



**ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ**  
ΔΙΑΤΜΗΜΑΤΙΚΟ ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ  
«ΣΥΣΤΗΜΑΤΑ ΑΥΤΟΜΑΤΙΣΜΟΥ»

Μεταπτυχιακή Εργασία

**Ανάπτυξη Συστήματος Πλοήγησης Ρομποτικού Οχήματος Εδάφους με χρήση  
κάμερας και GPS για εξερεύνηση σε Αμπελώνα**

Αυγουστάκης Γιάννης

*Επιβλέπων Καθηγητής: Τζαφέστας Κωνσταντίνος*

*Φουντάς Σπύρος*

ΑΘΗΝΑ 2021

## Περίληψη

Την τελευταία δεκαετία, υπάρχει ένα ολοένα και αυξανόμενο ενδιαφέρον για την ανάπτυξη αυτόνομων ρομποτικών οχημάτων στον αγροτικό τομέα λόγω των πολλαπλών αγροτικών χρήσεων που προσφέρουν όπως κλάδεμα, συγκομιδή ή τηλεπισκόπηση. Αποτελεί λοιπόν μείζον ζήτημα, για την οποιαδήποτε επίτευξη ενός στόχου ή μίας εργασίας, μία ρομποτική πλατφόρμα να διαθέτει ένα επίπεδο αυτονομίας χωρίς οποιαδήποτε ανθρώπινη παρέμβαση. Για την επίτευξη της επιζητούμενης αυτονομίας, το πρόβλημα του αυτοενοτισμού αποτελεί ένα από τα σημαντικότερα ανοιχτά προβλήματα. Στην παρούσα διπλωματική εργασία, δημιουργήθηκαν δύο υλοποιήσεις επίλυσης του αυτοενοτισμού βασισμένες στο Επεκταμένο φίλτρο Κάλμαν. Η παραλλαγή των δύο υλοποιήσεων στηρίζεται στην χρήση διαφορετικών αισθητήρων που χρησιμοποιήθηκαν καθώς στην πρώτη χρησιμοποιήθηκε ο αισθητήρας GPS ενώ στην δεύτερη χρησιμοποιήθηκε ένας οπτικός αισθητήρας. Για την σύγκριση και αξιολόγηση των δύο οδομετρικών εφαρμογών σχεδιάστηκε μία μεθοδολογία που υποδηλώνει την πραγματικότητα και χρησιμοποιείται ως μία αληθής μέτρηση η οποία επιτρέπει την σύγκριση διαφόρων εκτιμήσεων καθώς επίσης και ένα σύνολο πειραμάτων σε μηχανή προσομοίωσης αλλά και σε πραγματικές φυσικές συνθήκες. Για κάθε μία από τις δύο εκτιμήσεις εξετάστηκε η συνολική αλλά και η τοπική ακρίβεια σε σχέση με την πραγματικότητα. Τα ευρήματα που παρουσιάζονται στην παρούσα διπλωματική εργασία δείχνουν ότι ο οπτικός αισθητήρας παρουσιάζει παρόμοια αριθμητικά σφάλματα ως προς την συνολική ακρίβεια ενώ μικρότερα σφάλματα στην περίπτωση της τοπικής ακρίβειας.

## **Abstract**

In the last decade, there is a growing interest for autonomous robotic vehicles in the agricultural sector, due to the multiple agricultural uses such as pruning, harvesting or remote sensing. Ensuring a level of autonomy for the robotic platform is an important and challenging task that has been a goal of Roboticists and other visionaries since the emergence of this field. To achieve the desired level of autonomy, one must achieve of solving the localization principle meaning the constant monitoring of the route of the robot. In the present thesis, two odometry implementations based on the Extended Kalman Filter were created in order to solve the localization problem. The main difference of the two implementations is the different sensors used in each, as in first a GPS and in the second a Camera was used. For the comparison and evaluation of the two odometry estimations, a Groundtruth methodology was designed that compares the different two odometry estimations. A series of experiments was carried out, in a simulation environment but also in real open field scenario. In the last section of this thesis, a table that contains the error metrics shows that the camera has the same accuracy in global scale but in local scale the camera has better results.

## Ευχαριστίες

Αρχικά, θα ήθελα να ευχαριστήσω τον επιβλέποντα καθηγητή μου, κ. Τζαφέστα Κωνσταντίνο, για τις συμβουλές και την συνεχή υποστήριξη και εμπιστοσύνη που έδειξε στο πρόσωπο μου, αλλά και για την άριστη συνεργασία μας στα πλαίσια εκπόνησης της παρούσας διπλωματικής εργασίας.

Ένα μεγάλο ευχαριστώ, οφείλω στον συνεπιβλέποντα καθηγητή μου από το Γεωπονικό Πανεπιστήμιο Αθηνών, κ. Φουντά Σπύρο αλλά και τον επιβλέποντα και μέλος της ομάδας “Smart Farming Technology Group” Μυλωνά Νίκο που μου έδωσε την ευκαιρία να γνωρίσω, να ασχοληθώ και να αγαπήσω τον τομέα της ρομποτικής και ταυτόχρονα να βιώσω ανεκτίμητες εμπειρίες, μέσα σε ένα περιβάλλον εξερεύνησης, δημιουργικότητας και στοχασμού.

Τέλος, θα ήθελα να εκφράσω ένα μεγάλο ευχαριστώ στους γονείς μου, Ελένη και Παναγιώτης, όπως επίσης και την αδερφή μου Δάφνη, για την υπομονή και αμέριστη υποστήριξη τους, τόσο κατά την διάρκεια εκπόνησης της παρούσας διπλωματικής, αλλά και ακόμα περισσότερο, κατά την διάρκεια των σπουδών μου.

