

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ - ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ
ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ
ΥΠΟΛΟΓΙΣΤΩΝ

Τομέας: Τηλεπικοινωνιών και Τεχνολογίας Πληροφορίας (Τ.&Τ.Π.)
Ομάδα Απεικόνισης Πληροφορίας & Εικονικής Πραγματικότητας



ΠΑΝΕΠΙΣΤΗΜΙΟ
ΠΑΤΡΩΝ
UNIVERSITY OF PATRAS

Διπλωματική Εργασία

του φοιτητή του Τμήματος Ηλεκτρολόγων Μηχανικών και Τεχνολογίας
Υπολογιστών της Πολυτεχνικής Σχολής του Πανεπιστημίου Πατρών

Ιωάννη Κουτουλογένη του Ανδρέα

αριθμός μητρώου: 1019659

Θέμα

Σύστημα Υποβοήθησης Πλοήγησης για Άτομα με Προβλήματα Όρασης

Επιβλέπων

Κωνσταντίνος Μουστάκας
Αναπληρωτής Καθηγητής, Πανεπιστήμιο Πατρών

Αριθμός Διπλωματικής Εργασίας:

Πάτρα, Φεβρουάριος 2020

ΠΙΣΤΟΠΟΙΗΣΗ

Πιστοποιείται ότι η διπλωματική εργασία με θέμα

Σύστημα Υποβοήθησης Πλοήγησης για Άτομα με Προβλήματα Όρασης

του φοιτητή του Τμήματος Ηλεκτρολόγων Μηχανικών και Τεχνολογίας
Υπολογιστών

Ιωάννη Κουτουλογένη του Ανδρέα

Αριθμός Μητρώου: 1019659

παρουσιάστηκε δημόσια και εξετάστηκε στο τμήμα Ηλεκτρολόγων Μηχανικών και
Τεχνολογίας Υπολογιστών στις

20/02/2020

Ο Επιβλέπων

Κωνσταντίνος Μουστάκας
Αναπληρωτής Καθηγητής

Ο Διευθυντής του Τομέα

Ιωάννης Μουρτζόπουλος
Καθηγητής

Περίληψη

Η παρούσα διπλωματική εργασία έχει ως αντικείμενο τον σχεδιασμό ενός ολοκληρωμένου και διακριτικού συστήματος πλοήγησης για άτομα με προβλήματα όρασης, το οποίο θα λειτουργεί συνδυαστικά με παραδοσιακά βιοθήματα, όπως είναι το λευκό μπαστούνι, και θα παρέχει στον χρήστη την δυνατότητα ασφαλέστερης και γρηγορότερης μετακίνησης μέσα σε ένα αστικό περιβάλλον, εστιάζοντας περισσότερο στο πρόβλημα διάσχισης ενός δρόμου μέσω μιας διάβασης πεζών. Στόχος είναι η υλοποίηση μιας πειραματικής διάταξης η οποία θα είναι πλήρως λειτουργική, θα αυξάνει το βαθμό αυτονομίας του χρήστη και θα βασίζεται πάνω σε 3 βασικούς άξονες: την φορητότητα, την ελάχιστη επεμβατικότητα και την αξιοπιστία. Για την εξασφάλιση των παραπάνω χαρακτηριστικών χρησιμοποιήθηκαν τεχνικές επεξεργασίας εικόνας και αναπτύχθηκαν 3 διαφορετικοί αλγόριθμοι, που αφορούν την ανίχνευση διάβασης πεζών, την αναγνώριση της κατάστασης των φωτεινών σηματοδοτών και τον εντοπισμό εμποδίων, καθώς επίσης και μια εφαρμογή για Android smartphones, η οποία αναλαμβάνει την καθοδήγηση και αποτελεί την διεπιφάνεια αλληλεπίδρασης μεταξύ του συστήματος και του χρήστη. Η παροχή ανάδρασης στον χρήστη πραγματοποιείται μέσω του smartphone και τη χρήση 4 διαφορετικών μοτίβων δονήσεων, με σκοπό την αξιοποίηση του απτικού καναλιού και την αποσυμφόρηση του ακουστικού καναλιού του χρήστη. Η τελική πειραματική διάταξη περιλαμβάνει την ενσωμάτωση μιας κάμερας με αισθητήρα βάθους, ενός μικροϋπολογιστή, μιας φορητής μπαταρίας τύπου power bank και ένα smartphone.

Abstract

The current diploma thesis addresses the problem of designing a complete and discreet navigational system for visually impaired people, supporting the traditional assisting methods like white cane and providing the user with the ability to move safer and faster throughout an urban environment, especially when it comes to crossing a road through a crosswalk. The ultimate goal is to implement an experimental setup that will be totally functional, will increase user's autonomy and will follow 3 main principles; portability, low intrusiveness and reliability. For satisfying the above criteria we made use of diverse image processing techniques and developed 3 different algorithms, for crosswalk detection, for pedestrian lights recognition and for obstacle detection, while also developing an Android application, which is responsible for the guidance and the interaction between the system and the user. The feedback to the user is provided through his smartphone, by using 4 different vibration patterns, because we wanted to utilize the haptic channel and limit the cognitive load connected with the acoustic channel. The final experimental setup includes the incorporation of a camera with a depth sensor, a microcomputer, a portable battery (power bank) and a smartphone.

Ευχαριστίες

Ολοκληρώνοντας την διπλωματική μου εργασία, δεν μπορώ παρά να ευχαριστήσω τα άτομα που διαδραμάτισαν καθοριστικό ρόλο στην πορεία των φοιτητικών μου χρόνων και την επιτυχή περάτωση των σπουδών μου.

Καταρχάς, θα ήθελα να ευχαριστήσω θερμά τον επιβλέποντα καθηγητή της διπλωματικής μου εργασίας κ. Κωνσταντίνο Μουστάκα για τις συνεχείς συμβουλές και την πολύτιμη καθοδήγησή του καθ' όλη την διάρκεια εκπόνησης της εργασίας, καθώς επίσης και για την θετική ενέργεια την οποία μου μετέδιδε σε κάθε μας συνάντηση.

Παράλληλα, θα ήθελα να ευχαριστήσω από καρδιάς τα μέλη και το ερευνητικό προσωπικό της ομάδας Απεικόνισης Πληροφορίας & Εικονικής Πραγματικότητας για τις ώρες και τα σαββατοκύριακα που περάσαμε μαζί τους τελευταίους μήνες, ενώ ιδιαίτερη αναφορά θα ήθελα να κάνω στους Γιώργο Μιχαλάκη, Σωτήρη Αλεξίου και Άγγελο Χατζηκαλύμνιο για την συνεισφορά τους στην υλοποίηση του βίντεο επίδειξης της διπλωματικής μου εργασίας.

Τέλος, νιώθω την ανάγκη να ευχαριστήσω την οικογένειά μου για την αγάπη και τη στήριξη που μου παρείχε καθ' όλη την διάρκεια των σπουδών μου, όπως επίσης και τους φίλους μου που ήταν πάντα δίπλα μου και με ενέπνευσαν ο καθένας με τον τρόπο του.

*Στην γιαγιά μου,
για την αμέριστη συμπαράστασή της
όλα αυτά τα χρόνια*

Περιεχόμενα

Κατάλογος σχημάτων	xiii
Κατάλογος πινάκων	xvii
1 Εισαγωγή	1
1.1 Στατιστικά στοιχεία για άτομα με προβλήματα όρασης	1
1.2 Δυσκολίες πλούγησης σε αστικά περιβάλλοντα	2
1.3 Κίνητρα και σκοπός διπλωματικής εργασίας	3
1.4 Διάρθρωση διπλωματικής εργασίας	3
2 Προγενέστερη έρευνα στο πεδίο	5
2.1 Βοηθητικά μέσα υποστήριξης ατόμων με μειωμένη όραση	5
2.1.1 Συμβατικά μέσα καθοδήγησης	5
2.1.1.1 Λευκό μπαστούνι	5
2.1.1.2 Σκύλος οδηγός	6
2.1.2 Ηλεκτρονικές συσκευές υποβοήθησης ταξιδιού (ETAs)	7
2.1.2.1 Αντίληψη περιβάλλοντος	7
2.1.2.2 Ανάδραση με χρήση ηχητικών ειδοποιήσεων	7
2.1.2.3 Ανάδραση με χρήση απτικών ειδοποιήσεων	7
2.2 Κατηγορίες συστημάτων υποβοήθησης πλούγησης	8
2.2.1 RFID	8
2.2.2 Learning-based	11
2.2.2.1 Νευρωνικά δίκτυα	12
2.2.3 Vision-based	14
2.2.3.1 Ultrasonic sensors	15
2.2.3.2 LiDAR sensors	16
2.2.3.3 RGB-D Camera	17
3 Περιγραφή και ανάλυση του προτεινόμενου συστήματος	19
3.1 Προδιαγραφές ενός συστήματος πλούγησης για χρήστες με μειωμένη όραση	19
3.1.1 Δυνατότητα υποκατάστασης της όρασης με άλλες αισθήσεις (ακοή, αφή) . .	20
3.1.2 Αξιοπιστία συστήματος	20
3.1.3 Ανθρωποκεντρική σχεδίαση διεπαφής χρήστη	20
3.1.4 Ελαχιστοποίηση επεμβατικότητας του συστήματος	21
3.2 Εστίαση σε διαβάσεις πεζών	21
3.3 Περιγραφή του συστήματος	22
3.3.1 Γενική εικόνα	22
3.3.2 Αρχιτεκτονική	22
3.3.3 Πλούγηση σε αστικό περιβάλλον – δυσκολίες και περιορισμοί	24
3.3.4 Παράλληλη χρήση με το λευκό μπαστούνι	24
3.4 Απτική Ανάδραση	24

3.4.1	Αισθητηριακή υποκατάσταση	24
3.4.2	Υπάρχοντα κανάλια μετάδοσης χωρικής πληροφορίας	25
3.4.3	Γιατί απτική ανάδραση;	25
3.4.4	Haptic icons	25
3.4.5	Δυνατότητες και περιορισμοί	26
3.5	Εξοπλισμός - Υλικό (Hardware)	28
3.5.1	Μονάδα «αίσθησης» του περιβάλλοντος – Camera/Depth Sensor	28
3.5.1.1	Διαθέσιμοι αισθητήρες	28
3.5.1.2	Κριτήρια επιλογής κατάλληλου αισθητήρα	29
3.5.1.3	Intel RealSense D435i	30
3.5.2	Μονάδα επεξεργασίας εικόνας – Μικρούπολογιστής	31
3.5.2.1	Raspberry Pi 4	32
3.5.3	Συσκευή πλοϊγησης – Android Smartphone	33
3.5.3.1	Εφαρμογή πλοϊγησης	33
3.5.3.2	Προδιαγραφές εφαρμογής – Ελάχιστες απαιτήσεις	35
3.5.3.3	UX-UI Design	36
3.5.4	Συσκευή απτικής ανατροφοδότησης – Haptic Device	37
3.5.4.1	Υπάρχουσες συσκευές απτικής ανατροφοδότησης	37
3.5.4.2	Smartphone	37
3.6	Αλγόριθμοι πλοϊγησης & Λογισμικό (Software)	38
3.6.1	Χρησιμοποιηθέντα εργαλεία και γλώσσες προγραμματισμού	38
3.6.1.1	Android Studio - Java	38
3.6.1.2	Visual Studio - C++	39
3.6.1.3	OpenCV	40
3.6.1.4	Intel® RealSense™ SDK 2.0	41
3.6.1.5	Raspberry Pi - Linux OS	41
3.6.2	Αλγόριθμοι επεξεργασίας εικόνας σε πραγματικό χρόνο	42
3.6.2.1	Αναγνώριση διάβασης πεζών	44
3.6.2.1.1	Χαρακτηριστικά και περιορισμοί διάβασης πεζών	44
3.6.2.1.2	Η προτεινόμενη μέθοδος	45
3.6.2.1.3	Canny Edge Detector & Hough Transform	46
3.6.2.1.4	Block-Based Recognition	49
3.6.2.1.5	Synthesize	51
3.6.2.2	Αναγνώριση φωτεινού σηματοδότη	51
3.6.2.2.1	Spot Light Detection	52
3.6.2.2.2	Template Matching	54
3.6.2.2.3	Περισσότερα σχετικά με τους μορφολογικούς μετασχηματισμούς	56
3.6.2.3	Αλγόριθμος εντοπισμού εμποδίων	57
3.7	Επικοινωνία μεταξύ των συσκευών	59
3.7.1	Απαιτήσεις συστήματος	59
3.7.2	Socket Communication	60
3.7.2.1	Sockets	60
3.7.2.2	Τύποι sockets: TCP/UDP	60
3.7.3	Server/Client	60
3.7.3.1	Server	61
3.7.3.2	Client	61

4 Πειραματικός Έλεγχος Συστήματος	63
4.1 Απόδοση αλγορίθμων	63
4.1.1 Εντοπισμός διάβασης πεζών	63
4.1.2 Αναγνώριση φωτεινού σηματοδότη	64
4.1.3 Αποφυγή εμποδίων	66
4.2 Απτική ανάδραση	68
4.3 Τελική διάταξη	68
5 Συμπεράσματα - Προτάσεις για το μέλλον	69
5.1 Πλεονεκτήματα	69
5.2 Μειονεκτήματα	69
5.2.1 Ανίχνευση διάβασης πεζών	69
5.2.2 Αναγνώριση φωτεινού σηματοδότη	70
5.3 Μελλοντικές προτάσεις	71
5.3.1 Μετατροπέας Voice-to-Text	71
5.3.2 Αναγνώριση εμποδίων	71
5.3.3 Βελτιωμένο template φωτεινού σηματοδότη	71
5.3.4 Αυτονομία	71
5.3.5 Usability test	71
6 Παράρτημα	73
6.1 Φωτογραφίες εξοπλισμού	74
Βιβλιογραφία	77

Κατάλογος σχημάτων

1.1 Έκθεση του Παγκόσμιου Οργανισμού Υγείας για τα ποσοστά ανθρώπων με προβλήματα όρασης [1]	2
2.1 Παράδειγμα χρήσης λευκού μπαστουνιού	6
2.2 Σκύλος οδηγός	6
2.3 Παράδειγμα ολοκληρωμένου κυκλώματος ενός παθητικού RFID tag	9
2.4 Παράδειγμα ενός RFID reader	9
2.5 Περιγραφή του συστήματος που προτείνεται στο [2]	10
2.6 Παράδειγμα εκπαίδευσης ενός αλγορίθμου μάθησης	11
2.7 Απεικόνιση λειτουργίας του απλούστερου τεχνητού νευρωνικού δίκτυου, που ονομάζεται <i>Perceptron</i> και αποτελείται από έναν μόνο νευρώνα	13
2.8 Δομή επιπέδων σε ένα απλό τεχνητό νευρωνικό δίκτυο με ένα ενδιάμεσο κρυμμένο επίπεδο	13
2.9 Διάγραμμα εύρους ηχητικών συχνοτήτων	15
2.10 Αρχή λειτουργίας αισθητήρα υπερήχων	15
2.11 Αριστερά: Απεικόνιση του ηχο-εντοπισμού της νυχτερίδας. Δεξιά: Παράδειγμα ενός τυπικού αισθητήρα υπερήχων με έναν δέκτη και έναν πομπό.	16
2.12 3D ψηφιακή αναπαράσταση μιας τοποθεσίας μέσω LiDAR	16
2.13 Αισθητήρες LiDAR τοποθετούνται πάνω σε αυτόνομα οχήματα	17
2.14 Ο αισθητήρας LiDAR που χρησιμοποιείται στο [3]	17
2.15 Ο αισθητήρας Kinect	18
2.16 Αριστερά: Το μοτίβο από σημεία υπέρυθρου φωτός που εκπέμπεται. Δεξιά: Η παραγόμενη εικόνα βάθους ανάλογα με την παραμόρφωση της δομής του μοτίβου	18
3.1 Αρχιτεκτονική προτεινόμενου συστήματος	23
3.2 Αντιστοίχηση μοτίβων δονήσεων σε εντολές πλοήγησης	26
3.3 Μεταβολή του TPDT σε συνάρτηση με τα σημεία του σώματος [4]	27
3.4 ZED Stereo Camera	28
3.5 CamBoard pico flexx depth sensor	29
3.6 Αρχή λειτουργίας στερεοσκοπικής όρασης	30
3.7 Intel Realsense Depth Camera D435i	31
3.8 Intel Realsense D435i - Εσωτερική δομή	31
3.9 Raspberry Pi 4 Model B	32
3.10 Το powerbank <i>Intenso PD10000</i> που χρησιμοποιήθηκε	33
3.11 Η κατεύθυνση του χρήστη πρέπει να είναι μέσα σε ένα τόξο 60°για να θεωρείται έγκυρη. Άλλιώς παρέχεται κατάλληλη ανάδραση για αριστερή ή δεξιά στροφή.	34
3.12 Διάγραμμα ροής της Android εφαρμογής	34
3.13 Διεπιφάνεια επαφής χρήστη	36
3.14 Myo Armband	37
3.15 Προγραμματιστικό περιβάλλον Android Studio IDE	39
3.16 Προγραμματιστικό περιβάλλον Visual Studio IDE	40

3.17 Λογότυπο βιλβιοθήκης OpenCV	40
3.18 RealSense Viewer από το RealSense™ SDK 2.0	41
3.19 Ubuntu 19.10 with MATE Desktop Environment	42
3.20 Διάγραμμα ροής ενός frame	42
3.21 Παράδειγμα εφαρμογής του Spatial Filter βασιζόμενο στο [5]	43
3.22 Επιλογές φίλτρου εξάλεψης κενών pixels (hole filling)	44
3.23 Παράδειγμα εφαρμογής φίλτρων στο στάδιο του preprocessing (εξαιρείται το temporal filter, γιατί στην συγκεκριμένη λήψη δεν άλλαζε σημαντικά το αποτέλεσμα)	44
3.24 Κλασσική μορφή διάβασης πεζών τύπου ζέβρας [6]	45
3.25 Επιλογή περιοχής ενδιαφέροντος (ROI) για εντοπισμό διάβασης πεζών	46
3.26 Εφαρμογή Gaussian Filter	47
3.27 Εφαρμογή τελεστή Sobel για εύρεση κατευθυντικών παραγώγων	47
3.28 Εφαρμογή Non Maximum Suppresion	48
3.29 Εφαρμογή Double Thresholding	48
3.30 Εφαρμογή κατωφλίωσης με υστέρηση	49
3.31 Εφαρμογή Hough Transform	49
3.32 Παράδειγμα δημιουργίας πίνακα συσσώρευσης κατά τον μετασχηματισμό Hough [7]. Κάθε θέση στον πίνακα εκφράζει μια ευθεία και κάθε ψήφος σε μια θέση εκφράζει τον αριθμό των σημείων από τα οποία διέρχεται αυτή η ευθεία.	51
3.33 Overview του αλγόριθμου εντοπισμού διάβασης πεζών που παρουσιάζεται στην ερευνητική εργασία [6]	51
3.34 Overview του αλγόριθμου αναγνώρισης φωτεινών σηματοδοτών	52
3.35 Επιλογή περιοχής ενδιαφέροντος (ROI) για αναγνώριση φωτεινού σηματοδότη	52
3.36 Παράδειγμα ιδιότητας convexity	53
3.37 Παράδειγμα ιδιότητας inertia ratio	54
3.38 Overview της διαδικασίας template matching	55
3.39 Το αποτέλεσμα της διαδικασίας ολίσθησης βάσει της μετρικής Cross Correlation Normed (TM_CCORR_NORMED). Οι φωτεινές περιοχές υποδεικνύουν μεγαλύτερη ομοιότητα. Η θέση με τον κόκκινο κύκλο είναι αυτή με την υψηλότερη τιμή, γιατί η περιοχή με το περίγραμμα θεωρείται ως αυτή με το καλύτερο ταίριασμα.	55
3.40 Τα templates που χρησιμοποιήθηκαν στην παρούσα εργασία	55
3.41 Παράδειγμα μορφολογικών διεργασιών	57
3.42 Διαίρεση της εικόνας βάθους σε 3 τμήματα	58
3.43 Overview αλγορίθμου εντοπισμού εμποδίων	59
3.44 Μοντέλο Server/Client για μια σύνδεση TCP μέσω sockets	61
 4.1 Παράδειγμα εντοπισμού διάβασης πεζών. Η γκρίζα ευθεία γραμμή πάνω στη διάβαση, καθώς και η λευκή περιοχή κάτω από την διάβαση υποδηλώνουν το βέλτιστο μονοπάτι που πρέπει να ακολουθήσει ο χρήστης. Επίσης, το πρόγραμμα εμφανίζει και την κλίση των γραμμών της διάβασης σε μοίρες. Απλή διάβαση (αριστερά), βρεγμένη διάβαση (δεξιά).	64
4.2 Παράδειγμα εντοπισμού διάβασης πεζών με παρουσία εμποδίων (αριστερά) ή αντανάκλασης ηλιακού φωτός (δεξιά)	64
4.3 Παράδειγμα εντοπισμού διάβασης πεζών άλλων χρωματικών αποχρώσεων (αριστερά) ή σε σκοτεινό περιβάλλον (δεξιά)	64
4.4 Παράδειγμα αναγνώρισης κόκκινου φωτεινού σηματοδότη (κόκκινο περίγραμμα γύρω από το φανάρι). Οι μπλε κύκλοι υποδηλώνουν πιθανές περιοχές με φωτεινό σηματοδότη που εντοπίστηκαν από τον αλγόριθμο, αλλά δεν πέρασαν τη φάση της επικύρωσης (validation). Τα νούμερα που φαίνονται αριστερά στην κονσόλα είναι η τιμή του μέτρου σύγκρισης κατά την εφαρμογή του template matching.	65

4.5	Παραδείγματα αναγνώρισης κόκκινου φωτεινού σηματοδότη (κόκκινο περίγραμμα γύρω από το φανάρι).	65
4.6	Παράδειγμα αναγνώρισης πράσινου φωτεινού σηματοδότη (πράσινο περίγραμμα γύρω από το φανάρι).	65
4.7	Παραδείγματα αναγνώρισης πράσινου φωτεινού σηματοδότη (πράσινο περίγραμμα γύρω από το φανάρι). Οι μπλε κύκλοι υποδηλώνουν πιθανές περιοχές με φωτεινό σηματοδότη που εντοπίστηκαν από τον αλγόριθμο, αλλά δεν πέρασαν τη φάση της επικύρωσης (validation). Τα νούμερα που φαίνονται αριστερά στην κονσόλα είναι η τιμή του μέτρου σύγκρισης κατά την εφαρμογή του template matching.	66
4.8	Παράδειγμα εντοπισμού εμποδίων. Η κεντρική περιοχή της εικόνας βάθους είναι κενή, άρα το σύστημα ειδοποιεί τον χρήστη να προχωρήσει ελεύθερα.	66
4.9	Παράδειγμα εντοπισμού εμποδίων. Η κεντρική περιοχή της εικόνας βάθους είναι κατειλημμένη από εμπόδιο, άρα το σύστημα ειδοποιεί τον χρήστη να μετακινηθεί αριστερά στον κενό χώρο.	67
4.10	Παράδειγμα εντοπισμού εμποδίων. Τόσο η κεντρική, όσο και η αριστερή περιοχή της εικόνας βάθους είναι κατειλημμένες από εμπόδια, άρα το σύστημα ειδοποιεί τον χρήστη να μετακινηθεί δεξιά στον κενό χώρο.	67
4.11	Τελική διάταξη προτεινόμενου συστήματος	68
5.1	Παράδειγμα παραπλάνησης αλγορίθμου. Το πρόγραμμα αναγνωρίζει λανθασμένα κάποια σκαλοπάτια ως διάβαση πεζών. Αυτό συμβαίνει επειδή τα μοτίβα των σκαλοπατιών και της διάβασης είναι παρόμοια μεταξύ τους.	70
5.2	Παράδειγμα αστοχίας αλγορίθμου. Το πρόγραμμα αναγνωρίζει λανθασμένα κάποια σημεία της εικόνας ως φανάρια και τα ερμηνεύει με κόκκινο χρώμα. Αυτό συμβαίνει επειδή υπάρχουν έντονες αντανακλάσεις του ηλιακού φωτός, ενώ αντίστοιχα η ένταση της φωτεινότητας του φαναριού είναι χαμηλή.	70
6.1	Η κάμερα Intel Realsense 435i (πάνω) και το Powerbank Intenso PD10000 (κάτω) .	74
6.2	To Raspberry Pi 4 χωρίς θήκη (πάνω) και με προστατευτική θήκη (μέση) και τσαντάκι μέσης με όλο τον εξοπλισμό μέσα (κάτω)	75

Κατάλογος πινάκων

2.1	Λευκό μπαστούνι, θετικά και αρνητικά	5
2.2	Σκύλος οδηγός, θετικά και αρνητικά	6
3.1	10 βασικοί άξονες της ανθρωποκεντρικής σχεδίασης [8]	21
3.2	Χαρακτηριστικά ZED Stereo Camera [9]	28
3.3	Χαρακτηριστικά CamBoard pico flexx depth sensor [10]	29
3.4	Χαρακτηριστικά Intel Realsense D435i [11]	30
3.5	Χαρακτηριστικά Raspberry Pi 4 Model B [12]	32

1. Εισαγωγή

Στο κεφάλαιο αυτό γίνεται μια εισαγωγή στην κατάσταση που επικρατεί γύρω από τα άτομα με προβλήματα όρασης. Πιο συγκεκριμένα, παρατίθενται ορισμένα στατιστικά στοιχεία σχετικά με τον αριθμό των ατόμων με μειωμένη όραση, δίνεται ένας ορισμός της έννοιας της τύφλωσης και παρουσιάζονται μερικές από τις δυσκολίες που αναγκάζεται να αντιμετωπίσει ένα τέτοιο άτομο κατά την πλοιήγησή του σε μια πόλη. Τέλος, αναφέρονται τα κίνητρα τα οποία με οδήγησαν στην συγγραφή αυτής της διπλωματικής εργασίας, ενώ παράλληλα δίνεται μια σύντομη περιγραφή των κεφαλαίων που την συνθέτουν.

1.1 Στατιστικά στοιχεία για άτομα με προβλήματα όρασης

Για να προχωρήσουμε στον σχεδιασμό ενός συστήματος πλοιήγησης για τα άτομα με προβλήματα όρασης, πρέπει πρώτα να λάβουμε υπόψιν την κατηγοριοποίηση των ατόμων αυτών με βάση το ποσοστό τύφλωσής τους. Ο Παγκόσμιος Οργανισμός Υγείας (<https://www.who.int>), ο οποίος είναι ο πλέον αξιόπιστος φορέας σε θέματα που αφορούν την παγκόσμια υγεία, ορίζει την τύφλωση ως την μειωμένη ικανότητα όρασης σε βαθμό τέτοιο, ώστε τα προβλήματα δεν μπορούν να διορθωθούν με οποιοδήποτε μέσο, π.χ. γυαλιά, και παράλληλα προτείνει την ταξινόμηση των επιπέδων όρασης σε 4 βασικές κατηγορίες [13]:

- Φυσιολογική όραση – Normal vision
- Μέτρια Διαταραχή όρασης – Moderate visual impairment
- Σοβαρή Διαταραχή όρασης – Severe visual impairment
- Ολική Απώλεια όρασης – Blindness

Σύμφωνα με μια έρευνα που δημοσιεύτηκε το 2017 [14] το ποσοστό των ατόμων που παρουσιάζουν κάποιο πρόβλημα όρασης ή τύφλωση ανέρχεται στο 29,5% επί του συνόλου του παγκόσμιου πληθυσμού, δηλαδή το 1/3 του πληθυσμού παρουσιάζει κάποια δυσλειτουργία στην όραση. Στη ίδια έρευνα αναφέρεται ότι 36 εκατομμύρια άνθρωποι πάσχουν από ολική τύφλωση, ενώ 217 εκατομμύρια παρουσιάζουν κάποιας μορφής σοβαρή ή ήπια πάθηση, ενώ αξίζει να σημειώσουμε ότι στις περισσότερες περιπτώσεις η απώλεια όρασης θα μπορούσε να είχε αποφευχθεί αν υπήρχε η κατάλληλη πρόληψη και διενέργεια οφθαλμολογικών εξετάσεων. Δεδομένου ότι το ποσοστό των ανθρώπων με κάποιας μορφής τύφλωση είναι υψηλότερο στις χώρες με μέτριο ή χαμηλό εισόδημα, καταλαβαίνουμε ότι η πρόληψη είναι άμεσα συνδεδεμένη με το βιοτικό επίπεδο των πολιτών και την οικονομική τους δυνατότητα για πρόσβαση σε ιατρική περίθαλψη.



Σχήμα 1.1: Έκθεση του Παγκόσμιου Οργανισμού Υγείας για τα ποσοστά ανθρώπων με προβλήματα όρασης [1]

1.2 Δυσκολίες πλοιόγησης σε αστικά περιβάλλοντα

Τα περισσότερα άτομα με προβλήματα όρασης επιλέγουν να ζήσουν σε πόλεις, επειδή η οργάνωση και οι δυνατότητες που τους παρέχονται είναι πολύ καλύτερες σε σύγκριση με τη ζωή στην ύπαιθρο. Παρ' όλα αυτά, ένα από τα μεγαλύτερα εμπόδια που αντιμετωπίζουν είναι αυτό της μετακίνησης. Η δυνατότητα αυτών των ατόμων να μετακινούνται αυτόνομα σε ένα αστικό περιβάλλον είναι αυτή που τους επιτρέπει να έχουν μια αξιοπρεπή ζωή, χωρίς αυτήν το βιοτικό τους επίπεδο μειώνεται δραματικά. Σύμφωνα με έρευνες που έχουν πραγματοποιηθεί σε αυτό το πεδίο [15, 16], μερικές από τις προκλήσεις που αντιμετωπίζουν τα άτομα με μειωμένη όραση είναι:

- Συνεχής ανάγκη να ζητούν βοήθεια από περαστικούς
- Δυσκολία αναγνώρισης διαδρομών και χρήσης GPS
- Ρίσκο ατυχήματος
- Δυσκολία χρήσης των ανάγλυφων πλακιδίων για τυφλούς στα πεζοδρόμια
- Μη σωστή χρήση του λευκού μπαστουνιού οδηγεί σε ατυχήματα
- Φόβος χρήσης σκύλου-οδηγού
- Ελλιπής συντήρηση των πεζοδρομίων
- Απρόσεκτη συμπεριφορά πεζών

Ένα από τα σημαντικότερα προβλήματα μοιάζει να είναι η διάσχιση ενός δρόμου [17]. Φανταστείτε τον εαυτό σας μπροστά από ένα μεγάλο σταυροδρόμι ή στην άκρη μιας μεγάλης λεωφόρου την οποία πρέπει να διασχίσετε. Ένα άτομο με μειωμένη όραση πρέπει να λάβει υπόψιν του πολλούς παράγοντες πριν ξεκινήσει να διασχίζει τον δρόμο, όπως για παράδειγμα την ύπαρξη ή όχι διάβασης πεζών, την ταχύτητα με την οποία πλησιάζουν τα διερχόμενα οχήματα, τυχόν εμπόδια που υπάρχουν κατά μήκος της διαδρομής του, καθώς επίσης και τον χρόνο τον οποίο χρειάζεται για να διασχίσει τον δρόμο.

Κατανοούμε λοιπόν την ανάγκη ύπαρξης καλύτερων υποδομών, σωστής και επαρκούς σήμανσης στους δρόμους, καθώς και την ανάγκη να γίνουμε πιο προσεκτικοί απέναντι στους συμπολίτες μας με μειωμένη όραση.

1.3 Κίνητρα και σκοπός διπλωματικής εργασίας

Ο βασικότερος λόγος που με ώθησε να ασχοληθώ με το συγκεκριμένο θέμα είναι η έλλειψη κατάλληλων προϋποθέσεων, ώστε τα άτομα με προβλήματα όρασης να έχουν ίσες ευκαιρίες στην καθημερινότητά τους και να μην περιορίζονται από το ποσοστό αναπηρίας που μπορεί να έχουν. Η καθημερινότητά μας καθορίζεται σε μεγάλο ποσοστό από το πόσο εύκολα μπορούμε να πραγματοποιήσουμε εργασίες αυτόνομα και ανεξάρτητα, χωρίς να βασιζόμαστε αποκλειστικά σε άλλους. Για τους πολίτες δίχως προβλήματα όρασης το να βγουν από το σπίτι τους και να μετακινηθούν μέχρι το πάρκο για μια βόλτα ή να περπατήσουν μέχρι το τοπικό σούπερ μάρκετ για τα ψώνια τους είναι κάτι το αυτονόχτο και δεν αποτελεί επιπλέον κόπο. Αντιθέτως, για ένα άτομο με μειωμένη ή ολική απώλεια όρασης, ακόμα και το να βγει από το σπίτι του περπατώντας στην άκρη του πεζοδρομίου αποτελεί μια ενέργεια που ενέχει κινδύνους και απαιτεί έντονη συγκέντρωση και πνευματική προσπάθεια. Το δικαίωμα στην ελεύθερη, αυτόνομη και ασφαλή μετακίνηση είναι κάτι που θα έπρεπε να απολαμβάνουν όλοι οι άνθρωποι.

Η τεχνολογία σήμερα μας επιτρέπει να αλλάζουμε τον τρόπο με τον οποίο αντιμετωπίζουμε την πραγματικότητα και να δίνουμε λύσεις σε προβλήματα που είναι δύσκολα και περίπλοκα. Εχοντας υπόψιν τα παραπάνω, στόχος της παρούσας διπλωματικής εργασίας είναι ο σχεδιασμός και η υλοποίηση ενός συστήματος υποβοήθησης πλοήγησης που θα επιτρέπει στους χρήστες με προβλήματα όρασης να κινούνται με περισσότερη άνεση μέσα σε ένα αστικό περιβάλλον, συμβάλλοντας έτσι στην αύξηση της αυτονομίας τους. Θα πρέπει να τονιστεί ότι το προτεινόμενο σύστημα στόχο έχει να ενισχύσει την λειτουργία του λευκού μπαστουνιού, που χρησιμοποιείται κατά κόρον από τους τυφλούς, και όχι να την αντικαταστήσει. Παράλληλα, η παρούσα εργασία εστιάζει κυρίως στην παροχή βοήθειας στους χρήστες για την αποτελεσματικότερη και ασφαλέστερη διάσχιση μιας διάβασης πεζών, ενώ αξίζει να τονιστεί ότι η παροχή ανατροφοδότησης στον χρήστη κατά την πλοήγηση δίνεται μέσω της αφής, με σκοπό την αξιοποίηση του απτικού καναλιού και την μείωση του φόρτου στο κανάλι της ακοής.

1.4 Διάρθρωση διπλωματικής εργασίας

Εχοντας επίγνωση των αναγκών και της πολυπλοκότητας ενός συστήματος πλοήγησης για άτομα με προβλήματα όρασης, έχει γίνει προσπάθεια να παρουσιαστούν με τον πιο κατανοητό και περιεκτικό τρόπο όλες οι πλευρές ενός τέτοιου συστήματος. Για τον λόγο αυτό η παρούσα εργασία διαμορφώνεται ως εξής:

- Στο κεφάλαιο 2 «Προγενέστερη έρευνα στο πεδίο» γίνεται αναφορά στην προγενέστερη έρευνα που έχει γίνει στο αντίστοιχο ερευνητικό πεδίο. Πιο συγκεκριμένα, παρουσιάζονται τα βοηθητικά μέσα που έχουν στη διάθεσή τους τα άτομα με δυσκολία όρασης, γίνεται διαχωρισμός των

ηλεκτρονικών συστημάτων υποβοήθησης πλοϊγησης σε τρεις κύριες κατηγορίες και τέλος παρατίθενται μερικά από τα πιο αντιπροσωπευτικά συστήματα πλοϊγησης εξωτερικού χώρου που έχουν αναπτυχθεί.

- Στο κεφάλαιο 3 «Περιγραφή και ανάλυση του προτεινόμενου συστήματος» γίνεται μια αναλυτική περιγραφή της αρχιτεκτονικής του προτεινόμενου συστήματος, των ελάχιστων προδιαγραφών ενός συστήματος πλοϊγησης για άτομα με προβλήματα όρασης και παρουσιάζονται με λεπτομέρεια τα διάφορα μέρη που χρησιμοποιήθηκαν για την υλοποίησή του.
- Στο κεφάλαιο 4 «Πειραματικός Έλεγχος Συστήματος» παρουσιάζεται ο τρόπος με τον οποίο έγινε ο πειραματικός έλεγχος της ορθής λειτουργίας του συστήματος και αναφέρονται τα διάφορα αποτελέσματα που προέκυψαν.
- Στο κεφάλαιο 5 «Συμπεράσματα - Προτάσεις για το μέλλον» αναλύονται τα πλεονεκτήματα και τα μειονεκτήματα του προτεινόμενου συστήματος και παρατίθενται διάφορες προτάσεις για την βελτίωση της αποδοτικότητάς του.

2. Προγενέστερη έρευνα στο πεδίο

Για πολλά χρόνια οι άνθρωποι προσπαθούσαν να βρουν λύσεις που θα βελτίωναν την καθημερινότητα των ατόμων με μειωμένη όραση. Ο κύριος στόχος κάθε τέτοιας προσπάθειας ήταν να καταστήσουν πιο ασφαλή την πλοιόγηση του χρήστη στον χώρο και παράλληλα να αυξήσουν την ταχύτητα με την οποία ο τυφλός χρήστης λαμβάνει αποφάσεις κατά τη διάρκεια της πλοιόγησής του. Στο κεφάλαιο αυτό γίνεται μια αναδρομή, ξεκινώντας από τις κλασικές μεθόδους πλοιόγησης των ατόμων με μειωμένη όραση μέχρι και τις πιο σύγχρονες ηλεκτρονικές συσκευές υποβοήθησης ταξιδιού (ETAs). Επιπλέον, παρουσιάζεται η τρέχουσα τάση στο σχετικό ερευνητικό πεδίο και παρατίθενται οι διάφορες κατηγορίες των συστημάτων υποβοήθησης πλοιόγησης.

2.1 Βοηθητικά μέσα υποστήριξης ατόμων με μειωμένη όραση

Εδώ και εκατοντάδες χρόνια οι άνθρωποι χρησιμοποιούσαν διάφορα μέσα με σκοπό την καθοδήγηση των ατόμων με μειωμένη όραση. Τα πιο διαδεδομένα και ευρέως αποδεκτά από την κοινότητα των τυφλών είναι το λευκό μπαστούνι και ο σκύλος οδηγός [18]. Τις τελευταίες δεκαετίες, μετά την ραγδαία άνοδο της τεχνολογίας, ήρθαν στο προσκήνιο νέες συσκευές υποβοήθησης πλοιόγησης που "σκανάρουν" το εξωτερικό περιβάλλον και μεταδίδουν την πληροφορία αυτή στον τυφλό μέσω ηχητικών ή απτικών σημάτων.

2.1.1 Συμβατικά μέσα καθοδήγησης

Πρόκειται για τα μέσα που χρησιμοποιούνται κατά κόρον από τα άτομα με περιορισμένη όραση. Τα μέσα αυτά έχουν αποδειχθεί πολύ χρήσιμα στην καθημερινότητα των τυφλών, έχουν κερδίσει την εμπιστοσύνη της κοινότητας αυτής, αλλά πολύ συχνά η περιορισμένη ελευθερία κινήσεων που επιτρέπουν ή το κόστος τους αποτελούν ανασταλτικούς παράγοντες για την απόκτησή τους.

2.1.1.1 Λευκό μπαστούνι

Το λευκό μπαστούνι είναι ίσως το πιο διαδεδομένο μέσο καθοδήγησης των τυφλών. Προσφέρει αυτονομία και αίσθηση ασφάλειας στον χρήστη, καθώς του επιτρέπει να σκανάρει τον περιβάλλοντα χώρο και να εντοπίζει τυχόν εμπόδια ή επιφάνειες προσανατολισμού, με περιορισμένο ωστόσο εύρος κινήσεων. Παράλληλα, η χρήση του επιτρέπει στους υπόλοιπους περαστικούς να αναγνωρίσουν την ιδιότητα του τυφλού και συνεπώς να είναι πιο προσεκτικοί. Το λευκό του χρώμα έχει νομοθετηθεί και αποτελεί ένδειξη ότι το άτομο που το κρατάει έχει προβλήματα όρασης.

Θετικά	Αρνητικά
Εύκολη χρήση Φορητό ¹ Χαμηλό κόστος	Περιορισμένο εύρος κινήσεων Αδυναμία αναγνώρισης του είδους του εμποδίου

Πίνακας 2.1: Λευκό μπαστούνι, θετικά και αρνητικά



Σχήμα 2.1: Παράδειγμα χρήσης λευκού μπαστουνιού

2.1.1.2 Σκύλος οδηγός

Για πολλούς χρήστες με προβλήματα όρασης ο σκύλος οδηγός αποτελεί την καλύτερη λύση, αφού συνδυάζει την αίσθηση της ασφάλειας στην μετακίνηση με την αίσθηση της συντροφικότητας. Ο σκύλος οδηγός είναι εκπαιδευμένος να καθοδηγεί τον τυφλό με ασφάλεια, να αποφεύγει εμπόδια και να εντοπίζει στοιχεία όπως διαβάσεις, πόρτες, σκάλες, ταμεία και μέσα μαζικής μεταφοράς (π.χ., την είσοδο στο λεωφορείο). Η καθοδήγηση από έναν σκύλο οδηγό είναι πιο γρήγορη σε σύγκριση με το λευκό μπαστούνι, προσφέροντας μεγαλύτερη αυτονομία στον χρήστη. Παρ' όλα αυτά η εκπαίδευση ενός σκύλου οδηγού διαρκεί αρκετούς μήνες και απαιτεί αρκετούς οικονομικούς πόρους.

Θετικά	Αρνητικά
Ασφάλεια	Υψηλό κόστος αγοράς
Ταχύτητα	Κόστος συντήρησης
Συντροφικότητα	Ανάγκη εκπαίδευσης σκύλου

Πίνακας 2.2: Σκύλος οδηγός, θετικά και αρνητικά



Σχήμα 2.2: Σκύλος οδηγός

2.1.2 Ηλεκτρονικές συσκευές υποβοήθησης ταξιδιού (ETAs)

Όσο η τεχνολογία προχωρούσε με ραγδαίους ρυθμούς και η υπολογιστική ισχύς γινόταν όλο και πιο μεγάλη, έκαναν την εμφάνισή τους οι λεγόμενες ηλεκτρονικές συσκευές υποβοήθησης ταξιδιού ή αλλιώς ETAs (Electronic Travel Aids). Αν και τα πρώτα χρόνια εμφάνισή τους οι ETAs επικεντρώνονταν κυρίως στην ανίχνευση και αποφυγή εμποδίων, τα τελευταία χρόνια έχουν αναπτυχθεί επαρκώς ώστε να υποστηρίζουν πλήρη περιγραφή των εμποδίων, της διαδρομής του χρήστη αλλά και συνεχή ανάδραση. Οι ερευνητές που ασχολούνται με την ανάπτυξη αυτών των συσκευών αξιοποιούν τις πληροφορίες που λαμβάνουν από το περιβάλλον γύρω από τον χρήστη και χρησιμοποιώντας συγκεκριμένες τεχνικές τον ειδοποιούν για τυχόν εμπόδια ή για την διαδρομή που πρέπει να ακολουθήσει.

2.1.2.1 Αντίληψη περιβάλλοντος

Οι άνθρωποι χρησιμοποιούν τα διάφορα αισθητήρια κανάλια για να κατανοήσουν το περιβάλλον γύρω τους. Όταν μιλάμε για πλοϊγηση σε έναν χώρο τότε το κύριο αισθητήριο κανάλι που χρησιμοποιείται είναι αυτό της όρασης. Τα άτομα με μειωμένη ή καθόλου όραση τείνουν να αντικαθιστούν το οπτικό κανάλι με αυτό της ακοής ή της αφής. Συνήθως, το ακουστικό κανάλι είναι αυτό που χρησιμοποιείται κατά κύριο λόγο για να προσανατολιστούν και να αναγνωρίσουν τυχόν εμπόδια. Επιπρόσθετα, το απτικό κανάλι αξιοποιείται σε περίπτωση που τα άτομα αυτά θέλουν να κατανοήσουν με μεγαλύτερη λεπτομέρεια ένα εμπόδιο-αντικείμενο κοντά τους.

Επομένως, γίνεται κατανοητό ότι για να μπορέσουν οι ETAs να υποστηρίζουν τα άτομα με προβλήματα όρασης πρέπει αρχικά να είναι σε θέση να αντιληφθούν το περιβάλλον γύρω τους με επιτυχία και ακρίβεια. Τον ρόλο των ματιών παίζουν οι διάφοροι τύποι αισθητήρων, όπως οι αισθητήρες υπερήχων, τα ραντάρ και οι κάμερες. Με χρήση αισθητήρων όπως αυτοί των υπερήχων το σύστημα είναι δυνατό να εντοπίσει ένα εμπόδιο μέχρι κάποια ορισμένη απόσταση, χωρίς ωστόσο να δώσει απάντηση σχετικά με την μορφολογία του. Από την άλλη μεριά, η χρήση κάμερας δίνει την δυνατότητα για πλήρη καταγραφή του περιβάλλοντος και με κατάλληλη επεξεργασία της εικόνας ή χρήση τεχνικών μηχανικής μάθησης μπορεί να δώσει απαντήσεις σχετικά με την μορφολογία των εμποδίων, την απόστασή τους, την υφή τους και το βέλτιστο μονοπάτι που μπορεί να ακολουθήσει ο χρήστης. Περισσότερα για τις κατηγορίες των συστημάτων πλοϊγησης ανάλογα τον τρόπο "αίσθησης" του περιβάλλοντος θα αναλυθούν παρακάτω.

2.1.2.2 Ανάδραση με χρήση ηχητικών ειδοποιήσεων

Πέρα από την "αίσθηση" και την καταγραφή του περιβάλλοντος, ένα σύστημα πλοϊγησης για άτομα με μειωμένη όραση θα πρέπει να μπορεί να μεταδώσει την οπτική πληροφορία στον χρήστη με αποδοτικό τρόπο. Επειδή όπως είπαμε και παραπάνω τα άτομα αυτά αξιοποιούν την αίσθηση της ακοής για τον προσανατολισμό τους τα περισσότερα ηλεκτρονικά συστήματα πλοϊγησης χρησιμοποιούν το ακουστικό κανάλι για την μετάδοση της απαραίτητης πληροφορίας. Με άλλα λόγια, ο χρήστης ειδοποιείται μέσω ηχητικών σημάτων για την ύπαρξη κάποιου εμποδίου ή σε περίπτωση που υπάρχει ανάγκη για μια συγκεκριμένη δράση, όπως π.χ. να πραγματοποιήσει στροφή αριστερά/δεξιά ή να διασχίσει μια διάβαση πεζών. Τα ηχητικά αυτά σήματα μπορεί να είναι απλοί σύντομοι ή στιγμιαίοι ήχοι που ενημερώνουν τον χρήστη σχετικά με την απόσταση ενός εμποδίου, ή μπορεί να είναι σύντομες φράσεις/προτάσεις που να καθοδηγούν τον χρήστη σχετικά με την κατεύθυνση που πρέπει να ακολουθήσει.

2.1.2.3 Ανάδραση με χρήση απτικών ειδοποιήσεων

Πέρα από την ακοή, μια ακόμα αίσθηση που έχουν πολύ ανεπτυγμένη τα άτομα με προβλήματα όρασης είναι η αφή. Η αίσθηση της αφής είναι εκείνη που μας βοηθάει να αντιληφθούμε λεπτομέ-

ρειες σχετικά με τα αντικείμενα, όπως π.χ. την υφή των αντικειμένων. Τα τελευταία χρόνια έχει δοθεί ιδιαίτερη έμφαση στις συσκευές ETAs που βασίζονται στην ανάδραση με χρήση απτικών σημάτων, κυρίως επειδή ένα απτικό ερέθισμα είναι πολύ πιο διακριτικό σε σύγκριση με ένα ηχητικό και η αντίδραση του χρήστη σε αυτό είναι σχετικά πιο άμεση [19]. Οι διάφορες ETAs που έχουν κυκλοφορήσει κατά καιρούς χρησιμοποιούν ορισμένα μοτίβα δονήσεων για να ενεργοποιήσουν την αίσθηση της αφής στο χρήστη.

2.2 Κατηγορίες συστημάτων υποβοήθησης πλοήγησης

Η πλοήγηση ενός ατόμου μπορεί να διαιρεθεί σε δύο περιπτώσεις: α) πλοήγηση σε εσωτερικό χώρο, και β) πλοήγηση σε εξωτερικό χώρο. Και στις δύο περιπτώσεις για να έχουμε μια επιτυχμένη πλοήγηση είναι απαραίτητο να γνωρίζουμε τον προσανατολισμό που πρέπει να ακολουθήσει ο χρήστης, αλλά παράλληλα να ελέγχουμε ότι το μονοπάτι που ακολουθεί είναι ελεύθερο από εμπόδια. Συνεπώς, ένα ολοκληρωμένο σύστημα υποβοήθησης πλοήγησης για άτομα με μειωμένη όραση θα πρέπει να είναι σε θέση τόσο να παρέχει την κατεύθυνση του χρήστη σε σχέση με τον προορισμό του, όσο και να του εξασφαλίζει ένα ασφαλές μονοπάτι.

Οι πρώτες συσκευές ETAs επικεντρώνονταν περισσότερο στο δεύτερο σκέλος, αυτό της αποφυγής εμπόδιων, το οποίο επιτύχαναν κυρίως με τη χρήση αισθητήρων υπερήχων ή σόναρ που οποία εντόπιζαν ένα εμπόδιο σε μια συγκεκριμένη απόσταση και ειδοποιούσαν τον χρήστη για την ύπαρξη του. Με την ανάπτυξη της τεχνολογίας GPS (Global Positioning System) επήλθε και εντυπωσιακή πρόοδος όσον αφορά το πρώτο σκέλος των ETAs, αυτό της πλοήγησης. Το GPS λειτουργεί αξιοποιώντας την ύπαρξη 30+ δορυφόρων σε τροχιά γύρω από την γη, βάσει των οποίων μπορεί να εντοπίσει την τοποθεσία ενός δέκτη με ακρίβεια από μερικά μέτρα έως μερικά εκατοστά [20]. Η τεχνολογία αυτή επέτρεψε την κατασκευή συστημάτων πλοήγησης που όχι μόνο αναγνώριζαν πιθανά εμπόδια, αλλά μπορούσαν να κατευθύνουν τον χρήστη κατά την μετακίνησή του προς έναν προορισμό με αρκετά μεγάλη ακρίβεια. Ωστόσο, αν και το GPS λειτουργούσε αρκετά καλά στον εξωτερικό χώρο, η πλοήγηση στον εσωτερικό χώρο παρέμενε ακόμα ένα δύσκολο εγχείρημα, από την στιγμή που το σήμα του GPS εξασθενούσε σε μεγάλο βαθμό μέσα σε κτίρια. Για την αντιμετώπιση της πλοήγησης σε εσωτερικού χώρους έχουν δοθεί πολλαπλές λύσεις, με την πιο ενδιαφέρουσα από αυτές να αποτελεί η χρήση RFID (Radio Frequency IDentification), η οποία θα αναλυθεί παρακάτω.

Τέλος, με την αύξηση της επεξεργαστικής ισχύος και την δυνατότητα κατασκευής όλο και μικρότερων επεξεργαστών τα συστήματα πλοήγησης, τόσο για εσωτερικούς όσο και για εξωτερικούς χώρους, άρχισαν πλέον να βασίζονται σε κάμερες και αισθητήρες βάθους. Παρακάτω παρουσιάζονται μερικά συστήματα πλοήγησης για άτομα με προβλήματα όρασης που έχουν αναπτυχθεί από προηγούμενες προσπάθειες ερευνητών. Έχει γίνει διαχωρισμός τους με βάση την τεχνολογία την οποία αξιοποιούνται κατά κύριο λόγο για να πετύχουν την πλοήγηση του χρήστη ή την αναγνώριση των εμποδίων στην διαδρομή του.

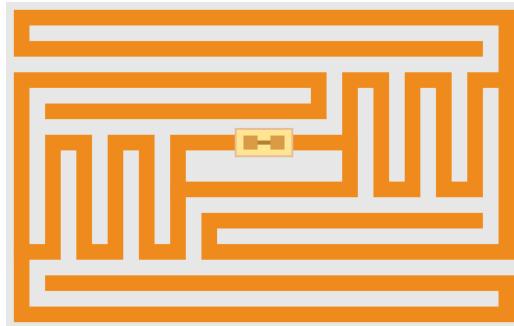
2.2.1 RFID

Η τεχνολογία RFID, στα ελληνικά ταυτοποίηση μέσω ραδιοσυχνοτήτων, αποτελεί την εξέλιξη των ραβδωτών κωδίκων barcode και χρησιμοποιείται ευρέως στην καθημερινή ζωή των ανθρώπων, κυρίως μέσω του εμπορίου. Ενα σύστημα RFID αποτελείται από δύο διαφορετικά μέρη:

1. τον **πομποδέκτη**, οποίος συχνά αναφέρεται ως ετικέτα RFID (RFID tag) και
2. τον **αναγνώστη ή αισθητήρα** (RFID reader).

Ένα RFID tag αποτελείται από ένα εσωτερικό ολοκληρωμένο κύκλωμα το οποίο περιέχει μια κεραία για την επικοινωνία με τους RFID readers και μία μνήμη, στην οποία μπορούν να αποθηκευτούν

διάφορα δεδομένα ή πληροφορίες για το αντικείμενο που φέρει την ετικέτα RFID. Ανάλογα την εφαρμογή στην οποία χρησιμοποιείται, η μνήμη μιας ετικέτας RFID μπορεί να είναι μόνο για ανάγνωση (*Read Only Memory - ROM*), επανεγγράψιμη μνήμη (*Read-Write Memory*), ή μνήμη μιας εγγραφής και πολλών αναγνώσεων (*Write Once, Read Many - WORM*) [21].



Σχήμα 2.3: Παράδειγμα ολοκληρωμένου κυκλώματος ενός παθητικού RFID tag

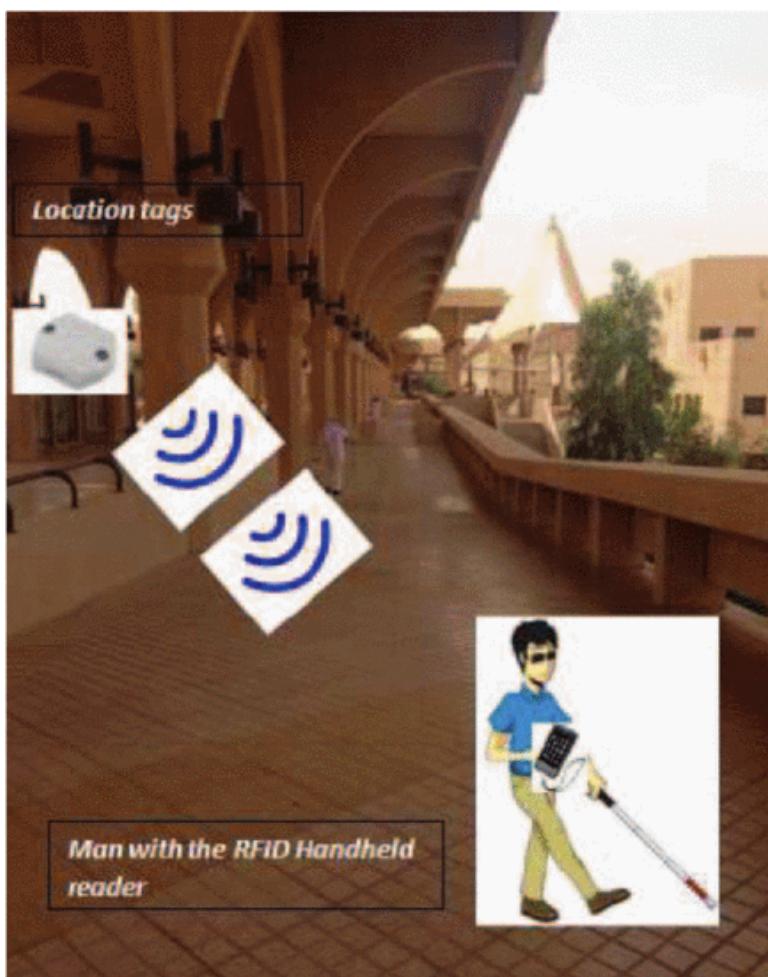
Αντίστοιχα, ένας RFID reader αποτελείται από μια κεραία, η οποία αναλαμβάνει την επικοινωνία με το RFID tag και μια μονάδα ελέγχου για την διαχείριση της ανταλλαγής δεδομένων με την εκάστοτε ετικέτα. Τα συστήματα RFID διακρίνονται σε ενεργητικά και παθητικά, ανάλογα με τον τρόπο επικοινωνίας της ετικέτας με τον αναγνώστη [22, 23]. Στα ενεργητικά συστήματα οι ετικέτες τροφοδοτούνται αυτόνομα από μια εξωτερική πηγή τροφοδοσίας, όπως π.χ. μια μπαταρία, και έχουν την δυνατότητα να μεταδώσουν την πληροφορία τους σε μεγαλύτερες αποστάσεις. Αντίθετα, στα παθητικά συστήματα οι ετικέτες δεν έχουν δική τους πηγή τροφοδοσίας, αλλά λαμβάνουν την απαραίτητη ενέργεια από τους αναγνώστες RFID μέσω επαγωγικής σύζευξης κατά την φάση επικοινωνίας τους. Αν και τα παθητικά συστήματα έχουν πολύ μικρότερη εμβέλεια, αποτελούν την πλειοψηφία των RFID συστημάτων σήμερα και είναι αυτά που χρησιμοποιούνται σε εφαρμογές πλοήγησης για άτομα με μειωμένη όραση. Ένα τυπικό παράδειγμα λειτουργίας ενός τέτοιου συστήματος είναι η χρήση της πιστωτικής μας κάρτας κατά τη διάρκεια ανέπαφης συναλλαγής σε ένα κατάστημα. Η κάρτα μας παίζει τον ρόλο του RFID tag, ενώ η συσκευή POS του καταστήματος παίζει τον ρόλο του RFID reader. Μόλις η κάρτα βρεθεί αρκετά κοντά στο POS, λαμβάνει μέσω επαγωγικής σύζευξης την απαραίτητη ενέργεια από αυτό και τροφοδοτείται το ενσωματωμένο της ολοκληρωμένο κύκλωμα. Με αυτό τον τρόπο ενεργοποιείται ουσιαστικά η μετάδοση πληροφορίας, από την κάρτα προς το POS, η οποία πραγματοποιείται μέσω της κεραίας που υπάρχει και στα δύο μέρη.



Σχήμα 2.4: Παράδειγμα ενός RFID reader

Η ενσωμάτωση των RFID συστημάτων σε συσκευές πλοϊγησης για άτομα με προβλήματα όρασης αποτελεί μια έξυπνη λύση κυρίως σε εσωτερικούς χώρους, καθώς μπορούμε να τοποθετήσουμε πολλά RFID tags σε διαφορετικά σημεία-κόμβους μέσα στον χώρο, ή στο κτίριο που θέλουμε να καλύψουμε, ή ακόμα και κατά μήκος μιας διαδρομής που ακολουθούν τα άτομα με προβλήματα όρασης. Χρησιμοποιώντας έναν RFID reader τα άτομα αυτά θα είναι σε θέση να λαμβάνουν πληροφορίες σχετικά με την τοποθεσία στην οποία βρίσκονται και να καθοδηγούνται προς μια διαφορετική τοποθεσία, η οποία έχει ένα αντίστοιχο RFID tag.

Πραγματοποιώντας μια επισκόπηση της βιβλιογραφίας βρίσκουμε αρκετά συστήματα πλοϊγησης που βασίζονται στην αξιοποίηση της τεχνολογίας RFID. Ένα αντιπροσωπευτικό παράδειγμα είναι αυτό που προτείνεται στο [2]. Το προτεινόμενο σύστημα λειτουργεί σε εξωτερικούς και εσωτερικούς χώρους και αποτελείται από έναν RFID reader ενσωματωμένο στο λευκό μπαστούνι του χρήστη και πολλαπλά RFID tags διασκορπισμένα σε καίρια σημεία της διαδρομής του. Το σύστημα εντοπίζει την τοποθεσία του χρήστη είτε μέσω GPS, είτε μέσω ανάγνωσης του κοντινότερου RFID tag χρησιμοποιώντας τον ενσωματωμένο RFID reader. Σε περίπτωση που ο χρήστης θελήσει να κατευθυνθεί προς μια προκαθορισμένη τοποθεσία (π.χ. γραφείο, ή τουαλέτα) τότε το σύστημα τον προτρέπει να κινηθεί προς την σωστή κατεύθυνση, η οποία υπολογίζεται με χρήση ηλεκτρονικής πυξίδας, GPS και μέσω της ανάγνωσης του επόμενου RFID tag στο μονοπάτι του χρήστη. Πιο συγκεκριμένα, γίνεται σύγκριση της τωρινής θέσης του χρήστη σε σχέση με τον προορισμό και το κοντινότερο RFID tag και μέσω αυτής υπολογίζεται η κατάλληλη κατεύθυνση. Για να μπορεί λαμβάνει αξιόπιστες αποφάσεις το σύστημα διατηρεί μια βάση δεδομένων με όλες τις τοποθεσίες των τοποθετημένων RFID tags, συμπεριλαμβανομένων των αποστάσεων μεταξύ τους, οι οποίες ποικίλλουν από 4 έως 15 μέτρα.

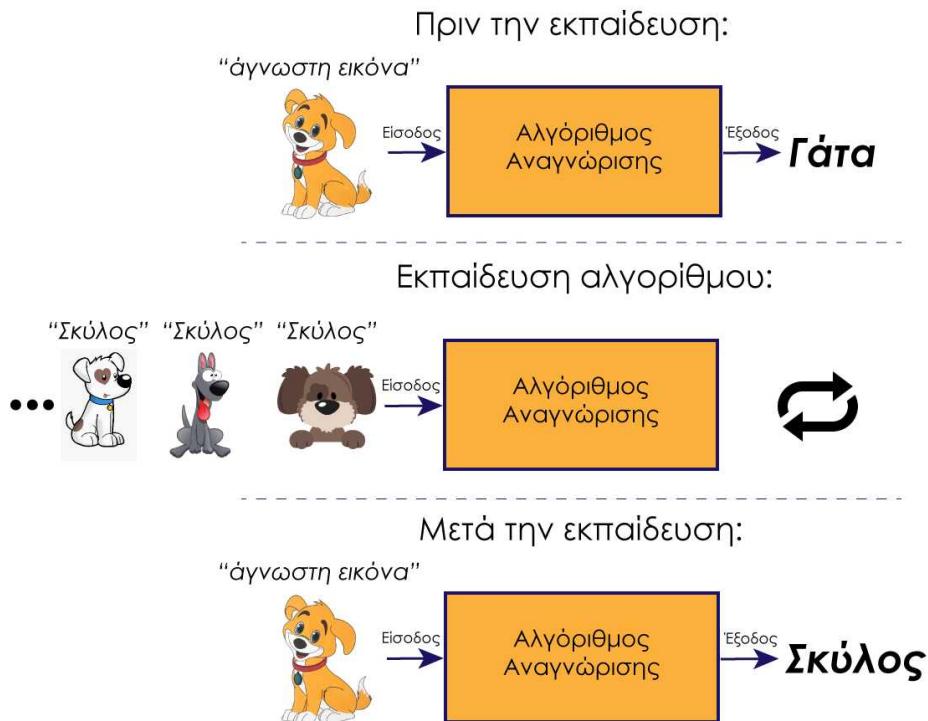


Σχήμα 2.5: Περιγραφή του συστήματος που προτείνεται στο [2]

Παρατηρούμε ότι το παραπάνω σύστημα πετυχαίνει εν μέρει το σκοπό του, δηλαδή την καθοδήγηση του χρήστη προς ένα σημείο ενδιαφέροντος, αλλά γίνεται σαφές ότι ένα σύστημα πλοήγησης που βασίζεται σε RFID tags μπορεί να έχει ουσιαστικό νόημα μόνο σε εσωτερικούς χώρους ή σε συγκεκριμένες τοποθεσίες ενδιαφέροντος. Η ελεύθερη πλοήγηση του χρήστη μέσα σε μια πόλη δεν μπορεί να πραγματοποιηθεί, διότι σε αυτήν την περίπτωση θα απαιτούνταν ένας τεράστιος αριθμός από RFID tags κατά μήκος κάθε πιθανής διαδρομής του χρήστη!

2.2.2 Learning-based

Μια άλλη κατηγορία συστημάτων πλοήγησης είναι αυτή που βασίζεται στην αναγνώριση αντικειμένων/εμποδίων αξιοποιώντας τεχνικές της Μηχανικής Μάθησης (Machine Learning). Η βασική ιδέα πίσω από τέτοιου είδους συστήματα είναι ότι εκπαιδεύουμε έναν αλγόριθμο να αναγνωρίζει με πολύ μεγάλη ακρίβεια ένα αντικείμενο ή μια ομάδα αντικειμένων. Με άλλα λόγια, για να μπορέσει ο αλγόριθμος να εντοπίσει ένα συγκεκριμένο μοτίβο μέσα σε μια εικόνα, είτε αυτό είναι πρόσωπο, είτε είναι αντικείμενο, όπως π.χ. φανάρι, δρόμος κλπ, είναι απαραίτητο να έχει προηγηθεί μια διαδικασία "μάθησης", κατά την οποία παρέχουμε στον αλγόριθμο έναν πολύ μεγάλο αριθμό από γνωστές εικόνες που περιέχουν το μοτίβο που αναζητούμε. Ο αλγόριθμος λαμβάνει υπόψιν του όλες τις εικόνες που έχουν προηγηθεί και "δημιουργεί" ένα μοντέλο του αντικειμένου που θέλουμε να αναγνωρίζει. Αφού έχει επέλθει η διαδικασία της μάθησης το σύστημα είναι εκπαιδευμένο να αναγνωρίζει το συγκεκριμένο μοτίβο σε οποιαδήποτε άλλη νέα/άγνωστη εικόνα του δώσουμε ως είσοδο.



Σχήμα 2.6: Παράδειγμα εκπαίδευσης ενός αλγορίθμου μάθησης

Η εκπαίδευση αλγορίθμων για βελτιστοποιημένη αναγνώριση αντικειμένων δεν είναι μια νέα ιδέα στον τομέα της Υπολογιστικής Όρασης (Computer Vision), αλλά αποτελεί εξέλιξη της Αναγνώρισης Προτύπων (Pattern Recognition) η οποία μέχρι πρότινος χρησιμοποιούσε τις παραδοσιακές

τεχνικές που σχετίζονται με την επεξεργασία εικόνας και σημάτων. Η ραγδαία αύξηση στην χρήση της μηχανικής μάθησης ως μέθοδο αναγνώρισης αντικειμένων συνέβη λόγω της αύξησης της διαθέσιμης επεξεργαστικής ισχύος τις τελευταίες δεκαετίες. Πλέον, ακόμα και ένα smartphone είναι ικανό να παρέχει επαρκή επεξεργαστική ισχύ για να εφαρμόσουμε μεθόδους μηχανικής μάθησης.

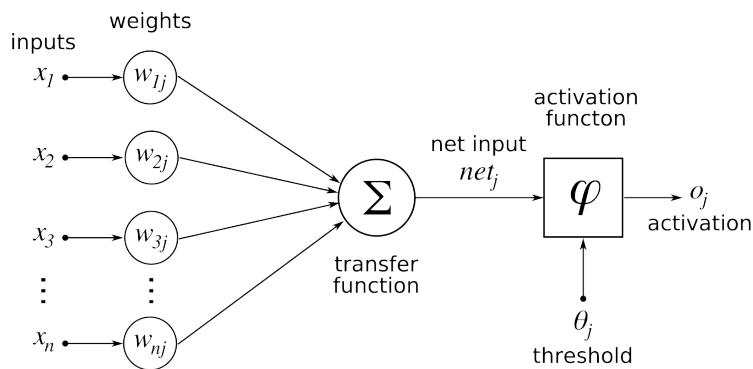
Οι εργασίες μηχανικής μάθησης χωρίζονται σε 3 (τρεις) μεγάλες κατηγορίες [24] :

- **Supervised Learning** (Επιβλεπόμενη Μάθηση): Ο αλγόριθμος μάθησης κατασκευάζει μια συνάρτηση που απεικονίζει τα δεδομένα εισόδου (σύνολο εκπαίδευσης) σε γνωστές επιθυμητές εξόδους. Η μάθηση χαρακτηρίζεται επιβλεπόμενη επειδή υπάρχει κάποιος "επιβλέπων" που παρέχει την σωστή τιμή εξόδου της συνάρτησης για το σύνολο εκπαίδευσης. Στόχος του αλγόριθμου είναι να "γενικεύσει" αυτήν την συνάρτηση, ώστε να μπορεί να προγνώσει με ακρίβεια την έξοδο του συστήματος σε δεδομένα εισόδου με αρχικά άγνωστη έξοδο. Μερικά παραδείγματα τεχνικών supervised learning είναι:
 - Δέντρα Απόφασης (Decision Trees)
 - Μάθηση Κανόνων (Rule Learning)
 - Μάθηση κατά Περίπτωση (Instance Based Learning)
 - Μάθηση κατά Bayes
 - Γραμμική Παρεμβολή (Linear Regression)
 - Νευρωνικά Δίκτυα (Neural Networks)
 - Μηχανές Διανυσμάτων Υποστήριξης (Support Vector Machines)
- **Unsupervised Learning** (Μη Επιβλεπόμενη Μάθηση): Ο αλγόριθμος αναλύει το σύνολο των δεδομένων εισόδου και κατασκευάζει ένα μοντέλο υπό μορφή συσχετίσεων ή ομαδοποιήσεων, χωρίς να γνωρίζει τις επιθυμητές εξόδους. Σαν αποτέλεσμα (έξοδος) προκύπτουν πρότυπα (περιγραφές), κάθε ένα από τα οποία περιγράφει ένα μέρος από τα δεδομένα. Τέτοιου τύπου αλγόριθμοι χρησιμοποιούνται κυρίως σε προβλήματα:
 - Ανάλυσης Συσχετισμών (Association Analysis)
 - Ομαδοποίησης (Clustering)
- **Reinforcement Learning** (Ενισχυτική Μάθηση): Αποτελεί την πιο "αυτόνομη" μορφή αλγορίθμων μάθησης, όπου το σύστημα αλληλεπιδρά άμεσα με το περιβάλλον και μαθαίνει μια στρατηγική ενεργειών, που αποσκοπεί στην επίτευξη ενός στόχου, π.χ. οδήγηση ενός οχήματος. Χρησιμοποιείται κυρίως σε προβλήματα Σχεδιασμού (Planning), όπως για παράδειγμα ο έλεγχος κίνησης ρομπότ και η βελτιστοποίηση εργασιών σε εργοστασιακούς χώρους.

Στα συστήματα πλοήγησης χρησιμοποιείται κυρίως το supervised learning και πιο συγκεκριμένα εκπαιδεύονται Νευρωνικά Δίκτυα για την αναγνώριση του περιβάλλοντος.

2.2.2.1 Νευρωνικά δίκτυα

Ένα τεχνητό νευρωνικό δίκτυο - TNΔ (Artificial Neural Network - ANN) στην μηχανική μάθηση είναι ένας αλγόριθμος εκμάθησης που εμπνέεται από τη δομή και τις λειτουργικές πτυχές των βιολογικών νευρωνικών δικτύων του ανθρώπινου εγκεφάλου. Με άλλα λόγια ο όρος "Νευρωνικά Δίκτυα" περιγράφει έναν αριθμό από διαφορετικά μαθηματικά μοντέλα που προσπαθούν να μιμηθούν τη συμπεριφορά των βιολογικών νευρώνων. Όπως τα νευρωνικά δίκτυα του εγκεφάλου περιέχουν έναν πολύ μεγάλο αριθμό από μεμονωμένα διακριτά στοιχεία, τους νευρώνες, έτσι και τα τεχνητά νευρωνικά δίκτυα αποτελούνται από έναν μεγάλο αριθμό απλών κόμβων (nodes) διασυνδεδεμένων μεταξύ τους, που παίζουν τον ρόλο των νευρώνων και οι οποίοι οργανώνονται σε στρώματα (layers).



Σχήμα 2.7: Απεικόνιση λειτουργίας του απλούστερου τεχνητού νευρωνικού δικτύου, που ονομάζεται *Perceptron* και αποτελείται από έναν μόνο νευρώνα

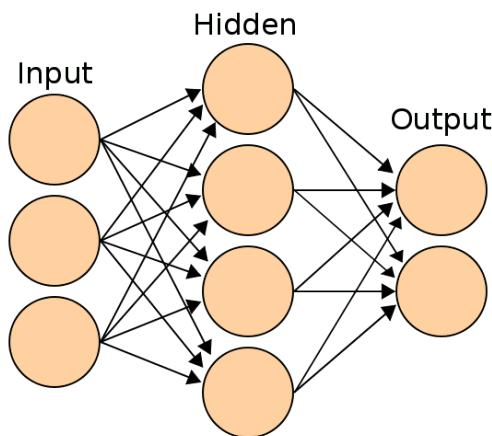
Κάθε τεχνητός νευρώνας αποτελείται από πολλές εισόδους x_i και μία μόνο έξοδο y . Κάθε είσοδος x_i πολλαπλασιάζεται με ένα βάρος w_i και αθροίζεται μέσω της συνάρτησης αθροίσματος (summation function) F , ως εξής:

$$F = \sum_i^n x_i w_i$$

Ο τεχνητός νευρώνας δίνει έξοδο μέσω της συνάρτησης ενεργοποίησης (activation function), μόνο όταν το σταθμισμένο άθροισμα F είναι μεγαλύτερο μιας ορισμένης τιμής κατωφλίου (threshold value) θ [25], δηλαδή όταν,

$$\sum_i^n x_i w_i - \theta > 0$$

Το απλούστερο νευρωνικό δίκτυο Perceptron [26] (σχήμα 2.7) αποτελείται από έναν μόνο νευρώνα και ένα επίπεδο. Συνήθως τα TNΔ είναι πιο σύνθετα και οργανώνονται σε πολλαπλά επίπεδα (layers) τα οποία καλούνται και στρώματα. Τα ενδιάμεσα επίπεδα ονομάζονται κρυμμένα επίπεδα (hidden layers) και δεν είναι απαραίτητο να υπάρχουν. Οι κόμβοι-νευρώνες που υπάρχουν σε κάθε επίπεδο είναι συνδεδεμένοι μεταξύ τους, ώστε κάθε κόμβος-νευρώνας να έχει συνδέσμους με πολλούς άλλους κόμβους του ίδιου ή άλλου επιπέδου. Το δίκτυο δέχεται τις εισόδους x_i στο επίπεδο εισόδου (input layer), το οποίο αλληλεπιδρά με τα κρυμμένα ενδιάμεσα επίπεδα, και παρέχει την έξοδό του μέσω του επιπέδου εξόδου (output layer).



Σχήμα 2.8: Δομή επιπέδων σε ένα απλό τεχνητό νευρωνικό δίκτυο με ένα ενδιάμεσο κρυμμένο επίπεδο

Σκοπός της παρούσας διπλωματικής εργασίας δεν είναι η αναλυτική περιγραφή των νευρωνικών δικτύων, γι' αυτό θα προχωρήσουμε κατευθείαν στην εφαρμογή τους όσον αφορά τα συστήματα πλοήγησης για άτομα με προβλήματα όρασης.

Η σημασία των νευρωνικών δικτύων στα συστήματα πλοϊγησης σχετίζεται με την ακρίβεια που πετυχαίνουν στην αναγνώριση αντικειμένων. Πιο συγκεκριμένα, τα τελευταία χρόνια χρησιμοποιούνται τα λεγόμενα Deep Neural Networks (Βαθιά Νευρωνικά Δίκτυα), τα οποία αποτελούνται από πολλά κρυμμένα επίπεδα και παρέχουν πολύ υψηλή ακρίβεια σε εφαρμογές όπως η αναγνώριση προσώπων, η κατηγοριοποίηση αντικειμένων κλπ. Ένα χαρακτηριστικό παράδειγμα αποτελεί το σύστημα που προτείνεται στο [27], το οποίο χρησιμοποιεί την κάμερα από το smartphone του χρήστη και έναν υπολογιστή ως server. Το κύριο χαρακτηριστικό της συγκεκριμένης έρευνας είναι ότι, καθώς ο χρήστης πλοιγείται στον χώρο, το smartphone τραβάει φωτογραφίες από τον χώρο μπροστά από τον χρήστη και τις στέλνει, μέσω διαδικτύου, στον server όπου γίνεται η επεξεργασία τους. Πιο συγκεκριμένα, πριν την αποστολή της φωτογραφίας, γίνεται επιλογή των χαρακτηριστικών (feature selection) που θα χρησιμοποιηθούν ως είσοδοι στο νευρωνικό δίκτυο στην πλευρά του server. Ο server "τρέχει" δύο διαφορετικούς αλγορίθμους ταξινόμησης εικόνων, ανάλογα την ταχύτητα της διαδικτυακής σύνδεσης του χρήστη. Όταν η ταχύτητα σύνδεσης είναι υψηλή εφαρμόζεται ο αλγόριθμος Faster R-CNN (Faster Region-based Convolutional Neural Network), ενώ όταν η ταχύτητα σύνδεσης είναι αργή εφαρμόζεται ο αλγόριθμος YOLO (You Only Look Once). Αφού ο κάθε αλγόριθμος έχει εκπαιδευτεί να αναγνωρίζει και να ταξινομεί μια ομάδα αντικειμένων, το σύστημα αποστέλλει τα features, που έχουν εξαχθεί κατά τη διαδικασία του feature selection, στον server και αυτός με την σειρά του ταξινομεί την δοθείσα φωτογραφία.

Όπως αναφέρεται και στην ίδια την έρευνα, αν και η χρήση των νευρωνικών δικτύων έχει ως αποτέλεσμα αρκετά υψηλή ακρίβεια, απαιτείται ιδιαίτερα υψηλή επεξεργαστική ισχύς, κάτι που δικαιολογεί και την χρήση του server ως εξωτερική μονάδα επεξεργασίας. Η χρήση του server και η αναγκαία μετάδοση των εικόνων από το smartphone προς αυτόν, συνεπάγεται την ύπαρξη μιας μικρής καθυστέρησης, η οποία μπορεί να γίνει σημαντικά μεγαλύτερη ανάλογα την ταχύτητα μετάδοσης και την πολυπλοκότητα του δικτύου που χρησιμοποιείται, έως ότου το σύστημα αποφανθεί για τον τύπο του εμποδίου που βρίσκεται μπροστά από τον χρήστη. Με άλλα λόγια, η έλλειψη επεξεργαστικής ισχύος οδηγεί σε προβλήματα όσον αφορά την real-time αναγνώριση αντικειμένων.

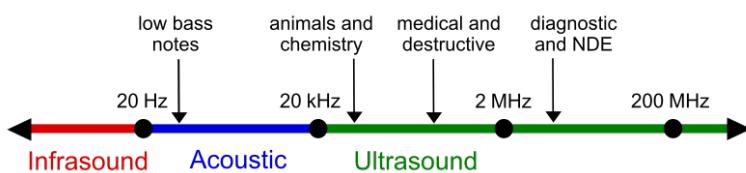
Καταλήγοντας, είναι αναγκαίο να τονίσουμε το γεγονός ότι τα βαθιά νευρωνικά δίκτυα αποτελούν πλέον μια πολύ αξιόπιστη λύση όσον αφορά τον εντοπισμό και την αναγνώριση εμποδίων σε συστήματα πλοϊγησης, καθώς παρέχουν μεγάλα ποσοστά ακρίβειας και είναι σε θέση να αναγνωρίζουν ταυτόχρονα πολλαπλά αντικείμενα σε μια φωτογραφία. Από την άλλη πλευρά, όμως, ο περιοριστικός παράγοντας των υψηλών απαιτήσεων σε επεξεργαστική ισχύ και η αναγκαιότητα εκπαίδευσης των νευρωνικών δικτύων εκ των προτέρων, αποτελούν τροχοπέδη στην πλήρη αξιοποίησή τους και εφαρμογή τους σε φορητά συστήματα πλοϊγησης για άτομα με μειωμένη όραση. Αξίζει, δε, να τονίσουμε ότι η εκπαίδευση των νευρωνικών δικτύων προϋποθέτει την ύπαρξη κατάλληλου σετ δεδομένων (dataset), για κάθε ξεχωριστό μοτίβο/αντικείμενο που θέλουμε να αναγνωρίζει το σύστημα, τα οποία δεν είναι πάντα διαθέσιμα. Αυτός είναι και ο λόγος που η παρούσα διπλωματική εργασία χρησιμοποιεί παραδοσιακές μεθόδους αναγνώρισης αντικειμένων, όπως αυτές που περιγράφονται παρακάτω.

2.2.3 Vision-based

Η κατηγορία αυτή των συστημάτων περιλαμβάνει τα συστήματα εκείνα που βασίζονται σε παραδοσιακές μεθόδους ανίχνευσης εμποδίων και αναγνώρισης αντικειμένων. Πιο συγκεκριμένα, η πλειοψηφία των συστημάτων πλοϊγησης χρησιμοποιεί διάφορους αισθητήρες απόστασης, όπως αισθητήρες υπερήχων, LiDaR κλπ, για την ανίχνευση της ύπαρξης εμποδίου μπροστά από τον χρήστη. Παράλληλα, η ανίχνευση και η αναγνώριση του είδους των αντικειμένων γίνεται με μεθόδους ψηφιακής επεξεργασίας εικόνας. Παρακάτω αναλύεται η λειτουργία των βασικών αισθητήρων που εντοπίζονται σε ένα τέτοιο σύστημα.

2.2.3.1 Ultrasonic sensors

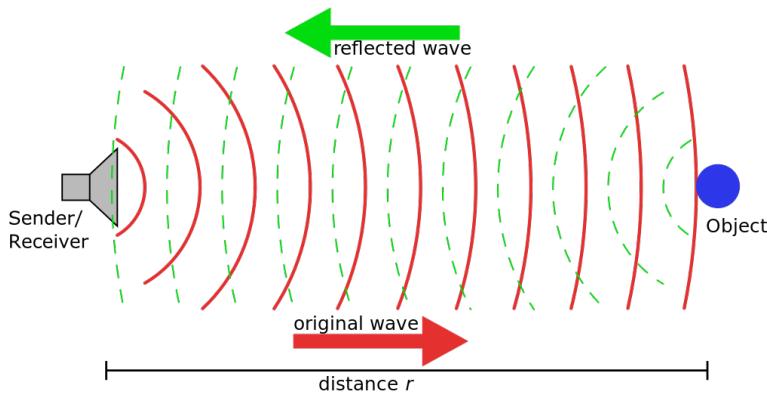
Η χρήση υπερήχων ανέκαθεν αποτελούσε μια από τις πιο γνωστές μεθόδους μέτρησης μικρών αποστάσεων και ανίχνευσης αντικειμένων, ενώ ακόμη και σήμερα αξιοποιούνται σε πολλές εφαρμογές, όπως είναι ο μη καταστροφικός έλεγχος (non destructive inspection) στη βιομηχανία ή η εξέταση υπερήχων στην απεικονιστική ιατρική. Ο άνθρωπος μπορεί να ακούσει ήχους σε ένα εύρος συχνοτήτων από 20 Hz έως 20 KHz (Σχήμα 2.9). Οι υπέρηχοι είναι ηχητικά κύματα που έχουν συχνότητες μεγαλύτερες από 20 KHz, άρα δεν γίνονται αντιληπτοί από έναν φυσιολογικό άνθρωπο. Η σημασία των υπερήχων στην πλοήγηση και τον ηχο-εντοπισμό (echolocation) ανακαλύφθηκε για πρώτη φορά από τον Ιταλό Lazzaro Spallanzani, ο οποίος μελέτησε την συμπεριφορά των νυχτερίδων και παρατήρησε ότι προσανατολίζονται χρησιμοποιώντας υπερήχους και όχι όραση [28].



Σχήμα 2.9: Διάγραμμα εύρους ηχητικών συχνοτήτων

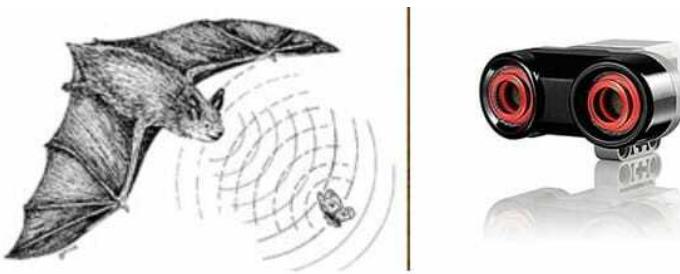
Οι αισθητήρες υπερήχων αποτελούνται από έναν πομπό και έναν δέκτη υπερήχων και εκτιμούν την απόσταση ενός στόχου λαμβάνοντας υπόψη τους την αντανάκλαση ενός ραδιοκύματος ή ενός ηχητικού σήματος πάνω στο στόχο (Σχήμα 2.10). Για να το επιτύχουν αυτό χρησιμοποιούν τον χρόνο που έκανε το σήμα για να καλύψει την απόσταση από τον αισθητήρα στο αντικείμενο και πίσω. Η απόσταση ενός αντικειμένου μπροστά από τον αισθητήρα υπερήχου δίνεται ως εξής [29]:

$$distance = \frac{Time_of_Flight \times Speed_of_Sound}{2}$$



Σχήμα 2.10: Αρχή λειτουργίας αισθητήρα υπερήχων

Χάρη στην ευκολία χρήσης τους και το χαμηλό τους κόστος, πολλές εφαρμογές πλοήγησης αξιοποιούν τους αισθητήρες υπερήχων για να εντοπίζουν εμπόδια κατά μήκος του μονοπατιού του χρήστη. Ένα αντιπροσωπευτικό παράδειγμα αποτελεί η σχετική υλοποίηση στο [30], όπου το προτεινόμενο σύστημα χρησιμοποιεί αποκλειστικά αισθητήρες υπερήχων για την καθοδήγηση του χρήστη. Χρησιμοποιώντας δύο αισθητήρες οι οποίοι εκπέμπουν παλμούς υπερήχων στα 40 KHz, είναι σε θέση να εντοπίζει συμπαγή αντικείμενα σε απόσταση 0.03 έως 6 μέτρα, και ανάλογα την εκτίμηση της απόστασης ειδοποιεί τον χρήστη μέσω δονήσεων.



Σχήμα 2.11: Αριστερά: Απεικόνιση του ηχο-εντοπισμού της νυχτερίδας. Δεξιά: Παράδειγμα ενός τυπικού αισθητήρα υπερήχων με έναν δέκτη και έναν πομπό.

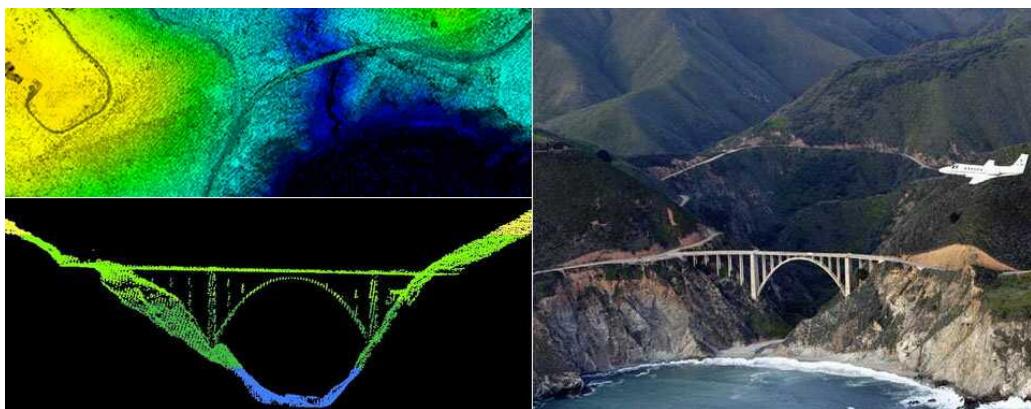
Αν και η χρήση υπερήχων λειτουργεί ικανοποιητικά σε εσωτερικούς χώρους, η αξιοποίησή τους σε εξωτερικά περιβάλλοντα έχει αρκετά μειονεκτήματα, κυρίως, λόγω της παρεμβολής από άλλα ηχητικά σήματα του ίδιου φάσματος. Επίσης, ένα μεγάλο μειονέκτημα αυτής της μεθόδου είναι η αδυναμία αναγνώρισης του είδους του αντικειμένου που ανιχνεύεται.

2.2.3.2 LiDAR sensors

Οι αισθητήρες Light Detection And Ranging, εν συντομίᾳ LiDAR, είναι συσκευές που λειτουργούν αξιοποιώντας το φως και την αντανάκλασή του πάνω σε επιφάνειες. Η αρχή λειτουργίας τους μοιάζει πολύ με εκείνη των αισθητήρων υπερήχων, μόνο που στην περίπτωση αυτή δεν εκπέμπουν υπερήχους, αλλά βασίζονται στην εκπομπή παλμικής ακτινοβολίας λέιζερ στην ατμόσφαιρα. Οι ακτίνες αντανακλούν πάνω στα αντικείμενα του περιβάλλοντος και ο αισθητήρας καταγράφει την ανακλώμενη ακτινοβολία λέιζερ, υπολογίζοντας την απόσταση των αντικειμένων ως εξής [31]:

$$distance = \frac{Time_of_Flight \times Speed_of_Light}{2}$$

Πιο συγκεκριμένα, οι αισθητήρες LiDAR εκπέμπουν ακτίνες φωτός στο υπεριώδες, στο ορατό, ή κοντά στο υπέρυθρο φάσμα ηλεκτρομαγνητικής ακτινοβολίας και είναι πολύ πιο ακριβείς από τους αντίστοιχους αισθητήρες υπερήχων, ενώ μπορούν επίσης να χρησιμοποιηθούν σε ένα μεγάλο εύρος υλικών, συμπεριλαμβανομένου μη-μεταλλικών αντικειμένων, βράχων, βροχής, ακόμα και σε μοριακό επίπεδο. Αυτή η μεγάλη ευελιξία των αισθητήρων LiDAR τους καθιστά πολύ χρήσιμους για την κατασκευή 3D αναπαραστάσεων των αντικειμένων-στόχων (3D reconstruction). Οι εφαρμογές τους περιλαμβάνουν την χρήση τους σε αεροφωτογραφίες, σε αυτόνομα οχήματα κ.α., ενώ αξίζει να τονίσουμε το γεγονός πως λειτουργούν καλύτερα σε εσωτερικά περιβάλλοντα, εξαιτίας την επίδρασης της υπέρυθρης ακτινοβολίας του ήλιου σε εξωτερικούς χώρους [32].



Σχήμα 2.12: 3D ψηφιακή αναπαράσταση μιας τοποθεσίας μέσω LiDAR



Σχήμα 2.13: Αισθητήρες LiDAR τοποθετούνται πάνω σε αυτόνομα οχήματα

Αν και η χρήση LiDAR δεν είναι τόσο διαδεδομένη σε φορητά συστήματα πλοιόγησης, κυρίως, λόγω του όγκου τους και του υψηλού κόστους, υπάρχουν προσπάθειες που αναδεικνύουν την χρησιμότητά τους. Για παράδειγμα, στο [3] προτείνεται ένα σύστημα πλοιόγησης για άτομα με προβλήματα όρασης που βασίζεται αποκλειστικά στη χρήση ενός 2D LiDAR αισθητήρα (Σχήμα 2.14), ο οποίος σκανάρει το περιβάλλον γύρω του σε ένα εύρος 270° και στέλνει τα δεδομένα σε μια μονάδα ψηφιακής επεξεργασίας σημάτων. Το αποτέλεσμα ανακοινώνεται στον χρήστη μέσω ηχητικών ειδοποιήσεων.



Σχήμα 2.14: Ο αισθητήρας LiDAR που χρησιμοποιείται στο [3]

2.2.3.3 RGB-D Camera

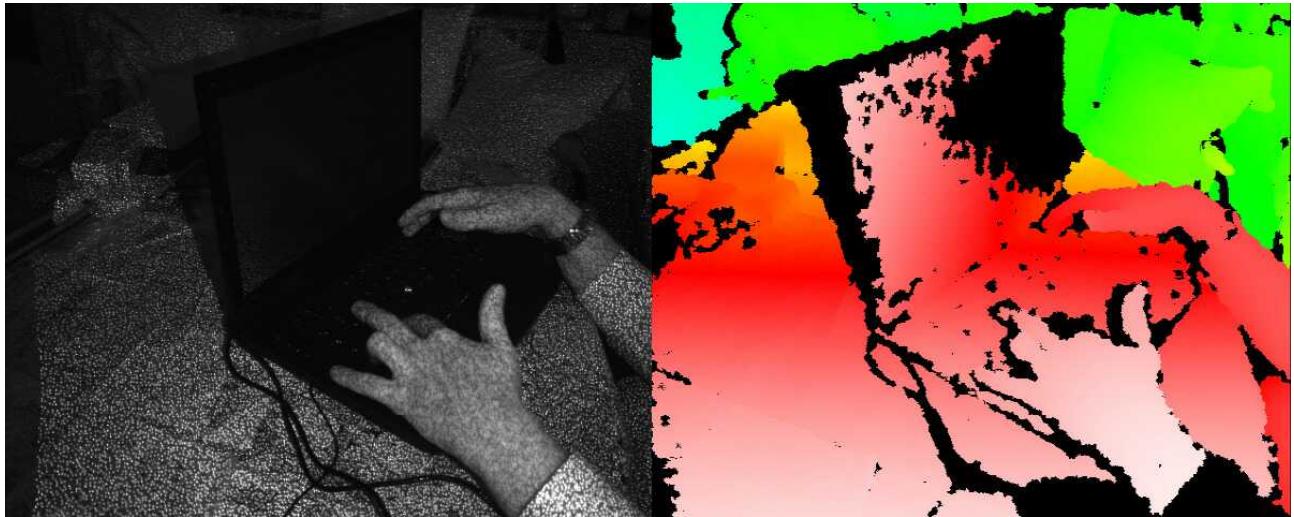
Η πλειοψηφία των συστημάτων πλοιόγησης για άτομα με προβλήματα όρασης χρησιμοποιεί κάμερες με ενσωματωμένο αισθητήρα βάθους, οι λεγόμενες RGB-D κάμερες. Τέτοια συστήματα είναι σε θέση να ανιχνεύουν τα εμπόδια μπροστά από τον χρήστη μέσω του αισθητήρα βάθους, όσο και να αναγνωρίζουν τον τύπο του εμποδίου, ή του αντικειμένου που μας ενδιαφέρει, μέσω της επεξεργασίας της RGB εικόνας που λαμβάνεται από την κάμερα.

Η πιο διαδεδομένη κάμερα με ενσωματωμένο αισθητήρα βάθους είναι η *Microsoft Kinect RGB-D Camera*, η οποία πέρα από την αρχική της χρήση για το XBOX 360, αποτελεί την λύση σε πολλές πειραματικές προσπάθειες για ανάπτυξη φορητών συστημάτων πλοιόγησης.



Σχήμα 2.15: Ο αισθητήρας Kinect

Αυτοί οι αισθητήρες αποτελούνται συνήθως από έναν πομποδέκτη υπέρυθρης ακτινοβολίας (Infrared - IR) και έναν αισθητήρα RGB. Η αρχή λειτουργίας τους είναι η εξής: ο πομπός υπερύθρων προβάλλει στον χώρο ένα μοτίβο από σημεία υπέρυθρου φωτός (Structured Light technique) και ο δέκτης δέχεται τις ανακλώμενες ακτίνες. Το μοτίβο σημείων που εκπέμπεται παραμορφώνεται ανάλογα την επιφάνεια και τα αντικείμενα στα οποία ανακλάται. Στη συνέχεια γίνεται μια σύγκριση της αρχικής δομής των σημείων φωτός και αυτής που λήφθηκε από τον δέκτη. Έτσι, ανάλογα την παραμόρφωση που έχει υποστεί η δομή αυτή, παράγεται μια εικόνα βάθους, στην οποία κάθε pixel αντικατοπτρίζει την καρτεσιανή απόσταση (σε χιλιοστά) από την επιφάνεια της κάμερας μέχρι το κοντινότερο αντικείμενο στις συγκεκριμένες x, y συντεταγμένες [33].



Σχήμα 2.16: Αριστερά: Το μοτίβο από σημεία υπέρυθρου φωτός που εκπέμπεται. Δεξιά: Η παραγόμενη εικόνα βάθους ανάλογα με την παραμόρφωση της δομής του μοτίβου

Ένα παράδειγμα συστήματος πλοιόγησης για άτομα με προβλήματα όρασης, που χρησιμοποιεί τον αισθητήρα Kinect, είναι αυτό που προτείνεται στο [34]. Πιο συγκεκριμένα, παρουσιάζεται η δημιουργία ενός λειτουργικού, χαμηλού κόστους και διακριτικού συστήματος πλοιόγησης, στο οποίο ο χρήστης φορά ένα γιλέκο με ένα ενσωματωμένο χωρικό πλέγμα από 16 κινητήρες δόνησης, που ενεργοποιούνται ανάλογα με τα εμπόδια που εντοπίζονται μπροστά από τον χρήστη. Η ανίχνευση των εμποδίων πραγματοποιείται μέσω του αισθητήρα Kinect.

3. Περιγραφή και ανάλυση του προτεινόμενου συστήματος

Η παρούσα διπλωματική εργασία εστιάζει στην υλοποίηση ενός φορητού συστήματος πλοήγησης για άτομα με προβλήματα όρασης, το οποίο θα χρησιμοποιείται παράλληλα με το παραδοσιακό λευκό μπαστούνι και θα επιτρέπει στα άτομα αυτά να μετακινούνται αυτόνομα και με ασφάλεια σε ένα αστικό περιβάλλον. Ένα από τα κύρια χαρακτηριστικά του είναι η χρήση δονήσεων ως μέσο ειδοποίησης του χρήστη. Επίσης, λαμβάνοντας υπόψιν την διαχρονική δυσκολία των χρηστών να υιοθετήσουν και να εμπιστευτούν ένα νέο σύστημα πλοήγησης [35, 36], αποφασίσαμε να δώσουμε έμφαση στην διακριτικότητα και την απλότητα του συστήματος, ώστε να μπορεί να υιοθετηθεί άμεσα από οποιονδήποτε μέσο χρήστη με προβλήματα όρασης. Στη συνέχεια αυτού του κεφαλαίου θα παρουσιαστούν αναλυτικά όλα τα επιμέρους υπο-συστήματα που απαρτίζουν το προτεινόμενο σύστημα, συμπεριλαμβανομένων του hardware και του software που χρησιμοποιήθηκαν.

3.1 Προδιαγραφές ενός συστήματος πλοήγησης για χρήστες με μειωμένη όραση

Έχουν γίνει πολλές προσπάθειες κατασκευής ενός συστήματος πλοήγησης που θα καλύπτει τις σύγχρονες ανάγκες των ατόμων με προβλήματα όρασης, αλλά καμία λύση έως τώρα δεν έχει καταφέρει να αντικαταστήσει πλήρως τα παραδοσιακά μέσα υποβοήθησης, όπως το λευκό μπαστούνι και ο σκύλος οδηγός. Πολύ συχνά οι τεχνολογικές λύσεις παρουσιάζονται ως πανάκεια, αλλά εν τέλει δεν καταφέρνουν να κερδίσουν την εμπιστοσύνη της κοινότητας των τυφλών, παρέχοντας απλά επιφανειακές λύσεις στα προβλήματα. Σύμφωνα με το [35], οι προσφερόμενες τεχνολογικές λύσεις πέφτουν συνήθως στις παρακάτω παγίδες:

1. **Η παγίδα του μηχανικού:** Συνήθως οι σχεδιαστές των συστημάτων πλοήγησης τα σχεδιάζουν βάσει των προσωπικών τους εμπειριών και υποκειμενικών κριτηρίων όσον αφορά τα προβλήματα που αντιμετωπίζουν οι χρήστες με μειωμένη όραση. Αντίστοιχα, η συγκεκριμένη τακτική πολλές φορές μας οδηγεί στην υλοποίηση μιας λύσης για ένα μη υπαρκτό πρόβλημα και καταλήγει να αποτελεί επιστήμη για την επιστήμη, χωρίς κάποια ουσιαστική εφαρμογή στον πραγματικό κόσμο.
2. **Ανεπαρκής γνώση:** Ο σχεδιασμός των συστημάτων πλοήγησης δεν λαμβάνει υπόψιν του, ή έχει ανεπαρκής γνώση, των αντιληπτικών και γνωστικών παραγόντων που σχετίζονται με την επεξεργασία μη οπτικής πληροφορίας. Με άλλα λόγια, οι σχεδιαστές οφείλουν να μελετάνε τον τρόπο με τον οποίο μπορεί να απορροφηθεί πιο αποτελεσματικά μια πληροφορίας μέσω απτικού ή ακουστικού καναλιού και να μην αρκούνται στην απλή μετατροπή του οπτικού σήματος σε ένα μη οπτικό ερέθισμα.
3. **Ελλειψη εξειδίκευσης:** Με τον όρο ”εξειδίκευση” εννοούμε την δημιουργία μιας λύσης που θα στοχεύει στην αντιμετώπιση ενός συγκεκριμένου προβλήματος που σχετίζεται με την κινητικότητα των ατόμων με προβλήματα όρασης. Πολλές από τις προσφερόμενες λύσεις χαρακτη-

ρίζονται ως βοηθήματα γενικού σκοπού, καταλήγοντας να προσφέρουν ανεπαρκείς λύσεις στις πραγματικές προκλήσεις των χρηστών. Επομένως, είναι ιδιαίτερα σημαντικό η λύση που προσφέρουμε να είναι εστιασμένη σε ένα συγκεκριμένο πρόβλημα και να ενσωματώνεται εύκολα στην καθημερινή ζωή του μέσου χρήστη.

3.1.1 Δυνατότητα υποκατάστασης της όρασης με άλλες αισθήσεις (ακοή, αφή)

Η όραση είναι η βασική αίσθηση του ανθρώπου κατά την πλοιόγησή του σε ένα περιβάλλον και μας επιτρέπει να φέρνουμε εις πέρας τις καθημερινές εργασίες γρήγορα και αποδοτικά. Είναι γεγονός ότι τα άτομα με προβλήματα όρασης βασίζονται στις υπόλοιπες αισθήσεις τους, τις οποίες έχουν αναπτύξει επαρκώς, ώστε να μπορούν να προσανατολίζονται και να αντιλαμβάνονται τον χώρο στον οποίο βρίσκονται. Οι δύο βασικές αισθήσεις που χρησιμοποιούνται ως υποκατάστατα της όρασης είναι η ακοή και η αφή, με την πρώτη να κατέχει την μερίδα του λέοντος στα συστήματα πλοιόγησης [37].

Τα άτομα με προβλήματα όρασης χρησιμοποιούν την τεχνική του ηχο-εντοπισμού (echolocation), όπως οι νυχτερίδες, για να προσανατολίζονται στον χώρο. Για παράδειγμα, χρησιμοποιώντας το λευκό μπαστούνι και "χτυπώντας" το πάνω σε επιφάνειες ή εμπόδια μπορούν να καταλάβουν το είδος και την υφή της επιφάνειας. Επίσης, το ακουστικό κανάλι χρησιμοποιείται για να αντιλαμβάνονται τα δυναμικά περιβάλλοντα, όπως την κίνηση των οχημάτων, την ύπαρξη πεζών κ.α. και, επομένως, γίνεται κατανοητό ότι αποτελεί έναν πολύ σημαντικό και κρίσιμο πυλώνα στην μετακίνηση των ατόμων αυτών και κάθε επιπλέον προσπάθεια για αξιοποίησή του από συσκευές πλοιόγησης οδηγεί στην υπερφόρτωσή του.

Από την άλλη πλευρά, το απτικό κανάλι, δηλαδή η αίσθηση της αφής, δεν αξιοποιείται στο έπακρο ως μέσο αντίληψης του περιβάλλοντος. Τα άτομα με προβλήματα όρασης χρησιμοποιούν την αφή για να αναγνωρίσουν αντικείμενα, να καταλάβουν την υφή τους και το σχήμα τους, αλλά μόνο όταν βρίσκονται σε πολύ κοντινή απόσταση. Κατά την μεγαλύτερη διάρκεια της πλοιόγησής τους σε μια πόλη δεν απαιτείται η χρήση της αφής για να αντιληφθούν το περιβάλλον τους, γεγονός που αφήνει πολλά περιθώρια αξιοποίησή της από τις νέες συσκευές πλοιόγησης [38], γι' αυτό και η παρούσα διπλωματική εργασία επικεντρώνεται στην απτική αλληλεπίδραση χρήστη-συστήματος, όπως θα αναλυθεί και παρακάτω.

3.1.2 Αξιοποιητία συστήματος

Η αξιοποιητία του προτεινόμενου συστήματος είναι βασική προϋπόθεση, ώστε να μπορέσει να γίνει αποδεκτό από την κοινότητα των ατόμων με προβλήματα όρασης. Είναι γεγονός πως πολλές καινοτόμες ιδέες που ευελπιστούν να βελτιώσουν την ζωή των ΑμΕΑ δεν πληρούν τις κατάλληλες προϋποθέσεις για να χαρακτηριστούν αποτελεσματικές λύσεις. Για να είναι αξιόπιστο ένα σύστημα πλοιόγησης πρέπει να χαρακτηρίζεται από επαναληψιμότητα, δηλαδή να δίνει παρόμοιες μετρήσεις με μικρή απόκλιση σε πειράματα με ίδιες συνθήκες, και από ακρίβεια, δηλαδή να παρέχει ικανοποιητικά αποτελέσματα [39].

Η παρούσα ερευνητική εργασία στοχεύει στην υλοποίηση ενός συστήματος το οποίο θα λειτουργεί επαρκώς καλά κάτω από συγκεκριμένες συνθήκες περιβάλλοντος. Λόγω του πειραματικού χαρακτήρα μιας τέτοιας εργασίας είναι φυσικό να υπάρχουν αποκλίσεις και να μην εγγυάται η σωστή λειτουργία του συστήματος κάτω από οποιεσδήποτε φυσικές συνθήκες.

3.1.3 Ανθρωποκεντρική σχεδίαση διεπαφής χρήστη

Ένα από τα σημαντικότερα χαρακτηριστικά ενός τεχνολογικού βοηθήματος που προορίζεται για ΑμΕΑ είναι η δυνατότητα αλληλεπίδρασης που έχει ο χρήστης με το σύστημα. Συχνά σχεδιάζονται συσκευές πλοιόγησης με εξαιρετικά αποτελέσματα ακρίβειας και αξιοπιστίας, αλλά παρόλα αυτά

δεν προχωράνε στην αγορά λόγω ελλιπούς σχεδιασμού διεπαφής χρήστη (User Interface). Η αλληλεπίδραση του χρήστη με το σύστημα θα πρέπει να είναι φυσική και όσο πιο απλή γίνεται, χωρίς να προσθέτει στον χρήστη επιπλέον πνευματικό φόρτο. Στην παρούσα εργασία προσανατολιζόμαστε στην σχεδίαση της εφαρμογής με βάση τις ανθρωποκεντρικές αρχές σχεδίασης [8], μελετώντας τις ανάγκες των ατόμων με προβλήματα όρασης και την ανατροφοδότησή τους σε άλλες προηγούμενες προσπάθειες υλοποίησης ενός συστήματος πλοιόγησης.

Αρχές ανθρωποκεντρικής σχεδίασης
Σχεδιασμός βασιζόμενος στους χρήστες και τις ανάγκες τους
Διατήρηση συνοχής
Απλοί και φυσικοί διάλογοι
Μείωση της περιττής πνευματικής προσπάθειας από τον χρήστη
Παροχή επαρκούς ανάδρασης στον χρήστη
Παροχή επαρκών μηχανισμών πλοιόγησης μεταξύ διαφορετικών διεπιφανειών
Εξάλειψη των περιορισμών από τον χρήστη
Ξεκάθαρη παρουσίαση πληροφοριών στον χρήστη
Παροχή βοήθειας όπου απαιτείται
Αποτελεσματική διαχείριση σφαλμάτων

Πίνακας 3.1: 10 βασικοί άξονες της ανθρωποκεντρικής σχεδίασης [8]

3.1.4 Ελαχιστοποίηση επεμβατικότητας του συστήματος

Τέλος, ένα σύστημα πλοιόγησης που προορίζεται για καθημερινή χρήση από άτομα με προβλήματα όρασης οφείλει να λαμβάνει υπόψιν του τις ανάγκες των χρηστών για ελάχιστη επεμβατικότητα στις συνήθειές τους και στους τρόπους που συμπεριφέρονται. Λύσεις που απαιτούν την ενδυμασία με συγκεκριμένα ρούχα, π.χ. ειδικό γιλέκο, ή λύσεις που είναι αισθητικά αποτρεπτικές για χρήση σε εξωτερικό περιβάλλον, π.χ. ογκώδεις συσκευές, δεν είναι ελκυστικές για τον τελικό χρήστη και τον κάνουν να αισθάνεται άβολα όταν τις χρησιμοποιεί. Τα προτεινόμενα συστήματα πρέπει να είναι φορητά, εύκολα στη χρήση και άμεσα διαθέσιμα. Στα πλαίσια της διπλωματικής εργασίας, εστιάσαμε στην φορητότητα του προτεινόμενου συστήματος και την εύκολη ενσωμάτωσή του στην καθημερινή ζωή του χρήστη.

3.2 Εστίαση σε διαβάσεις πεζών

Η έννοια της κινητικότητας των ατόμων με προβλήματα όρασης είναι αρκετά γενική και περιλαμβάνει ένα πολύ μεγάλο εύρος εφαρμογών και υπο-περιπτώσεων. Η αρχική μας ιδέα περιελάμβανε την υλοποίηση ενός συστήματος πλοιόγησης γενικού σκοπού, που θα ήταν σε θέση να βοηθάει τον χρήστη σε κάθε πιθανή διαδρομή μέσα σε ένα αστικό περιβάλλον. Ωστόσο, μετά από τις αρχικές προσπάθειες έγινε φανερό ότι ο στόχος αυτός ήταν μη ρεαλιστικός, λόγω του χρονικού περιορισμού και του πειραματικού χαρακτήρα μιας διπλωματικής εργασίας. Επομένως, αποφασίσαμε να εστιάσουμε σε ένα πιο συγκεκριμένο πρόβλημα που αντιμετωπίζουν τα άτομα αυτά κατά την μετακίνησή τους, διατηρώντας ωστόσο την καθολικότητα ενός τέτοιου συστήματος. Στα πλαίσια της βιβλιογραφικής έρευνας σχετικά με τις ανάγκες των χρηστών με προβλήματα όρασης [40, 15], βρέθηκε ότι μία από τις μεγαλύτερες προκλήσεις είναι η διάσχιση ενός δρόμου, είτε αυτός είναι μια λεωφόρος είτε ένας μικρότερος δρόμος. Για να διασχίσουν έναν δρόμο μέσω μια διάβασης πεζών τα άτομα με προβλήματα όρασης χρειάζονται κατά μέσο όρο 3 φορές περισσότερο χρόνο από ένα άτομο με φυσιολογική όραση [41], κυρίως επειδή απαιτείται, αρχικά, ο εντοπισμός του προσανατολισμού της διάβασης, ένα χρονικό διάστημα εξοικείωσης με την ροή των οχημάτων στο δρόμο, μια εκτίμηση του ρίσκου διάσχισης της διάβασης και η εκτίμηση της κατάστασης του φωτεινού σηματοδότη για πεζούς [42]. Ειδικά

σε περιπτώσεις όπου δεν παρέχεται φωνητική υποβοήθηση από τον φωτεινό σηματοδότη, η διαδικασία αξιολόγησης της διάβασης γίνεται ιδιαίτερα χρονοβόρα. Έτσι, καταλήξαμε να υλοποιήσουμε ένα σύστημα πλοήγησης που θα εστιάζει στην αναγνώριση διαβάσεων πεζών και θα συμβάλλει στην ασφαλέστερη και πιο γρήγορη διάσχισή τους από τον χρήστη με προβλήματα όρασης. Ουσιαστικά, ο χρήστης εξακολουθεί να έχει την ευθύνη για την διάσχιση ή όχι του δρόμου, αλλά το σύστημα τον βοηθάει στην λήψη της σωστής απόφασης, παρέχοντάς του κατάλληλη ανάδραση.

3.3 Περιγραφή του συστήματος

3.3.1 Γενική εικόνα

Το προτεινόμενο σύστημα πλοήγησης αποτελείται από 3 επιμέρους μέρη:

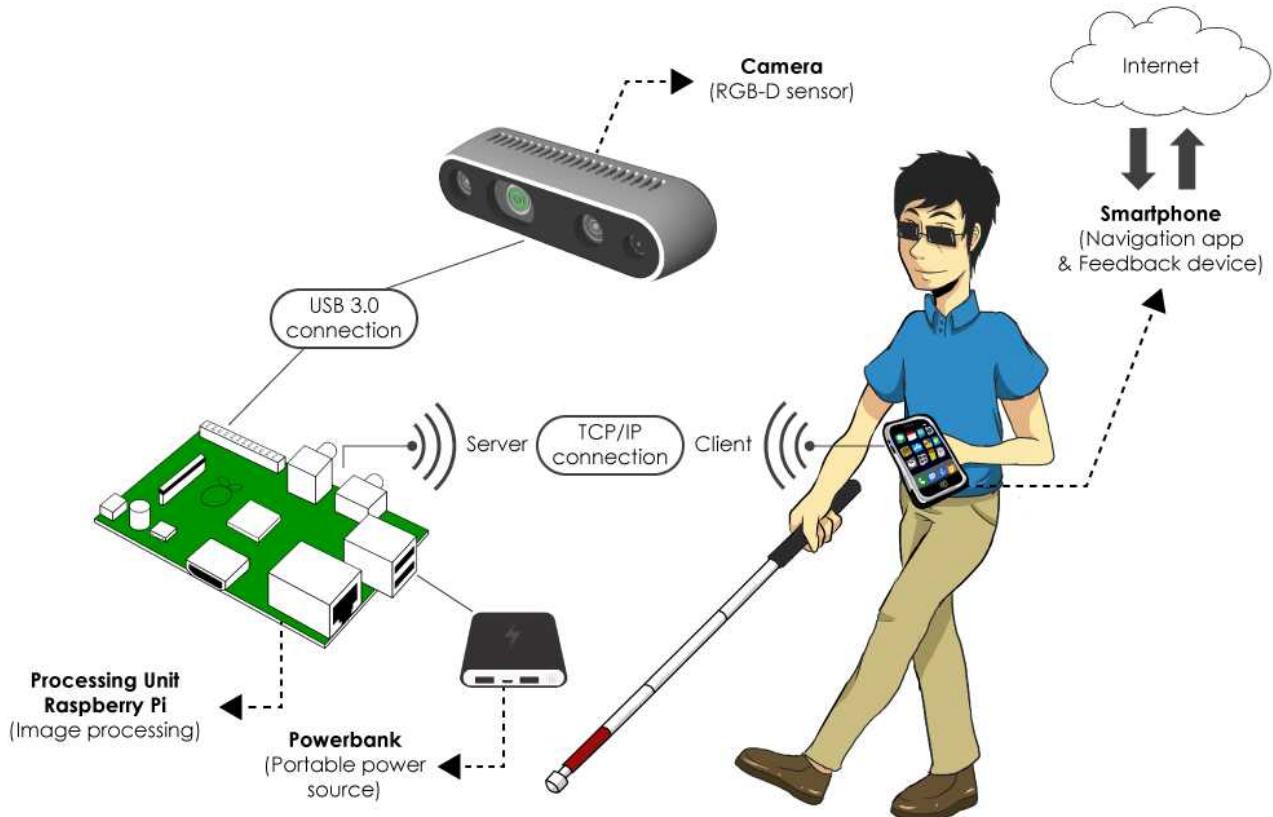
1. Αλγόριθμος αναγνώρισης διάβασης πεζών & κατάστασης φωτεινού σηματοδότη (Crosswalk and Pedestrian Light recognition)
2. Αλγόριθμος εντοπισμού εμποδίων μπροστά από τον χρήστη (Obstacle detection)
3. Εφαρμογή καθοδήγησης του χρήστη προς τον επιλεγμένο προορισμό (Navigation app)

Κάθε επιμέρους μέρος θεωρείται αυτόνομο και μπορεί να λειτουργήσει ανεξάρτητα από τα υπόλοιπα. Αυτό σημαίνει ότι ο χρήστης μπορεί να απενεργοποιήσει προσωρινά την αναγνώριση διαβάσεων και φαναριών και να χρησιμοποιεί μόνο την εφαρμογή καθοδήγησης με την αναγνώριση εμποδίων, ή το αντίστροφο. Οι δύο προτεινόμενοι αλγόριθμοι έχουν υλοποιηθεί σε ένα Raspberry Pi 4, ώστε να επιτευχθεί η φορητότητα του συστήματος, ενώ η εφαρμογή καθοδήγησης έχει υλοποιηθεί σε ένα android smartphone. Το σύστημα χρησιμοποιεί μια κάμερα με ενσωματωμένο αισθητήρα βάθους (RGB-D), η οποία τοποθετείται στο ύψος της μέσης του χρήστη και ο ρόλος της οποίας είναι να "βλέπει" το περιβάλλον μπροστά από τον χρήστη και να στέλνει την εικόνα στην επεξεργαστική μονάδα του συστήματος.

Ο χρήστης εισάγει τον επιθυμητό προορισμό μέσω της εφαρμογής smartphone και αυτή υπολογίζει την κοντινότερη διαδρομή από την τωρινή τοποθεσία του χρήστη προς τον επιλεγμένο προορισμό. Το σύστημα εισέρχεται σε λειτουργία πυξίδας και καθοδηγεί τον χρήστη προς το κατάλληλο μονοπάτι, παρέχοντάς του απτική ανάδραση για το πότε και πού πρέπει να κατευθυνθεί. Πιο συγκεκριμένα, όταν απαιτείται μια στροφή αριστερά ή δεξιά ο χρήστης δέχεται ένα αντίστοιχο μοτίβο δονήσεων στο κινητό του, που τον ενημερώνουν για το είδος της στροφής που απαιτείται. Αντίστοιχα, όταν, κατά τη διάρκεια της πλοήγησης, εντοπιστεί ένα εμπόδιο μπροστά από τον χρήστη, το σύστημα τον ειδοποιεί με χρήση δονήσεων για το αν πρέπει να μετακινηθεί δεξιά ή αριστερά, ανάλογα με τον διαθέσιμο χώρο. Τέλος, όταν η εφαρμογή καθοδηγήσει τον χρήστη μπροστά από μια διάβαση πεζών, ο αντίστοιχος αλγόριθμος αναγνωρίζει την διάβαση και το χρώμα του φαναριού και ενημερώνει τον χρήστη μέσω κατάλληλης απτικής ανάδρασης σχετικά με το αν μπορεί να διασχίσει τον δρόμο ή όχι.

3.3.2 Αρχιτεκτονική

Η αρχιτεκτονική του προτεινόμενου συστήματος φαίνεται στο σχήμα 3.1. Αρχικά, το κίνητρο για την υλοποίηση της συγκεκριμένης αρχιτεκτονικής δόθηκε από τις παρατηρήσεις και τις προτάσεις που παρουσιάζονται στο [43]. Στη συνέχεια, δόθηκε έμφαση στην απλότητα του συστήματος και έγινε ένα πρώτο πλάνο των συστατικών που θα το αποτελούν.



Σχήμα 3.1: Αρχιτεκτονική προτεινόμενου συστήματος

Η αρχιτεκτονική του συστήματος πλοιόγησης που προτείνεται στην παρούσα εργασία βασίζεται πάνω σε 3 κύριους πυλώνες:

- **Μια κάμερα RGB-D με ενσωματωμένο αισθητήρα βάθους:** Στην λύση που προτείνουμε χρησιμοποιείται μια κάμερα ως μονάδα "αίσθησης" του εξωτερικού περιβάλλοντος μπροστά από τον χρήστη, η οποία έχει τη δυνατότητα καταγραφής βίντεο σε έγχρωμη μορφή (RGB) και παράλληλα δίνει ως έξοδο έναν χάρτη βάθους (depth map). Ουσιαστικά η κάμερα παίζει τον ρόλο των ματιών και δίνει την δυνατότητα υπολογισμού της απόστασης διαφόρων αντικειμένων από το άτομο με προβλήματα όρασης.
- **Ένα smartphone με εγκατεστημένη την εφαρμογή πλοιόγησης:** Ως μέσο αλληλεπίδρασης του χρήστη με το σύστημα επιλέχθηκε το κινητό του. Ο λόγος είναι στην σημερινή εποχή η πλειοψηφία των ανθρώπων, ακόμα κι αυτών με προβλήματα όρασης, είναι εξοικειωμένοι με την χρήση ενός smartphone. Πρόκειται, δηλαδή, για μια συσκευή που ο καθένας έχει ανά πάσα στιγμή πάνω του και δεν προσθέτει επιπλέον φόρτο εργασίας στον χρήστη. Η εφαρμογή που αναπτύχθηκε στα πλαίσια της διπλωματικής εργασίας επιτρέπει στον χρήστη να εισάγει έναν προορισμό και η ίδια αναλαμβάνει την εύρεση της κατάλληλης διαδρομής, λαμβάνοντας υπόψιν την τωρινή τοποθεσία του χρήστη και τα δεδομένα από τους χάρτες πλοιόγησης. Γίνεται αξιοποίηση του ενσωματωμένου GPS του κινητού και της εσωτερικής του πυξίδας, που περιλαμβάνει μαγνητόμετρο, γυροσκόπιο και επιταχυνσιόμετρο.
- **Την κύρια μονάδα επεξεργασίας εικόνας:** Για να λειτουργήσει ένα σύστημα πλοιόγησης πρέπει να υπάρχει μια μονάδα η οποία θα λαμβάνει το βίντεο από την κάμερα και θα είναι υπεύθυνη για την επεξεργασία εικόνας που απαιτείται. Στην παρούσα εργασία επιλέχθηκε η χρήση ενός Raspberry Pi 4, το οποίο παρέχει την απαραίτητη επεξεργαστική ισχύ που απαιτείται για την επεξεργασία της ροής βίντεο, καθώς επίσης έχει και πολύ μικρό μέγεθος (λίγο μεγαλύτερο από

πιστωτική κάρτα), που συμβάλλει στην φορητότητα του συστήματος. Πιο συγκεκριμένα, στο Raspberry Pi ”τρέχουν” όλοι οι αλγόριθμοι εντοπισμού και αναγνώρισης εικόνας.

3.3.3 Πλοήγηση σε αστικό περιβάλλον – δυσκολίες και περιορισμοί

Όπως έχει ήδη αναφερθεί, στην παρούσα διπλωματική εργασία παρουσιάζεται ένα σύστημα πλοήγησης που προορίζεται για εξωτερικό χώρο και πιο συγκεκριμένα για αστικά περιβάλλοντα. Αν και οι δυσκολίες μετακίνησης είναι εξίσου έντονες και σε εσωτερικού χώρους, επιλέχθηκε η εστίαση σε εξωτερικά περιβάλλοντα, γιατί συνήθως αυτά είναι άγνωστα στους χρήστες και οι κίνδυνοι που μπορεί να αντιμετωπίσουν είναι πιο απρόβλεπτοι. Κατά την υλοποίηση του προτεινόμενου συστήματος πλοήγησης λήφθηκαν υπόψιν οι εξής παρατηρήσεις:

- Ο εντοπισμός της θέσης του χρήστη είναι πιο εύκολος λόγω της διαθεσιμότητας του GPS. Αντίθετα, σε έναν εσωτερικό χώρο η χρήση του GPS είναι προβληματική λόγω χαμηλής στάθμης του σήματος.
- Σε ένα αστικό περιβάλλον, όπως μεγάλες πόλεις με ψηλά κτήρια, υπάρχει πιθανότητα να έχουμε ελάττωση της στάθμης σήματος του GPS, λόγω της παρεμβολής των κτιρίων. Αυτό οδηγεί συνήθως σε μειωμένη ακρίβεια του GPS, κάτι που είναι κρίσιμο για την πλοήγηση ενός ατόμου με μειωμένη όραση. Συχνά, αυτό το πρόβλημα εξαλείφεται με την κατασκευή όλο και καλύτερων δεκτών GPS, ή αλγορίθμικά με τη χρήση τεχνικών εντοπισμού θέσης μέσω της κάμερας που φέρει ο χρήστης. Στην παρούσα εργασία δεν θα ασχοληθούμε με την επίλυση του συγκεκριμένου ζητήματος, καθώς είναι κάτι που ξεφεύγει από τους στόχους.
- Κατά την πλοήγησή του ο χρήστης θα βρεθεί, συχνά, αντιμέτωπος με διάφορα φυσικά αντικείμενα, που λειτουργούν ως εμπόδια στην διαδρομή του. Επίσης, η ελλιπής συντήρηση των υποδομών, όπως η ύπαρξη σπασμένων πλακιδίων ή φυτών στα πεζοδρόμια και η έλλειψη ανάγλυφων οδηγών για τους τυφλούς, δυσχεραίνει την μετακίνησή των ατόμων αυτών.
- Τέλος, τα συστήματα πλοήγησης που βασίζονται σε κάμερες ή αισθητήρες υπερύθρων, επηρεάζονται από τις ακτίνες του ήλιου, οι οποίες δημιουργούν παρεμβολές και περιορίζουν την εφαρμοσιμότητα των αλγορίθμων.

3.3.4 Παράλληλη χρήση με το λευκό μπαστούνι

Αν και η τεχνολογία έχει πραγματοποιήσει τεράστια άλματα προόδου τις τελευταίες δεκαετίες, είναι γεγονός πως δεν έχει σχεδιαστεί κάποια εφαρμογή πλοήγησης για τυφλούς που να ικανοποιεί τις ανάγκες τους πλήρως και να αντικαταστήσει τις παραδοσιακές μεθόδους υποβοήθησης πλοήγησης [44]. Συνεπώς, η κοινότητα των ατόμων με προβλήματα όρασης εξακολουθεί δικαίως να χρησιμοποιεί το λευκό μπαστούνι, ως το κύριο βοήθημα μετακίνησης, κάτι το οποίο δεν μπορεί να αλλάξει από τη μια στιγμή στην άλλη. Για αυτόν τον λόγο, το σύστημα που παρουσιάζεται στην παρούσα εργασία προορίζεται να λειτουργεί υποστηρικτικά στη χρήση του λευκού μπαστούνιού, βελτιώνοντας ταυτόχρονα την εμπειρία μετακίνησης των χρηστών και επιτρέποντάς τους την χρήση των εργαλείων που εμπιστεύονται και με τα οποία έχουν εξοικειωθεί.

3.4 Απτική Ανάδραση

3.4.1 Αισθητηριακή υποκατάσταση

Η έννοια της αισθητηριακής υποκατάστασης (sensory substitution) αναφέρεται στην τροποποίηση ορισμένων χαρακτηριστικών ενός αντιληπτικού συστήματος (π.χ. όραση, ακοή, αφή, όσφρηση,

γεύση) και η μετατροπή τους σε ερεθίσματα για ένα διαφορετικό αντιληπτικό σύστημα [45]. Με άλλα λόγια, η κεντρική ιδέα της αισθητηριακής υποκατάστασης είναι ότι οι πληροφορίες, οι οποίες δεν είναι διαθέσιμες σε κάποιον εξαιτίας μιας αισθητηριακής διαταραχής, μπορούν να γίνουν αντιληπτές μέσω ενός άλλου αντιληπτικού συστήματος. Η μετατροπή, ωστόσο, των ερεθισμάτων από ένα αντιληπτικό σύστημα σε ένα άλλο είναι μια πολύπλοκη διαδικασία με πολλές παραμέτρους, η οποία αποτελεί αντικείμενο έρευνας τόσο στον ιατρικό όσο και στον τεχνολογικό τομέα. Στα πλαίσια της διπλωματικής εργασίας θα αξιοποιήσουμε πολύ βασικές γνώσεις που αφορούν την αντικατάσταση μιας αίσθησης με μια άλλη, χωρίς να εμβαθύνουμε περαιτέρω.

3.4.2 Υπάρχοντα κανάλια μετάδοσης χωρικής πληροφορίας

Οι άνθρωποι με φυσιολογική όραση χρησιμοποιούν την όραση για να προσδιορίσουν την θέση τους και το μονοπάτι στο οποίο θα κινηθούν. Τα άτομα με προβλήματα όρασης έχουν αναπτύξει τις υπόλοιπες αισθήσεις τους ώστε να μπορούν να αντιλαμβάνονται τον χώρο στον οποίο βρίσκονται. Όπως είναι φυσιολογικό τα δύο αισθητήρια κανάλια που αξιοποιούνται περισσότερο είναι το ακουστικό και το απτικό, το οποίο έχει επιβεβαιωθεί κι από αντίστοιχες έρευνες [46]. Χρησιμοποιώντας το ακουστικό κανάλι τα άτομα με προβλήματα όρασης είναι σε θέση να αντιλαμβάνονται τις αλλαγές που συμβαίνουν γύρω τους και να προσανατολίζονται, χρησιμοποιώντας διάφορες πηγές ήχου ως σημεία αναφοράς, π.χ. ο θόρυβος που κάνουν τα οχήματα τους επιτρέπει να γνωρίζουν πού βρίσκεται ο δρόμος, ενώ η ηχητική ειδοποίηση των φαναριών τους βοηθάει να καταλάβουν αν είναι πράσινο ή κόκκινο. Αντίθετα, η χωρική πληροφορία είναι δύσκολο να μεταφερθεί άμεσα μέσω του απτικού καναλιού. Η χρησιμότητά του είναι ότι μέσω της αφής ο χρήστης αντιλαμβάνεται περισσότερες λεπτομέρειες σχετικά με την υφή ή το είδος ενός φυσικού αντικειμένου. Για παράδειγμα, όταν ο τυφλός χρησιμοποιεί το λευκό μπαστούνι και αυτό πέσει πάνω σε ένα εμπόδιο, τότε, αναλόγως την δύναμη που ασκείται στον μπαστούνι, μπορεί να αντιληφθεί αν το αντικείμενο αυτό είναι κάτι σκληρό και αμετακίνητο (π.χ. μια κολόνα δημοτικού φωτισμού), ή κάτι πιο εύκαμπτο και μετακινήσιμο (π.χ. μια καρέκλα).

3.4.3 Γιατί απτική ανάδραση;

Στην παρούσα διπλωματική εργασία υλοποιείται ένα σύστημα υποβοήθησης πλοήγησης που βασίζεται στην απτική αλληλεπίδραση με τον χρήστη. Πιο συγκεκριμένα, υλοποιήθηκε ένα σύστημα που ειδοποιεί τον χρήστη μέσω δονήσεων για το πότε και πού πρέπει να κινηθεί, καθώς επίσης και όταν απαιτείται η διάσχιση ενός δρόμου. Η επιλογή του απτικού καναλιού ως μέσο μετάδοσης της χωρικής πληροφορίας έγινε για τους εξής δύο λόγους:

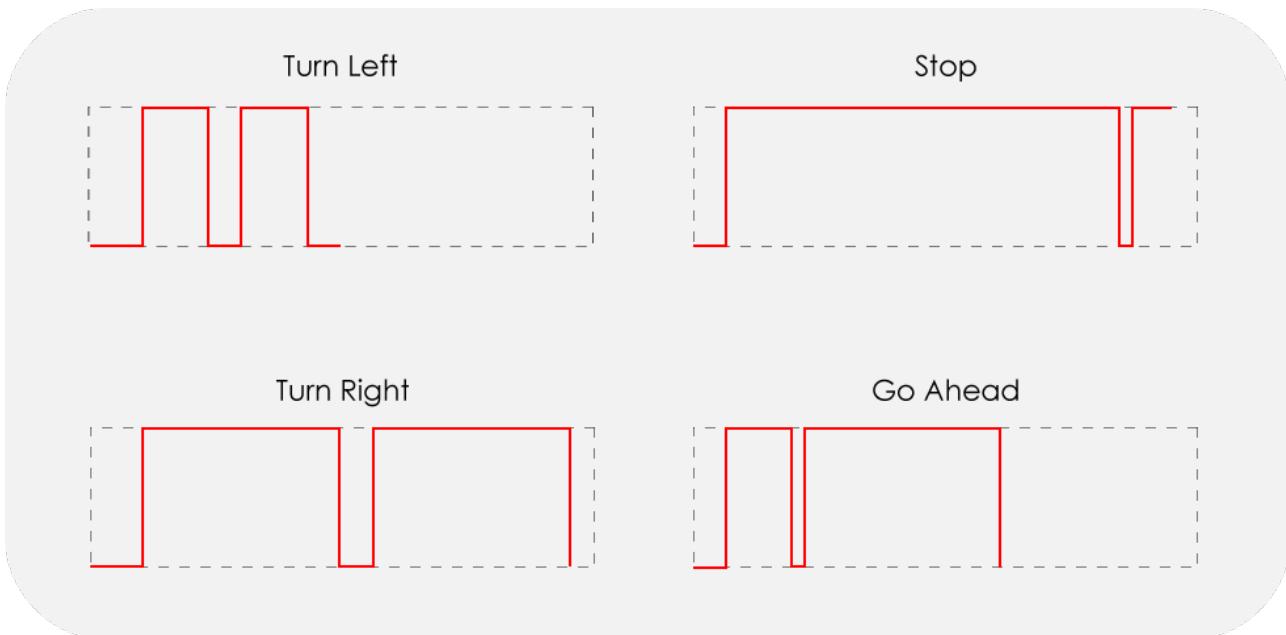
1. Το ακουστικό κανάλι είναι ήδη αρκετά επιβαρυμένο με άλλες λειτουργίες [47, 48] και απαιτείται μια αποσυμφόρηση, ώστε το άτομο με προβλήματα όρασης να μπορεί να αντιλαμβάνεται άλλες κρίσιμες πληροφορίες μέσω της ακοής, π.χ. τον ήχο από τα αυτοκίνητα ή κάποιον περαστικό. Επίσης, η παροχή ανάδρασης μέσω ήχου θα μπορούσε να συγχύσει και να κουράσει πνευματικά τον χρήστη, λόγω πολλών παράλληλων ακουστικών ερεθισμάτων.
2. Η χρήση δονήσεων αποτελεί έναν πιο διακριτικό τρόπο να ενημερώνεται ο χρήστης για πληροφορίες που αφορούν το εξωτερικό περιβάλλον. Η διακριτικότητα και η ευκολία είναι από τα βασικά χαρακτηριστικά του προτεινόμενου συστήματος και η αξιοποίηση του απτικού καναλιού είναι απαραίτητη για να επιτευχθεί κάτι τέτοιο.

3.4.4 Haptic icons

Ως haptic icons ορίζονται στην βιβλιογραφία τα διάφορα απτά ερεθίσματα στα οποία έχουμε αποδώσει συγκεκριμένα νοήματα (π.χ. ένα απτικό ερεθίσμα μπορεί να είναι "συνδεδεμένο" με την

έκφραση χαράς, λύπης κλπ.) [49]. Σχετικές έρευνες [50, 51] έχουν μελετήσει τη σχέση μεταξύ των απτικών ερεθισμάτων που δέχεται ένας άνθρωπος και των συναισθημάτων που του προκαλούν. Συνήθως συστήματα που βασίζονται στην αξιοποίηση του απτικού καναλιού στοχεύουν στην μετάδοση πληροφοριών στον χρήστη μέσω της χρήσης διαφορετικών μοτίβων δονήσεων, τα οποία είναι συνδεδεμένα με διαφορετικές εντολές πλοϊγησης. Τα διαφορετικά μοτίβα δονήσεων μπορεί να περιλαμβάνουν διαφορές στην ένταση της δόνησης, στην αλληλουχία των δονήσεων και στην συχνότητά τους. Για τις ανάγκες της συγκεκριμένης διπλωματικής εργασίας χρησιμοποιήθηκε η λογική της αντιστοιχησης ενός μοτίβου δονήσεων, τα οποία παίζουν τον ρόλο των haptic icons, για κάθε διαφορετική εντολή που απαιτείται από τον χρήστη. Κάθε μοτίβο έχει το ίδιο επίπεδο έντασης, καθώς θεωρήθηκε ότι ο αριθμός των εντολών που έπρεπε να κωδικοποιηθούν δεν ήταν τόσο μεγάλος, ώστε να αξιοποιηθεί και η ένταση των δονήσεων για περισσότερη διακριτοποίηση. Πιο συγκεκριμένα:

- Εντολή **"Στρίψε Αριστερά"**: Η εντολή αυτή ειδοποιεί τον χρήστη ότι πρέπει να πραγματοποιήσει στροφή αριστερά 90°. Το μοτίβο δονήσεων που αντιστοιχεί σε αυτήν την εντολή είναι "σύντομη δόνηση - μικρή παύση - σύντομη δόνηση".
- Εντολή **"Στρίψε Δεξιά"**: Η εντολή αυτή ειδοποιεί τον χρήστη ότι πρέπει να πραγματοποιήσει στροφή δεξιά 90°. Το μοτίβο δονήσεων που αντιστοιχεί σε αυτήν την εντολή είναι "μακρά δόνηση - μικρή παύση - μακρά δόνηση".
- Εντολή **"Σταμάτα"**: Η εντολή αυτή ειδοποιεί τον χρήστη ότι πρέπει να σταματήσει μπροστά από μια διάβαση, επειδή το ο φωτεινός σηματοδότης είναι κόκκινος. Το μοτίβο δονήσεων που αντιστοιχεί σε αυτήν την εντολή είναι "πολύ μακρά δόνηση - πολύ σύντομη παύση".
- Εντολή **"Προχώρα"**: Η εντολή αυτή ειδοποιεί τον χρήστη ότι μπορεί να διασχίσει μια διάβαση, επειδή το ο φωτεινός σηματοδότης είναι πράσινος. Το μοτίβο δονήσεων που αντιστοιχεί σε αυτήν την εντολή είναι "σύντομη δόνηση - πολύ σύντομη παύση - μακρά δόνηση - πολύ σύντομη παύση".

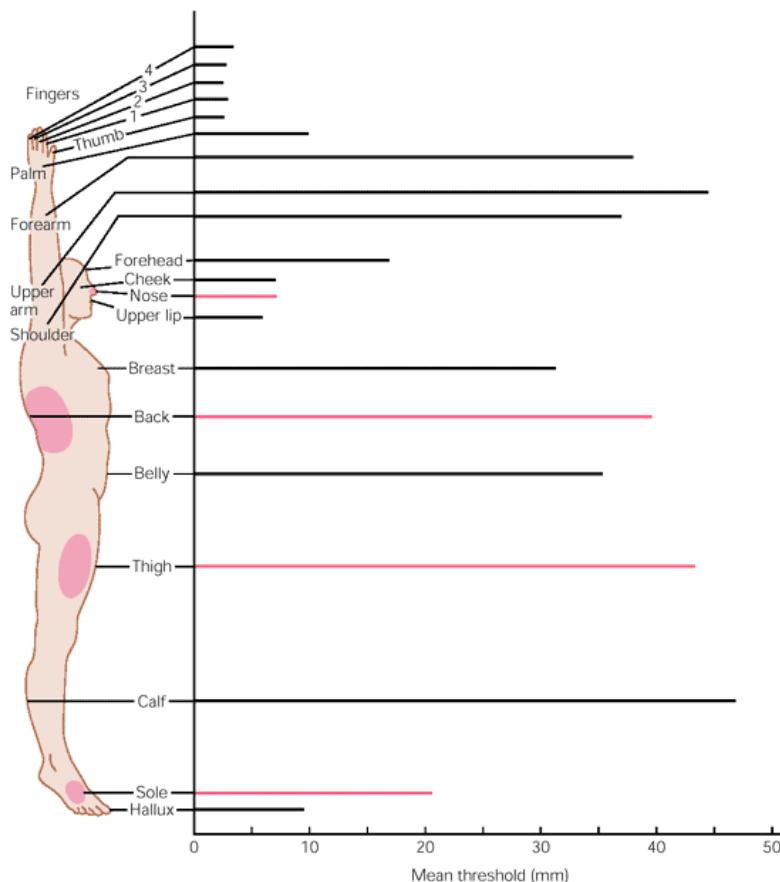


Σχήμα 3.2: Αντιστοιχηση μοτίβων δονήσεων σε εντολές πλοϊγησης

3.4.5 Δυνατότητες και περιορισμοί

Όπως έχουμε ήδη αναφέρει η χρήση της αφής ως μέσο μετάδοσης πληροφορίας έχει πολύ μεγάλο πλεονέκτημα την διακριτικότητα και την αμεσότητα στην αντίληψη της πληροφορίας. Ωστόσο, συ-

γκριτικά με την ακοή, προβάλει ένα σημαντικό μειονέκτημα, το οποίο είναι η χαμηλή διακριτότητα της πληροφορίας που μπορεί να μεταδοθεί [52]. Με απλά λόγια, αν θέλουμε να κωδικοποιήσουμε πολλές διαφορετικές εντολές με χρήση μοτίβων δονήσεων, τότε θα πρέπει να δημιουργήσουμε ένα διαφορετικό μοτίβο για κάθε ξεχωριστή εντολή. Ωστόσο, η δυνατότητα του ανθρώπου να αντιλαμβάνεται διαφορετικά μοτίβα δονήσεων με μικρές διαφορές μεταξύ τους μειώνεται όσο αυξάνεται η πολυπλοκότητα και ο αριθμός αυτών των μοτίβων. Έρευνες έχουν δείξει επίσης, ότι η διακριτότητα της πληροφορίας που μπορεί να αντιληφθεί ένα άνθρωπος μέσω της αφής εξαρτάται από το σημείο του δέρματος που δέχεται τη διέγερση, για παράδειγμα τα δάκτυλα έχουν πολύ μεγαλύτερη δυνατότητα διάκρισης σε σχέση με τον αγκώνα. Το κατώφλι διάκρισης μεταξύ δυο σημείων (two-point discrimination threshold (TPDT)) ορίζεται ως το μέτρο το οποίο αναπαριστά πόσο πρέπει να απέχουν δυο σημεία πίεσης ώστε να θεωρηθούν από το δέρμα ως διακριτά [53].



Σχήμα 3.3: Μεταβολή του TPDT σε συνάρτηση με τα σημεία του σώματος [4]

3.5 Εξοπλισμός - Υλικό (Hardware)

Στην ενότητα αυτή θα γίνει μια πλήρης περιγραφή του εξοπλισμού που χρησιμοποιήθηκε στα πλαίσια της παρούσας διπλωματικής εργασίας.

3.5.1 Μονάδα «αίσθησης» του περιβάλλοντος – Camera/Depth Sensor

3.5.1.1 Διαθέσιμοι αισθητήρες

Υπάρχουν πολλοί διαθέσιμοι αισθητήρες-κάμερες που θα μπορούσαν να αξιοποιηθούν στα πλαίσια ενός συστήματος πλοϊγησης για άτομα με προβλήματα όρασης. Η επιλογή του κατάλληλου αισθητήρα εξαρτάται κυρίως από τις ανάγκες τις εκάστοτε εφαρμογής, π.χ. αν πρόκειται για εξωτερικό ή εσωτερικό περιβάλλον κλπ. Στην παρούσα εργασία χρησιμοποιήσαμε την κάμερα *Intel Realsense D435i*, ωστόσο πριν προχωρήσουμε στην τελική επιλογή κάμερας εξετάσαμε τις εξής εναλλακτικές επιλογές:

- **ZED Stereo Camera:** Πρόκειται για μια κάμερα που λειτουργεί με την μέθοδο της στερεοσκοπικής όρασης, δηλαδή χρησιμοποιεί δύο διαφορετικούς φακούς με οριζόντια απόσταση ο ένας από τον άλλο, και το βάθος/απόσταση υπολογίζεται συγκρίνοντας την μετατόπιση των δύο διαφορετικών εικόνων.

Technology:	Stereo Depth Sensing
Field of View (FoV):	Max. 90°(H) x 60°(V) x 100°(D)
RGB Sensor Type:	1/3" 4MP CMOS
Depth Range:	0.5m to 25m
Depth FPS:	Up to 100Hz
Depth Accuracy:	<= 2% up to 3m, <= 4% up to 15m
Dimensions:	175 x 30 x 33 mm
Power:	380mA / 5V USB Powered
SDK Provided:	YES

Πίνακας 3.2: Χαρακτηριστικά ZED Stereo Camera [9]



Σχήμα 3.4: ZED Stereo Camera

- **CamBoard pico flexx depth sensor:** Ο αισθητήρας αυτός είναι από τους πιο μικρούς που υπάρχουν και χρησιμοποιεί την τεχνολογία Time-of-Flight (ToF), με εκπομπή υπέρυθρης ακτινοβολίας, για να υπολογίσει την απόσταση των αντικειμένων, χωρίς ωστόσο να παρέχει δυνατότητα RGB εικόνας.

Technology:	Time-of-Flight (ToF)
Field of View (FoV):	Max. 62°(H) x 45°(V)
RGB Sensor Type:	Δεν υποστηρίζεται
Depth Range:	0.1m to 4m
Depth FPS:	Up to 45fps
Depth Accuracy:	<= 1% of distance (0.5 – 4m @ 5fps), <= 2% of distance (0.1 – 1m @ 45fps)
Dimensions:	68 x 17 x 7.35 mm
Power:	300mW / USB2.0 compliant
SDK Provided:	YES

Πίνακας 3.3: Χαρακτηριστικά CamBoard pico flexx depth sensor [10]



Σχήμα 3.5: CamBoard pico flexx depth sensor

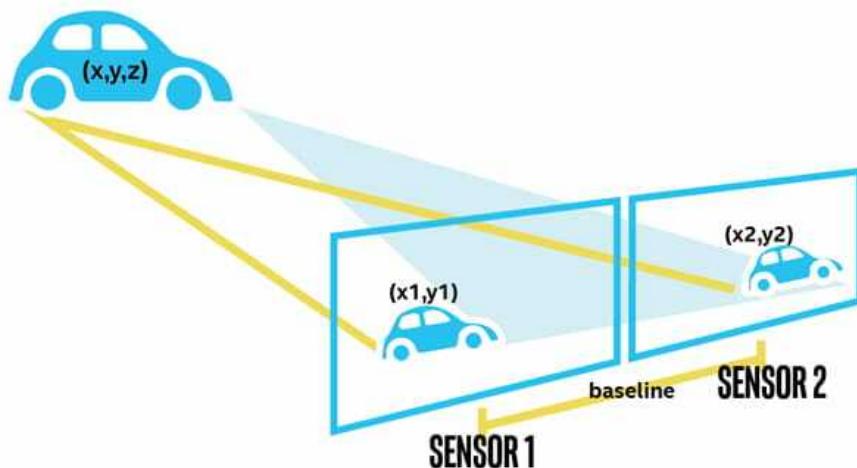
3.5.1.2 Κριτήρια επιλογής κατάλληλου αισθητήρα

Η επιλογή του κατάλληλου αισθητήρα έγινε λαμβάνοντας υπόψιν τα παρακάτω κριτήρια:

1. **Διαστάσεις κάμερας:** Στόχος είναι η φορητότητα και η διακριτικότητα του συστήματος, άρα θέλουμε έναν όσο γίνεται μικρότερο σε διαστάσεις αισθητήρα.
2. **Ταυτόχρονη υποστήριξη RGB & Depth Frames:** Η προτεινόμενη εφαρμογή αξιοποιεί τόσο την έγχρωμη εικόνα RGB για την αναγνώριση της διάβασης πεζών και των φωτεινών σηματοδοτών, όσο και τον χάρτη βάθους (depth map) για τον εντοπισμό φυσικών εμποδίων.
3. **Ανθεκτικότητα στις συνθήκες φωτισμού:** Λόγω της φύσης του συστήματος, είναι απαραίτητο η κάμερα να αποδίδει επαρκώς καλά τόσο σε εσωτερικά, όσο και σε εξωτερικά περιβάλλοντα. Ιδιαίτερα η ακρίβεια των μετρήσεων της κάμερας σε εξωτερικό περιβάλλον επηρεάζεται από τον φωτισμό του περιβάλλοντος και τον προσανατολισμό της σε σχέση με τις ακτίνες του ήλιου.
4. **Επαρκή υποστήριξη με SDK (Software Development Kit):** Κατά την υλοποίηση μιας εργασίας είναι πολύ σημαντικό να υπάρχουν διαθέσιμα παραδείγματα κώδικα και έτοιμες συναρτήσεις που διευκολύνουν την ανάπτυξη των αλγορίθμων. Τα SDKs είναι πακέτα λογισμικού, που περιέχουν βελτιστοποιημένες συναρτήσεις και βασικά παραδείγματα για την ορθή χρήση του εκάστοτε αισθητήρα. Η ύπαρξη μιας ευρύτερης κοινότητας γύρω από έναν συγκεκριμένο αισθητήρα συμβάλλει στην πιο γρήγορη και αποτελεσματική υλοποίηση, καθώς και στην πιο εύκολη επίλυση των σφαλμάτων.
5. **Ακρίβεια και εύρος μέτρησης απόστασης:** Η ακρίβεια κατά την μέτρηση των αποστάσεων, καθώς και το εύρος της μέτρησης (ελάχιστη/μέγιστη μετρήσιμη απόσταση) είναι κρίσιμα κριτήρια για την επιλογή του κατάλληλου αισθητήρα βάθους, ο οποίος θα καλύπτει τις ανάγκες ενός πεζού χρήστη με προβλήματα όρασης.

3.5.1.3 Intel RealSense D435i

Στα πλαίσια της παρούσας διπλωματικής εργασίας χρησιμοποιήθηκε η κάμερα *Intel Realsense D435i*, η οποία περιέχει έναν RGB φακό και έναν ενσωματωμένο αισθητήρα βάθους, ο οποίος λειτουργεί με στερεοσκοπική όραση. Η αρχή λειτουργίας της στερεοσκοπικής όρασης είναι παρόμοια με τον τρόπο που αντιλαμβάνεται το βάθος ο ανθρώπινος εγκέφαλος μέσω των ματιών, δηλαδή υπάρχουν δύο φακοί-αισθητήρες τοποθετημένοι σε μικρή απόσταση μεταξύ τους και λαμβάνουν δύο διαφορετικές εικόνες ενός αντικειμένου. Συγκρίνοντας τις δύο εικόνες και γνωρίζοντας την απόσταση μεταξύ των δύο φακών εξάγεται η πληροφορία που αφορά το βάθος (Σχήμα 3.6). Οι στερεοσκοπικές κάμερες λειτουργούν αρκετά καλά κάτω από οποιεσδήποτε συνθήκες φωτισμού, αφού το βάθος υπολογίζεται αποκλειστικά από την σύγκριση εικόνων. Η κάμερα *Intel Realsense D435i* χρησιμοποιεί επιπλέον έναν πομπό υπέρυθρης ακτινοβολίας (infrared projector), ώστε έχει μεγαλύτερη ακρίβεια μετρήσεων σε περιβάλλοντα χαμηλού φωτισμού. Τέλος, ενσωματώνει κι έναν IMU (Inertial Measurement Sensor) sensor, έναν αισθητήρα που περιλαμβάνει επιταχυνσιόμετρο και γυροσκόπιο, για να μετράει την περιστροφή και τον προσανατολισμό της κάμερας με 6 βαθμούς ελευθερίας.



Σχήμα 3.6: Αρχή λειτουργίας στερεοσκοπικής όρασης

Technology:	Active IR Stereo
Field of View (FoV):	Max. 90°(H) x 59°(V) x 98°(D)
RGB Sensor:	1920 x 1080, 30fps
Depth Range:	0.105m to 10m
Depth FPS:	Up to 90fps
Dimensions:	90 x 25 x 25 mm
Power:	USB3.0
SDK Provided:	YES

Πίνακας 3.4: Χαρακτηριστικά *Intel Realsense D435i* [11]



Σχήμα 3.7: Intel Realsense Depth Camera D435i

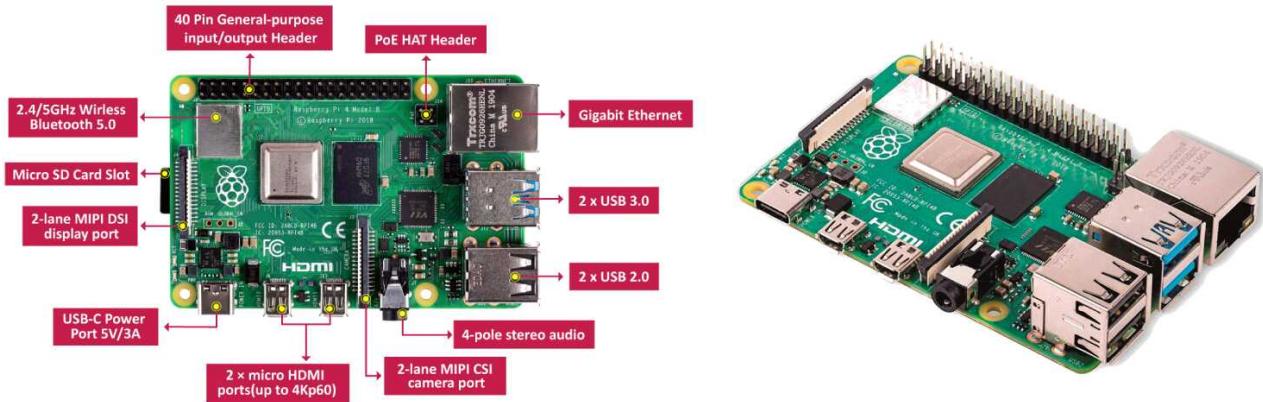


Σχήμα 3.8: Intel Realsense D435i - Εσωτερική δομή

3.5.2 Μονάδα επεξεργασίας εικόνας – Μικροϋπολογιστής

Οι αλγόριθμοι εντοπισμού και αναγνώρισης αντικειμένων χρησιμοποιούν τεχνικές επεξεργασίας εικόνας για να καταφέρουν να αποκωδικοποιήσουν μια φωτογραφία ή ένα βίντεο. Η επεξεργασία εικόνας και βίντεο είναι μια απαιτητική εργασία τόσο σε μνήμη όσο και σε επεξεργαστική ισχύ, επομένως είναι απαραίτητη η ύπαρξη ενός υπολογιστή ή κάποιας μονάδας επεξεργασίας που θα μπορεί να αντεπεξέλθει στις απαιτήσεις αυτές.

3.5.2.1 Raspberry Pi 4



Σχήμα 3.9: Raspberry Pi 4 Model B

Για τις ανάγκες της διπλωματικής εργασίας επιλέχθηκε η χρήση του μικροεπεξεργαστή *Raspberry Pi 4* (RPi4). Το RPi4 είναι ένας πλήρης υπολογιστής σε μέγεθος πιστωτικής κάρτας, στον οποίο μπορούν να συνδεθούν διάφορες περιφερειακές συσκευές, όπως πληκτρολόγιο, οθόνη κλπ. Πιο συγκεκριμένα, χρησιμοποιήθηκε η έκδοση Model B του Raspberry Pi 4, το οποίο έχει βελτιωμένα χαρακτηριστικά, όπως υποστήριξη USB3.0, μεγαλύτερη μνήμη RAM (4GB) και καλύτερο επεξεργαστή γραφικών. Στον πίνακα 3.5 παρουσιάζονται αναλυτικά τα χαρακτηριστικά του Raspberry Pi 4 που χρησιμοποιήσαμε.

Processor:	Broadcom BCM2711, quad-core Cortex-A72 (ARM v8) 64-bit SoC @ 1.5GHz
Memory:	4GB LPDDR4
Connectivity:	WLAN, Ethernet, Bluetooth 5.0, BLE
USB2.0:	2 ports
USB3.0:	2 ports
HDMI:	2 micro HDMI ports
Micro SD:	YES
Input power:	5V DC via USB-C connector (minimum 3A)

Πίνακας 3.5: Χαρακτηριστικά Raspberry Pi 4 Model B [12]

Ουσιαστικά, στα πλαίσια της παρούσας εργασίας αξιοποιήθηκαν κυρίως η μεγάλη χωρητικότητα μνήμης RAM και η ύπαρξη θύρας USB3.0, μιας και η κάμερα που επιλέξαμε υποστήριζε αποκλειστικά USB3.0 interface. Ταυτόχρονα, η μεγάλη επεξεργαστική ισχύς συνέβαλε στο να τρέχει ο αλγόριθμος σε real-time, ενώ η WiFi συνδεσιμότητα επέτρεψε την σύνδεση με τα άλλα μέρη του συστήματος, όπως θα δούμε παρακάτω. Η τροφοδοσία του RPi4 μπορεί να γίνει με δύο τρόπους: είτε με μια συστοιχία από μπαταρίες ίοντων λιθίου, είτε μέσω ενός powerbank (φορητή μπαταρία). Επιλέχθηκε η δεύτερη λύση του powerbank, καθώς ήταν πολύ πιο εύκολο να το αποκτήσουμε και δεν απαιτεί ιδιαίτερους χειρισμούς όσον αφορά την επαναφόρτισή του. Το RPi4 απαιτεί τροφοδοσία 5V/3A, επομένως χρησιμοποιήθηκε ένα powerbank που έχει ως έξοδο 5V/3A και χωρητικότητα μπαταρίας 10.000mAh (Σχήμα 3.10). Η αυτονομία του συστήματος με την χρήση της συγκεκριμένης φορητής μπαταρίας φτάνει έως και την 1.5 ώρα.



Σχήμα 3.10: Το powerbank *Intenso PD10000* που χρησιμοποιήθηκε

3.5.3 Συσκευή πλοϊγησης – Android Smartphone

Η επιλογή του κινητού του χρήστη ως κύρια συσκευή πλοϊγησης έγινε, διότι πρόκειται για μια συσκευή με την οποία οι χρήστες είναι αρκετά εξοικειωμένοι, είναι άμεσα διαθέσιμη σχεδόν σε όλους και είναι πολύ εύκολο να αναπτυχθούν εφαρμογές για smartphone.

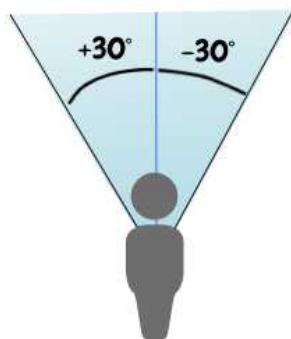
3.5.3.1 Εφαρμογή πλοϊγησης

Στα πλαίσια της διπλωματικής αναπτύχθηκε ειδική εφαρμογή για κινητά σε λειτουργικό Android. Η εφαρμογή προορίζεται να εγκαθίσταται στο κινητό του χρήστη και αποτελεί την μοναδική διεπιφάνεια αλληλεπίδρασης του χρήστη με το σύστημα πλοϊγησης. Το κομμάτι της υλοποίησης της εφαρμογής περιλαμβάνει την ανάπτυξη κώδικα που αξιοποιεί τους ενσωματωμένους αισθητήρες GPS, IMU και πυξίδα για να μπορεί να εντοπίζεται η τοποθεσία του χρήστη.

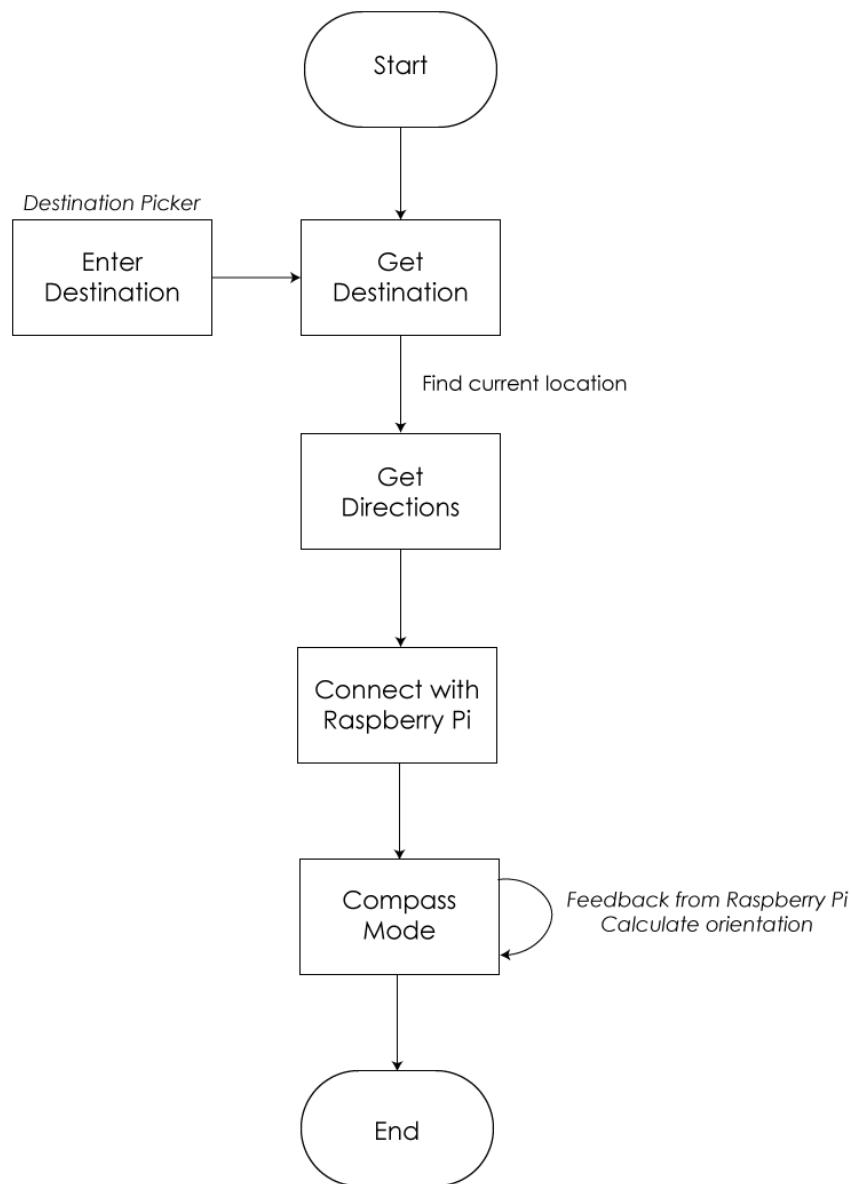
Αρχή λειτουργίας:

Η εφαρμογή ενσωματώνει έναν χάρτη της ευρύτερης περιοχής χάρη στη χρήση του *Google Maps SDK* <https://developers.google.com/maps/documentation/android-sdk/intro>. Ο χρήστης εισάγει έναν προορισμό στην εφαρμογή και η αντίστοιχη τοποθεσία βρίσκεται μέσω του *Google Places API* (<https://developers.google.com/places/web-service/intro>). Αφού ο χρήστης εισάγει έναν προορισμό, το σύστημα βρίσκει την βέλτιστη διαδρομή, χρησιμοποιώντας το *Google Directions API* (<https://developers.google.com/maps/documentation/directions/intro>) και επιστρέφονται ως δεδομένα οι οδηγίες κατεύθυνσης για πεζούς. Το format των οδηγιών είναι JSON, οπότε αναπτύχθηκε μια εσωτερική συνάρτηση στην εφαρμογή που αποκωδικοποιεί τα δεδομένα κατεύθυνσης και τα εξάγει ως μια λίστα από απλές εντολές στροφής (turn-by-turn navigation). Καθόσον η εφαρμογή βρίσκεται σε λειτουργία πυξίδας, επικοινωνεί με το Raspberry Pi ασύρματα και λαμβάνει εντολές για το αν έχει βρεθεί κάποιο εμπόδιο ή έχει εντοπιστεί διάβαση πεζών. Η εφαρμογή αντιλαμβάνεται τον προσανατολισμό του χρήστη εκμεταλλευόμενη τους εσωτερικούς αισθητήρες του κινητού και υπολογίζει την γωνία απόκλισης που έχει σε σχέση με τον προσανατολισμό του μονοπατιού που ακολουθεί. Αν η απόκλιση είναι μεγαλύτερη των 30°(σχήμα 3.11), τότε τον ειδοποιεί

κατάλληλα (με χρήση δονήσεων για στροφή αριστερά ή δεξιά).



Σχήμα 3.11: Η κατεύθυνση του χρήστη πρέπει να είναι μέσα σε ένα τόξο 60° για να θεωρείται έγκυρη. Άλλιως παρέχεται κατάλληλη ανάδραση για αριστερή ή δεξιά στροφή.



Σχήμα 3.12: Διάγραμμα ροής της Android εφαρμογής

3.5.3.2 Προδιαγραφές εφαρμογής – Ελάχιστες απαιτήσεις

Κατά την ανάπτυξη της εφαρμογής προέκυψαν διάφορες ανάγκες όσον αφορά την χρήση της από άτομα με προβλήματα όρασης. Βασικός μας στόχος ήταν να είναι όσο πιο απλή γίνεται με έμφαση στη λειτουργικότητα. Μια τέτοια εφαρμογή πρέπει:

1. Να χαρακτηρίζεται από την απλότητα στη διεπιφάνεια χρήστη
2. Να είναι διαισθητική, δηλαδή να είναι αυτονόητο αυτό που προσφέρει μην απαιτείται πολύς χρόνος εκπαίδευσης σε νέους χρήστες
3. Να χρησιμοποιεί τεχνολογία πρόσβασης για άτομα με μειωμένη όραση, όπως χρήση δονήσεων, μετατροπή λόγου σε κείμενο και αντίστροφα.

Από τις παραπάνω απαιτήσεις, στα πλαίσια της παρούσας εργασίας καταφέραμε να υλοποιήσουμε το 1, 2 και, εν μέρει, το 3. Λόγω περιορισμένου χρόνου δεν υλοποιήθηκε η μετατροπή από προφορικό λόγο σε κείμενο.

3.5.3.3 UX-UI Design

Παρακάτω παρουσιάζονται μερικά στιγμιότυπα από τα διάφορα screens της εφαρμογής.



Σχήμα 3.13: Διεπιφάνεια επαφής χρήστη

Στο σχήμα 3.13 παρουσιάζονται με κόκκινο πλαίσιο οι βασικοί τρόποι με τους οποίους αλληλεπιδρά ο χρήστης με το σύστημα. Η βασική ιδέα είναι ότι ο χρήστης, αφού εισάγει τον προορισμό του, θα τοποθετήσει το κινητό στη τσέπη του (ή θα το κρατάει στο χέρι του) και αυτό θα τον καθιδηγεί καθ' όλη την διάρκεια της πλοιήγησης, μειώνοντας στο ελάχιστο την αλληλεπίδραση του χρήστη με

το σύστημα. Η εφαρμογή εντοπίζει αυτόματα τον προσανατολισμό του χρήστη με τη βοήθεια των αισθητήρων επιταχυνσιομέτρου, γυροσκοπίου και πυξίδας. Είναι γεγονός ότι ο χρήστης που έχει προβλήματα όρασης θα εισάγει τον προορισμό του μέσω ενός συστήματος αναγνώρισης και μετατροπής ομιλίας σε κείμενο. Δυστυχώς, λόγω έλλειψης χρόνου δεν υλοποιήθηκε το συγκεκριμένο χαρακτηριστικό στα πλαίσια της παρούσας διπλωματικής εργασίας.

3.5.4 Συσκευή απτικής ανατροφοδότησης – Haptic Device

3.5.4.1 Υπάρχουσες συσκευές απτικής ανατροφοδότησης

Ως συσκευές απτικής ανατροφοδότησης μπορούν να χαρακτηριστούν όλες εκείνες οι συσκευές που παρέχουν κάποιου είδους πίεση στο δέρμα του χρήστη, π.χ. μέσω δόνησης. Την τελευταία δεκαετία έχουν εμφανιστεί πολλές τέτοιες συσκευές που είναι πολλά υποσχόμενες για την αλληλεπίδραση ανθρώπου-μηχανής. Μια πολύ ενδιαφέρουσα εναλλακτική που εξετάστηκε στα πλαίσια της διπλωματικής είναι το βραχιόλι Myo Armband [54], το οποίο τοποθετείται κάτω από τον αγκώνα και αποτελείται από πολλαπλούς αισθητήρες που διαβάζουν ηλεκτρομυογραφικά σήματα τα οποία ερμηνεύονται ως συγκεκριμένες κινήσεις. Αν και το Myo Armband παρέχει τη δυνατότητα δόνησης στο χέρι του χρήστη, δεν επιλέχθηκε τελικά λόγω του ότι η εταιρεία που το είχε έκλεισε και δεν υπήρχε η απαραίτητη υποστήριξη από την κοινότητα. Επιπλέον, η χρησιμοποίηση της συγκεκριμένης συσκευής θα πρόσθετε ακόμα ένα βάρος στον χρήστη, κάτι που εναντιώνεται στο κριτήριο της απλότητας και της παρεμβατικότητας που αναφέραμε νωρίτερα.



Σχήμα 3.14: Myo Armband

3.5.4.2 Smartphone

Μετά από εξέταση όλων των εναλλακτικών, καταλήξαμε στην επιλογή του smartphone ως το μέσο διάδοσης της απτικής ανάδρασης στο χρήστη. Αρχικά, η επιλογή αυτή μας εξυπηρετεί επειδή

δεν απαιτείται η ένταξη μιας άλλης συσκευής στο σύστημα, η οποία πιθανόν να είχε προβλήματα συμβατότητας με τα υπόλοιπα μέρη. Επίσης, ο χρήστης δεν επιβαρύνεται με περιττά αντικείμενα, που θα δυσκόλευαν την εξωτερική του μετακίνηση και την εικόνα του προς τον υπόλοιπο κόσμο. Τέλος, το smartphone συνδέεται με το υπόλοιπο σύστημα και λαμβάνει κατάλληλα σήματα σχετικά με τον εντοπισμό και την αναγνώριση διαβάσεων και φαναριών. Ανάλογα το είδος του σήματος που φτάνει στο smartphone, αυτό με τη σειρά του παρέχει το κατάλληλο μοτίβο δόνησης στον χρήστη. Στα πλαίσια της παρούσας διπλωματικής εργασίας χρησιμοποιήθηκε το μοντέλο *Redmi Note 4X*.

3.6 Αλγόριθμοι πλοιόγησης & Λογισμικό (Software)

Στην ενότητα αυτή παρουσιάζονται οι αλγόριθμοι που χρησιμοποιήθηκαν για την υλοποίηση του συστήματος πλοιόγησης, καθώς και το λογισμικό που αξιοποιήθηκε για την ανάπτυξή του. Λαμβάνοντας υπόψιν το εύρος της υλοποίησης, ήταν αναγκαίος ο συνδυασμός πολλών διαφορετικών γνώσεων και εργαλείων, γεγονός που πολλές φορές δυσκόλεψε την υλοποίηση λόγω μη συμβατότητας. Παρόλα αυτά, η ενασχόληση του συγγραφέα της εργασίας με τόσα διαφορετικά εργαλεία και γλώσσες προγραμματισμού τον βοήθησε να αποκτήσει μια πιο ολοκληρωμένη άποψη για το συγκεκριμένο ερευνητικό πεδίο και να αναπτύξει πολλές νέες δεξιότητες.

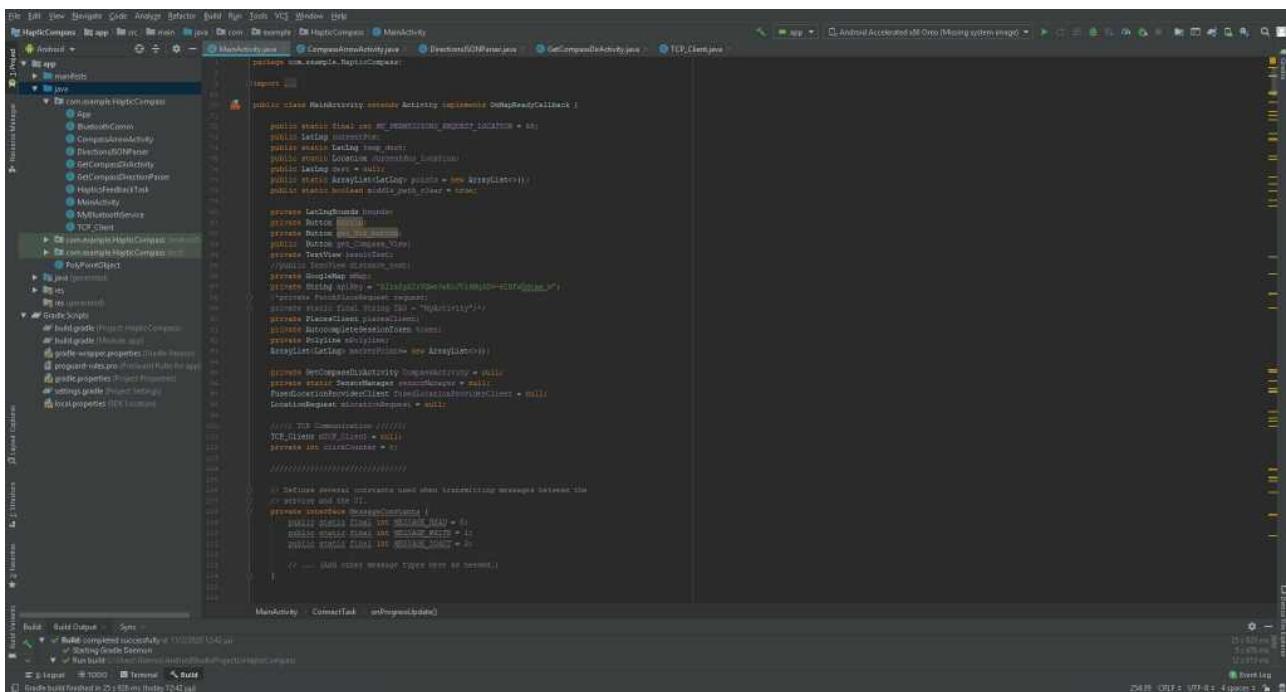
3.6.1 Χρησιμοποιηθέντα εργαλεία και γλώσσες προγραμματισμού

Για την ανάπτυξη της εφαρμογής πλοιόγησης σε λειτουργικό Android, επιλέχθηκε το λογισμικό της Google, *Android Studio*, το οποίο υποστηρίζει συγγραφή κώδικα σε Java, ενώ για την συγγραφή του κώδικα που αφορά τους αλγορίθμους επεξεργασίας εικόνας χρησιμοποιήθηκε το *Visual Studio* της Microsoft και η συγγραφή έγινε στη γλώσσα C++. Καθ' όλη την υλοποίηση των αλγορίθμων έγινε εκτεταμένη χρήση της γνωστής βιβλιοθήκης *OpenCV* (Open Computer Vision) και του Intel® *RealSense™ SDK* (Software Development Kit) που παρέχεται από την Intel.

3.6.1.1 *Android Studio - Java*

To *Android Studio IDE* (Intergrated Development Environment) [55] είναι ίσως το πιο διαδεδομένο λογισμικό ανάπτυξης εφαρμογών για *Android* κινητά. Υποστηρίζεται επίσημα από την Google και μπορεί να εγκατασταθεί στις περισσότερες συσκευές (Windows, Mac, Linux). Το λογισμικό αυτό διευκολύνει τον σχεδιασμό του γραφικού περιβάλλοντος της εφαρμογής, δηλαδή της διεπιφάνειας χρήστη, και παρέχει τη δυνατότητα αξιοποίησης όλων των διαθέσιμων υπηρεσιών σε ένα smartphone, π.χ. GPS, IMU sensors, Bluetooth κ.α. Παράλληλα, ένα μεγάλο πλεονέκτημα του *Android Studio* είναι ότι επιτρέπει το γρήγορο και άμεσο testing της εφαρμογής υπό ανάπτυξη στο κινητό του χρήστη, ή σε κάποιον προσομοιωτή (emulator).

Πιο συγκεκριμένα, χρησιμοποιήθηκε η έκδοση *Android Studio 3.5.1*, η οποία παρέχει την δυνατότητα συγγραφής κώδικα σε Java [56] και σε *Kotlin* [57], η οποία είναι μια παραλλαγή της Java. Όπως προαναφέρθηκε, επιλέχθηκε η ανάπτυξη σε Java, λόγω της μεγαλύτερης και ευρύτερης υποστήριξης από την κοινότητα.

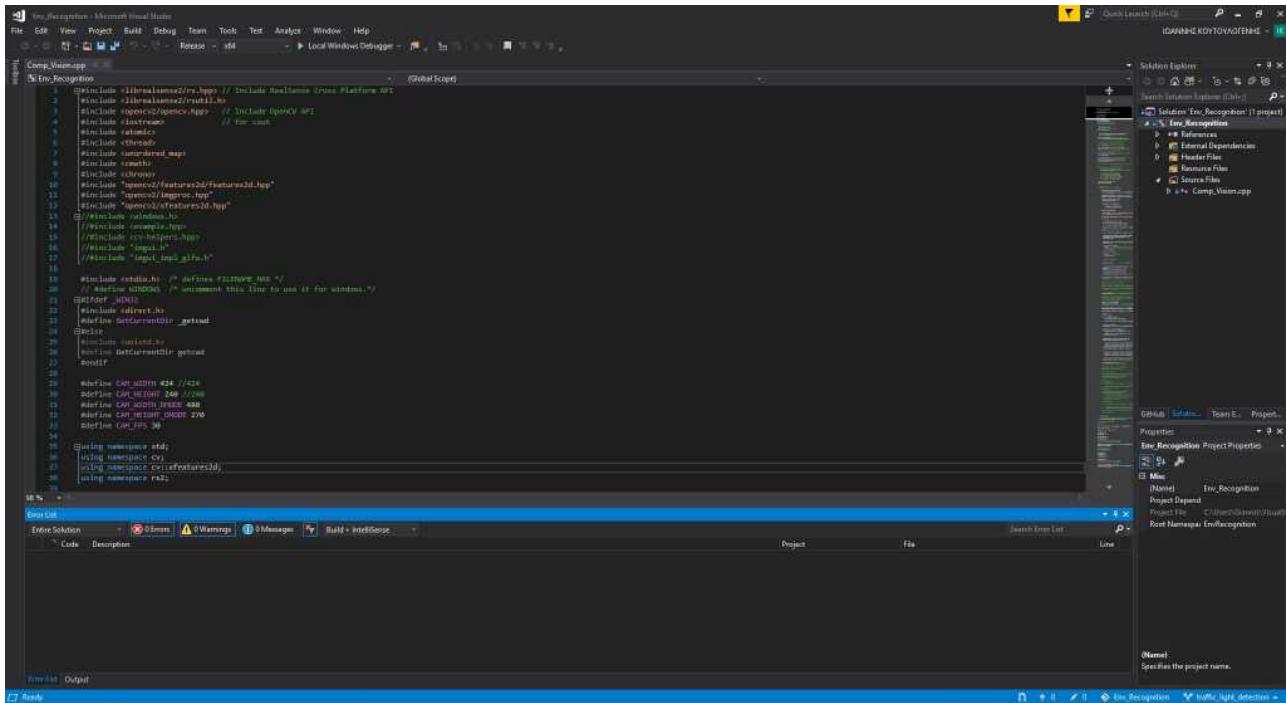


Σχήμα 3.15: Προγραμματιστικό περιβάλλον Android Studio IDE

3.6.1.2 Visual Studio - C++

Η κύρια ανάπτυξη κώδικα έγινε σε C++ στο περιβάλλον του *Visual Studio 2017 IDE* [58]. Πρόκειται για ένα από τα καλύτερα IDE, που προσφέρουν αρκετές διευκολύνσεις κατά την ανάπτυξη κώδικα, όπως είναι η αυτόματη συμπλήρωση, εντοπισμός σφαλμάτων, debugging κ.α. Ένα IDE (Integrated Development Environment), ή αλλιώς ολοκληρωμένο περιβάλλον ανάπτυξης, είναι μια σουίτα λογισμικού που βοηθάει στην ανάπτυξη προγραμμάτων υπολογιστή και περιλαμβάνει κάποιον επεξεργαστή πηγαίου κώδικα, έναν μεταγλωττιστή, εργαλεία αυτόματης παραγωγής κώδικα, debugger, linker, version control systems (git) και εργαλεία κατασκευής γραφικών διασυνδέσεων χρήστη.

Ως εναλλακτική επιλογή υπήρχε η ανάπτυξη του κώδικα στη γλώσσα Python [59], η οποία είναι πολύ διαδεδομένη τα τελευταία χρόνια και υποστηρίζεται σε πολύ μεγάλο βαθμό από την κοινότητα, ωστόσο απορρίφθηκε επειδή στόχος της διπλωματικής ήταν η υλοποίηση κώδικα για embedded συστήματα, όπως το Raspberry Pi, για τα οποία η C++ είναι πιο κατάλληλη γλώσσα. Επίσης, η Python είναι εν γένει πιο αργή σε σύγκριση με την C++, καθώς η πρώτη ακολουθεί μια διαδικασία "μετάφρασης", ενώ η δεύτερη γίνεται compile κατευθείαν σε γλώσσα μηχανής. Τέλος, ο κώδικας που γράφτηκε στο Visual Studio μεταφέρθηκε στο Raspberry Pi αυτούσιος, με μικρές μόνο τροποποιήσεις, ώστε να είναι συμβατός με το λειτουργικό σύστημα Linux του Raspberry Pi.



Σχήμα 3.16: Προγραμματιστικό περιβάλλον Visual Studio IDE

3.6.1.3 OpenCV

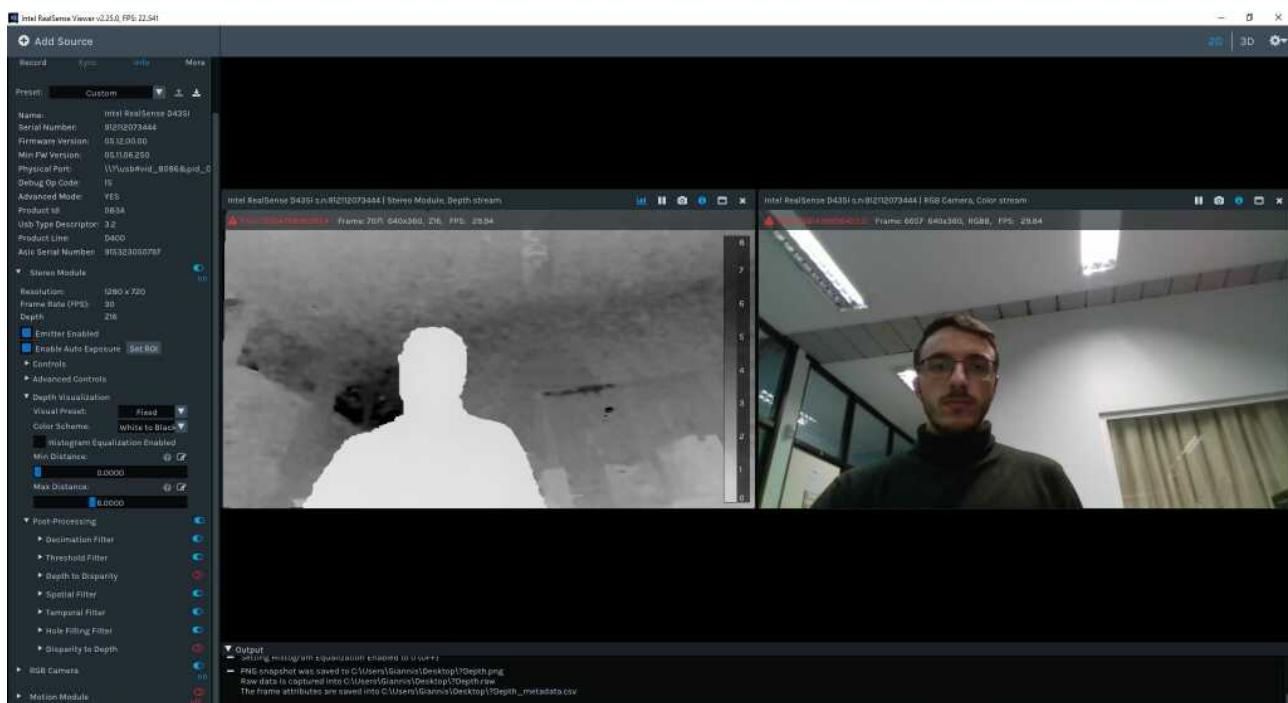
Στα πλαίσια της διπλωματικής χρησιμοποιήθηκε η έκδοση *OpenCV 4.2.0*. Η βιβλιοθήκη OpenCV [60] είναι μια συλλογή από βελτιστοποιημένες συναρτήσεις που χρησιμοποιούνται κυρίως στο πεδίο της μηχανικής όρασης, ενώ παράλληλα παρουσιάζει συμβατότητα με τις περισσότερες πλατφόρμες και χρησιμοποιείται κάτω από την άδεια ανοιχτού κώδικα. Ο λόγος που επιλέχθηκε η αξιοποίηση της OpenCV είναι η ενσωμάτωση πολλών έτοιμων φίλτρων και συναρτήσεων για επεξεργασία εικόνας, που είναι ήδη βελτιστοποιημένες για πιο αποδοτική χρήση, επιτρέποντας στον προγραμματιστή να ασχοληθεί με πιο αφηρημένες έννοιες όσον αφορά τον σχεδιασμό του συστήματος. Τέλος, η υποστήριξη της OpenCV από την κοινότητα είναι καθολική και είναι πολύ πιο εύκολο να διορθωθούν τυχόν σφάλματα στον κώδικα. Η κύρια γλώσσα προγραμματισμού καθώς και η γλώσσα στην οποία είναι γραμμένη η OpenCV είναι η C++.



Σχήμα 3.17: Λογότυπο βιλβιοθήκης OpenCV

3.6.1.4 Intel® RealSense™ SDK 2.0

Παράλληλα με την OpenCV αξιοποιήθηκε και η βιβλιοθήκη LibRealsense [61] που παρέχετε από το SDK της κάμερας Intel Realsense. Πιο συγκεκριμένα, η κάμερα που χρησιμοποιήθηκε έρχεται μαζί με ένα πακέτο συναρτήσεων που υλοποιούν βασικά κομμάτια επικοινωνίας της κάμερας με το πρόγραμμα και διευκολύνουν την πρόσβαση στα δεδομένα της κάμερας, π.χ. depth maps και rgb frames. Παράλληλα, μέσα από το συγκεκριμένο SDK δίνεται η δυνατότητα παρέμβασης στις εσωτερικές προεπιλεγμένες ρυθμίσεις της κάμερας και παρέχονται βασικά παραδείγματα και χρήσιμα εργαλεία αποσφαλμάτωσης (debugging).



Σχήμα 3.18: RealSense Viewer από το RealSense™ SDK 2.0

3.6.1.5 Raspberry Pi - Linux OS

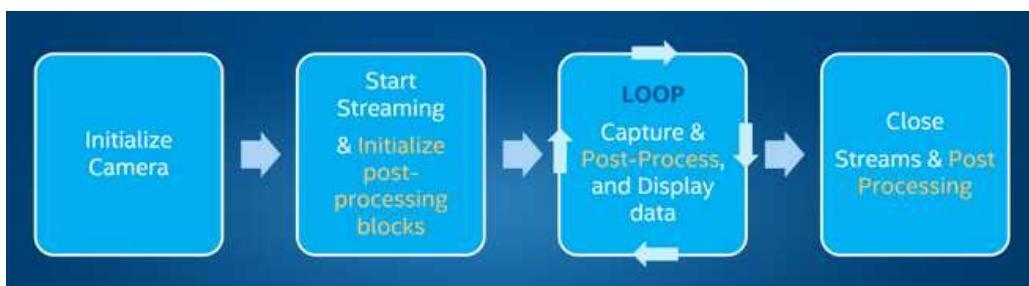
Το λειτουργικό που χρησιμοποιήθηκε στο Raspberry Pi 4 είναι το *Ubuntu Server 19.10 64-bit* [62], με κωδική ονομασία *Eoan Ermite*, το οποίο υποστηρίζει πλήρως τις δυνατότητες του RPi4. Επειδή το λειτουργικό σύστημα Ubuntu Server δεν έχει ενσωματωμένο γραφικό περιβάλλον, εγκαταστάθηκε το MATE Desktop Environment [63].



Σχήμα 3.19: Ubuntu 19.10 with MATE Desktop Environment

3.6.2 Αλγόριθμοι επεξεργασίας εικόνας σε πραγματικό χρόνο

Παρακάτω αναλύονται οι αλγόριθμοι που χρησιμοποιήθηκαν στην παρούσα διπλωματική εργασία, αποφεύγοντας όσο γίνεται τις πολύ τεχνικές λεπτομέρειες που αφορούν την προγραμματιστική υλοποίησή τους. Αρχικά, είναι σημαντικό να αναφέρουμε ότι το πρόγραμμά μας αποτελείται από δύο διαφορετικά threads, δηλαδή κομμάτια κώδικα που τρέχουν το ένα ανεξάρτητα από το άλλο, όπου το πρώτο thread (εφεζής *processing thread*) τρέχει στο παρασκήνιο και είναι υπεύθυνο για να "παραλαμβάνει" τα frames από την κάμερα και να εφαρμόζει ένα αρχικό φιλτράρισμα μόνο στα depth frames, ενώ στο δεύτερο thread (εφεζής *main thread*) γίνεται η κυρίως επεξεργασία και ανάλυση των εικόνων και τρέχει σε ένα συνεχή βρόχο για όσο είναι ανοιχτό το πρόγραμμα. Ο λόγος που χρησιμοποιούμε δύο διαφορετικά thread είναι για να αποφύγουμε καταστάσεις όπου το πρόγραμμα "κολλάει" αναμένοντας κάποιο frame να φτάσει από την κάμερα.

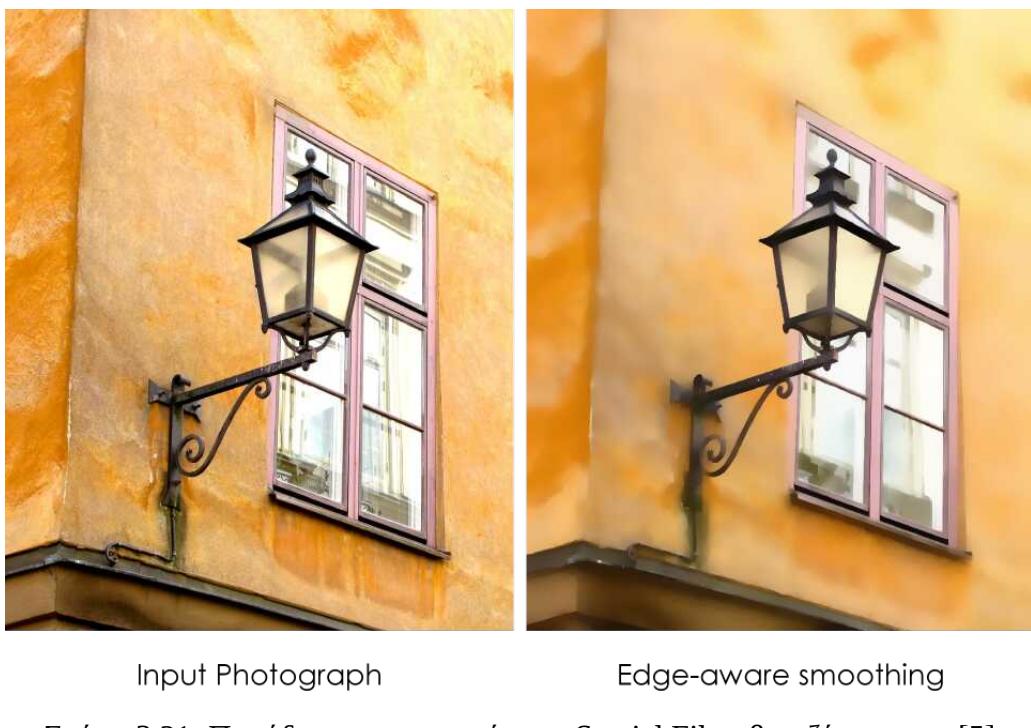


Σχήμα 3.20: Διάγραμμα ροής ενός frame

Το πρόγραμμα ρυθμίζεται να λειτουργεί με RGB frames ανάλυσης 424x240 στα 30fps και Depth frames ανάλυσης 480x270 στα 30fps. Οι συγκεκριμένες αναλύσεις επιλέχθηκαν ώστε να κάνει την επεξεργασία πιο γρήγορη, μιας και δεν είναι αναγκαία μεγαλύτερη ενάλυση για την εφαρμογή που

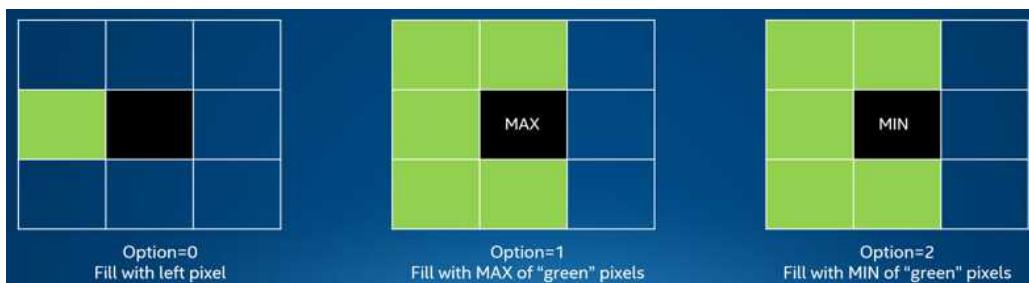
θέλουμε. Στη συνέχεια, αφού παραληφθεί κάποιο depth frame στο processing thread, εφαρμόζονται τα εξής φίλτρα με σειρά προτεραιότητας [64]:

- Decimation filter:** Φίλτρο που μειώνει την πολυπλοκότητα της εικόνας, εφαρμόζοντας ουσιαστικά υπο-δειγματοληψία (downsampling) της αρχικής εικόνας. Πιο συγκεκριμένα, εφαρμόζεται συνέλιξη του αρχικού frame με ένα πίνακα kernel 2x2, χρησιμοποιώντας την μέση τιμή των pixels που καλύπτονται από τον kernel. Αντίστοιχα, το μέγεθος της εικόνας μειώνεται αναλογικά και στις δύο διαστάσεις για να διατηρηθεί η αρχική αναλογία διαστάσεων. Η διαδικασία αυτή μπορεί να αυξήσει την ταχύτητα της μετέπειτα επεξεργασίας μέχρι 4 φορές, ενώ παράλληλα συμβάλλει στην εξάλειψη τυχόν κενών pixels (black holes).
- Spatial Edge-Preserving filter:** Φίλτρο που εξομαλύνει τον θόρυβο βάθους (depth noise), διατηρεί τις γωνίες/άκρες και κάνει τις επιφάνειες πιο επίπεδες [5].



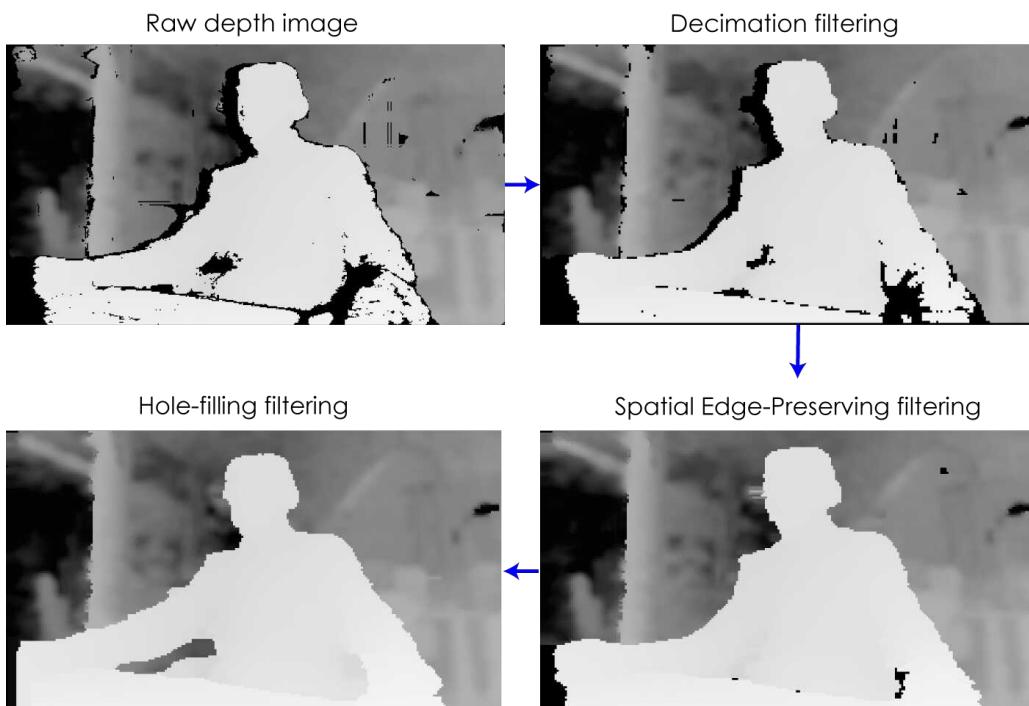
Σχήμα 3.21: Παράδειγμα εφαρμογής του Spatial Filter βασιζόμενο στο [5]

- Temporal filter:** Φίλτρο που βελτιώνει τα δεδομένα βάθους με βάση τα αντίστοιχα frames σε προηγούμενη χρονική περίοδο. Πιο συγκεκριμένα, για κάθε pixel που είναι κενό ή έχει λανθασμένη τιμή κοιτάζει το ιστορικό και χρησιμοποιεί την τιμή που είχε το pixel αυτό σε προηγούμενο frame. Ο κανόνας που χρησιμοποιείται είναι ότι η τιμή ενός κενού pixel αντικαθίσταται από την τελευταία έγκυρη τιμή, αν αυτή είναι έγκυρη στα τελευταία 2 από τα 4 frames.
- Hole-Filling filter:** Φίλτρο που χρησιμοποιείται για εξάλειψη τυχόν εναπομείναντων κενών pixels (holes). Το φίλτρο κάνει αναζήτηση σε μια γειτονιά ενός pixel και η διαδικασία που ακολουθείται είναι ότι κάθε κενό pixel αντικαθίσταται από την τιμή του γειτονικού pixel με την κοντινότερη απόσταση από τον αισθητήρα (nearest from around).



Σχήμα 3.22: Επιλογές φίλτρου εξάλειψης κενών pixels (hole filling)

Στη συνέχεια, τα RGB και Depth frames προωθούνται στο main thread και ξεκινάει η εφαρμογή των αλγορίθμων που θα περιγραφεί παρακάτω.



Σχήμα 3.23: Παράδειγμα εφαρμογής φίλτρων στο στάδιο του preprocessing (εξαιρείται το temporal filter, γιατί στην συγκεκριμένη λήψη δεν άλλαζε σημαντικά το αποτέλεσμα)

3.6.2.1 Αναγνώριση διάβασης πεζών

Η αναγνώριση διάβασης πεζών βασίζεται σε μια παραλλαγή του αλγορίθμου που προτείνεται από τους Wu, X., Hu, R., & Bao, Y. (2019), στη δημοσίευση με τίτλο *Block-Based Hough Transform for Recognition of Zebra Crossing in Natural Scene Images* [6]. Βασική προϋπόθεση είναι να τρέχει σε πραγματικό χρόνο (real-time) και να παρέχει αξιόπιστα αποτελέσματα. Ο προτεινόμενος αλγόριθμος βασίζεται στην εφαρμογή του Μετασχηματισμού Hough σε συγκεκριμένα τμήματα της αρχικής εικόνας που καλούνται blocks και χωρίζεται σε δύο φάσεις:

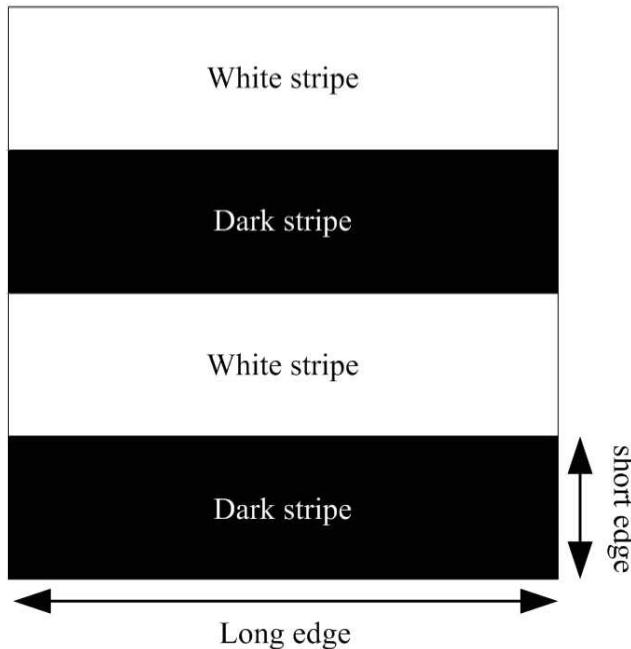
1. την φάση της αναγνώρισης ανά block (Block-Based Recognition), και
2. την φάση της σύνθεσης (Synthesize).

3.6.2.1.1 Χαρακτηριστικά και περιορισμοί διάβασης πεζών Η κλασσική διάβαση πεζών αποτελείται από λευκές και μαύρες λωρίδες που εναλλάσσονται μεταξύ τους και θυμίζουν το ασπρόμαυρο

μοτίβο της ζέβρας (Σχήμα 3.24). Το σύστημα πλοϊγησης που παρουσιάζεται στη συγκεκριμένη εργασία θέτει ως βασική προϋπόθεση την ύπαρξη διάβασης πεζών τύπου ζέβρας, ώστε να μπορεί να βοηθήσει τον χρήστη να διασχίσει ένα δρόμο. Τέτοιου τύπου διαβάσεις χαρακτηρίζονται από τις 2 παρακάτω ιδιότητες:

- Οι μεγάλες πλευρές των λωρίδων είναι παράλληλες
- Η χρωματική ένταση έχει διπολικά χαρακτηριστικά (άσπρο-μαύρο)

Ωστόσο, πολλές φορές οι παραπάνω ιδιότητες εκφυλίζονται όταν πρόκειται για διαβάσεις στο πραγματικό περιβάλλον, επειδή άλλοι παράγοντες (όπως συνθήκες φωτισμού, αντανακλάσεις, ξεθώριασμα, γεωμετρία δρόμου, οπτική παρεμπόδιση) επηρεάζουν σημαντικά το πως φαίνονται οι διαβάσεις.



Σχήμα 3.24: Κλασσική μορφή διάβασης πεζών τύπου ζέβρας [6]

3.6.2.1.2 Η προτεινόμενη μέθοδος Επειδή συνήθως οι διαβάσεις πεζών βρίσκονται στο κάτω μέρος της εικόνας όπως στέκεται ένας πεζός, καταρχήν περιορίζουμε την αρχική εικόνα ορίζοντας μια περιοχή ενδιαφέροντος, η οποία περιλαμβάνει το κάτω μισό της αρχικής εικόνας (50% του ύψους) και το 80% του αρχικού πλάτους. Έπειτα αυτή η περιοχή διαιρείται σε πολλαπλά επικαλυπτόμενα blocks που έχουν ίδιο ύψος με την περιοχή ενδιαφέροντος, αλλά μικρότερο πλάτος. Αφού οριστεί το μέγεθος κάθε block (στη παρούσα εργασία επιλέχθηκε το πλάτος κάθε block ίσο με το 1/3 του πλάτους της περιοχής ενδιαφέροντος), το πρόγραμμα μπαίνει σε έναν βρόχο όπου εφαρμόζεται ο Μετασχηματισμός Hough σε κάθε επιμέρους block. Όπως αναφέρθηκε, τα blocks είναι επικαλυπτόμενα, δηλαδή σε κάθε επανάληψη του βρόχου μπαίνει ένα νέο block το οποίο είναι μετατοπισμένο κατά μερικά pixels πιο δεξιά στον άξονα x σε σχέση με το προηγούμενο, μέχρι να φτάσουμε στο τέλος της περιοχής ενδιαφέροντος.

Ο αριθμός των pixels που μετατοπίζεται το block σε κάθε επανάληψη εξαρτάται από την επιλογή του χρήστη και επηρεάζει άμεσα την ποιότητα αλλά και την ταχύτητα του αλγορίθμου.

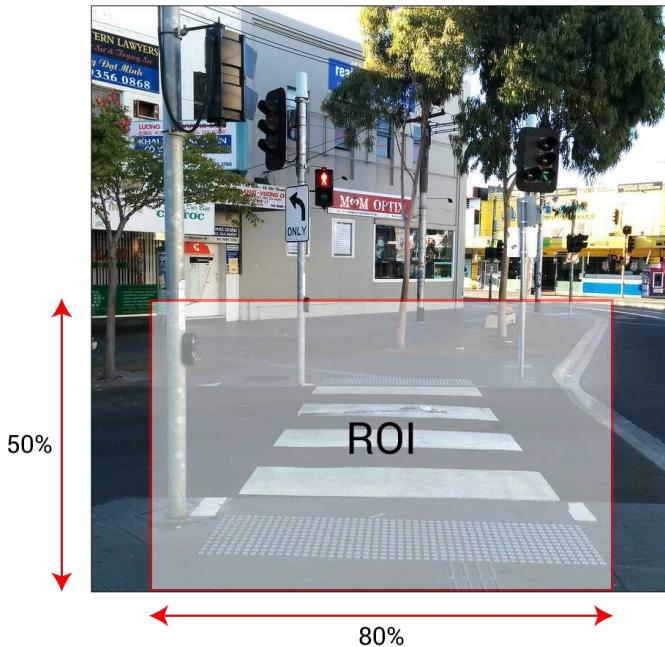
Όσο μικρότερη είναι η μετατόπιση του block τόσο πιο αναλυτικός γίνεται ο αλγόριθμος και παράγει καλύτερα αποτελέσματα, σε βάρος όμως της ταχύτητας. Αντίθετα, όσο μεγαλύτερη γίνεται η μετατόπιση τόσο πιο γρήγορος γίνεται ο αλγόριθμος με αντίστοιχη μείωση της ποιότητας του αποτελέσματος. Στη παρούσα εργασία, μετά από αρκετές δοκιμές, επιλέχθηκε εμπειρικά η μετατόπιση κατά 75

pixels σε κάθε επανάληψη, που αποτελεί έναν ”καλό” συμβιβασμό ανάμεσα στην ποιότητα και την ταχύτητα.

Ο αλγόριθμος χρησιμοποιεί μια λογική όπου σε κάθε pixel ανατίθεται ένα σκορ, το οποίο στην αρχή είναι μηδέν. Εάν εντοπιστούν παράλληλες γραμμές σε ένα block, τότε εκτιμάται η γωνία τους και μεταβάλεται το σκορ των pixels που περιέχονται σε αυτό το block κατά μία συγκεκριμένη φόρμουλα που παρουσιάζεται παρακάτω:

$$\text{score}[pixel] = 0.2 \times \text{score}[pixel] + 0.8 \times \text{number_of_pixels_in_parallel_lines}$$

Παρατηρούμε ότι το σκορ ενός pixel του block εξαρτάται κατά 20% από την προηγούμενη τιμή του και κατά 80% από το πόσο ισχυρή είναι η παρουσία των παράλληλων γραμμών στο εκάστοτε block. Αφού η διαδικασία αυτή γίνει για κάθε block, στο τέλος υπολογίζεται ο μέσος όρος των γωνιών, ώστε να βρεθεί η κατεύθυνση της διάβασης, και τα αθροιστικά σκορ συνθέτονται για να βρεθεί η θέση της διάβασης στην εικόνα.



Σχήμα 3.25: Επιλογή περιοχής ενδιαφέροντος (ROI) για εντοπισμό διάβασης πεζών

3.6.2.1.3 Canny Edge Detector & Hough Transform Στην υλοποίηση χρησιμοποιήθηκαν δύο πολύ διαδεδομένες τεχνικές στο χώρο της επεξεργασίας εικόνας, ο μετασχηματισμός Hough και το φίλτρο Canny:

- **Canny Edge Detector:** Το φίλτρο Canny αποτελεί μια από τις πιο γνωστές μεθόδους εξαγωγής των ακμών από μια εικόνα [65]. Ως ακμή ορίζεται το όριο μεταξύ περιοχών με σχετικά διακριτές τιμές χρωματικών πυκνοτήτων. Με τον όρο ακμές για μια ασπρόμαυρη εικόνα, αναφερόμαστε σε αλλαγές της φωτεινότητας μεταξύ γειτονικών περιοχών της. Τα βήματα του φίλτρου Canny είναι τα ακόλουθα:
 - Βήμα 1: Εφαρμογή του γκαουσιανού φίλτρου (Gaussian Filtering) για εξομάλυνση της εικόνας και ελαχιστοποίηση της επίδρασης του θορύβου.



Σχήμα 3.26: Εφαρμογή Gaussian Filter

- Βήμα 2: Εφαρμογή του τελεστή διαφόρισης στην εικόνα (συνέλιξη τελεστή με εικόνα A) για να βρεθεί η κατευθυντική παράγωγος της εικόνας. Το φίλτρο Canny που υλοποιείται στην βιβλιοθήκη OpenCV χρησιμοποιεί τον τελεστή Sobel [66]:

$$G_x = \begin{bmatrix} +1 & 0 & -1 \\ +2 & 0 & -2 \\ +1 & 0 & -1 \end{bmatrix} * A, G_y = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * A$$

Το μέτρο της παραγώγου δίνεται από την $G = \sqrt{G_x^2 + G_y^2}$ και η κατεύθυνση από την $\Theta = \arctan(\frac{G_y}{G_x})$.



Σχήμα 3.27: Εφαρμογή τελεστή Sobel για εύρεση κατευθυντικών παραγώγων

- Βήμα 3: Non Maximum Suppresion (απαλοιφή των μη-μεγίστων) για την διατήρηση μόνο των έντονων ακμών.



Σχήμα 3.28: Εφαρμογή Non Maximum Suppresion

- Βήμα 4: Εφαρμογή κατωφλίωσης (double thresholding) για τον περαιτέρω διαχωρισμό των αχνών ακμών.



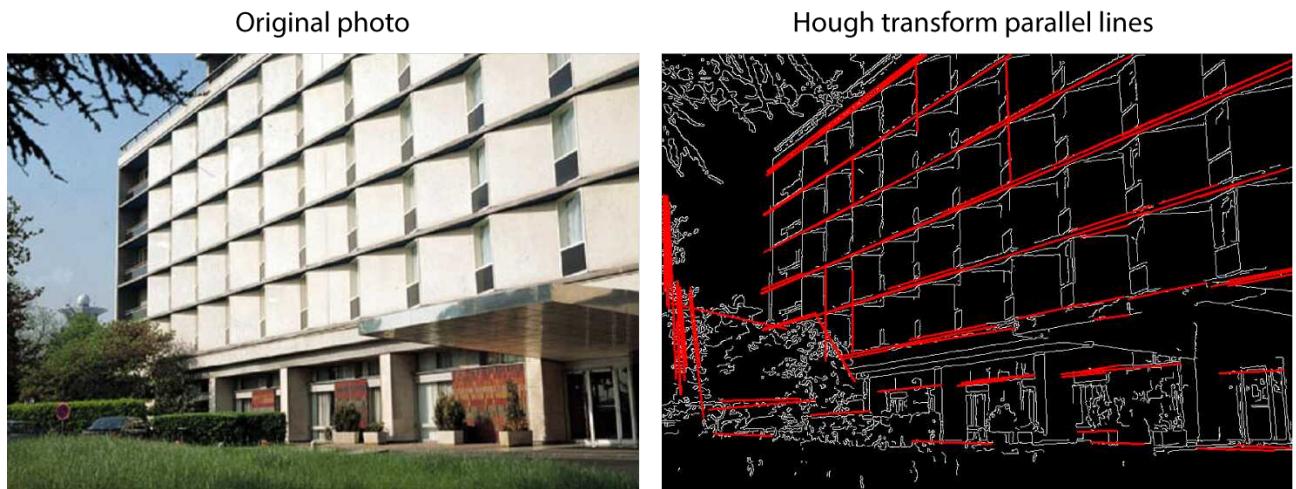
Σχήμα 3.29: Εφαρμογή Double Thresholding

- Βήμα 5: Αφού έχουμε διαχωρίσει τις έντονες από τις αχνές ακμές, εφαρμόζουμε κατωφλίωση των ακμών με υστέρηση, δηλαδή διατήρηση μόνο των έντονων ακμών και των αχνών ακμών που συνδέονται με έντονες. Αυτό παράγει μια εικόνα στην οποία φαίνονται μόνο οι ακμές που αντιστοιχούν σε πραγματικές γραμμές.



Σχήμα 3.30: Εφαρμογή κατωφλίωσης με υστέρηση

- *Hough Transform*: Ο μετασχηματισμός Hough είναι μια ευρύτερη μέθοδος εντοπισμού χαρακτηριστικών συγκεκριμένου σχήματος μέσα σε μια εικόνα [67]. Επειδή τα επιθυμητά χαρακτηριστικά πρέπει να βρίσκονται σε κάποια παραμετρική μορφή, ο κλασσικός μετασχηματισμός Hough χρησιμοποιείται συχνά για τον εντοπισμό καμπύλων, όπως ευθείες γραμμές, κύκλοι, ελλείψεις κλπ. Τα βασικότερα πλεονεκτήματα του μετασχηματισμού αυτού είναι ότι είναι αρκετά αποτελεσματικός ακόμα και σε επικαλυπτόμενα αντικείμενα (δηλαδή είναι ανεκτικός στην έλλειψη κομματιών από μια ευθεία) και παράλληλα δεν επηρεάζεται πολύ από τον θόρυβο που υπεισέρχεται στην εικόνα. Στα πλαίσια της παρούσας διπλωματικής επιλέχθηκε η χρήση του μετασχηματισμού Hough για εντοπισμό παράλληλων γραμμών. Η βασική ιδέα της μεθόδου είναι ότι μετατρέπει τα σημεία από καρτεσιανές συντεταγμένες (x, y) σε πολικές (ρ, θ) και έπειτα χρησιμοποιεί μια διαδικασία ψηφοφορίας για την ανάδειξη των υποψήφιων σχημάτων στην εικόνα.



Σχήμα 3.31: Εφαρμογή Hough Transform

3.6.2.1.4 Block-Based Recognition Το κομμάτι του Block-Based Recognition περιλαμβάνει την προ-επεξεργασία και τον εντοπισμό παράλληλων γραμμών. Χρησιμοποιείται για να υπολογιστούν οι γωνίες των παράλληλων γραμμών και για να καταγραφούν τα σκορ κάθε pixel στα blocks.

1) *Προ-επεξεργασία:* Η φάση αυτή περιλαμβάνει τα εξής 4 βήματα:

- Βήμα 1: Μετατροπή του block από έγχρωμο σε κλίμακα του γκρι (grayscale), επειδή οι γραμμές της διάβασης σε μια εικόνα κλίμακας του γκρι φαίνονται ως σκούρες και φωτεινές λωρίδες.
- Βήμα 2: Εφαρμόζεται η μέθοδος προσαρμοστικού κατωφλίου (adaptive thresholding), η οποία αποδίδει σε ένα pixel μια δυαδική τιμή (0 ή 1) ανάλογα με το αν η τιμή του pixel είναι μεγαλύτερη από ένα μεταβλητό κατώφλι. Η διαδικασία αυτή βοηθάει να περιοριστεί το φαινόμενο της σκίασης στις εικόνες.
- Βήμα 3: Εφαρμογή των μορφολογικών μετασχηματισμών Διαστολής (Dilation) και Διάβρωσης (Erosion) για εξομάλυνση των ορίων.
- Βήμα 4: Εφαρμογή του φίλτρου Canny Edge Detector για να ανιχνευθούν οι ακμές που υπάρχουν στο block.

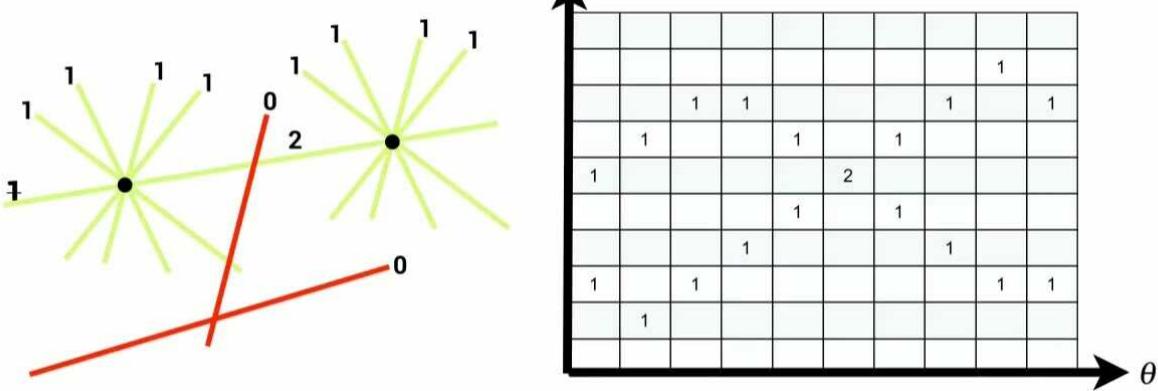
2) *Εντοπισμός παράλληλων γραμμών:* Εφαρμόζεται ο Μετασχηματισμός Hough για να εντοπιστούν οι παράλληλες γραμμές στο block. Ορίζεται τοπικά ως αρχή των αξόνων η πάνω αριστερή γωνία του block και κάθε μη-μηδενικό pixel που ανήκει σε ακμή μετασχηματίζεται από το καρτεσιανό στο πολικό σύστημα συντεταγμένων με την παρακάτω εξίσωση:

$$\rho = x \cos \theta + y \sin \theta,$$

όπου η συντεταγμένη ρ αναπαριστά την απόσταση της αρχής των αξόνων από την ευθεία γραμμή που διέρχεται από το (x, y) , και το η συντεταγμένη θ αναπαριστά την γωνία μεταξύ του κάθετου στη γραμμή διανύσματος (normal) και του άξονα x. Η αναζήτηση περιορίζεται σε παράλληλες γραμμές που βρίσκονται στο εύρος $[-30^\circ, +30^\circ]$, δηλαδή ψάχνουμε μόνο για $\theta \in [60^\circ, 120^\circ]$, επιλέγοντας διακριτές τιμές του θ . Ανάλογα με τις πολικές συντεταγμένες (ρ, θ) που ικανοποιούν κάθε pixel ακμής, δημιουργείται ένας πίνακας συσσώρευσης (Σχήμα 3.32). Το πλεονέκτημα του Block-Based Μετασχηματισμού Hough είναι ότι φιλτράρει τυχόν σημεία που έχουν υπεισέλθει στις ακμές λόγω θορύβου. Περιληπτικά, τα βήματα που ακολουθούνται για την εξαγωγή των παράλληλων γραμμών είναι τα παρακάτω:

- Βήμα 1: Σε κάθε block, ακολουθείτε μια διαδικασία ψηφιοφορίας κατά την οποία κάθε ζεύγος (ρ, θ) , που αντιστοιχεί σε ένα pixel ακμής, αντιστοιχίζεται σε μια θέση στον πίνακα συσσώρευσης (Σχήμα 3.32). Κάθε φορά που ένα τέτοιο ζεύγος εντοπίζεται, η τιμή της αντίστοιχης θέσης στον πίνακα συσσώρευσης αυξάνεται κατά 1. Με άλλα λόγια κάθε pixel ακμής "ψηφίζει" για τις αντίστοιχες θέσεις (ρ, θ) που ικανοποιούν την εξίσωση του. Κάθε θέση με παραπάνω από 1 ψήφους θεωρείται μια ευθεία γραμμή στις καρτεσιανές συντεταγμένες.
- Βήμα 2: Επιλέγονται οι 10 θέσεις με τις περισσότερες ψήφους. Η γωνία θ με τις περισσότερες ψήφους επιλέγεται ως η γωνία ανάμεσα στο κάθετο διάνυσμα της υποψήφιας παράλληλης γραμμής και του άξονα x.
- Βήμα 3: Εξάγονται όλες οι γραμμές των οποίων η γωνία είναι ίση με την γωνία που επιλέχθηκε στο προηγούμενο βήμα και η παραλληλία επιβεβαιώνεται εάν υπάρχουν τουλάχιστον 3 ευθείες γραμμές, των οποίων το πλήθος των ψήφων είναι μεγαλύτερο από 60.

Hough voting

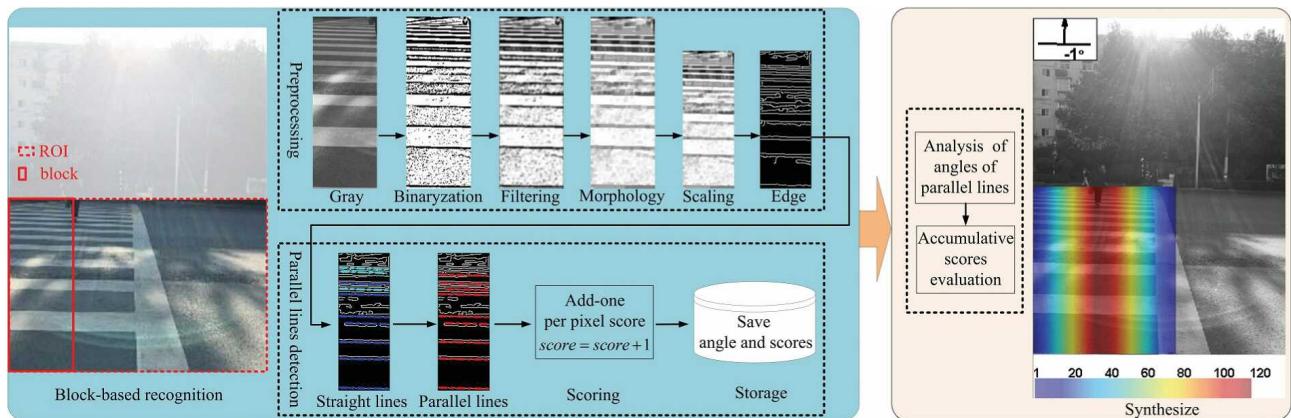


- For every point (u, v) that lies on a line we vote for **all** the lines (ρ, θ) such that

$$u \sin \theta + v \cos \theta = \rho$$

Σχήμα 3.32: Παράδειγμα δημιουργίας πίνακα συσσώρευσης κατά τον μετασχηματισμό Hough [7]. Κάθε θέση στον πίνακα εκφράζει μια ευθεία και κάθε ψήφος σε μια θέση εκφράζει τον αριθμό των σημείων από τα οποία διέρχεται αυτή η ευθεία.

3.6.2.1.5 Synthesize Αφού έχουν αναλυθεί όλα τα blocks, η τελική γωνία θ μεταξύ των παράλληλων γραμμών και του άξονα x υπολογίζεται ως αυτή με τους περισσότερους ψήφους ανά block. Τα σκορ των pixels αξιολογούνται για να εξαχθεί ένα ασφαλές μονοπάτι που να αντιστοιχεί στη διάβαση. Τα pixels με μεγαλύτερο σκορ αντιπροσωπεύουν μονοπάτι με μεγαλύτερη ασφάλεια.

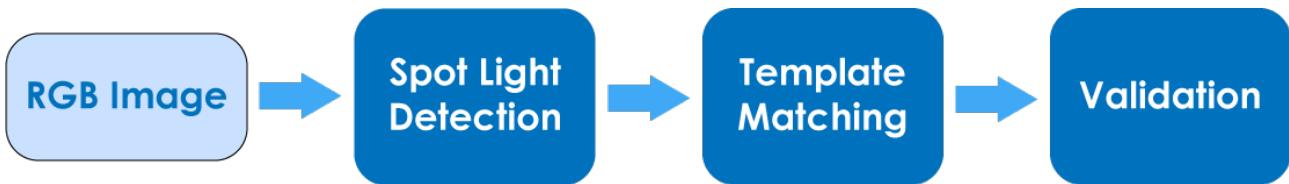


Σχήμα 3.33: Overview του αλγόριθμου εντοπισμού διάβασης πεζών που παρουσιάζεται στην ερευνητική εργασία [6]

3.6.2.2 Αναγνώριση φωτεινού σηματοδότη

Αν το σύστημα ανιχνεύσει διάβαση πεζών τότε καλείται ο αλγόριθμος αναγνώρισης φωτεινού σηματοδότη. Ο αλγόριθμος αυτός είναι υπεύθυνος για τον εντοπισμό του φαναριού στην εικόνα και την αναγνώριση του χρώματος (πράσινο ή κόκκινο). Η βασική ιδέα του αλγορίθμου βασίζεται στην μέθοδο του *Spot Light Detection* (ανίχνευση σημειακών πηγών φωτός), η οποία εκμεταλλεύεται το γεγονός ότι οι τα φανάρια είναι αυτόφωτα, δηλαδή είναι πηγές φωτός, για να μπορέσει να εντοπίσει

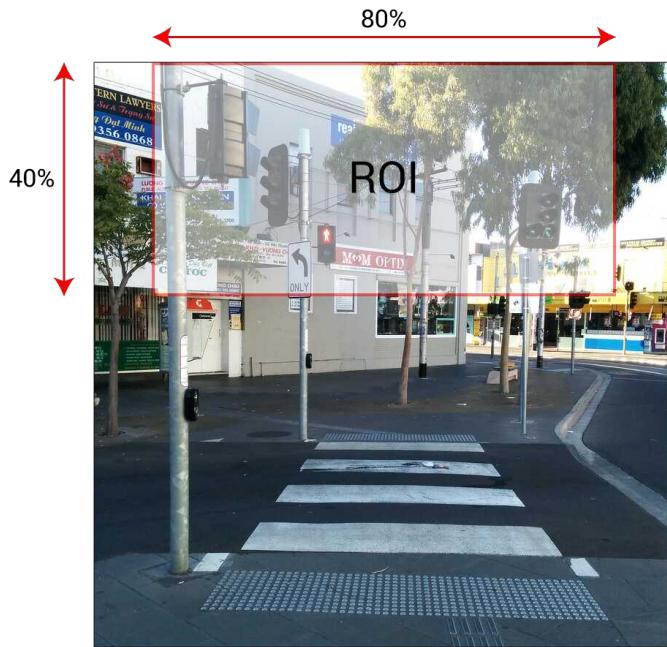
τη θέση τους μέσα σε μια εικόνα. Στη συνέχεια, αφού εντοπιστούν οι θέσεις της εικόνας με τους υποψήφιους φωτεινούς σηματοδότες, ορίζεται ένα bounding box (περίγραμμα) γύρω από αυτά και κάθε τέτοιο bounding box συγκρίνεται με ένα προκαθορισμένο template φωτεινού σηματοδότη. Αν το αποτέλεσμα τη σύγκρισης είναι μεγαλύτερο από ένα κατώφλι (φάση επαλήθευσης), που ορίζεται από τον χρήστη, τότε το συγκεκριμένο bounding box, και συνεπώς η συγκεκριμένη θέση στην εικόνα, αναπαριστά έναν φωτεινό σηματοδότη. Για να γίνει αναγνώριση του κόκκινου και του πράσινου χρώματος, έχουν οριστεί δύο διαφορετικά templates που αντιπροσωπεύουν το κόκκινο και το πράσινο φανάρι.



Σχήμα 3.34: Overview του αλγόριθμου αναγνώρισης φωτεινών σηματοδοτών

3.6.2.2.1 Spot Light Detection Σκοπός του συγκεκριμένου βήματος είναι η εξαγωγή των περιοχών της αρχικής εικόνας, που είναι υποψήφιες να περιέχουν φωτεινό σηματοδότη. Η υλοποίηση του αλγορίθμου βασίστηκε σε ορισμένες ερευνητικές δημοσιεύσεις σχετικά με τον εντοπισμό φωτεινών σηματοδοτών σε εφαρμογές αυτόνομης οδήγησης [68, 69] και έχει την ακόλουθη δομή:

- Βήμα 1: Αρχικά, όπως και στον αλγόριθμο εντοπισμού διάβασης πεζών, έτσι κι εδώ διαιρούμε την αρχική εικόνα ορίζοντας μια περιοχή ενδιαφέροντος που αντιστοιχεί στο πάνω 40% τμήμα της αρχικής εικόνας και στο 80% πλάτος. Αυτό συμβαίνει για να επιταχύνουμε τον αλγόριθμο, γνωρίζοντας ότι οι φωτεινοί σηματοδότες βρίσκονται σχεδόν πάντα στο πάνω κεντρικό τμήμα μιας εικόνας.



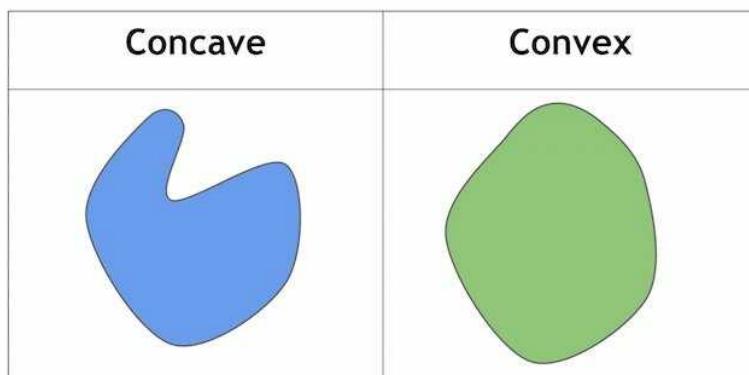
Σχήμα 3.35: Επιλογή περιοχής ενδιαφέροντος (ROI) για αναγνώριση φωτεινού σηματοδότη

- Βήμα 2: Στη συνέχεια μετατρέπουμε την εικόνα από έγχρωμη σε κλίμακα του γκρι.
- Βήμα 3: Αφαίρεση θορύβου με εφαρμογή του μορφολογικού μετασχηματισμού Top-hat [70]. Γενικά, η αρχική εικόνα μπορεί να περιέχει αρκετό θόρυβο που προέρχεται από παραμόρφωση του φακού, αντανάκλαση του ηλιακού φωτός, δόνηση της κάμερας και αντικείμενα φόντου

με παρόμοιο χρώμα. Αυτός ο θόρυβος μπορεί να περιοριστεί με την εφαρμογή μορφολογικών μετασχηματισμών. Πιο συγκεκριμένα, μετασχηματισμός Top-hat ανήκει σε μια ευρύτερη οικογένεια μορφολογικών μετασχηματισμών που χρησιμοποιούν ένα δομικό στοιχείο (structuring element) το οποίο εφαρμόζεται στην αρχική εικόνα και βάσει αυτού οι τιμές των pixels αλλάζουν ανάλογα με τις τιμές των γειτονικών pixels. Γενικά το δομικό στοιχείο είναι απλού γεωμετρικού σχήματος και μεγέθους μικρότερου από το σχήμα το οποίο θα επεξεργαστεί. Υπάρχουν διάφορα σχήματα δομικών στοιχείων που μπορούν να εφαρμοστούν, τα πιο γνωστά είναι το τετράγωνο, ο σταυρός, ο δίσκος και η γραμμή. Οι μορφολογικοί τελεστές έχουν διάφορες εφαρμογές. Απλές εφαρμογές τους είναι το φιλτράρισμα, η ανίχνευση χαρακτηριστικών, η λέπτυνση, η ενίσχυση της αντίθεσης κ.α. Στην προκειμένη περίπτωση χρησιμοποιείται ένα δομικό στοιχείο σε σχήμα τετραγώνου και επιλέχθηκε ο μετασχηματισμός White Top-hat, ο οποίος επιστρέφει μια εικόνα που περιέχει τα "αντικείμενα" που είναι "μικρότερα" από το δομικό στοιχείο και πιο φωτεινά από το περιβάλλον τους. Με αυτόν τον τρόπο δίνεται έμφαση στις φωτεινές πηγές της εικόνας, όπως οι φωτεινοί σηματοδότες.

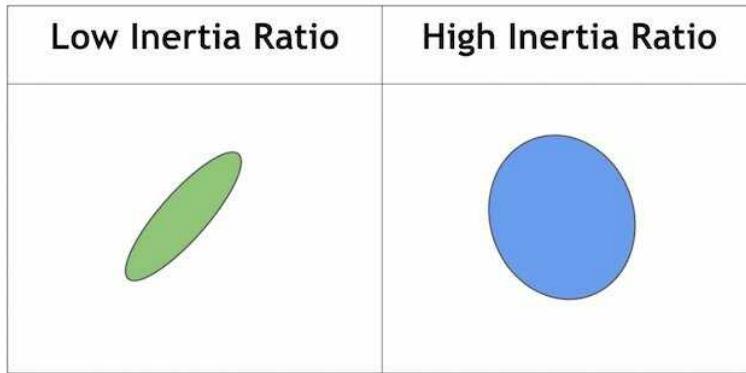
- Βήμα 4: Τέλος, γίνεται εξαγωγή των υποψήφιων περιοχών με χρήστη της μεθόδου Blob Detection. Ως blob ορίζεται ένα γκρουπ από συνδεδεμένα pixels που μοιράζονται κάποια κοινά χαρακτηριστικά, π.χ. έχουν την ίδια χρωματική τιμή, και στόχος της μεθόδου blob detection είναι να αναγνωρίσει και να προσδιορίσει αυτές τις περιοχές. Η OpenCV περιέχει μια κλάση, την *SimpleBlobDetector*, η οποία υλοποιεί την μέθοδο Blob Detection. Ο αλγόριθμος που χρησιμοποιείται από την OpenCV, προσδιορίζει τα blobs με βάση τις παρακάτω παραμέτρους:

- Μέγεθος: φιλτράρισμα των περιοχών με βάση το εμβαδόν τους σε pixels.
- Circularity: παράμετρος που καθορίζει την ομοιότητα του αντικειμένου με κύκλο.
- Convexity: με τον όρο αυτόν εννοούμε την κυρτότητα του αντικειμένου και το φίλτρο αυτό καθορίζει τον ελάχιστο και μέγιστο βαθμό κυρτότητας.



Σχήμα 3.36: Παράδειγμα ιδιότητας convexity

- Inertia Ratio: αυτή η παράμετρος μετράει το πόσο επίμηκες είναι ένα αντικείμενο.

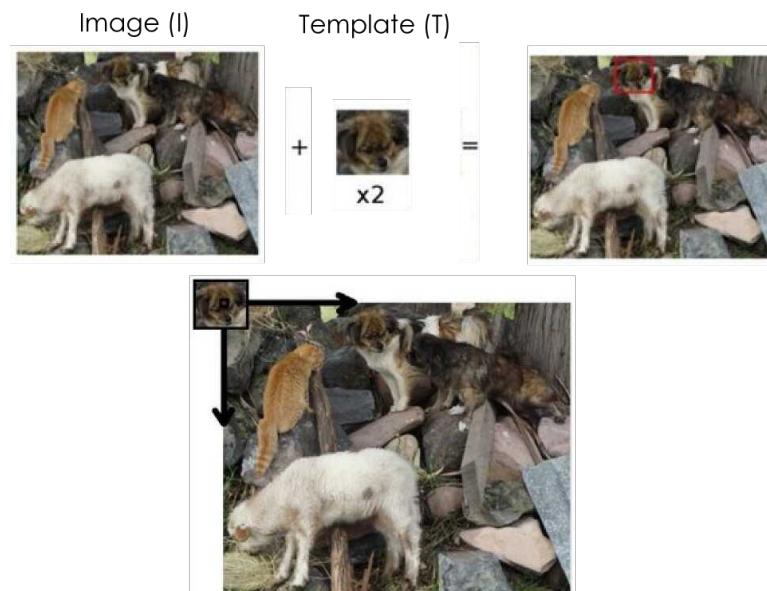


Σχήμα 3.37: Παράδειγμα ιδιότητας inertia ratio

3.6.2.2.2 Template Matching Template matching ονομάζεται η τεχνική εύρεσης περιοχών μιας εικόνας, που ταιριάζουν, ή είναι παρόμοιες, με μια δοθείσα εικόνα-πρότυπο. Αφού έχουν εντοπιστεί οι περιοχές που είναι υποψήφιες να περιέχουν κάποιον φωτεινό σηματοδότη, λαμβάνει χώρα το επόμενο στάδιο το οποίο περιλαμβάνει σύγκριση των περιοχών αυτών με κάποια προεπιλεγμένα templates φωτεινών σηματοδοτών. Σε αυτό το σημείο πρέπει να γίνει ξεκάθαρο ότι η σύγκριση γίνεται ανάμεσα στα πρότυπα και στην έγχρωμη αρχική εικόνα. Οι υποψήφιες περιοχές που βρέθηκαν προηγουμένως αντιστοιχούνται στην αρχική έγχρωμη εικόνα. Μετά από πειραματικές προσπάθειες βρέθηκε ότι τα templates που παρουσιάζουν μεγαλύτερη αποτελεσματικότητα είναι αυτά που φαίνονται στο σχήμα 3.40. Προϋπόθεση για να γίνει το template matching είναι το δοθέν πρότυπο να είναι μικρότερο σε διαστάσεις από την εικόνα προς σύγκριση. Πριν πραγματοποιηθεί η σύγκριση, εφαρμόζεται γκαουσιανό φίλτρο για την εξομάλυνση του προτύπου και το φιλτράρισμα τυχόν θορύβου. Στα πλαίσια της παρούσας διπλωματικής, στόχος ήταν να βρούμε την ομοιότητα του κόκκινου και πράσινου προτύπου με μια περιοχή της εικόνας. Αν μια υποψήφια περιοχή είχε βαθμό ομοιότητας με οποιοδήποτε πρότυπο πάνω από ένα συγκεκριμένο κατώφλι, τότε λέμε ότι αναπαριστά έναν φωτεινό σηματοδότη. Επιπλέον, συγκρίνεται η ομοιότητα με το πράσινο και το κόκκινο πρότυπο ξεχωριστά για να γίνει η αναγνώριση του χρώματος του φαναριού. Η διαδικασία ελέγχου του αποτελέσματος των συγκρίσεων με τα προκαθορισμένα κατώφλια καλείται **φάση επαλήθευσης (validation)**.

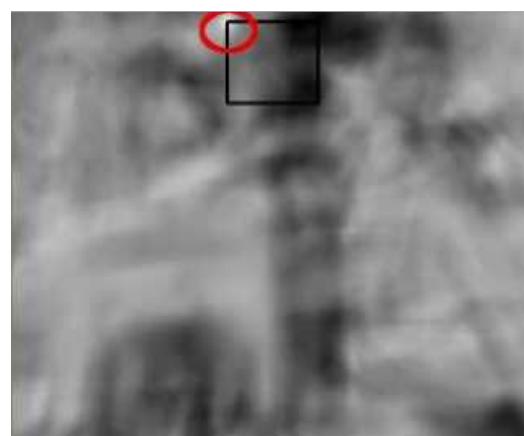
Η σύγκριση γίνεται ολισθαίνοντας την εικόνα-πρότυπο πάνω από την αρχική εικόνα. Με τον όρο ολίσθηση υποδηλώνεται η μετακίνηση της εικόνας-προτύπου κατά ένα pixel τη φορά (από τα αριστερά στα δεξιά, από πάνω προς τα κάτω). Σε κάθε θέση υπολογίζεται ένα μέτρο ομοιότητας που αντιπροσωπεύει το ποσοστό της ομοιότητας του προτύπου T με την συγκεκριμένη περιοχή της αρχικής εικόνας I . Η τιμή αυτή του μέτρου για κάθε pixel αποθηκεύεται σε μια νέα εικόνα R , η οποία αποτελεί την εικόνα μετρήσεων (σχήμα 3.39). Υπάρχουν διάφορες διαθέσιμες μέθοδοι μέτρησης της ομοιότητας [71]. Στην προκειμένη περίπτωση χρησιμοποιήθηκε η μέθοδος *Sum of Square Differences (SSD) Normed* η οποία υπολογίζει την ομοιότητα βάσει της παρακάτω εξίσωσης:

$$R(x, y) = \frac{\sum_{x',y'} (T(x', y') - I(x + x', y + y'))^2}{\sqrt{\sum_{x',y'} T(x', y')^2 \cdot \sum_{x',y'} I(x + x', y + y')^2}}$$

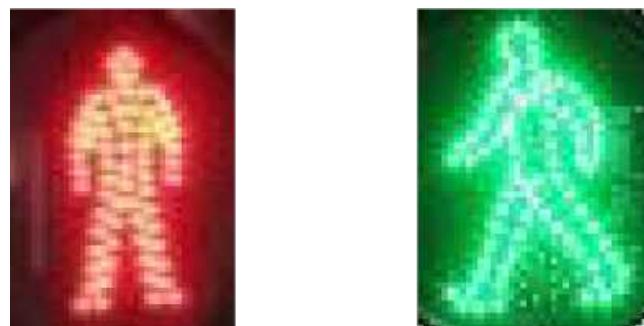


Σχήμα 3.38: Overview της διαδικασίας template matching

Result Image (R)



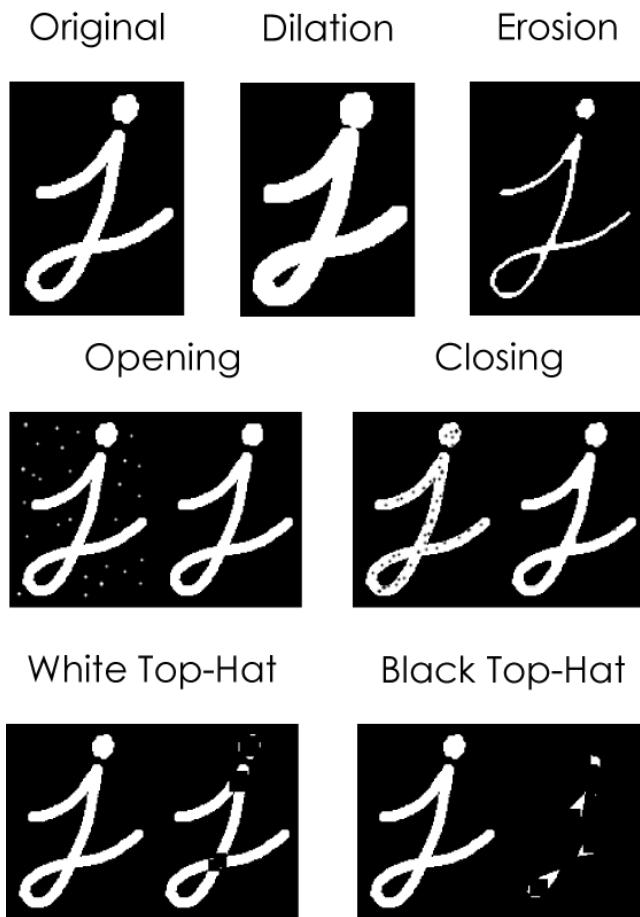
Σχήμα 3.39: Το αποτέλεσμα της διαδικασίας ολίσθησης βάσει της μετρικής Cross Correlation Normed (TM_CCORR_NORMED). Οι φωτεινές περιοχές υποδεικνύουν μεγαλύτερη ομοιότητα. Η θέση με τον κόκκινο κύκλο είναι αυτή με την υψηλότερη τιμή, γιατί η περιοχή με το περίγραμμα θεωρείται ως αυτή με το καλύτερο ταίριασμα.



Σχήμα 3.40: Τα templates που χρησιμοποιήθηκαν στην παρούσα εργασία

3.6.2.2.3 Περισσότερα σχετικά με τους μορφολογικούς μετασχηματισμούς Οι μορφολογικοί μετασχηματισμοί είναι κάποιες διεργασίες που αποσκοπούν στην εξαγωγή ορισμένων χαρακτηριστικών από τις εικόνες στις οποίες εφαρμόζονται [72]. Όπως προαναφέρθηκε, σε αυτές τις διεργασίες υπάρχει ένα δομικό στοιχείο (structuring element) το οποίο είναι ουσιαστικά ένας πίνακας οποιασδήποτε διάστασης που περιέχει 0 και 1. Γενικά, το δομικό στοιχείο είναι απλού γεωμετρικού σχήματος (τετράγωνο, σταυρός, δίσκος, γραμμή) και μεγέθους μικρότερου από το σχήμα το οποίο θα επεξεργαστεί. Κάποιες από τις βασικές διεργασίες μορφολογικών μετασχηματισμών είναι οι παρακάτω [73]:

- **Διαστολή (Dilation):** Η διαστολή επεκτείνει τα αντικείμενα της εικόνας και παράλληλα κλείνει τις τρύπες που υπάρχουν. Γι' αυτό το λόγο είναι επεκτατικό φίλτρο. Αυτό που συμβαίνει είναι ότι η διαστολή τοποθετεί το δομικό στοιχείο στην εικόνα και το μετακινεί μέσα σε αυτή με τρόπο παρόμοιο με αυτόν της συνέλιξης.
- **Συστολή (Erosion):** Η συστολή είναι το δυαδικό συμπλήρωμα της διαστολής. Η διαδικασία που εκτελεί είναι παρόμοια με τη διαστολή, μόνο που αυτό αντί να επεκτείνει τα αντικείμενα της εικόνας, τα μικρύνει, κόβοντας παράλληλα της κορυφές των περιγραμμάτων των σχημάτων και γι' αυτό το λόγο είναι μη επεκτατικό φίλτρο.
- **Άνοιγμα (Opening):** Το φίλτρο ανοίγματος αποτελεί συνδυασμό της διαστολής και της συστολής. Πιο συγκεκριμένα, υποδηλώνει μια διεργασία στην οποία ένα αντικείμενο υπόκειται πρώτα σε συστολή και αμέσως μετά σε διαστολή. Ανήκει στα μη επεκτατικά φίλτρα και χρησιμοποιείται για την αφαίρεση των μικρών αντικειμένων (που συνήθως είναι τα πιο φωτεινά) από το προσκήνιο.
- **Κλείσιμο (Closing):** Αντίστοιχα, το φίλτρο κλεισίματος αποτελεί τον ανάποδο συνδυασμό από το άνοιγμα, δηλαδή πρώτα διαστολή και αμέσως μετά συστολή. Ανήκει στα επεκτατικά φίλτρα και χρησιμοποιείται για την αφαίρεση των μικρών οπών (συνήθως μαύρου χρώματος) από το προσκήνιο.
- **Top-Hat:** Όπως εξηγήθηκε και πιο πάνω, ο μετασχηματισμός Top-Hat χρησιμοποιείται στην επεξεργασία εικόνας για την εξαγωγή χαρακτηριστικών, την εξίσωση του φόντο, την ενίσχυση της αντίθεσης της εικόνας, την απομάκρυνση αντικειμένων από μια εικόνα κ.α. Στον Top-Hat μετασχηματισμό γίνεται είτε η αφαίρεση του αρχικού σήματος με το άνοιγμα του εαυτού του ($I - Opening(I)$), το οποίο ονομάζεται και white Top-Hat transform, ή η αφαίρεση του κλεισίματος του σήματος με το αρχικό σήμα ($Closing(I) - I$), το οποίο ονομάζεται και ως black Top-Hat transform.



Σχήμα 3.41: Παράδειγμα μορφολογικών διεργασιών

3.6.2.3 Αλγόριθμος εντοπισμού εμποδίων

Το τρίτο και τελευταίο μέρος του προτεινόμενου συστήματος πλοιόγησης για άτομα με προβλήματα όρασης είναι ο αλγόριθμος εντοπισμού φυσικών εμποδίων. Ο αλγόριθμος αυτός τρέχει καθ' όλη την διάρκεια της πλοιόγησης του χρήστη και χρησιμοποιεί τον ενσωματωμένο αισθητήρα βάθους της κάμερας για να αποφασίσει σχετικά με την ύπαρξη ή όχι εμποδίων μπροστά από τον χρήστη.

Η βασική ιδέα πίσω από την υλοποίηση τους συγκεκριμένου αλγόριθμου είναι αρκετά απλή. Αρχικά, λαμβάνεται η εικόνα βάθους από την κάμερα και αφού περάσει τη φάση της προ-επεξεργασίας, που αναφέραμε στην αρχή της ενότητας, εφαρμόζεται ακόμα ένα στάδιο υπο-δειγματοληψίας, με σκοπό να μειώσουμε κι άλλο το υπολογιστικό κόστος. Πιο συγκεκριμένα, οι διαστάσεις τις εικόνας μειώνονται στο μισό. Ο λόγος που εφαρμόζουμε υπο-δειγματοληψία είναι ότι στον συγκεκριμένο αλγόριθμο δεν παίζει σημαντικό ρόλο η λεπτομέρεια στην εικόνα βάθους. Το ζητούμενο είναι να εντοπιστούν μεγάλοι, ή σχεδόν μεγάλοι, όγκοι αντικειμένων μπροστά από τον χρήστη, που εμποδίζουν την διέλευσή του από το συγκεκριμένο μονοπάτι. Έτσι, η ελάττωση της λεπτομέρειας της εικόνας βάθους που επάγεται από την υπο-δειγματοληψία δεν επηρεάζει την αποτελεσματικότητα του αλγορίθμου. Τα εμπόδια που εντοπίζονται έχουν μέγιστη απόσταση από την κάμερα 1 μέτρο. Αυτό υλοποιείται εφαρμόζοντας ένα φίλτρο αποκοπής, όπου μηδενίζει όλα τα pixels που έχουν τιμή κάτω από ένα όριο, το οποίο αντιστοιχεί σε απόσταση 1μ.

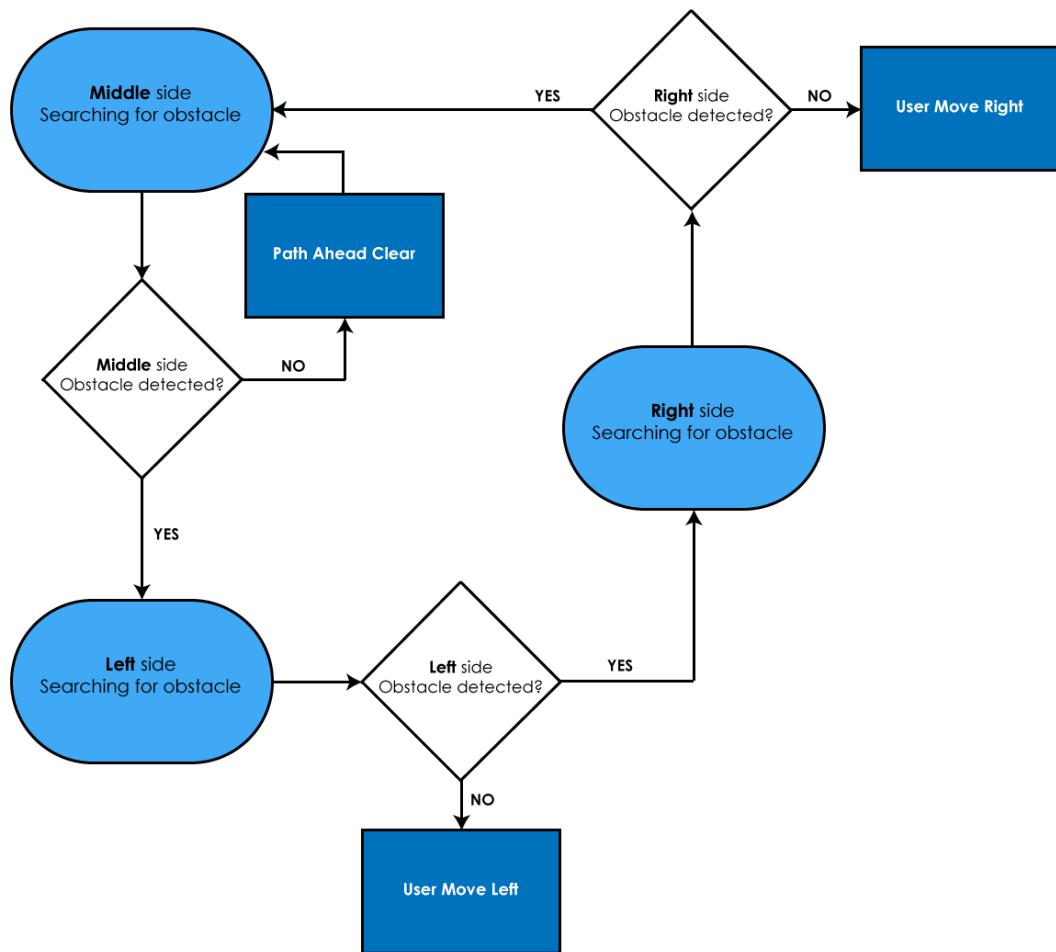
Στη συνέχεια, η εικόνα διαιρείτε σε 3 υπο-περιοχές: την αριστερή, την κεντρική και την δεξιά περιοχή. Ο αλγόριθμος ελέγχει συνεχώς την κεντρική περιοχή για τυχόν εμπόδια. Αν βρεθεί εμπόδιο στην κεντρική περιοχή τότε εξετάζει την αριστερή περιοχή και αν είναι ελεύθερη από εμπόδια, ειδοποιεί τον χρήστη να μετακινθεί προς τα αριστερά. Σε περίπτωση που υπάρχει εμπόδιο και στην αριστερή περιοχή, τότε ελέγχεται η δεξιά πλευρά και ακολούθως ο χρήστης ειδοποιείται κατάλληλα. Εάν όλες οι υπο-περιοχές είναι κατειλημμένες από εμπόδια, τότε ο αλγόριθμος φτάνει σε αδιέξodo

και περιμένει την ελευθέρωση κάποιας περιοχής. Συνήθως τέτοια εμπόδια είναι δυναμικής φύσεως, δηλαδή πρόκειται για κάποιον περαστικό, ο οποίος μετά από μερικά δευτερόλεπτα θα έχει φύγει από το οπτικό πεδίο της κάμερας κι έτσι ο αλγόριθμος θα ανακαλύψει το ελεύθερο μονοπάτι που υπάρχει. Η προτεραιότητα που δίνεται στην αριστερή περιοχή είναι καθαρά αυθαίρετη επιλογή και μπορεί κάλλιστα να τροποποιηθεί.

Η ανίχνευση των εμποδίων σε κάποια υπο-περιοχή γίνεται χρησιμοποιώντας την μέθοδο blob detection, που περιγράφηκε στην προηγούμενη ενότητα. Πιο συγκεκριμένα, η μοναδική παράμετρος που χρησιμοποιείται για την ανίχνευση εμποδίου είναι το μέγεθός του. Δεδομένου ότι η εικόνα βάθους που λαμβάνεται από την κάμερα είναι σε κλίμακα του γκρι και οι κοντινότερες περιοχές αναπαριστώνται με φωτεινότερες αποχρώσεις, η μέθοδος blob detection απομονώνει τα pixels που έχουν τιμή μεγαλύτερη από ένα συγκεκριμένο κατώφλι (π.χ. μεγαλύτερη από 100) και στη συνέχεια αναζητά ομάδες από pixels που σχηματίζουν ένα ενιαίο αντικείμενο με εμβαδόν ανάμεσα στα όρια που έχουν θεσπιστεί στον αλγόριθμο. Αν εντοπιστεί κάποιο αντικείμενο που να πληρεί τα συγκεκριμένα κριτήρια, τότε συνεπάγεται ότι αυτό είναι εμπόδιο για τον χρήστη.



Σχήμα 3.42: Διαίρεση της εικόνας βάθους σε 3 τμήματα



Σχήμα 3.43: Overview αλγορίθμου εντοπισμού εμποδίων

3.7 Επικοινωνία μεταξύ των συσκευών

Σύμφωνα με την αρχιτεκτονική του προτεινόμενου συστήματος, που παρουσιάστηκε στο σχήμα 3.1, απαιτείται η επικοινωνία μεταξύ του smartphone του χρήστη και της μονάδας επεξεργασίας, δηλαδή του Raspberry Pi. Το smartphone αποτελεί το μέσο παροχής ανάδρασης στον χρήστη, μέσα από το οποίο δέχεται ειδοποιήσεις σχετικά με την διαδρομή που πρέπει να ακολουθήσει, την ύπαρξη εμποδίων και την δυνατότητα διάσχισης μιας διάβασης πεζών. Επομένως, είναι απαραίτητη η διασύνδεση της επεξεργαστικής μονάδας, που υλοποιεί τους αλγορίθμους επεξεργασίας εικόνας, με την μονάδα παροχής ανάδρασης. Πιο συγκεκριμένα, το Raspberry Pi στέλνει συγκεκριμένα σήματα ελέγχου στην smartphone, τα οποία ερμηνεύονται κατάλληλα από την εφαρμογή Android και μεταφράζονται στα αντίστοιχα μοτίβα δονήσεων.

3.7.1 Απαιτήσεις συστήματος

Η βασική προϋπόθεση είναι η ασύρματη επικοινωνία, η οποία να επιτρέπει την αμφίδρομη αποστολή/λήψη μηνυμάτων από τις δύο πλευρές. Η κύρια ροή μηνυμάτων είναι από το Raspberry Pi προς το smartphone, ωστόσο είναι αναγκαία και η αντίστροφη διαδρομή, από το smartphone προς το Raspberry Pi, για την λήψη μηνυμάτων επιβεβαίωσης (confirmation messages).

Η αρχική σκέψη του συγγραφέα ήταν η υλοποίηση μιας αμφίδρομης σύνδεσης Bluetooth μεταξύ των δύο συσκευών. Κάτι τέτοιο αποδείχθηκε αρκετά προβληματικό, ωστόσο, αφού τόσο το Raspberry Pi 4 όσο και το smartphone χρησιμοποιούσαν το πρωτόκολλο Bluetooth 5, το οποίο είναι αρκετά δύσχρηστο για περιπτώσεις όπου χρειάζεται η υλοποίηση μιας απλής σύνδεσης. Επιπλέον, η

ξεχωριστή υλοποίηση της Bluetooth σύνδεσης τόσο σε C++ (Raspberry Pi) όσο και σε Java (Android App), αποδείχθηκε πολύ χρονοβόρα, οπότε απορρίφθηκε η λύση του Bluetooth. Η εναλλακτική που επιλέχθηκε ήταν η χρήση του δικτυακού προγραμματισμού (*socket programming*).

3.7.2 Socket Communication

Η υλοποίηση μιας ασύρματης σύνδεσης με χρήση των λεγόμενων sockets απαιτεί την ύπαρξη ενός τοπικού δικτύου, στο οποίο να είναι συνδεδεμένες οι συσκευές που πρόκειται να επικοινωνήσουν. Για την εγκαθίδρυση ενός τοπικού δικτύου συνδέουμε και τις δύο συσκευές (smartphone & raspberry pi) σε ένα δίκτυο Wifi. Για τις ανάγκες της διπλωματικής εργασίας οι δύο συσκευές συνδέθηκαν σε ένα δίκτυο wifi-hotspot που παρεχόταν από ένα τρίτο smartphone. Μέσω του ίδιου δικτύου οι συσκευές είχαν πρόσβαση στο διαδίκτυο. Παρακάτω αναλύεται περισσότερο η δομή μιας σύνδεσης μέσω sockets [74].

3.7.2.1 Sockets

Τα sockets (ή αλλιώς υποδοχές στα ελληνικά) παρέχουν σημειακή, αμφίδρομη επικοινωνία μεταξύ διεργασιών, συνήθως -αλλά όχι απαραίτητα- μέσω δικτύου. Αποτελούν μια πολύ ευέλικτη και θεμελιώδη μέθοδο δικτυακής επικοινωνίας μεταξύ διεργασιών ή συστημάτων. Ένα socket είναι μια τερματική σύνδεση η οποία μπορεί να λάβει ένα όνομα (name). Επίσης έχει ένα τύπο (type) και σχετίζεται με μια ή περισσότερες διεργασίες. Τα sockets, ανάλογα με την εφαρμογή που εξυπηρετούν, χαρακτηρίζονται από ένα πρωτόκολλο. Όταν τα sockets χρησιμοποιούνται για επικοινωνία μέσω ενός δικτύου, όπως το internet, αξιοποιούν το πρωτόκολλο επικοινωνίας TCP/IP. Η διεύθυνση ενός socket στο internet αποτελείται από μια διεύθυνση IP και έναν αριθμό θύρας (port).

3.7.2.2 Τύποι sockets: TCP/UDP

Για να γίνει επικοινωνία μεταξύ δύο τερματικών σημείων (endpoints) πρέπει αυτά να χρησιμοποιούν sockets του ίδιου τύπου. Οι δύο βασικοί και πιο διαδεδομένοι τύποι sockets είναι οι ακόλουθοι:

- *Stream socket*: Παρέχει αμφίδρομη, ακολουθιακή, αξιόπιστη και μη-επαναλαμβανόμενη ροή δεδομένων χωρίς καθορισμένα όρια μηνύματος. Η επικοινωνία μοιάζει με τηλεφωνική επικοινωνία και χρησιμοποιεί το πρωτόκολλο TCP.
- *Datagram socket*: Υποστηρίζει αμφίδρομη ροή μηνυμάτων (datagrams). Η σειρά αποστολής και παραλαβής μηνυμάτων μπορεί να μην είναι ακολουθιακή. Τα όρια των μηνυμάτων είναι καθορισμένα. Η επικοινωνία μοιάζει με ανταλλαγή επιστολών μέσω ταχυδρομείου και χρησιμοποιεί το πρωτόκολλο UDP.

Η σύνδεση που υλοποιήθηκε στα πλαίσια της διπλωματικής χρησιμοποιεί τα stream sockets, που βασίζονται στο TCP, επειδή είναι αναγκαία η αξιόπιστη και μη-επαναλαμβανόμενη μετάδοση των σημάτων ελέγχου από το Raspberry Pi προς το smartphone.

3.7.3 Server/Client

Η επικοινωνία ακολουθεί το μοντέλο εξυπηρετητή/πελάτη (server/client), δηλαδή το ένα τερματικό σημείο λειτουργεί ως εξυπηρετητής και είναι υπεύθυνο για την δημιουργία της σύνδεσης, ενώ το άλλο λειτουργεί ως πελάτης και συνδέεται στην σύνδεση που έχει δημιουργηθεί από έναν εξυπηρετητή. Είναι δυνατόν να υπάρχουν πολλοί διαφορετικοί πελάτες που συνδέονται στον ίδιο εξυπηρετητή και αφού εγκαθιδρυθεί η σύνδεση, τότε η επικοινωνία γίνεται και προς τις δύο πλευρές. Μια σύνδεση μέσω sockets μπορεί να τερματιστεί και από τις δύο πλευρές.

3.7.3.1 Server

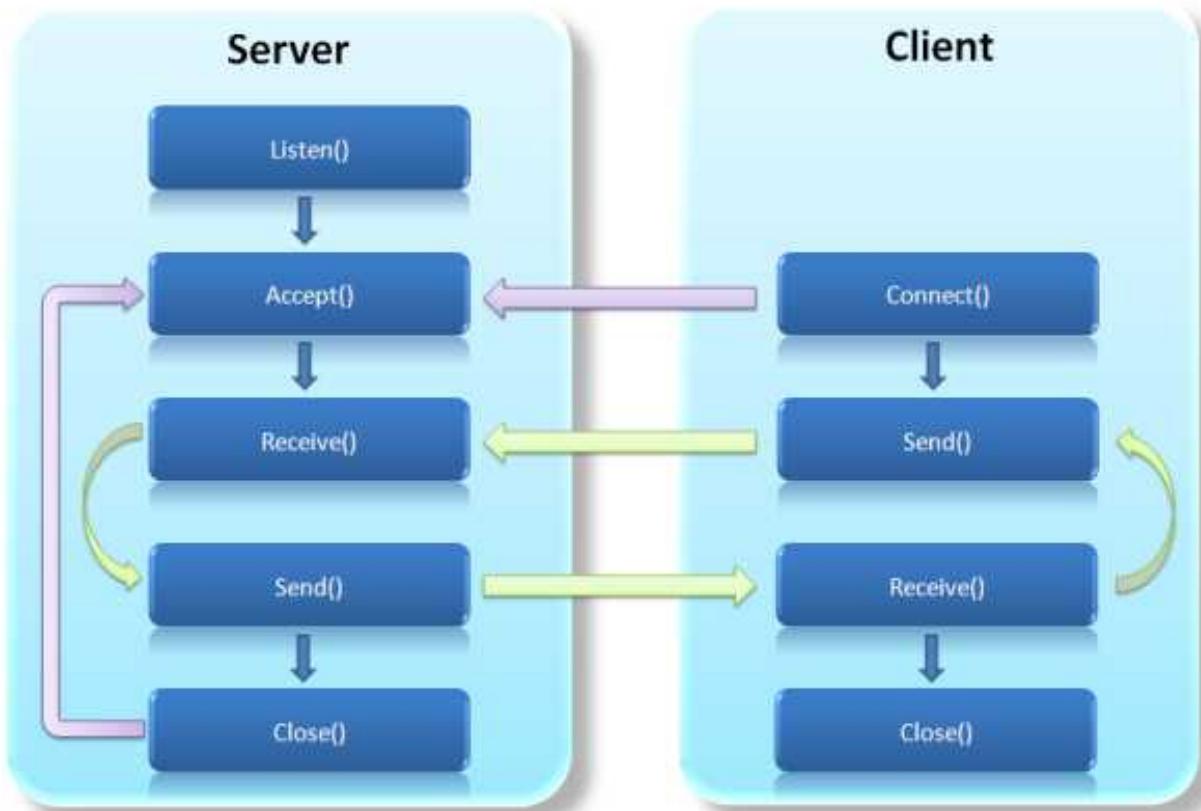
Τον ρόλο του server λαμβάνει το Raspberry Pi, δηλαδή είναι υπεύθυνο για την δημιουργία της σύνδεσης. Τα βήματα που υλοποιεί ένας server για τη σύνδεση είναι τα εξής:

1. Δημιουργεί ένα socket.
2. Δεσμεύει το socket με την επιθυμητή δικτυακή διεύθυνση του τελικού σημείου.
3. Ζητά από το λειτουργικό σύστημα να “ακούει” για εισερχόμενες συνδέσεις.
4. Εισέρχεται σε κατάσταση όπου αναμένει διαρκώς να του παραδώσει μια σύνδεση το λειτουργικό σύστημα.
5. Αν λάβει αίτημα εισερχόμενης σύνδεσης, τότε στέλνει πίσω μήνυμα επιβεβαίωσης και δημιουργεί την σύνδεση.

3.7.3.2 Client

Τον ρόλο του μοναδικού client παίζει το smartphone του χρήστη με την εγκατεστημένη εφαρμογή Android. Τα βήματα που ακολουθεί ο client για να συνδεθεί σε ένα socket είναι τα εξής:

1. Δημιουργεί ένα socket.
2. Επιλέγει την διεύθυνση IP του server και προσδιορίζει την επιθυμητή θύρα (port).
3. Υποβάλλει αίτηση σύνδεσης στον server.
4. Με το που λάβει το μήνυμα επιβεβαίωσης εγκαθιδρύεται η σύνδεση.



Σχήμα 3.44: Μοντέλο Server/Client για μια σύνδεση TCP μέσω sockets

4. Πειραματικός Έλεγχος Συστήματος

Στην ενότητα αυτή παρουσιάζονται μερικά αποτελέσματα που αφορούν τον πειραματικό έλεγχο ορθής λειτουργίας του συστήματος πλοϊγησης που προτάθηκε. Αρχικά, πρέπει να το τονιστεί το γεγονός πως οι χρονικοί περιορισμοί της διπλωματικής εργασίας δεν επέτρεψαν την διενέργεια ενός επίσημου τεστ χρηστικότητας σε πραγματικούς χρήστες με προβλήματα όρασης. Παρόλα αυτά έγινε μια προσπάθεια ελέγχου κάθε ξεχωριστού τμήματος της διπλωματικής πριν την ενσωμάτωσή του στο υπόλοιπο σύστημα, όπως επίσης και συνολικός έλεγχος ολόκληρου του συστήματος. Για την πληρέστερη περιγραφή έγινε καταγραφή ενός βίντεο επίδειξης (demo), στο οποίο παρουσιάζονται οι δυνατότητες του υλοποιηθέντος συστήματος πλοϊγησης. Το βίντεο μπορείτε να το βρείτε στο ακόλουθο link <https://youtu.be/2TiHW27NSLQ>.

4.1 Απόδοση αλγορίθμων

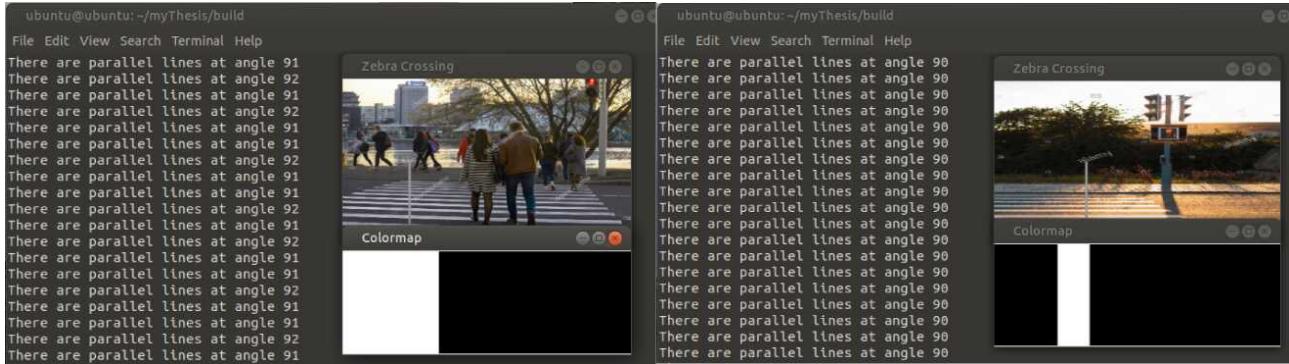
Κάθε αλγόριθμος που παρουσιάστηκε στην ενότητα της υλοποίησης δοκιμάζεται και ελέγχεται κατά πόσο λειτουργεί αποτελεσματικά και σύμφωνα με τις προδιαγραφές που ορίστηκαν. Είναι σημαντικό να τονίσουμε ότι, αν και στην αρχή το πρόγραμμα είχε πολύ χαμηλά fps (frames per second), μετά από κατάλληλες τροποποιήσεις στον κώδικα και στην λογική των αλγορίθμων η ταχύτητα κυμαίνεται στο εύρος 20-30fps, το οποίο είναι ένα πολύ καλό νούμερο για εφαρμογές πραγματικού χρόνου.

4.1.1 Εντοπισμός διάβασης πεζών

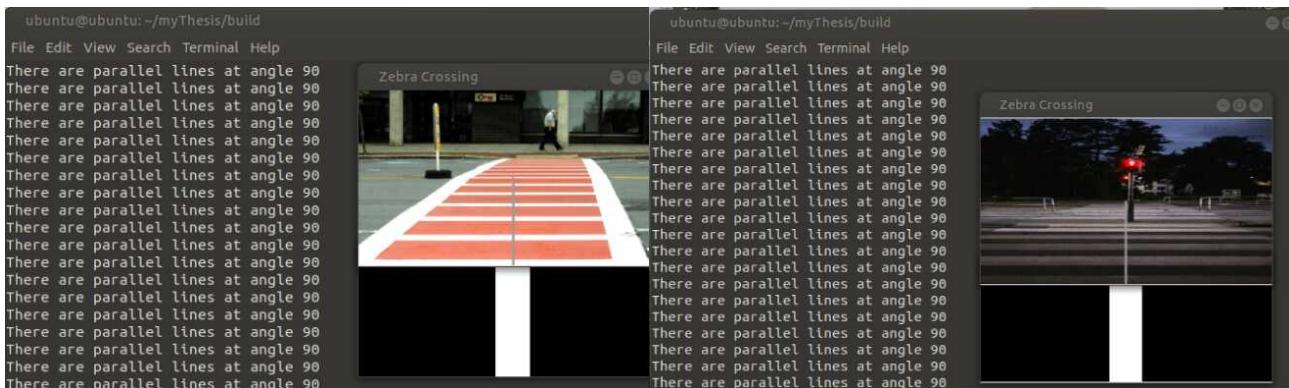
Ο αλγόριθμος εντοπισμού διαβάσεων λειτουργεί αρκετά αποτελεσματικά κάτω από φυσιολογικές συνθήκες φωτισμού και τα πάει αρκετά καλά ακόμα και σε πιο ακραίες συνθήκες, όπως βρεγμένος δρόμος ή αντανακλάσεις από τον ήλιο. Είναι γεγονός πως σε συνθήκες πολύ χαμηλού φωτισμού, όπως π.χ. όταν επικρατεί σκοτάδι, δεν μπορεί να λειτουργήσει αποδοτικά. Παρακάτω, παρατίθενται παραδείγματα διαβάσεων που αναγνωρίζονται από τον προτεινόμενο αλγόριθμο, καθώς και μερικές περιπτώσεις στις οποίες ”εξαπατάται” από το οπτικό περιβάλλον, όπως για παράδειγμα στην ύπαρξη σκαλιών.



Σχήμα 4.1: Παράδειγμα εντοπισμού διάβασης πεζών. Η γκρίζα ευθεία γραμμή πάνω στη διάβαση, καθώς και η λευκή περιοχή κάτω από την διάβαση υποδηλώνουν το βέλτιστο μονοπάτι που πρέπει να ακολουθήσει ο χρήστης. Επίσης, το πρόγραμμα εμφανίζει και την κλίση των γραμμών της διάβασης σε μοίρες. Απλή διάβαση (αριστερά), βρεγμένη διάβαση (δεξιά).



Σχήμα 4.2: Παράδειγμα εντοπισμού διάβασης πεζών με παρουσία εμποδίων (αριστερά) ή αντανάκλασης ηλιακού φωτός (δεξιά)

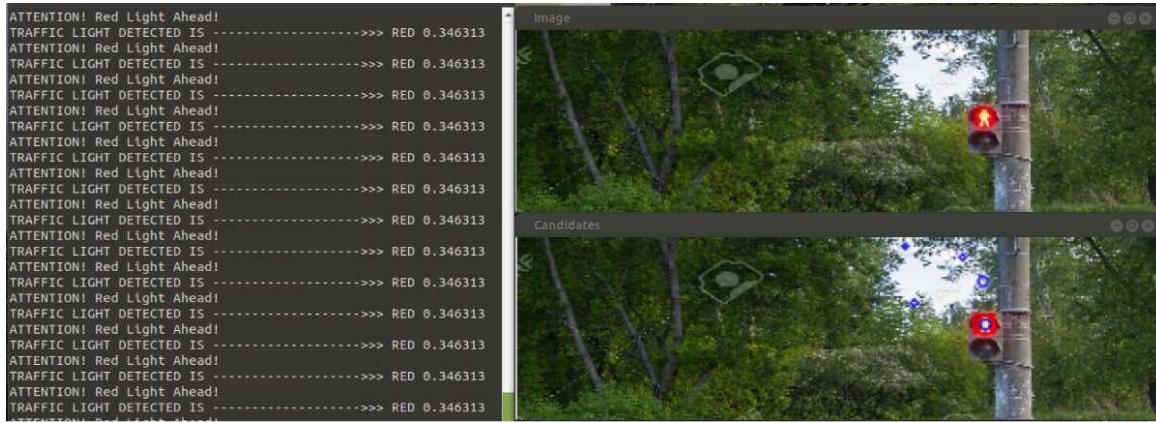


Σχήμα 4.3: Παράδειγμα εντοπισμού διάβασης πεζών άλλων χρωματικών αποχρώσεων (αριστερά) ή σε σκοτεινό περιβάλλον (δεξιά)

4.1.2 Αναγνώριση φωτεινού σηματοδότη

Ο αλγόριθμος αναγνώρισης φωτεινού σηματοδότη λειτουργεί επαρκώς καλά κάτω από συγκεκριμένες συνθήκες. Η αποτελεσματικότητά του επηρεάζεται άμεσα από την ένταση της φωτεινότητας του φαναριού, την ανάλυση της κάμερας, την γωνία λήψης της φωτογραφίας και τις εξωτερικές

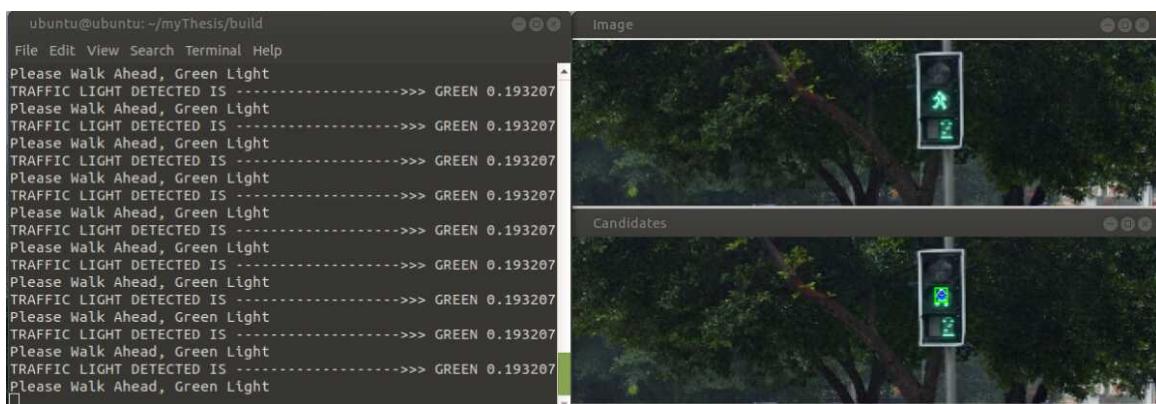
συνθήκες περιβάλλοντος, π.χ. συννεφιά, βροχή κλπ. Ωστόσο, η χρήση του κάτω από συνηθισμένες συνθήκες δίνει ικανοποιητικά αποτελέσματα.



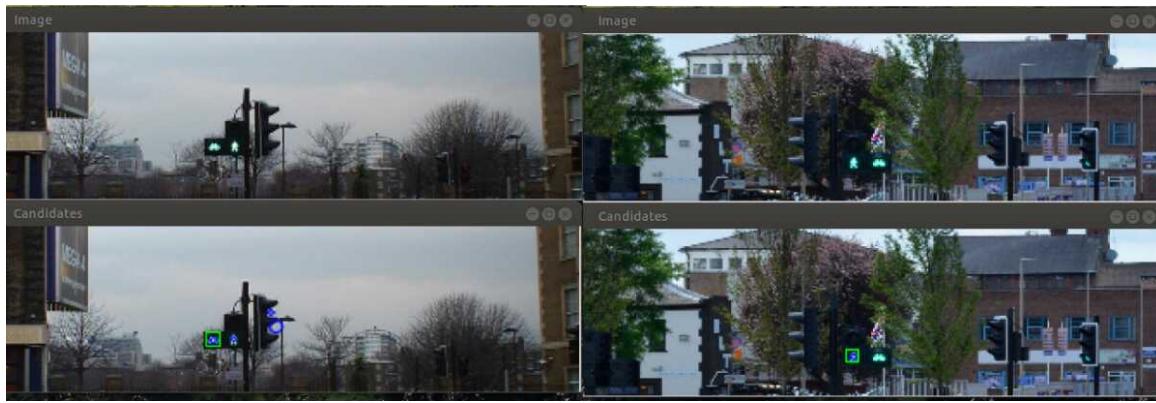
Σχήμα 4.4: Παράδειγμα αναγνώρισης κόκκινου φωτεινού σηματοδότη (κόκκινο περίγραμμα γύρω από το φανάρι). Οι μπλε κύκλοι υποδηλώνουν πιθανές περιοχές με φωτεινό σηματοδότη που εντοπίστηκαν από τον αλγόριθμο, αλλά δεν πέρασαν τη φάση της επικύρωσης (validation). Τα νούμερα που φαίνονται αριστερά στην κονσόλα είναι η τιμή του μέτρου σύγκρισης κατά την εφαρμογή του template matching.



Σχήμα 4.5: Παραδείγματα αναγνώρισης κόκκινου φωτεινού σηματοδότη (κόκκινο περίγραμμα γύρω από το φανάρι).



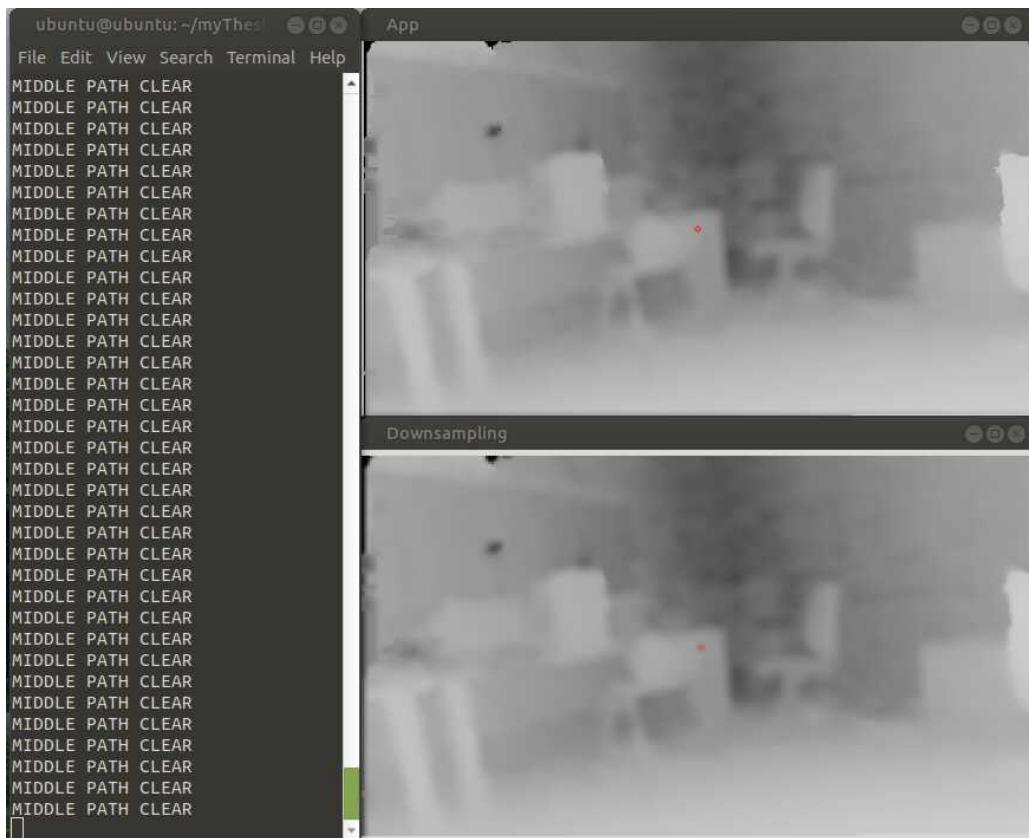
Σχήμα 4.6: Παράδειγμα αναγνώρισης πράσινου φωτεινού σηματοδότη (πράσινο περίγραμμα γύρω από το φανάρι).



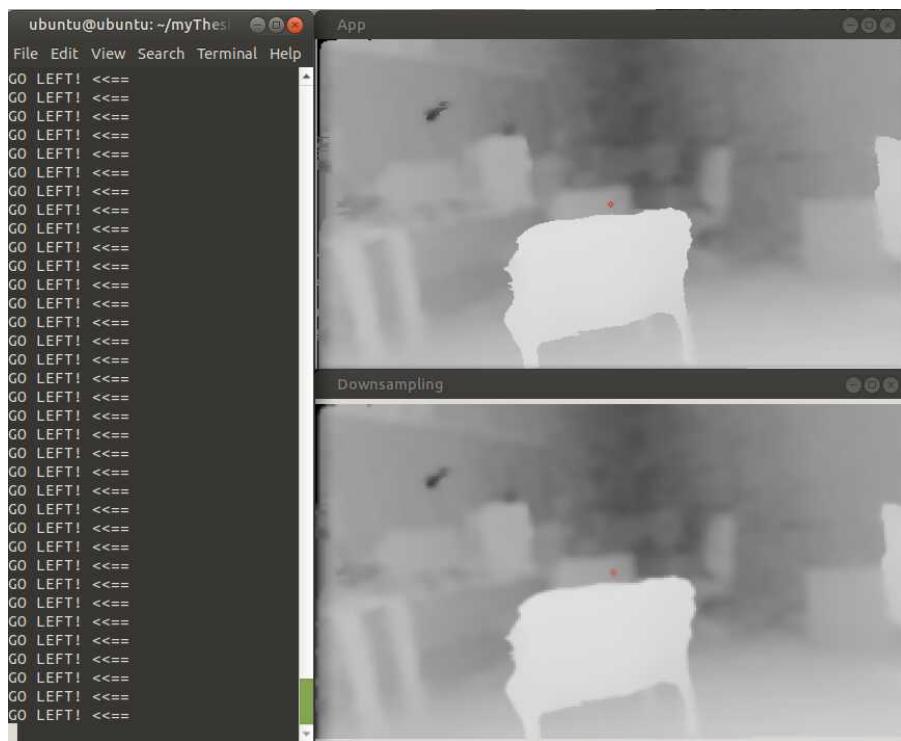
Σχήμα 4.7: Παραδείγματα αναγνώρισης πράσινου φωτεινού σηματοδότη (πράσινο περίγραμμα γύρω από το φανάρι). Οι μπλε κύκλοι υποδηλώνουν πιθανές περιοχές με φωτεινό σηματοδότη που εντοπίστηκαν από τον αλγόριθμο, αλλά δεν πέρασαν τη φάση της επικύρωσης (validation). Τα νούμερα που φαίνονται αριστερά στην κονσόλα είναι η τιμή του μέτρου σύγκρισης κατά την εφαρμογή του template matching.

4.1.3 Αποφυγή εμποδίων

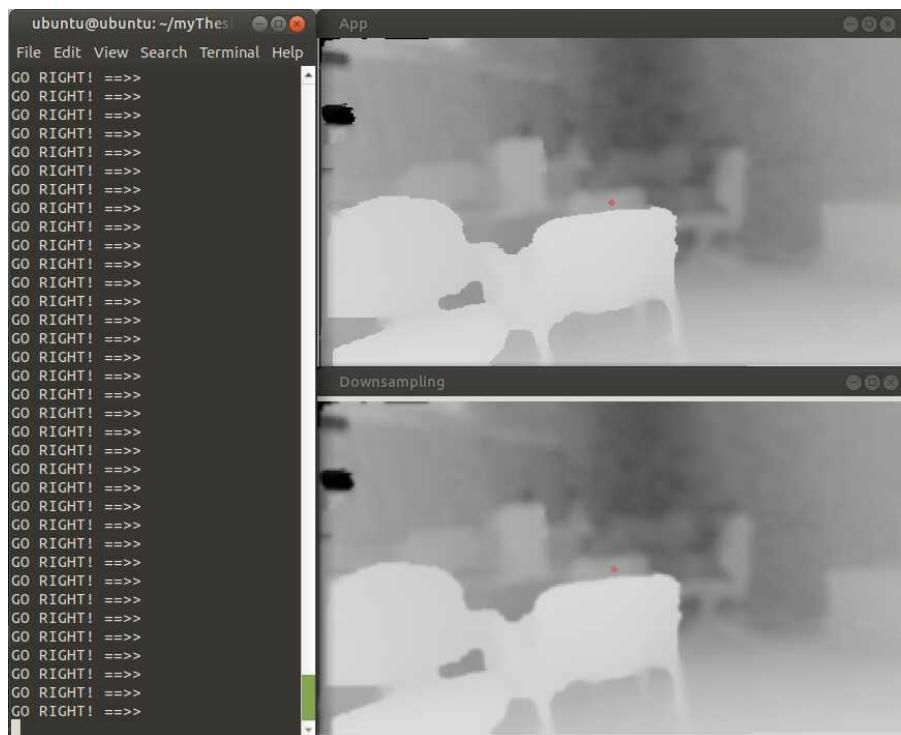
Ο αλγόριθμος αποφυγής εμποδίων είναι ο πιο απλός αλγόριθμος του συστήματος. Λειτουργεί αποτελεσματικά στις περισσότερες περιπτώσεις, ενώ είναι ιδιαίτερα ανθεκτικός σε συνθήκες φωτισμού και εξωτερικού περιβάλλοντος.



Σχήμα 4.8: Παράδειγμα εντοπισμού εμποδίων. Η κεντρική περιοχή της εικόνας βάθους είναι κενή, άρα το σύστημα ειδοποιεί τον χρήστη να προχωρήσει ελεύθερα.



Σχήμα 4.9: Παράδειγμα εντοπισμού εμποδίων. Η κεντρική περιοχή της εικόνας βάθους είναι κατειλημμένη από εμπόδιο, άρα το σύστημα ειδοποιεί τον χρήστη να μετακινηθεί αριστερά στον κενό χώρο.



Σχήμα 4.10: Παράδειγμα εντοπισμού εμποδίων. Τόσο η κεντρική, όσο και η αριστερή περιοχή της εικόνας βάθους είναι κατειλημμένες από εμπόδια, άρα το σύστημα ειδοποιεί τον χρήστη να μετακινηθεί δεξιά στον κενό χώρο.

4.2 Απτική ανάδραση

Η χρήση απτικής ανάδρασης για την ειδοποίηση του χρήστη είναι από τα πιο δυνατά χαρακτηριστικά του συστήματος. Παρέχει διακριτικότητα και δίνει στον χρήστη την άνεση να μετακινείται στον χώρο χωρίς να απαιτείται η χρήση κάποιου ακουστικού. Τα μοτίβα δονήσεων (haptic icons), όπως παρουσιάστηκαν στην ενότητα 3.4.4, είναι ευδιάκριτα μεταξύ τους χάρη στον περιορισμένο αριθμό τους και στην χρήση κατάλληλων συνδυασμών παύσης, δόνησης και έντασης σε κάθε ξεχωριστό μοτίβο. Το γεγονός αυτό καθιστά την εκμάθησή τους εύκολη και γρήγορη, ενώ αναδεικνύεται η καταλληλότητά τους για ένα σύστημα πλοιόγησης για άτομα με προβλήματα όρασης.

4.3 Τελική διάταξη

Η τελική διάταξη του πειραματικού συστήματος περιλαμβάνει την ενσωμάτωση όλου του εξοπλισμού (Raspberry Pi 4, Camera, Powerbank) σε ένα τσαντάκι μέσης, που λειτουργεί ως μέσο στήριξης, για την εύκολη και άνετη χρήση του από το άτομο με προβλήματα όρασης. Στο σχήμα 4.11 φαίνεται η τελική μορφή του συστήματος πάνω στον χρήστη. Ο χρήστης, πέρα από το τσαντάκι μέσης με τον εξοπλισμό, κρατάει και το smartphone από το οποίο δέχεται ανάδραση.



Σχήμα 4.11: Τελική διάταξη προτεινόμενου συστήματος

5. Συμπεράσματα - Προτάσεις για το μέλλον

Στην ενότητα αυτή παρουσιάζονται συνοπτικά τα πλεονεκτήματα και τα μειονεκτήματα του συστήματος πλοϊγησης που προτείνεται στην εργασία μας. Παράλληλα, προτείνονται ορισμένες βελτιώσεις που αφορούν κάποια χαρακτηριστικά του συστήματος με σκοπό την συνολική αύξηση της απόδοσης, αλλά και της χρηστικότητάς του. Στόχος της παρούσας διπλωματικής εργασίας ήταν η υλοποίηση ενός διακριτικού συστήματος πλοϊγησης για άτομα με προβλήματα όρασης, το οποίο θα υποβοηθούσε τον χρήστη να διασχίζει τους δρόμους και να μετακινείται με περισσότερη άνεση και ασφάλεια μέσα σε ένα αστικό περιβάλλον με παράλληλη χρήση του λευκού μπαστουνιού. Ο στόχος αυτός επετεύχθη και, παρ' όλες τις ατέλειες του συστήματος, η τελική πειραματική υλοποίησή του είναι λειτουργική και έρχεται πολύ κοντά στις αρχικές φιλοδοξίες του συγγραφέα.

5.1 Πλεονεκτήματα

Το κύριο πλεονέκτημα του συστήματος είναι η ενσωμάτωση διαφορετικών λειτουργιών και η φορητότητά του. Πιο συγκεκριμένα, συνδυάζει τρία επίπεδα καθοδήγησης, την πλοϊγηση turn-by-turn, την αποφυγή εμποδίων και την υποβοήθηση διάσχισης διάβασης πεζών. Οι τρεις αυτές λειτουργίες παρέχονται ταυτόχρονα στον χρήστη καθ' όλη την διάρκεια πλοϊγησής του. Επιπλέον, το σύστημα χαρακτηρίζεται από διακριτικότητα όσον αφορά τον όγκο του εξοπλισμού που χρησιμοποιεί και ελάχιστη επεμβατικότητα στις συνήθειες και τη συμπεριφορά του χρήστη, καθώς όλα τα μέρη ενσωματώνονται σε ένα τσαντάκι μέσης.

Παράλληλα, η υλοποίηση που προτείνεται είναι αρκετά φιλική προς τον χρήστη και υποστηρίζει την ελάχιστη παρεμβατικότητα, ενώ αξιοποιεί την λογική plug-and-play, δηλαδή ο χρήστης απλά φοράει το τσαντάκι μέσης και ενεργοποιεί το σύστημα μέσω του κινητού του, χωρίς να χρειάζεται περαιτέρω ρύθμιση. Η χρήση των μοτίβων δονήσεων (haptic icons) είναι αρκετά απλή και κατανοητή, μιας και ο αριθμός τους είναι μικρός, μόλις 4 διαφορετικά μοτίβα, και η διαδικασία εκμάθησής τους γρήγορη και εύκολη.

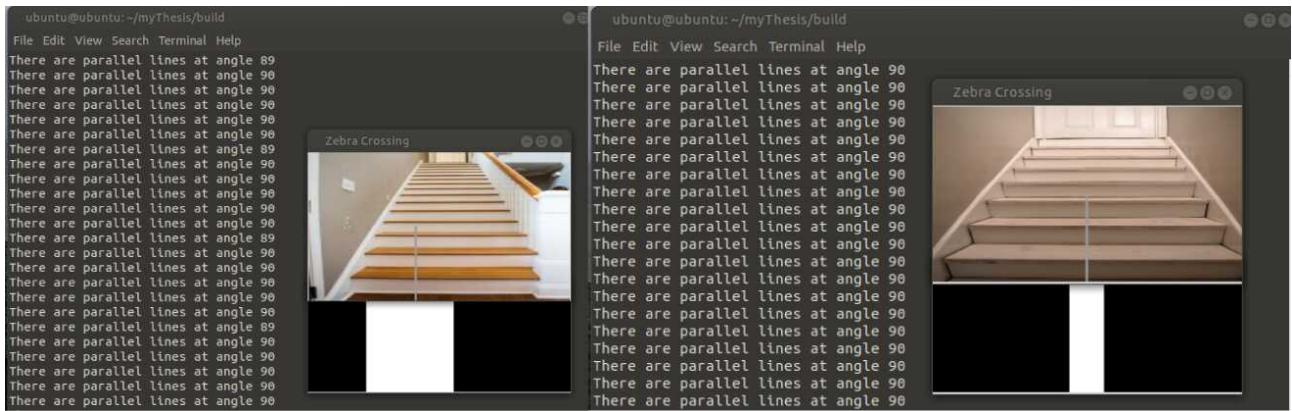
5.2 Μειονεκτήματα

Στα πλαίσια εκπόνησης μιας διπλωματικής εργασίας δεν υπάρχει ο απαραίτητος χρόνος για την υλοποίηση ενός πλήρους και ολοκληρωμένου συστήματος πλοϊγησης που θα λειτουργεί απρόσκοπτα χωρίς προβλήματα. Παρόλο, λοιπόν, που το προτεινόμενο σύστημα είναι λειτουργικό, παρουσιάζει όπως είναι φυσικό ορισμένα μειονεκτήματα τα οποία είναι κυρίως αλγορίθμικής φύσεως. Από την άλλη πλευρά, το μοναδικό ίσως μειονέκτημα δομικής φύσεως είναι η χαμηλή απόδοση του συστήματος σε συνθήκες νύχτας, μιας και βασίζεται στην χρήση κάμερας και η ύπαρξη φωτισμού είναι αναγκαία.

5.2.1 Ανίχνευση διάβασης πεζών

Καταρχάς, ο αλγόριθμος ανίχνευσης διάβασης πεζών φαίνεται να είναι επιρρεπής στις περιβαλλοντικές συνθήκες, στις συνθήκες του οδοστρώματος (βρεγμένο ή όχι), ενώ εξαρτάται επίσης και από

τον βαθμό συντήρησης της διάβασης πεζών (κατά πόσο έντονη είναι η διαγράμμιση). Μετά από κάποιες δοκιμές, ωστόσο, βρέθηκε ότι η μεγαλύτερη αδυναμία του συγκεκριμένου αλγορίθμου είναι η παραπλάνησή του από εικόνες που περιέχουν σκαλοπάτια, αντί για διάβαση πεζών. Εξαιτίας του γεγονότος ότι αξιοποιείται το εναλλασσόμενο μοτίβο των δύο γραμμών για την αναγνώριση της διάβασης, ο αλγόριθμος βγάζει λάθος συμπεράσματα όταν επεξεργάζεται φωτογραφίες που απεικονίζουν σκαλοπάτια, μιας και οι τελευταίες ενσωματώνουν περίπου το ίδιο εναλλασσόμενο μοτίβο, όπως φαίνεται στο σχήμα 5.1.



Σχήμα 5.1: Παράδειγμα παραπλάνησης αλγορίθμου. Το πρόγραμμα αναγνωρίζει λανθασμένα κάποια σκαλοπάτια ως διάβαση πεζών. Αυτό συμβαίνει επειδή τα μοτίβα των σκαλοπατιών και της διάβασης είναι παρόμοια μεταξύ τους.

5.2.2 Αναγνώριση φωτεινού σηματοδότη

Όσον αφορά τον αλγόριθμο αναγνώρισης φωτεινού σηματοδότη, η βασική του αδυναμία έγκειται στο γεγονός πως το τελικό αποτέλεσμα εξαρτάται σε πολύ μεγάλο ποσοστό από τον τύπο των templates που χρησιμοποιούνται για την αντιστοίχηση του χρώματος του φαναριού. Τέλος, η φύση του αλγορίθμου αυτού τον καθιστά αρκετά ευάλωτο στις περιβαλλοντικές συνθήκες φωτισμού. Σχετικό παράδειγμα παρουσιάζεται στο σχήμα 5.2.



Σχήμα 5.2: Παράδειγμα αστοχίας αλγορίθμου. Το πρόγραμμα αναγνωρίζει λανθασμένα κάποια σημεία της εικόνας ως φανάρια και τα ερμηνεύει με κόκκινο χρώμα. Αυτό συμβαίνει επειδή υπάρχουν έντονες αντανακλάσεις του ηλιακού φωτός, ενώ αντίστοιχα η ένταση της φωτεινότητας του φαναριού είναι χαμηλή.

5.3 Μελλοντικές προτάσεις

Οι παρακάτω προτάσεις αφορούν την βελτίωση ορισμένων χαρακτηριστικών του συστήματος τα οποία θα αυξήσουν άμεσα την απόδοση και την αποτελεσματικότητά του και τα οποία δεν μπόρεσαν να υλοποιηθούν, κυρίως, λόγω έλλειψης χρόνου.

5.3.1 Μετατροπέας Voice-to-Text

Μια αναγκαία προσθήκη που απαιτείται για να κάνει το παραπάνω σύστημα πλήρως προσβάσιμο σε άτομα με προβλήματα όρασης είναι η ενσωμάτωση ενός μετατροπέα φωνής σε κείμενο (Voice-to-Text) και, αντίστροφα, κειμένου σε φωνή (Text-to-Voice). Με τον τρόπο αυτό ο χρήστης θα μπορεί να πλοηγείται στην εφαρμογή πλήρως ανεξάρτητα, ενώ η εισαγωγή του προορισμού θα γίνεται μέσω φωνητικών εντολών του χρήστη. Για την υλοποίηση του τελευταίου, υπάρχει ήδη διαθέσιμο Google API για android development που χρησιμοποιεί τον αλγόριθμο αναγνώριση φωνής της Google.

5.3.2 Αναγνώριση εμποδίων

Δεύτερη σημαντική προσθήκη είναι η βελτίωση του υπάρχοντος αλγορίθμου εντοπισμού εμποδίων. Πιο συγκεκριμένα, για να είναι πιο αποδοτικός θα πρέπει να χρησιμοποιεί πιο εξελιγμένη λογική στην ανάλυση της εικόνας βάθους και να ενσωματώνει λειτουργία αναγνώρισης του είδους των εμποδίων που ανιχνεύονται, π.χ. άνθρωπος, κολόνα, αυτοκίνητο κλπ.

5.3.3 Βελτιωμένο template φωτεινού σηματοδότη

Ο αλγόριθμος αναγνώρισης φωτεινού σηματοδότη βασίζεται στην χρήση ενός συγκεκριμένου προτύπου. Ένας τρόπος να αυξήσουμε την απόδοση του αλγορίθμου αυτού είναι να βρεθεί ένα καλύτερο πρότυπο, το οποίο να αντιπροσωπεύει πιο αποτελεσματικά τους φωτεινούς σηματοδότες. Κάτι τέτοιο απαιτεί αντίστοιχη έρευνα για την εξακρίβωση των χαρακτηριστικών εκείνων που παίζουν μεγαλύτερο ρόλο στο template matching.

5.3.4 Αυτονομία

Η αυτονομία ενός συστήματος πλοϊγησης είναι ένα από τα κύρια χαρακτηριστικά που ενδιαφέρει τον τελικό χρήστη, καθώς έχει ανάγκη την, όσο το δυνατόν, μεγαλύτερη διάρκεια συνεχόμενης χρήσης. Η αυτονομία του συστήματος εξαρτάται τόσο από το μέγεθος της φορητής μπαταρίας (powerbank) όσο και από την κατανάλωση του συστήματος. Θεωρώντας δεδομένη την κατανάλωση, θα προτείναμε να γίνει χρήση μιας μεγαλύτερης σε χωρητικότητα μπαταρίας, η οποία θα επιτρέπει την συνεχόμενη χρήση τουλάχιστον 2 ωρών. Για τις ανάγκες της διπλωματικής εργασίας, η αυτονομία της πειραματικής διάταξης ήταν παραπάνω από αρκετή.

5.3.5 Usability test

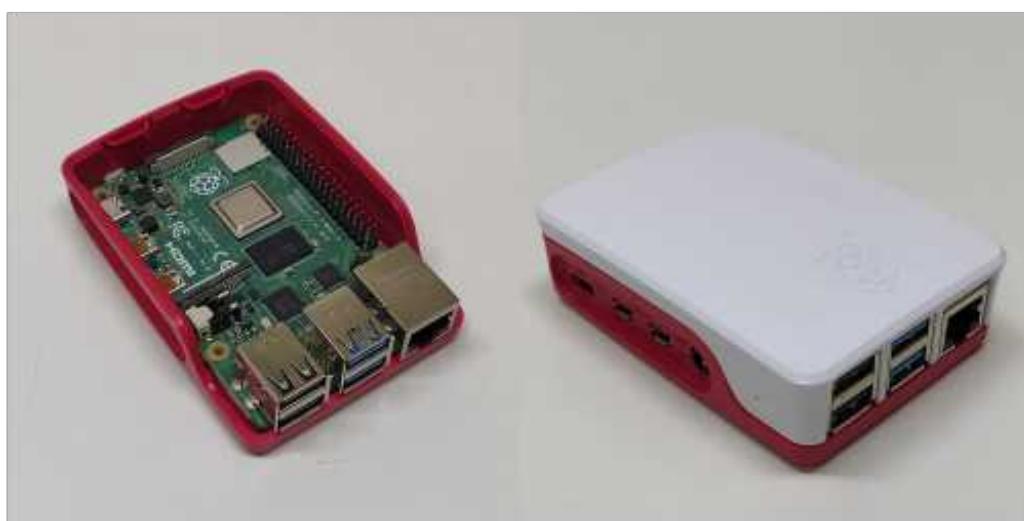
Τέλος, κάθε σύστημα που προορίζεται για χρήση από τελικούς χρήστες είναι απαραίτητο να περνάει από έναν έλεγχο χρηστικότητας (usability test). Συνήθως ο έλεγχος αυτός περιλαμβάνει την χρήση του προς εξέταση συστήματος από τους πραγματικούς χρήστες, δηλαδή από άτομα που πάσχουν από προβλήματα μειωμένης όρασης, και υπάρχουν διάφορες μέθοδοι που μπορούν να αξιοποιηθούν για τον υπολογισμό μιας μετρικής χρηστικότητας, όπως είναι οι 10 ευρετικοί κανόνες του Nielsen [75].

6. Παράρτημα

6.1 Φωτογραφίες εξοπλισμού



Σχήμα 6.1: Η κάμερα Intel Realsense 435i (πάνω) και το Powerbank Intenso PD10000 (κάτω)



Σχήμα 6.2: Το Raspberry Pi 4 χωρίς θήκη (πάνω) και με προστατευτική θήκη (μέση) και τσαντάκι μέσης με όλο τον εξοπλισμό μέσα (κάτω)

Βιβλιογραφία

- [1] “What exactly is the world report on vision?! - iapb,” <https://www.iapb.org/news/what-exactly-is-the-world-report-on-vision/>, (Accessed on 02/12/2020).
- [2] R. Sammouda and A. Alrjoub, “Mobile blind navigation system using RFID,” in *2015 Global Summit on Computer Information Technology (GSCIT)*, Jun. 2015, pp. 1–4.
- [3] M. Miles and J. Jetter, “A lidar-based navigation system for the visually impaired.” International Foundation for Telemetering, 2016.
- [4] S. Weinstein, “Intensive and extensive aspects of tactile sensitivity as a function of body part, sex and laterality,” *The skin senses*, 1968.
- [5] E. S. L. Gastal and M. M. Oliveira, “Domain transform for edge-aware image and video processing,” *ACM TOG*, vol. 30, no. 4, pp. 69:1–69:12, 2011, proceedings of SIGGRAPH 2011.
- [6] X. Wu, R. Hu, and Y. Bao, “Block-Based Hough Transform for Recognition of Zebra Crossing in Natural Scene Images,” *IEEE Access*, vol. 7, pp. 59 895–59 902, 2019.
- [7] “Finding line features | lesson | robot academy,” <https://robotacademy.net.au/lesson/finding-line-features/>, (Accessed on 02/12/2020).
- [8] “10 key principles of user centered design | cognitiveclouds blog,” <https://www.cognitiveclouds.com/insights/key-principles-of-user-centered-design/>, (Accessed on 02/06/2020).
- [9] “Zed stereo camera | stereolabs,” <https://www.stereolabs.com/zed/>, (Accessed on 02/10/2020).
- [10] “Camboard pico flexx,” <https://pmdtec.com/picofamily/flexx/>, (Accessed on 02/10/2020).
- [11] “Depth camera d435 – intel® realsense™ depth and tracking cameras,” <https://www.intelrealsense.com/depth-camera-d435/>, (Accessed on 02/10/2020).
- [12] “Buy a raspberry pi 4 model b – raspberry pi,” <https://www.raspberrypi.org/products/raspberry-pi-4-model-b/>, (Accessed on 02/10/2020).
- [13] W. H. Organization, “Vision impairment and blindness,” October 2019, [Last accessed: January 10, 2020]. [Online]. Available: <https://www.who.int/en/news-room/fact-sheets/detail/blindness-and-visual-impairment>
- [14] R. R. Bourne, S. R. Flaxman, T. Braithwaite, M. V. Cicinelli, A. Das, J. B. Jonas, J. Keeffe, J. H. Kempen, J. Leasher, H. Limburg *et al.*, “Magnitude, temporal trends, and projections of the global prevalence of blindness and distance and near vision impairment: a systematic review and meta-analysis,” *The Lancet Global Health*, vol. 5, no. 9, pp. e888–e897, 2017.
- [15] A. Riazi, F. Riazi, R. Yoosfi, and F. Bahmeei, “Outdoor difficulties experienced by a group of visually impaired Iranian people,” *Journal of Current Ophthalmology*, vol. 28, no. 2, pp. 85–90, Jun. 2016. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S245223251530055X>

- [16] J. Parkin and N. Smithies, “Accounting for the needs of blind and visually impaired people in public realm design,” *Journal of Urban Design*, vol. 17, no. 1, pp. 135–149, 2012. [Online]. Available: <https://doi.org/10.1080/13574809.2012.646139>
- [17] S. R. A. W. Alwi and M. N. Ahmad, “Survey on outdoor navigation system needs for blind people,” in *2013 IEEE Student Conference on Research and Developement*. IEEE, 2013, pp. 144–148.
- [18] H. Social Sciences, “LibGuides: Blind/Visual Impairment: Common Assistive Technologies,” [Last accessed: January 16, 2020]. [Online]. Available: <https://guides.library.illinois.edu/c.php?g=526852&p=3602299>
- [19] A. W. Ng and A. H. Chan, “Finger response times to visual, auditory and tactile modality stimuli,” in *Proceedings of the international multiconference of engineers and computer scientists*, vol. 2, 2012, pp. 1449–1454.
- [20] N. S. Place, “How does gps work?” June 2019, [Last accessed: January 27, 2020]. [Online]. Available: <https://spaceplace.nasa.gov/gps/en/>
- [21] W. F. Kevin Bonsor, “How rfid works,” November 2007, [Last accessed: February 1, 2020]. [Online]. Available: <https://electronics.howstuffworks.com/gadgets/high-tech-gadgets/rfid3.htm>
- [22] ElProCus, “Basic of rfid system - types and working example of rfid application,” November 2018, [Last accessed: February 1, 2020]. [Online]. Available: <https://www.elprocus.com/rfid-basic-introduction-simple-application/>
- [23] M. Kaur, M. Sandhu, N. Mohan, and P. S. Sandhu, “Rfid technology principles, advantages, limitations & its applications,” *International Journal of Computer and Electrical Engineering*, vol. 3, no. 1, p. 151, 2011.
- [24] S. J. Russell and P. Norvig, *Artificial intelligence: a modern approach*. Malaysia; Pearson Education Limited., 2016.
- [25] E. Grossi and M. Buscema, “Introduction to artificial neural networks,” *European journal of gastroenterology hepatology*, vol. 19, pp. 1046–54, 01 2008.
- [26] F. Rosenblatt, “Principles of neurodynamics. perceptrons and the theory of brain mechanisms,” Cornell Aeronautical Lab Inc Buffalo NY, Tech. Rep., 1961.
- [27] B.-S. Lin, C.-C. Lee, and P.-Y. Chiang, “Simple smartphone-based guiding system for visually impaired people,” *Sensors*, vol. 17, no. 6, p. 1371, 2017.
- [28] L. Spallanzani and A. M. Vassalli, *Lettere sopra il sospetto di un nuovo senso nei pipistrelli dell’abate Lazzaro Spallanzani... con le risposte dell’abate Antonmaria Vassalli..* Nella stamperia reale, 1794.
- [29] “The working principle, applications and limitations of ultrasonic sensors,” <https://www.microcontrollertips.com/principle-applications-limitations-ultrasonic-sensors-faq/>, (Accessed on 02/04/2020).
- [30] M. Bousbia-Salah, A. Redjati, M. Fezari, and M. Bettayeb, “An ultrasonic navigation system for blind people,” in *2007 IEEE International Conference on Signal Processing and Communications*. IEEE, 2007, pp. 1003–1006.

- [31] “www.lidar-uk.com/how-lidar-works/,” <http://www.lidar-uk.com/how-lidar-works/>, (Accessed on 02/05/2020).
- [32] Wikipedia contributors, “Lidar — Wikipedia, the free encyclopedia,” <https://en.wikipedia.org/w/index.php?title=Lidar&oldid=938995714>, 2020, [Online; accessed 5-February-2020].
- [33] ——, “Kinect — Wikipedia, the free encyclopedia,” <https://en.wikipedia.org/w/index.php?title=Kinect&oldid=938679274>, 2020, [Online; accessed 5-February-2020].
- [34] K. Dimitrios, “Development of haptic navigation system for blind people,” Ph.D. dissertation, 2017.
- [35] N. A. Giudice, “Navigating without vision: Principles of blind spatial cognition,” in *Handbook of behavioral and cognitive geography*. Edward Elgar Publishing, 2018.
- [36] N. A. Giudice and G. E. Legge, “Blind navigation and the role of technology,” *The engineering handbook of smart technology for aging, disability, and independence*, vol. 8, pp. 479–500, 2008.
- [37] L. J. K. R. G. NA, “Sensory substitution of vision: importance of perceptual and cognitive processing in: Assistive technology for blindness and low vision manduchi r. kurniawan s.(eds) pp. 162–191,” 2012.
- [38] M. M. Billah, Z. M. Yusof, K. Kadir, and A. M. M. Ali, “Sensory substitution for visual impairments: A technological review,” in *Visual Impairment and Blindness*. IntechOpen, 2019.
- [39] A. Bruton, J. H. Conway, and S. T. Holgate, “Reliability: what is it, and how is it measured?” *Physiotherapy*, vol. 86, no. 2, pp. 94–99, 2000.
- [40] C. Saitis and K. Kalimeri, “Identifying urban mobility challenges for the visually impaired with mobile monitoring of multimodal biosignals,” in *International Conference on Universal Access in Human-Computer Interaction*. Springer, 2016, pp. 616–627.
- [41] D. H. Ashmead, D. Guth, R. S. Wall, R. G. Long, and P. E. Ponchillia, “Street crossing by sighted and blind pedestrians at a modern roundabout,” *Journal of Transportation Engineering*, vol. 131, no. 11, pp. 812–821, 2005.
- [42] “Accessible pedestrian signals: Understanding how blind pedestrians cross at signalized intersections,” http://www.apsguide.org/appendix_d_understanding.cfm, (Accessed on 02/07/2020).
- [43] P. Angin, B. K. Bhargava *et al.*, “Real-time mobile-cloud computing for context-aware blind navigation,” *International Journal of Next-Generation Computing*, vol. 2, no. 2, pp. 405–414, 2011.
- [44] “Low vision and the white cane: A tool for fall prevention - visionaware,” <https://www.visionaware.org/info/essential-skills-2/an-introduction-to-orientation-and-mobility-skills/low-vision-and-the-white-cane-a-tool-for-fall-prevention/235>, (Accessed on 02/08/2020).
- [45] Wikipedia contributors, “Sensory substitution — Wikipedia, the free encyclopedia,” https://en.wikipedia.org/w/index.php?title=Sensory_substitution&oldid=927706045, 2019, [Online; accessed 8-February-2020].
- [46] S. Schmidt, C. Tinti, M. Fantino, I. C. Mammarella, and C. Cornoldi, “Spatial representations in blind people: The role of strategies and mobility skills,” *Acta psychologica*, vol. 142, no. 1, pp. 43–50, 2013.

- [47] A. Bharadwaj, “Tactile navigation: An additional processing channel for environments of high sensory load,” Ph.D. dissertation, 2017.
- [48] C. A. Shingledecker and E. Foulke, “A human factors approach to the assessment of the mobility of blind pedestrians,” *Human factors*, vol. 20, no. 3, pp. 273–286, 1978.
- [49] A. Chan, K. MacLean, and J. McGrenere, “Designing haptic icons to support collaborative turn-taking,” *International Journal of Human-Computer Studies*, vol. 66, no. 5, pp. 333–355, 2008.
- [50] H. Seifi and K. E. MacLean, “Exploiting haptic facets: Users’ sensemaking schemas as a path to design and personalization of experience,” *International Journal of Human-Computer Studies*, vol. 107, pp. 38–61, 2017.
- [51] M.-D. Shieh and Z.-B. Wu, “Tactile icon design using a vibration actuator in mobile devices,” in *Asia-Pacific Conference on Computer Human Interaction*. Springer, 2008, pp. 240–247.
- [52] K. A. Kaczmarek, J. G. Webster, P. Bach-y Rita, and W. J. Tompkins, “Electrotactile and vibrotactile displays for sensory substitution systems,” *IEEE transactions on biomedical engineering*, vol. 38, no. 1, pp. 1–16, 1991.
- [53] J.-P. Koo, S.-H. Kim, H.-J. An, O.-G. Moon, J.-H. Choi, Y.-D. Yun, J.-H. Park, and K.-O. Min, “Two-point discrimination of the upper extremities of healthy koreans in their 20’s,” *Journal of physical therapy science*, vol. 28, no. 3, pp. 870–874, 2016.
- [54] “Myo gesture control armband tech specs – welcome to myo support,” <https://support.getmyo.com/hc/en-us/articles/202648103-Myo-Gesture-Control-Armband-tech-specs>, (Accessed on 02/11/2020).
- [55] “Download android studio and sdk tools | android developers,” <https://developer.android.com/studio>, (Accessed on 02/11/2020).
- [56] Wikipedia contributors, “Java (programming language) — Wikipedia, the free encyclopedia,” [https://en.wikipedia.org/w/index.php?title=Java_\(programming_language\)&oldid=939818655](https://en.wikipedia.org/w/index.php?title=Java_(programming_language)&oldid=939818655), 2020, [Online; accessed 11-February-2020].
- [57] ——, “Kotlin (programming language) — Wikipedia, the free encyclopedia,” [https://en.wikipedia.org/w/index.php?title=Kotlin_\(programming_language\)&oldid=939832300](https://en.wikipedia.org/w/index.php?title=Kotlin_(programming_language)&oldid=939832300), 2020, [Online; accessed 11-February-2020].
- [58] “Visual studio ide, code editor, azure devops, & app center - visual studio,” <https://visualstudio.microsoft.com/>, (Accessed on 02/11/2020).
- [59] Wikipedia contributors, “Python (programming language) — Wikipedia, the free encyclopedia,” [https://en.wikipedia.org/w/index.php?title=Python_\(programming_language\)&oldid=939861405](https://en.wikipedia.org/w/index.php?title=Python_(programming_language)&oldid=939861405), 2020, [Online; accessed 11-February-2020].
- [60] “Opencv - wikipedia,” <https://en.wikipedia.org/wiki/OpenCV>, (Accessed on 02/11/2020).
- [61] “Intelrealsense/librealsense: Intel® realsense™ sdk,” <https://github.com/IntelRealSense/librealsense>, (Accessed on 02/11/2020).
- [62] “Install ubuntu server on a raspberry pi 2, 3 or 4 | ubuntu,” <https://ubuntu.com/download/raspberry-pi>, (Accessed on 02/11/2020).
- [63] “Mate desktop environment | mate,” <https://mate-desktop.org/>, (Accessed on 02/11/2020).

- [64] “Depth post-processing for intel® realsense™ d400 depth cameras,” <https://dev.intelrealsense.com/docs/depth-post-processing>, (Accessed on 02/11/2020).
- [65] Wikipedia contributors, “Canny edge detector — Wikipedia, the free encyclopedia,” https://en.wikipedia.org/w/index.php?title=Canny_edge_detector&oldid=930556285, 2019, [Online; accessed 12-February-2020].
- [66] ——, “Sobel operator — Wikipedia, the free encyclopedia,” https://en.wikipedia.org/w/index.php?title=Sobel_operator&oldid=938419054, 2020, [Online; accessed 12-February-2020].
- [67] ——, “Hough transform — Wikipedia, the free encyclopedia,” https://en.wikipedia.org/w/index.php?title=Hough_transform&oldid=936206574, 2020, [Online; accessed 12-February-2020].
- [68] R. De Charette and F. Nashashibi, “Real time visual traffic lights recognition based on spot light detection and adaptive traffic lights templates,” in *2009 IEEE Intelligent Vehicles Symposium*. IEEE, 2009, pp. 358–363.
- [69] H. Tae-Hyun, J. In-Hak, and C. Seong-Ik, “Detection of traffic lights for vision-based car navigation system,” in *Pacific-Rim Symposium on Image and Video Technology*. Springer, 2006, pp. 682–691.
- [70] Wikipedia contributors, “Top-hat transform — Wikipedia, the free encyclopedia,” https://en.wikipedia.org/w/index.php?title=Top-hat_transform&oldid=936434061, 2020, [Online; accessed 13-February-2020].
- [71] “Template matching — opencv 2.4.13.7 documentation,” https://docs.opencv.org/2.4/doc/tutorials/imgproc/histograms/template_matching/template_matching.html, (Accessed on 02/13/2020).
- [72] L. Zacharia and G. Matsopoulos, “Ανάπτυξη Μεθόδων Για Τμηματοποίηση Οδοντιατρικών Δεδομένων.” B.S. thesis, 2013.
- [73] “Opencv: Morphological transformations,” https://docs.opencv.org/3.4/d9/d61/tutorial_py_morphological_ops.html, (Accessed on 02/13/2020).
- [74] Wikipedia contributors, “Network socket — Wikipedia, the free encyclopedia,” https://en.wikipedia.org/w/index.php?title=Network_socket&oldid=940267123, 2020, [Online; accessed 14-February-2020].
- [75] J. Nielsen and R. Molich, “Heuristic evaluation of user interfaces,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI ’90. New York, NY, USA: Association for Computing Machinery, 1990, p. 249–256. [Online]. Available: <https://doi.org/10.1145/97243.97281>

Πανεπιστήμιο Πατρών, Πολυτεχνική Σχολή
Τμήμα Ηλεκτρολόγων Μηχανικών & Τεχνολογίας Υπολογιστών
Ιωάννης Κουτουλογένης του Ανδρέα
© Φεβρουάριος 2020 – Με επιφύλαξη παντός δικαιώματος
