



Διπλωματική Εργασία

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

Τρύφωνα Χριστόδουλου του Πέτρου

Αριθμός Μητρώου: 1046922

Θέμα

**«Αυτόματη αποφυγή σύγκρουσης κινούμενων
οχημάτων με μεθόδους μηχανικής όρασης»**

Επιβλέπων

Ευάγγελος Δερματάς

Αριθμός Διπλωματικής Εργασίας :

Πάτρα, Οκτώβρης 2020

ΠΙΣΤΟΠΟΙΗΣΗ

Πιστοποιείται ότι η Διπλωματική Εργασία με θέμα

«Αυτόματη αποφυγή σύγκρουσης κινούμενων οχημάτων με μεθόδους μηχανικής όρασης»

Του φοιτητή του Τμήματος Ηλεκτρολόγων Μηχανικών και
Τεχνολογίας Υπολογιστών

Τρύφωνα Χριστόδουλου του Πέτρου

Αριθμός Μητρώου: 1046922

Παρουσιάστηκε δημόσια και εξετάστηκε στο Τμήμα
Ηλεκτρολόγων Μηχανικών και Τεχνολογίας Υπολογιστών στις
...../...../.....

Ο Επιβλέπων

Ευάγγελος Δερματάς
Αναπληρωτής Καθηγητής

Ο Διευθυντής του Τομέα

Κυριάκος Σγάρμπας
Αναπληρωτής Καθηγητής

Αριθμός Διπλωματικής Εργασίας:

Θέμα: «Αυτόματη αποφυγή σύγκρουσης κινούμενων οχημάτων με μεθόδους μηχανικής όρασης»

Φοιτητής

Επιβλέπων

Τρύφωνας Χριστόδουλος

Ευάγγελος Δερματάς

Περίληψη

Αντικείμενο της παρούσας εργασίας είναι η ανάπτυξη ενός συστήματος πρόβλεψης σύγκρουσης αυτοκινήτων, βασιζόμενο στην μηχανική όραση καθώς και σε τεχνικές αναγνώρισης προτύπων. Το αρχικό πρόβλημα χωρίζεται σε δύο σκέλη : α) αναγνώριση του προπορευόμενου αυτοκινήτου και β) εντοπισμός της περιοχής του οδοστρώματος.

Στο 1^ο σκέλος εκμεταλλευόμαστε το χαρακτηριστικό του χρώματος για την τμηματοποίηση των κόκκινων αντικειμένων της εικόνας και στην συνέχεια βασιζόμαστε στην ιδιότητα της συμμετρίας, της σχετικής θέσης και της ομοιότητας, για να λάβουμε μόνο τα πίσω φώτα των προπορευόμενων αυτοκινήτων και να αποκλείσουμε τα υπόλοιπα.

Στο 2^ο σκέλος επιχειρούμε να τμηματοποιήσουμε την εικόνα ώστε να εκτιμήσουμε την περιοχή του δρόμου. Πρόκειται λοιπόν για πρόβλημα κατηγοριοποίησης των pixels μίας εικόνας σε δύο κλάσεις: i) δρόμος και ii) μη-δρόμος. Αυτό επιτυγχάνεται μέσω του αλγορίθμου Random Walker σε συνδυασμό με την οπτική ροή και το προηγούμενο αποτέλεσμα ώστε να καθοδηγήσουμε σωστά τον αλγόριθμο και να δώσει αξιόπιστο αποτέλεσμα. Η ποιοτική αξιολόγηση του συστήματος γίνεται με εικόνες και βίντεο που

έχουν ληφθεί από δρόμους της Πάτρας ή από το διαδίκτυο. Απαραίτητη είναι και η ποσοτική αξιολόγηση ώστε να συγκρίνουμε τις επιδόσεις του συστήματος με άλλες προσεγγίσεις που συναντάμε στην βιβλιογραφία. Γι' αυτό χρησιμοποιούμε δημόσιες βάσεις δεδομένων ειδικά σχεδιασμένων για προβλήματα μηχανικής όρασης. Σε κάθε σκέλος παρουσιάζονται αναλυτικά τα βήματα υλοποίησης και παρατίθενται πειραματικά αποτελέσματα, παρατηρήσεις και πιθανές βελτιώσεις.

Abstract

The purpose of this dissertation is to create a system that performs Preceding Vehicle Detection and Road Area Recognition based on computer vision and pattern recognition methods. At the first part, our subsystem relies on cues like color, symmetry and relevant position that are common to all rear car lights, making it accurate even in difficult scenarios like urban driving. The road area segmentation is performed using the powerful Random Walker algorithm, combined with the a priori information of camera's optical view and the calculation of optical flow between two consecutive frames. Despite the challenging part of taking the input from just a monocular camera our system was evaluated both qualitative and quantitative, giving impressively high scores in completeness and precision.

Ευχαριστίες

Η παρούσα διπλωματική εργασία εκπονήθηκε κατά την διάρκεια του τελευταίου έτους φοίτησής μου στο Τμήμα των Ηλεκτρολόγων Μηχανικών και Τεχνολογίας Υπολογιστών του Πανεπιστημίου Πατρών. Κατά την ενασχόλησή μου με το αντικείμενο είχα την ευκαιρία να διευρύνω το γνωστικό μου πεδίο και να αποκτήσω νέες δεξιότητες. Αρχικά θα ήθελα να ευχαριστήσω τον επιβλέποντα αναπληρωτή καθηγητή κ. Ευάγγελο Δερματά που με παρότρυνε να ασχοληθώ με ένα τόσο ενδιαφέρον θέμα, καθώς και για τις πολύτιμες συμβουλές και υποδείξεις του. Έπειτα είμαι παραπάνω από ευγνώμων για την στήριξη που έλαβα από την οικογένειά μου από τα πρώτα χρόνια της ζωής μου έως τώρα. Τέλος θα ήθελα να ευχαριστήσω τους φίλους μου για τις αναμνήσεις που δημιουργήσαμε όλα αυτά τα χρόνια.

Περιεχόμενα

1. Εισαγωγή.....	8
2. Εργαλεία.....	9
2.1 Βάσεις δεδομένων για την αξιολόγηση του συστήματος.....	9
2.1.1 Ποιοτικά	9
2.1.2 Ποσοτικά.....	9
2.2 Γλώσσα και βιβλιοθήκες	11
2.3 Λειτουργικό Σύστημα.....	12
3. Θεωρία – αλγόριθμοι	13
3.1 Χρωματικοί χώροι.....	13
3.1.1 RGB	13
3.1.2 CIEL ab	14
3.1.3 HSV.....	14
3.1.4 c1c2c3	15
3.2 Ποσοτικά μέτρα αξιολόγησης.....	16
3.3 K-means	17
3.3.1 Μέθοδος γονάτου	18
3.4 Fast Radial Symmetry Transform	19
3.5 Otsu's Thresholding	23
3.6 Οπτική ροή	24
3.6.1 Πυκνή οπτική ροή - Αλγόριθμος Farneback.....	25
3.6.2 Αραιή οπτική ροή.....	27
3.7 Random Walker	31
3.8 Εκτίμηση Ομοιότητας	34
3.8.1 Mean Absolute Error (MAE)	34
3.8.2 Structural Similarity Index (SSIM).....	35
3.8.3 Διαφορές MAE-SSIM.....	35
4. Εντοπισμός προπορευόμενου οχήματος	36
4.1 Εισαγωγή.....	36
4.2 Σχετική βιβλιογραφία.....	36
4.3 Προτινόμενο σύστημα	37

4.4 Πειραματικά αποτελέσματα.....	50
4.5 Συμπεράσμα	55
5. Αναγνώριση Δρόμου.....	56
5.1 Εισαγωγή.....	56
5.2 Σχετική βιβλιογραφία.....	56
5.3 Προτινόμενο σύστημα	59
5.3.1 Προεπεξεργασία.....	61
5.3.2 Οπτική ροή	64
5.3.3 Τοποθέτηση των seeds	70
5.3.4 Segmentation - Random Walker	76
5.4 Πειραματικά αποτελέσματα.....	78
5.4.1 Ποιοτική αξιολόγηση	78
5.4.2 Ποσοτική αξιολόγηση.....	81
5.5 Συμπεράσμα	85
6. Επίλογος.....	86
6.1 Ανασκόπηση του παρόντος συστήματος	86
6.2 Μελλοντικές μελέτες	86
7. Παράρτημα.....	88
8. Βιβλιογραφία.....	93

Πίνακες

Πίνακας 1. Βάσεις δεδομένων Caltech 1999 2001.	9
Πίνακας 2. Βάση δεδομένων DIPLODOC	10
Πίνακας 3. Χαρακτηριστικά συστήματος.....	12
Πίνακας 4. Αξιολόγηση του συστήματος στις βάσεις Caltech 1999,2001	50
Πίνακας 5. Διαφορετικές περιπτώσεις υποδειγματοληψίας της αρχικής εικόνας και αντίστοιχοι χρόνοι εκτέλεσης του αλγορίθμου.	62
Πίνακας 6. Πλήθος Αληθώς Θετικών (TP), Λανθασμένα Θετικών (FP) και Λανθασμένα Αρνητικών (FN) για το σύστημα μας.	84
Πίνακας 7. Ποσοστά πληρότητας (completeness), ορθότητας (correctness) και ποιότητας (quality) του συστήματος.	84

Παράρτημα

7. 1 Συνάρτηση υπολογισμού του σημείου γονάτου	88
7. 2 Συνάρτηση εύρεσης των ζευγών υποψήφιων φαναριών.....	88
7. 3 Συνάρτηση υπολογισμού του δείκτη ομοιότητας SSIM.	89
7. 4 Συνάρτηση δημιουργίας της περιοχής ενδιαφέροντος (ROI).	89
7. 5 Συνάρτηση δημιουργίας του τραπεζίου για την τοποθέτηση των seeds δρόμου.	91
7. 6 Δημιουργία των seeds μη – δρόμου.	92
7. 7 Δημιουργία των seeds δρόμου	92

1. Εισαγωγή

Η ανάγκη του ανθρώπου για μετακίνηση γίνεται συνεχώς πιο έντονη από την στιγμή που εγκατέλειψε τον αγροτικό τρόπο ζωής και εγκαταστάθηκε στα μεγάλα αστικά κέντρα. Ως αποτέλεσμα παρατηρείται η συνεχής αύξηση του παγκόσμιου στόλου οχημάτων που κυκλοφορούν στους δρόμους. Πιο συγκεκριμένα στην ΕΕ το 2012, είχαν καταγραφεί 243 εκατομμύρια επιβατικά αυτοκίνητα με τον αριθμό αυτόν να αυξάνεται στα 257 εκ. το 2016. Είναι επόμενο λοιπόν, να σημειωθεί ανάλογη αύξηση στα ατυχήματα που συμβαίνουν καθημερινά, με πολλά εξ' αυτών να είναι θανάσιμα. Σύμφωνα με την στατιστική υπηρεσία Eurostat, το 2018 στην Ε.Ε. έχασαν την ζωή τους 23400 συνάνθρωποί μας, με το 45% αυτών να είναι επιβάτες ή οδηγοί οχημάτων και ένα άλλο 21% να είναι πεζοί. Εξετάζοντας αυτά τα στοιχεία η αρμόδια επιτροπή έλαβε την πρωτοβουλία να ιδρύσει το 1997 το πρόγραμμα Euro NCAP με σκοπό την αξιολόγηση της ασφάλειας κάθε νέου μοντέλου αυτοκινήτου πρωτού βγει στην αγορά. Αυτή η νέα νομοθεσία ανάγκασε τους κατασκευαστές να σχεδιάζουν αυτοκίνητα με πιο ανθεκτικό σασί, περισσότερους αερόσακους και καλύτερες ζώνες ασφαλείας. Αυτή η φιλοσοφία εστιάζει στους παθητικούς μηχανισμούς ασφαλείας, με την έννοια ότι δρουν αφού γίνει το ατύχημα με σκοπό την μείωση των τραυματισμών των επιβατών. Δυστυχώς με την πάροδο του χρόνου η εξέλιξη τους έφτασε σε έναν κορεσμό, γεγονός που έστρεψε το ενδιαφέρον των κατασκευαστών στην ανάπτυξη ενεργητικών συστημάτων ασφαλείας τα οποία θα ενεργοποιούνται ώστε να αποτρέψουν ένα ενδεχόμενο ατύχημα. Τέτοια συστήματα εντάσσονται στην ευρεία κατηγορία των Προηγμένων Συστημάτων Υποβοήθησης της Οδήγησης ή με τον αγγλικό όρο ADAS (Advance Driving Assistance Systems). Ενδεικτικά παρουσιάζουμε τα πιο διαδεδομένα συστήματα που συναντάμε σε πολλά εμπορικά μοντέλα αυτοκινήτων και εκτελούν τις εξής λειτουργίες: α) Προειδοποίηση αλλαγής λωρίδας (Lane Departure System), β) Παρακολούθηση των πεζών (Pedestrian Monitor System), γ) Παρακολούθηση της «νεκρής» γωνίας (Blind Spot Monitors), δ) Ειδοποίηση σύγκρουσης με το προπορευόμενο όχημα (Forward Collision Warning). Τα προηγμένα αυτά συστήματα, επεξεργάζονται τις εισόδους που λαμβάνονται από μία ομάδα αισθητήρων και καμερών και στην συνέχεια καλούνται να λάβουν μία απόφαση αν θα επέμβουν ώστε να διορθώσουν κάποιο σφάλμα του οδηγού.

Η έρευνα και η κατασκευή των ADAS είναι ιδιαίτερα δύσκολη αν αναλογισθούμε ότι πρέπει να λειτουργούν αξιόπιστα ακόμη και στις πιο δύσκολες συνθήκες οδήγησης όπως με βαριά βροχόπτωση, ομίχλη ή σε πυκνό οδικό δίκτυο. Επίσης ο χρόνος ενεργοποίησης τους πρέπει να είναι ο ελάχιστος δυνατός και μάλιστα πολύ μικρότερος από τον χρόνο αντίδρασης του ανθρώπινου εγκεφάλου σε κάποιο ερέθισμα. Σε διαφορετική περίπτωση η αξιοπιστία του συστήματος μειώνεται και τίθενται αυτόματα σε κίνδυνο οι ζωές των επιβατών. Συνδυάζοντας τους δύο μηχανισμούς ασφαλείας που αναφέραμε προηγουμένως και σε συνυπολογίζοντας την αφύπνηση των πολιτών ως προς την σωστή οδηγική συμπεριφορά, παρατηρείται μείωση της τάξεως 37% στον ετήσιο αριθμό θανάσιμων ατυχημάτων στην ΕΕ.

2. Εργαλεία

2.1 Βάσεις δεδομένων για την αξιολόγηση του συστήματος

2.1.1 Ποιοτικά

Στα πρώτα στάδια της ανάπτυξης ενός συστήματος είναι χρήσιμο να αξιολογείται ποιοτικά ώστε να διορθωθούν τυχόν αστοχίες και να γίνουν οι απαραίτητες βελτιώσεις. Στην παρούσα εργασία χρησιμοποιούμε ένα σετ από 15 βίντεο που έχουν ληφθεί από κάμερα τοποθετημένη ψηλά στο παρμπρίζ του αυτοκινήτου. Η δειγματοληψία έγινε σε δρόμους της Πάτρας υπό αίθριες καιρικές συνθήκες, την ημέρα. Κάθε frame έχει διαστάσεις 1080x1920 και κάθε βίντεο έχει διάρκεια μερικά δευτερόλεπτα. Το σετ αυτό είναι αρκετά πλήρες καθώς περιέχει διάφορες κατηγορίες δρόμων, από ταχείας κυκλοφορίας με αυξημένη κίνηση έως επαρχιακό δίκτυο αραιής κίνησης. Επίσης συναντώνται πλήρως συντηρημένα οδικά δίκτυα με σωστές διαγραμμίσεις καθώς και οδοστρώματα με ανύπαρκτη διαγράμμιση και διασδιάκριτα όρια. Επί προσθέτως, ένα εκ των βίντεο αφορά κίνηση του οχήματος σε δρόμο που είναι υπο κατασκευή και στον οποίο εκτελούνται έργα. Σε αυτήν την περίπτωση εμφανίζονται αναπάντεχα εμπόδια όπως πινακίδες υποχρεωτικής κατεύθυνσης, προσωρινές διαχωριστικές μπαριέρες και κορδέλες που ανεβάζουν τον πήχη δυσκολίας του προβλήματος αναγνώρισης δρόμου. Τέλος παρατηρείται σε πολλές εικόνες έντονη ηλιοφάνεια ή υπερβολική σκίαση στον δρόμο θέτωντας έτσι άλλη μία πρόκληση για το σύστημα μας.

2.1.2 Ποσοτικά

Για το σκέλος της αναγνώρισης προπορευόμενου οχήματος χρησιμοποιούνται δύο βάσεις δεδομένων [1]. Η Caltech 1999 περιέχει εικόνες σταθμευμένων οχημάτων με υπόβαθρο που ποικίλει όπως κτήρια, βλάστηση καθώς και άλλα σταθμευμένα οχήματα. Η δεύτερη βάση είναι η Caltech 2001 η οποία αποτελείται από στιγμιότυπα οδικών δικτύων σε πραγματικές συνθήκες με οχήματα που βρίσκονται εν κινήσει. Στα frames συναντάμε διαφορετικά επίπεδα φωτεινότητας και πολυπλοκότητας του περιβάλλοντος και έχουν ληφθεί από κάμερα η οποία είναι τοποθετημένη στο πάνω μέρος του παρμπρίζ του αυτοκινήτου (ego vehicle).

	Frames	Ανάλυση	Υπόβαθρο	Συνθήκες
Caltech 1999	126	892x592	Στάθμευση	Ηλιοφάνεια - ημέρα
Caltech 2001	526	360x240	Οδικό δίκτυο	Ηλιοφάνεια - ημέρα

Πίνακας 1. Βάσεις δεδομένων Caltech 1999 2001.

Η αξιολόγηση του συστήματος αναγνώρισης δρόμου πραγματοποιήθηκε με την δημόσια βάση δεδομένων DIPLODOC [2] η οποία αναπτύχθηκε υπό την συνεργασία του ιδρύματος FBK, της FIAT και του Πανεπιστημίου του Trento της Ιταλίας. Αποτελείται από 5 βίντεο που καλύπτουν ποικίλα σενάρια οδήγησης όπως δρόμοι ταχείας κυκλοφορίας, ημιαστικό δίκτυο, δρόμοι με στροφές. Ως προς το ευρύτερο περιβάλλον συναντώνται πεζοδρόμια, διαβάσεις, διασταυρώσεις, χώροι στάθμευσης και εμπόδια.

Τα τεχνικά χαρακτηριστικά της βάσης DIPLODOC είναι τα εξής:

Όνομα	Κάμερα	Ανάλυση	fps	Διάρκεια	Περιβάλλον	Καιρός	Περιοχή
DIPLODOC	Videre design MEGA-D Stereo Camera	320x240	15	0:56	Μεικτό	Αίθριος	Trento- Ιταλία

Πίνακας 2. Βάση δεδομένων DIPLODOC

Για τον ακριβή προσδιορισμό των περιοχών δρόμου και την αποφυγή των παρερμηνεύσεων θεωρείται ότι ως δρόμος λαμβάνεται οπουδήποτε μπορεί να μεταβεί ένα αυτοκίνητο χωρίς να ανέβει επίπεδο. Συνολικά αποκτήθηκαν 865 frames με χρήση στερεοσκοπικής κάμερας δύο καναλιών (δεξί-αριστερό) που τοποθετήθηκε στο παρμπρίζ του αυτοκινήτου. Κάθε ζεύγος εικόνων συνοδεύεται από ένα αρχείο text που περιέχει τις κανονικοποιημένες συντεταγμένες $[0,1]$ των γωνιών-κόμβων του πολυγώνου ιδανικού δρόμου (road) και των πολυγώνων των εμποδίων (occl). Η δομή ενός τυπικού πολυγώνου είναι η ακόλουθη:

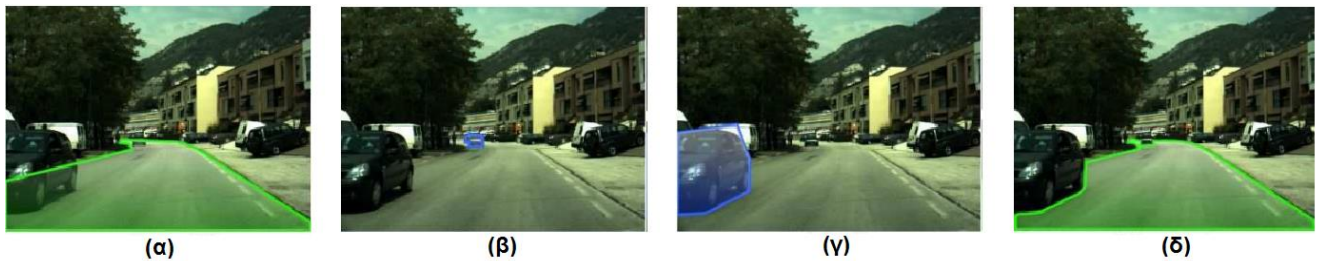
$$type\ N\ x_1\ y_1\ x_2\ y_2\ \dots\ \dots\ \dots\ x_N\ y_N$$

όπου το type λαμβάνει 2 τιμές, δρόμος (road) ή εμπόδιο (occl), N είναι το πλήθος των σημείων που οριοθετούν το πολύγωνο και στην συνέχεια σημειώνονται με την σειρά οι κανονικοποιημένες συντεταγμένες κάθε κόμβου ως προς τις διαστάσεις του frame (x_1, y_1) (x_2, y_2) ... κ.ο.κ.

Αυτά τα σημεία έχουν οριστεί χειροκίνητα και αποτελούν την βάση αληθείας σύμφωνα με την οποία θα κάνουμε την αξιολόγηση του συστήματός μας. Παρακάτω, στην Εικόνα 1 φαίνεται ένα αρχείο text και στην Εικόνα 2 η αντίστοιχη οπτική αναπαράσταση των πολυγώνων.

```
road 11 0 0.773611 0.415625 0.636111 0.40625 0.613889 0.379167 0.606944 0.38125 0.591667 0.470833 0.601389 0.566667 0.604167
0.628125 0.631944 0.998958 0.959722 0.998958 0.998611 0 0.998611 0.417708 0.563889 0.402083 0.580556 0.404167 0.598611
0.41875 0.6 0.419792 0.625 0.4625 0.620833 0.461458 0.570833 0.2375 0.675 0.194792 0.526389 0 0.558333 0 0.930556 0.0895833
0.922222 0.235417 0.820833]
```

Εικόνα 1. Μορφή του αρχείου diplo000202-L.txt με τις συντεταγμένες των πολυγώνων



Εικόνα 2. (α) Περιοχή ιδανικού δρόμου αγνοώντας τα εμπόδια, (β) Πρώτο εμπόδιο, (γ) Δεύτερο εμπόδιο (δ) Ορατή περιοχή δρόμου.

2.2 Γλώσσα και βιβλιοθήκες

Η Python είναι μία από τις πλέον διαδεδομένες και αναπτυσσόμενες γλώσσες προγραμματισμού, καλύπτοντας ένα ευρύ φάσμα εφαρμογών όπως γραφική αναπαράσταση δεδομένων, δημιουργία ιστοσελίδων, αυτοματοποίηση διεργασιών κ.α. Μερικά από τα πλεονεκτήματα της, όπως παρουσιάζονται στο [3], είναι α) εύκολο και απλό συντακτικό, β) πληθώρα βιβλιοθηκών με χρήσιμες συναρτήσεις, γ) υποστήριξη πολλών τύπων δεδομένων (strings, numbers, lists, dictionaries), δ) δυναμική δήλωση των μεταβλητών στο σημείο χρήσης τους, ε) η διαχείριση της απαιτούμενης μνήμης γίνεται αυτόματα σε αντίθεση με άλλες γλώσσες όπως η C που γίνεται από τον χρήστη και στ) είναι συμβατή με τα περισσότερα λειτουργικά συστήματα όπως Windows, Linux, Os και Unix.

Ένα μειονέκτημα που συναντάται συχνά είναι η μειωμένη ταχύτητα απόκρισης και αυτό διότι δεν είναι άμεσα συνδεδεμένη με το υλικό του υπολογιστή. Σε real time εφαρμογές αυτός είναι ένας σοβαρός περιορισμός, ωστόσο στο σημείο αυτό μας ενδιαφέρει κυρίως η ακρίβεια των προτινόμενων αλγορίθμων και όχι τόσο η ταχύτητα. Για τους παραπάνω λόγους επιλέχθηκε η γλώσσα Python για την ανάπτυξη του παρόντος συστήματος.

Στην συνέχεια περιγράφουμε τις βοηθητικές βιβλιοθήκες που χρησιμοποιήθηκαν με κυρίαρχο ρόλο να κατέχει η OpenCV [4]. Πρόκειται για μία βιβλιοθήκη ανοιχτού κώδικα που περιέχει συναρτήσεις επεξεργασίας εικόνας καθώς και αλγορίθμους μηχανικής μάθησης. Μερικές μόνο από τις λειτουργίες που μπορούμε να εκτελέσουμε είναι η αναγνώριση προσώπων, εντοπισμών ακμών, τμηματοποίηση μίας εικόνας σε 2 ή περισσότερες κλάσεις, ανίχνευση κινούμενων αντικειμένων, εντοπισμό χρώματος, σύγκριση εικόνων ως προς την ομοιότητα κ.α. Επίσης είναι ανοιχτού κώδικα άρα συνεχώς εξελισσόμενη ενώ είναι συμβατή με τις γλώσσες προγραμματισμού Matlab, C, C++, Python και Java.

Ιδιαίτερα χρήσιμη είναι και η βιβλιοθήκη Scikit-image ή αλλιώς Skimage [5], ανοιχτού κώδικα, η οποία προσφέρει πληθώρα συναρτήσεων για την επεξεργασία και την διαχείριση εικόνων. Περιλαμβάνει αλγορίθμους για τμηματοποίηση εικόνας και μετασχηματισμούς, τροποποιήσεις στο χρώμα, φιλτράρισμα, feature detection κ.α. Αναπτύχθηκε για την Python και διαχειρίζεται τις εικόνες ως δεδομένα τύπου numpy. Στην μελέτη μας χρησιμοποιούμε το module skimage.segmentation που περιέχει τον supervised αλγόριθμο Random Walker τον οποίο θα αναλύσουμε στο επόμενο κεφάλαιο.

Μία άλλη σημαντική βιβλιοθήκη που πρέπει να αναφέρουμε είναι η Scikit-learn [6] καθώς περιλαμβάνει αλγορίθμους μηχανικής μάθησης για προεπεξεργασία εικόνων, τμηματοποίηση, ταξινόμηση και μοντελοποίηση. Μερικοί εξ αυτών είναι ο Support Vector Machines, Random Forests, K- nearest neighbors καθώς και τις μαθηματικές επιστημονικές βιβλιοθήκες NumPy και SciPy. Όσον αφορά το σύστημά μας, μέσω του module sklearn.cluster αποκτάμε πρόσβαση στον k-means [κεφ. 3.3] με τον οποίο κάνουμε ομαδοποίηση σημείων με συγκεκριμένα χαρακτηριστικά. Με αυτόν τον τρόπο το πρόβλημα απλοποιείται και απαλλασσόμαστε από επιπλέον πληροφορία καθώς εστιάζουμε σε αντιπροσωπευτικά κέντρα της κάθε ομάδας σημείων.

2.3 Λειτουργικό Σύστημα

Οι αλγορίθμοι μηχανικής μάθησης και επεξεργασίας εικόνας έχουν μεγάλο υπολογιστικό κόστος και κατ' επέκταση υψηλές απαιτήσεις σε υπολογιστικούς πόρους. Επομένως για να έχουμε άμεση απόκριση των προγραμμάτων, πρέπει να διαθέτουμε επεξεργαστή με μεγάλο κύκλο ρολογιού συνδυασμένο με μνήμη που έχει την βέλτιστη αρχιτεκτονική και προσπελάζεται γρήγορα.

Παρακάτω παραθέτουμε τα βασικά χαρακτηριστικά του συστήματός μας

CPU	Intel Core i5 – 2.4 GHz
Processor Graphics	Intel HD Graphics 4400
RAM	8 GB
System Type	64 bit
Operating System	Windows 8.1

Πίνακας 3. Χαρακτηριστικά συστήματος

Μία προσέγγιση που χρησιμοποιείται ευρέως για περαιτέρω μείωση του χρόνου εκτέλεσης ενός προγράμματος είναι η παράλληλη επεξεργασία. Αυτό επιτυγχάνεται μέσω των πολυπύρηνων επεξεργαστών (multicore processors) οι οποίοι διαθέτουν επιμέρους υπολογιστικές μονάδες στις οποίες γίνεται ταυτόχρονη εκτέλεση εντολών ή διεργασιών, αυξάνοντας θεαματικά την ταχύτητα εκτέλεσης των προγραμμάτων. Στην μελέτη μας δεν χρησιμοποιούμε τεχνικές παράλληλης επεξεργασίας, καθώς στα πρώτα στάδια ανάπτυξης του συστήματος πρέπει να εστιάσουμε στην σωστή και ομαλή λειτουργία του. Ύστερα, αν θέλουμε να πετύχουμε γρήγορη απόκριση σε πραγματικό χρόνο, τότε θα εξετάσουμε το ενδεχόμενο της παράλληλης υλοποίησης.

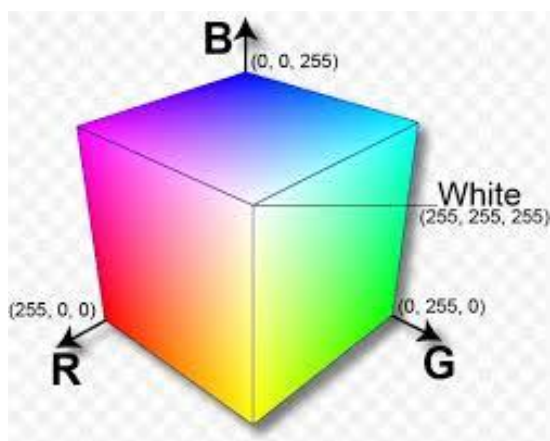
3. Θεωρία – αλγόριθμοι

3.1 Χρωματικοί χώροι

Οι εφαρμογές μηχανικής όρασης αναπτύσσονται με σκοπό να είναι λειτουργικές υπό διάφορες συνθήκες φωτεινότητας, δίχως να χάνουν την αξιοπιστία τους. Για παράδειγμα ένα σύστημα αναγνώρισης σηματοδοτών ρύθμισης κυκλοφορίας τοποθετημένο σε ένα αυτοκίνητο, απαιτούμε να λειτουργεί σωστά σε αντίξοες καιρικές συνθήκες όπως συννεφιά βροχή και ομίχλη. Για να επιτευχθεί αυτό το ζητούμενο, πρέπει σε πρώτο στάδιο να περιληφθεί στην αναπαράσταση της εικόνας η παράμετρος της φωτεινότητας. Στην διεθνή βιβλιογραφία εξετάζονται οι χώροι HSV, CIELab, YCbCr οι οποίοι είναι περισσότερο ανθεκτικοί στις αλλαγές της φωτεινότητας συγκριτικά με τον RGB. Παρακάτω παρουσιάζονται οι βασικές συνιστώσες των χώρων που χρησιμοποιούνται στην εργασία.

3.1.1 RGB

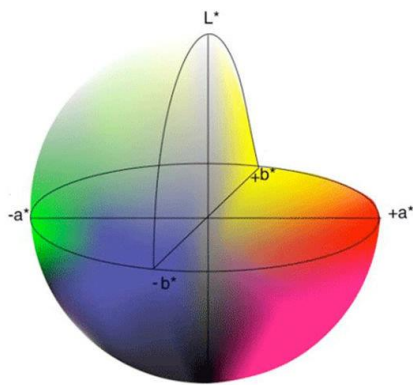
Κατά το πρότυπο του RGB [7] κάθε pixel αποτελείται από 3 ακέραιες τιμές από 0 έως 255 όπου κάθε μία αφορά τα βασικά χρώματα κόκκινο (R), πράσινο (G) και μπλε (B) αντιστοίχα. Ως γνωστών συνδυάζοντας κατάλληλα τις τιμές αυτές μπορεί να κατασκευαστεί οποιοδήποτε χρώμα του φάσματος. Το μεγάλο μειονέκτημα του RGB είναι ότι δεν διαχωρίζεται η χρωματική πληροφορία από την πληροφορία της έντασης (intensity) με αποτέλεσμα να είναι ευαίσθητο στις απότομες μεταβολές της φωτεινότητας.



Εικόνα 3. Τρισδιάστατη αναπαράσταση του χώρου RGB.

3.1.2 CIEL ab

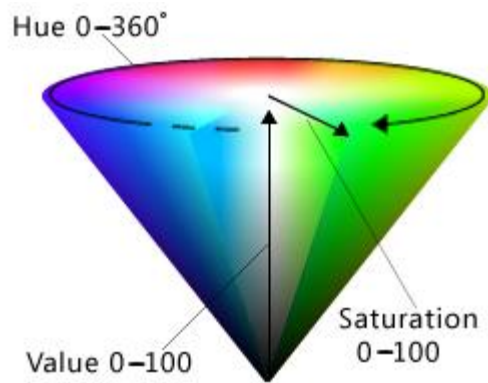
Η αναπαράσταση αυτή θεωρείται από την International Commission on Illumination ως η πληρέστερη. Περιγράφει όλα τα χρώματα που είναι ορατά από το ανθρώπινο μάτι με χρήση τριών καναλιών. Το πρώτο είναι το κανάλι της φωτεινότητας L^* με τιμές από 0 (μάυρο-χαμηλή) έως 100 (άσπρο-υψηλή). Το κανάλι a^* αντιπροσωπεύει την θέση του χρώματος σε μια κλίμακα πράσινου-κόκκινου, όπου -127 υποδεικνύει καθαρό πράσινο και +127 υποδεικνύει καθαρό κόκκινο. Το κανάλι b^* αντιπροσωπεύει την θέση του χρώματος σε μια κλίμακα κίτρινου-μπλε, όπου -127 υποδεικνύει καθαρό μπλε και +127 υποδεικνύει καθαρό κίτρινο. Τα πλεονεκτήματα του CIELab είναι ότι αφενώς διαχωρίζει την φωτεινότητα της εικόνας από το χρώμα και αφετέρου το κανάλι L έχει σχεδιαστεί ώστε να ταιριάζει με τον τρόπο που αντιλαμβάνεται την φωτεινότητα ο άνθρωπος [7], [8].



Εικόνα 4. Τρισδιάστατη αναπαράσταση του χώρου CIELab.

3.1.3 HSV

Ο χώρος HSV [7], [9] αποτελείται από τρεις συνιστώσες. Η απόχρωση (Hue) αφορά τον κωδικό του χρώματος (κόκκινο, πορτοκαλί, μπλε) και παίρνει τιμές από 0 - 360°. Ο κορεσμός (saturation) υποδεικνύει το βάθος του χρώματος (μπλέ,μπλε σκούρο) και παίρνει τιμές από 0% (ανοιχτό) έως 100% (πλήρης κορεσμός). Τέλος η τιμή (Value) αφορά την φωτεινότητα δηλαδή πόσο ανοιχτό ή σκούρο είναι το χρώμα από 0% (μάυρο) έως 100% (λευκό).



Εικόνα 5. Τρισδιάστατη αναπαράσταση του χώρου HSV.

3.1.4 c1c2c3

Το μοντέλο c1c2c3 [10] χρησιμοποιείται συχνά σε περιβάλλοντα με έντονη παρουσία σκιών καθώς μειώνει τον θόρυβο που προέρχονται από αυτές. Αποτελείται από 3 κανάλια όπου το καθένα είναι συνάρτηση των συνιστωσών του RGB χώρου σύμφωνα με τις σχέσεις:

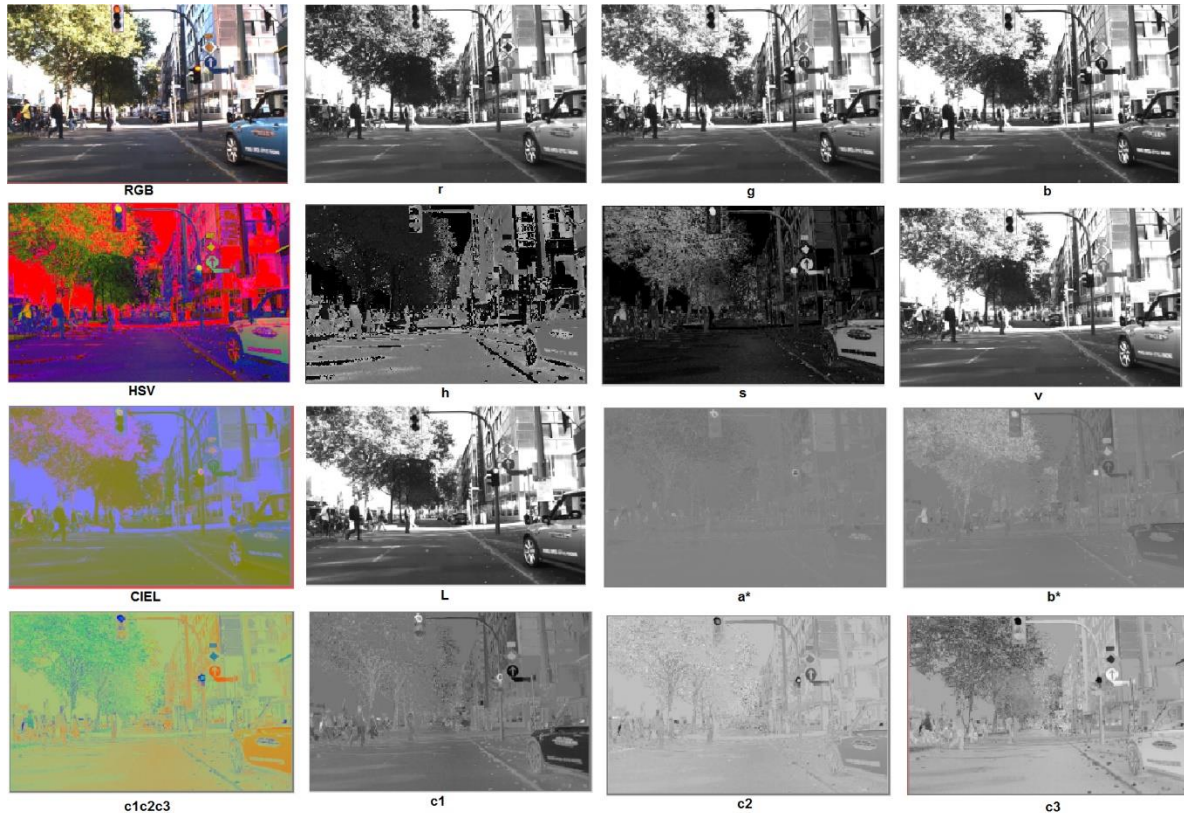
$$c_1 = \tan^{-1} \frac{R}{\max(G, B)}$$

$$c_2 = \tan^{-1} \frac{G}{\max(R, B)}$$

$$c_3 = \tan^{-1} \frac{B}{\max(R, G)}$$

Παρακάτω στην Εικόνα 6, παρουσιάζονται οι διαφορετικές αναπαραστάσεις της ίδιας εικόνας στα επιμέρους κανάλια των χώρων που προαναφέρθηκαν. Συμπεραίνουμε ότι σε πραγματικές συνθήκες το RGB δεν μας δίνει καμία πληροφορία ως προς την φωτεινότητα των pixels και αυτό ωφείλεται στην μεγάλη συσχέτιση μεταξύ των τριών καναλιών του. Αντίθετα οι χώροι HSV, CIELab και c1c2c3 διαχωρίζουν την πληροφορία της φωτεινότητας από το χρώμα και παρατηρούμε ότι καθε κανάλι λειτουργεί σαν μία «μάσκα» αφού δίνει

έμφαση σε ορισμένα χαρακτηριστικά της εικόνας. Για παράδειγμα το κανάλι a^* και το κανάλι $c1$ τονίζουν τους φωτεινούς σηματοδότες και τις πινακίδες κυκλοφορίας (κόκκινο-κίτρινο χρώμα). Η εικόνα περιέχει έναν δρόμο ο οποίος είναι σκιερός με μερικά φωτεινά σημεία στο κέντρο. Αξίζει να σημειωθεί ο χώρος $c1c2c3$ και ειδικά το κανάλι $c1$ διαχειρίζεται πολύ καλά των θόρυβο των σκιών δίνοντας έμφαση στα φωτεινά αντικείμενα.