## Πίνακας Περιεχομένων

1. Εισαγωγή .....	1
1.1 Ρομποτική στην Γεωργία .....	1
1.2 Σκοπός εργασίας.....	1
1.3 Διάρθρωση της Διπλωματικής.....	2
2. Στοιχεία Θεωρίας.....	4
2.1. Μηχανισμοί Κίνησης .....	4
2.2. Διαφορική Οδήγηση .....	6
2.3. Αισθητήρες .....	9
2.3.1. Οπτικοί Κωδικοποιητές .....	10
2.3.2. Αισθητήρες Θέσης/Προσανατολισμού .....	12
2.3.3. GPS-RTK.....	17
2.3.4. LiDAR.....	19
2.3.5. Μονοσκοπική Όραση.....	21
2.3.6. Στερεοσκοπική Όραση .....	24
3. Αυτόνομη Πλοήγηση σε άγνωστο Περιβάλλον.....	30
3.1. Αυτοεντοπισμός, Χαρτογράφηση και Σχεδιασμός Δρόμου ....	Error! Bookmark not defined.
3.2. Οδομετρία.....	31
3.3. Οπτική Οδομετρία .....	32
3.3.1. Χαρακτηριστικά Εικόνας.....	34
3.3.2. Περιγραφείς Χαρακτηριστικών.....	36
3.3.3. Αντιστοίχιση Χαρακτηριστικών.....	38
3.3.4. Εκτίμηση Κίνησης.....	40
3.4. Φίλτρα Κάλμαν .....	43
3.5. Επεκταμένο Φίλτρο Κάλμαν.....	46
4. Αρχιτεκτονική Συστήματος σε επίπεδο υλοποίησης.....	48
4.1. ROS - Robot Operating System.....	48
4.2. Μηχανή Προσομοίωσης Gazebo.....	52
4.3. Robot Localization .....	53
4.4. RTAB-Map.....	54
5. Πείραμα και Αποτελέσματα.....	57
5.1. Δημιουργία Groundtruth .....	58
5.2. Προσομοίωση Πειράματος στο Gazebo.....	61
5.3. Πείραμα σε Αμπελώνα.....	64

<b>5.4. Αποτελέσματα.....</b>	<b>66</b>
<b>6. Συμπεράσματα.....</b>	<b>74</b>
<b>7. Μελλοντική Εργασία .....</b>	<b>76</b>
<b>Βιβλιογραφία .....</b>	<b>77</b>

## Κατάλογος Εικόνων

Εικόνα 1. Οι τέσσερις βασικοί τύποι τροχών. a) Τυπικός τροχός (standard wheel), b) Περιστρεφόμενος Τροχός Castor (castor wheel), c) Πανκατευθυντικός τροχός (Swedish wheel), d) Σφαιρικός τροχός (Spherical wheel). Πηγή [12].....	5
Εικόνα 2. Κινηματική διαφορικής οδήγησης. Πηγή [15] .....	7
Εικόνα 3. Το ρομπότ Husky από την εταιρεία Clearpath. Πηγή [16].....	9
Εικόνα 4. Κωδικοποίηση δίσκου από διαφανείς και μη διαφανείς επιφάνειες και παραγόμενες ακολουθίες. Πηγή [2].....	11
Εικόνα 5. Έξοδος καναλιών για τις τέσσερις διαφορετικές καταστάσεις για την μέτρηση CPR. Πηγή [19] .....	11
Εικόνα 6. Μηχανικό Γυροσκόπιο και δύναμη Coriolis. Πηγή [20].....	12
Εικόνα 7. Διάταξη ηλεκτρομηχανικού γυροσκοπίου. Πηγή [20].....	13
Εικόνα 8. Μηχανικό μοντέλο επιταχυνσιόμετρου. Πηγή [20].....	14
Εικόνα 9. Μοντέλο Πιεζοηλεκτρικού Επιταχυνσιόμετρου. Πηγή [20] .....	15
Εικόνα 10. Σχηματικό που περιγράφει την γενική λειτουργική δομή μίας IMU. Πηγή [21] .....	16
Εικόνα 11. Μονάδα Αδρανειακής Μέτρησης με ονομασία UM6 από την εταιρεία CH Robotics. Πηγή [22].....	16
Εικόνα 12. Προσδιορισμός της θέσης με χρήση ενός δορυφόρου (α), με χρήση δύο δορυφόρων (β) και τρεις δορυφόρους (γ). Πηγή [23]. .....	17
Εικόνα 13. Τοπική διόρθωση (αριστερά) και διόρθωση ευρύς κάλυψης (δεξιά) σε συστήματα DGPS. Πηγή [23].....	18
Εικόνα 14. Δέκτης GPS με ονομασία Reach RS2 από την εταιρεία Emlid. Πηγή [21] .....	19
Εικόνα 15. LiDAR LMS 10-1000 από την εταιρεία Sick (αριστερά), Σχηματικό λειτουργίας (δεξιά). ....	21
Εικόνα 16 Ένας αισθητήρας CCD μεγέθους 2048 x 2048 (Αριστερά), Κάμερα CMOS με φακό (Δεξιά). Πηγή [12].....	22
Εικόνα 17. Μοντελοποίηση ενός φωτογραφικού συστήματος με ονομασία "Κάμερα Μικροσκοπικής Οπής". Πηγή [33].....	22
Εικόνα 18. Μοντέλο Κάμερας Σημειακής Οπής. ....	24
Εικόνα 19. Μια μεμονωμένη λήψη όπως αυτή της εικόνας ενός ανθρώπου που κρατά ψηλά τον Πύργο της Πίζα οδηγεί σε διαφορετικά σενάρια όσον αφορά την αίσθηση του βάθους. Πηγή [34] .....	25
Εικόνα 20. Γενικευμένο γεωμετρικό μοντέλο διοπτικής κάμερας. [35].....	26
Εικόνα 21. Εικόνα με Χρώμα (αριστερά), Εικόνα Βάθους (δεξιά). ....	28
Εικόνα 22 Διάφορες όψεις της κάμερας Intel D435i. Πηγή [37] .....	28
Εικόνα 23. Τεχνικά Χαρακτηριστικά κάμερα Intel Realsense D435i. Πηγή [36].....	29
Εικόνα 24. Ονομασία κάθε προβλήματος ανάλογα με τον συνδυασμό των προβλημάτων “Σχεδίασης” “Έντοπισμού” και “Χαρτογράφησης”. Πηγή [38].....	30
Εικόνα 25 Μέτρηση ταχύτητας τροχών για τον προσδιορισμό της θέσης. ....	31
Εικόνα 26. Απεικόνιση του προβλήματος της Οπτικής Οδομετρίας. Πηγή [39].....	33
Εικόνα 27. Διάγραμμα Ροής που περιγράφει τα δομικά στοιχεία μίας Οπτικής Οδομετρίας. Πηγή [40] .....	34
Εικόνα 28 Περιοχές ιδιοτιμών στις οποίες χαρακτηρίζεται ένα σημείο ως επιθυμητό χαρακτηριστικό. Πηγή [50].....	36
Εικόνα 29 Αποτελέσματα από την ταυτοποίηση των χαρακτηριστικών με βάση την εκάστοτε στρατηγική. Πηγή [51] .....	38

Εικόνα 30. Ποσοστά αναγνώρισης χρησιμοποιώντας και τις πέντε στρατηγικές δειγματοληψίας. Πηγή [51].....	38
Εικόνα 31. Μπλε διαδρομή χωρίς την χρήση του RANSAC, κόκκινη διαδρομή με χρήση RANSAC .....	40
Εικόνα 32. Εκτίμηση Θέσης κάμερας με βάση την 3D-2D αντιστοίχιση χαρακτηριστικών. ....	42
Εικόνα 33. Διάγραμμα Ροής του αλγορίθμου ενός φίλτρου Κάλμαν. Πηγή [62] .....	44
Εικόνα 34. Ψευδοκώδικας που υλοποιεί τα βήματα ενός Φίλτρου Κάλμαν. Πηγή [61] .....	45
Εικόνα 35. Απεικόνιση φίλτρων Κάλμαν: (α) αρχική πεποίθηση, (β) μέτρηση (με έντονα γράμματα) με τη σχετική αβεβαιότητα, (γ) πεποίθηση μετά την ενσωμάτωση της μέτρησης χρησιμοποιώντας μετά την χρήση του φίλτρου. Πηγή [61] .....	45
Εικόνα 36. Ψευδοκώδικας που υλοποιεί τα βήματα ενός Εκτεταμένου Φίλτρου Κάλμαν. Πηγή [63] .....	46
Εικόνα 37. Γράφος κόμβων που δείχνει την κατεύθυνση των δεδομένων με βάση τον “ρόλο” του κάθε κόμβου.....	49
Εικόνα 38. Δομή δεδομένων για την δημιουργία ενός μηνύματος οδομετρίας στο ROS. [65] .....	50
Εικόνα 39. Σχέδιο που περιγράφει την δενδροειδή ένωση δύο συνδέσμων μέσω μίας άρθρωσης. Πηγή [67].....	51
Εικόνα 40. Το αρχικό GUI της μηχανής προσομοίωσης Gazebo. Πηγή [69] .....	52
Εικόνα 41. Στιγμιότυπο από την παρουσίαση του πακέτου robot_localization στο συνέδριο ROSCon. [71] .....	53
Εικόνα 42. Φωτογραφικό καρτέ στο οποίο φαίνονται με κόκκινο χρώμα τα σημεία που εξήχθησαν για το πρώτο βήμα για την δημιουργία μιας οπτικής οδομετρίας με χρήση του πακέτου RTAB-Map. ....	56
Εικόνα 43. Ταύτιση χαρακτηριστικών εικόνας σε δύο διαδοχικά καρτέ, με κόκκινο χρώμα φαίνονται τα σημεία τα οποία θεωρούνται outliers.....	56
Εικόνα 44. Αισθητήρια δεδομένα που χρησιμοποιήθηκαν ανά περίπτωση.....	57
Εικόνα 45. Σύνολο αντικειμένων τα οποία εντοπίζονται από τον αισθητήρα LiDAR. ....	58
Εικόνα 46. Τεχνητά σημεία αναφοράς (landmarks) το οποίο είναι τυλιγμένο με ανακλαστική ταινία (άσπρη και κόκκινη στρώση).....	59
Εικόνα 47. Σύνολο αντικειμένων (λευκά σημεία) τα οποία εντοπίζονται από το LiDAR(αριστερά), φιλτράρισμα με βάση την ανακλαστική ιδιότητα των αντικειμένων (δεξιά).....	59
Εικόνα 48. Κάτοψη που περιέχει την θέση της ρομποτικής πλατφόρμας σε δύο διαδοχικές στιγμές ως προς το σημείο αναφοράς.....	60
Εικόνα 49. Ρομποτική πλατφόρμα στην μηχανή προσομοίωσης Gazebo. ....	61
Εικόνα 50. Ρομποτική πλατφόρμα με εγκατεστημένα τις βάσεις και τους αισθητήρες στην μηχανή προσομοίωσης Gazebo. ....	62
Εικόνα 51. Περιβάλλον προσομοίωσης αμπελώνα με σημεία αναφοράς στο περιβάλλον Gazebo.....	63
Εικόνα 52. Αρχική θέση ρομποτικής πλατφόρμας. ....	64
Εικόνα 53. Πίνακες Συνδιακύμανσης για P και Q για την παραμετροποίηση του πρώτου Επεκταμένου Φίλτρου Κάλμαν. ....	65
Εικόνα 54. Πίνακες Συνδιακύμανσης για P και Q για την παραμετροποίηση του δεύτερου Επεκταμένου Φίλτρου Κάλμαν. ....	66
Εικόνα 55. Η διαδικασία της αξιολόγησης ενός δρόμου ξεκινάει με την ευθυγράμμιση με την πραγματική εκτίμηση και στην συνέχεια υπολογίζεται το σφάλμα με συγκεκριμένη μεθοδολογία μέτρησης σφάλματος. Πηγή [81] .....	67
Εικόνα 56. Απεικονίσεις του Απόλυτου Σφάλματος Δρόμου και Σχετικού Σφάλματος. Πηγή [81].....	68
Εικόνα 57 Γράφημα που περιέχει τους δρόμους που δημιουργήθηκαν από τις διάφορες εκτιμήσεις αλλά και της πραγματικής θέσης στο χώρο. ....	69



Εικόνα 58. Γράφημα που περιέχει τους ευθυγραμμισμένους δρόμους που δημιουργήθηκαν από τις διάφορες εκτιμήσεις αλλά και της πραγματικής θέσης στο χώρο με τον αλγόριθμο Uneyama. ....	70
Εικόνα 59. Μέγιστη και ελάχιστη τιμή, μέσος όρος και διάμεσος τιμών, διακύμανση και μέσο τετραγωνικού σφάλματος του μετρητή ATE (Absolute Trajectory Error) για την εκτίμηση με χρήση GPS (πράσινο χρώμα) και για την εκτίμηση με χρήση Κάμερας (μπλε χρώμα). ....	71
Εικόνα 60. Μέγιστη και ελάχιστη τιμή, μέσος όρος και διάμεσος τιμών, διακύμανση και μέσο τετραγωνικό σφάλμα για το RE με βάση την απόσταση 0.5 μέτρων. ....	72
Εικόνα 61. Μέγιστη και ελάχιστη τιμή, μέσος όρος και διάμεσος τιμών, διακύμανση και μέσο τετραγωνικό σφάλμα για το RE με βάση την απόσταση 1 μέτρου. ....	72
Εικόνα 62. Μέγιστη και ελάχιστη τιμή, μέσος όρος και διάμεσος τιμών, διακύμανση και μέσο τετραγωνικό σφάλμα για το RE με βάση την απόσταση 1.5 μέτρου. ....	73

## Κατάλογος Πινάκων

Πίνακας 1. Μορφή πίνακα Συνδιακύμανσης στο πακέτο robot_localization.....	54
Πίνακας 2. Πίνακας που περιέχει τις τιμές της μέγιστης και ελάχιστης τιμής, τον μέσο όρο και την διάμεσο τιμών, την διακύμανση και το μέσο τετραγωνικό σφάλμα του μετρητή ATE (Absolute Trajectory Error) για την εκτίμηση με χρήση GPS και για την εκτίμηση της οπτικής οδομετρίας.....	71
Πίνακας 3. Μέση τιμή και την διακύμανση για τον μετρητή RE, ανά διαφορετική παράμετρο και οδομετρική εκτίμηση.....	73

# Κεφάλαιο 1

## Εισαγωγή

### 1.1 Ρομποτική στην Γεωργία

Ο τομέας της ρομποτικής αποτελεί ένα σχετικά νέο επιστημονικό πεδίο το οποίο βασίζεται σε ένα ευρύ φάσμα επιστημονικών πεδίων όπως η μηχανική, η ηλεκτρονική, η επιστήμη των υπολογιστών με κύριο στόχο την σχεδίαση και την δημιουργία ρομποτικών μηχανισμών για την αυτοματοποίηση ενός συνόλου εργασιών.

Η κύρια χρήση της ρομποτικής εφαρμόζονταν κυρίως στις βιομηχανίες, μέσω των ρομποτικών βραχιόνων για την αυτοματοποίηση της γραμμής παραγωγής προϊόντων σε εργοστάσια. Ένας νέος και ανερχόμενος τομέας στον οποίο η ρομποτική βρίσκει εφαρμογές είναι αυτός της Γεωπονίας και πιο συγκεκριμένα στο κομμάτι των Ψεκασμών σε καλλιέργειες. Είναι γνωστό ότι ψεκασμοί χρησιμοποιούνται κυρίως για την προστασία των καλλιεργειών απέναντι σε ασθένειες και έντομα καθώς επίσης και για την ανάπτυξη της καλλιέργειας αυτής καθ' αυτής. Παρ' όλα τα οικονομικά οφέλη που προσφέρουν, οι Ψεκασμοί συνοδεύονται και από κάποια αρνητικά χαρακτηριστικά. Κάποια από τα βασικότερα συνδέονται με τις περιβαλλοντικές επιπτώσεις, την αυξημένη ζήτηση για εργατικό δυναμικό καθώς επίσης και τις βλαβερές για τον άνθρωπο συνέπειες λόγω των τοξικών δραστικών ουσιών που χρησιμοποιούνται στους ψεκασμούς. Επίσης σημαντικό γεγονός αποτελεί ότι η επαναλαμβανόμενη χρήση των τρακτέρ για εργασίες ψεκασμού, μπορεί να απελευθερώσει περισσότερα από 4.310 κιλά CO<sub>2</sub>/εκτάριο.

Γι' αυτούς τους λόγους οι γεωργικοί ψεκασμοί στρέφονται τα τελευταία χρόνια στην ανάπτυξη και σχεδίαση αυτόνομων ρομποτικών συστημάτων ικανών να αυτοματοποιήσουν και να διευκολύνουν ολόκληρη την διαδικασία του ψεκασμού. Τα γεωργικά ρομπότ έχουν τη δυνατότητα να αντικαταστήσουν την ανθρώπινη παρουσία κατά την διάρκεια των ψεκασμών όπως επίσης λόγω της ελαφριάς και αρθρωτής κατασκευής τους να συμβάλλουν θετικά σε περιβαλλοντικά και πρακτικά χαρακτηριστικά.

Ενδεικτικό παράδειγμα της χρήσης κινητών ρομποτικών συστημάτων αλλά και αυτόνομων τρακτέρ αποτελούν δύο Ευρωπαϊκά Έργα με ονόματα “Agrobofood” [1] και “Robots for Protecting the Crops” [2], στα οποία βλέπουμε την απτή εφαρμογή και χρήση μέσω πιλοτικών δοκιμών των αγροτικών ρομπότ στις γεωργικούς ψεκασμούς. Τέλος μεγάλο ενδιαφέρον παρουσιάζουν και οι εταιρείες οι οποίες σχεδιάζουν αγροτικά ρομπότ τα οποία χρησιμοποιούνται στους ψεκασμούς. Κάποιες γνωστές ρομποτικές αγροτικές πλατφόρμες αποτελούν το ρομπότ Ceol [3], το Robboti [4] αλλά και η πλατφόρμα Farmdroid [5], οι οποίες δείχνουν ότι τα αγροτικά ρομπότ ολοένα κερδίζουν έδαφος στο κομμάτι της γεωργίας.

### 1.2 Σκοπός εργασίας

Ένα από τα πιο σημαντικά χαρακτηριστικά το οποίο επιβάλλεται να έχει μία κινητή αγροτική ρομποτική πλατφόρμα είναι αυτό της αυτόνομης πλοήγησης σε ένα εξωτερικό περιβάλλον. Αυτό το χαρακτηριστικό ανάγεται σε ένα σύνολο προβλημάτων τα οποία όπως θα δούμε σε αυτήν την διπλωματική

εργασία λύνονται με πολλές και διαφορετικές προσεγγίσεις. Για να μπορέσει να λυθεί το πρόβλημα της αυτόνομης πλοήγησης, ικανή και αναγκαία προϋπόθεση αποτελεί η λύση του προβλήματος του αυτοεντοπισμού. Όπως θα δούμε με μεγαλύτερη λεπτομέρεια, ο αυτοεντοπισμός πρόκειται για δυνατότητα της ρομποτικής πλατφόρμας να γνωρίζει πού βρίσκεται ανά πάσα ώρα και στιγμή κατά την διάρκεια λειτουργίας της.

Η πιο κοινή λύση για τις αγροτικές ρομποτικές πλατφόρμες οι οποίες χρησιμοποιούνται σε εξωτερικό περιβάλλον είναι η χρήση του συστήματος GPS (Global Navigation Satellite System) ([6], [7]). Ωστόσο σε πολλούς γεωργικούς και δασικούς χώρους προκύπτει ότι, η έλλειψη δορυφορικού σήματος σε τέτοιες περιοχές αποτελεί ένα κύριο πρόβλημα [7,8], καθιστώντας τη χρήση του GPS αναξιόπιστη. Κάτω από αυτό το πλαίσιο, αποτελεί μείζονος σημασία η έρευνα αλλά και η ανάπτυξη έξυπνων λύσεων οι οποίες χρησιμοποιούν διαφορετικές προσεγγίσεις για την επίλυση του προβλήματος του αυτοεντοπισμού χωρίς την χρήση δορυφορικών σημάτων.

Μία διαφορετική προσέγγιση του προβλήματος του αυτοεντοπισμού, αποτελεί η χρήση μίας υπολογιστικής εκτίμησης της πόζας του ρομπότ, με χρήση οπτικών αισθητήρων η οποία ονομάζεται **Οπτική Οδομετρία** (Visual Odometry). Η Οπτική Οδομετρία αποτελεί μία πολύ δημοφιλή τεχνική στον τομέα της κινητής ρομποτικής λόγω του ότι μπορεί να χρησιμοποιηθεί σε περιβάλλοντα που υπάρχει έλλειψη σήματος GPS καθώς επίσης και λόγω της χαμηλής τιμής των αισθητήρων που μπορούν να την δημιουργήσουν. Γι' αυτόν τον λόγο μία πολύ μεγάλη ερευνητική δραστηριότητα έχει επικεντρωθεί στην επίλυση του προβλήματος του αυτοεντοπισμού σε εξωτερικό αγροτικό περιβάλλον χρησιμοποιώντας αυτήν την τεχνική ([8], [9], [10], [11]).

Κύριο λοιπόν αντικείμενο αυτής της διπλωματικής εργασίας αποτελεί η συγκριτική μελέτη της οπτικής οδομετρίας σε σχέση με την οδομετρία η οποία δημιουργήθηκε με χρήση GPS, σαν κύριες πηγές για την επίλυση του προβλήματος του αυτοεντοπισμού σε εξωτερικό αγροτικό περιβάλλον και συγκεκριμένα σε έναν αμπελώνα. Για να μπορέσει να γίνει αυτή η σύγκριση των δύο διαφορετικών προσεγγίσεων δημιουργήθηκε μία μεθοδολογία για την εκτίμηση της θέσης στην πραγματικότητα (GroundTruth) χρησιμοποιώντας αισθητήρα LiDAR. Τέλος, εκτός της διεξαγωγής πειράματος σε πραγματικό αμπελώνα, σχεδιάστηκε προσομοίωση στην μηχανή προσομοίωσης Gazebo για την επαλήθευση της μεθοδολογίας πριν οποιαδήποτε εκτέλεση της πραγματικής πειραματικής διαδικασίας.

### 1.3 Διάρθρωση της Διπλωματικής

Στα επόμενα κεφάλαια θα παρουσιαστεί αναλυτικά η ανάπτυξη και σύγκριση των δύο εκτιμήσεων του αυτοεντοπισμού της ρομποτικής πλατφόρμας Clearpath Husky-a200 αλλά και η σύγκριση και παρουσίαση των αποτελεσμάτων για κάθε εκτίμηση ως προς την πραγματική θέση. Συγκεκριμένα:

- Στο **Κεφάλαιο 2** παρουσιάζεται η θεωρία για τους βασικούς αισθητήρες αλλά και η μοντελοποίηση και περιγραφή του κινηματικού μοντέλου όσον αφορά τη μετάδοση κίνησης των κινητήρων στους τροχούς της ρομποτικής πλατφόρμας Clearpath Husky-a200.
- Στο **Κεφάλαιο 3** περιγράφεται το κύριο πρόβλημα του αυτοεντοπισμού ως ένα υποσύνολο του προβλήματος της Χαρτογράφησης και Πλοήγησης, το οποίο και αποτελεί αναπόσπαστο μέρος στον τομέα της κινητής ρομποτικής. Επιπρόσθετα, παρουσιάζουμε μία βασική μαθηματική προσέγγιση η οποία δίνει λύση στο πρόβλημα του αυτοεντοπισμού βασιζόμενη στην εφαρμογή ενός Επεκταμένου Φίλτρου Κάλμαν (Extended Kalman Filter - EKF).

- Στο **Κεφάλαιο 4**, παρουσιάζεται η αρχιτεκτονική του συστήματος λογισμικού που υλοποιήθηκε για την ρομποτική πλατφόρμα Clearpath Husky-a200, όπως επίσης και τα εργαλεία που χρησιμοποιήθηκαν, με έμφαση στο meta-λειτουργικό σύστημα ROS (Robotic Operating System) αλλά και την μηχανή προσομοίωσης Gazebo.
- Στο **Κεφάλαιο 5** παρουσιάζεται το πειραματικό μέρος, τόσο σε επίπεδο προσομοίωσης, όσο και σε φυσικό επίπεδο, όπως επίσης και η παρουσίαση των αποτελεσμάτων για κάθε εκτίμηση με χρήση GPS αλλά και κάμερα για τον αυτοεντοπισμό της ρομποτικής πλατφόρμας Clearpath Husky-a200, που στοχεύουν στην επίτευξη της αυτόνομης λειτουργίας της.
- Στο **Κεφάλαιο 6** γίνεται μία έκθεση των συμπερασμάτων που προέκυψαν βάσει της ανάλυσης και των πειραμάτων που πραγματοποιήθηκαν για κάθε ένα από τα υποσυστήματα της ρομποτικής πλατφόρμας, ενώ, τελικά, αναφέρεται και ένα σύνολο ιδεών και προτάσεων για την βελτίωση της υπάρχουσας υλοποίησης, τόσο ως προς το υλικό, όσο και ως προς τους αλγορίθμους και το αντίστοιχο λογισμικό.

# Κεφάλαιο 2

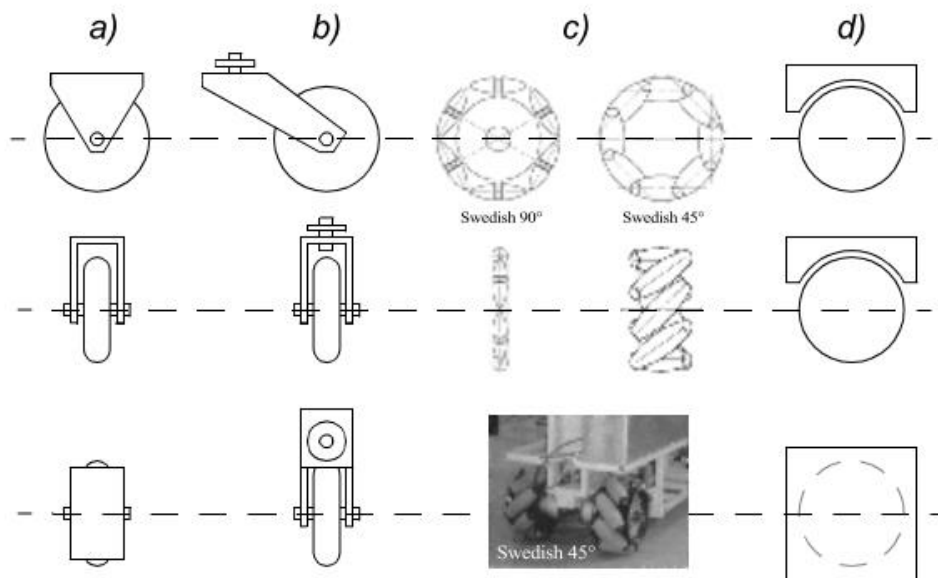
## Στοιχεία Θεωρίας

Σκοπός αυτού του κεφαλαίου είναι να εισάγει τον αναγνώστη στο βασικό θεωρητικό υπόβαθρο, στο οποίο βασίστηκε η μεθοδολογία και κατ' επέκταση και ο πειραματικός σχεδιασμός αυτής της διπλωματικής. Γι' αυτό τον λόγο αναφέρονται σε πρώτη φάση οι διάφοροι μηχανισμοί κίνησης αλλά και η κινηματική ανάλυση για την ρομποτική πλατφόρμα Husky-a200. Στην συνέχεια, γίνεται μία περιγραφή των διαφόρων αισθητήριων διατάξεων οι οποίοι χρησιμοποιούνται για την επίλυση του προβλήματος του αυτοπροσδιορισμού της πόζας μίας ρομποτικής πλατφόρμας. Για κάθε κατηγορία αισθητήρα στην οποία αναφερόμαστε, θα δούμε την αρχή λειτουργίας της, καθώς επίσης παρουσιάζονται και οι αισθητήριες διατάξεις που χρησιμοποιήθηκαν στην παρούσα διπλωματική εργασία.

### 2.1 Μηχανισμοί Κίνησης

Μία κινητή ρομποτική πλατφόρμα χρειάζεται έναν καθορισμένο μηχανισμό μετακίνησης ο οποίος της επιτρέπει την κίνηση στο περιβάλλον λειτουργίας της. Ο τρόπος κίνησης (locomotion) της ρομποτικής πλατφόρμας αποτελεί ένα σημαντικό αντικείμενο, στο οποίο συναντάται μία μεγάλη ποικιλία διαφορετικών υλοποιήσεων ανάλογα με τις απαιτήσεις του εκάστοτε συστήματος. Οι περισσότεροι από τους μηχανισμούς μετακίνησης στην πραγματικότητα είναι εμπνευσμένοι από βιολογικά μοντέλα που έχουν αναπτυχθεί από την φύση. Κάτω από αυτό το πλαίσιο έχουν αναπτυχθεί ερευνητικές ή εμπορικές κινητές ρομποτικές πλατφόρμες οι οποίες μπορούν να περπατήσουν, να πηδήξουν, να τρέξουν, να γλιστρήσουν, να κολυπήσουν, να πετάξουν και, φυσικά, να κυλήσουν. Φυσικά η ανάλυση και ο σχεδιασμός ενός μηχανισμού κίνησης αποτελεί ένα πολύπλοκο και πολυπαραγοντικό ζήτημα και γι' αυτό το λόγο τα περισσότερα κινητά ρομπότ βασίζονται στην κίνηση με κάποιο τροχήλατο μηχανισμό (Wheeled Locomotion) ή χρησιμοποιώντας μικρό αριθμό αρθρωτών σκελών (Legged Locomotion). Στα πλαίσια αυτής διπλωματικής θα αναφέρουμε κάποιες περαιτέρω πληροφορίες για την πρώτη εκ των δύο κατηγοριών (Wheeled Locomotion) καθώς η ρομποτική πλατφόρμα που χρησιμοποιήθηκε ανήκει σε αυτήν την κατηγορία.

Ο τροχός αποτελεί τον πιο δημοφιλή μηχανισμό μετακίνησης στην κινητή ρομποτική αλλά και γενικά στα τεχνητά οχήματα. Σε γενικές γραμμές, ένας ρομποτικός μηχανισμός ο οποίος χρησιμοποιεί τροχούς καταναλώνει λιγότερη ενέργεια και κινείται γρηγορότερα σε σχέση με άλλους μηχανισμούς κίνησης (π.χ. ρομπότ με πόδια). Επίσης από την πλευρά του ελέγχου ενός τέτοιου μηχανισμού, το πρόβλημα ανάγεται σε ευκολότερη ανάλυση, λόγω των απλών μηχανισμών και των μειωμένων προβλημάτων σταθερότητας που προσδίδει ένας τροχός. Παρ' όλ' αυτά στην κατηγορία των τροχοφόρων ρομπότ, υπάρχει μία πληθώρα διαφορετικών υλοποιήσεων οι οποίοι κυρίως διαφέρουν ως προς γεωμετρία των τροχών αλλά και στην επιλογή του αριθμού και τον τύπο του τροχού αυτού καθ' αυτού. Στην **Εικόνα 1** συνοψίζουμε όλους τους βασικούς τύπους των τροχών τους οποίους συναντάμε στις τροχοφόρες ρομποτικές πλατφόρμες.



Εικόνα 1. Οι τέσσερις βασικοί τύποι τροχών. a) Τυπικός τροχός (standard wheel), b) Περιστρεφόμενος Τροχός Castor (castor wheel), c) Πανκατευθυντικός τροχός (Swedish wheel), d) Σφαιρικός τροχός (Spherical wheel). Πηγή [12]

Αποτελεί γεγονός ότι, κάθε ένας τύπος τροχού έχει διαφορετική επίδραση στο τελικό κινηματικό μοντέλο του μηχανισμού. Σε περίπτωση που χρησιμοποιηθούν οι δύο πρώτες κατηγορίες τροχών σε μία ρομποτική πλατφόρμα (**Εικόνα 1**(a)(b) , standard wheel και castor wheel), για οποιαδήποτε ζητούμενη αλλαγή κατεύθυνση του οχήματος θα πρέπει πρώτα να περιστραφούν οι τροχοί κατά μήκος ενός κατακόρυφου άξονα. Αυτοί κινηματικοί περιορισμοί, οι οποίοι δεν επιτρέπουν την οποιαδήποτε κίνηση πάνω στο διδιάστατο επίπεδο ονομάζονται “**Ολόνομοι**” (holonomic) περιορισμοί. Στις επόμενες δύο κατηγορίες της **Εικόνα 1**(c)(d) (Swedish και Spherical wheels) ανήκουν τροχοί οι οποίοι προσδίδουν μία μεγαλύτερη ευελιξία ως προς την κίνησή τους στον χώρο. Η Πανκατευθυντικοί τροχοί έχουν σχεδιασμένους κυλίνδρους γύρω από την περιφέρεια του τροχού οι οποίοι και επιτρέπουν την κίνηση και σε παραπάνω κατευθύνσεις. Τέλος οι σφαιρικοί τροχοί όπως φαίνονται στην **Εικόνα 1**(d) έχουν προσαρτημένη μία μπάλα της οποίας η περιστροφή καθορίζει και την κίνηση ολόκληρου του μηχανισμού. Αξίζει να αναφερθεί πως, οι τελευταίες δύο κατηγορίες ονομάζονται “**Πανκατευθυντικοί**” (omnidirectional ή non-holonomic) και έχουν την δυνατότητα της κίνησης του ρομποτικού μηχανισμού ως προς οποιαδήποτε κατεύθυνση ανεξάρτητα από τον προσανατολισμό στον οποίο βρίσκεται η πλατφόρμα.

## 2.2. Διαφορική Οδήγηση

Εν γένει, οι σχεδιαστές τροχοφόρων ρομποτικών οχημάτων τείνουν να εστιάζουν σε τρία βασικά προβλήματα: την ικανότητα πρόσφυσης και σταθερότητας ολόκληρου του μηχανισμού, την ικανότητα αλλαγής κατεύθυνσης και τον έλεγχο της κίνησης. Αξίζει να αναφερθεί ότι, μία συνήθης κατηγοριοποίηση αυτών των τροχήλατων οχημάτων με βάση το σύστημα κίνησης είναι η παρακάτω:

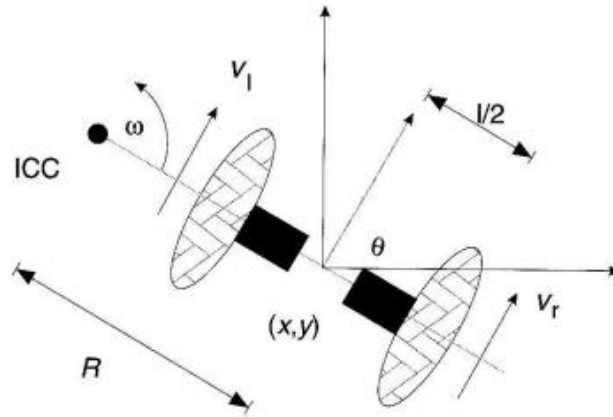
1. Διαφορικά οδηγούμενα οχήματα (*Differential Drive*)
2. Οδήγηση κατά Ackerman (*Car-type Drive*)
3. Πανκατευθυντική οδήγηση (*Omnidirectional Drive*)
4. Σύγχρονη οδήγηση (*Synchro Drive*)

Κάθε μία από τις παραπάνω κατηγορίες έχει διαφορετική προσέγγιση και ταυτόχρονα διαφορετικούς περιορισμούς ως προς την ικανότητα πρόσφυσης και σταθερότητας, την ικανότητα αλλαγής κατεύθυνσης αλλά και τον έλεγχο των τροχών του μηχανισμού.

Ενδεικτικά αναφέρουμε ότι η πανκατευθυντική οδήγηση επιτρέπει την κίνηση ενός ρομποτικού μηχανισμού στο επίπεδο ανεξάρτητα από τον προσανατολισμό του οχήματος. Για να μπορέσει ένα όχημα να χαρακτηριστεί ως πανκατευθυντικό θα πρέπει οπωσδήποτε να χρησιμοποιηθούν σουηδικοί ή σφαιρικοί τροχοί. Στην δεύτερη κατηγορία υπάρχει η οδήγηση κατά Ackerman, η οποία και αποτελεί και την δημοφιλέστερη επιλογή οδήγησης στην αυτοκινητοβιομηχανία μέχρι και σήμερα. Σκοπός του κινηματικού μοντέλου Ackermann είναι η αποφυγή της πλευρικής ολίσθησης των τροχών, ενός τετράτροχου οχήματος κατά την ακολούθηση καμπύλων τροχιών. Για την εξυπηρέτηση αυτού του σκοπού, λοιπόν, το κινηματικό μοντέλο Ackermann, στηρίζεται σε μία συνθήκη, την λεγόμενη συνθήκη Ackermann, μεταξύ των τροχών στρέψης ενός οχήματος, που αν ικανοποιείται προβλέπει την κίνηση των τροχών χωρίς πλευρική ολίσθηση [13]. Τέλος η σύγχρονη οδήγηση είναι μία κατηγορία στην οποία αν και υπάρχουν τρεις ή περισσότεροι τροχοί χρησιμοποιούνται μόνο δύο μοτέρ κίνησης. Το ένα εκ των δύο ορίζει την ταχύτητα και των τριών τροχών, ενώ το δεύτερο είναι υπεύθυνο για την περιστροφή για κάθε έναν από τους διαθέσιμους τροχούς.

Στην παρούσα διπλωματική εργασία ο μηχανισμός οδήγησης της ρομποτικής πλατφόρμας Husky-a200 αποτελείται με διαφορική οδήγηση. Τα διαφορικά οδηγούμενα έντροχα οχήματα αποτελούνται από δύο κινητήριους τροχούς τοποθετημένους σε έναν κοινό άξονα και κάθε τροχός μπορεί να οδηγηθεί ανεξάρτητα είτε προς τα εμπρός είτε προς τα πίσω. Μεταβάλλοντας την ταχύτητα κάθε τροχού, ένα τέτοιο όχημα μπορεί να εκτελέσει μία κίνηση κύλισης ικανή να το περιστρέψει γύρω από ένα σημείο που βρίσκεται κατά μήκος του κοινού άξονα του αριστερού και δεξιού τροχού. Αυτό το σημείο ονομάζεται “**Στιγμιαίο Κέντρο Περιστροφής**”(Instantaneous Center of Rotation-ICR [14]) και αποτελεί το κέντρο της στιγμιαίας κυκλικής τροχιάς που ακολουθεί το όχημα.





Εικόνα 2. Κινηματική διαφορικής οδήγησης. Πηγή [15]

Όπως αναφέρθηκε και νωρίτερα μεταβάλλοντας τις ταχύτητες των δύο τροχών, διαφοροποιούνται οι τροχιές που ακολουθεί το ρομπότ. Επειδή ο ρυθμός περιστροφής  $\omega$  για το ICC πρέπει να είναι ο ίδιος και για τους δύο τροχούς, μπορούμε να γράψουμε τις ακόλουθες εξισώσεις:

$$\omega * \left( R + \frac{l}{2} \right) = V_r \quad (2.1)$$

$$\omega * \left( R - \frac{l}{2} \right) = V_l \quad (2.2)$$

όπου  $l$  είναι η απόσταση μεταξύ των κέντρων των δύο τροχών,  $V_r$ ,  $V_l$  είναι οι ταχύτητες του δεξιού και του αριστερού τροχού κατά μήκος του εδάφους και  $R$  είναι η απόσταση από το ICC έως την μέση μεταξύ των δύο τροχών. Για κάθε χρονική στιγμή λύνοντας ως προς την μεταβλητή  $R$  και  $\omega$  έχουμε:

$$R = \frac{l}{2} * \frac{V_l + V_r}{V_r - V_l} \quad (2.3)$$

$$\omega = \frac{V_r - V_l}{l} \quad (2.4)$$

Από την παραπάνω εξίσωση προκύπτουν τα παρακάτω συμπεράσματα:

- Για να κινηθεί το ρομπότ σε ευθεία γραμμή, και οι δύο τροχοί θα πρέπει να κινούνται με ίδια κατεύθυνση και ταχύτητα. Το κέντρο περιστροφής του ρομπότ βρίσκεται σε άπειρη απόσταση από το ρομπότ ( $R \rightarrow \infty$ ) και η γωνία περιστροφής της πλατφόρμας είναι μηδέν ( $\omega = 0$ ).

- