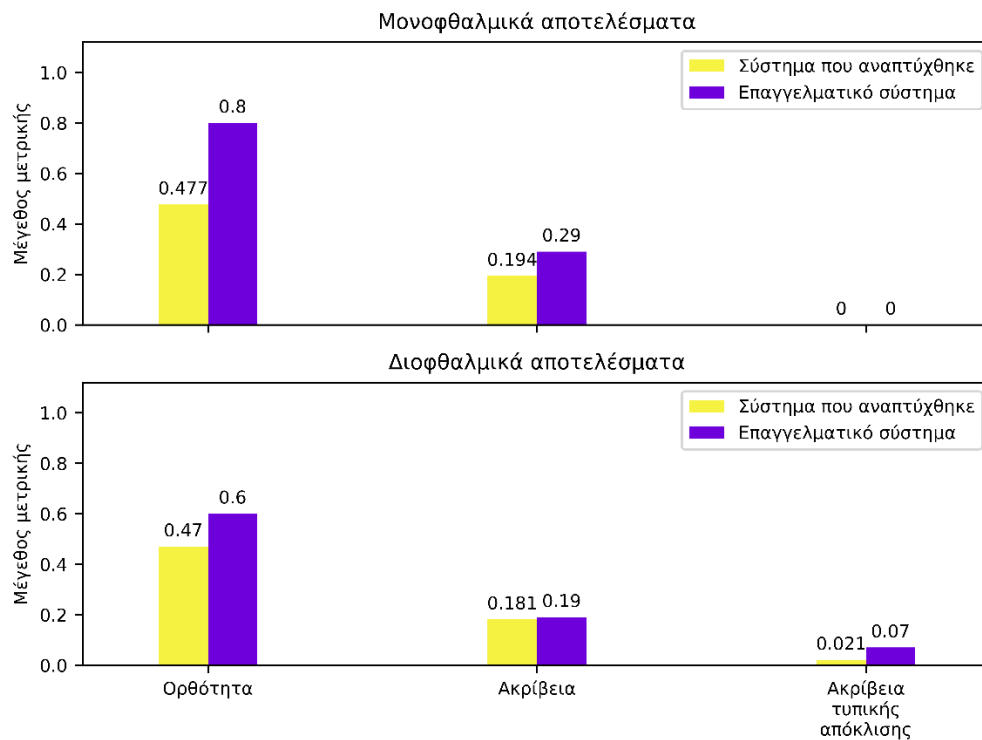


Σχήμα 11 Μονοφθαλμικά (επάνω) και διοφθαλμικά (κάτω) αποτελέσματα αξιολόγησης του συστήματός μας με το κεφάλι του χρήστη μετατοπισμένο προς τα επάνω, συγκριτικά με επαγγελματικό σύστημα

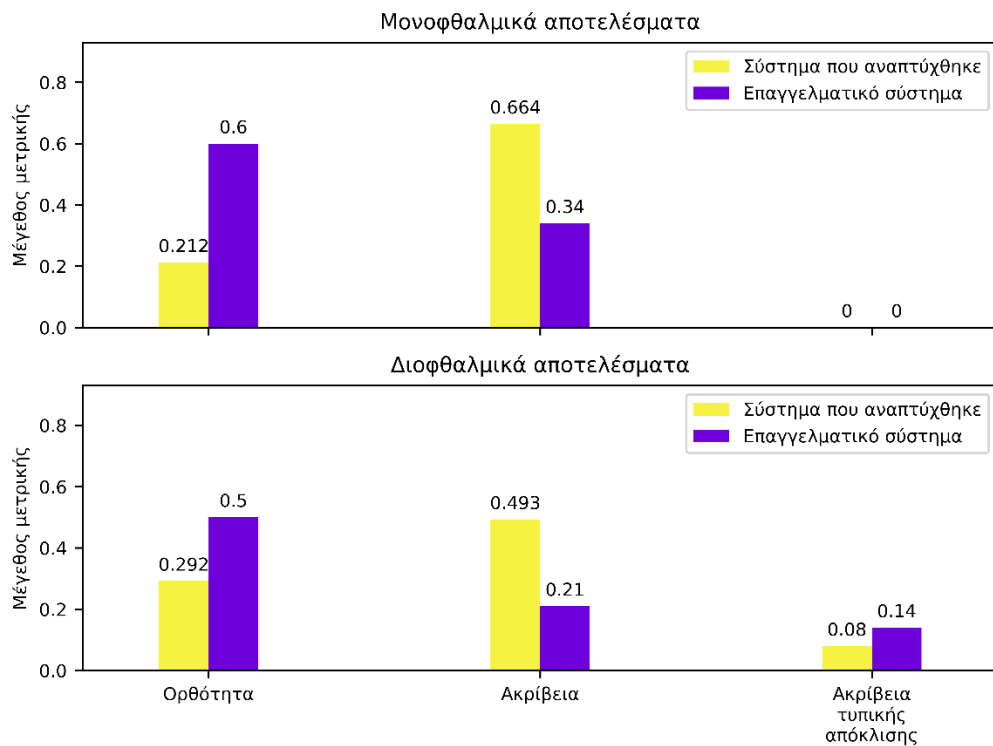


Σχήμα 12 Μονοφθαλμικά (επάνω) και διοφθαλμικά (κάτω) αποτελέσματα αξιολόγησης του συστήματός μας με το κεφάλι του χρήστη μετατοπισμένο προς τα κάτω, συγκριτικά με επαγγελματικό σύστημα

Τέλος, στον έλεγχο της απόδοσης των συστημάτων με βάση την απόσταση, παρατηρούμε ότι το σύστημά μας προσφέρει καλύτερες μετρικές εκτός από την ακρίβεια στην περίπτωση που ο χρήστης βρίσκεται μακριά από την κάμερα.



Σχήμα 13 Μονοφθαλμικά (επάνω) και διοφθαλμικά (κάτω) αποτελέσματα αξιολόγησης του συστήματός μας με το κεφάλι του χρήστη είναι τοποθετημένο πιο κοντά στην οθόνη, συγκριτικά με επαγγελματικό σύστημα



Σχήμα 14 Μονοφθαλμικά (επάνω) και διοφθαλμικά (κάτω) αποτελέσματα αξιολόγησης του συστήματός μας με το κεφάλι του χρήστη είναι τοποθετημένο πιο μακριά από την οθόνη, συγκριτικά με επαγγελματικό σύστημα

Κεφάλαιο 5ο

5. Αξιοποίηση του συστήματος σε διαδικτυακή σελίδα

5.1 Βασική λειτουργία της ιστοσελίδας

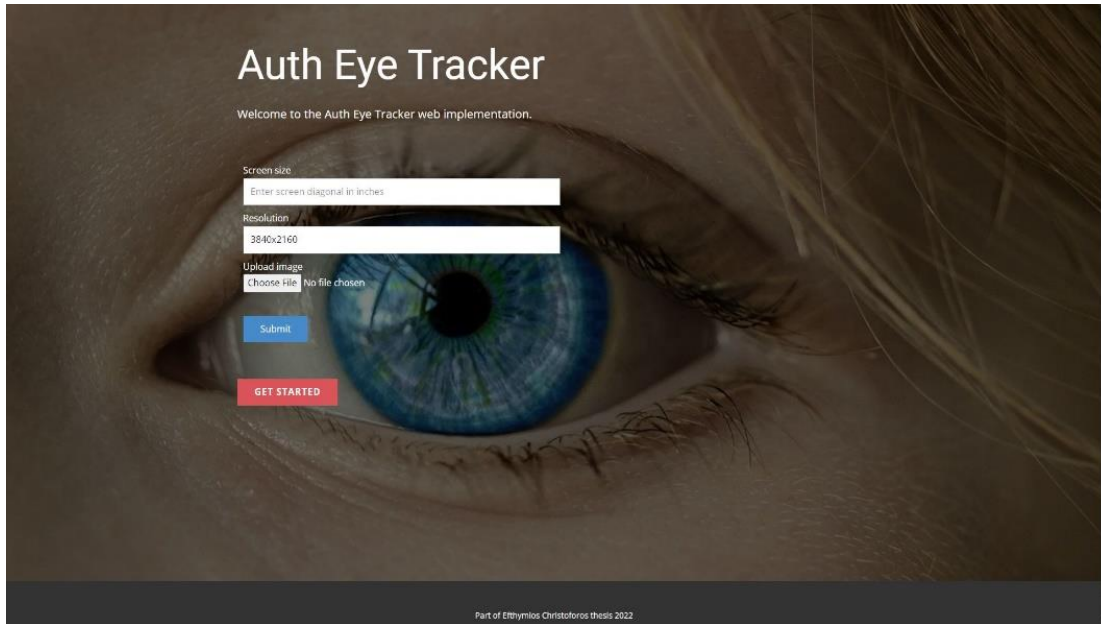
Για την επίδειξη της χρησιμότητας του συστήματος και των αποτελεσμάτων του, καθώς και για την ελεύθερη χρήση του από το κοινό δημιουργήθηκε μία ιστοσελίδα η οποία κάνει χρήση του υλοποιημένου συστήματος εντοπισμού βλέμματος. Η βασική λειτουργία της εφαρμογής περιλαμβάνει την καταγραφή του βλέμματος του χρήστη από την κάμερα που διαθέτει για ένα χρονικό διάστημα, καθώς προβάλλεται μία εικόνα σε πλήρη οθόνη. Μετά το χρονικό αυτό διάστημα, γίνεται επεξεργασία των δεδομένων της καταγραφής και προβάλλονται ορισμένες οπτικοποιήσεις τους επάνω στην εικόνα η οποία προβλήθηκε προηγουμένως, αποκαλύπτοντας σε ποια σημεία της εστίασε ο χρήστης το βλέμμα του για μεγαλύτερο χρονικό διάστημα και με ποια σειρά. Μία τέτοια ιστοσελίδα θα είχε χρησιμότητα σε αξιολόγηση περιβάλλοντος διεπαφής χρήστη και μακετών ιστοσελίδων και εφαρμογών υπολογιστή, για να γνωρίζουμε καλύτερα το πόσο ο χρήστης αξιοποιεί και αντιλαμβάνεται τα στοιχεία της διεπαφής που του παρουσιάζονται, και ποια από αυτά αγνοούνται. Φυσικά η παραπάνω χρήση είναι ένα μόνο παράδειγμα, καθώς στην πραγματικότητα η ιστοσελίδα επιτρέπει ανέμβαση της φωτογραφίας που παρατηρείται από τον χρήστη, καθώς και κατέβαση των θέσεων στις οποίες κοίταξε αναλυτικά σε αρχείο CSV. Ο μόνος περιορισμός της ιστοσελίδας είναι ότι δεν παρέχει ζωντανό feed πληροφοριών εστίασης στην οθόνη, αλλά παρέχει όλες τις πληροφορίες μετά το τέλος της μαγνητοσκόπησης, όπως θα δούμε και παρακάτω.

5.2 Αναλυτική περιγραφή

Η ιστοσελίδα στήθηκε, όμοια με το σύστημα εντοπισμού βλέμματος, σε γλώσσα Python στο περιβάλλον PyCharm με χρήση των βιβλιοθηκών Flask και Socketio. Οι βιβλιοθήκες αυτές βοήθησαν στο στήσιμο του διακομιστή ο οποίος θα φιλοξενήσει τον ιστότοπο, καθώς και στην εύκολη και γρήγορη επικοινωνία του με τον κάθε χρήστη. Η επικοινωνία αυτή περιλαμβάνει αρχικά τις απαραίτητες κλήσεις για την πλοήγηση του χρήστη στις διάφορες σελίδες του ιστότοπου, αλλά και την αποστολή και λήψη ζωντανής ροής βίντεο. Αρχικά γίνεται αποστολή ροής βίντεο από την κάμερα του χρήστη προς τον διακομιστή. Εκεί πραγματοποιείται επεξεργασία του κάθε καρέ από το σύστημα εντοπισμού βλέμματος, καταγραφή των αποτελεσμάτων και επιστρέφεται η επεξεργασμένη ροή εικόνων πίσω στο πρόγραμμα περιήγησης του χρήστη. Για να λειτουργήσει το σύστημα εντοπισμού βλέμματος όμως, είναι απαραίτητη η λήψη ορισμένων πληροφοριών του συστήματος του χρήστη, καθώς και η βαθμονόμηση του. Και οι δύο λειτουργίες αυτές πραγματοποιούνται ζωντανά στον ιστότοπο και δεν υπάρχει ανάγκη για κάποια διαδικασία προετοιμασίας πριν τη χρήση του. Μετά την καταγραφή, τα αποτελέσματα παρουσιάζονται με τη μορφή ορισμένων οπτικοποιήσεων επάνω στην αρχική εικόνα που ανέβασε και παρατήρησε ο χρήστης. Οι οπτικοποιήσεις αυτές περιλαμβάνουν θερμοχάρτη, χάρτη εστίασης, μονοπάτι εστίασης και συνδιασμό των δυο τελευταίων. Τα αποτελέσματα παρουσιάζονται και με τη μορφή αρχείου CSV στο οποίο περιλαμβάνονται αναλυτικά οι θέσεις της οθόνης στις οποίες κοίταξε ο χρήστης, μαζί με τις αντίστοιχες χρονικές στιγμές. Η αναλυτική ροή και λειτουργία του ιστότοπου παρουσιάζεται στα παρακάτω υποκεφάλαια.

5.2.1 Αρχική σελίδα

Στην αρχική σελίδα του ιστότοπου γίνεται καλωσόρισμα του χρήστη και ζητείται η εισαγωγή κάποιων στοιχείων του συστήματός του, όπως το μέγεθος και την ανάλυση της οθόνης.



Εικόνα 20 Αρχική σελίδα του ιστότοπου

Ζητείται επίσης ανέβασμα της εικόνας την οποία θέλει να παρατηρήσει ο χρήστης κατά το στάδιο της μαγνητοσκόπησης. Η εικόνα αυτή θα πρέπει ιδανικά να έχει αναλογία ανάλυσης ίδια με αυτή της οθόνης του έτσι ώστε να μπορεί να προβληθεί σε πλήρη οθόνη χωρίς να περικόπτεται κάποιο μέρος της. Οι παραπάνω απαραίτητες πληροφορίες στέλνονται στον διακομιστή με τη μορφή φόρμας πατώντας το μπλε κουμπί SUBMIT, ενώ πατώντας το κόκκινο κουμπί GET STARTED προχωράμε στην επόμενη ιστοσελίδα για την πραγματοποίηση της βαθμονόμησης.

5.2.2 Σελίδα βαθμονόμησης

Στην δεύτερη σελίδα πραγματοποιείται βαθμονόμηση των παραμέτρων του συστήματος εντοπισμού βλέμματος που αναπτύξαμε. Εδώ αφού ζητηθεί άδεια για πρόσβαση στην κάμερα του χρήστη, ξεκινά η αποστολή και λήψη της ροής βίντεο. Συγκεκριμένα στέλνονται καρέ της κάμερας με ανάλυση 640x480 και ρυθμό 30 καρέ το δευτερόλεπτο από το πρόγραμμα περιήγησης στον διακομιστή. Η ανάλυση αυτή και ο ρυθμός των καρέ επιλέχθηκε για την καλύτερη ισορροπία μεταξύ της ακρίβειας και της ταχύτητας/αποκριτικότητας του συστήματος. Το κάθε καρέ που φτάνει στον διακομιστή περνάει από την επεξεργασία του συστήματος εντοπισμού βλέμματος όπου του τοποθετούνται κάποια διαγνωστικά γραφικά. Αυτά περιλαμβάνουν τους άξονες προσανατολισμού του προσώπου, κάποια σημεία 'άγκυρες' του προσώπου και οι εντοπισμένες θέσεις της ίριδας του κάθε ματιού. Το σύστημα εντοπισμού βλέμματος θα μπορεί να υπολογίζει τις συντεταγμένες του σημείου της οθόνης στο οποίο εστιάζει ο χρήστης μετά το στάδιο της βαθμονόμησης.



Εικόνα 21 Σελίδα βαθμονόμησης του ιστότοπου

Αφού επιτευχθεί η επικοινωνία ροών βίντεο ζητείται από τον χρήστη να θέσει το πρόγραμμα περιήγησής του σε λειτουργία πλήρους οθόνης. Στη συνέχεια του ζητείται να διατηρήσει ουδέτερο ύφος χωρίς μορφασμούς, να σταθεί σε απόσταση περίπου 60 εκατοστών από την κάμερά του και να ακολουθήσει ορισμένες οδηγίες οι οποίες περιγράφονται παρακάτω στην ιστοσελίδα. Οι οδηγίες αυτές βρίσκονται στα κείμενα 6 μπλε κουμπιών τα οποία ο χρήστης πρέπει να πατήσει με τη σειρά. Οι οδηγίες του κάθε κουμπιού περιγράφουν διαφορετικές θέσεις των ματιών και προσανατολισμούς του προσώπου. Ενώ ο χρήστης βρίσκεται στις θέσεις αυτές πρέπει να πατήσει το αντίστοιχο μπλε κουμπί. Με το κάθε πάτημα του κουμπιού ο διακομιστής αποθηκεύει πληροφορίες οι οποίες προέρχονται από την επεξεργασία του καρέ που βρισκόταν στο σύστημα όταν πατήθηκε το κουμπί και χρησιμοποιούνται για την εξαγωγή αποτελεσμάτων του συστήματος εντοπισμού εστίασης του βλέμματος. Οι πράσινοι κύκλοι υπάρχουν για να μπορεί ο χρήστης με ευκολία να βρει το κέντρο της οθόνης και άλλα σημεία που του ζητείται να κοιτάξει. Αφού ολοκληρώσει το πάτημα και των 6 κουμπιών μπορεί να προχωρήσει στο στάδιο της καταγραφής πατώντας το κόκκινο κουμπί START RECORDING.

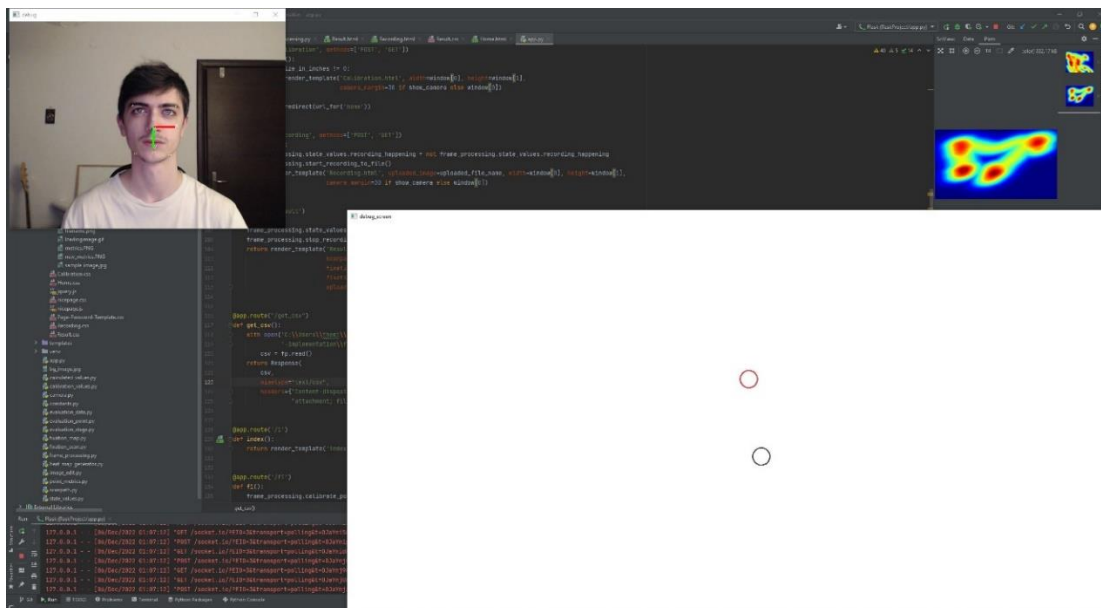
5.2.3 Σελίδα καταγραφής

Στη σελίδα καταγραφής του βλέμματος προβάλλεται σε πλήρη οθόνη η εικόνα την οποία ανέβασε ο χρήστης στην φόρμα της αρχικής σελίδας. Προβάλλεται επίσης η επεξεργασμένη ροή καρέ που λαμβάνει ο περιηγητής από τον διακομιστή.



Εικόνα 22 Σελίδα καταγραφής του ιστότοπου

Για τη διαπίστωση της ομαλής λειτουργίας του διακομιστή δημιουργούνται διαγνωστικά παράθυρα τα οποία προβάλλουν την επεξεργασμένη ροή της κάμερας του χρήστη και το σημείο στο οποίο κοιτάει στην οθόνη του, όπως φαίνεται στην παρακάτω εικόνα.



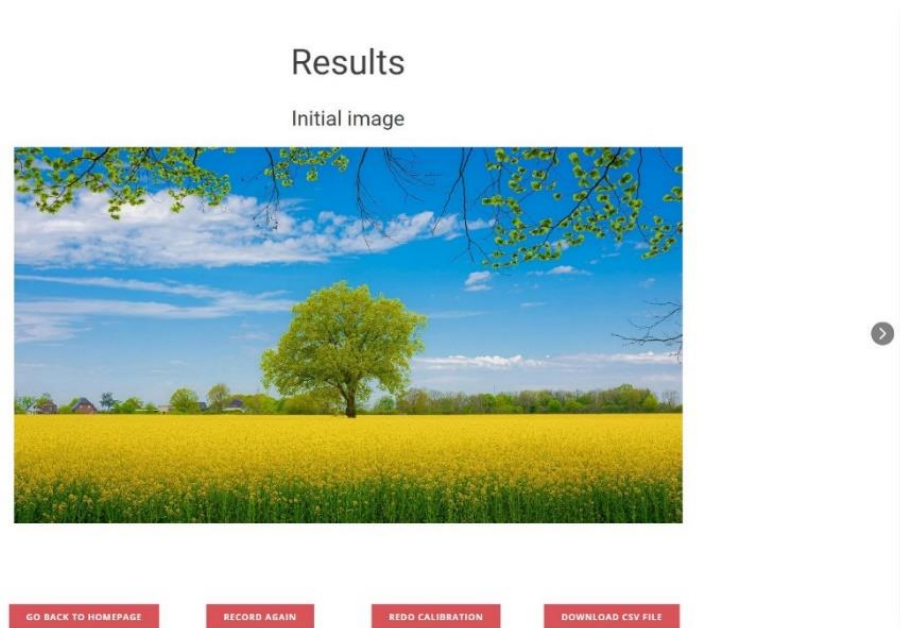
Εικόνα 23 Διαγνωστικά παράθυρα στον διακομιστή του ιστότοπου

Στο στάδιο αυτό ο διακομιστής διαθέτει όλα τα απαραίτητα δεδομένα για να μπορέσει από κάθε καρέ της κάμερας του χρήστη που λαμβάνει να εξάγει τις συντεταγμένες του σημείου

της οθόνης στο οποίο κοιτά ο χρήστης. Έτσι, από κάθε καρέ της κάμερας του χρήστη που λαμβάνει και επεξεργάζεται με το σύστημα εντοπισμού βλέμματος που αναπτύξαμε, υπολογίζει τις συντεταγμένες αυτές και τις αποθηκεύει (στον διακομιστή) σε ένα αρχείο CSV. Κάθε γραμμή του αρχείου αυτού περιλαμβάνει τις συντεταγμένες X και Y του σημείου στην οθόνη πριν και μετά την εξομάλυνση καθώς και την χρονική στιγμή στην οποία έγινε η καταγραφή. Η διαδικασία αυτή συνεχίζεται μέχρι να πατήσει ο χρήστης το κόκκινο κουμπί STOP RECORDING, με το πάτημα του οποίου πλοηγούμαστε στην τελευταία σελίδα του ιστοτόπου μας, αυτή της παρουσίασης των αποτελεσμάτων.

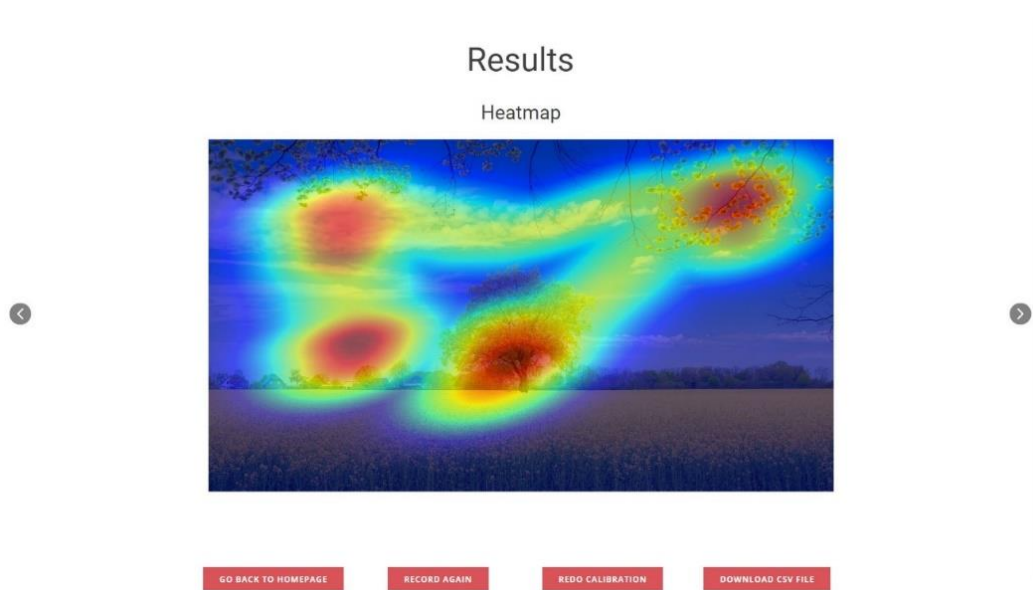
5.2.4 Σελίδα αποτελεσμάτων / οπτικοποιήσεων

Σε αυτήν την ιστοσελίδα παρουσιάζονται τα αποτελέσματα της καταγραφής οπτικοποιημένα με 4 διαφορετικές μεθόδους σε μία γκαλερί εικόνων. Στην πρώτη θέση της γκαλερί, έχουμε την αρχική εικόνα που μεταφόρτωσε ο χρήστης στην αρχική σελίδα, ως σημείο αναφοράς.



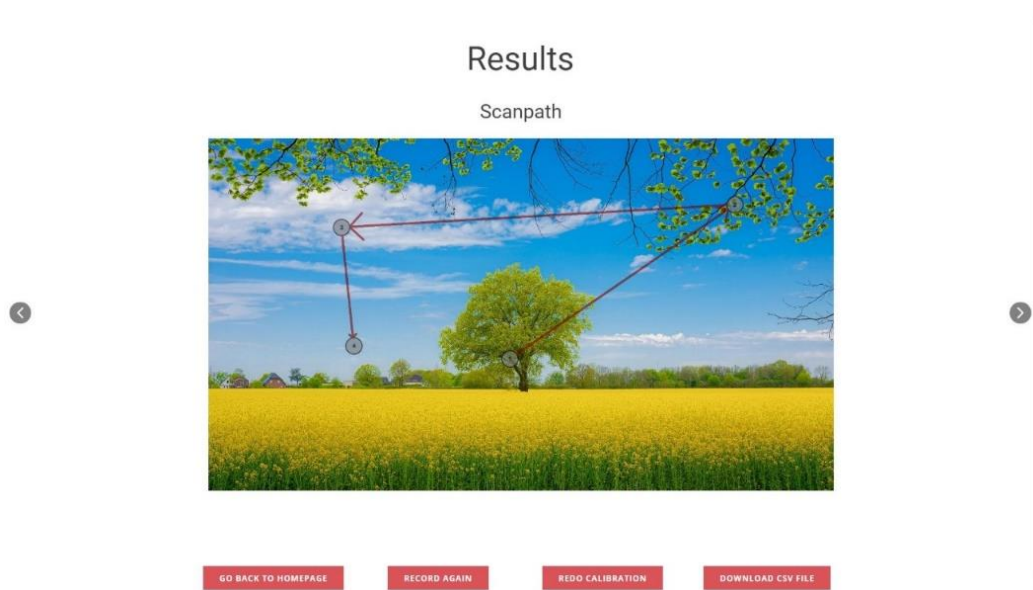
Εικόνα 24 Προβολή αρχικής εικόνας παρατήρησης για σημείο αναφοράς στη σελίδα αποτελεσμάτων του ιστοτόπου

Στη δεύτερη θέση έχουμε ένα θερμοχάρτη των σημείων του βλέμματος του χρήστη που καταγράφηκαν στο προηγούμενο στάδιο, τοποθετημένο πάνω από την αρχική εικόνα. Με την οπτικοποίηση αυτή βλέπουμε σε ποια σημεία εστίασε ο χρήστης και για πόση ώρα, αλλά δεν γνωρίζουμε το μονοπάτι που ακολούθησε το βλέμμα του.



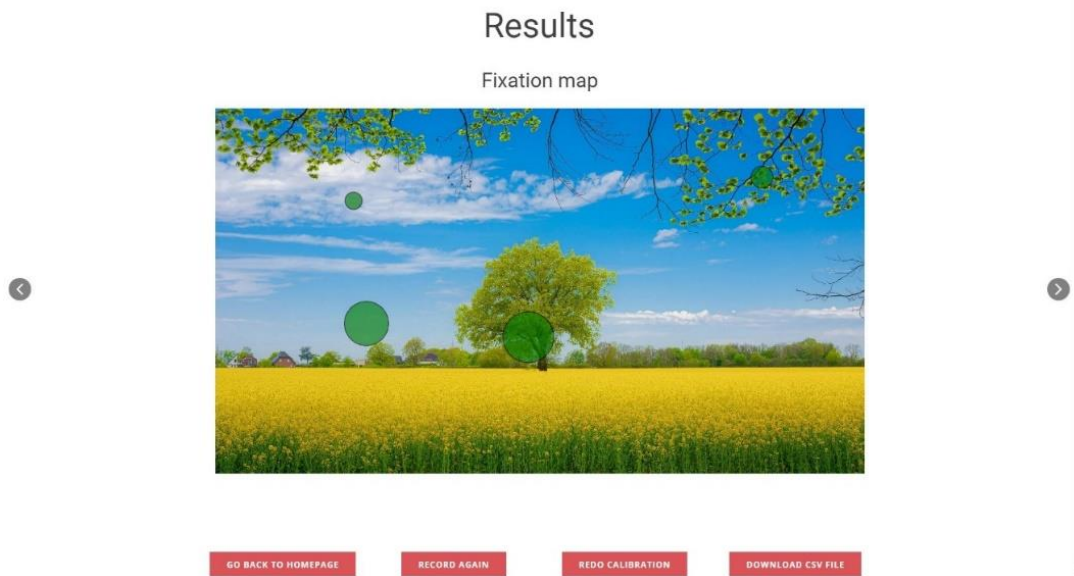
Εικόνα 25 Προβολή θερμοχάρτη στη σελίδα αποτελεσμάτων του ιστότοπου

Στην τρίτη θέση της γκαλερί έχουμε το μονοπάτι εστίασης στο οποίο βλέπουμε σε ποια σημεία εστίασε ο χρήστης και με ποια σειρά, αλλά δεν έχουμε πληροφορία για το πόση ώρα εστίασε στο κάθε σημείο.



Εικόνα 26 Προβολή μονοπατιού εστίασης στη σελίδα αποτελεσμάτων του ιστότοπου

Στην τέταρτη εικόνα έχουμε έναν χάρτη εστίασης όπου το μέγεθος του κάθε κύκλου είναι ανάλογο του χρόνου για τον οποίο εστίασε ο χρήστης στο αντίστοιχο σημείο, χωρίς να ξέρουμε την σειρά εστίασης των σημείων.



Εικόνα 27 Προβολή χάρτη εστίασης στη σελίδα αποτελεσμάτων του ιστότοπου

Τέλος στην πέμπτη εικόνα της γκαλερί έχουμε ένα συνδυασμό του μονοπατιού και του χάρτη εστίασης, όπου γίνεται φανερό σε ποια σημεία εστίασε ο χρήστης, για πόση ώρα και με ποια σειρά.



Εικόνα 28 Προβολή συνδυασμού μονοπατιού και χάρτη εστίασης στη σελίδα αποτελεσμάτων του ιστότοπου

Ο χρήστης μπορεί να αποθηκεύσει οποιαδήποτε από τις οπτικοποιήσεις σαν αρχείο εικόνας στη συσκευή του, ή να κατεβάσει και το αρχείο CSV με τις αναλυτικές πληροφορίες εστίασής του πατώντας το τελευταίο κόκκινο κουμπί DOWNLOAD CSV FILE. Τα υπόλοιπα κόκκινα κουμπιά της σελίδας δίνουν τη δυνατότητα στο χρήστη να επιστρέψει στην αρχική σελίδα (όπου μπορεί να αλλάξει τις ρυθμίσεις του και να ανεβάσει άλλη εικόνα παρατήρησης), να επαναλάβει την βαθμονόμηση ή να κάνει καινούρια καταγραφή με τις ίδιες ρυθμίσεις και βαθμονόμηση.

Κεφάλαιο 6ο

6. Συμπεράσματα

6.1 Αποτελέσματα - Συμπεράσματα

Έχουμε διαπιστώσει τη χρησιμότητα μιας συσκευής ιχνηλάτησης οφθαλμών, λόγω της εφαρμογής της σε πολυάριθμα πεδία όπως την αλληλεπίδραση ανθρώπου-υπολογιστή, την ψυχολογία, την ιατρική και την έρευνα αγοράς. Στην παρούσα εργασία επιλέξαμε να εστιάσουμε στην ευκολία χρήσης και στην ενσωμάτωση του συστήματος που θα αναπτύξουμε σε οποιαδήποτε περίπτωση χρήσης, χωρίς επιπλέον υλικό. Αναπτύξαμε, λοιπόν, αλγόριθμο για τον εντοπισμό του βλέμματος και την αποτύπωσή του στην οθόνη του χρήστη, με συμβατική συσκευή καταγραφής. Ο παραπάνω αλγόριθμος ενσωματώθηκε σε εφαρμογή που υλοποιήσαμε για την εμφάνιση διαγνωστικών στοιχείων, και την εκτέλεση διαδικασιών όπως η βαθμονόμηση, η αξιολόγηση του συστήματος, καθώς και η καταγραφή της εξόδου του συστήματος σε ορισμένο χρονικό διάστημα, η οποία συνοδεύεται από την οπτικοποίηση των δεδομένων που καταγράφηκαν. Για την χρήση και την αξιοποίηση του συστήματος που υλοποιήθηκε από το κοινό, δημιουργήθηκε ιστοσελίδα στην οποία ο χρήστης μπορεί να επιλέγει και να ανεβάζει την εικόνα που θέλει να παρατηρήσει και, ομοίως με την προηγούμενη υλοποίηση, να καταγράφει το βλέμμα του μέσω της διαδικτυακής του κάμερας και να λαμβάνει τόσο τα δεδομένα της καταγραφής σε μορφή αρχείου CSV, όσο και κάποιες οπτικοποιήσεις που του προσφέρουν άμεση κατανόηση των δεδομένων ιχνηλάτησης.

Λόγω της επιλογής που κάναμε να μη χρησιμοποιήσουμε επιπλέον υλικό για την οπτική καταγραφή και να αρκεστούμε στη χρήση απλής συσκευής καταγραφής, είχαμε εξ αρχής μειωμένες προσδοκίες για την ακρίβεια του ανεπτυγμένου συστήματος. Από το στάδιο αξιολόγησης στην εφαρμογή του αλγορίθμου αντλήσαμε σημαντικές πληροφορίες για την απόδοση του συστήματος σε σύγκριση με αξιόπιστο εμπορικό σύστημα το οποίο κάνει χρήση εξειδικευμένου υλικού και αισθητήρων, καθώς και λογισμικού το οποίο αναπτύχθηκε και εξελίσσεται από ομάδα ανάπτυξης λογισμικού. Παρατηρήσαμε ότι παρόλο που οι μετρικές της εμπορικής συσκευής ήταν καλύτερες σχεδόν σε όλες τις συνθήκες αξιολόγησης, γεγονός το οποίο ήταν αναμενόμενο, το σύστημά μας παρήγαγε αξιοπρεπή αποτελέσματα για περιπτώσεις χρήσης στις οποίες η απόλυτη ακρίβεια δεν είναι ύψιστης σημασίας.

6.2 Καινοτομία και συνεισφορά της εργασίας

Με την ολοκλήρωση της παρούσας εργασίας, έχουμε επιτύχει στη δημιουργία ενός χρήσιμου και προσβάσιμου εργαλείου εντοπισμού και ιχνηλάτησης βλέμματος. Με την απαίτηση του ελάχιστου δυνατού εξοπλισμού, μιας απλής κάμερας κι ενός υπολογιστή, το σύστημά μας μπορεί να αξιοποιηθεί στην πλειοψηφία των περιπτώσεων χρήσης. Κάποιες από αυτές είναι η χρήση σε εργαστήρια διεξαγωγής διαφόρων ειδών έρευνας (ιατρικής,

ακαδημαϊκής, μάρκετινγκ), η χρήση για την αλληλεπίδραση μεταξύ ανθρώπου και υπολογιστή, βοηθητική χρήση στην καθημερινότητα ανθρώπων με ειδικές ανάγκες και η χρήση για βελτίωση μέσων ψυχαγωγίας όπως βιντεοπαιχνιδιών. Το γεγονός ότι το υλοποιημένο σύστημα αποτελείται μόνο από λογισμικό, συνεπάγεται ότι το μόνο κόστος της υλοποίησης αυτής είναι ο χρόνος της ανάπτυξης και της συντήρησης του κώδικα. Επομένως, το σύστημα μπορεί να διανεμηθεί στο ευρύ κοινό με μηδενικό κόστος. Μία πραγματοποίηση της διανομής αυτής είναι η ιστοσελίδα που αναπτύξαμε. Μέσω αυτής μπορεί ο κάθε χρήστης που κατέχει έναν υπολογιστή και μια οποιαδήποτε διαδικτυακή κάμερα, να έχει στη διάθεσή του ένα σύστημα οφθαλμικής ιχνηλάτησης, του οποίου η απόδοση θα είναι, ως ένα βαθμό, ανάλογη της ποιότητας του συστήματός του. Σε σύγκριση με άλλα συστήματα τα οποία υλοποιούνται χωρίς τη χρήση εξειδικευμένου υλικού, έχουμε πετύχει μικρότερη εξάρτηση από την ποιότητα αυτή του συστήματος του χρήστη, με την εξάλειψη των διαφορών εισόδου. Η εξάλειψη αυτή βοήθησε, επίσης, στην ανεξαρτητοποίηση από τις συνθήκες φωτισμού. Επιπλέον, λόγω των γεωμετρικών μεθόδων που χρησιμοποιήσαμε, επιτρέπουμε μεγαλύτερη ελευθερία κινήσεων του κεφαλιού του χρήστη (και πάλι σε σύγκριση με άλλα συστήματα που χρησιμοποιούν μόνο λογισμικό) όσον αφορά τόσο την μετατόπιση, όσο και την περιστροφή του.

6.3 Μελλοντικές εξελίξεις

Υπάρχουν πολλές υλοποιήσεις οι οποίες ξεφεύγουν από το θέμα της παρούσας εργασίας, αλλά θα μπορούσαν να βελτιώσουν σε μεγάλο βαθμό το ήδη υλοποιημένο σύστημα, ή και να προσθέσουν χρήσιμες λειτουργικότητες σε αυτό. Όσον αφορά τη βελτίωση της ήδη υπάρχουσας υλοποίησης, ένα σημείο το οποίο θα μπορούσε να γίνει πιο αποδοτικό είναι το στάδιο της βαθμονόμησης. Πιο συγκεκριμένα, στην παρούσα κατάσταση, τα σημεία τα οποία χρησιμοποιούνται για την λήψη χρήσιμων παραμέτρων για τη λειτουργία του αλγορίθμου, περιορίζονται στο κέντρο, το δεξί και το κάτω άκρο της οθόνης. Θεωρούμε ότι με τη χρήση παραπάνω σημείων τα οποία εκτείνονται σε όλη την οθόνη, όπως τις γωνίες και τα υπόλοιπα άκρα της, θα είναι δυνατή η βελτίωση της ακρίβειας του συστήματος και της απόδοσής του σε όλη την έκταση της επιφάνειάς της. Ένα άλλο μέρος της υλοποίησής μας το οποίο θα μπορούσε να βελτιωθεί και να επεκταθεί σε μεγάλο βαθμό είναι ο ιστότοπος. Εδώ θα μπορούσε να γίνει καλύτερος ο τρόπος με τον οποίο διαχειριζόμαστε τα δεδομένα στον διακομιστή. Ιδανικά θα μπορούσε να υπάρχει βάση δεδομένων στην οποία να μπορούν να δημιουργούνται χρήστες. Έτσι, θα ήταν δυνατή η μόνιμη αποθήκευση όλων των καταγραφών που έχει πραγματοποιήσει ένας χρήστης και η εύρεσή τους σε οποιαδήποτε στιγμή χωρίς να χρειάζεται να αποθηκεύσει τίποτα στην συσκευή του. Η υλοποίηση αυτή προϋποθέτει προφανώς και την αντίστοιχη ανάπτυξη του client του ιστοτόπου, με τη δημιουργία ιστοσελίδων για την εγγραφή ή είσοδο του χρήστη και την προβολή του προφίλ και των καταγραφών του. Θα μπορούσαν τέλος να γίνουν βελτιώσεις στη διάταξη των υπάρχοντων ιστοσελίδων για την ευκολότερη περιήγηση διεξαγωγή ενεργειών του χρήστη.

Όπως αναφέραμε είναι σε μεγάλο βαθμό δυνατή η επεκτασιμότητα του συστήματός μας. Αρχικά θα μπορούσαν να γίνουν ενέργειες για την διεύρυνση των δυνατών περιπτώσεων χρήσης του συστήματος ιχνηλάτησης οφθαλμού. Συγκεκριμένα, με την ανακατασκευή του συστήματός μας ώστε να μπορεί να χρησιμοποιηθεί από μικροεπεξεργαστές ή και κινητά τηλέφωνα θα γίνεται δυνατή η χρήση του συστήματος σε διάφορων ειδών περιβάλλοντα. Ένας άλλος τρόπος με τον οποίο θα μπορούσε να επιτραπεί η ευρύτερη χρήση του αλγορίθμου που αναπτύξαμε, αλλά και η ανάπτυξη νέων εφαρμογών με αυτόν από άλλους προγραμματιστές, είναι η δημιουργία βιβλιοθήκης η οποία θα υποστηρίζει όσο το

δυνατόν περισσότερες γλώσσες προγραμματισμού. Έτσι, θα μπορεί να ενσωματωθεί σε μεγάλο αριθμό έργων και να βοηθήσει στην υλοποίηση εφαρμογών σχετικών με την ιχνηλάτηση οφθαλμών με πιο εξειδικευμένο στόχο.

Βιβλιογραφικές πηγές και Αναφορές

Βιβλιογραφικές αναφορές

- Camilleri, S. C. (2018). Unobtrusive and pervasive video-based eye-gaze tracking. *Image and Vision Computing*, 21-40.
- Cognolato, M. a. (2018). Head-mounted eye gaze tracking devices: An overview of modern devices and recent advances. *Journal of Rehabilitation and Assistive Technologies Engineering*, 205566831877399.
- Fakotakis, E. S. (2015). On visual gaze tracking based on a single low cost camera. *Signal Processing: Image Communication*, 29-42.
- J. K., J. J. (2017). *Facial recognition using histogram of gradients and support vector machines*.
- Janarthanan, A. B. (2013). Classifying Electrooculogram to Detect Directional Eye Movements. *Procedia Technology*, 67-75.
- Jiaqi Jiang, Z. Z. (2019). Appearance-Based Gaze Tracking: A Brief Review. Στο Z. Z. Jiaqi Jiang, *Intelligent Robotics and Applications* (σσ. 629-640). Cham: Springer International Publishing.
- Kim, H.-C. a. (2014). Eye Detection for Gaze Tracker with Near Infrared Illuminator. *2014 IEEE 17th International Conference on Computational Science and Engineering*, (σσ. 458-464).
- Kumari, N. a. (2021). Mobile Eye-Tracking Data Analysis Using Object Detection via YOLO v4. *Sensors*, 7668.
- Martínez, A. L. (2019). Video-oculography eye tracking towards clinical applications: A review. *Computers in Biology and Medicine*, 57-66.
- O. Déniz, G. B. (2011). *Face recognition using Histograms of Oriented Gradients*. Elsevier B.V.
- Phase, T. (2019, 11). Face Detection and Eye Detection in Image and Video using Pre-trained HAAR-Cascade Classifier. *International Journal for Research in Applied Science and Engineering Technology*, σσ. 652-659.
- Robinson, D. A. (1963). A Method of Measuring Eye Movement Using a Scieral Search Coil in a Magnetic Field. *IEEE Transactions on Bio-medical Electronics*, 137-145.
- Viola, P. a. (2001). *Rapid object detection using a boosted cascade of simple features*. Kauai, HI, USA: Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001.
- Yu, M. a. (2018). An eye detection method based on convolutional neural networks and support vector machines. *Intelligent Data Analysis*, 345-362.

Λίστα Εικόνων

Εικόνα 1 Διαφορετικές τοποθετήσεις ηλεκτροδίων κατά το ηλεκτροφθαλμογράφημα	10
Εικόνα 2 Πηνία αναζήτησης σκληρού χιτώνα σε μορφή φακών επαφής	11
Εικόνα 3 Εμφάνιση αντανάκλασης κερατοειδούς χιτώνα.....	11
Εικόνα 4 Συσκευή ιχνηλάτησης οφθαλμού τοποθετημένη στο κεφάλι του χρήστη. Φαίνονται οι όψεις της κάμερας για τους οφθαλμούς και της κάμερας για το περιβάλλον.....	12
Εικόνα 5 Αντιστοίχιση Haar Features σε στοιχεία προσώπου από τον αλγόριθμο Viola - Jones.....	13
Εικόνα 6 Προσανατολισμός των gradient σε σημεία του προσώπου	14
Εικόνα 7 Ευθεία που ορίζεται από το κέντρο του ματιού (πράσινο σημείο) και την επιφάνειά του στο σημείο της ίριδας (μπλε σημείο), η οποία συναντά την επιφάνεια της οθόνης στο σημείο που ψάχνουμε (κίτρινο σημείο).....	18
Εικόνα 8 Εφαρμογή μεθόδου ομογραφίας επιπέδου σε δισδιάστατη εικόνα	20
Εικόνα 9 Εφαρμογή μεθόδου ομογραφίας επιπέδου σε τρισδιάστατα σημεία του προσώπου	21
Εικόνα 10 Εύρεση ακτίνας της σφαίρας του ματιού με αλλαγή του προσανατολισμού του προσώπου	22
Εικόνα 11 Διαχωρισμός εικόνας σε RGB πίνακα	24
Εικόνα 12 Πλέγμα προσώπου που προσφέρει η βιβλιοθήκη Mediaripe	25
Εικόνα 13 Αρχιτεκτονική βιβλιοθήκης Mediaripe.....	26
Εικόνα 14 Διαγνωστικά στοιχεία παραθύρου κάμερας. Φαίνονται τα σημεία των κέντρων των ματιών, της κάθε ίριδας και τα σημεία – άγκυρες του προσώπου	27
Εικόνα 15 Διαγνωστικά στοιχεία παραθύρου οθόνης. Φαίνονται το σημείο βλέμματος του δεξιού και αριστερού ματιού (κόκκινο και πράσινο σημείο αντίστοιχα), το τελικό συνολικό σημείο του βλέμματος (κίτρινο σημείο) και το σημείο προσανατολισμού και θέσης του προσώπου (μπλε και λευκό σημείο αντίστοιχα)	28
Εικόνα 16 Διάταξη παραθύρου οθόνης στο στάδιο της βαθμονόμησης. Φαίνονται κάτω δεξιά τα τέσσερα μπλε κουμπιά, και τα βοηθητικά σημεία στο κέντρο και τις άκρες της οθόνης.	30
Εικόνα 17 Ο θερμοχάρτης που παράγεται από την καταγραφή τοποθετημένος πάνω από την εικόνα που παρατηρήθηκε	32
Εικόνα 18 Ιδανικά σημεία παρατήρησης για το στάδιο της αξιολόγησης	35
Εικόνα 19 Ακραία σημεία παρατήρησης για το στάδιο της αξιολόγησης	36
Εικόνα 20 Αρχική σελίδα του ιστότοπου	52
Εικόνα 21 Σελίδα βαθμονόμησης του ιστότοπου.....	53
Εικόνα 22 Σελίδα καταγραφής του ιστότοπου	54
Εικόνα 23 Διαγνωστικά παράθυρα στον διακομιστή του ιστότοπου.....	54
Εικόνα 24 Προβολή αρχικής εικόνας παρατήρησης για σημείο αναφοράς στη σελίδα αποτελεσμάτων του ιστότοπου	55
Εικόνα 25 Προβολή θερμοχάρτη στη σελίδα αποτελεσμάτων του ιστότοπου	56
Εικόνα 26 Προβολή μονοπατιού εστίασης στη σελίδα αποτελεσμάτων του ιστότοπου	56
Εικόνα 27 Προβολή χάρτη εστίασης στη σελίδα αποτελεσμάτων του ιστότοπου.....	57
Εικόνα 28 Προβολή συνδυασμού μονοπατιού και χάρτη εστίασης στη σελίδα αποτελεσμάτων του ιστότοπου	57

Λίστα Πινάκων

Πίνακας 1 Μονοφθαλμικά και διοφθαλμικά αποτελέσματα αξιολόγησης σε ιδανικές συνθήκες	38
Πίνακας 2 Μονοφθαλμικά και διοφθαλμικά αποτελέσματα αξιολόγησης σε ακραία σημεία της οθόνης. 38	
Πίνακας 3 Μονοφθαλμικά και διοφθαλμικά αποτελέσματα αξιολόγησης σε συνθήκες χαμηλού φωτισμού	38
Πίνακας 4 Μονοφθαλμικά και διοφθαλμικά αποτελέσματα αξιολόγησης σε συνθήκες έντονου φωτισμού	39
Πίνακας 5 Μονοφθαλμικά και διοφθαλμικά αποτελέσματα αξιολόγησης με το κεφάλι του χρήστη μετατοπισμένο προς τα αριστερά.....	39
Πίνακας 6 Μονοφθαλμικά και διοφθαλμικά αποτελέσματα αξιολόγησης με το κεφάλι του χρήστη μετατοπισμένο προς τα δεξιά	39
Πίνακας 7 Μονοφθαλμικά και διοφθαλμικά αποτελέσματα αξιολόγησης με το κεφάλι του χρήστη μετατοπισμένο προς τα επάνω	40
Πίνακας 8 Μονοφθαλμικά και διοφθαλμικά αποτελέσματα αξιολόγησης με το κεφάλι του χρήστη μετατοπισμένο προς τα κάτω.....	40
Πίνακας 9 Μονοφθαλμικά και διοφθαλμικά αποτελέσματα αξιολόγησης με το κεφάλι του χρήστη μετατοπισμένο πιο κοντά στην οθόνη	40
Πίνακας 10 Μονοφθαλμικά και διοφθαλμικά αποτελέσματα αξιολόγησης με το κεφάλι του χρήστη μετατοπισμένο πιο μακριά από την οθόνη	40

Λίστα Σχημάτων

Σχήμα 1 Δεδομένα εξόδου του συστήματος πριν την ομαλοποίηση παρελθοντικών τιμών (επάνω σχήμα) και μετά την ομαλοποίηση (κάτω σχήμα)	29
Σχήμα 2 Φαίνονται (πάνω) τα δεδομένα εξόδου του συστήματος και (κάτω) ο θερμοχάρτης που παράγεται από το Gaussian KDE	32
Σχήμα 3 Διαφορά ακρίβειας (precision) και ορθότητας (accuracy)	33
Σχήμα 4 Συνολικό γράφημα αποτελεσμάτων αξιολόγησης	37
Σχήμα 5 Μονοφθαλμικά (επάνω) και ιοφθαλμικά (κάτω) αποτελέσματα αξιολόγησης του συστήματός μας σε ιδανικές συνθήκες, συγκριτικά με επαγγελματικό σύστημα	41
Σχήμα 6 Μονοφθαλμικά (επάνω) και ιοφθαλμικά (κάτω) αποτελέσματα αξιολόγησης του συστήματός μας σε ακραία σημεία της οθόνης, συγκριτικά με επαγγελματικό σύστημα	42
Σχήμα 7 Μονοφθαλμικά (επάνω) και ιοφθαλμικά (κάτω) αποτελέσματα αξιολόγησης του συστήματός μας σε συνθήκες χαμηλού φωτισμού, συγκριτικά με επαγγελματικό σύστημα	43
Σχήμα 8 Μονοφθαλμικά (επάνω) και ιοφθαλμικά (κάτω) αποτελέσματα αξιολόγησης του συστήματός μας σε συνθήκες υψηλού φωτισμού, συγκριτικά με επαγγελματικό σύστημα	44
Σχήμα 9 Μονοφθαλμικά (επάνω) και ιοφθαλμικά (κάτω) αποτελέσματα αξιολόγησης του συστήματός μας με το κεφάλι του χρήστη μετατοπισμένο προς τα αριστερά, συγκριτικά με επαγγελματικό σύστημα	45
Σχήμα 10 Μονοφθαλμικά (επάνω) και ιοφθαλμικά (κάτω) αποτελέσματα αξιολόγησης του συστήματός μας με το κεφάλι του χρήστη μετατοπισμένο προς τα δεξιά, συγκριτικά με επαγγελματικό σύστημα	46
Σχήμα 11 Μονοφθαλμικά (επάνω) και ιοφθαλμικά (κάτω) αποτελέσματα αξιολόγησης του συστήματός μας με το κεφάλι του χρήστη μετατοπισμένο προς τα επάνω, συγκριτικά με επαγγελματικό σύστημα	47
Σχήμα 12 Μονοφθαλμικά (επάνω) και ιοφθαλμικά (κάτω) αποτελέσματα αξιολόγησης του συστήματός μας με το κεφάλι του χρήστη μετατοπισμένο προς τα κάτω, συγκριτικά με επαγγελματικό σύστημα	48
Σχήμα 13 Μονοφθαλμικά (επάνω) και ιοφθαλμικά (κάτω) αποτελέσματα αξιολόγησης του συστήματός μας με το κεφάλι του χρήστη είναι τοποθετημένο πιο κοντά στην οθόνη, συγκριτικά με επαγγελματικό σύστημα	49
Σχήμα 14 Μονοφθαλμικά (επάνω) και ιοφθαλμικά (κάτω) αποτελέσματα αξιολόγησης του συστήματός μας με το κεφάλι του χρήστη είναι τοποθετημένο πιο μακριά από την οθόνη, συγκριτικά με επαγγελματικό σύστημα	50