Εικόνα 6. Αναπαράσταση μίας εικόνας στους χώρους RGB, CIELab, HSV, $c1c2c3$. Στην πρώτη στήλη είναι η συνολική απεικόνιση του κάθε χώρου και δίπλα τα επιμέρους κανάλια του.

3.2 Ποσοτικά μέτρα αξιολόγησης

Για την μέτρηση της απόδοσης του συστήματος χρησιμοποιείται λογική παρόμοια με αυτήν της ταξινόμησης δειγμάτων σε 2 ομάδες π.χ θετικό ή αρνητικό. Μετά το πέρας της διαδικασίας υπάρχουν 4 δυνατές περιπτώσεις για το καθένα: α) το δείγμα να έχει ταξινομηθεί ως σωστό και να είναι πράγματι σωστό (Αληθώς Θετικό), β) να έχει τοποθετηθεί στην κατηγορία των θετικών ενώ στην πραγματικότητα είναι αρνητικό (Λανθασμένα Θετικό), γ) να έχει τοποθετηθεί στην κατηγορία των αρνητικών ενώ στην πραγματικότητα είναι θετικό (Λανθασμένα Αρνητικό) και τέλος να κατηγοριοποιείται ως αρνητικό και να είναι πράγματι

αρνητικό (Αληθώς Αρνητικό). Αυτά τα μέτρα ορίζονται στην διεθνή βιβλιογραφία με την σειρά ως True Positive (TP), False Positive (FP), False Negative (FN), True Negative (TN) και για λόγους ευκολίας θα αναφέρονται στην συνέχεια με τα αρχικά τους.

Υστερα ορίζονται οι ποσότητες Recall (Πληρότητα), Precision (Ακρίβεια) και Quality (Ποιότητα). Η πρώτη εκφράζει τον λόγο των αληθώς θετικών δειγμάτων προς το άθροισμα των αληθώς θετικών και λανθασμένα αρνητικών (πληρότητα ταξινόμησης). Η δεύτερη εκφράζει τον λόγο των αληθώς θετικών δειγμάτων προς εκείνα που έχουν κατηγοριοποιηθεί ως θετικά είτε σωστά είτε εσφαλμένα (ακρίβεια ταξινόμησης). Τέλος η ποσότητα Quality είναι ο λόγος των αληθώς θετικών δειγμάτων προς το άθροισμα όλων των θετικών και των λανθασμένα αρνητικών.

- Recall - Completeness (Πληρότητα) : $R = TP / (TP + FN)$
- Correctness - Precision (Ακρίβεια) : $P = TP / (TP + FP)$
- Quality (Ποιότητα) : $Q = TP / (TP + FP + FN)$

3.3 K-means

Πρόκειται για έναν από τους δημοφιλέστερους αλγόριθμους [11] για την ομαδοποίηση δειγμάτων – παρατηρήσεων και είναι unsupervised δηλαδή δεν χρειάζεται να δοθούν ως είσοδοι εκ των προτέρων γνωστά δείγματα, αλλά επεξεργάζεται απευθείας τα ήδη διαθέσιμα. Η ομαδοποίηση των παρατηρήσεων σε k ομάδες (clusters) γίνεται με τέτοιο τρόπο ώστε το άθροισμα των αποστάσεων των σημείων της κάθε ομάδας από το αντίστοιχο κέντρο να είναι η ελάχιστη δυνατή. Αυτή η ποσότητα αναφέρεται στην διεθνή βιβλιογραφία ως intra cluster distance και σκοπός είναι να ελαχιστοποιηθεί. Ως είσοδος δίνονται τα δεδομένα όπου αναπαριστώνται ως διανύσματα με τις συντεταγμένες τους στον χώρο και ο επιθυμητός αριθμός των ομάδων K. Ο αλγόριθμος εκτελεί τα παρακάτω βήματα :

1. Ορισμός k τυχαίων κέντρων, ένα για κάθε ομάδα.
2. Τοποθέτηση της κάθε παρατήρησης σε μία ομάδα :
Αυτό επιτυγχάνεται υπολογίζοντας για κάθε σημείο την Ευκλείδεια απόσταση μεταξύ αυτού και του κάθε κέντρου c_i και τοποθετείται στην ομάδα με την μικρότερη απόσταση από το κέντρο της. Η παραπάνω πρόταση μαθηματικά εκφράζεται ως εξής:

$$\operatorname{argmin}_{c_i \in C} (\sqrt{c_i - x_i}^2)$$

3. Ενημέρωση των κέντρων

Λόγω των αλλαγών στις ομάδες που έγιναν στο προηγούμενο βήμα υπολογίζονται τα νέα κέντρα c_i' των ομάδων :

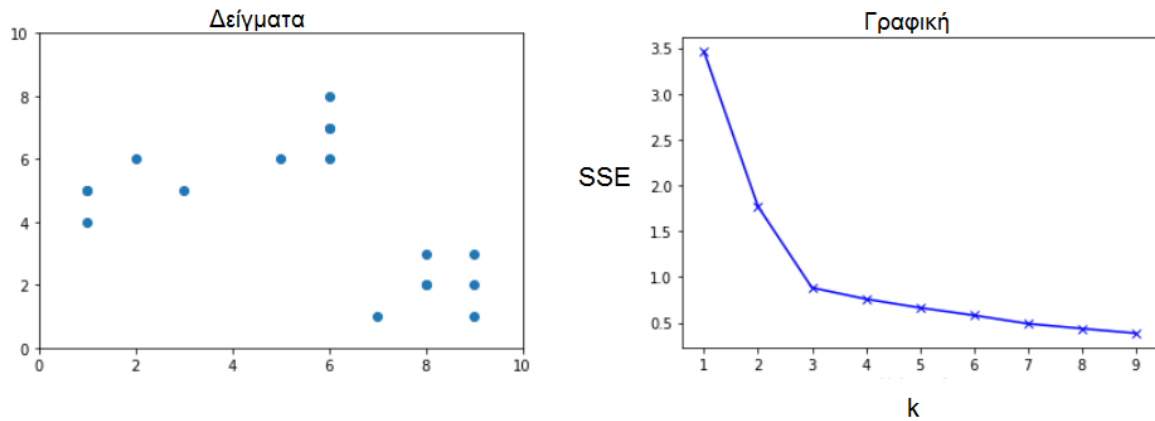
$$c_i' = \frac{1}{|S_i|} \sum_{x_i \in S_i} x_i$$

όπου S_i η i -οστή ομάδα και $|S_i|$ είναι το πλήθος των στοιχείων της.

4. Επαναλαμβάνονται τα βήματα (2) και (3) μέχρις ότου τα νέα κέντρα να μετατοπίζονται κατά αμελητέα ποσότητα $\overrightarrow{ds}$ η οποία ορίζεται από τον χρήστη ή να ολοκληρωθεί ένας συγκεκριμένος αριθμός επαναλήψεων.

3.3.1 Μέθοδος γονάτου

Ένα σημαντικό σημείο του αλγορίθμου είναι η σωστή επιλογή του αριθμού των ομάδων (clusters). Η παράμετρος αυτή δίνεται από τον χρήστη και θεωρητικά μπορεί να πάρει τιμές από 1 έως και N όπου N το πλήθος των δειγμάτων. Ωστόσο δεν είναι πάντα γνωστό εκ των προτέρων η διάταξη των σημείων στον χώρο και κατ' επέκταση λανθασμένη επιλογή του k μπορεί να οδηγήσει σε άστοχη ομαδοποίηση των σημείων. Μία λύση σε αυτό το πρόβλημα δίνεται με την μέθοδο γονάτου [12] ή στην διεθνή ορολογία Elbow Method η οποία δοκιμάζει επαναληπτικά τιμές του k από 1 έως N και εκτελεί τον αλγόριθμο k -means. Σε κάθε επανάληψη υπολογίζεται το άθροισμα των τετραγώνων των αποστάσεων των σημείων κάθε ομάδας από το κέντρο της (Sum of Squared Errors – SSE) . Όταν εξετασθούν όλα τα k σχηματίζεται μία γραφική παράσταση με τον οριζόντιο άξονα να έχει τις τιμές k και τον κατακόρυφο την ποσότητα SSE που υπολογίστηκε για το αντίστοιχο k . Είναι λογικό για $k=1$ όλα τα σημεία ανήκουν στην ίδια ομάδα που σημαίνει ότι το SSE θα είναι μεγάλο. Αντίθετα αν επιλέξω πολύ μεγάλο k , έστω ίσο με το πλήθος των σημείων, τότε το SSE θα είναι ίσο με 0 εφόσον κάθε cluster αποτελείται από μόλις ένα σημείο το οποίο είναι και κέντρο. Η γραφική παράσταση στην αρχή έχει μεγάλη αρνητική κλίση μέχρι ένα σημείο k -optimal (σημείο γονάτου) και στην συνέχεια η κλίση μειώνεται αισθητά. Επιλέγεται για τον αλγόριθμο το k που αντιστοιχεί στο σημείο γονάτου και είναι αυτό για το οποίο θα γίνει η βέλτιστη ομαδοποίηση των δεδομένων (Εικόνα 7).



Εικόνα 7. Αριστερά φαίνονται τα σημεία στον χώρο και δεξιά η γραφική $SSE(k)$ με το σημείο γονάτου να είναι το $k=3$.

3.4 Fast Radial Symmetry Transform

Σε πολλές εφαρμογές της μηχανικής όρασης κρίνεται αναγκαίο να γίνει αναγνώριση των συμμετρικών αντικειμένων μίας εικόνας και ο διαχωρισμός τους από αντικείμενα με τυχαίο ακανόνιστο σχήμα. Στα πλαίσια της εργασίας και ειδικότερα στο σκέλος της αναγνώρισης του προπορευόμενου οχήματος, μία μέτρηση του βαθμού συμμετρίας των αντικειμένων θα μπορούσε να αξιοποιηθεί για τον εντοπισμό των πίσω φαναριών του αυτοκινήτου. Στη συνέχεια η διαδικασία ταυτοποίησης θα απλοποιούνταν, εφόσον το ενδιαφέρον θα εστιζόταν στα συμμετρικά αντικείμενα της εικόνας ενώ θα απορρίπτονταν άλλα λιγότερο συμμετρικά που αποτελούν θόρυβο. Πρέπει λοιπόν να χρησιμοποιηθεί ένας αλγόριθμος εύρεσης συμμετριών σε μία εικόνα και γι' αυτό επιλέγουμε τον αλγόριθμο Fast Radial Symmetry Transform [13].

Στο σημείο αυτό είναι σημαντικό να εισάγουμε το μέγεθος image gradient το οποίο υπολογίζεται για ένα pixel και εκφράζει την μεταβολή της έντασης της εικόνας προς τις δύο κατευθύνσεις x και y και περιγράφεται από την σχέση :

$$g = \nabla I = \begin{bmatrix} \frac{\partial I}{\partial x} \\ \frac{\partial I}{\partial y} \end{bmatrix} \quad \text{όπου } \frac{\partial I}{\partial x}, \frac{\partial I}{\partial y} \text{ είναι ο ρυθμός μεταβολής της έντασης της}$$

εικόνας I

ως προς x και y αντίστοιχα.

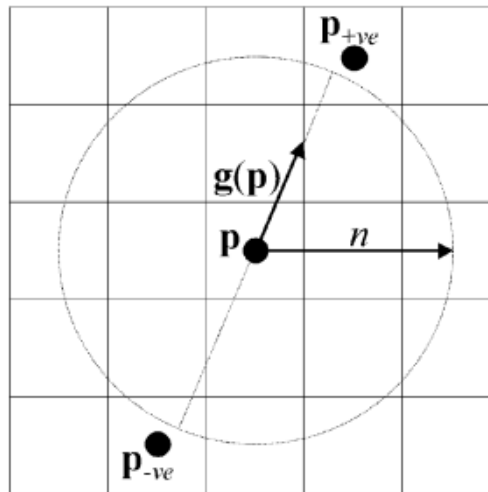
Ο αλγόριθμος FRST δέχεται ως είσοδο μία εικόνα ενός καναλιού και έναν φυσικό αριθμό N και ψάχνει συμμετρικά αντικείμενα με ακτίνα έυρους $[1, N]$.

Εκτελούνται τα εξής βήματα για κάθε ακτίνα n από 1 έως N :

- i. Για κάθε pixel p υπολογίζει το διάνυσμα g του image gradient και αυτομάτως ορίζονται ένα positive affected pixel $p_{+ve}(p)$ και ένα negative affected pixel $p_{-ve}(p)$ τα οποία απέχουν απόσταση n από το εξεταζόμενο pixel p . Οι συντεταγμένες τους ορίζονται από την κατεύθυνση του διανύσματος g και συγκεκριμένα αναχωρεί από το $p_{-ve}(p)$ και κατευθύνεται προς το $p_{+ve}(p)$ (Εικόνα 8) και μαθηματικά περιγράφονται από τις σχέσεις :

$$p_{+ve}(p) = p + \text{round}\left(\frac{g(p)}{\|g(p)\|} n\right)$$

$$p_{-ve}(p) = p - \text{round}\left(\frac{g(p)}{\|g(p)\|} n\right)$$



Εικόνα 8. Απεικόνιση των positively και negatively affected σημείων στον χώρο για δεδομένο pixel p .

- ii. Γνωρίζοντας τις συντεταγμένες των positively και negatively affected pixels ορίζονται δύο νέες εικόνες, η εικόνα προσανατολισμού O_n και η εικόνα του πλάτους M_n που αρχικά είναι κενές - όλα τα pixel τιμή 0. Για κάθε νέο ζεύγος $(p_{+ve}(p), p_{-ve}(p))$ που υπολογίζεται, οι πίνακες τροποποιούνται σύμφωνα με τις σχέσεις :

$$O_n(p_{+ve}(p)) = O_n(p_{+ve}(p)) + 1$$

$$O_n(p_{-ve}(p)) = O_n(p_{-ve}(p)) - 1$$

$$M_n(p_{+ve}(p)) = M_n(p_{+ve}(p)) + ||g(p)||$$

$$M_n(p_{-ve}(p)) = M_n(p_{-ve}(p)) + ||g(p)||$$

- iii. Σχηματίζονται ακόμη δύο πίνακες σύμφωνα με τις σχέσεις:

$$\widetilde{O}_n(p) = \begin{cases} O_n(p) & , \text{ αν } O_n(p) < k_n \\ k_n & , \text{ διαφορετικά} \end{cases}$$

$$F_n = \frac{M_n(p)}{k_n} \left(\frac{\widetilde{O}_n(p)}{k_n} \right)^a$$

Όπου η παράμετρος a εκφράζει την αυστηρότητα του μετασχηματισμού (radii strictness) που σημαίνει ότι όσο πιο μεγάλο το a τόσο δεν λαμβάνονται υπ' όψη τα μη συμμετρικά αντικείμενα ενώ το k_n χρησιμοποιείται για την κανονικοποίηση των εικόνων O_n και M_n .

- iv. Τέλος ορίζεται ο πίνακας S_n που εκφράζει την συνεισφορά της ακτίνας n στην συμμετρία της εικόνας και προκύπτει από την σχέση:

$$S_n = F_n * A_n$$

όπου ο A_n είναι μία Gaussian δισδιάστατη συνάρτηση πυκνότητας πιθανότητας που στην διεθνή ορολογία αναφέρεται και ως Gaussian Blur. Η τιμή του μετασχηματισμού για δεδομένη ακτίνα n εκφράζει την συνεισφορά των gradients σε απόσταση n από το εξεταζόμενο pixel.

Επειδή στις περισσότερες περιπτώσεις εξετάζουμε ένα σενάριο ακτινών $[1, N]$ τότε για να λάβουμε υπ' όψη την συνεισφορά κάθε μίας ακτίνας n υπολογίζουμε την μέση τιμή :

$$S = \frac{1}{N} \sum_{n=1}^N S_n$$

Διευκρινίζεται ότι όλοι οι πίνακες που ορίστηκαν προηγουμένως αφορούν μία μόνο ακτίνα $n \in [1, N]$ ενώ ο τελικός πίνακας S είναι μία γενίκευση των αποτελεσμάτων καθώς αφορά όλο το σετ ακτινών $n \in [1, N]$.

Το ζήτημα που προκύπτει σε αυτό το σημείο είναι η εύρεση του βέλτιστου χώρου στον οποίο θα μετατρέψουμε την εικόνα ώστε να πάρουμε το καλύτερο αποτέλεσμα από τον αλγόριθμο FRST.

Σε προηγούμενη παράγραφο αναφερθήκαμε στην σημασία του χώρου CIEL ab και στην ιδιότητά του να κωδικοποιεί σε διαφορετικό κανάλι την ισορροπία μεταξύ πράσινου και κόκκινου χρώματος. Εκμεταλλευόμενοι λοιπόν το χαρακτηριστικό του κόκκινου χρώματος των πίσω φαναριών των αυτοκινήτων που είναι κοινό σε όλα, δοκιμάζουμε να δώσουμε ως είσοδο στον FRST το κανάλι a του CIELab. Στην Εικόνα 9 βλέπουμε 3 frames στον RGB χώρο (1^η γραμμή), στο κανάλι a^* (2^η γραμμή) και το αποτέλεσμα του αλγορίθμου (3^η γραμμή). Αυτό που παρατηρούμε είναι ότι τα πίσω φώτα έχουν αναγνωρισθεί επιτυχώς ως σημεία υψηλής συμμετρίας και γι' αυτό επιλέγουμε να εργαστούμε στο κανάλι a .



Εικόνα 9. Εφαρμογή του αλγορίθμου σε 3 εικόνες από τις βάσεις Caltech 1999, 2001. Πρώτη γραμμή η αρχική εικόνα RGB, στην δεύτερη γραμμή το κανάλι a^* και στην τρίτη γραμμή το αποτέλεσμα του FRST μαρκάροντας τα σημεία με υψηλή συμμετρία με πράσινο χρώμα.

3.5 Otsu's Thresholding

Ένα συχνό ζήτημα στην μηχανική όραση είναι η τμηματοποίηση μίας εικόνας, δηλαδή ο διαχωρισμός των pixels σε 2 ή περισσότερες κατηγορίες (classes). Η πιο δημοφιλής τεχνική είναι αλγόριθμος του Otsu [14] ο οποίος λαμβάνει μία εικόνα ενός καναλιού (πχ grayscale) και χωρίζει τα pixels σε 2 κατηγορίες βάσει μίας τιμής κατωφλίου. Το πλεονέκτημα του Otsu έγκυται στο ότι η τιμή του κατωφλίου δεν είναι προκαθορισμένη αλλά σε κάθε επανάληψη επιλέγεται μία τιμή, χωρίζονται τα pixels σε δύο κατηγορίες και μετράται η διασπορά της κάθε κλάσης. Ως κατώφλι επιλέγεται εκείνη που ελαχιστοποιεί τις διαφορές των κλάσεων ή ισοδύναμα μεγιστοποιεί την συνδιασπορά (απο κοινού διασποράς) των κλάσεων. Για την μαθηματική περιγραφή θεωρούμε μία εικόνα με L pixels τα οποία βάσει του τρέχοντος κατωφλίου th χωρίζονται σε 2 ομάδες ανάλογα με το αν η τιμή του κάθε pixel είναι πάνω ή κάτω από το κατώφλι:

Class 1: $p_1 / \omega_1(th)$, $p_2 / \omega_1(th)$, $p_3 / \omega_1(th)$, ... , $p_k / \omega_1(th)$

Class 2: $p_{k+1} / \omega_2(th)$, $p_{k+2} / \omega_2(th)$, $p_{k+3} / \omega_2(th)$, ... , $p_L / \omega_2(th)$

όπου $\omega_1(th) = \sum_{i=1}^k p_i$ και $\omega_2(th) = \sum_{i=k+1}^L p_i$

Υπολογίζονται οι κατανομές των 2 κλάσεων από τις παραπάνω σχέσεις και κατόπιν η intra-class variance προκύπτει από την σχέση :

$$\sigma_{\omega}^2(th) = \omega_1(th) \sigma_1^2(th) + \omega_2(th) \sigma_2^2(th)$$

όπου $\sigma_1^2(th)$ και $\sigma_2^2(th)$ οι διασπορές των κλάσεων 1 και 2 αντίστοιχα. Ο αλγόριθμος εκτελεί τα παραπάνω βήματα για διάφορα κατώφλια και επιλέγει εκείνο το οποίο ελαχιστοποιεί την ποσότητα $\sigma_{\omega}^2(th)$. Στην συνέχεια παρουσιάζονται δύο περιπτώσεις εφαρμογής του αλγόριθμου του Otsu για μία εικόνα σε grayscale μορφή καθώς και στο κανάλι a^* του χώρου CIELab (Εικόνα 10). Παρατηρούμε ότι ο αλγόριθμος δεν συμπεριφέρεται το ίδιο σε κάθε κανάλι και αυτό είναι λογικό εφόσον κάθε χρωματικός χώρος αναλύει την εικόνα με διαφορετικές συνιστώσες, όπου κάθε μία δίνει έμφαση σε διαφορετικό χαρακτηριστικό της όπως φωτεινότητα, αντίθεση, αποόχρωση κ.α. Αυτό που συμπεραίνουμε είναι ότι ο αλγόριθμος του Otsu, λόγω τ'ότι έχει μεταβλητό κατώφλι υπερτερεί έναντι άλλων αλγορίθμων με σταθερό κατώφλι. Όσον αφορά τον χρωματικό χώρο επιλέγεται κάθε φορά ο βέλτιστος για το αντίστοιχο πρόβλημα.



Εικόνα 10. Εφαρμογή του αλγορίθμου τμηματοποίησης του Otsu σε εικόνα της βάσης Caltech 1999. (a) αρχική εικόνα RGB, (b) grayscale, (c) Otsu στην grayscale, (d) κανάλι a^* του χώρου CIE Lab (e) Otsu στο κανάλι a^* .

3.6 Οπτική ροή

Σε πολλές εφαρμογές της μηχανικής όρασης χρειάζεται να συμπεριλάβουμε στους υπολογισμούς μας την συνιστώσα του χρόνου ώστε να εντοπίσουμε την κίνηση και να διαχωρίσουμε τα κινούμενα στοιχεία από τα στατικά. Ένα σχετικό παράδειγμα είναι η καταγραφή της κυκλοφορίας οχημάτων σε ένα οδικό δίκτυο. Εισάγουμε λοιπόν τον όρο της οπτικής ροής [15] ο οποίος περιγράφει την κίνηση των αντικειμένων μεταξύ διαδοχικών frames που ωφείλεται είτε στην μετατόπιση της κάμερας είτε των ίδιων των αντικειμένων. Κάθε pixel (x, y) της εικόνας μία δεδομένη χρονική στιγμή t έχει ένταση $I(x, y, t)$ ενώ ύστερα από απειροστό χρόνο dt θα έχει ένταση $I(x+dx, y+dy, t+dt)$. Βασικός περιορισμός που τίθεται στον ορισμό της οπτικής ροής είναι ότι η ένταση γειτονικών pixels δεν μεταβάλλεται σε διαδοχικές χρονικές στιγμές και μαθηματικά εκφράζεται από την ισότητα:

$$\bullet \quad I(x, y, t) = I(x + dx, y + dy, t + dt) \quad 3.6.1.$$

όπου αναπτύσσοντας τον δεύτερο όρο κατά Taylor παίρνουμε :

$$\bullet \quad I(x + dx, y + dy, t + dt) = I(x, y, t) + \frac{\partial I}{\partial x} \Delta x + \frac{\partial I}{\partial y} \Delta y + \frac{\partial I}{\partial t} \Delta t + O(\bullet) \quad 3.6.2.$$

Αν οι μετατοπίσεις Δx και Δy είναι μικρές (κάτω των 3 pixels) τότε οι όροι ανώτερης τάξης μπορούν να παραληφθούν. Διαιρώντας τα δύο μέλη με Δt μέσω της (3.6.1) παίρνουμε την γραμμική εξίσωση οπτικής ροής :

$$\bullet \quad \frac{\partial I}{\partial x} v_x + \frac{\partial I}{\partial y} v_y + \frac{\partial I}{\partial t} = 0 \quad 3.6.3.$$

με τις ποσότητες $\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y}, \frac{\partial I}{\partial t}$ είναι τα image gradients για κάθε pixel (x,y) τα οποία είναι γνωστά και σκοπός είναι η εύρεση των v_x και v_y , δηλαδή των ταχυτήτων κατά x και y σε κάθε θέση. Παρατηρούμε ότι η εξίσωση (3.6.3) δεν μπορεί να επιλυθεί με συμβατικές μεθόδους γραμμικών διαφορικών εξισώσεων καθώς έχουμε μία σχέση με δύο αγνώστους. Στην συνέχεια θα εισάγουμε δύο βασικούς αλγορίθμους για την αντιμετώπιση αυτού του ζητήματος και την εύρεση της οπτικής ροής.

3.6.1 Πυκνή οπτική ροή - Αλγόριθμος Farneback

Ο Gunnar Farneback εισάγει στο [16] έναν αλγόριθμο για τον υπολογισμό της πυκνής οπτικής ροής, δηλαδή την εκτίμηση της μετατόπισης κάθε pixel της εικόνας μεταξύ δύο διαδοχικών frames. Αρχικά για κάθε pixel εκτιμάται η γειτονιά του και στα δύο frames μέσω του πολυωνυμικού αναπτύγματος επιλέγοντας την χρήση του τετραγωνικού πολυωνύμου:

$$f(x) = x^T A x + b^T x + c \quad 3.6.4.$$

όπου A συμμετρικός πίνακας, b διάνυσμα και c σταθερά τα οποία υπολογίζονται από μέθοδο σταθμισμένων ελαχίστων τετραγώνων.

Στην συνέχεια εξετάζεται τι θα συμβεί στο πολυώνυμο αυτό αν υποστεί μία καθολική μετατόπιση d . Έστω το αρχικό πολυώνυμο

$$f_1(x) = x^T A x + b^T x + c \quad 3.6.5.$$

και $f_2(x)$ το μετατοπισμένο πολυώνυμο κατά d τότε :

$$f_2(x) = f_1(x - d) = (x - d)^T A_1 (x - d) + b_1^T (x - d) + c_1 = x^T A_2 x + b_2^T x + c_2 \quad 3.6.6.$$

Εξισώνοντας τις σχέσεις (3.6.5) και (3.6.6) έχουμε ότι

$$d = -\frac{1}{2} A_1^{-1} (b_2 - b_1) \quad 3.6.7.$$

και βρίσκουμε λοιπόν την μετατόπιση κάθε γειτονιάς της εικόνας.

Σε πρακτικές εφαρμογές μία εικόνα δεν μπορεί να περιγραφεί από ένα και μόνο πολυωνυμικό ανάπτυγμα, όπως και δεν μπορεί η μετατόπιση d να είναι καθολική για κάθε υποπεριοχή της εικόνας. Διαφορετικά εισάγονται σφάλματα στην μελέτη. Για πιο αξιόπιστο αλγόριθμο γίνεται προσέγγιση της εικόνας τοπικά, χρησιμοποιώντας συντελεστές $A_k(x)$, $b_k(x)$, $c_k(x)$ και μετατόπιση $d_k(x)$ εξαρτώμενα από το x με $k=1,2$ για τα δύο διαδοχικά frames. Ορίζουμε :

$$A(x) = \frac{A_1(x)+A_2(x)}{2} \quad 3.6.8.$$

$$\Delta b(x) = -\frac{1}{2}(b_2(x) - b_1(x)) \quad 3.6.9.$$

και τελικά προκύπτει η ισοδύναμη σχέση της (3.6.7) αλλά αναφερόμενη σε τοπικά μεγέθη :

$$A(x)d(x) = \Delta b(x) \quad 3.6.10.$$

Το πεδίο μετατόπισης $d(x)$ θεωρείται ότι μεταβάλλεται με αργό ρυθμό και η τιμή του προσδιορίζεται για μία γειτονιά I του x ώστε να ικανοποιεί όσο το δυνατόν καλύτερα την εξίσωση 3.6.10. Με άλλα λόγια αναζητείται το $d(x)$ εκείνο που ελαχιστοποιεί την σχέση :

$$\sum_{\Delta x \in I} \omega(\Delta x) ||A(x + \Delta x)d(x) - \Delta b(x + \Delta x)||^2 \rightarrow \min \quad 3.6.11.$$

με $\omega(\Delta x)$ να είναι μία σταθμισμένη συνάρτηση για την κάθε γειτονιά.

Το ελάχιστο προκύπτει για την τιμή

$$d(x) = (\sum \omega A^T A)^{-1} (\sum \omega A^T \Delta b) \quad 3.6.12.$$

όπου είναι και το τελικό πεδίο μετατόπισης για κάθε γειτονιά της εικόνας.

Όλες οι παραπάνω διεργασίες είναι ενσωματωμένες σε μία συνάρτηση της OpenCV η οποία δέχεται ως είσοδο τα δύο διαδοχικά frames, το παράθυρο επεξεργασίας καθώς και κάποιες άλλες παραμέτρους και επιστρέφει ένα tuple με την υπολογισμένη ροή για κάθε pixel υπό μορφή μιγαδικού στο καρτεσιανό επίπεδο [17]. Για καλύτερη διαχείριση μπορούμε να την μετατρέψουμε σε πολική μορφή και να πάρουμε το πλάτος και την γωνία.

Αντιλαμβανόμαστε ότι η μέθοδος αυτή θα υστερεί σε ταχύτητα καθώς ο υπολογισμός της οπτικής ροής για κάθε pixel της εικόνας αυξάνει την πολυπλοκότητα του αλγορίθμου.

3.6.2 Αραιή οπτική ροή

Βασιζόμενοι στο μειονέκτημα της παραπάνω προσέγγισης εισάγουμε την έννοια της αραιής οπτικής ροής. Σύμφωνα με αυτή ο υπολογισμός δεν γίνεται για όλα τα pixels της εικόνας αλλά μόνο για μερικά σημεία ενδιαφέροντος προσπαθώντας έτσι να μειώσουμε την πολυπλοκότητα. Σε μία εικόνα ως σημεία ενδιαφέροντος ορίζουμε εκείνα που είναι πιθανότερο να εμφανίσουν υψηλή ροή δηλαδή ακμές και γωνίες αντικειμένων. Στην συνέχεια παρουσιάζουμε συνοπτικά δύο τεχνικές προσδιορισμού των σημείων αυτών.

3.6.2.a Harris corner detection

Η τεχνική αυτή [18], [19] προτάθηκε από τους Harris και Stephens και βασίζεται στον υπολογισμό των gradients μίας εικόνας. Συγκεκριμένα λαμβάνεται μία εικόνα gray-scale μορφής και προσδιορίζονται για κάθε pixel οι τιμές των μερικών παραγώγων της έντασης ως προς x και y. Για τον χαρακτηρισμό ενός σημείου ως γωνία ορίζεται ο παράγοντας αξιολόγησης

$$R = \det(M) - k \cdot (\text{trace}(M))^2 \quad 3.6.13.$$

όπου k είναι μία εμπειρική σταθερά και

$$M = \begin{bmatrix} f_x^2 & f_x f_y \\ f_x f_y & f_y^2 \end{bmatrix} \quad 3.6.14.$$

με $f(x,y)$ η τιμή της εικόνας στην θέση (x,y) .

Αν λ_1 και λ_2 είναι οι ιδιοτιμές του M τότε $\det(M) = \lambda_1 \lambda_2$ και $\text{trace}(M) = \lambda_1 + \lambda_2$ τότε το κριτήριο απόφασης παίρνει την μορφή :

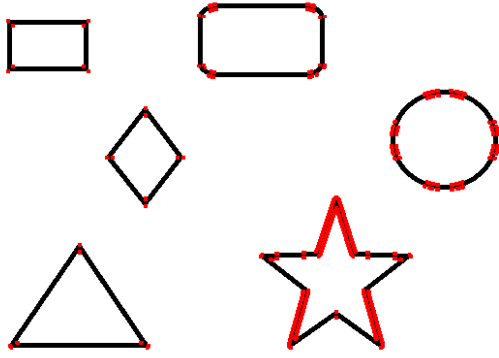
$$R = \lambda_1 \lambda_2 - k \cdot (\lambda_1 + \lambda_2)^2 \quad 3.6.15.$$

Και διακρίνουμε 3 περιπτώσεις :

Αν R είναι μεγάλο τότε το σημείο χαρακτηρίζεται ως γωνία.

- a) Αν $R < 0$ τότε $\lambda_1 \ll \lambda_2$ ή $\lambda_2 \ll \lambda_1$ και το σημείο χαρακτηρίζεται ως ακμή.
- b) Αν $|R|$ είναι μικρό τότε το σημείο είναι ομαλό.

Ένα μειονέκτημα της τεχνικής αυτής είναι η αδυναμία να διαχειριστεί περίπλοκες και false positive γωνίες. Η μέθοδος αυτή υλοποιείται μέσω κατάλληλης συνάρτησης στην OpenCV και απλή εφαρμογή της φαίνεται στην Εικόνα 11.



Εικόνα 11. Απλή εφαρμογή της τεχνικής *Harris Corners Detection* σε γεωμετρικά σχήματα.

3.6.2.b Shi Tomasi corner detection

Η τεχνική των Shi Tomasi [20] βασίζεται στον αλγόριθμο του Moravec ο οποίος λειτουργεί ως εξής: Δημιουργεί ένα παράθυρο στην εικόνα και το μετακινεί ελαφρώς προς όλες τις κατευθύνσεις, υπολογίζοντας κάθε φορά την μέση ισχύ του. Για κάθε μετάβαση ελέγχει αν η διαφορά των ισχύων υπερβαίνει ένα κατώφλι και τότε το pixel στο κέντρο του παραθύρου χαρακτηρίζεται ως γωνία. Έστω pixel (x,y) με ένταση $f(x,y)$ τότε αν μετακινηθεί κατά (u,v) τότε η ένταση εκφράζεται ως :

$$E_{u,v}(x, y) = \sum_{u,v} \omega_{u,v} [f(x + u, y + v) - f(x, y)]^2 \quad 3.6.16.$$

Όπου $\omega_{u,v}$ είναι μία Gaussian συνάρτηση για την αντοχή στον θορύβο που εισάγουν οι μερικές παράγωγοι πρώτης τάξης.

Το μειονέκτημα του αλγορίθμου του Moravec είναι ότι η μετατόπιση γίνεται προς 8 βασικές κατευθύνσεις με αποτέλεσμα να μην λειτουργεί σωστά όταν η εικόνα περιστραφεί κατά μικρή γωνία. Με βάση αυτή την φιλοσοφία οι Shi Tomasi τροποποιούν την σχέση (3.6.16) ως εξής :

$$E_{u,v}(x, y) = \sum_{u,v} \omega_{u,v} [(u, v)[M](u, v)^T]$$

όπου

$$M = \begin{bmatrix} f_x^2 & f_x f_y \\ f_x f_y & f_y^2 \end{bmatrix}$$

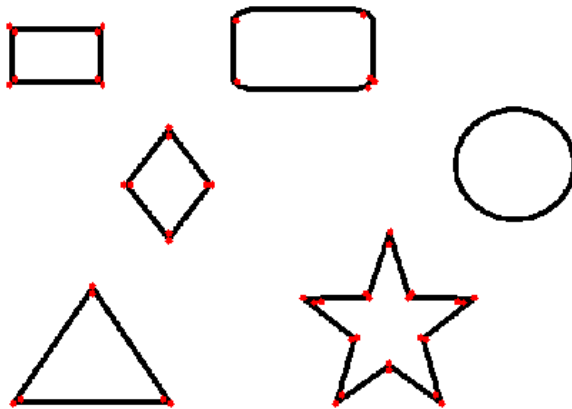
με $f(x,y)$ η τιμή της εικόνας στην θέση (x,y) .

Η βασική διαφοροποίηση του από την τεχνική του Harris είναι ο διαφορετικός παράγοντας αξιολόγησης R . Έστω λ_1 και λ_2 οι ιδιοτιμές του πίνακα M τότε ορίζουμε το $R = \min(\lambda_1, \lambda_2)$

καθώς και ένα κατώφλι λ_{min} [19]. Ανάλογα με την τιμή του R διακρίνουμε τις εξής περιπτώσεις :

- a) Αν το R μεγαλύτερο του κατωφλίου τότε το σημείο χαρακτηρίζεται ως γωνία.
- b) Αν μόνο ένα εκ των λ_1, λ_2 είναι μικρότερα του κατωφλίου τότε το σημείο είναι ακμή.
- c) Αν και τα δύο μικρότερα του κατωφλίου τότε είναι ομαλό σημείο.

Παρακάτω στην Εικόνα 12 έχουμε το αποτέλεσμα της τεχνικής Shi Tomasi σε απλή εικόνα γεωμετρικών σχημάτων.



Εικόνα 12. Εφαρμογή της τεχνικής Shi Tomasi σε γεωμετρικά σχήματα.

3.6.2.c Αλγόριθμος Lucas Kanade

Ο αλγόριθμος αυτός χρησιμοποιείται για την εκτίμηση της αραιής οπτικής ροής μεταξύ 2 διαδοχικών frames [22], [23], [24]. Έστω ένα pixel με συντεταγμένες (x,y) στο πρώτο frame και ένταση $I(x,y)$ τότε σύμφωνα με την βασική εξίσωση της οπτικής ροής (3) που περιγράψαμε στην παράγραφο 2.5 θα ισχύει :

$$I_x u + I_y v = -I_t \quad 3.6.17.$$

όπου $[u \ v]$ είναι το διάνυσμα της ροής. Η εξίσωση (3.6.17) δεν έχει μοναδική λύση αλλά άπειρες οπότε για την απόκτηση περισσότερης πληροφορίας ο αλγόριθμος θεωρεί μία γειτονιά $N \times N$ γύρω από το (x,y) όπου η ροή είναι ομαλή και πρακτικά παραμένει ίδια. Τότε η (3.6.17) θα πάρει την μορφή:

$$Au = b \quad 3.6.18.$$

$$\text{όπου } A = \begin{bmatrix} I_{x1} & I_{y1} \\ \vdots & \vdots \\ I_{xN} & I_{yN} \end{bmatrix}, \quad u = \begin{bmatrix} u \\ v \end{bmatrix}, \quad b = \begin{bmatrix} -I_{t1} \\ \vdots \\ -I_{tN} \end{bmatrix} \quad 3.6.19.$$

Το σύστημα αυτό $N \times 2$, έχει άπειρες λύσεις και για να βρεθεί το διάνυσμα $[u \ v]^T$ πρέπει να υποβιβαστεί σε πρόβλημα 2×2 . Η τεχνική που χρησιμοποιείται από τους Lucas - Kanade είναι η μέθοδος προσαρμογής ελαχίστων τετραγώνων (Least Square Fit) η οποία εφαρμόζεται συχνά για υπερ-ορισμένα προβλήματα (overdetermined problems) καθώς και για την εύρεση ενός μαθηματικού μοντέλου προσέγγισης μίας ομάδας παρατηρήσεων. Σκοπός είναι να ελαχιστοποιηθεί η παράσταση $\|Au-b\|^2$ και τελικά μέσω της μεθόδου ελαχίστων τετραγώνων προκύπτει:

$$(A^T A)u = A^T b \quad 3.6.20.$$

και αντικαθιστώντας τα A και b έχουμε :

$$\begin{bmatrix} \sum I_x I_x & \sum I_x I_y \\ \sum I_x I_y & \sum I_y I_y \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = - \begin{bmatrix} \sum I_x I_t \\ \sum I_y I_t \end{bmatrix} \quad 3.6.21.$$

Το αρχικό σύστημα N εξισώσεων με 2 αγνώστους έχει μετατραπεί σε απλό σύστημα 2×2 και μπορούμε πλέον να υπολογίσουμε το διάνυσμα ροής $[u \ v]$ για το σημείο (x,y) . Ένα ερώτημα που τίθεται είναι με ποιον τρόπο θα επιλεγούν τα pixels για τα οποία θα υπολογιστεί η ροή με την παραπάνω διαδικασία. Τα σημεία αυτά επιλέγονται βάσει κάποιας μεθόδου εύρεσης σημείων ενδιαφέροντος πχ τεχνική Harris ή Shi Tomasi και θα είναι γωνίες αντικειμένων ή τμήματα ακμών τα οποία είναι πιθανότερο να έχουν υψηλή οπτική ροή. Το αποτέλεσμα του αλγορίθμου είναι το διάνυσμα της οπτικής ροής $[u \ v]$ στα σημεία ενδιαφέροντος της εικόνας και το οποίο δίνει μία εκτίμηση της μετατόπισης των αντικειμένων. Πιο αναλυτικά προσπαθεί να εκτιμήσει προς ποια κατεύθυνση έχει κινηθεί η κάθε γειτονιά εξετάζοντας τις μεταβολές στην ένταση, οι οποίες μπορούν να εξηγηθούν από τα γνωστά gradients (I_x, I_y, I_t) αυτής της γειτονιάς.

Οι αλγόριθμοι εντοπισμού αντικειμένων ή ανίχνευσης κίνησης πρέπει να χαρακτηρίζονται από ακρίβεια (accuracy) και ευαισθησία (robustness). Για την ανίχνευση μεγάλων κινήσεων μία λύση είναι να αυξηθεί το μέγεθος του παραθύρου η οποία ωστόσο οδηγεί σε μείωση της ακρίβειας. Για να εξασφαλιστεί μία ισορροπία μεταξύ ακρίβειας και ευαισθησίας εισάγουμε την τεχνική της αναπαράστασης των εικόνων σε μορφή πυραμίδας [25].

Έστω μίας εικόνα I διαστάσεων $n_x \times n_y$. Ορίζουμε στο μηδενικό επίπεδο - βάση το I^0 που ταυτίζεται με την αρχική εικόνα και έχει διαστάσεις $n_{x0} = n_x$ και $n_{y0} = n_y$. Κάθε ανώτερο επίπεδο L που σχηματίζεται εξαρτάται από το αμέσως προηγούμενο ως εξής :

$$\begin{aligned}
\bullet \quad I^L(x, y) = & \frac{1}{4} I^{L-1}(2x, 2y) + \\
& + \frac{1}{8} [I^{L-1}(2x-1, 2y) + I^{L-1}(2x+1, 2y) + I^{L-1}(2x, 2y-1) + I^{L-1}(2x, 2y+1) \\
& + \frac{1}{16} [I^{L-1}(2x-1, 2y-1) + I^{L-1}(2x+1, 2y-1) + I^{L-1}(2x-1, 2y+1) + \\
& + I^{L-1}(2x+1, 2y+1)]
\end{aligned}
\tag{3.6.22}$$

και με διαστάσεις

$$n_x^L \leq \frac{n_x^{L-1} + 1}{2}, \quad n_y^L \leq \frac{n_y^{L-1} + 1}{2} \tag{3.6.23}$$

όπου $L=1,2,3,\dots,K$ το τρέχον επίπεδο και $L-1$ το αμέσως προηγούμενο. Το ύψος της πυραμίδας επιλέγεται ευρετικά ανάλογα με το πρόβλημα, ενώ από τις σχέσεις (3.6.23) παρατηρούμε ότι για κάθε νέο επίπεδο που δημιουργείται οι διαστάσεις μειώνονται περίπου κατά το ήμισυ. Ο αλγόριθμος Lucas Kanade εφαρμόζεται πρώτα στην εικόνα I^K ύστερα το αποτέλεσμα περνά στο αμέσως κατώτερο στάδιο I^{K-1} και η διαδικασία αυτή συνεχίζεται ως την βάση της πυραμίδας I^0 , δηλαδή την αρχική εικόνα. Όσο βρισκόμαστε σε ανώτερα επίπεδα στην πυραμίδα η ανάλυση είναι μικρή με αποτέλεσμα να ανιχνεύονται οι μεγάλες κινήσεις. Ένα σημαντικό πλεονέκτημα του αλγορίθμου Lucas Kanade είναι ότι δεν επεξεργάζεται όλα τα pixels των εικόνων αλλά μόνο τα σημεία ενδιαφέροντος μειώνοντας θεαματικά την πολυπλοκότητα και τον χρόνο εκτέλεσης.

3.7 Random Walker

Ο αλγόριθμος Random Walker [26] χρησιμοποιείται για image segmentation δηλαδή για την κατηγοριοποίηση των pixels μίας εικόνας σε M κατηγορίες. Εκτός από τα άγνωστα δεδομένα, πρέπει να του δοθούν ως είσοδος και μερικά γνωστά pixels (seeds) τα οποία είναι εκ των προτέρων τοποθετημένα σε μία από τις M κατηγορίες. Επομένως είναι ένας supervised αλγόριθμος που λειτουργεί με την εξής φιλοσοφία : Για κάθε unlabeled pixel της εικόνας ξεκινάει ένας τυχαίος περπατητής και διαγράφει μία τυχαία πορεία. Η κατηγορία στην οποία θα ενταχθεί το pixel αναχώρησης είναι εκείνη με την μεγαλύτερη πιθανότητα πρώτης άφιξης. Η εύρεση αυτής της πιθανότητας ισοδυναμεί με την επίλυση του προβλήματος Dirichlet με οριακές συνθήκες στις θέσεις των seeds [27].

Η παραπάνω διαδικασία προφανώς δεν είναι απολύτως τυχαία καθώς τότε δεν θα είχε πρακτικό νόημα. Επιλέγεται λοιπόν η μετάβαση μεταξύ γειτονικών pixels να εξαρτάται από

μία πιθανότητα η οποία υπολογίζεται μέσω της διαφοράς έντασης μεταξύ αυτών των pixels. Αυτό γίνεται περισσότερο κατανοητό αναπαριστώντας το πρόβλημα με την θεωρία γράφων. Ορίζουμε ένα δικοτευθυντικό γράφο $G=(V,E)$ με n κόμβους V και m ακμές E . Δύο κόμβοι u_i και u_j θα ενώνονται με την ακμή e_{ij} η οποία έχει ένα βάρος w_{ij} και αυτό θα εξαρτάται από τις εντάσεις των δύο κόμβων :

$$w_{ij} = \exp(-\beta(g_i - g_j)^2)$$

με g_i και g_j να είναι οι εντάσεις των 2 pixels και β μία παράμετρος που ορίζεται από τον χρήστη. Έχοντας ορίσει τα βάρη κάθε ακμής πλέον η μετάβαση μεταξύ των pixels δεν είναι απολύτως τυχαία αλλά εξαρτάται από την διαφορά των εντάσεών τους. Κατόπιν της μετατροπής της εικόνας σε γράφο τα βήματα που εκτελεί ο αλγόριθμος είναι τα εξής :

i. Κατασκευάζεται ένας Laplacian πίνακας :

$$L_{ij} = \begin{cases} d_i, & \text{αν } i = j \\ -w_{ij}, & \text{αν } u_i \text{ και } u_j \text{ γειτονικοί} \\ 0, & \text{αλλού} \end{cases}$$

όπου $d_i = \sum w_{ij}$ και ονομάζεται βαθμός του κόμβου u_i .

ii. Ορίζεται ο edge-node incidence πίνακας

$$A_{e_{ij}u_k} = \begin{cases} +1, & \text{αν } i = k \\ -1, & \text{αν } j = k \\ 0, & \text{αλλού} \end{cases}$$

με διαστάσεις $m \times n$ που μπορεί να ερμηνευτεί ως combinatorial gradient operator και ο A^T ως combinatorial divergence operator. Κατόπιν ορίζεται ο διαγώνιος πίνακας C (constitutive matrix) διαστάσεων $m \times m$ ο οποίος περιέχει στην κύρια διαγώνιο τα βάρη όλων των ακμών και τέλος ο Laplacian πίνακας $L=A^T A$.

iii. Σχηματίζεται το ολοκλήρωμα Dirichlet :

$$D[x] = \frac{1}{2}(Ax)^T C(Ax) = \frac{1}{2}x^T Lx = \frac{1}{2}\sum_{e_{ij}} w_{ij}(x_i - x_j)^2 \quad 3.7.3$$

όπου η x είναι μία αρμονική (combinatorial harmonic function) που ελαχιστοποιεί την (3.7.3) και e_{ij} η ακμή που συνδέει τους κόμβους u_i και u_j

- iv. Διαχωρίζονται οι κόμβοι σε δύο κατηγορίες V_M (marked/seeded) και V_U (unseeded) έτσι ώστε $V_M \cup V_U = V$ και $V_M \cap V_U = \emptyset$. Αυθαίρετα θεωρούμε ότι στα L και x είναι τοποθετημένοι πρώτα οι seeds κόμβοι και μετά οι unseeded, χωρίς αυτή η θεώρηση να παραβιάζει την γενικότητα. Κατόπιν του διαχωρισμού η 3.7.3 γράφεται :

$$D[x_U] = \frac{1}{2} [x_M^T \ x_U^T] \begin{bmatrix} L_M & B \\ B^T & L_U \end{bmatrix} \begin{bmatrix} x_M \\ x_U \end{bmatrix} \quad 3.7.4$$

με τα L_M και L_U να είναι οι τιμές των seeded και unseeded κόμβων του L , B είναι αποτέλεσμα του decomposition και ουσιαστικά περιέχει την σχέση μεταξύ seeded και unseeded κόμβων αντιστοιχα, ενώ x_M και x_U είναι οι πιθανότητες των seeded και unseeded κόμβων. Αν διαφορίσουμε την (3.7.4) ως προς x_U παίρνουμε :

$$L_U x_U = -B^T x_M$$

- v. Έστω πρόβλημα κατηγοριοποίησης σε K κλάσεις και x_i^s η πιθανότητα ένας unlabeled κόμβος να ανήκει στην κλάση s όπου $s \in Z$ με $0 < s \leq K$. Ορίζεται η συνάρτηση $Q(u_j) = s$ η οποία ορίζει τα labels για κάθε seeded κόμβο καθώς και η ποσότητα :

$$m_j^s = \begin{cases} 1, & \text{αν } Q(u_j) = s \\ 0, & \text{αν } Q(u_j) \neq s \end{cases}$$

διάστασης $|V_M| \times 1$ για κάθε κόμβο $u_j \in V_M$. Το πρόβλημα Dirichlet για την κλάση s επιλύεται μέσω της σχέσης :

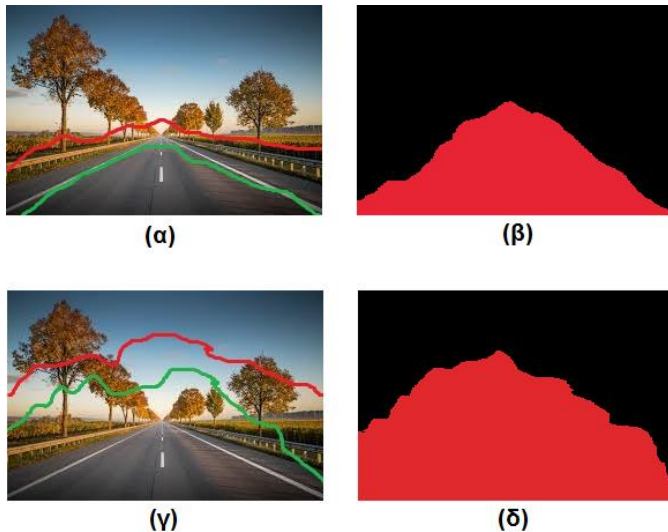
$$L_U x_s = -B^T m_s$$

και πιο συγκεκριμένα η η πιθανότητα ο unlabeled κόμβος x_i^s να φτάσει πρώτη φορά στην κλάση s δίνεται από την σχέση:

$$L_U x_i^s = -B^T m_s$$

Ένα σημαντικό σημείο που πρέπει να τονιστεί είναι ότι για να πάρουμε αξιόπιστα αποτελέσματα θα πρέπει η τοποθέτηση των αρχικών seeds να γίνει πολύ προσεκτικά από τον χρήστη. Αυτό προϋποθέτει να γίνει σωστή επεξεργασία της εικόνας και να εξαχθούν όσο το δυνατόν περισσότερα χαρακτηριστικά της (ακμές, συμμετρίες κ.α). Για παράδειγμα στην Εικόνα 13 συναντάται ένα πρόβλημα κατηγοριοποίησης δύο κλασέων (δρόμος και μη δρόμος). Αν η τοποθέτηση των seeds (α) είναι ακριβής τότε το αποτέλεσμα είναι αξιόπιστο και ανταποκρίνεται στην πραγματικότητα (β). Αντίθετα μία λιγότερο καλή τοποθέτηση (γ)

οδηγεί σε φτωχό αποτέλεσμα (δ).



Εικόνα 13. Στην 1η στήλη δύο περιπτώσεις διαφορετικής τοποθέτησης των *road seeds* (πράσινη γραμμή) και των *non-road seeds* (κόκκινη γραμμή). Στην 2η στήλη τα αντίστοιχα αποτελέσματα του *Random Walker*.

3.8 Εκτίμηση Ομοιότητας

Συχνά στις εφαρμογές μηχανικής όρασης καλούμαστε να εκτιμήσουμε κατά πόσο μία εικόνα ή κάποιο τμήμα της είναι όμοιο με ένα πρωτότυπο (*template*). Παρακάτω παρουσιάζονται τα πιο δημοφιλή και ευρέως γνωστά μέτρα που συναντώνται στην βιβλιογραφία.

3.8.1 Mean Absolute Error (MAE)

Όπως περιγράφεται και στο [28], για δύο εικόνες A και B ίδιων διαστάσεων $h \times w$ τότε ορίζεται το μέσο απόλυτο σφάλμα ως εξής:

$$MAE = \frac{1}{h \times w} \sum_{i=1}^h \sum_{j=1}^w |A_{ij} - B_{ij}|$$

Πρακτικά υπολογίζεται η διαφορά των τιμών των pixels στις αντίστοιχες θέσης και διαιρείται με το πλήθος των pixels της εικόνας για να προκύψει η μέση τιμή. Ο δείκτης τείνει στο 0 αν η εικόνα ταυτίζεται με την πρωτότυπη, ενώ αν διαφέρουν σε μεγάλο βαθμό τότε τείνει στο άπειρο. Εξ' ορισμού φαίνεται ότι το μέτρο αυτό υστερεί όταν δύο εικόνες είναι παρόμοιες αλλά η μία έχει περιστραφεί ελαφρώς, καθώς σε αυτή τη περίπτωση το αποτέλεσμα θα βγαίνει πολύ μεγάλο, κάτι που δεν ανταποκρίνεται στην πραγματικότητα.

3.8.2 Structural Similarity Index (SSIM)

Το μέσο απόλυτο σφάλμα που ορίστηκε παραπάνω δέχεται ως παράμετρο μόνο τις τιμές των pixels με αποτέλεσμα υπό ορισμένες συνθήκες να κρίνεται αναξιόπιστο. Ορίζεται λοιπόν ο δείκτης Structure Similarity Index [29] ο οποίος όπως φανερώνει και το όνομα εμπριέχει την πληροφορία της δομής των εικόνων. Συγκεκριμένα αν έχουμε μία εικόνα C ενός καναλιού και ένα υπόδειγμα T (template), τότε ο δείκτης SSIM λαμβάνει υπ'οψη την μέση τιμή των pixels της κάθε εικόνας μ_c και μ_T , τις τυπικές αποκλίσεις σ_c και σ_T καθώς και την συνδιασπορά των δύο εικόνων σ_{cT} . Περιγράφεται από την σχέση :

$$SSIM = \frac{(2\mu_c\mu_T + K1)(2\sigma_{cT} + K2)}{(\mu_c^2 + \mu_T^2 + K1)(\sigma_c^2 + \sigma_T^2 + K2)}$$

όπου οι $K1$ και $K2$ είναι σταθερές οι οποίες τοποθετούνται για να μην μηδενιστεί ο παρονομαστής και ορίζονται ως $K1 = (k1L)^2$ και $K2 = (k2L)^2$ όπου L είναι το εύρος των τιμών των pixels της εικόνας και $k1, k2$ απειροστά μικρές σταθερές. Τέλος τιμές του δείκτη κοντά στο 0 υποδηλώνουν μικρή ομοιότητα ενώ κοντά στο 1 απόλυτη ομοιότητα.

3.8.3 Διαφορές MAE-SSIM

Οι δείκτες που περιγράψαμε παραπάνω δεν παρουσιάζουν ίδια συμπεριφορά, λόγω του διαφορετικού μαθηματικού ορισμού τους [30], [31]. Ο MAE εφαρμόζεται καθολικά σε ολόκληρη την εικόνα και παρουσιάζει ασταθή συμπεριφορά καθώς μεγάλες αποστάσεις στις τιμές των pixels δεν σημαίνουν απαραίτητα ότι τα περιεχόμενα των pixels διαφέρουν σημαντικά. Για παράδειγμα στην Εικόνα 14, ο δείκτης θα λάβει μεγάλη τιμή ενώ στην πραγματικότητα πρόκειται για αντικείμενα ίδιου σχήματος που διαφέρουν ελάχιστα στην απόχρωση. Επίσης έχει μεγάλο εύρος καθώς οι τιμές που λαμβάνει είναι από 0 (τέλεια ομοιότητα) έως και κάποια πολύ μεγάλη τιμή (μικρή ομοιότητα).

Αντίθετα ο SSIM υπολογίζει τις αλλαγές στην δομή της εικόνας. Εξετάζονται επαναλαμβανόμενα υποπεριοχές, δηλαδή γειτονικά pixels με ισχυρή αλληλοεξάρτηση που δίνουν σημαντικές πληροφορίες για την δομή των αντικειμένων στο frame. Επιπλέον ο δείκτης SSIM παίρνει τιμές στο κλειστό διάστημα $[-1,1]$, με το 1 να αντιπροσωπεύει την απόλυτη ομοιότητα των 2 εικόνων.



Εικόνα 14. Παράδειγμα αδυναμίας του δείκτη ομοιότητας MAE

4. Εντοπισμός προπορευόμενου οχήματος

4.1 Εισαγωγή

Μία από τις σημαντικότερες λειτουργίες ενός συστήματος υποβοήθησης της οδήγησης είναι η έγκαιρη και αξιόπιστη αναγνώριση του προπορευόμενου οχήματος. Με τον τρόπο αυτό αφενώς διασφαλίζεται η τήρηση των προβλεπόμενων αποστάσεων ασφαλείας και αφετέρου σε περίπτωση απότομου φρεναρίσματος του μπροστινού οχήματος οδηγός μπορεί να επέμβει ώστε να ακινητοποιήσει το όχημα και να αποφευχθεί η σύγκρουση. Το σύστημα που αναπτύχθηκε στα πλαίσια της εργασίας βασίζεται κυρίως στο χαρακτηριστικό του χρώματος και της συμμετρίας.

4.2 Σχετική βιβλιογραφία

Το ζήτημα της ανίχνευσης του προπορευόμενου έχει γίνει αντικείμενο εκτενούς μελέτης από την επιστημονική κοινότητα και κατ' επέκταση έχουν υιοθετηθεί πολλές διαφορετικές προσεγγίσεις. Στις δημοσιεύσεις [32], [33] το πρόβλημα αντιμετωπίζεται από την σκοπιά της μηχανικής όρασης. Σε τέτοιες περιπτώσεις το πρόβλημα χωρίζεται σε 2 σκέλη: α) Hypothesis Generation και β) Hypothesis Verification. Για το Hypothesis Generation έχουν προταθεί διάφορες προσεγγίσεις που χωρίζονται σε τρεις γενικές κατηγορίες: 1) knowledge based 2) stereo based 3) motion based.

Στην πρώτη αξιοποιούνται χαρακτηριστικά όπως το χρώμα, η συμμετρία, οι ακμές, η σκιά και τα πίσω φώτα του οχήματος. Στην δεύτερη κατηγορία γίνεται προσπάθεια στερεοσκοπικής απεικόνισης της δισδιάστατης εικόνας που λαμβάνεται στην είσοδο και τέλος στην τρίτη κατηγορία γίνεται υπολογισμός του μεγέθους της οπτικής ροής για την εκτίμηση της σχετικής κίνησης. Όσον αφορά το στάδιο Hypothesis Verification η επαλήθευση των υποψηφίων γίνεται με δύο τρόπους. Συγκρίνεται το αποτέλεσμα με μία εικόνα αναφοράς (template matching) ή με χρήση μεθόδων αναγνώρισης προτύπων κατασκευάζεται ένα σύστημα κατηγοριοποίησης ενός δείγματος σε «όχημα» ή «μη-όχημα».

Οι Gao, Duan, Zheng και Yang [32] εκμεταλλεύονται τις διαγραμμίσεις του δρόμου και χαρακτηριστικά των οχημάτων. Συγκεκριμένα για την εύρεση υποψηφίων βασίζονται στις διαγραμμίσεις του δρόμου και με χρήση ενός threshold εντοπίζουν την σκιά που δημιουργείται κάτω από το προπορευόμενο όχημα (self shadow). Ύστερα γίνεται χρήση του μη γραμμικού φίλτρου SUSAN για τον εντοπισμό των ακμών και την δημιουργία πιο αξιόπιστων περιοχών ενδιαφέροντος. Στο τελευταίο στάδιο της επαλήθευσης υποψηφίων λαμβάνεται υπ' όψη η συμμετρία του πίσω μέρος των οχημάτων ως προς την μεσοκάθετο. Αυτή η προσέγγιση λόγω της στενής σύνδεσης της με τις διαγραμμίσεις του δρόμου και την σκιά είναι ευάλωτη σε συνθήκες χαμηλής φωτεινότητας καθώς και για επαρχιακά οδικά δίκτυα όπου η διαγράμμιση έχει αλλοιωθεί ή δεν υπάρχει. Παρόμοια φιλοσοφία έχει και η

μέθοδος Kaniktar, Bharti και Hivarkar [34] στην οποία το στάδιο Hypothesis Generation τις σκίες των οχημάτων (self shadows) ενώ η επαλήθευση γίνεται με χρήση appearance based μεθόδων. Συγκεκριμένα εφαρμόζεται ο μετασχηματισμός Harr Wavelets για την ανίχνευση των ακμών του αυτοκινήτου και ύστερα τροφοδοτείται η περιοχή ενδιαφέροντος σε ένα νευρωνικό δίκτυο. Εξαιτίας της χρήσης περιορισμένων χαρακτηριστικών του οχήματος αυτή η προσέγγιση παρουσιάζει αδυναμία διάκρισης αντικειμένων της εικόνας από πραγματικά οχήματα.

Βασίζόμενοι στην παρατήρηση ότι η μεγάλη ποικιλία χρωμάτων και σχημάτων στα οχήματα δεν ευνοεί την χρήση μεθόδων μηχανικής όρασης οι Kato, Ninomiya και Masaki στην [35] στράφηκαν αποκλειστικά σε appearance based τεχνικές και συγκεκριμένα σε έναν ταξινομητή αντικειμένων. Για την εκπαίδευση του χρησιμοποιούνται δύο σέτ εικόνων, μία που περιέχει «οχήματα» και μία που αποτελείται από «μη-οχήματα». Στο στάδιο της δοκιμής λαμβάνεται ένα frame και σκανάρεται ανά περιοχές με χρήση ενός παραθύρου, ύστερα γίνεται κανονικοποίηση του παραθύρου αυτού και μετατρέπεται σε ένα διάνυσμα x_{mxm} το οποίο θα αποτελεί και την είσοδο του ταξινομητή. Κάθε υποπεριοχή της εικόνας χαρακτηρίζεται ως όχημα ή μη-όχημα και αφού ολοκληρωθεί η σάρωση σε όλη την εικόνα προσδιορίζονται όλα τα προπορευόμενα οχήματα. Το μειονέκτημα των μεθόδων που βασίζονται στην μηχανική μάθηση είναι η έλλειψη επαρκών training data που να καλύπτουν μεγάλο αριθμό διαφορετικών οχημάτων και καιρικών συνθηκών.

Στην δημοσίευση [33] γίνεται χρήση knowledge based μεθόδων στο σκέλος του Hypothesis Generation και appearance based για το Hypothesis Verification. Αναλυτικότερα αξιοποιείται το στοιχείο της συμμετρίας και του κόκκινου χρώματος των πίσω φαναριών για την δημιουργία των περιοχών ενδιαφέροντος (ROI) και η επαλήθευση γίνεται με template matching των 2 μισών της υποψήφιας ROI ως προς την μεσοκάθετο. Η έμφαση που δίνεται στα πίσω φώτα των οχημάτων καθιστά το σύστημα λειτουργικό για οδήγηση σε δύσκολες καιρικές συνθήκες καθώς και την νύχτα.

Τέλος οι Kim και Song στο [36] συνδυάζουν την μηχανική όραση, το radar και την οπτική ροή για να αποφανθούν αν ένα αντικείμενο είναι «όχημα» ή «μη όχημα».

4.3 Προτινόμενο σύστημα

Το σύστημα που αναπτύχθηκε στα πλαίσια της εργασίας βασίζεται κυρίως στην χρωματική πληροφορία και σε χαρακτηριστικά τα οποία είναι κοινά σε όλα τα οχήματα όπως το κόκκινο χρώμα των πίσω φαναριών, η συμμετρία που παρουσιάζουν καθώς και η σχετική απόσταση μεταξύ τους. Η είσοδος στο σύστημα λαμβάνεται από κάμερα η οποία είναι τοποθετημένη στο επάνω μέρος του παρμπριζ του αυτοκινήτου. Η εκ των προτέρων (a-priori) πληροφορία του οπτικού πεδίου της κάμερας είναι σημαντική ώστε να περιορίσουμε την περιοχή αναζήτησης υποψήφιων φαναριών στην ελάχιστη δυνατή. Για παράδειγμα στις πρώτες γραμμές ενός frame θα βρίσκεται κατά πάσα πιθανότητα ο ουρανός ενώ η αναζήτηση υποψηφίων αυτοκινήτων θα περιοριστεί στο κέντρο και χαμηλά του frame.

Για την καλύτερη προσέγγιση του προβλήματος η μελέτη χωρίζεται σε δύο στάδια καθένα από τα οποία αποτελείται από επιμέρους εργασίες :

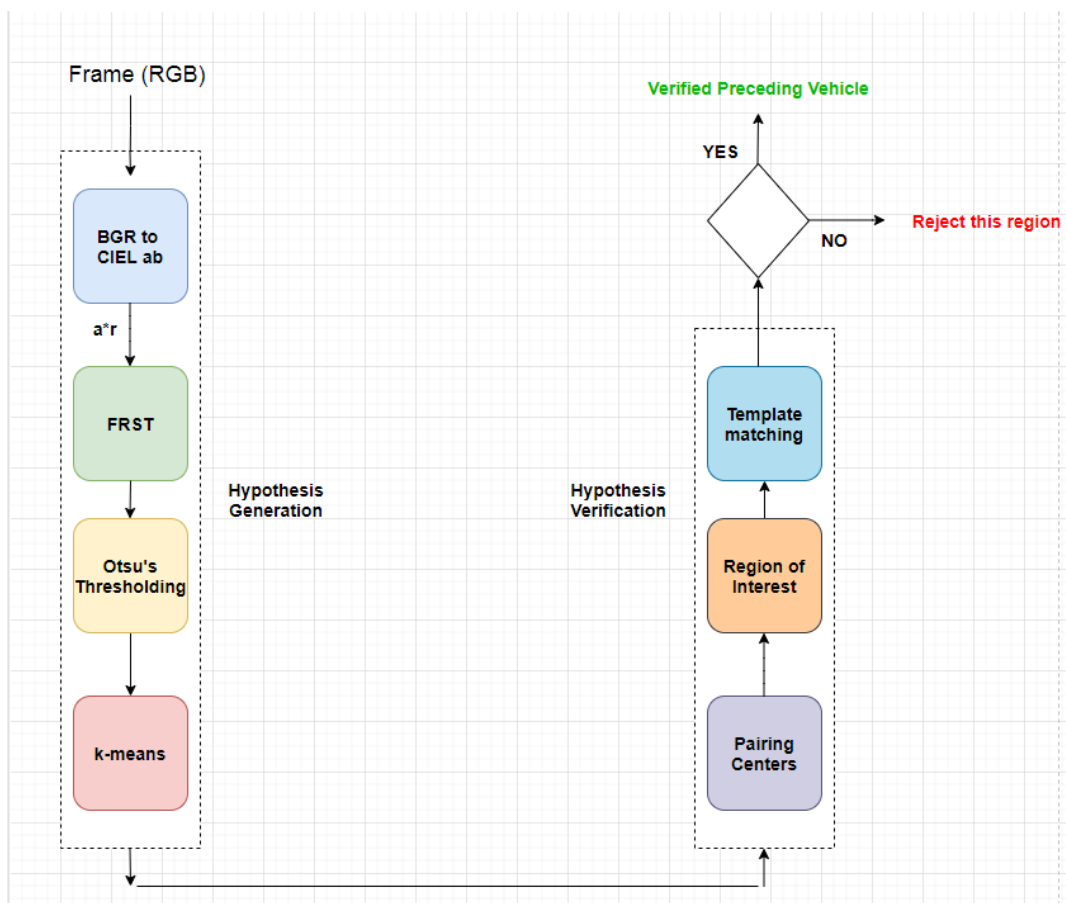
1) Εύρεση υποψηφίων (*Hypothesis Generation*)

Περιλαμβάνει την επεξεργασία της αρχικής εικόνας με σκοπό την εύρεση σημείων κόκκινου χρώματος και υψηλής συμμετρίας, χαρακτηριστικά που αυτομάτως τα καθιστούν υποψήφια πίσω φώτα οχημάτων.

2) Επαλήθευση των υποψηφίων (*Hypothesis Verification*)

Στο στάδιο αυτό ελέγχονται όλα τα υποψήφια σημεία που βρέθηκαν στο προηγούμενο σκέλος ώστε να εξακριβωθεί αν κάποια εξ' αυτών αποτελούν ζεύγη φαναριών. Στην συνέχεια σχηματίζεται μία περιοχή ενδιαφέροντος γύρω από αυτά και λαμβάνοντας υπ' όψη την υψηλή συμμετρία του πίσω μέρους ενός αυτοκινήτου η περιοχή αυτή σημειώνεται ως προπορευόμενο αυτοκίνητο, αλλιώς απορρίπτεται.

Για εύκολη και γρήγορη κατανόηση της φιλοσοφίας του συστήματος κατασκευάζουμε ένα διάγραμμα ροής των εργασιών που επιτελούνται σε κάθε στάδιο (Εικόνα 15).



Εικόνα 15. Διάγραμμα ροής των λειτουργιών της αναγνώρισης προπορευόμενου οχήματος.

4.3.1 Hypothesis Generation

Σε αυτό το στάδιο γίνεται η επεξεργασία της εικόνας ώστε να εξαχθούν τα σημεία που είναι υποψήφια πίσω φανάρια αυτοκινήτου. Αρχικά αναζητούμε τις περιοχές του frame που έχουν υψηλή συγκέντρωση κόκκινου χρώματος. Στο κεφάλαιο 3 παρουσιάστηκαν οι πιο δημοφιλείς χρωματικοί χώροι και τα χαρακτηριστικά του καθενός. Μεταξύ αυτών επιλέγουμε να εργαστούμε στον χώρο CIEL ab καθώς διαχωρίζει την φωτεινότητα από την χρωματική πληροφορία και διαθέτει κανάλι (a^* channel) που αφορά αποκλειστικά το πράσινο και κόκκινο χρώμα.

Γνωρίζουμε ότι τα πίσω φώτα κάθε αυτοκινήτου ανεξάρτητα από το σχήμα τους το οποίο ποικίλει ανάλογα το μοντέλο και τον κατασκευαστή, χαρακτηρίζονται από υψηλή συμμετρία. Αυτό είναι το δεύτερο χαρακτηριστικό που πρέπει να αξιοποιήσουμε ώστε να εξάγουμε τα σημεία της εικόνας που μας ενδιαφέρουν. Εφαρμόζουμε τον αλγόριθμο Fast Radial Symmetry Transform στο κανάλι a^* με σκοπό την εύρεση συμμετρικών αντικειμένων με εύρος ακτινών από 1 έως N . Η επιλογή της μέγιστης ακτίνας N διαφέρει ανάλογα με την θέση των προπορευόμενων αυτοκινήτων στο frame. Για παράδειγμα αν βρίσκεται κάποιο όχημα ακριβώς μπροστά του ego-vehicle τα φανάρια (σημεία συμμετρίας) θα έχουν μεγάλη ακτίνα ενώ αν βρίσκεται σε απόσταση 15 μέτρων είναι λογικό η ακτίνα τους να είναι μικρή (Εικόνα 16). Για να λάβουμε υπ όψη όλα τα πιθανά σενάρια ορίζουμε το N να είναι μεγάλο παρόλο που αυξάνει το χρόνο εκτέλεσης του αλγορίθμου.



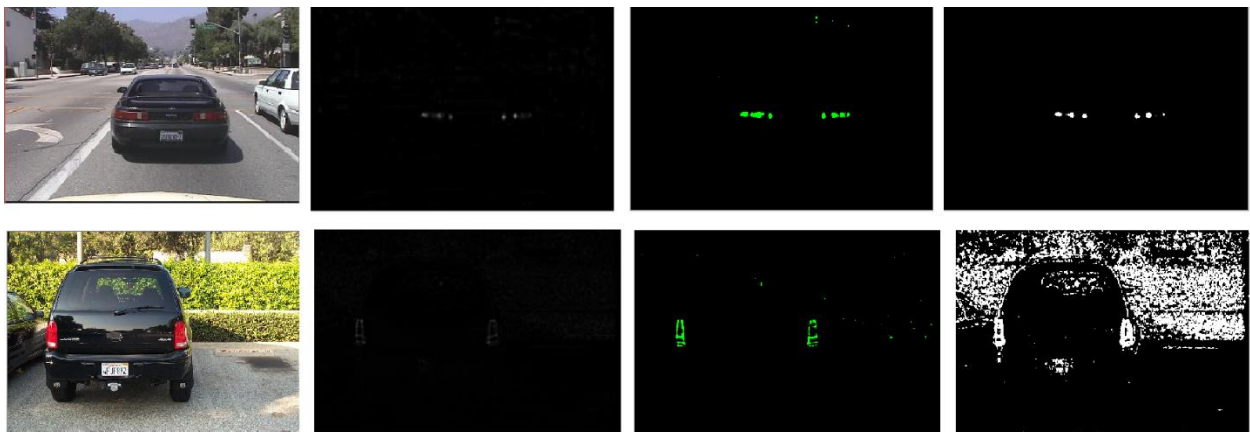
Εικόνα 16. Εξαιτίας της συνεχώς μεταβαλλόμενης απόστασης του προπορευόμενου οχήματος από το ego vehicle τα φανάρια ενός οχήματος εμφανίζονται με διαφορετική ακτίνα σε συνεχόμενα frames.

Ως αποτέλεσμα λαμβάνουμε μία νέα εικόνα που για κάθε pixel (i,j) έχει μία τιμή στο διάστημα $(0,1)$ και όσο πιο υψηλή είναι τόσο μεγαλύτερη συμμετρία υπάρχει στην περιοχή εκείνη.

Στο σημείο αυτό το πρόβλημα απλοποιείται σημαντικά αν μετατρέψουμε την εικόνα σε δυαδική. Η διαδικασία αυτή μπορεί να γίνει επιλέγοντας είτε μία καθολική τιμή κατωφλίου (global thresholding) είτε μέσω κάποια δυναμικής μέθοδου όπου η τιμή κατωφλίου μεταβάλλεται ανά περιοχή της εικόνας. Η πιο γνωστή μέθοδος μη-καθολικού thresholding είναι ο αλγόριθμος του Otsu [κεφ. 3.5] που χωρίζει τα pixels σε 2 κλάσεις βάσει μίας τιμής κατωφλίου, μέχρις ότου να βρεθεί εκείνη η τιμή που ελαχιστοποιεί τις διασπορές των

κλάσεων ή ισοδύναμα μεγιστοποιεί την συνδιασπορά (απο κοινού διασποράς) των κλάσεων. Η χρήση στατικού threshold συνδέεται άμεσα με τα χαρακτηριστικά της κάμερας που χρησιμοποιείται και επηρεάζεται πολύ από τις καιρικές συνθήκες και την φωτεινότητα του περιβάλλοντος. Αντίθετα η μέθοδος του Otsu είναι unsupervised και κάθε φορά αναζητά την βέλτιστη τιμή κατωφλίου δοθέντων των εκάστοτε συνθηκών με αποτέλεσμα να λειτουργεί ικανοποιητικά ανεξάρτητα της φωτεινότητας και των τεχνικών χαρακτηριστικών της κάμερας.

Ελέγχοντας την αποτελεσματικότητα του συστήματος μέχρι αυτό το σημείο, προέκυψε κάτι παράδοξο. Ενώ θεωρητικά η επιλογή της μεθόδου του Otsu για την διαδικασία του image binarization μοιάζει ιδανική για αστικό περιβάλλον, όταν στο υπόβραθρο υπάρχει βλάστηση δεν ανταποκρίνεται ικανοποιητικά (Εικόνα 17). Η συμπεριφορά αυτή ωφείλεται στο κανάλι a^* το οποίο ποσοτικοποιεί την παρουσία κόκκινου και πράσινου χρώματος με αποτέλεσμα όταν το δίνουμε ως είσοδο στον FRST ανιχνεύονται τα συμμετρικά αντικείμενα και των δύο χρωμάτων.



Εικόνα 17. Σύγκριση της μεθόδου του Otsu και του στατικού threshold σε 2 frames. Το 1ο frame είναι από αστικό περιβάλλον ενώ το 2ο έχει έντονη βλάστηση. Σε κάθε γραμμή διακρίνονται η αρχική RGB εικόνα (1η στήλη), το κανάλι a^* (2η στήλη), FRST και global thresholding στο κανάλι a^* (3η στήλη) και FRST και Otsu thresholding στο κανάλι a^* (4η στήλη).

Στις παραπάνω εικόνες παρατηρούμε ότι στο αστικό περιβάλλον οι διαφορές των δύο μεθόδων είναι αμελητέες με την μέθοδο του Otsu να είναι ελαφρώς καλύτερη, καθώς απορρίπτει τον θόρυβο από συμμετρικά αντικείμενα του υποβάθρου (Εικόνα 17, 1^η γραμμή εικόνες 3,4). Σε περιβάλλον με βλάστηση η επιλογή καθολικού threshold λειτουργεί ικανοποιητικά σε αντίθεση με την μέθοδο του Otsu που όχι απλά να δεν αποκόπτει τον θόρυβο αλλά να τον ενισχύει (Εικόνα 17, 2^η γραμμή εικόνες 3,4).

Σκοπός ενός συστήματος υποβοήθησης της οδήγησης είναι να λειτουργεί αξιόπιστα υπό πολλές συνθήκες και σε διαφορετικά περιβάλλοντα. Επίσης σημαντικό χαρακτηριστικό του συστήματος είναι η φορητότητα δηλαδή το λογισμικό που αναπτύσσουμε να μπορεί να τοποθετηθεί σε οποιοδήποτε hardware ανεξάρτητα των τεχνικών προδιαγραφών του. Για τον λόγο αυτό η εφαρμογή καθολικού-στατικού κατωφλίου μοιάζει απαγορευτική καθώς θα έπρεπε κάθε φορά να ρυθμίζουμε την τιμή του threshold ανάλογα με τις καιρικές συνθήκες

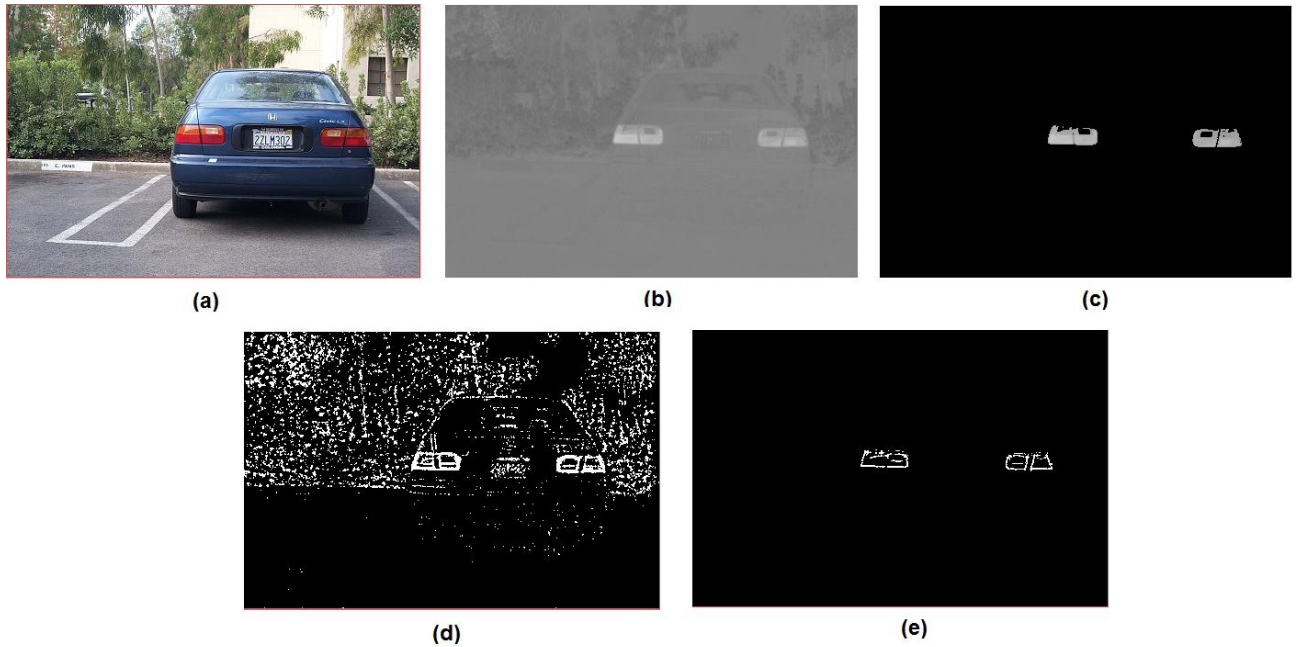
και την κάμερα που χρησιμοποιείται στο εκάστοτε αυτοκίνητο. Από την άλλη πλευρά για να εφαρμόσουμε την τεχνική του Otsu θα πρέπει να ξεπεράσουμε το πρόβλημα της ανίχνευσης συμμετρικών αντικειμένων και των δύο χρωμάτων που αναφέραμε προηγουμένως.

Το κανάλι a^* του CIEL ab κωδικοποιεί την ισορροπία μεταξύ πράσινου και κόκκινου χρώματος. Σύμφωνα με την θεωρία [κεφ 3.1.2] οι τιμές του καναλιού a^* κυμαίνονται μεταξύ -127 (καθαρό πράσινο) και +127 (καθαρό κόκκινο). Στην πράξη ωστόσο κατά την μετατροπή της εικόνας στην OpenCV από RGB σε CIEL ab το κανάλι a^* έχει τιμές στο διάστημα [80-210] με μικρές αποκλίσεις. Για τον διαχωρισμό των περιοχών κόκκινου χρώματος χωρίζουμε το διάστημα στην μέση και λαμβάνουμε το άνω τμήμα το οποίο θα αποκαλούμε «κόκκινο υποκανάλι» του a^* και θα το συμβολίσουμε με a^*_r (Εικόνα 18).



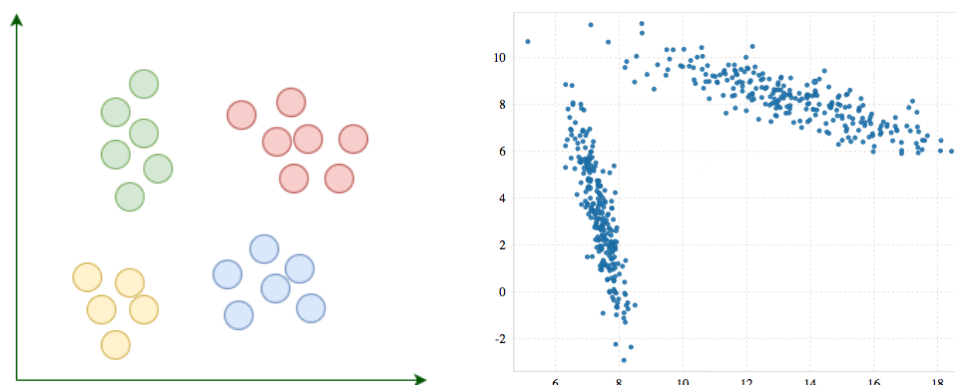
Εικόνα 18. Η αρχική εικόνα RGB (1η στήλη), το κανάλι a^* (2η στήλη) και το κόκκινο υποκανάλι a^*_r (3η στήλη).

Για να απαλλαγούμε από τα αντικείμενα πράσινου χρώματος -που στην μελέτη μας εισάγουν θόρυβο στο σύστημα- θα τροφοδοτήσουμε στον FRST το κόκκινο υποκανάλι a^*_r και κατόπιν με χρήση της τεχνικής του Otsu θα μετατρέψουμε το αποτέλεσμα σε δυαδική εικόνα. Όπως φαίνεται και στην Εικόνα 19, απορρίπτουμε τον θόρυβο που προκαλείται από τις φυλλώδεις περιοχές του frame και εστιάζουμε στα πίσω φανάρια του προπορευόμενου οχήματος.



Εικόνα 19. Σύγκριση της αποτελεσματικότητας της εφαρμογής FRST και Otsu thresholding. (a) Αρχικό frame, (b) channel a^* , (c) channel a^*r , (d) FRST & Otsu στο channel a^* , (e) FRST & Otsu στο «κόκκινο» κανάλι a^*r .

Εφόσον έχουμε κάνει το segmentation της εικόνας, επόμενος στόχος είναι να προσδιορίσουμε τα κέντρα των αντικειμένων υψηλής συμμετρίας ώστε να δοθούν στο στάδιο της επαλήθευσης υποψηφίων και να βρεθούν τα ζεύγη φαναριών. Στην binary εικόνα που δημιουργήσαμε προηγουμένως (Εικόνα 19(e)) παρατηρούμε ότι δημιουργούνται νέφη - ομάδες σημείων. Αν βρούμε το κέντρο της κάθε ομάδας θα προσδιοριστεί αυτομάτως και το κέντρο συμμετρίας του κάθε αντικειμένου. Ο πλέον δημοφιλής αλγόριθμος ομαδοποίησης είναι ο k-means σύμφωνα με τον οποίο N παρατηρήσεις ομαδοποιούνται σε k κλάσεις κατά τέτοιο τρόπο ώστε να ελαχιστοποιείται το άθροισμα των αποστάσεων των σημείων από το αντίστοιχο κέντρο της κλάσης στην οποία ανήκουν. Ο αλγόριθμος αυτός είναι supervised καθώς απαιτείται να καθοριστεί από τον χρήστη το πλήθος (k) των ομάδων-clusters, που διαφέρει ανάλογα με το πώς είναι διατεταγμένα τα σημεία στο χώρο (Εικόνα 20).



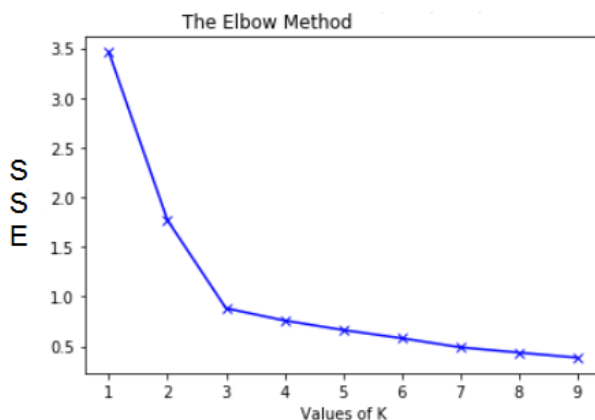
Εικόνα 20. Στο αριστερό γράφημα η καλύτερη δυνατή ομαδοποίηση είναι σε 4 clusters ενώ στο δεξί γράφημα σε 2 clusters.

Προφανώς αυτό πρέπει να ληφθεί υπ' όψη στο σύστημα που αναπτύσσεται πόσο μάλλον αν σκεφτούμε ότι η διαδικασία τη οδήγησης είναι πλήρως δυναμική και τα πιθανά σενάρια που έχουμε να αντιμετωπίσουμε είναι πάρα πολλά. Για παράδειγμα στο οπτικό πεδίο της κάμερας ενδέχεται να υπάρχουν ένα, δύο ή περισσότερα οχήματα σε τυχαίες θέσεις, ενώ συχνό φαινόμενο είναι κάποιο από τα πίσω φανάρια των προπορευόμενων οχημάτων να επικαλύπτεται από κάποιο εμπόδιο (Εικόνα 21).



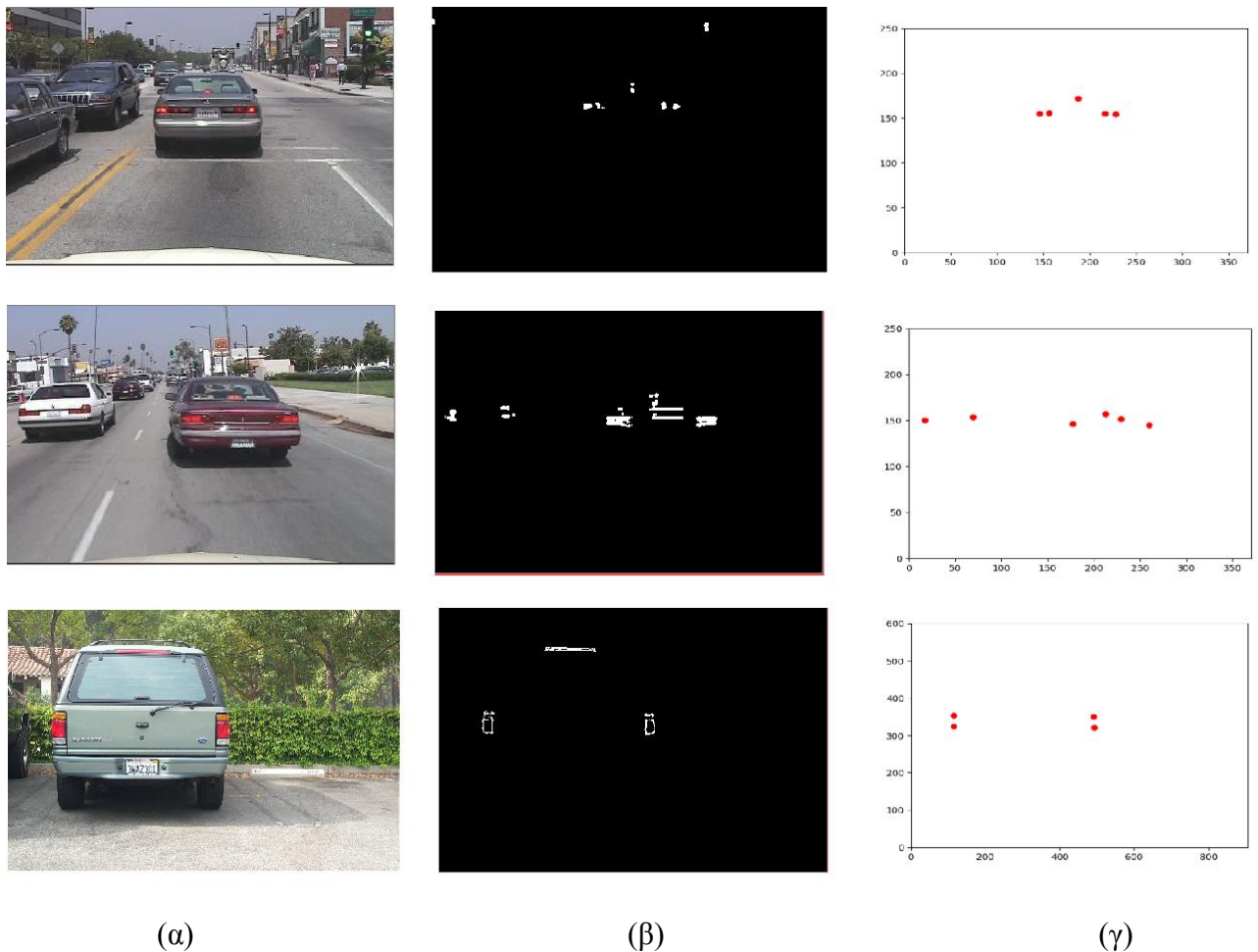
Εικόνα 21. Τρία σενάρια διαφορετικής τιμής του βέλτιστου k . Στο frame (a) ιδανικά θα χρειαστούμε 2 κλάσεις για την ομαδοποίηση, στο (b) 4 κλάσεις ενώ στο (c) 3 κλάσεις.

Οι συνθήκες αυτές οδηγούν στην αδυναμία επιλογής ενός στατικού k εκ των προτέρων και το πρόβλημα μετατρέπεται σε πλήρως δυναμικό. Για τον προσδιορισμό του βέλτιστου πλήθους των clusters εφαρμόζουμε την μέθοδο γονάτου [κεφ. 3.3.1] η οποία εκτελεί τον k -means επαναληπτικά και μας επιστρέφει το k για το οποίο ο διαχωρισμός των παρατηρήσεων είναι ο βέλτιστος δυνατός. Όπως προαναφέρθηκε και στην αναλυτική περιγραφή της μεθόδου αναζητούμε το σημείο γονάτου στο οποίο αλλάζει η κλίση της συνάρτησης από απότομη σε σχεδόν μηδενική (Εικόνα 22). Στο Παράρτημα 7.1 επισυνάπτεται η συνάρτηση που δημιουργήσαμε και υλοποιεί την μέθοδο.



Εικόνα 22. Το elbow point είναι στο $K=3$ δηλαδή η βέλτιστη επιλογή για την ομαδοποίηση των δειγμάτων είναι 3 clusters.

Πρακτικά στην συνάρτηση που κατασκευάστηκε εκτελείται επαναληπτικά ο k -means και υπολογίζεται το άθροισμα τετραγώνων των αποστάσεων (SSE) για K από 1 έως N . Σε κάθε επανάληψη ελέγχεται αν τα δύο διαδοχικά σημεία, δηλαδή οι τιμές SSE που υπολογίστηκαν διαφέρουν πάνω από 30%. Αν ναι τότε σημαίνει ότι η κλίση παραμένει απότομη και το τρέχον k καταχωρείται ως το βέλτιστο. Την πρώτη φορά που δεν θα ικανοποιηθεί η συνθήκη σημαίνει ότι αυτό είναι το σημείο γονάτου άρα το ζητούμενο βρέθηκε και σταματάω την αναζήτηση. Τέλος δίνουμε ως είσοδο στον k -means τον αριθμό των clusters και την δυαδική segmented εικόνα και λαμβάνουμε ως αποτέλεσμα τα κέντρα των σχημάτων τα οποία αυτομάτως θεωρούμε ως υποψήφια πίσω φώτα αυτοκινήτου (Εικόνα 23).



Εικόνα 23. α) RGB εικόνα, β) σημεία υψηλής συμμετρίας όπως προέκυψαν από τον FRST, γ) Τα κέντρα των σημείων όπως προέκυψαν από τον αλγόριθμο ομαδοποίησης k -means.

4.3.2 Hypothesis Verification

Στο προηγούμενο σκέλος εντοπίστηκαν τα κέντρα των περιοχών κόκκινου χρώματος που παρουσιάζουν υψηλή συμμετρία. Τα δύο αυτά χαρακτηριστικά ωστόσο δεν επαρκούν για να επαληθευτεί η παρουσία ενός οχήματος στο frame καθώς στα οδικά δίκτυα και συχνότερα σε πυκνό αστικό περιβάλλον, εμφανίζονται αντικείμενα (φωτεινοί σηματοδότες, σήματα οδικής κυκλοφορίας ή και διαφημιστικές πινακίδες) που πληρούν τα κριτήρια της συμμετρίας και του κόκκινου χρώματος αλλά συνιστούν false alarm για το σύστημα.

Στο σκέλος αυτό, πρέπει σε πρώτο στάδιο (Lights Pairing) να απομονώσουμε τα φώτα των αυτοκινήτων από τα υπόλοιπα αντικείμενα που αποτελούν θόρυβο, εκμεταλλευόμενοι την ιδιότητα ότι εμφανίζονται ανά ζεύγη. Σε δεύτερο στάδιο γίνεται η επαλήθευση δημιουργώντας μία περιοχή ενδιαφέροντος γύρω από τα ζεύγη και λαμβάνοντας υπ' όψη την συμμετρία του πίσω μέρους του οχήματος ως προς την μεσοκάθετο.

4.3.2.a Lights Pairing

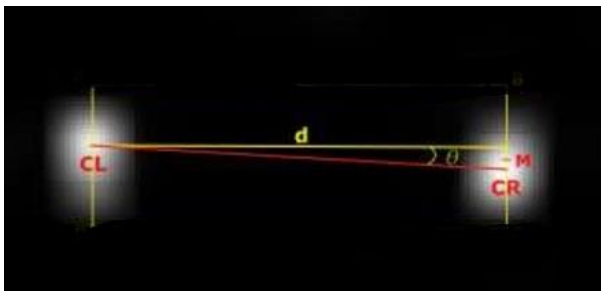
Σε πραγματικές συνθήκες οδήγησης μπορεί να εμφανιστούν αρκετά αντικείμενα κόκκινου χρώματος που ταυτόχρονα παρουσιάζουν υψηλή συμμετρία όπως φωτεινοί σηματοδότες, διαφημιστικές πινακίδες κ.α. Ωστόσο τα πίσω φώτα αυτοκινήτου έχουν κάποια χαρακτηριστικά που μας επιτρέπουν να τα διακρίνουμε από τα υπόλοιπα. Πρώτον υπάρχουν ανά δύο σε ένα όχημα, δεύτερον βρίσκονται στο ίδιο ύψος και τρίτον η οριζόντια απόσταση τους κυμαίνεται εντός κάποιων ορίων ανάλογα με το μέγεθος του αυτοκινήτου και την απόστασή του από το ego-vehicle. Στην συνέχεια εξετάζουμε μερικά σενάρια όπου η σχετική θέση των πίσω φαναριών ενός αυτοκινήτου διαφέρει. Στην Εικόνα 24 (α) τα φώτα είναι στο ίδιο ύψος και η απόστασή τους είναι συγκρίσιμη με το πλάτος (width) του frame διότι το προπορευόμενο όχημα είναι πολύ κοντά στο ego-vehicle. Αντίθετα στην εικόνα 24 (β) το όχημα απέχει αρκετά από το ego vehicle και κατ' επέκταση τα φώτα έχουν μικρή απόσταση ως προς πλάτος του frame. Τέλος στην εικόνα 24 (γ) το προπορευόμενο αυτοκίνητο είναι στραμμένο ελαφρώς προς τα δεξιά και γι αυτό τα φώτα δεν είναι στην ίδια ευθεία αλλά αποκλίνουν από τον οριζόντιο άξονα κατα μικρή γωνία θ .



Εικόνα 24. Τρία σενάρια διαφορετικής θέσης του προπορευόμενου οχήματος ως προς το ego vehicle.

Λαμβάνοντας υπόψη όλα τα πιθανά σενάρια καταλήγουμε στο ότι δεν μπορεί να εφαρμοστεί ένα καθολικό κριτήριο για τον χαρακτηρισμό δύο σημείων ως ζεύγος αλλά θα πρέπει η συνθήκη ελέγχου να εξαρτάται από τις διαστάσεις του frame. Όσον αφορά την κάθετη απόσταση των φαναριών αυτή πρέπει να είναι αμελητέα θεωρώντας αποδεκτή μία απόκλιση της τάξεως των 5° από τον οριζόντιο άξονα. Για την οριζόντια απόσταση, αυτή θα εξαρτάται από το πλάτος (w) του frame και κατόπιν πολλών δοκιμών καταλήξαμε στο συμπέρασμα ότι θέλουμε να κυμαίνεται μεταξύ $\frac{1}{8}$ και $\frac{6}{10}$ του πλάτους (Εικόνα 25). Το εύρος αυτό θετεί αυτομάτως περιορισμούς στην ικανοποιητική λειτουργία του συστήματος για οχήματα πολύ μακριά ή πολύ κοντά στο ego-vehicle. Για παράδειγμα αν ένα αυτοκίνητο βρίσκεται στο βάθος της εικόνας τότε τα πίσω φώτα του θα απέχουν λιγότερο από $\frac{1}{8}$ με αποτέλεσμα να μην αναγνωριστεί επιτυχώς. Το ίδιο θα συμβεί και για ένα όχημα που βρίσκεται σε πολύ κοντινή απόσταση και τα πίσω φώτα του απέχουν περισσότερο από $\frac{6}{10}$ του πλάτους του frame. Αυτές οι δύο περιπτώσεις ωστόσο δεν θα πρέπει να μας προβληματίζουν καθώς το σύστημα που προτίνεται στα πλαίσια της εργασίας θα μπορούσε με το κατάλληλο hardware και κάποιες τροποποιήσεις να δέχεται video και να κάνει την αναγνώριση σε πραγματικό χρόνο.

Στην περίπτωση αυτή τα αυτοκίνητα που δεν έχουν ανιχνευθεί στο frame t είναι πολύ πιθανό να είχαν ήδη αναγνωριστεί επιτυχώς σε προηγούμενα frames ή να αναγνωριστούν σε επόμενα. Στο Παράρτημα 7.2 παρουσιάζουμε την συνάρτηση που κατασκευάσαμε για την δημιουργία του πίνακα των ζευγών υποψηφίων φαναριών.



Εικόνα 25. Η οριζόντια απόσταση d δύο κέντρων πρέπει να είναι μεταξύ $1/8$ και $6/10$ του πλάτους του frame ενώ η γωνία θ από τον οριζόντιο άξονα να μην υπερβαίνει τις 5° .

4.3.2.b Δημιουργία περιοχής ενδιαφέροντος

Για το Hypothesis Verification δεν γίνεται χρήση κάποιου ταξινομητή αλλά επιλέγουμε template based προσέγγιση. Συγκεκριμένα δημιουργούμε μία περιοχή ενδιαφέροντος γύρω από κάθε ζεύγος σημείων που προσδιορίστηκαν στο προηγούμενο βήμα. Στη συνέχεια εκμεταλλευόμαστε την συμμετρία που παρουσιάζει η πίσω όψη των αυτοκινήτων ως προς τον κατακόρυφο άξονα και συγκρίνουμε τις δύο υποπεριοχές χρησιμοποιώντας κάποιον

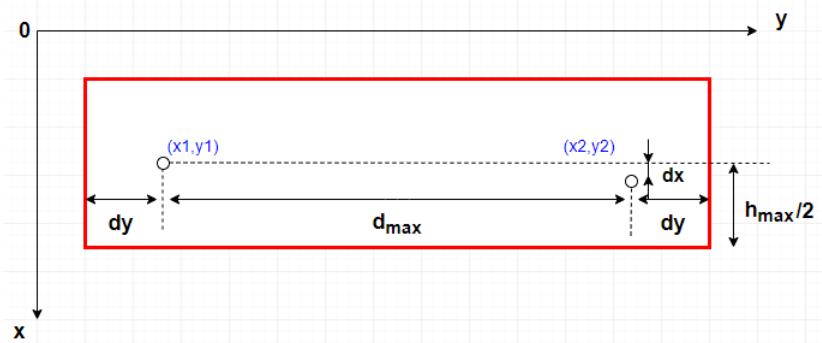
δείκτη ομοιότητας. Επιλέγοντας ένα κατώφλι θα είμαστε σε θέση να επιβεβαιώσουμε αν πράγματι η περιοχή ενδιαφέροντος περιέχει ένα προπορευόμενο όχημα ή σε αντίθετη περίπτωση να την απορρίψουμε.

Για την δημιουργία της περιοχής ενδιαφέροντος εξετάζουμε κάθε ζεύγος σημείων-κεντρών όπως ορίστηκαν παραπάνω λαμβάνοντας ως αναφορά εκείνο με το μικρότερο x . Ορίζουμε τις αποστάσεις :

- $d_{max} = |y_1 - y_2|$
- $h_{max} = d_{max} / 3$
- $dy = d_{max} / 10$
- $dx = |x_1 - x_2|$

όπου στον οριζόντιο άξονα είναι το y , στον κάθετο το x και το θ στην πάνω αριστερή γωνία σύμφωνα με την σύμβαση της OpenCV.

Το πλάτος της περιοχής ενδιαφέροντος ορίζεται ως η οριζόντια απόσταση των σημείων προσαυξανόμενη κατά dy . Για τις κάθετες πλευρές εξετάζουμε 4 περιπτώσεις ανάλογα με την σχετική θέση των σημείων και στον υπολογισμό εμπλέκονται οι ποσότητες x_1 , x_2 , $h_{max}/2$ και dx (Εικόνα 26).



Εικόνα 26. Δημιουργία περιοχής ενδιαφέροντος μεταξύ ενός ζεύγους κεντρών.

Παρακάτω στην Εικόνα 27, παρουσιάζονται μερικά αποτελέσματα της προτινόμενης τεχνικής για την δημιουργία της περιοχής ενδιαφέροντος και η οποία υλοποιείται με την συνάρτηση του Παραρτήματος 7.4. Διακρίνουμε τα εξής σενάρια οδήγησης:

- 1) Προπορευόμενο όχημα δίπλα σε κίτρινη διαχωριστική λωρίδα.
- 2) Δύο προπορευόμενα οχήματα το ένα εκ των οποίων δεν εμπεριέχεται όλο στο frame.
- 3) Αυτοκίνητο με ξεθωριασμένα πίσω φανάρια.



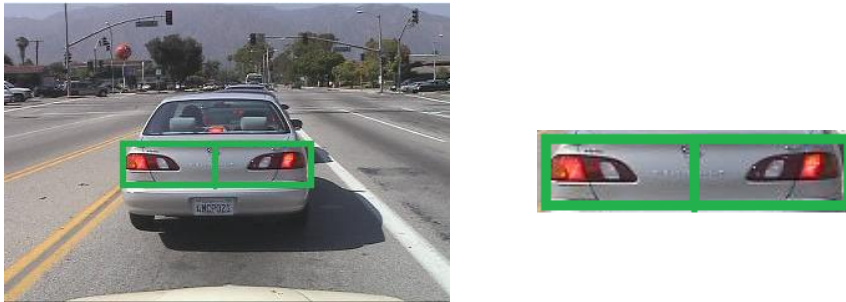
Εικόνα 27. Πρώτη στήλη η RGB εικόνα με τις περιοχές ενδιαφέροντος σχηματισμένες, δεύτερη στήλη το κόκκινο υποκανάλι $a*r$ και τρίτη στήλη η δυαδική εικόνα κατά Otsu.

Όπως προαναφέραμε είναι λογικό εκτός των αποδεκτών περιοχών γύρω από τα πίσω φώτα του αυτοκινήτου να σχηματίζονται και άλλες -false negative, εξαιτίας αντικειμένων της εικόνας που πληρούν τα κριτήρια του χρώματος, της συμμετρίας και των αποστάσεων και χαρακτηρίζονται ως ζεύγη. Για παράδειγμα στην πρώτη εικόνα η κίτρινη διαχωριστική γραμμή είναι η αιτία εμφάνισης λανθασμένης περιοχής ενδιαφέροντος ενώ στις άλλες δύο το σύστημα παραπλανάται από ένα όχημα που δεν βρίσκεται όλο εντός του frame. Είναι φανερό λοιπόν ότι πρέπει να αυξήσουμε την αξιοπιστία του συστήματος προσθέτοντας ένα επιπλέον στάδιο ελέγχου με σκοπό να απορρίψουμε τις false negative περιοχές και να διατηρήσουμε μόνο εκείνες που μας ενδιαφέρουν.

4.3.2.c Template Matching

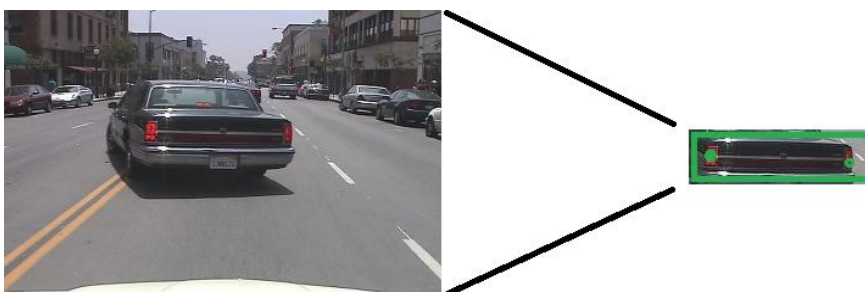
Στο στάδιο αυτό εξετάζεται κάθε περιοχή ενδιαφέροντος ώστε να παραμείνουν μόνο εκείνες οι οποίες πράγματι περιέχουν κάποιο όχημα. Αυτό θα επιτευχθεί λαμβάνοντας υπόψη το κοινό χαρακτηριστικό που έχει το πίσω μέρος όλων των αυτοκινήτων και αυτό είναι η συμμετρία ως προς τον κατακόρυφο άξονα (Εικόνα 28).

Χωρίζουμε κάθε περιοχή ενδιαφέροντος σε δύο υποεικόνες δεξιά και αριστερή ίδιων διαστάσεων και ελέγχουμε τον βαθμό ομοιότητας τους. Οι περιοχές για τις οποίες το δεξί και αριστερό μέρος παρουσιάζουν μεγάλη ομοιότητα θεωρούνται αποδεκτές ενώ οι υπόλοιπες απορρίπτονται .



Εικόνα 28. Συμμετρία της πίσω όψης ενός αυτοκινήτου ως προς τον κατακόρυφο άξονα, από frame της βάσης δεδομένων Caltech 1999.

Στο κεφάλαιο 3.8 παρουσιάστηκαν οι δύο επικρατέστεροι δείκτες ομοιότητας καθώς και οι διαφορές μεταξύ τους. Στην παρούσα εργασία επιλέγουμε τον δείκτη SSIM για τρεις βασικούς λόγους. Πρώτον οι τιμές του είναι φραγμένες στο $[-1,1]$ κάνοντας ευκολότερη την επιλογή ενός κατωφλίου. Δεύτερον ο SSIM δεν επηρεάζεται από την γωνία των αντικειμένων καθώς εξετάζει μία εικόνα ανά περιοχές και αντλεί πληροφορίες σχετικά με την δομή της. Αυτό είναι ιδιαίτερα χρήσιμο καθώς επιτρέπει την πετυχημένη ανίχνευση προπορευόμενων οχημάτων που στρίβουν ή βρίσκονται στις άκρες του frame (Εικόνα 29). Τρίτον η χρήση του MAE είναι απαγορευτική καθώς παρουσιάζει ασταθή συμπεριφορά λόγω τ' ότι εφαρμόζεται καθολικά σε ολόκληρη την εικόνα με αποτέλεσμα μεγάλες αποστάσεις στις τιμές των pixels να μην ανταποκρίνονται σε πραγματικά μεγάλες διαφορές στο περιεχόμενο των εικόνων. Με χρήση της συνάρτησης του Παραρτήματος 7.3, υπολογίζουμε και επιστρέφουμε τον δείκτη SSIM για 2 υποεικόνες.



Εικόνα 29. Περίπτωση πιθανής αδυναμίας του MAE λόγω εξέτασης αποκλειστικά των τιμών των αντίστοιχων pixels.

4.4 Πειραματικά αποτελέσματα

Η απόδοση του συστήματος αξιολογείται με χρήση των βάσεων δεδομένων Caltech 1999 και Caltech 2001 κάτω από ορισμένες προϋποθέσεις. Πρώτον ως προπορευόμενα όχηματα θεωρούνται εκείνα που βρίσκονται στην ίδια λωρίδα ή σε ακριβώς διπλανή και απέχουν το πολύ 10-15 μέτρα από το ego-vehicle. Δεύτερον λαμβάνοντας υπ' όψη τον σχετικό κανονισμό πρέπει όλα τα αυτοκίνητα να έχουν ευδιάκριτα πίσω φώτα, κόκκινου χρώματος και πατώντας το φρένο (stop) να φωτίζονται αμφότερα. Λόγω αυτών των απαιτήσεων έχουν αποκλειστεί 5 αυτοκίνητα από την βάση Caltech 2001 που εμφανίζονται σε 15 frames (Εικόνα 30) .



Εικόνα 30. Αποκλεισμένα αυτοκίνητα λόγω διαδιάκριτων φαναριών ή καμμένης λυχνίας του stop.

Οι επιδόσεις του συστήματος αναγνώρισης προπορευόμενου οχήματος που αναπτύχθηκε, για τις δύο -ενημερωμένες- βάσεις δεδομένων, φαίνονται στον Πίνακα 4. Παρατηρείται ότι στην Caltech 1999 τα ποσοστά είναι αρκετά υψηλά με πολύ καλή αναλογία μεταξύ πληρότητας (Detection) και ακρίβειας (Precision). Όσον αφορά την Caltech 2001 η ακρίβεια διατηρείται σε ικανοποιητικά επίπεδα ενώ η πληρότητα είναι στο 67%, ποσοστό που δεν θεωρείται χαμηλό ωστόσο σίγουρα υπάρχουν περιθώρια βελτίωσης.

	Detection Rate	Precision Rate
Caltech 1999	85.8%	83%
Caltech 2001	67%	78%

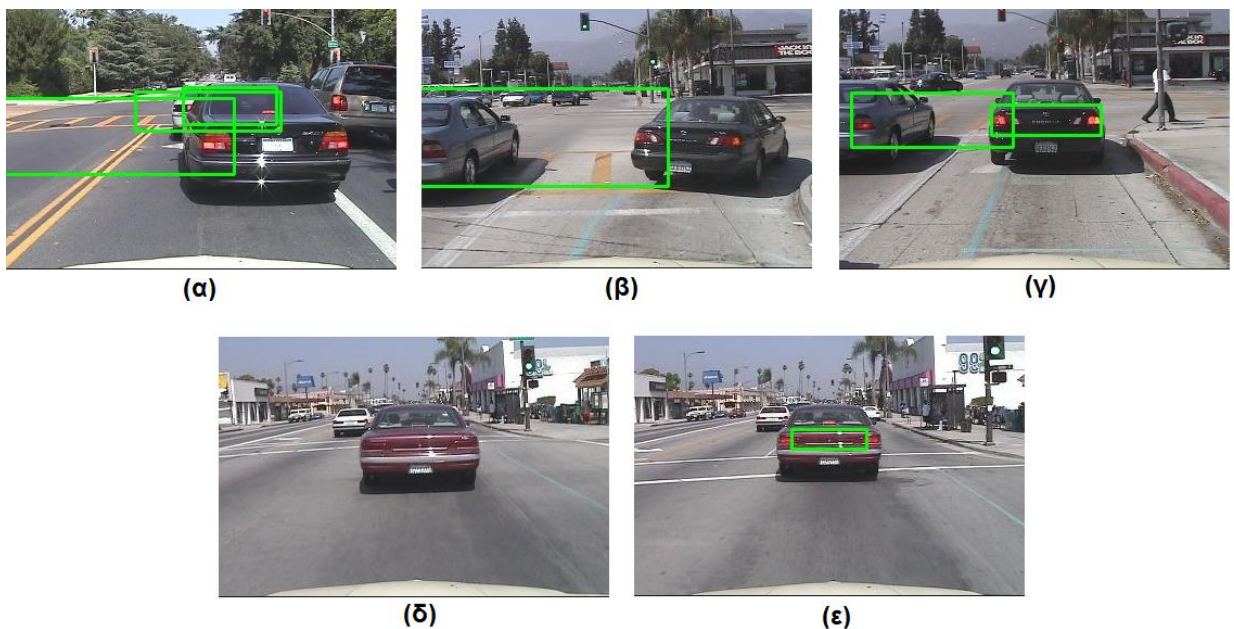
Πίνακας 4. Αξιολόγηση του συστήματος στις βάσεις Caltech 1999,2001

Στην συνέχεια παρουσιάζονται κάποια αποτελέσματα για frames των βάσεων Caltech 2001 (Εικόνα 31 & 32) και Caltech 1999 (Εικόνα 33) σε διάφορες συνθήκες οδήγησης και με παρουσία πολλών διαφορετικών οχημάτων. Παρατηρούμε ότι το προτινόμενο σύστημα ανταποκρίνεται πολύ καλά σε αρκετές περιπτώσεις με υψηλή ηλιοφάνεια (Εικόνα 31 (α), (ε), (θ)) καθώς και σε σκιερούς δρόμους με χαμηλή φωτεινότητα (Εικόνα 31 (β)). Επίσης περιβάλλοντα με διαφημιστικές πινακίδες, φωτεινούς σηματοδότες, προσόψεις καταστημάτων κ.α εισάγουν θόρυβο στο σύστημα καθώς αποτελούν σημεία υψηλής συγκέντρωσης κόκκινου χρώματος. Παρόλ' αυτά το σύστημα ανταποκρίνεται επιτυχώς, όπως φαίνεται στα (ζ), (η) και (θ) της Εικόνας 31. Μία μικρή αδυναμία φαίνεται να υπάρχει όταν το προπορευόμενο όχημα είναι κόκκινου χρώματος. Σε αυτή την περίπτωση σχηματίζεται μία μεγάλη περιοχή σημείων ενδιαφέροντος που περιλαμβάνει όλη την πίσω όψη του οχήματος με αποτέλεσμα ο k-means να τοποθετεί τα κέντρα σε θέσεις που δεν συμπίπτουν με τα πραγματικά κέντρα των φαναριών του οχήματος. Για τον λόγο αυτό η περιοχή ενδιαφέροντος δεν σχηματίζεται με μεγάλη ακρίβεια (Εικόνα 31 (στ)). Τέλος αξίζει να αναφερθεί η επιτυχής αναγνώριση όχι μόνο του μπροστινού οχήματος αλλά και εκείνων που βρίσκονται σε διπλανή λωρίδα εφόσον φυσικά εμπεριέχονται ολόκληρα μέσα στο frame (Εικόνα 31 (ζ), (η)).



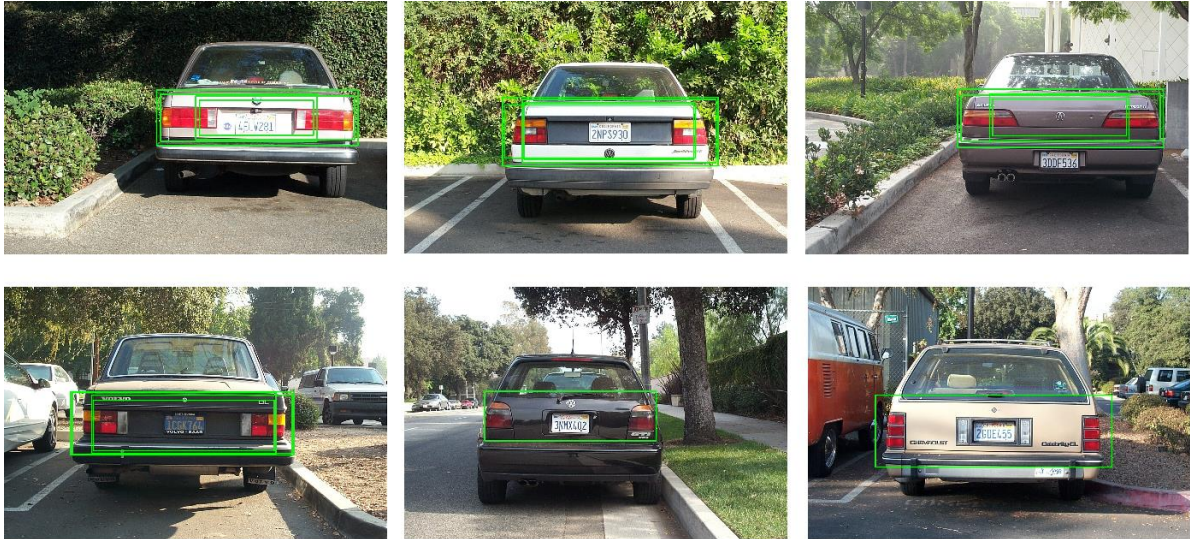
Εικόνα 31. Παραδείγματα εφαρμογής του αλγορίθμου αναγνώρισης προπορευόμενου οχήματος σε πυκνό αστικό περιβάλλον υπό διάφορες συνθήκες (Caltech 2001).

Βασικός λόγος για τον οποίον ενδέχεται το σύστημα να οδηγηθεί σε λανθασμένο ή ανακριβές αποτέλεσμα είναι η παρουσία τμήματος ενός αυτοκινήτου δίπλα στο προπορευόμενο με αποτέλεσμα να δημιουργηθούν περιοχές ενδιαφέροντος με κέντρα τα φανάρια διαφορετικών οχημάτων (Εικόνα 32 (β), (γ)). Μία άλλη συχνή αιτία false positive είναι τα αντικείμενα κόκκινου χρώματος όπως διαφημιστικές πινακίδες ή διαγραμμίσεις του δρόμου (Εικόνα 32 (α) – κάτω αριστερά). Παρατηρείται επίσης πως αυξάνεται η πιθανότητα σωστής και ακριβούς ανίχνευσης του μπροστινού οχήματος όταν είναι πατημένα τα stop. Για παράδειγμα στην Εικόνα 32 (δ) αρχικά με σβηστό το stop έχουμε false negative ενώ στο επόμενο frame της Εικόνας 32 (ε) ο οδηγός πατάει φρένο, ανάβουν τα στοπ και η ανίχνευση του οχήματος είναι πετυχημένη. Αυτή η συμπεριφορά είναι φυσιολογική εφόσον ο αλγόριθμος που αναπτύχθηκε βασίζεται στην χρωματική πληροφορία της εικόνας. Σε επόμενο στάδιο ανάπτυξης του υπάρχοντος συστήματος, η ξαφνική εμφάνιση μίας περιοχής ενδιαφέροντος για το ίδιο όχημα και μεταξύ διαδοχικών frames (Εικόνα 32 (δ), (ε)), θα μπορούσε να αξιοποιηθεί ώστε να ειδοποιεί τον οδηγό ότι το προπορευόμενο όχημα επιβραδύνει. Η λειτουργία αυτή είναι πολύ σημαντική σε ένα σύστημα υποβοήθησης οδήγησης καθώς κρατά τον οδηγό σε επαγρύπνηση οσον αφορά την προβλεπόμενη τήρηση των αποστάσεων και μειώνει την πιθανότητα σύγκρουσης σε περίπτωση απότομου φραναρίσματος.



Εικόνα 32. Παραδείγματα αστοχιών του συστήματος σε πυκνό αστικό περιβάλλον υπό διάφορες συνθήκες (Caltech 2001).

Στην συνέχεια παρουσιάζονται ορισμένα αποτελέσματα για σταθμευμένα αυτοκίνητα από την βάση Caltech 1999 (Εικόνα 33). Το υπόβαθρο σε πολλά frames είναι πολύπλοκο όπως δέντρα ή φυλλώματα θάμνων τα οποία συνιστούν σημεία υψηλής συμμετρίας και πράσινου χρώματος. Εδώ λοιπόν φαίνεται πόσο επιτακτική είναι η ανάγκη διαχωρισμού των δύο τμημάτων του καναλιού a^* και χρήσης μόνο του κόκκινου τμήματος καθώς τα σημεία αυτά πρέπει να απομονωθούν και να επικεντρωθούν αποκλειστικά στα πίσω φώτα.



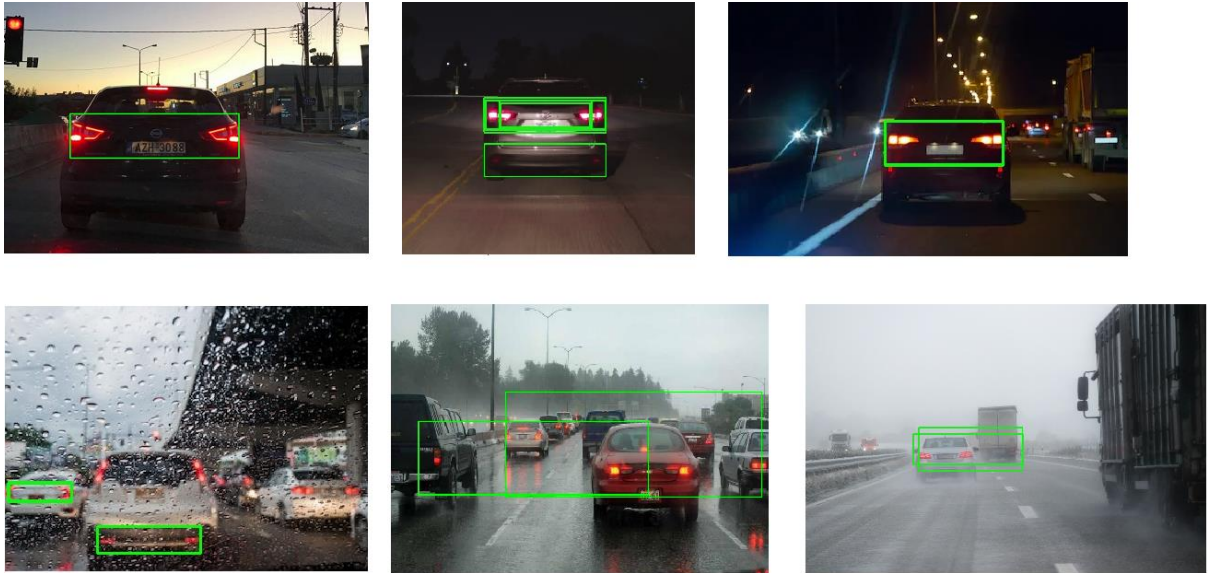
Εικόνα 33. Παράθεση μερικών επιτυχημένων παραδειγμάτων εφαρμογής του αλγορίθμου αναγνώρισης προπορευόμενου οχήματος σε σταθμευμένα αυτοκίνητα (Caltech 1999).

Το σύστημα λειτουργεί με μεγάλη ακρίβεια σε frames με σταθμευμένα οχήματα παρά το περίπλοκο υπόβαθρο που ενδεχομένως να υπάρχει (θάμνοι, δέντρα). Μία αδυναμία που εντοπίστηκε είναι οι false positive περιοχές εξαιτίας αντικειμένων κόκκινου χρώματος δίπλα στο μπροστινό αυτοκίνητο κάτι που παρατηρήσαμε και στο αστικό περιβάλλον (Εικόνα 34 (α)). Επίσης μία σπάνια περίπτωση αποτυχημένης ανίχνευσης είναι όταν το πίσω μέρος του αυτοκινήτου δεν είναι πλήρως συμμετρικό. Αυτό έχει ως αποτέλεσμα οι 2 υποπεριοχές ως προς την μεσοκάθετο να παρουσιάζουν μικρή ομοιότητα και η περιοχή ενδιαφέροντος να απορρίπτεται (false negative) (Εικόνα 34 (β)).



Εικόνα 34. Πιθανές περιπτώσεις αστοχιών του αλγορίθμου για σταθμευμένα αυτοκίνητα (Caltech1999).

Για την εξαγωγή ασφαλών συμπερασμάτων ως προς την αξιοπιστία του συστήματος πρέπει αυτό να αξιολογηθεί σε δύσκολες συνθήκες οδήγησης όπως την νύχτα, σε συννεφιά και υπό βροχόπτωση. Για τις συνθήκες αυτές ο νόμος επιβάλλει όλα τα οχήματα να έχουν αναμμένα τα φώτα τους κάτι που λειτουργεί θετικά για το σύστημα μας καθώς βασίζεται στην χρωματική πληροφορία. Η αξιολόγηση θα γίνει ποιοτικά σε εικόνες που έχουν ληφθεί από το διαδίκτυο λόγω έλλειψης πλήρων βάσεων δεδομένων για αντίξοες συνθήκες. Όσον αφορά την οδήγηση την νύχτα και ειδικά σε επαρχιακό δίκτυο, το σύστημα φαίνεται να λειτουργεί ικανοποιητικά καθώς το έντονο κόκκινο χρώμα των πίσω φαναριών επικρατεί των αντικειμένων που εισάγουν θόρυβο (π.χ σηματοδότες, πινακίδες), διευκολύνοντας έτσι τον σχεδιασμό της περιοχής ενδιαφέροντος (Εικόνα 35). Αντίθετα σε πυκνό αστικό περιβάλλον το σύστημα ενδέχεται να υστερεί ελαφρώς σε ακρίβεια εμφανίζοντας false positive περιοχές. Αυτό συμβαίνει διότι αυξάνεται ο θόρυβος τόσο από προπορευόμενα οχήματα που δεν βρίσκονται ολόκληρα στο frame όσο και από διαφημιστικές πινακίδες ή φωτεινούς σηματοδότες. Η οδήγηση στην βροχή αυξάνει τον βαθμό δυσκολίας του προβλήματος λόγω των σταγόνων νερού που παραμένουν στο παρμπρίζ και παρεμβάλλονται στο οπτικό πεδίο της κάμερας. Επίσης δημιουργούνται είδωλα στον δρόμο από τα αναμμένα φώτα των προπορευόμενων οχημάτων, τα οποία ικανοποιούν όλες τις συνθήκες και άρα είναι πολύ πιθανό να χαρακτηριστούν από το σύστημά μας ως False Positive περιοχές ενδιαφέροντος.



Εικόνα 35. Αποτελέσματα για νυχτερινή οδήγηση και οδήγηση στην βροχή.

4.5 Συμπεράσμα

Η ανάπτυξη ενός συστήματος αναγνώρισης προπορευόμενου οχήματος είναι πολύ σημαντική για ένα σύστημα υποβοήθησης της οδήγησης. Στο κεφάλαιο αυτό παρουσιάστηκε ένα σύστημα το οποίο βασίζεται στο στοιχείο του χρώματος και της συμμετρίας για την εύρεση υποψηφίων. Ύστερα προσδιορίζονται τα σημεία που αποτελούν ζεύγη και σχεδιάζεται μία περιοχή ενδιαφέροντος γύρω από αυτά. Τέλος αποδεκτές είναι μόνο οι περιοχές που είναι συμμετρικές ως προς την μεσοκάθετο. Το σύστημα ελέγχθηκε υπό διάφορα επίπεδα φωτεινότητας (ηλιοφάνεια, συννεφιά, σκιερός δρόμος) αρχικά σε απλό περιβάλλον με το μπροστινό όχημα να είναι σταθμευμένο και στην συνέχεια σε πολύπλοκο αστικό δίκτυο με πυκνή κυκλοφορία. Τα ποσοστά που πέτυχε στην πρώτη περίπτωση είναι αρκετά ικανοποιητικά ενώ στην δεύτερη περίπτωση το σύνθετο αστικό περιβάλλον οδηγεί σε μείωση του ποσοστού ανίχνευσης με την ακρίβεια ωστόσο να παραμένει υψηλή. Τέλος η μεγάλη εξάρτηση του συστήματος μας από το κόκκινο χρώμα των πίσω φαναριών, δίνει προοπτικές για την χρήση του υπό βροχόπτωση καθώς και την νύχτα. Σε επόμενο στάδιο το παρόν σύστημα μπορεί να εξελιχθεί ώστε βασιζόμενο στην ενεργοποίηση των πίσω φαναριών με το πάτημα του φρένου, να ειδοποιεί τον οδηγό ότι το μπροστινό όχημα επιβραδύνει διασφαλίζοντας την τήρηση των προβλεπόμενων αποστάσεων και την αποφυγή μία πιθανή σύγκρουσης.

5. Αναγνώριση Δρόμου

5.1 Εισαγωγή

Στο κεφάλαιο αυτό περιγράφεται το σύστημα αναγνώρισης δρόμου που αναπτύχθηκε στα πλαίσια της εργασίας και το οποίο βασίζεται στον αλγόριθμο Automated Random Walker [κεφ. 3.7], για την κατάταξη των pixels μίας εικόνας σε δύο κατηγορίες «δρόμος» ή «μη-δρόμος». Για την καθοδήγηση του αλγορίθμου αρχικά εκτιμούνται οι σπόροι «δρόμου» και «μη δρόμου». Σε ένα τυπικό σενάριο οδήγησης η περιοχή στην οποία είναι πιο πιθανό να βρεθεί ο δρόμος είναι ένα τραπέζιο ακριβώς μπροστά από το ego-vehicle το οποίο θέτω ως σπόρους «δρόμου», ενώ οι σπόροι «μη-δρόμου» προσδιορίζονται συνδυάζοντας τα σημεία υψηλής οπτικής ροής με την εκ των προτέρων πληροφορία για την θέση του ουρανού σε ένα frame.

Παρουσιάζονται επίσης οι πιο γνωστές προσεγγίσεις του προβλήματος που συναντώνται στην διεθνή βιβλιογραφία και στο τέλος καταγράφονται τα συμπεράσματα ως προς την αποτελεσματικότητα καθώς και το περιθώριο πιθανών βελτιώσεων και πρακτικών εφαρμογών του συστήματός μας.

5.2 Σχετική βιβλιογραφία

Σε πολλές εφαρμογές στην βιβλιογραφία συναντώνται μέθοδοι που βασίζονται αποκλειστικά στις διαγραμμίσεις του οδοστρώματος. Στην προσέγγιση [37] των Low, Zamzuri και Mazlan η εικόνα που λαμβάνεται από την κάμερα υποδειγματοληπτείται για μείωση του χρόνου επεξεργασίας και μετατρέπεται σε grayscale μορφή. Ύστερα μέσω του αλγορίθμου Canny εντοπίζονται οι ακμές-γραμμές και στην συνέχεια χρησιμοποιείται ο μετασχηματισμός Hough για να ενώσει τυχόν ασυνέχειες και να διαχωρίσει τις γραμμές σε δύο ομάδες (αριστερά και δεξιά) ανάλογα με την θέση τους. Τέλος αν εμφανιστούν πολλές γραμμές στην κάθε ομάδα τότε υπολογίζεται μία μέση τιμή που είναι και η αντιπροσωπευτική της. Τέτοιες μέθοδοι είναι πολύ γρήγορες στην υλοποίηση ωστόσο έχουν σοβαρά μειονεκτήματα. Πρώτον δεν μπορούν να λειτουργήσουν σε κακοσυντηρημένους δρόμους στους οποίους η διαγράμμιση έχει φθαρεί ή δεν υπάρχει καθόλου. Δεύτερον παρουσιάζει αδυναμία όταν οι καιρικές συνθήκες είναι αντίξοες, καθώς τότε ακόμη και αν ο δρόμος είναι σε καλή κατάσταση, οι διαγραμμίσεις δεν είναι ευδιάκριτες λόγω μειωμένης ορατότητας ή αντανάκλασεων. Οι παραπάνω περιορισμοί έστρεψαν το ενδιαφέρον σε υλοποιήσεις που δεν βασίζονται στην πληροφορία των διαγραμμίσεων προσδίδοντας ευελιξία και αξιοπιστία στο σύστημα.

Στο [38] η αρχική εικόνα επεξεργάζεται με ένα median value filter για την μείωση του θορύβου, μετατρέπεται σε δυαδική με την τεχνική two-peak και τον αλγόριθμο Otsu πολλαπλών κατωφλίων. Κατόπιν εντοπίζονται οι ακμές μέσω του Canny edge detector και τέλος χρησιμοποιείται η μέθοδος Mathematical morphology για την απόρριψη των περιττών

ακμών και την ενίσχυση των κυρίαρχων. Η μέθοδος αυτή υλοποιείται γρήγορα και είναι μικρού υπολογιστικού κόστους ωστόσο αποτυγχάνει όταν ο δρόμος έχει κάποια ανωμαλία όπως λακούβα, λιμνάζοντα νερά ή σκιερά σημεία. Αυτή η αδυναμία ωφείλεται σε μεγάλο βαθμό στην απουσία της χρωματικής πληροφορίας της εικόνας. Μία άλλη προσέγγιση του προβλήματος της εξαγωγής του δρόμου είναι η αξιοποίηση του χρώματος και της φωτεινότητας της εικόνας, με κάποια επιπλέον πληροφορία όπως οι ακμές.

Οι συγγραφείς του [39] προτείνουν ένα σύστημα για την αναγνώριση κακοσυντηρημένων δρόμων το οποίο δεν χρειάζεται πληροφορία από προηγούμενο frame, ούτε ένα μοντέλο πρόβλεψης του δρόμου και βασίζεται σε ταξινομητή με επιπλέον πληροφορία για το χρώμα και τις ακμές. Για κάθε νέα εικόνα στην είσοδο ομαδοποιεί τα pixels με παρόμοια χρώματα ανά περιοχές και τέλος επιλέγει τις περιοχές εκείνες που αν συνδυαστούν δίνουν το αναμενόμενο σχήμα του δρόμου. Το σύστημα είναι ανθεκτικό στις μεταβολές της φωτεινότητας καθώς τα pixels ομαδοποιούνται ανεξάρτητα για κάθε εικόνα ενώ η χρήση των τριών καναλιών αντί του ενός (π.χ grayscale) επιτρέπει την διατήρηση όλης της πληροφορίας οδηγώντας σε αποτελεσματικότερη τμηματοποίηση του δρόμου.

Στο [40] η μελέτη εστιάζεται στην αναγνώριση δρόμου σε αστικό και ημι-αστικό περιβάλλον το οποίο ως επί των πλείστον χαρακτηρίζεται από α) ανεπαρκή διαγράμμιση, β) δισδιάκριτα όρια του οδοστρώματος, γ) παρουσία πολλών οχημάτων και πεζών. Το πρόβλημα χωρίζεται σε δύο σκέλη. Αρχικά από την μονοχρωματική αναπαράσταση της εικόνας εκτιμώνται τα άκρα του δρόμου και στην συνέχεια δίνονται ως είσοδος σε μία διεργασία εύρεσης του οδοστρώματος η οποία βασίζεται στην πλήρη χρωματική αναπαράσταση της εικόνας. Σε παρόμοια φιλοσοφία κινείται και το [41] όπου συνδυάζονται ένα morphological temporal filter για μείωση θορύβου, αλγόριθμος ανίχνευσης ακμών και τέλος ο μετασχηματισμός Watershed πάνω στην gradient εικόνα. Στην προσέγγιση των P. Lombardi, M. Zanin και S. Messelodi [42] τα frames λαμβάνονται από στερεοσκοπική κάμερα και υπολογίζεται ο χάρτης disparity map, που εκφράζει τις αποστάσεις των αντίστοιχων pixel μεταξύ αριστερού και δεξιού frame. Με τον τρόπο αυτό παράγεται από την αρχική εικόνα μία νέα στην οποία τονίζονται οι ακμές όπως άκρα του δρόμου, εμπόδια και τυχόν διαγραμμίσεις. Στην συνέχεια ένας model based αλγόριθμος εργάζεται πάνω στην πληροφορία αυτή και τμηματοποιεί την περιοχή του δρόμου. Κατά την αξιολόγηση το σύστημα είχε πολύ καλή συμπεριφορά σε κακοσυντηρημένους δρόμους με ανεπαρκή διαγράμμιση.

Μία άλλη κατηγορία συστημάτων είναι εκείνα που αξιοποιούν την οπτική ροή. Στην μελέτη των Γ.Σίογκα και Ε.Δερματά [43] τα pixels διαχωρίζονται σε δρόμος και μη-δρόμος με χρήση του supervised αλγορίθμου Automated Random Walker. Ο προσδιορισμός των σπόρων δρόμου και μη-δρόμου γίνεται συνδυάζοντας την εκ των προτέρων πληροφορία (π.χ ουρανός) με ένα προσχέδιο δρόμου το οποίο προκύπτει από το αποτέλεσμα της προηγούμενης ανίχνευσης. Η επεξεργασία των frames ανά δύο επιτρέπει την χρήση της οπτικής ροής για την ακριβέστερη τοποθέτηση των σπόρων μη-δρόμου οδηγώντας σε καλύτερα αποτελέσματα. Ο συνδυασμός πολλών χαρακτηριστικών, χωρικών και χρονικών επιτρέπει στο σύστημα να λειτουργεί σε επαρχιακούς δρόμους και σε κακές καιρικές

συνθήκες ενώ η χρήση του καναλιού c1 του χώρου c1c2c3 εξασφαλίζει ανθεκτικότητα σε σκιές και ανωμαλίες του οδοστρώματος.

Στην βιβλιογραφία συναντώνται συστήματα που συνδυάζουν την μηχανική όραση με νευρωνικά δίκτυα για την εύρεση της περιοχής του δρόμου σε ένα frame. Στο [44] χρησιμοποιείται deep neural network για την εκμάθηση της υφής του δρόμου και ένα μοντέλο έγχρωμων γραμμών (color lines) για την εκμάθηση της μορφής των μη-συντηρημένων δρόμων. Παρόλο την αξιοσημείωτη εξέλιξη των νευρωνικών δικτύων το μειόνεκτημα αυτής της προσέγγισης είναι ο μεγάλος χρόνος εκπαίδευσης και η έλλειψη δημόσιων βάσεων δεδομένων από οδικά δίκτυα.

Όλες οι παραπάνω μέθοδοι χρησιμοποιούν αποκλειστικά frames που λαμβάνονται μία κάμερα τοποθετημένη στο παρμπρίζ του αυτοκινήτου. Η επίδοση τους είναι ικανοποιητική για τυπικά σκηνικά οδήγησης, ωστόσο παρατηρούνται αστοχίες σε δυσχερείς συνθήκες (βροχή ομίχλη), την νύχτα καθώς και όταν διερχόμενο όχημα έχει έντονα φώτα. Βασιζόμενοι σε αυτά μερικοί ερευνητές εστίασαν στην χρήση αισθητήρων radar για το πρόβλημα της αναγνώρισης του δρόμου. Στο [45] προτίνεται μία μέθοδος εύρεσης των ορίων της λωρίδας με χρήση χαρακτηριστικών Range-Doppler ενώ στο [46] αναπτύσσεται ένα σύστημα βασισμένο στα Vehicle Sensor Networks (VSN) τα οποία γίνονται ολό ένα και πιο δημοφιλή. Συγκεκριμένα χρησιμοποιείται ένα μοντέλο Bayesian δικτύου με τρία επίπεδα για την εξαγωγή των χαρακτηριστικών από την είσοδο των αισθητήρων, την ανάλυση του περιβάλλοντος και της συμπεριφοράς των οχημάτων.

Μελετώντας την βιβλιογραφία τα συμπεράσματα που προκύπτουν είναι τα εξής:

- a) Αλγόριθμοι που εστιάζουν στην εύρεση ακμών και δεν λαμβάνουν υπ' όψη την χρωματική πληροφορία της εικόνας είναι γρήγοροι στην υλοποίηση αλλά αποτυγχάνουν σε περιπτώσεις όπου ο δρόμος έχει ανωμαλίες όπως σκιές, λακούβες ή λιμνάζοντα νερά. Επίσης είναι ευαίσθητοι σε αλλαγές φωτεινότητας.
- b) Ο συνδυασμός περισσότερων χαρακτηριστικών όπως χρώμα, ακμές, οπτική ροή και πληροφορία προηγούμενου αποτελέσματος, αυξάνουν αφενός την πολυπλοκότητα αλλά προσφέρουν μεγάλη ακρίβεια και υψηλότερη απόδοση ακόμη και σε αντίξοες συνθήκες οδήγησης. Τέτοια συστήματα είναι κατάλληλα να εφαρμοστούν για αστικά ή επαρχιακά δίκτυα ανεξαρτήτου κατάστασης (καλοσυντηρημένα ή όχι).
- c) Η χρήση νευρωνικών δικτύων για το κομμάτι του προσδιορισμού μίας περιοχής ως δρόμος ή μη δεν είναι ιδιαίτερα δημοφιλής εξαιτίας της χρονοβόρας εκπαίδευσης των δικτύων και την έλλειψη επαρκών βάσεων με δείγματα προς εκπαίδευση. Οι επιδόσεις τέτοιων συστημάτων ωστόσο είναι αρκετά υψηλές.
- d) Ανεξαρτήτως της προσέγγισης που ακολουθείται η απόδοση των συστημάτων εξαρτάται σε μεγάλο βαθμό από την πολυπλοκότητα του σκηνικού. Για παράδειγμα

στο αστικό περιβάλλον το πρόβλημα γίνεται αρκετά περίπλοκο λόγω της αυξημένης κίνησης και του σύνθετου υπόβαθρου.

- ε) Η απόδοση του συστήματος επηρεάζεται από πολλούς παράγοντες οι οποίοι συναντώνται συχνά στην οδήγηση και είναι απρόβλεπτοι όπως παρεμβολή των υαλοκαθαριστήρων στο οπτικό πεδίο της κάμερας, αντανάκλασεις του εσωτερικού του αυτοκινήτου στο παρμπρίζ, είδωλα που δημιουργούνται από λιμνάζοντα νερά του δρόμου κ.α.

Στην συνέχεια αποσκοπούμε στην ανάπτυξη ενός συστήματος που να λειτουργεί όσο το δυνατόν πιο αξιόπιστα υπό διάφορες καιρικές συνθήκες και ανεξάρτητα της ποιότητας του οδοστρώματος και της πολυπλοκότητας του περιβάλλοντος.

5.3 Προτινόμενο σύστημα

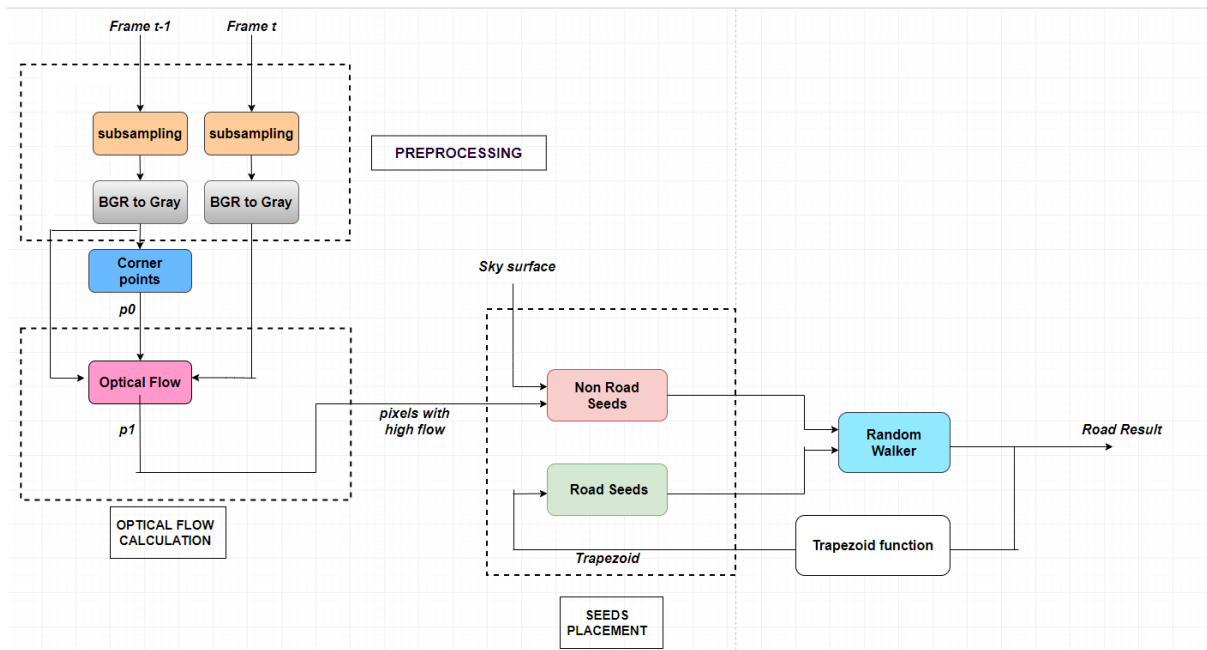
Για την ανάπτυξη του συστήματος βασιζόμαστε στην μηχανική όραση λόγω της πληθώρας τεχνικών και αλγορίθμων που υπάρχουν στην βιβλιογραφία καθώς και για να είναι δυνατή η ενοποίηση του, με το σύστημα αναγνώριση προπορευόμενου οχήματος του προηγούμενου κεφαλαίου. Με τον τρόπο αυτό θα έχουμε ένα ολοκληρωμένο μοντέλο ικανό να χρησιμοποιηθεί για την υποβοήθηση της οδήγησης ή και για αυτόνομη πλοήγηση ενός οχήματος.

Η είσοδος λαμβάνεται από κάμερα που είναι τοποθετημένη εσωτερικά του αυτοκινήτου στο πάνω μέρος του παρμπρίζ ώστε να προστατεύεται από τα καιρικά φαινόμενα. Το βασικό ζήτημα εδώ είναι η κατάταξη των pixels σε δύο κλάσεις, δρόμο ή μη δρόμο. Για την τμηματοποίηση της εικόνας χρησιμοποιούμε τον supervised αλγόριθμο Random Walker [3.7] δίνοντας του μερικά seeds για τις δύο κλάσεις. Όπως αναφέραμε στο κεφάλαιο 3.7 η «τυχαία» φύση του αλγορίθμου οδηγεί σε λανθασμένο αποτέλεσμα αν τα seeds δεν προσδιοριστούν με ακρίβεια. Για να αποτρέψουμε κάτι τέτοιο προσπαθούμε να αντλήσουμε όση περισσότερη πληροφορία μπορούμε από την εικόνα εισόδου και να προσδιορίσουμε όσο το δυνατόν περισσότερα seeds με ικανοποιητική ακρίβεια.

Για τα σημεία μη-δρόμου εκμεταλλευόμαστε την οπτική ροή και την εκ των προτέρων πληροφορία της τοποθέτησης της κάμερας, ενώ για τα σημεία δρόμου λαμβάνουμε υπ' όψη την περιοχή δρόμου που προέκυψε από προηγούμενη έξοδο του αλγορίθμου. Με τον τρόπο αυτό εξασφαλίζουμε ότι η τοποθέτηση των seeds θα είναι πάντα δυναμική και εξαρτώμενη μόνο από την τρέχουσα εισόδο t και το προηγούμενο αποτέλεσμα για είσοδο $t-1$.

Στην Εικόνα 36 παρουσιάζεται το διάγραμμα ροής των λειτουργιών του αλγορίθμου που αναπτύξαμε και αποτελείται από τα παρακάτω βήματα:

- 1) Λαμβάνεται το frame (t-1), υποδειγματοληπτείται για την μείωση του χρόνου επεξεργασίας και μετατρέπεται σε grayscale μορφή.
- 2) Προσδιορίζονται τα σημεία ενδιαφέροντος -διάνυσμα p_0 - μέσω της τεχνικής Shi-Tomasi [3.6.2.b]. Τα σημεία αυτά είναι γωνίες της εικόνας για τα οποία στην συνέχεια θα υπολογιστεί η οπτική ροή.
- 3) Λαμβάνεται το επόμενο frame (t), υποδειγματοληπτείται και μετατρέπεται σε grayscale μορφή.
- 4) Υπολογίζεται η οπτική ροή μέσω της μεθόδου Lucas-Kanade [3.6.2.c] έχοντας ως είσοδο τα 2 μονοχρωματικά frames και τα σημεία ενδιαφέροντος p_0 που ορίστηκαν για t-1. Ο αλγόριθμος δίνει στην έξοδο τις θέσεις των νέων σημείων ενδιαφέροντος -διάνυσμα p_1 - για την στιγμή t.
- 5) Ορίζονται ως non-road seeds τα pixels που βρίσκονται ψηλά στο frame και αποτελούν την περιοχή του ουρανού καθώς και τα σημεία με υψηλή οπτική ροή (p_1) που υπολογίστηκαν από τον Lucas-Kanade.
- 6) Με βάση το προηγούμενο αποτέλεσμα του αλγορίθμου για τα frames t-2 και t-1, δημιουργείται ένα τραπέζιο το οποίο τροφοδοτείται στην νέα ανίχνευση και εντός της περιοχής του τοποθετούνται όλα τα seeds δρομου.
- 7) Το σύνολο των seeds δίνεται ως είσοδος στον Random Walker, υπολογίζεται για κάθε unlabeled pixel η πιθανότητα να ανήκει σε δρόμο ή μη δρόμο και καταχωρείται στην κλάση με την μεγαλύτερη πιθανότητα πρώτης εμφάνισης. Ως εξοδος επιστρέφεται μία δυαδική εικόνα με άσπρο στην περιοχή του δρόμου και μαύρο αλλού.
- 8) Πριν μεταβούμε στην επόμενη επανάληψη ενημερώνουμε την μεταβλητή του frame t-1 με το τρέχον frame t και αναθέτουμε στο p_0 το p_1 .
- 9) Επαναλαμβάνουμε τα βήματα 2-8 μέχρις ότου επεξεργαστούν όλα τα frames.



Εικόνα 36. Διάγραμμα ροής του συστήματος αναγνώρισης δρόμου.

Για την καλύτερη κατανόηση της λειτουργίας του συστήματος χωρίζουμε τις επιμέρους ενέργειες σε σε τρία βασικά στάδια τα οποία είναι: α) προεπεξεργασία των εικόνων, β) εύρεση της οπτικής ροής και γ) τοποθέτηση των seeds. Παρακάτω εμβαθύνουμε σε καθένα από αυτά και παρουσιάζουμε εναλλακτικές προσεγγίσεις που δοκιμάστηκαν.

5.3.1 Προεπεξεργασία

Η είσοδος του συστήματος προέρχεται από κάμερα που είναι τοποθετημένη στο παρμπρίζ του αυτοκινήτου και σε σημείο που δεν εμποδίζει το οπτικό πεδίο του οδηγού. Για να αξιοποιήσουμε την πληροφορία της οπτικής ροής επεξεργαζόμαστε τα frames του βίντεο ανά δύο, σε αντίθεση με το πρώτο σκέλος της αναγνώρισης προπορευόμενου οχήματος όπου μας αρκούσε ένα frame για την εξαγωγή του αποτελέσματος. Κατά την εκτέλεση του προγράμματος παρατηρήθηκε μεγάλη καθυστέρηση, σε αντίθεση με ένα αξιόπιστο σύστημα σύστημα υποβοήθησης της οδήγησης όπου η διαδικασία αυτή γίνεται σε πραγματικό χρόνο. Η αιτία αυτή της καθυστέρησης πηγάζει κατά πάσα πιθανότητα από την φύση του Random Walker. Όπως αναφέραμε και προηγουμένως, ο αλγόριθμος έχει μεγάλη πολυπλοκότητα εφόσον καλείται να υπολογίσει τις πιθανότητες πρώτης άφιξης για κάθε unlabeled pixel, εμπλέκοντας τα βάρη όλων των ενδιάμεσων σημείων. Μία λύση του προβλήματος είναι να υποδειγματοληπτήσουμε την αρχική εικόνα και ύστερα να την επεξεργαστούμε. Στον Πίνακα 5 φαίνονται οι διάφορες κλίμακες που δοκιμάστηκαν με τους αντίστοιχους χρόνους απόκρισης του αλγορίθμου. Παρατηρούμε ότι διατηρώντας αναλλοίωτες τις διαστάσεις, η τελική δυαδική εικόνα με την περιοχή του δρόμου, παράγεται σε 4 λεπτά και 40

δευτερόλεπτα. Κατόπιν αν μειώσουμε την εικόνα στο 70% της αρχικής ο χρόνος ολοκλήρωσης του αλγορίθμου υποτετραπλασιάζεται και γίνεται 58 δευτερόλεπτα. Με το ίδιο σκεπτικό για υποδειγματοληψία στο 50% ο χρόνος εκτέλεσης είναι 24 δευτερόλεπτα ενώ για επιπλέον μείωση των διαστάσεων στο 30% η τελική δυαδική εικόνα με την περιοχή δρόμου προκύπτει σε μόλις 12 sec.

Κλίμακα (ποσοστό αρχική εικόνας)	Χρόνος απόκρισης
100 %	4': 40"
70 %	58"
50 %	24"
30 %	12"
10 %	10"

Πίνακας 5. Διαφορετικές περιπτώσεις υποδειγματοληψίας της αρχικής εικόνας και αντίστοιχοι χρόνοι εκτέλεσης του αλγορίθμου.

Η υποδειγματοληψία της εισόδου επομένως είναι σημαντική και αναμφίβολα ελαττώνει τον απαιτούμενο χρόνο εκτέλεσης ωστόσο πρέπει να διερευνήσουμε αν επηρεάζεται η ακρίβεια του αποτελέσματος. Στην Εικόνα 37, φαίνονται τα αποτελέσματα του αλγορίθμου για τις παραπάνω κλίμακες της αρχικής εικόνας. Αν δεν μεταβάλλουμε καθόλου τις διαστάσεις του frame (Εικόνα 37 (β)) τότε το αποτέλεσμα θεωρείται ικανοποιητικό αλλά αργεί σημαντικά η εκτέλεση (Πίνακας 5, 1^η γραμμή). Στα (γ) και (δ) η εικόνα μειώνεται κατά 70% και 50% αντίστοιχα μειώνοντας σημαντικά τον χρόνο ενώ και το αποτέλεσμα είναι καλύτερο εφόσον το πεζοδρόμιο στα δεξιά αρχίζει να ξεχωρίζει από την περιοχή δρόμου. Στο (ε) η δειγματοληψία είναι στο 30% ο χρόνος έχει μειωθεί στα μόλις 12 δευτερόλεπτα ενώ το η τμηματοποίηση είναι καλύτερη από τις προηγούμενες περιπτώσεις διαχωρίζοντας πλήρως πεζοδρόμιο με δρόμο. Τέλος στο (ζ) για 10%, η αρχική εικόνα φαίνεται θολή και έχει χαθεί σημαντική πληροφορία όσον αφορά το ακριβές σχήμα και την θέση των αντικειμένων. Επίσης η απόκριση του αλγορίθμου δεν βελτιώνεται σημαντικά αλλά μόλις κατά 2 δευτερόλεπτα συγκριτικά με την περίπτωση (ε). Για το σύστημα που αναπτύσσουμε και για τους λόγους που αναφέρθηκαν παραπάνω, επιλέγουμε να υποδειγματοληπτήσουμε τα frames στο 30% εξασφαλίζοντας έτσι μία ισορροπία μεταξύ αξιοπιστίας και ταχύτητας στην εκτέλεση.



(α)



(β)



(γ)



(δ)



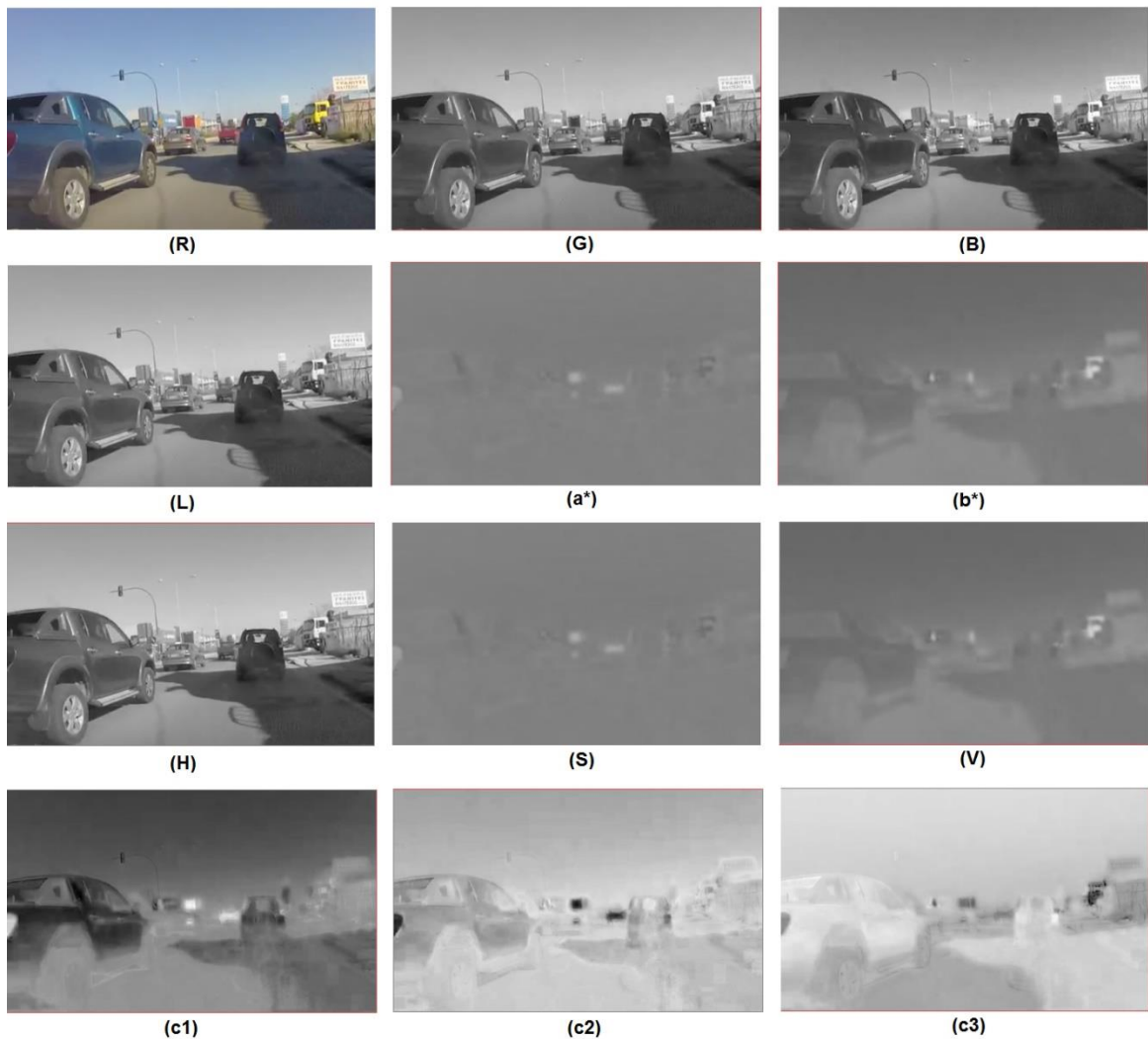
(ε)



(ζ)

Εικόνα 37. Αποτελέσματα αναγνώρισης δρόμου για διάφορες περιπτώσεις υποδειγματοληψίας. Η διαφορετική απόχρωση στα β, γ, δ, ε και ζ ωφείλεται σε παράμετρο της μεθόδου απεικόνισης plot της βιβλιοθήκης matplotlib.

Επόμενο βήμα είναι να καθορίσουμε τον χρωματικό χώρο που θα χρησιμοποιηθεί προκειμένου να αντλήσουμε την μέγιστη δυνατή πληροφορία από την εικόνα. Αρχικά, είναι σημαντικό να επισημάνουμε ότι ο αλγόριθμος υπολογισμού της ροής δέχεται εικόνα ενός καναλιού ώστε κάθε pixel να προσδιορίζεται από μία και μόνο τιμή. Όπως είδαμε στην βιβλιογραφία μία συχνή αδυναμία των συστημάτων εμφανίζεται όταν το οδόστρωμα έχει σκιερά σημεία, λιμνάζοντα νερά και γενικότερα ασυνέχειες στην υφή του. Για να αποφασίσουμε που θα εργαστούμε αρκεί να συγκρίνουμε τους βασικούς χώρους και τα επιμέρους κανάλια τους, σε μία εικόνα με περιοχές σκιών (Εικόνα 38). Βλέπουμε ότι στις τρεις συνιστώσες του RGB καθώς και στα κανάλια L και H των CIE Lab και HSV αντιστοίχα, οι σκιές δεν εξασθενούν αλλά παραμένουν σε ίδια ένταση. Ύστερα τα a^* και S δεν είναι κατάλληλα καθώς το σχήμα του δρόμου χάνεται εντελώς, ενώ στα b^* και V είναι διασδιάκριτο το ακριβές σχήμα των οχημάτων. Τέλος μεταξύ των $c1$, $c2$ και $c3$ το κανάλι εκείνο που προσφέρει την βέλτιστη απόσβεση των σκιών και ταυτόχρονα διατηρεί ευδιάκριτες τις ακμές των αντικειμένων και του δρόμου είναι το $c1$.



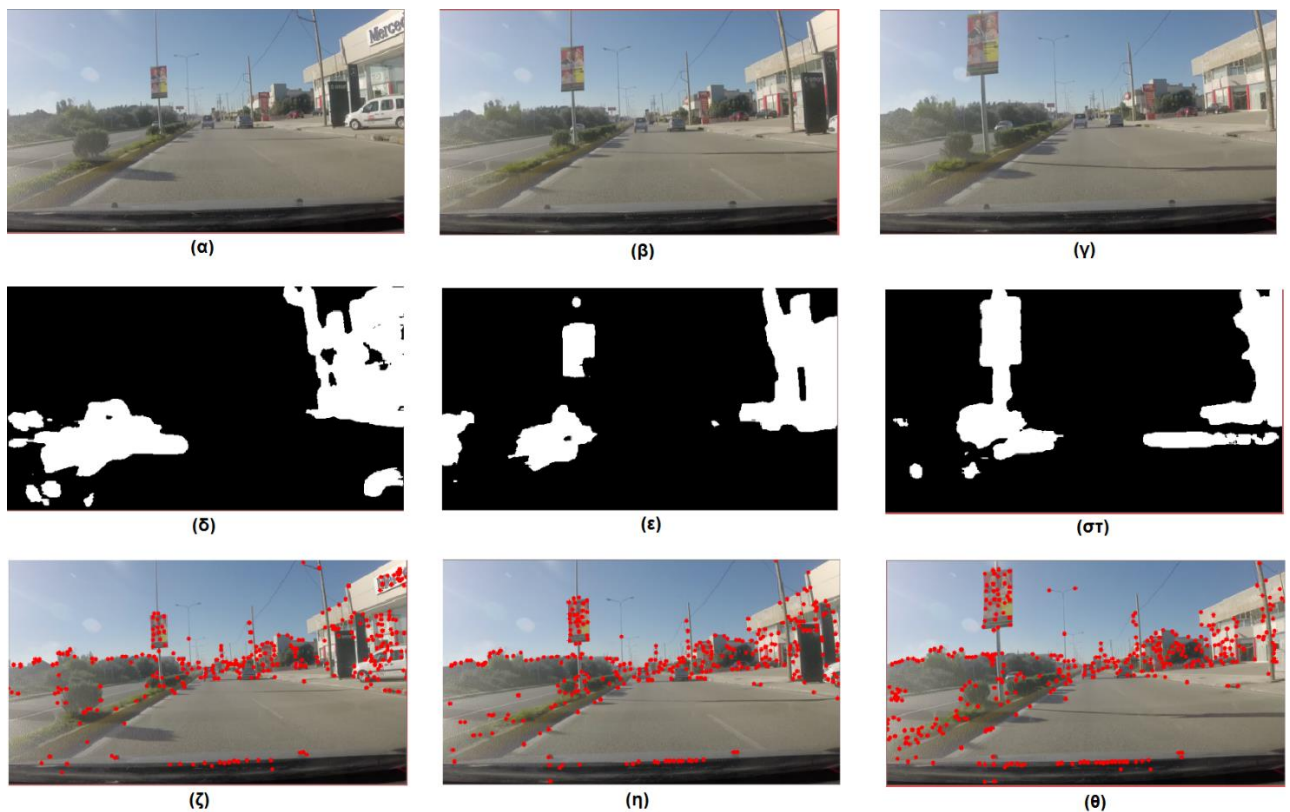
Εικόνα 38. Σύγκριση των χρωματικών χώρων RGB, CIELab, HSV και $c1c2c3$ ως προς την αντοχή στις ασυνέχειες της υφής του δρόμου.

Το κανάλι $c1$ είναι και εκείνο που προτίνεται στην δημοσίευση [43], βελτιώνοντας σημαντικά την αντοχή του αλγορίθμου και αυξάνοντας την αξιοπιστία του. Δεδομένου ότι κάθε σύστημα έχει διαφορετική δομή και φιλοσοφία πρέπει να συγκρίνουμε τα αποτελέσματα του αλγορίθμου μας για κάθε υπονήγιο χρωματικό κανάλι και στο τέλος να επιλέξουμε ανάλογα. Επίσης η επιλογή μας αυτή θα εξαρτάται και από την τεχνική υπολογισμού οπτικής ροής που θα χρησιμοποιήσουμε.

5.3.2 Οπτική ροή

Το προτινόμενο σύστημα προσπαθεί να αντλήσει όσο περισσότερη πληροφορία μπορεί από τις εικόνες εισόδου ώστε να γίνει η βέλτιστη τοποθέτηση των seeds. Κατά την εμπρόσθια κίνηση, ο οδηγός ενός οχήματος -στην δική μας περίπτωση η κάμερα- βλέπει το

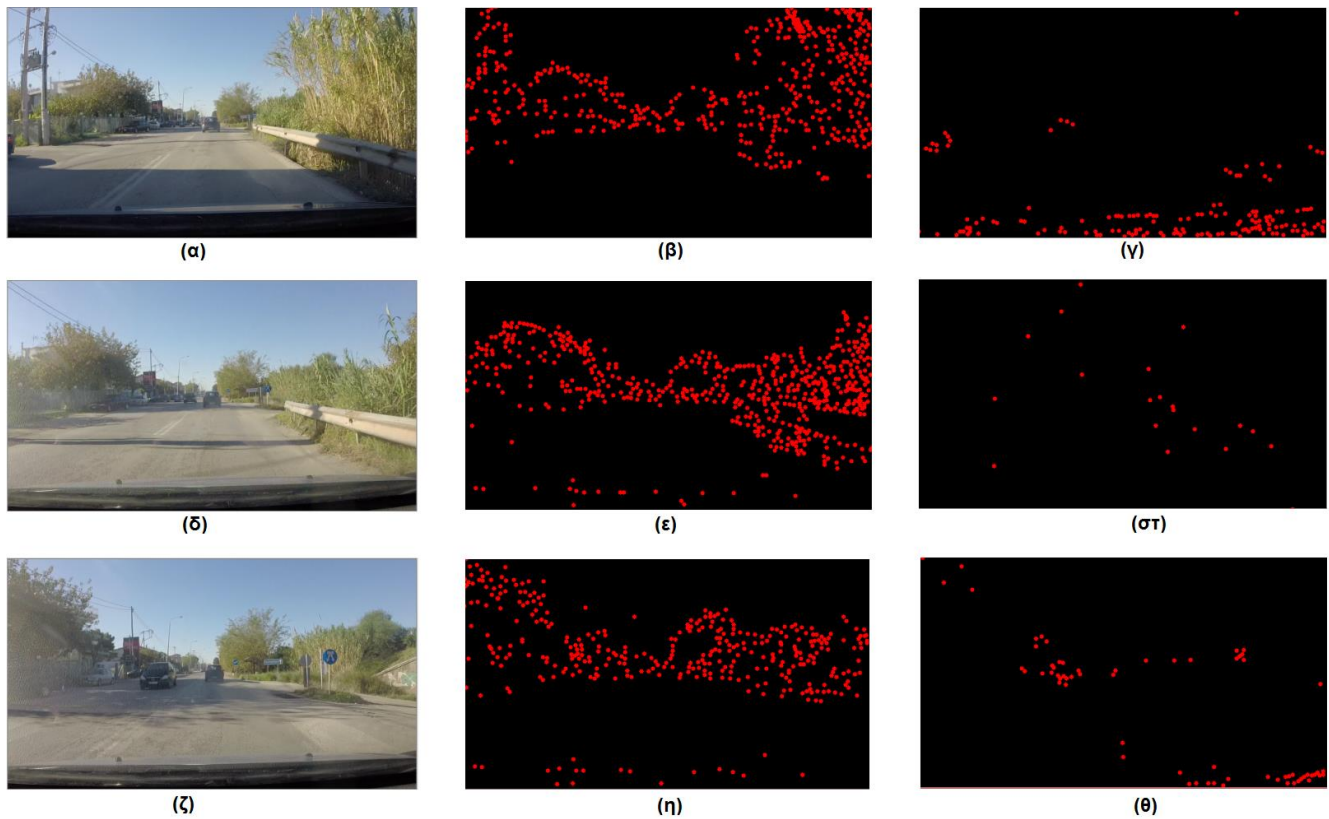
τοπίο να κινείται με ταχύτητα προς το μέρος του ενώ το οδόστρωμα φαίνεται στάσιμο λόγω της λείας υφής του. Αυτό μπορούμε να το εκμεταλλευτούμε για να οριοθετήσουμε την περιοχή του δρόμου δηλαδή να προσδιορίσουμε τα road seeds. Ένα χρήσιμο μέτρο που δείχνει την μετατόπιση-κίνηση των αντικειμένων μεταξύ διαδοχικών frames είναι η οπτική ροή. Στο κεφάλαιο 3.6 διακρίναμε την αραιή με την πυκνή οπτική ροή και περιγράψαμε τα χαρακτηριστικά της καθεμίας. Το ζητούμενο για το σύστημα μας είναι να εξασφαλιστεί μία ισορροπία μεταξύ αξιοπιστίας και ταχύτητας. Η αραιή οπτική ροή εστιάζει σε μερικά σημεία ενδιαφέροντος της εικόνας, υπολογίζει την ροή τους και προσπαθεί να τα “ανιχνεύσει” στο επόμενο frame. Με τον τρόπο αυτό ο χρόνος εκτέλεσης μειώνεται αλλά ίσως επηρεάζεται η ακρίβεια. Από την άλλη πλευρά, η πυκνή οπτική ροή εργάζεται πάνω σε κάθε pixel της εικόνας αυξάνοντας την πολυπλοκότητα και το φόρτο εργασίας του αλγορίθμου, ενώ οδηγεί σε πληρέστερη εκτίμηση της ροής τοπικά. Για άλλη μία φορά είμαστε αντιμέτωποι με την εύρεση του βέλτιστου συνδυασμού ταχύτητας – ακρίβειας που θα μας δώσει ένα ικανοποιητικό αποτέλεσμα. Παρακάτω, στην Εικόνα 39 δοκιμάζουμε και τις δύο τεχνικές υπολογισμού της οπτικής ροής. Στην 1^η γραμμή έχουμε 3 διαδοχικά frames σε RGB μορφή από δρόμο της Πάτρας. Στην 2^η γραμμή παρατίθενται τα αποτελέσματα που λάβαμε χρησιμοποιώντας τον αλγόριθμο Farneback [κεφ. 3.6.1], ενώ στην 3^η γραμμή τα αποτελέσματα που λάβαμε χρησιμοποιώντας τον αλγόριθμο Lucas Kanade [κεφ 3.6.2].



Εικόνα 39. Υπολογισμός της οπτικής ροής με χρήση των αλγορίθμων Farneback και Lucas Kanade

Παρατηρούμε ότι η αραιή ροή δίνει πολύ καλύτερο αποτέλεσμα καθώς ορίζει με ακρίβεια τα άκρα του δρόμου και ανιχνεύει την σχετική κίνηση των σημαντικότερων αντικειμένων του υποβάθρου (κτήρια, πινακίδες, οχήματα). Αντίθετα με την πυκνή ροή, ενώ θεωρητικά θα περιμέναμε καλύτερο αποτέλεσμα αυτό δεν συμβαίνει στην πράξη καθώς τα άκρα του δρόμου δεν εντοπίζονται όπως και τα κινούμενα αντικείμενα στο βάθος του frame. Ένα άλλο σημαντικό πλεονέκτημα της αραιής ροής είναι ότι ο χρόνος επεξεργασίας κάθε ζεύγους frames είναι περίπου μισός σε σχέση με την πυκνή ροή. Αυτό είναι πολύ σημαντικό καθώς στην τελική αναπτύσσουμε ένα εν δυνάμει σύστημα πραγματικού χρόνου. Επί προσθέτως η επιλογή της πυκνής οπτικής ροής φαίνεται να περιπλέκει το πρόβλημα καθώς η πληροφορία που παράγει είναι πλεονάζουσα. Μην ξεχνάμε ότι το ζητούμενο είναι ο απλός διαχωρισμός των «κινούμενων» αντικειμένων από την επιφάνεια δρόμου, άρα η εύρεση της οπτικής ροής μόνο στα σημεία ενδιαφέροντος μας ικανοποιεί χωρίς να επιβαρύνει το σύστημα.

Για τους παραπάνω λόγους καταφεύγουμε στην επιλογή της αραιής οπτικής ροής μέσω του αλγορίθμου Lucas – Kanade, ο οποίος υποστηρίζεται από την OpenCV με κατάλληλη συνάρτηση. Εφόσον επιλάξαμε τον αλγόριθμο απομένει να προσδιορίσουμε σε ποιο κανάλι θα εργαστούμε. Στο κεφάλαιο 5.3.1 συγκρίναμε τους βασικούς χώρους και καταλήξαμε ότι η grayscale μορφή δεν ενισχύει τις σκιές ούτε τις αποσβένει ενώ c1 προκαλεί σημαντική εξασθένηση των σκιών. Πρωτού προχωρήσουμε καλό θα ήταν να συγκρίνουμε την συμπεριφορά του αλγορίθμου στο καθένα κανάλι. Στην Εικόνα 40 στην πρώτη στήλη έχουμε 3 διαφορετικά frames, ενώ στην δεύτερη και τρίτη στήλη φαίνεται το αποτέλεσμα υπολογισμού της οπτικής ροής με χρήση του Lucas Kanade για grayscale μορφή και κανάλι c1 αντίστοιχα. Η χρήση του c1 (Εικόνα 40 (γ), (στ), (θ)) δεν μας προσφέρει κάποια πληροφορία καθώς τα χρωματισμένα σημεία δεν αντιστοιχούν στα πραγματικά σημεία υψηλής ροής. Αντίθετα με χρήση της grayscale μορφής, το αποτέλεσμα είναι πολύ πιο ακριβές καθώς εντοπίζονται πολλά σημεία του τοπίου που εμφανίζουν υψηλή ροή όπως δέντρα, πλαϊνές διαχωριστικές μπαριέρες καθώς και διερχόμενα οχήματα. Επίσης παρόλο που σε όλα τα frames είναι έντονη η παρουσία σκιών, το σύστημα φαίνεται να μην επηρεάζεται. Αυτό είναι πολύ σημαντικό καθώς οι περιοχές αυτές δεν αναγνωρίζονται ως σημεία υψηλής ροής, επιτρέποντας στην συνέχεια την σωστή καταχώρηση τους ως seeds δρόμου. Μία άλλη χρήσιμη παρατήρηση είναι ότι εντοπίζονται επιτυχώς μέχρι και τα σημεία στο βάθος του frame επιτρέποντας στο σύστημα να «γνωρίζει» το σχήμα του δρόμου και τα πιθανά εμπόδια για μελλοντικές χρονικές στιγμές. Τέλος αξίζει να σημειωθεί ότι στα frames (ε) και (η) έχουν σημειωθεί εσφαλμένα με κόκκινο χρώμα μερικά σημεία στην βάση του frame. Αυτά ανήκουν στο καπό του ego-vehicle και θα τα διαχειριστούμε διαφορετικά στην συνέχεια.



Εικόνα 40. Υπολογισμός της οπτικής ροής για frames από δρόμο της Πάτρας (1η στήλη) με χρήση grayscale εισόδων (2η στήλη) και εισόδων καναλιού c1 (3η στήλη).

Τα ανωτέρω πειράματα μας οδηγούν στο συμπέρασμα ότι η grayscale μορφή της εικόνας και ο αλγόριθμος οπτικής ροής Lucas Kanade είναι ένας καλός συνδυασμός για το σύστημα μας και τα αποτελέσματα μέχρι στιγμής φαίνονται αξιόπιστα. Στην συνέχεια μπορούμε να προχωρήσουμε στην ανάλυση της διαδικασίας εύρεσης της οπτικής ροής.

5.3.2.a Σημεία ενδιαφέροντος

Όπως προαναφέραμε η προσέγγιση των Lucas Kanade εκτιμά την αραιή οπτική ροή. Αυτό σημαίνει ότι δεν ενεργεί σε όλα τα pixels του frame αλλά εστιάζει σε μερικά σημεία ενδιαφέροντος. Σε πρώτο στάδιο λοιπόν πρέπει να εντοπίσουμε αυτά τα σημεία τα οποία συνήθως είναι γωνίες και ακμές αντικειμένων. Η python διαθέτει κατάλληλη συνάρτηση για αυτό, η οποία βασίζεται στην τεχνική Shi-Tomasi. Η συνάρτηση δέχεται ως παράμετρος :

- Το πλήθος των επιθυμητών σημείων ενδιαφέροντος που θέλουμε να εντοπίσουμε, ενώ αν βρεθούν περισσότερα τότε επιστρέφονται τα επικρατέστερα.
- Την ελάχιστη αποδεκτή ποιότητα σημείων που επιθυμούμε να λάβουμε.
- Την ελάχιστη απόσταση των σημείων μεταξύ τους.

Ορίζοντας τις κατάλληλες τιμές των παραπάνω παραμέτρων μας δίνεται η δυνατότητα να ρυθμίσουμε την ευαισθησία της μεθόδου και την ποιότητα των επιστρεφόμενων σημείων. Κατά την εκτέλεση του αλγορίθμου Lucas Kanade ωστόσο, παρατηρήσαμε ότι τα αρχικά N σημεία ενδιαφέροντος που ανιχνεύσαμε στο πρώτο frame, σταδιακά χάνονται. Αυτό είναι λογικό καθώς το σκηνικό μεταβάλλεται γρήγορα με αποτέλεσμα τα αντικείμενα στα οποία είχαν βρεθεί γωνίες για το frame $t=0$, πλέον δεν υπάρχουν στο οπτικό πεδίο της κάμερας. Μία λύση σε αυτό το ζήτημα είναι να ανανεώνεται συχνά η ομάδα των σημείων ενδιαφέροντος ώστε να ανταποκρίνονται στα νέα σχήματα της εικόνας. Για να το εξασφαλίσουμε αυτό στο σύστημα μας, καλούμε την συνάρτηση εύρεσης γωνιών και ακμών όταν το πλήθος το σημείων ενδιαφέροντος μειωθεί στο μισό του αρχικού. Με την προσέγγιση αυτή παρατηρήθηκε ότι κατά μέσο όρο γίνεται ανανέωση κάθε 2 δευτερόλεπτα το οποίο είναι πολύ ικανοποιητικό, αν λάβουμε υπ' όψη μας μία τυπική ταχύτητα δειγματοληψίας της κάμερας των 25 fps.

5.3.2.b Lucas-Kanade

Εφόσον καθορίσαμε τον τρόπο ορισμού των σημείων ενδιαφέροντος, μπορούμε να προχωρήσουμε στον υπολογισμό της οπτικής ροής. Η Python διαθέτει κατάλληλη συνάρτηση που υλοποιεί τον αλγόριθμο Lucas Kanade και δέχεται τις εξής εισόδους :

- Το προηγούμενο frame $t-1$ σε μορφή grayscale
- Το τρέχον frame t σε μορφή grayscale
- Το διάνυσμα p_0 των σημείων ενδιαφέροντος που ανιχνεύθηκαν με την μέθοδο Shi Tomasi εφαρμοζόμενη στο frame $t-1$. Το διάνυσμα αυτό έχει μήκος N και κάθε στοιχείο του έχει δύο διαστάσεις x και y .
- Ένα σετ παραμέτρων (criteria) που σχετίζονται με το μέγεθος της γειτονιάς γύρω από το εξεταζόμενο pixel, το πλήθος των χρησιμοποιούμενων πυραμίδων καθώς και τα κριτήρια τερματισμού του επαναληπτικού αλγορίθμου αναζήτησης.

Ως έξοδο επιστρέφει τρία διανύσματα :

- a) Το νέο διάνυσμα $p1$ των σημείων ενδιαφέροντος που εντοπίστηκαν στο τρέχον frame t και το οποίο θα έχει ίδιες διαστάσεις με το $p0$.
- b) Ένα διάνυσμα κατάστασης (status) που υποδηλώνει για κάθε σημείο ενδιαφέροντος αν η ροή υπολογίστηκε επιτυχώς.
- c) Ένα διάνυσμα (error) με το σφάλμα υπολογισμού για κάθε σημείο.

Η παραπάνω διαδικασία γίνεται για όλα τα frames του video stream της κάμερας και υπολογίζεται κάθε φορά η οπτική ροή για ένα ζεύγος διαδοχικών εικόνων. Επίσης δεν ξεχνάμε εντός του βρόχου επανάληψης να συμπεριλάβουμε την ανανέωση των σημείων ενδιαφέροντος, ώστε να συμβαδίζουν με την συνεχή εναλλαγή των frames. Στην συνέχεια παραθέτουμε μερικά αποτελέσματα εκτίμησης της οπτικής ροής χρησιμοποιώντας την παραπάνω προσέγγιση (Εικόνα 41). Παρατηρούμε ότι συμπεριφέρεται ικανοποιητικά καθώς ανιχνεύει την πλειοψηφία των αντικειμένων που παρουσιάζουν σχετική κίνηση ως προς το ego-vehicle, ενώ δεν επηρεάζεται ιδιαίτερα από τις σκιές.



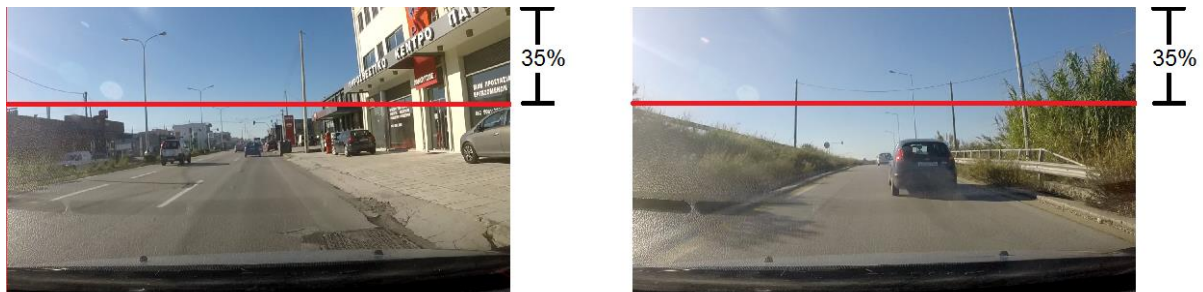
Εικόνα 41. Εφαρμογή του Lucas Kanade για την εύρεση της οπτικής ροής σε δύο βίντεο από δρόμους της Πάτρας. Σε κάθε σειρά υπάρχουν frames του ίδιου βίντεο τα οποία έχουν ληφθεί με μικρή χρονική διαφορά.

5.3.3 Τοποθέτηση των seeds

Το πρόβλημα που έχουμε είναι η τμηματοποίηση της εικόνας σε δύο κλάσεις, δρόμος ή μη-δρόμος. Ο αλγόριθμος ταξινόμησης Random Walker είναι supervised δηλαδή χρειάζεται να του δοθούν στην είσοδο σημεία (seeds), με ήδη γνωστή ταυτότητα δηλαδή να ανήκουν σε κάποια κλάση. Για την εξασφάλιση ενός αξιόπιστου αποτελέσματος πρέπει αφενός να τοποθετηθούν όσο το δυνατόν περισσότερα seeds και αφετέρου να είναι ορισμένα με σχετικά μεγάλη ακρίβεια.

5.3.3.a Non road seeds

Εκμεταλλευόμαστε την θέση της κάμερας στο παρμπρίζ και το γνωστό οπτικό της πεδίο για να προσδιορίσουμε την εκ των προτέρων (a-priori) πληροφορία για την θέση του ουρανού. Σε ένα τυπικό σενάριο οδήγησης η πιθανή θέση του δρόμου είναι ακριβώς μπροστά από το ego-vehicle και εκτείνεται από χαμηλά έως την μέση του frame. Άρα η περιοχή στο πάνω μέρος δεν εξετάζεται και μπορούμε με ασφάλεια να πούμε ότι τα pixels στις θέσεις αυτές θα οριστούν ως μη-δρόμος. Μελετώντας πολλά στιγμιότυπα καταλήξαμε ότι ο ουρανός καταλαμβάνει περίπου το 35% του ύψους του frame (Εικόνα 42).



Εικόνα 42. Περιοχή ουρανού για δύο τυπικά frames από δρόμους της Πάτρας.

Σημειώνουμε ότι η πληροφορία που λαμβάνουμε είναι στατική, δηλαδή ίδια για κάθε frame και εξαρτάται αποκλειστικά από την αρχική εγκατάσταση της κάμερας. Σε περίπτωση που τοποθετηθεί σε άλλο όχημα ή γίνει αλλαγή εξοπλισμού τότε πρέπει να επανεξετάσουμε την θέση του ουρανού με βάση το νέο οπτικό πεδίο.

Στα προαναφερθέντα seeds μη-δρόμου πρέπει να συμπεριλάβουμε το περιβάλλον που βρίσκεται στα πλάγια του frame (πεζοδρόμια δέντρα, κτίρια) καθώς και τα προπορευόμενα οχήματα. Τα αντικείμενα αυτά θα έχουν σχετική κίνηση ως προς το ego-vehicle με αποτέλεσμα να εμφανίζουν υψηλή οπτική ροή. Σε αντίθεση με πριν, εδώ η πληροφορία είναι δυναμική και προκύπτει από την εξέταση δύο διαδοχικών frames. Ο αλγόριθμος Lucas – Kanade θα επιστρέφει κάθε φορά τις νέες θέσεις των pixels με υψηλή οπτική ροή, τα οποία και αυτομάτως θα ορίζονται ως seeds μη-δρόμου. Στην Εικόνα 43 έχουμε δύο διαδοχικά RGB frames (a) και (c) ενώ στα (b) και (d) αντίστοιχα με κόκκινο χρώμα σημειώνονται τα seeds μη-δρόμου. Αξίζει να προσέξουμε την μεγάλη ποσότητα πληροφορίας που

λαμβάνουμε, αν γνωρίζουμε εκ των προτέρων το οπτικό πεδίο της κάμερας. Με την προσέγγιση αυτή (Εικόνα 43 (b), (d)) καταφέρνουμε να περιορίσουμε την περιοχή αναζήτησης, μειώνοντας αυτομάτως το υπολογιστικό κόστος του συστήματός μας.

Παρόλο που ακόμη βρισκόμαστε σε πρώιμο στάδιο, έχουμε ήδη λάβει αρκετή πληροφορία καθώς με μία απλή επισκόπηση της εικόνας μπορούμε να εκτιμήσουμε το σχήμα του δρόμου όπου βρίσκεται στο κέντρο του frame, είναι μαύρου χρώματος και περιβάλλεται από κόκκινα σημεία. Ο κώδικας για την τοποθέτηση των non road seeds φαίνεται στο Παράρτημα 7.6.



(a)



(b)



(c)



(d)

Εικόνα 43. 1η στήλη (a,c) έχουμε 2 διαδοχικά frames από δρόμους της Πάτρας και στην 2η στήλη (b,d) τα αντιστοιχα seeds δρόμου.

5.3.3.b Road seeds

Για την σωστή καθοδήγηση του αλγορίθμου Random Walker πρέπει εκτός των seeds μη-δρόμου να ορίσουμε και εκείνα που θα ανήκουν στο δρόμο. Γνωρίζουμε ότι το οδόστρωμα ενός καλοσυντηρημένου δρόμου παρουσιάζει ομαλή επιφάνεια και λεία υφή. Αναμένουμε λοιπόν ότι η οπτική ροή σε αυτή την περιοχή της εικόνας θα είναι χαμηλή και κατ' επέκταση τα αντίστοιχα pixels θα κατηγοριοποιηθούν ως σπόροι δρόμου. Με την κάμερα τοποθετημένη ψηλά στο κέντρο του παρμπρίζ η περιοχή ενδιαφέροντος θα έχει σχήμα τραπέζιου (Εικόνα 44).



Εικόνα 44. Περιοχή ενδιαφέροντος σχήματος τραπεζίου για τοποθέτηση των seeds δρόμου.

Αν εφαρμόσουμε έναν καθολικό κανόνα σχηματισμού του τραπεζίου, το σύστημα μας δεν θα είναι αξιόπιστο καθώς το σκηνικό μεταβάλλεται συνεχώς και ο δρόμος αλλάζει σχήμα (στροφές, διασταυρώσεις κ.α). Επομένως είναι απαραίτητο αυτή η διαδικασία να γίνει δυναμικά και γι' αυτό θα βασιστούμε στο αποτέλεσμα της προηγούμενης ανίχνευσης. Η μέθοδος που προτίνεται στο [43] – στην οποία θα βασιστούμε - έχει το πλεονέκτημα ότι προσαρμόζεται στο σχήμα του δρόμου ενώ λειτουργεί σωστά και σε στροφές. Για την εξαγωγή των παραμέτρων του τραπεζίου χρησιμοποιείται η δυαδική εικόνα που προέκυψε από την αμέσως προηγούμενη αναγνώριση δρόμου για τα frames $t-2$ και $t-1$ (Εικόνα 45). Η εικόνα αυτή έχει τιμή 0 (μαύρο χρώμα) στα pixels μη-δρόμου και τιμή 255 (άσπρο χρώμα) στα pixels δρόμου.



Εικόνα 45. Δυαδική εικόνα με άσπρο στις περιοχές δρόμου και μαύρο στις περιοχές μη-δρόμου ως αποτέλεσμα του αλγορίθμου τμηματοποίησης.

Επίσης παρατηρώντας ότι στο οπτικό πεδίο της κάμερας παραμβάλλεται το καπό του ego-vehicle (Εικόνα 46), ορίζουμε ότι το 12% των γραμμών στο κάτω μέρος του frame δεν θα λαμβάνονται ως δρόμος αλλά σαν μία «ουδέτερη ζώνη». Θα εξηγήσουμε αναλυτικότερα σε επόμενη παράγραφο πως θα διαχειριστούμε αυτά τα σημεία μέσω του Random Walker.



I_{12%}



I_{12%}

Εικόνα 46. Περιοχή του οπτικού πεδίου της κάμερας στην οποία περιεμβάλλεται το καπό του αυτοκινήτου

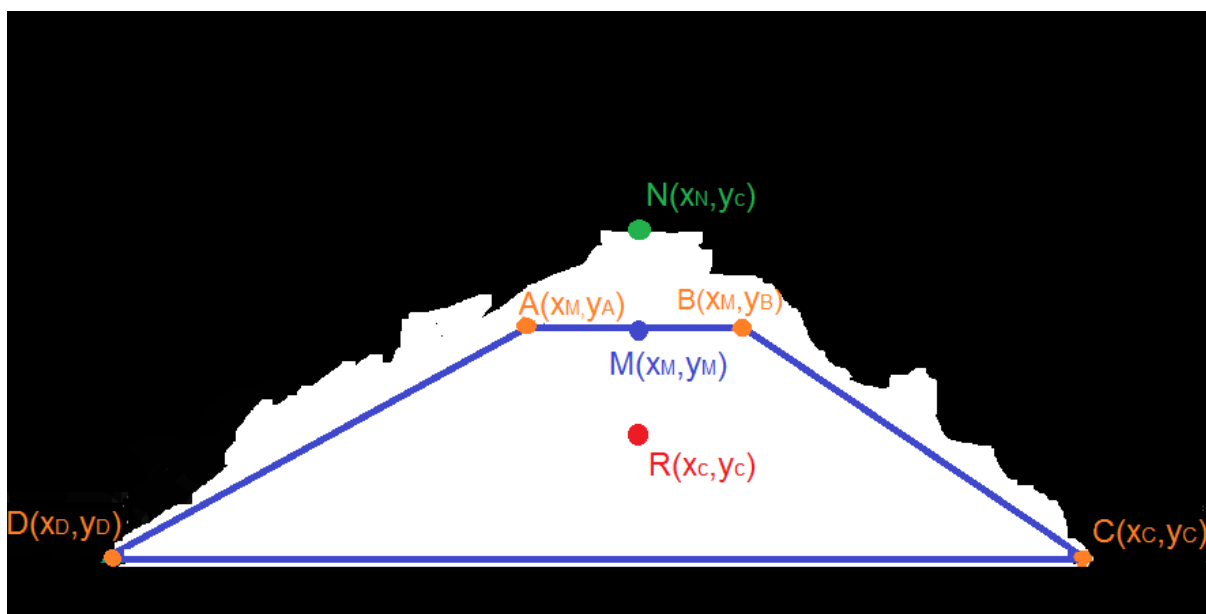
Για τον προσδιορισμό της περιοχής ενδιαφέροντος του δρόμου εργαζόμαστε πάνω στο αποτέλεσμα προηγούμενης ανίχνευσης (Εικόνα 45) και ακολουθούμε τα εξής βήματα :

- a) Υπολογίζουμε το κέντρο της περιοχής $R(x_c, y_c)$ με απλό μέσο όρο των συντεταγμένων των σημείων άσπρου χρώματος.
- b) Συμβολίζουμε ως $N(x_N, y_c)$ το σημείο δρόμου που βρίσκεται ψηλότερα στο frame και είναι κάθετα στο κέντρο R. Το N θα έχει το μικρότερο x αλλά ίδιο y με το κέντρο.
- c) Τοποθετούμε την κάτω βάση CD του τραπεζίου ώστε να συμπίπτει με την γραμμή που βρίσκεται στο 88% του ύψους (εφόσον αφαιρούμε την περιοχή του καπό 12%) και υπολογίζουμε το πλάτος της εξετάζοντας 4 περιπτώσεις :
 1. Αν το πρώτο και το τελευταίο pixel της γραμμής έχουν τιμή 255 τότε ο προηγούμενος δρόμος είναι εκτός ορίων του frame και τότε ορίζουμε την βάση CD ίση με το πλάτος του frame.
 2. Αν το πρώτο pixel της γραμμής είναι 0 και το τελευταίο 255 τότε ο δρόμος εξέρχει του frame μόνο από τα δεξιά. Ορίζουμε λοιπόν το σημείο C ως το πρώτο που θα έχει τιμή 255 πηγαίνοντας από αριστερά προς δεξιά και το σημείο D το τελευταίο pixel της γραμμής.
 3. Αν το πρώτο pixel της γραμμής είναι 255 και το τελευταίο 0 τότε ο δρόμος εξέρχει του frame μόνο από τα αριστερά. Ορίζουμε τότε το σημείο C ως το πρώτο της pixel της γραμμής και το σημείο D ως το πρώτο pixel με τιμή 0 πηγαίνοντας από αριστερά προς τα δεξιά.
 4. Αν το πρώτο και το τελευταίο pixel της γραμμής είναι 0 τότε το πλάτος του προηγούμενου δρόμου είναι μικρότερο από το πλάτος του frame. Ξεκινώντας

από τα αριστερά προς δεξιά ορίζουμε το C ως το πρώτο pixel με τιμή 255 και το D ως το τελευταίο pixel με τιμή 255.

- d) Προσδιορίζουμε το ύψος του τραπεζίου ώστε η άνω βάση AB να ισαπέχει από τα σημεία R και N που είναι γνωστά και αυτόμάτως προκύπτει το μέσο $M(x_M, y_M)$ της AB το οποίο είναι στην ίδια κάθετη διεύθυνση με τα R και N (ίδιο y).
- e) Ορίζουμε το μήκος της AB να είναι ποσοστιαίο της απόστασης RN και κατόπιν δοκιμών καταλήξαμε στο 80%.

Παρακάτω στην Εικόνα 47 παρουσιάζουμε ένα παράδειγμα ορισμού των σημείων του τραπεζίου σε υποθετική δυαδική εικόνα που προέκυψε από αμέσως προηγούμενη εκτέλεση του αλγόριθμου. Ο αντίστοιχος κώδικας που δημιουργεί το τραπέζιο με βάση την προηγούμενο αποτέλεσμα φαίνεται στο Παράρτημα 7.5. Παρατηρούμε ότι τα σημεία C και D της βάσης του τραπεζίου δεν φθάνουν έως το μέγιστο ύψος του frame καθώς λαμβάνουμε υπ' όψη το 12% που καταλαμβάνει το καπό του ego-vehicle. Το σημαντικό πλεονέκτημα αυτής της τεχνικής είναι ότι η τοποθέτηση των seeds καθίσταται πλήρως δυναμική και εξαρτώμενη από το την προηγούμενη χρονική στιγμή. Επίσης η περιοχή ενδιαφέροντος προσαρμόζεται εύκολα στο σχήμα του δρόμου και το πλάτος του.



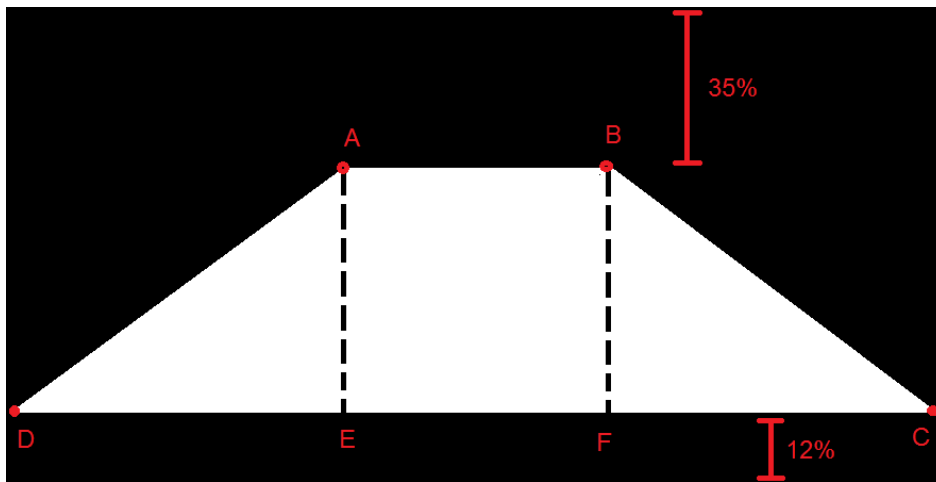
Εικόνα 47. Ορισμός τραπεζίου με βάση το προηγούμενο αποτέλεσμα δρόμου.

Κατά την πρώτη εκτέλεση του αλγορίθμου προφανώς δεν διαθέτουμε προηγούμενο αποτέλεσμα οπότε αναγκαστικά θα δημιουργήσουμε μία ιδανική μορφή δρόμου για να ξεκινήσουμε. Από την τοποθέτηση της κάμερας και γνωρίζοντας το οπτικό της πεδίο παρατηρούμε ότι ο δρόμος είναι τραπεζοειδούς μορφής. Επίσης αν θεωρήσουμε ότι το όχημα είναι εν κινήσει, βρίσκεται στο κέντρο του οδοστρώματος και δεν υπάρχουν μπροστινά εμπόδια τότε μπορούμε να κατασκευάσουμε την απλουστερη μορφή, δηλαδή ένα τέλειο τραπέζιο. Στην συνέχεια μέσω αυτό του τεχνητού «αποτελέσματος» -θεωρητικά για στιγμές $t = -1$ & $t = 0$ - θα λάβουμε ένα νέο τραπέζιο το οποίο θα δοθεί ως περιοχή seeds δρόμου στην πρώτη εκτέλεση του αλγορίθμου.

Τα βήματα που ακολουθούμε για την κατασκευή του 1^{ου} τραpezίου είναι τα εξής:

1. Δημιουργούμε μία μονοχρωματική εικόνα μεγέθους ίσο με το υποδειγματοληπτούμενο frame και τιμή 0 (μαύρο) σε όλα τα pixels.
2. Τοποθετούμε την κάτω βάση στο 88% του ύψους του frame λαμβάνοντας και εδώ ότι το 12% δεσμεύεται από το καπό του αυτοκινήτου. Το μήκος της επιλέγεται ίσο με το πλάτος του frame.
3. Τοποθετούμε την άνω βάση στο 35% του ύψους του frame και κατόπιν πολλών δοκιμών ορίζουμε το μήκος της ίσο με $\frac{1}{4}$ της κάτω βάσης.
4. Ως περιοχή δρόμου και με τιμή 255 (άσπρο) ορίζονται :
 - Όλα τα σημεία που είναι βρίσκονται μεταξύ των 2 βάσεων, δηλαδή στο παραλληλόγραμμο ABFE (Εικόνα 48).
 - Τα σημεία που βρίσκονται εντός των τριγώνων ADE και BCF (Εικόνα 48).

Βασιζόμενοι σε αυτό το ιδανικό σχήμα δρόμου μπορούμε να εκκινήσουμε την επαναληπτική διαδικασία (Εικόνα 48). Η θεώρηση αυτή δεν παραβιάζει την γενικότητα του συστήματος καθώς με την πάροδο του χρόνου η μορφή του τραpezίου μεταβάλλεται ακολουθώντας το μοτίβο του δρόμου.



Εικόνα 48. Κατασκευή της πρώτης δυαδικής εικόνας δρόμου δια χειρός.

Εφόσον ορίσαμε πλήρως πως δημιουργείται η περιοχή ενδιαφέροντος στην συνέχεια τοποθετούμε τα seeds δρόμου. Ελέγχουμε κάθε pixel στο εσωτερικό του τραpezίου και αν είναι ήδη καταχωρημένο ως seed μη-δρόμου τότε το αφήνουμε ως έχει, διαφορετικά το ορίζουμε ως seed δρόμου. Με αυτό τον τρόπο τυχόν περιοχές που έχουν υψηλή οπτική ροή και βρίσκονται εντός της περιοχής του τραpezίου (π.χ προπορευόμενο όχημα) δεν αλλοιώνονται και κατά την τμηματοποίηση θα αναγνωρισθούν επιτυχώς ως εμπόδια. Συνολικά στο Παράρτημα 7.7 παραθέτουμε τον κώδικα που αναπτύξαμε για την πλήρη τοποθέτηση των seeds δρόμου.

5.3.4 Segmentation - Random Walker

Έχοντας προσδιορίσει πλήρως τις περιοχές που θα τοποθετήσουμε τα seeds δρόμου και μη δρόμου, καλούμαστε να βρούμε έναν τρόπο να τα τροφοδοτήσουμε στον αλγόριθμο τμηματοποίησης Random Walker. Αρχικά δημιουργούμε μία μονοχρωματική εικόνα εν ονόματι “seeds”, ίδιων διαστάσεων με το υποδειγματοληπτημένο frame εισόδου και θέτουμε την τιμή όλων των pixels της ίση με 0. Στη συνέχεια σε κάθε κατηγορία των seeds θα αναθέσουμε μία τιμή από 0 έως 255 και θα κατασκευάσουμε τα labels με τα οποία θα κατευθύνουμε τον αλγόριθμο.

Οι ενέργειες που εκτελούνται είναι οι εξής :

- Συνδυάζουμε την εκ των προτέρων πληροφορία της θέσης του ουρανού μαζί με την πληροφορία της οπτικής ροής και λαμβάνουμε τα seeds μη δρόμου, τα οποία θα έχουν τιμή 255.

- Τα seeds δρόμου τοποθετούνται εντός της περιοχής του τραπεζίου που προέκυψε από το προηγούμενο αποτέλεσμα και την τιμή αυτών θέτουμε ίση με 100.
- Τα pixels που απαρτίζουν το 12% των κάτω γραμμών του frame χαρακτηρίζονται ως ουδέτερη ζώνη και θα έχουν τιμή 50.

Παρατήρηση

Η περιοχή του καπό στο κάτω μέρος του frame απαιτεί διαφορετική διαχείριση καθώς ανήκει στο ίδιο το όχημα μας. Από την μία πλευρά αν την λάβουμε ως δρόμο θα είμαστε λάθος εξ' ορισμού. Από την άλλη αν θεωρηθεί ως μη δρόμος τότε θα συγκεντρωθεί μεγάλη ομάδα non road seeds στην περιοχή αυτή με αποτέλεσμα να επηρεάζεται η τμηματοποίηση. Για παράδειγμα τα unlabeled pixels χαμηλά στα πλάγια και εκτός του τραπεζίου θα καταχωρούνται ως μη δρόμος ενώ ανήκουν στον δρόμο.

Συνοψίζοντας, μέχρι στιγμής έχουμε κατασκευάσει μία μονοχρωματική εικόνα τα pixels της οποίας έχουν 3 δυνατές τιμές 50, 100, 255 για κάθε μία κατηγορία των seeds. Πρωτού την δώσουμε ως είσοδο στον Random Walker πρέπει να την μετατρέψουμε σε πιο διαχειρίσιμη μορφή. Αυτό το πετυχαίνουμε δημιουργώντας μία νέα εικόνα “markers” και κάθε τιμή της (i,j) προκύπτει ως εξής :

- Αν $seeds[i,j] = 50$ τότε $markers[i,j] = -1$
- Αν $seeds[i,j] = 100$ τότε $markers[i,j] = 1$
- Αν $seeds[i,j] = 255$ τότε $markers[i,j] = 2$
- Αν το αντίστοιχο pixel είναι unlabeled παραμένει ως έχει δηλαδή 0

Παρακάτω παρουσιάζουμε μία πιθανή μορφή των markers που δίνονται ως είσοδος στον Random Walker για την εκκίνηση της διαδικασίας τμηματοποίησης και την εξαγωγή του δρόμου (Εικόνα 49).



Εικόνα 49. Παράδειγμα των seeds που δίνονται στον Random Walker. Με κίτρινο χρώμα μη δρόμου, με ροζ χρώμα δρόμου, με μωβ χρώμα η ουδέτερη ζώνη και μαύρο τα unlabeled pixels.

Η βιβλιοθήκη Scikit-image της python διαθέτει συνάρτηση που υλοποιεί τον Random Walker η οποία δέχεται ως είσοδο τον πίνακα markers που ορίσαμε παραπάνω καθώς και μία παράμετρο β , η οποία ορίζεται από τον χρήστη και δείχνει τον βαθμό ελευθερίας κατά την μετάβαση του περπατητή σε γειτονικά pixels.

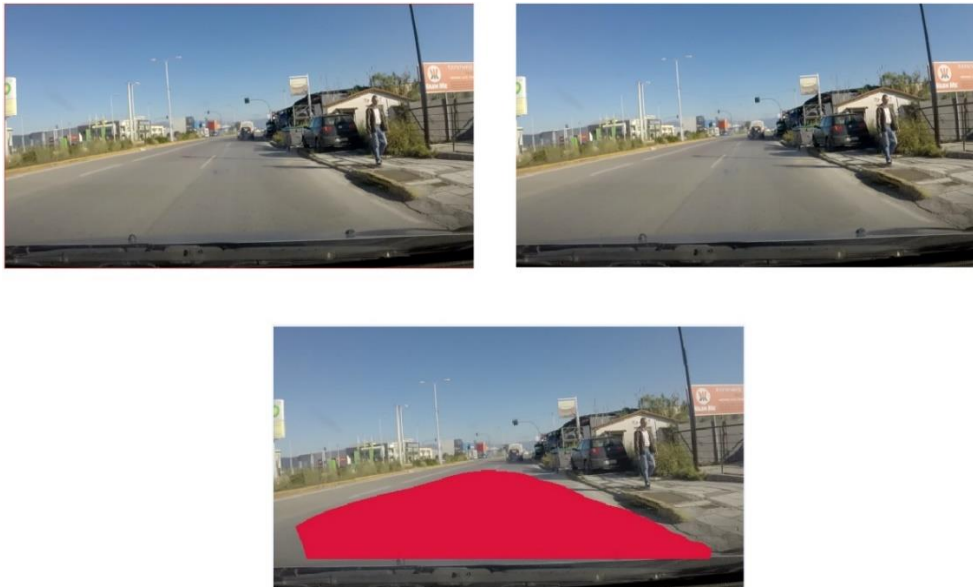
5.4 Πειραματικά αποτελέσματα

Για την αξιολόγηση του συστήματος αναγνώρισης δρόμου χρησιμοποιούνται 2 σετ δεδομένων:

- i. Το dataset Patras [κεφ. 2.1.1] για ποιοτική αξιολόγηση, το οποίο περιέχει έναν συνδυασμό τυπικών και απαιτητικών σεναρίων οδήγησης σε δρόμους της Πάτρας, την ημέρα και με καλές καιρικές συνθήκες.
- ii. Το dataset DIPLODOC [2.1.2] για ποσοτική αξιολόγηση, το οποίο αποτελείται από 5 βίντεο σε διαφορετικές τοποθεσίες του Trento συνοδευόμενα από βάση αληθείας με τα σημεία που οριοθετούν τον δρόμο και τα εμπόδια για κάθε frame. Η οδήγηση γίνεται σε αίθριο καιρό τις πρωινές ώρες της ημέρας.

5.4.1 Ποιοτική αξιολόγηση

Ο αλγόριθμος που κατασκευάσαμε επεξεργάζεται τα frames ανά ζεύγη και ως αποτέλεσμα της τμηματοποίησης, λαμβάνεται η αρχική εικόνα RGB χρωματισμένη με κόκκινο στην περιοχή καθαρού-ορατού δρόμου (Εικόνα 50). Κατά την επεξεργασία, κάθε νέο αποτέλεσμα που παράγεται προστίθεται σε αντίστοιχο αρχείο εικόνων ώστε να είναι διαθέσιμο για περαιτέρω εξέταση.



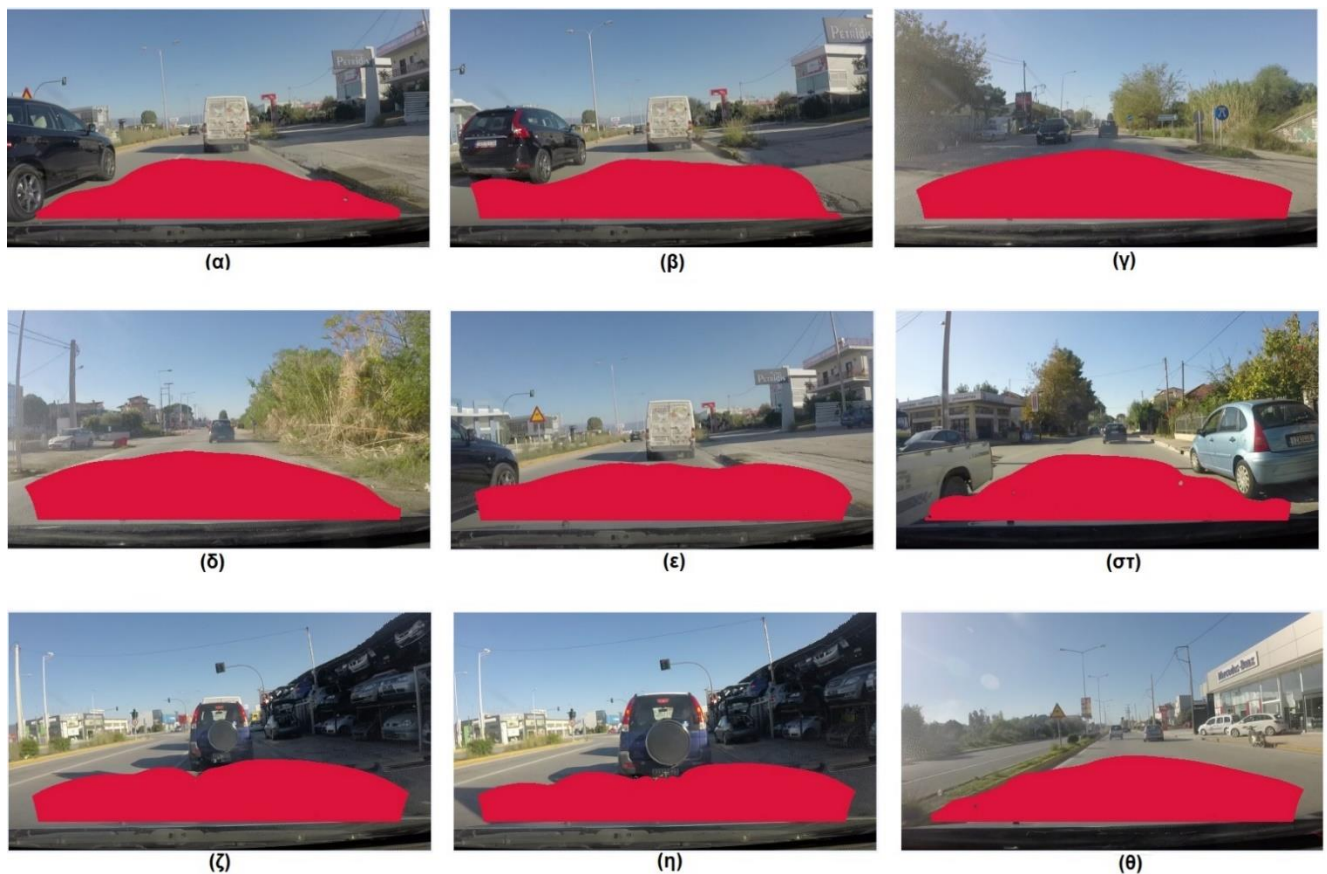
Εικόνα 50. Στην πρώτη γραμμή φαίνονται 2 διαδοχικά frames και στην δεύτερη γραμμή το αποτέλεσμα αναγνώρισης δρόμου με κόκκινο χρώμα.

Η ποιοτική αξιολόγηση προβλέπει την αρχική επισκόπηση των αποτελεσμάτων και τον χαρακτηρισμό της απόδοσης του συστήματος βάσει της εμπειρίας μας και των προδιαγραφών που έχουμε θέσει. Στην συνέχεια, ανάλογα με την συμπεριφορά του μπορούμε να προχωρήσουμε σε τροποποιήσεις του κώδικα ή της λογικής για διόρθωση των αδυναμιών που συναντώνται. Το σετ που χρησιμοποιούμε περιέχει πολλά διαφορετικά σενάρια που είναι πολύ πιθανό να εμφανιστούν, όπως πυκνή κίνηση, κακοσυντηρημένοι δρόμοι χωρίς διαγράμμιση, σημάδια στο οδόστρωμα, υψηλή ηλιοφάνεια καθώς και διακοπή της ροής του δρόμου προς διεξαγωγή έργων. Μελετώντας όλα τα παραπάνω, μπορούμε να κατασκευάσουμε ένα αξιόπιστο σύστημα ικανό να ανταπεξέλθει στις περισσότερες αντίξοες συνθήκες.

Παρατηρώντας προσεκτικά τα αποτελέσματα για τα 15 βίντεο του dataset «Patras» καταλήγουμε στα εξής συμπεράσματα:

- 1) Όταν ένα **όχημα κινείται** παράλληλα στο ego-vehicle, αυτό αναγνωρίζεται ως αντικείμενο υψηλής οπτικής ροής, με αποτέλεσμα η περιοχή δρόμου να σχηματίζεται περιμετρικά του οχήματος χωρίς να το επικαλύπτει (Εικόνα 51 (α), (β)). Παρόμοια συμπεριφορά παρατηρήθηκε και για **σταθμευμένα οχήματα** στα άκρα του δρόμου, με την διαφορά ότι σε αυτή την περίπτωση η ροή ωφείλεται αποκλειστικά στην κίνηση του ego-vehicle (Εικόνα 51 (στ)).
- 2) Οι **σκιές** και τα σημάδια του δρόμου δεν αποπροσανατολίζουν τον αλγόριθμο αλλά σε ορισμένες περιπτώσεις αλλοιώνουν στιγμιαία το αποτέλεσμα. Όταν πάνε να υφίστανται, τότε ο αλγόριθμος επανέρχεται γρήγορα σε ομαλή λειτουργία (Εικόνα 51 (γ)).
- 3) Η εξάρτηση της περιοχής ενδιαφέροντος του δρόμου από το προηγούμενο αποτέλεσμα, προσδίδει ευελιξία στο σύστημα και του επιτρέπει να προσαρμόζεται στις αλλαγές. Με τον τρόπο αυτό το σύστημα ανταποκρίνεται ικανοποιητικά σε **δρόμους με καμπυλότητα** (Εικόνα 52 (i)-(iii)) όπως και σε περίπτωση που το ego-vehicle πλησιάζει πολύ κοντά στο προπορευόμενο όχημα (Εικόνα 51 (ζ), (η)).
- 4) Σε συνθήκες **υψηλής ηλιοφάνειας** δημιουργούνται αντανakλάσεις στο παρμπρίζ από το εσωτερικό του αυτοκινήτου, ενώ εμφανίζονται είδωλα του ηλίου σε επίπεδες λείες επιφάνειες του περιβάλλοντος (πχ. σε διαφημιστικές πινακίδες ή στην επιφάνεια μπροστινών οχημάτων). Το σύστημα ωστόσο φαίνεται να μην επηρεάζεται από την μεγάλη έκθεση των εικόνων και να διατηρεί την αξιοπιστία του (Εικόνα 51 (θ)).
- 5) Το περιβάλλον εκατέρωθεν του δρόμου είναι πολύ σημαντικό καθώς οριοθετεί την περιοχή του οδοστρώματος διευκολύνοντας έτσι την διαδικασία της τμηματοποίησης. Αν τα **όρια** αυτά δεν είναι ευδιάκριτα, τότε είναι φυσικό το σύστημα να οδηγηθεί σε ανακριβή αποτελέσματα. Πιο συγκεκριμένα παρατηρήθηκε ότι αν η περιοχή μη-δρόμου που βρίσκεται στο σύνορο παρουσιάζει ασυνέχειες, τότε το σύστημα δεν επηρεάζεται καθώς τα σημεία αυτά θα έχουν υψηλή οπτική ροή (Εικόνα 51 (δ)). Αντίθετα αν η περιοχή μη-δρόμου έχει λεία υφή (π.χ πεζοδρόμιο χωρίς εμπόδια) τότε τα όρια είναι ασθενή και τμήμα του πεζοδρομίου θα κατηγοριοποιείται λανθασμένα ως δρόμος (Εικόνα 51 (ε)).

- 6) Ένα σημαντικό στοιχείο ενός συστήματος αναγνώρισης δρόμου, είναι η ικανότητα να τίθεται σε πλήρη λειτουργία όσο τον δυνατόν γρηγορότερα. Ιδανικά επιθυμούμε η διαδικασία τμηματοποίησης, να ξεκινά από την στιγμή που μόλις αρχίσει να κινείται ώστε να εξασφαλίζεται η μέγιστη δυνατή ασφάλεια. Το σύστημα που προτίνεται στην παρούσα εργασία, τίθεται σε **πλήρη λειτουργία** μόλις από τα πρώτα frames που επεξεργάζονται και προσαρμόζεται άμεσα στο σχήμα του δρόμου (Εικόνα 52 (iv)-(vi)).
- 7) Η **διεξαγωγή έργων** για την συντήρηση ή την κατασκευή ενός νέου τμήματος του δρόμου, αυξάνουν την πολυπλοκότητα της διαδικασίας τμηματοποίησης. Αυτό συμβαίνει διότι το συνεργείο τοποθετεί κώνους, προειδοποιητικές πινακίδες και κορδέλες, που αφενώς αλλοιώνουν την δομή και το σχήμα του οδοστρώματος και αφετέρου εμποδίζουν την ομαλή κυκλοφορία των οχημάτων. Εξετάζοντας την συμπεριφορά του συστήματος σε αντίστοιχο δρόμο της Πάτρας, παρατηρήσαμε ότι κάθε φορά αναγνωρίζεται η μέγιστη δυνατή περιοχή δρόμου χωρίς να επηρεάζεται από τα εμπόδια και τις απότομες αλλαγές κατεύθυνσης (Εικόνα 52 (vii)-(ix)). Η υψηλή αντοχή του συστήματος ακόμη και σε δύσκολες περιπτώσεις οδικού δικτύου ωφείλεται στην εξάρτηση του από το προηγούμενο αποτέλεσμα και την ενσωμάτωση της πληροφορίας της οπτικής ροής.



Εικόνα 51. Αποτελέσματα του συστήματος αναγνώρισης δρόμου για το dataset Patras. (α)&(β) Διερχόμενο όχημα δίπλα από το ego-vehicle, (γ) παρουσία έντονων σκιών στον δρόμο, (δ)&(ε) διαφορετικές περιπτώσεις δισδιάκριτων ορίων μεταξύ δρόμου και μη-δρόμου, (στ) σταθμευμένο όχημα στα άκρα του δρόμου, (ζ)&(η) πολύ μικρή απόσταση από το προπορευόμενο όχημα, (θ) έντονη ηλιοφάνεια.



Εικόνα 52. Αποτελέσματα του συστήματος αναγνώρισης δρόμου για το dataset Patras. (i)-(iii) Δρόμος με στροφή, (iv)-(vi) εκκίνηση οχήματος από στάση με παρκαρισμένο όχημα ακριβώς μπροστά, (vii)-(ix) κυκλοφορία σε δρόμο που είναι υπό κατασκευή και οριοθετείται από κολωνάκια, πινακίδες και διαχωριστικές μπαριέρες.

5.4.2 Ποσοτική αξιολόγηση

Για την ποσοτική αξιολόγηση του συστήματος χρησιμοποιείται η βάση δεδομένων DIPODOC, η οποία αποτελείται από 5 βίντεο (865 frames) που έχουν ληφθεί από στερεοσκοπική κάμερα, σε διαφορετικές περιοχές του Trento της Ιταλίας. Η λήψη έγινε τις πρωινές ώρες με αίθριο καιρό και περιλαμβάνει από επαρχιακό δρόμο μέχρι εθνική οδό καθώς και ποικίλλα, σύνθετα περιβάλλοντα όπως πεζοδρόμια, διαβάσεις, περιοχές στάθμευσης. Όπως αναφέρθηκε και στο [κεφ. 2.1.2] κάθε αριστερή εικόνα συνοδεύεται από ένα αρχείο κειμένου που περιέχει ανά γραμμή τις συντεταγμένες ενός πολυγώνου (δρόμου ή εμποδίου). Αυτές οι συντεταγμένες έχουν ορισθεί χειροκίνητα και θα χρησιμοποιηθούν ως οδηγός για την αξιολόγηση του συστήματος μας. Για να γίνει η διαδικασία περισσότερο κατανοητή, είναι χρήσιμο να οπτικοποιήσουμε την βάση αληθείας, δηλαδή αξιοποιήσουμε την πληροφορία των γωνίων των πολύγωνων και να τα απεικονίσουμε σε μία κενή (black) εικόνα ιδιών διαστάσεων με τα frames της βάσης δεδομένων. Με αυτόν τον τρόπο θα

κατασκευάσουμε 865 εικόνες αναφοράς τις οποίες θα συγκρίνουμε με τα αποτελέσματα του συστήματός μας.

Για την κατασκευή του φακέλου με τις εικόνες αναφοράς -templates- εκτελούνται τα εξής βήματα:

- 1) Διαβάζεται το αρχείο κειμένου ανά γραμμή.
- 2) Διαχωρίζονται οι λέξεις κάθε γραμμής με βάση τον κενό χαρακτήρα και εκείνες που αφορούν συντεταγμένες πολυγώνων μετατρέπονται από αλφαριθμητικά σε πραγματικούς αριθμούς.
- 3) Οι συντεταγμένες μέχρι αυτό το σημείο είναι κανονικοποιημένες στο διάστημα $[0,1]$ οπότε πολλαπλασιάζοντας με τις πραγματικές διαστάσεις του frame (240x320) μετατρέπονται σε ακέραιους αριθμούς.
- 4) Καταγράφουμε την σειρά εμφάνισης της κατηγορίας των πολυγώνων. Για παράδειγμα 1^η γραμμή εμπόδιο, 2^η δρόμος κοκ. Η διάκριση αυτή γίνεται στην αρχή κάθε γραμμής με τις λέξεις «road» για τον δρόμο και «occl» για εμπόδιο.
- 5) Ανάλογα με την σειρά τοποθέτησης τους στο αρχείο, διακρίνουμε περιπτώσεις ώστε σε κάθε περίπτωση να ζωγραφίσουμε πρώτα το πολύγωνο ιδανικού δρόμου με κόκκινο χρώμα και στην συνέχεια όλα τα εμπόδια με μπλέ χρώμα.
- 6) Αποθηκεύουμε την τελική εικόνα σε φάκελο.

Τα βήματα 1-6 επαναλαμβάνονται για καθένα από τα 865 text αρχεία της βάσης DIPLODOC και τελικά προκύπτει ο φάκελος με τις εικόνες αναφοράς (Εικόνα 53) που θα αποτελούν της βάση αληθείας για την αξιολόγηση.



Εικόνα 53. Αριστερά το frame *diplo000400-L* και δεξιά η οπτική αναπαράσταση των πολυγώνων δρόμου (κόκκινο χρώμα) και εμποδίων (μπλε χρώμα).

Έχοντας ολοκληρώσει την παραπάνω διαδικασία μπορούμε να προχωρήσουμε στην δοκιμή του συστήματός μας. Τα δεδομένα εισόδου που διαθέτουμε είναι σταθερές εικόνες, ενώ το προτινόμενο σύστημα δέχεται video stream. Αυτό το ζήτημα λύνεται εύκολα ομαδοποιώντας τα 865 left frames σε 5 κατηγορίες όπως υποδεικνύεται στο [2] και με

κατάλληλη συνάρτηση τα διατάσουμε σε σειρά και κατασκευάζουμε τα 5 βίντεο. Για την εξαγωγή των αποτελεσμάτων ακολουθούμε την διαδικασία που περιγράψαμε στο κεφάλαιο [5.3] με μερικές τροποποιήσεις :

- Δεν λαμβάνουμε υπόψη το 12% των pixels της βάσης του frame ως ουδέτερη ζώνη, αλλά ως τμήμα του δρόμου. Ο λόγος είναι ότι η κάμερα έχει τοποθετηθεί κατά ιδανικό τρόπο ώστε το ταμπλό του ego-vehicle να μην παρεμβάλλεται στο οπτικό πεδίο.
- Τα frames που επεξεργάζονται δεν υποδειγματοποιούνται καθώς οι διαστάσεις τους είναι σχετικά μικρές (320x240) και δεν καθυστερούν την εκτέλεση του αλγορίθμου. Αντίθετα, κατά την ποιοτική αξιολόγηση η υποδειγματοληψία ήταν απαραίτητη για την εξασφάλιση μικρού χρόνου απόκρισης καθώς επεξεργαζόμασταν εικόνες του σετ Patras με διαστάσεις 1080x1920.
- Η περιοχή κόκκινου χρώματος δεν σχεδιάζεται πάνω στην αρχική RGB εικόνα αλλά σε μία μαύρη εικόνα διαστάσεων 320x240 (Εικόνα 54) ώστε να μπορούμε να την συγκρίνουμε με το αντίστοιχο template που κατασκευάσαμε προηγουμένως.

Τα αποτελέσματα του συστήματος μας για τα 865 frames τα αποθηκεύουμε σε φάκελο καθώς θα τα χρειαστούμε στην συνέχεια.



Εικόνα 54. Αριστερά το frame *diplo000400-L* και δεξιά το αποτέλεσμα του προτινόμενου συστήματος, όπου με κόκκινο χρώμα σημειώνεται η περιοχή δρόμου.

Ολοκληρώνοντας επιτυχώς τα παραπάνω βασικά στάδια επεξεργασίας, έχουμε στην διάθεσή μας δύο φακέλους ίδιου μεγέθους όπου στον έναν περιέχονται οι εικόνες αναφοράς (templates) ενώ στον άλλον τα αποτελέσματα. Συγκρίνουμε κάθε αποτέλεσμα A με το αντίστοιχο template G και ορίζουμε τα παρακάτω μεγέθη :

- **TP** : αριθμός των pixels που έχουν κόκκινο χρώμα στο G και στο A .
- **FP** : αριθμός των pixels κόκκινου χρώματος που συναντώνται στο A αλλά όχι στο G .
- **FN** : αριθμός των pixels κόκκινου χρώματος που συναντώνται στο G αλλά όχι στο A .

Εκτελώντας τις συγκρίσεις των 865 frames προκύπτουν τα συνολικά μεγέθη TP,FP,FN και στην συνέχεια μετράμε την απόδοση με τα μεγέθη completeness, correctness, quality όπως έχουν περιγραφεί στο κεφάλαιο [3.2].

Το σύστημα που αναπτύχθηκε στα πλαίσια της παρούσας εργασίας σημείωσε συνολικά τα εξής νούμερα :

TP	FP	FN
13,580,173	448,655	5,154,729

Πίνακας 6. Πλήθος Αληθώς Θετικών (TP), Λανθασμένα Θετικών (FP) και Λανθασμένα Αρνητικών (FN) για το σύστημα μας.

τα οποία αν τοποθετήσουμε στις σχέσεις του κεφ 3.2, θα λάβουμε τα τρία ζητούμενα ποσοστά :

	completeness	correctness	quality
Απόδοση (%)	72.5%	96.8%	70.8%

Πίνακας 7. Ποσοστά πληρότητας (completeness), ορθότητας (correctness) και ποιότητας (quality) του συστήματος.

Η πληρότητα (completeness) εκφράζει το ποσοστό δεδομένων αναφοράς που εξηγούνται από τα δεδομένα που εξάχθηκαν. Η ορθότητα (correctness) – ακρίβεια (precision) εκφράζει το ποσοστό των σωστών αποτελεσμάτων που προέκυψαν και τέλος η ποιότητα (quality) είναι ένας συνδυασμός των δύο μέτρων. Όλα τα παραπάνω παίρνουν μέγιστη τιμή 100%.

Για να χαρακτηρίσουμε την απόδοση του συστήματός μας πρέπει να συγκρίνουμε τα ποσοστά αυτά με παρόμοιες μελέτες. Στο [42] αναπτύσσεται ένα ενιαίο σύστημα για την αναγνώριση δρόμου και εμποδίων βασισμένο στην στερεοσκοπική όραση. Για την αξιολόγηση χρησιμοποιήθηκε η βάση δεδομένων DIPLODOC με τον δείκτη quality για τα 5 βίντεο να κυμαίνεται μεταξύ 85-90%. Όσον αφορά τα μεγέθη πληρότητας και ορθότητας παρατηρήθηκε ότι δεν έπεσαν κάτω του 90%. Στο [43] ακολουθείται λεπτομερής διαδικασία αξιολόγησης για κάθε βίντεο της βάσης και στο τέλος καταγράφεται η μέση τιμή τους. Η ποιότητα σημειώθηκε ίση με 92.8% ενώ η ορθότητα και η πληρότητα στο 95.7% και 96.9% αντίστοιχα.

Παρατηρούμε ότι το σύστημα μας πετυχαίνει πολύ ικανοποιητική ορθότητα (correctness) που ξεπερνά το 96% και είναι συγκρίσιμη με εκείνη των μεθόδων της βιβλιογραφίας. Η πληρότητα (completeness) ωστόσο είναι κοντά στο 73%, δηλαδή περίπου 20 μονάδες χαμηλότερα από τις τυπικές αξιολογήσεις των υπόλοιπων συστημάτων. Στο σημείο αυτό θα προσπαθήσουμε να ερμηνέσουμε για ποιο λόγο υστερεί το σύστημά μας συγκριτικά με τα [42] και [43] της βιβλιογραφίας. Αρχικά παρατηρούμε ότι η χαμηλή πληρότητα ωφείλεται

στο μεγάλο πλήθος των FN, δηλαδή περιοχών που είναι πράγματι δρόμος αλλά δεν ανιχνεύθηκαν επιτυχώς. Επίσης όπως αναφέραμε στο κεφάλαιο 3.7, ο αλγόριθμος Random Walker περιέχει εξ'ορισμού το στοιχείο της τυχειότητας και για να είναι αξιόπιστος πρέπει να δοθούν με όσο το δυνατόν μεγαλύτερη ακρίβεια τα seeds των 2 κατηγοριών. Άρα αν επιλέξουμε μία διαφορετική τεχνική ορισμού των seeds, τότε ίσως θα καθοδηγούταν καλύτερα ο αλγόριθμος και θα χαρακτηρίζονταν περισσότερες περιοχές ως δρόμος, δηλαδή λιγότερα FN.

Μία λύση θα ήταν να προσθέσουμε και την πληροφορία για τα άκρα του δρόμου όπως στο [37] ώστε να προσδιοριστεί καλύτερα η περιοχή ενδιαφέροντος και να τοποθετηθούν περισσότερα seeds.

5.5 Συμπεράσμα

Το ζήτημα της αναγνώρισης του δρόμου είναι μεγίστης σημασίας για την ανάπτυξη ενός ολοκληρωμένου συστήματος υποβοήθησης της οδήγησης. Πολλές από τις προσεγγίσεις που συναντάμε στην βιβλιογραφία βασίζονται στην ανίχνευση των διαγραμμίσεων του οδοστρώματος, γεγονός που αυτομάτως τις καθιστά αναξιόπιστες για κακοσυντηρημένα ή επαρχιακά οδικά δίκτυα. Στην συνέχεια αρκετοί ερευνητές έστρεψαν το ενδιαφέρον τους στην χρωματική πληροφορία, σε συνδυασμό με την εύρεση των ακμών για την τμηματοποίηση του δρόμο σε ένα frame. Τέλος κάποιες μελέτες βασίστηκαν σε στοχαστικές μεθόδους όπως ταξινομητές ή νευρωνικά δίκτυα μέσω των οποίων αποφασίζουν αν μία περιοχή της εικόνας θα χαρακτηριστεί ως δρόμος ή όχι. Το σύστημα που αναπτύχθηκε στα πλαίσια της εργασίας επεξεργάζεται τα frames σε ζεύγη και αξιοποιεί το μέτρο της οπτικής ροής για τον ακριβή προσδιορισμό μερικών pixels – seeds που ανήκουν στις κατηγορίες «δρόμος» ή «μη-δρόμος». Ύστερα με βάση το προηγούμενο αποτέλεσμα δημιουργείται μία περιοχή ενδιαφέροντος όπου τροφοδοτείται στον αλγόριθμο ταξινόμησης Random Walker, μαζί με τα seeds και παράγει την νέα περιοχή του δρόμου.

Η αξιολόγηση έγινε τόσο ποιοτικά όσο και ποσοτικά χρησιμοποιώντας τις βάσεις δεδομένων του κεφαλαίου 2.1. Οι επιδόσεις του συστήματος μας ήταν αρκετά υψηλές σε σημείο που ανταγωνίζεται αντίστοιχες προσεγγίσεις της βιβλιογραφίας. Φυσικά υπάρχουν επιπλέον περιθώρια βελτίωσης αν προσδιοριστεί καλύτερα η περιοχή ενδιαφέροντος ή επιλεγθεί διαφορετικός αλγόριθμος ταξινόμησης.

6. Επίλογος

6.1 Ανασκόπηση του παρόντος συστήματος

Η παρούσα εργασία αποτελεί μία προσπάθεια κατασκευής ενός συστήματος υποβοήθησης της οδήγησης που θα εκτελεί τις εξής λειτουργίες: α) ανίχνευση του προπορευόμενου οχήματος, β) εντοπισμό της περιοχής του δρόμου. Για την κατανόηση του προβλήματος έγινε μία εκτενής μελέτη στις αντίστοιχες προσεγγίσεις που υπάρχουν στην βιβλιογραφία, προσπαθώντας να αντλήσουμε χρήσιμα στοιχεία και να κατανοήσουμε τους περιορισμούς που τίθενται από την φύση του προβλήματος.

Το σκέλος της αναγνώρισης του προπορευόμενου οχήματος βασίζεται στην συμμετρία, την χρωματική πληροφορία και την προβλεπόμενη σχετική απόσταση κάθε ζεύγους φαναριών. Η είσοδος του συστήματος λαμβάνεται από κάμερα ενός φακού και επεξεργάζεται ένα frame κάθε φορά. Αρχικά η εικόνα μετατρέπεται στο κόκκινο υποκανάλι του CIE Lab και σε συνδυασμό με τον αλγόριθμο FRST εντοπίζονται τα κόκκινα αντικείμενα υψηλής συμμετρίας. Έπειτα ο K means χωρίζει τα σημεία σε ομάδες, ορίζει αντιπροσωπευτικά κέντρα και δημιουργούνται τα ζεύγη υποψηφίων φαναριών με κριτήριο την σχετική τους θέση. Τέλος σχηματίζεται η περιοχή ενδιαφέροντος και συγκρίνεται η αριστερή με την δεξιά υποπεριοχή ως προς την μεσοκάθετο. Το σύστημα αξιολογήθηκε υπό διάφορες συνθήκες τόσο σε αστικό όσο και υπαίθριο περιβάλλον. Τα ποσοστά ακρίβειας και πληρότητας που πετυχαίνει είναι αρκετά ικανοποιητικά και στις δύο περιπτώσεις, με την μόνη αδυναμία να παρουσιάζεται σε πολύ πυκνό αστικό περιβάλλον λόγω του εκτεταμένου θορύβου (διαφημιστικές πινακίδες, φανάρια, διαγραμμίσεις, άλλα οχήματα).

Στο 2^ο σκέλος της τμηματοποίησης του δρόμου η προσέγγιση που κάναμε λαμβάνει πάλι είσοδο από κάμερα ενός φακού αλλά εδώ τα frames επεξεργάζονται σε ζεύγη ώστε να λάβουμε την πληροφορία της οπτικής ροής. Η διαδικασία της κατηγοριοποίησης κάθε pixel της εικόνας σε δρόμο ή μη δρόμο, γίνεται με την βοήθεια του αλγορίθμου Random Walker ο στον οποίο τροφοδοτούμε ήδη γνωστά σημεία που προκύπτουν από την a-priori πληροφορία σχετικά με την θέση της κάμερας, το προηγούμενο αποτέλεσμα καθώς και την οπτική ροή.

Οι επιδόσεις του συστήματος μας ήταν αρκετά υψηλές σε σημείο που ανταγωνίζεται αντίστοιχες προσεγγίσεις της βιβλιογραφίας.

6.2 Μελλοντικές μελέτες

Όσον αφορά το σύστημα που αναπτύξαμε, η πιο χρήσιμη εξέλιξή του θα ήταν αφενώς να ενοποιηθούν οι 2 υποεργασίες (αναγνώριση οχήματος και δρόμου) υπό ένα σύστημα και αφετέρου να είναι σε θέση να λειτουργήσει σε πραγματικό χρόνο. Αυτό θα μπορούσε να επιτευχθεί γράφοντας εκ νέου τον κώδικα, αλλά αυτή τη φορά με χρήση τεχνικών παράλληλου προγραμματισμού οι οποίες μειώνουν αισθητά τον χρόνο εκτέλεσης ενός προγράμματος. Με τον τρόπο αυτό το λογισμικό θα μπορεί να εγκατασταθεί σε έναν

μικροεπεξεργαστή τύπου Raspberry προσδίδοντας στο σύστημά μας, το χαρακτηριστικό της φορητότητας. Έπειτα με μικρές τροποποιήσεις θα είμαστε σε θέση να το τοποθετήσουμε σε οποιοδήποτε αυτοκίνητο παρέχοντας ένα προηγμένο σύστημα υποβοήθησης της οδήγησης.

Η προσπάθεια που πραγματοποιήθηκε και περιγράφεται σε αυτήν την εργασία εκτείνεται σε διάφορους τομείς της σύγχρονης τεχνολογίας ADAS. Η χρήση πληροφοριών που προέρχονται μόνο από έγχρωμη κάμερα με μόνο έναν φακό, αποτελεί μία πρόκληση ωστόσο προσθέτει επιπλέον αξία στο σύστημα, καθώς όλα αυτά μπορούν να αξιοποιηθούν ως τα θεμέλια πιο εξελιγμένων ADAS που θα συνδυάζουν και επιπλέον πληροφορίες από άλλες πηγές. Για παράδειγμα μία ιδέα είναι να τοποθετηθεί ένα radar στο μπροστινό μέρος του αυτοκινήτου και να μετράει την απόσταση από μπροστινά εμπόδια στον δρόμο. Η πληροφορία της αποστασης θα μπορούσε να ενσωματωθεί στο υποσύστημα αναγνώρισης του προπορευόμενου οχήματος ώστε να βελτιώνει την ακρίβεια της ανίχνευσης καθώς και για να ενημερώνει τον οδηγό πόσο απέχει το κάθε όχημα από το ego vehicle. Επίσης η ακρίβεια του συστήματος μπορεί να αυξηθεί αν εφαρμόζεται αυστηρότερος έλεγχος από την τροχαία για οχήματα που κυκλοφορούν με ξεθωριασμένα πίσω φώτα, σε συνδυασμό με την θέσπιση αυστηρότερων κανόνων για τους κατασκευαστές, ως προς το σχημα και το χρώμα των φαναριών για τα νέα μοντέλα αυτοκινήτων.

Σε επόμενο στάδιο θα μπορούσε κανείς να προσθέσει στο υπάρχον σύστημα, μία τρίτη λειτουργία που αφορά την ανάλυση του προσώπου του οδηγού, με σκοπό να εντοπιστούν σημάδια κόπωσης ή μειωμένης συγκέντρωσης. Αυτό θα επιτευχθεί με μία επιπλέον κάμερα τοποθετημένη στο εσωτερικού του οχήματος και στραμμένη προς την θέση του οδηγού. Το προγραμματιστικό σκέλος του νέου υποσυστήματος δεν είναι περίπλοκο, καθώς για την αναγνώριση του προσώπου θα χρησιμοποιηθούν τα στοιχεία του χρώματος και της συμμετρίας. Αυτό σημαίνει ότι αρκετοί αλγόριθμοι που χρειάζονται, έχουν ήδη μελετηθεί σε αυτή την εργασία. Τέλος, με την ενοποίηση των 3 υποσυστημάτων θα υπήρχε η δυνατότητα όταν ο οδηγός δείχνει καταβεβλημένος να μεταβάλλονται στιγμιαία τα κατώφλια στα υποσυστήματα εντοπισμού προπορευόμενου οχήματος και δρόμου. Σε τέτοιες περιπτώσεις που οι αντιδράσεις του οδηγού είναι μειωμένες, η αύξηση της ευαισθησίας του ADAS κρίνεται αναγκαία ώστε να τον ειδοποιεί πιο γρήγορα και να αποτρέψει ένα ενδεχόμενο ατύχημα.

Στην προσπάθεια ανάπτυξης του υπάρχοντος συστήματος και μελετώντας την βιβλιογραφία καταλήξαμε στο συμπέρασμα ότι δεν υπάρχουν επαρκείς βάσεις δεδομένων για την αξιολόγηση σε κακές καιρικές συνθήκες. Θα μπορούσαν λοιπόν οι ερευνητές να δημιουργήσουν μεγάλα databases με βίντεο σε ακραίες συνθήκες οδήγησης με μειωμένη ορατότητα, ομίχλη ή βροχή. Καλό θα ήταν οι βάσεις αυτές να συνοδεύονταν με πληροφορίες σχετικά με εμπόδια, περιοχές δρόμου, φωτεινούς σηματοδότες ή πινακίδες ώστε να είναι ευκολότερη η ποσοτική αξιολόγηση επόμενων συστημάτων που θα κατασκευαστούν.

7. Παράρτημα

7.1 Συνάρτηση υπολογισμού του σημείου γονάτου

```
# I consider that big slope is when the gradient between 2 consecutive points is above 30%.
# At the beginning the slope is very steep (above 30%) so in each iteration for example points x(i-1) and x(i)
# take as elbow the second point (x(i)).
# The first time that I find a slope less than 30%, break my loop and as the final elbow point is consider
# the last that was assigned in the just previous loop (where slope was greater than 30%).
def elbow(X):
    Sosqd = [] # Array Sum of squared distances
    i = 0
    km = KMeans(n_clusters=1) # calculate for K is 1 to initialise the array Sosqd
    km = km.fit(X)
    Sosqd.append(km.inertia_)
    K = range(2, 10) # check k from 1 to 10 in total
    opt = 1 # Before the entering the loop the optimal k is 1
    for k in K:
        i = i + 1
        km = KMeans(n_clusters=k)
        km = km.fit(X)
        Sosqd.append(km.inertia_)
        if ((Sosqd[i - 1] - Sosqd[i]) / Sosqd[i - 1]) > 0.3:
            opt = k
    return opt
```

7.2 Συνάρτηση εύρεσης των ζευγών υποψήφιων φαναριών.

```
# This function returns 1 if a pair of high symmetrical objects obey these limitations and so is verified as
# a pair of back lights.

import math
def pairs(centers,n,h,w):
    pair = []
    for k in range(0,n):
        if k+1 > (n-1):
            break
        for v in range(k+1,n):
            hmaxi = abs(centers[k][1]-centers[v][1])*0.087 # 0.087 is tan of 5 degrees as siogkas proposed
            if abs(centers[k][0]-centers[v][0]) <= hmaxi and abs(centers[k][1]-centers[v][1]) >= math.floor(w/8) \
            and abs(centers[k][1]-centers[v][1]) <= math.floor(0.6*w):
                pair.append([k,v])
    return pair
# In this function the centers[i][0] is x-coordinate, and centers[i][1] is y-coordinate
```

7.3 Συνάρτηση υπολογισμού του δείκτη ομοιότητας SSIM.

```
from skimage.metrics import structural_similarity as ssim
import math
...
def sim(img,h,w): #takes 3 arguments : the source img, the height h and the width w and returns SSIM index
    if w % 2==0: # If the width is even number then the middle point does not fall to any y
        left = img[0:h,0:int(w/2)]
        right = img[0:h, int(w/2):w]
        s = ssim(left,right)

        # Attention!!! If i solve this problem to the paper i must compare y of left image to y'=w-1-y of the right image
        # but when i define the right image automatically x starts from 0 and goes to h-1 and y from 0 to w/2-1.
        # In this way i just compare x and y of each image

    else: # If the width is odd number then the middle point is falling at y=floor(w/2)
        w2 = math.floor(w/2)
        left = img[0:h, 0:w2]
        right = img[0:h, w2+1:w]
        s = ssim(left,right)

    return s
```

7.4 Συνάρτηση δημιουργίας της περιοχής ενδιαφέροντος (ROI).

```
def drawroi(img,h1,w1,x1,x2,y1,y2):
    # h1,w1 is height and width of img
    # 1st point is (x1,y1)
    # 2nd point is (x2,y2)
    ssimthresh = 0.13
    hmax = math.floor(abs(y1 - y2) / 3) # height of the ROI is defined as 1/3 of the horizontal centers distance
    dy = math.floor((abs(y1 - y2)) / 10) # dy margin that is added at the y1-y2 horizontal distance
    dx = math.floor(abs(x1 - x2)) # vertical distance of the centers
    hmax = math.floor(abs(y1 - y2) / 3)
    # height of the ROI is defined as 1/3 of the horizontal centers distance

    if x1 <= x2 and y1 < y2:
        # print("Entered the 1st loop")
        # The next part of the code prevents the roi to exceed the dimensions of the image h and w
        # print(roi.shape[0], roi.shape[1])
        if x1 - math.floor(hmax / 2) < 0 or x2 + math.floor(hmax / 2) - dx > h1:
            s=0
        elif y1 - dy < 0:
            roi = img[x1 - math.floor(hmax / 2): x2 + math.floor(hmax / 2) - dx, 0: y2 + dy]
        elif y2 + dy > w1:
            roi = img[x1 - math.floor(hmax / 2): x2 + math.floor(hmax / 2) - dx, y1 - dy: w1]
        else:
            roi = img[x1 - math.floor(hmax / 2): x2 + math.floor(hmax / 2) - dx, y1 - dy: y2 + dy]

        groi = cv2.cvtColor(roi, cv2.COLOR_BGR2GRAY)
        # be careful that each ROI is defined as [x1:x2,y1:y2] where x1<x2 and y1<y2
        # print(roi.shape[0],roi.shape[1])
        # a = MAE(groi, roi.shape[0], roi.shape[1])
        s = sim(groi, roi.shape[0], roi.shape[1])
```

```

elif x1 >= x2 and y1 < y2:
    # print("Entered the 2nd loop")
    if x1 - math.floor(hmax / 2) - dx < 0 or x2 + math.floor(hmax / 2) > h1:
        s=0
    elif y1 - dy < 0:
        roi = img[x1 - math.floor(hmax / 2) - dx: x2 + math.floor(hmax / 2), 0: y2 + dy]
    elif y2 + dy > w1:
        roi = img[x1 - math.floor(hmax / 2) - dx: x2 + math.floor(hmax / 2), y1 - dy: w1]
    else:
        roi = img[x1 - math.floor(hmax / 2) - dx: x2 + math.floor(hmax / 2), y1 - dy: y2 + dy]

    # print(roi.shape[0], roi.shape[1])
    groi = cv2.cvtColor(roi, cv2.COLOR_BGR2GRAY)
    # roi = img[x2 - math.floor(hmax / 2) - dx : x1 + math.floor(hmax / 2), y1 - dy: y2 + dy]
    # be careful that each ROI is defined as [x1:x2,y1:y2] where x1<x2 and y1<y2

    # print(roi.shape[0], roi.shape[1])

    # a = MAE(groi, roi.shape[0], roi.shape[1])
    s = sim(groi, roi.shape[0], roi.shape[1])
elif x1 <= x2 and y1 > y2:
    # print("Entered the 3rd loop")
    if x2 - math.floor(hmax / 2) - dx < 0 or x1 + math.floor(hmax / 2) > h1:
        s=0
    elif y2 - dy < 0:
        roi = img[x2 - math.floor(hmax / 2) - dx: x1 + math.floor(hmax / 2), 0: y1 + dy]
    elif y1 + dy > w1:
        roi = img[x2 - math.floor(hmax / 2) - dx: x1 + math.floor(hmax / 2), y2 - dy: w1]
    else:
        roi = img[x2 - dx - math.floor(hmax / 2): x1 + math.floor(hmax / 2), y2 - dy: y1 + dy]

    # print(roi.shape[0], roi.shape[1])
    groi = cv2.cvtColor(roi, cv2.COLOR_BGR2GRAY)
    # be careful that each ROI is defined as [x1:x2,y1:y2] where x1<x2 and y1<y2

    s = sim(groi, roi.shape[0], roi.shape[1])
else:
    # print("Entered the 4th loop")
    if x2 - math.floor(hmax / 2) < 0 or x1 + math.floor(hmax / 2) - dx > h1:
        s=0
    elif y2 - dy < 0:
        roi = img[x2 - math.floor(hmax / 2): x1 + math.floor(hmax / 2) - dx, 0: y1 + dy]
    elif y1 + dy > w1:
        roi = img[x2 - math.floor(hmax / 2): x1 + math.floor(hmax / 2) - dx, y2 - dy: w1]
    else:
        roi = img[x2 - math.floor(hmax / 2): x1 + math.floor(hmax / 2) - dx, y2 - dy: y1 + dy]

    # print(roi.shape[0], roi.shape[1])
    groi = cv2.cvtColor(roi, cv2.COLOR_BGR2GRAY)
    # be careful that each ROI is defined as [x1:x2,y1:y2] where x1<x2 and y1<y2

    s = sim(groi, roi.shape[0], roi.shape[1])

return s

```

7.5 Συνάρτηση δημιουργίας του τραπεζίου για την τοποθέτηση των seeds δρόμου.

```
def trapezoid(img):
    Xn = 0
    h = img.shape[0]
    w = img.shape[1]
    sumx=0
    sumy=0
    count=0
    flag=0
    for i in range(h):
        for j in range(w):
            if img[i,j]==255 :
                sumx = sumx+i
                sumy = sumy+j
                count=count+1
                if flag==0 :
                    Xn=i          # Inside the loops check to find the highest point in the image N(Xn,Yc)
                    flag=1

    Xc = int(sumx/count)
    Yc = int(sumy/count)

# At this part i try to find the width of the lower base of the trapezoid which is placed at the 88% of the height
# taking into account 4 cases:
    b = int((88/100)*h) # indicates the x coordinate of the down base of the trapezoid
    right=0
    left=0
    if img[b-1,0]==255 and img[b-1,w-1]==255:      # 1st case the road limits are out of frame bounds
        left = 0
        right = w-1
    elif img[b-1,0]==255 and img[b-1,w-1]==0 :      # 2nd case: left limit out, right limit in
        left=0
        for j in range(w):
            if img[b-1, j] == 0 :
                right = j
                break
    elif img[b-1,0]==0 and img[b-1,w-1]==255 :      # 3rd case: left limit in, right limit out
        right=w-1
        for j in range(w):
            if img[b-1, j] == 255 :
                left = j
                break
    else:
        count=0
        index=0
        for j in range(w):                          # 4th case: both limits in the frame
            if img[b - 1, j] == 255:
                index = j
                count=count+1
        right = index
        left = right-count

    Xm = Xn+int((Xc-Xn)/2) # So the X coordinate of median of upper base AB is Xn+(Xc-Xn)/2
    Xa = Xm # A and B have the same X with M where M has the same distance from N and R
    Xb = Xm
    AB = 0.8*(Xc-Xn) # Length of upper base AB is proportional to the length RN
    Ya = Yc-int(AB/2) # Ya and Yb have the same distance from the median M of AB
    Yb = Yc+int(AB/2)
    Yd = left
    Yc = right

    return Xa,Xb,Ya,Yb,Yc,Yd
```

7. 6 Δημιουργία των seeds μη – δρόμου.

```
# ----- Create the non road seeds-----

sky_limit = int(35 / 100 * hh)
seeds = np.zeros((hh,ww), dtype=np.uint8)
seeds[0:sky_limit] = 255 # place the sky surface (a-priori)

for i in range(int_good_new.shape[0]): # int_good_new contains the coordinates of points with high flow so 255
    x = int_good_new[i][1]
    y = int_good_new[i][0]
    seeds[x,y] = 255
```

7. 7 Δημιουργία των seeds δρόμου

```
# ----- Create the road seeds-----
# After marking the non road seeds(result of high flow + sky) I mark the road seeds which are inside the
# trapezoid. This trapezoid is defined from the previous result of Random Walker for frames(t-2,t-1).
# previousroad is a binary frame with 0(non road) and 255(road) with smaller dimensions (at scale%)

xa, xb, ya, yb, yc, yd = trapezoid(previousroad) # Get the trapezoid's points(A,B,C,D)
xd = int(0.88 * hh) - 1
xc = int(0.88 * hh) - 1

dash_limit = int((88 / 100) * hh) # Consider that ego-car's dashboard is visible at the bottom and takes
# 12% of total height hh so it starts at 88% of the height and ends at 100% of height
lamda_ad = float((xa - xd) / (ya - yd))
lamda_bc = float((xb - xc) / (yb - yc))

for x in range(dash_limit, hh): # the 12% of the image is not road but the ego-car's dashboard(a-priori info)
    seeds[x, :] = 50 # so i set it equal to 50 in order to eliminate these from road
for x in range(xa, dash_limit):
    for y in range(ya, yb):
        if seeds[x, y] != 255:
            seeds[x, y] = 100
    for y in range(yd, ya):
        if seeds[x, y] != 255:
            if x > (lamda_ad * (y - yd) + xd):
                seeds[x, y] = 100
    for y in range(yb, yc):
        if seeds[x, y] != 255:
            if x > (lamda_bc * (y - yc) + xc):
                seeds[x, y] = 100

#-----
```

8. Βιβλιογραφία

- [1] Computational Vision archive : Discription of databases Caltech 1999 & 2001, California Institute of Technology
- [2] Michele Zanin, Stefano Messelodi, Carla Maria Modena, Fondazione Bruno Kessler, "Diplodoc road stereo sequence", FBK Technical Report #ID 164010, April 2013.
- [3] Guide/Overview, Python.org
- [4] About, OpenCV
- [5] Getting started with Scikit-image, Scikit-image.org
- [6] Machine learning in Python, Scikit-learn.org
- [7] Noor A. Ibraheem, Mokhtar M. Hasan, Rafiqul Z. Khan, Pramod K. Mishra, "Understanding Color Models: A Review", *ARPN Journal of Science and Technology*, Vol. 2, No. 3, April 2012.
- [8] P. Ganesan, V. Rajini and R. I. Rajkumar, "Segmentation and edge detection of color images using CIELAB color space and edge detectors," *INTERACT-2010*, Chennai, 2010, pp. 393-397.
- [9] C. Junhua and L. Jing, "Research on Color Image Classification Based on HSV Color Space," *2012 Second International Conference on Instrumentation, Measurement, Computer, Communication and Control*, Harbin, 2012, pp. 944-947.
- [10] D. Song, W. Sun, Z. Ji, G. Hou, X. Li and L. Liu, "Color model selection for underwater object recognition," *2014 International Conference on Information Science, Electronics and Electrical Engineering*, Sapporo, 2014, pp. 1339-1342.
- [11] S. Na, L. Xumin and G. Yong, "Research on k-means Clustering Algorithm: An Improved k-means Clustering Algorithm," *2010 Third International Symposium on Intelligent Information Technology and Security Informatics*, Jingtangshan, 2010, pp. 63-67.
- [12] Purnima Bholowalia, Arvind Kumar, "EBK-Means: A Clustering Technique based on Elbow Method and K-Means in WSN", *International Journal of Computer Applications* (0975 – 8887) Volume 105 – No. 9, November 2014.

- [13] G. Loy and A. Zelinsky, "Fast radial symmetry for detecting points of interest," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 8, pp. 959-973, Aug. 2003.
- [14] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," in *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62-66, Jan. 1979.
- [15] L. Ziyun and L. Wei, "The Compensated HS Optical Flow Estimation Based on Matching Harris Corner Points," *2010 International Conference on Electrical and Control Engineering*, Wuhan, 2010, pp. 2279-2282.
- [16] Gunnar Farneback, "Two-Frame Motion Estimation Based on Polynomial Expansion", Computer Vision Laboratory - Linkoping University, Sweden.
- [17] "Motion Analysis and Object Tracking", OpenCV documentation.
- [18] Jie Chen, Li-hui Zou, Juan Zhang and Li-hua Dou, "The Comparison and Application of Corner Detection Algorithms", *Journal of Mutlimedia*, Vol. 4, No. 6, Dec. 2009.
- [19] "Harris Corner Detection", OpenCV documentation.
- [20] Z. Mu and Z. Li, "A Novel Shi-Tomasi Corner Detection Algorithm Based on Progressive Probabilistic Hough Transform," *2018 Chinese Automation Congress (CAC)*, Xi'an, China, 2018, pp. 2918-2922.
- [21] "Shi-Tomasi Corner Detector & Good Features to Track", OpenCV documentation.
- [22] L. Y. Siong, S. S. Mokri, A. Hussain, N. Ibrahim and M. M. Mustafa, "Motion detection using Lucas Kanade algorithm and application enhancement," *2009 International Conference on Electrical Engineering and Informatics*, Selangor, 2009, pp. 537-542.
- [23] B. D. Lucas, & Kanade. "An Iterative Image Registration Technique with an Application to Stereo Vision," *DARPA Image Understanding Workshop*, pp 121-130, 1981.
- [24] Prof. Dr. Raul Rojas, "Lucas-Kanade in a Nutshell".
- [25] Jean-Yves Bouguet, "Pyramidal Implementation of the Lucas Kanade Feature Tracker Description of the algorithm", *Intel Corporation Microprocessor Research Labs*.
- [26] L. Grady, "Random Walks for Image Segmentation," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 11, pp. 1768-1783, Nov. 2006.

- [27] Dirichlet : R. Courant and D. Hilbert, "Methods of Math. Physics", vol. 2. John Wiley and Sons, 1989.
- [28] Weijie Wang, Yanmin Lu, "Analysis of the Mean Absolute Error (MAE) and the Root Mean Square Error (RMSE) in Assessing Rounding Model", *5th International Conference on Mechanical Engineering, Materials Science and Civil Engineering* 15–16 December 2017, Kuala Lumpur, Malaysia.
- [29] Zhou Wang, Alan C. Bovik, Hamid R. Sheikh, Eero P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity", *IEEE transactions on image processing*, Vol. 13 No. 4 April 2004.
- [30] Adrian Rosebrock, "Python Compare Two Images", *Pyimagesearch*, 15 Sept. 2014.
- [31] Swati A.Gandhi, C.V.kulkarni, "MSE vs SSIM", *International Journal of Scientific & Engineering Research*, Vol. 4, Issue 7, July 2013.
- [32] D. Gao, J. Duan, B. Zheng and X. Yang, "Preceding Vehicles Detection Based on Vehicle Features," *2008 Second International Symposium on Intelligent Information Technology Application*, Shanghai, 2008, pp. 408-412.
- [33] E. Skodras, G. Siogkas, E. Dermatas and N. Fakotakis, "Rear lights vehicle detection for collision avoidance," *2012 19th International Conference on Systems, Signals and Image Processing (IWSSIP)*, Vienna, 2012, pp. 134-137.
- [34] A. Kanitkar, B. Bharti and U. N. Hivarkar, "Vision based preceding vehicle detection using self shadows and structural edge features," *2011 International Conference on Image Information Processing*, Shimla, 2011, pp. 1-6.
- [35] T. Kato, Y. Ninomiya and I. Masaki, "Preceding vehicle recognition based on learning from sample images," in *IEEE Transactions on Intelligent Transportation Systems*, vol. 3, no. 4, pp. 252-260, Dec. 2002.
- [36] H. Kim and B. Song, "Vehicle recognition based on radar and vision sensor fusion for automatic emergency braking," *2013 13th International Conference on Control, Automation and Systems (ICCAS 2013)*, Gwangju, 2013, pp. 1342-1346.
- [37] C. Y. Low, H. Zamzuri and S. A. Mazlan, "Simple robust road lane detection algorithm," *2014 5th International Conference on Intelligent and Advanced Systems (ICIAS)*, Kuala Lumpur, 2014, pp. 1-4.

- [38] W. Wu and G. ShuFeng, "Research on Unstructured Road Detection Algorithm Based on the Machine Vision," *2009 Asia-Pacific Conference on Information Processing*, Shenzhen, 2009, pp. 112-115.
- [39] J. D. Crisman and C. E. Thorpe, "UNSCARF-a color vision system for the detection of unstructured roads," *Proceedings. 1991 IEEE International Conference on Robotics and Automation*, Sacramento, CA, USA, 1991, pp. 2496-2501 vol.3.
- [40] Yinghua He, Hong Wang and Bo Zhang, "Color-based road detection in urban traffic scenes," in *IEEE Transactions on Intelligent Transportation Systems*, vol. 5, no. 4, pp. 309-318, Dec. 2004.
- [41] S. Beucher and M. Bilodeau, "Road segmentation and obstacle detection by a fast watershed transformation," *Proceedings of the Intelligent Vehicles '94 Symposium*, Paris, France, 1994, pp. 296-301.
- [42] P. Lombardi, M. Zanin and S. Messelodi, "Unified stereovision for ground, road, and obstacle detection," *IEEE Proceedings. Intelligent Vehicles Symposium, 2005.*, Las Vegas, NV, USA, 2005, pp. 783-788.
- [43] G. K. Siogkas and E. S. Dermatas, "Random-Walker Monocular Road Detection in Adverse Conditions Using Automated Spatiotemporal Seed Selection," in *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 2, pp. 527-538, June 2013.
- [44] S. Yadav, S. Patra, C. Arora and S. Banerjee, "Deep CNN with color lines model for unmarked road segmentation," *2017 IEEE International Conference on Image Processing (ICIP)*, Beijing, 2017, pp. 585-589.
- [45] Z. Feng, M. Stolz, M. Li, M. Kunert and W. Wiesbeck, "Verification of a Lane Detection Method with Automotive Radar based on a New Type of Road Marking," *2018 IEEE MTT-S International Conference on Microwaves for Intelligent Mobility (ICMIM)*, Munich, 2018, pp. 1-4.
- [46] Y. Zeng, K. Xiang and D. Li, "Applying behavior recognition in road detection using vehicle sensor networks," *2012 International Conference on Computing, Networking and Communications (ICNC)*, Maui, HI, 2012, pp. 751-755.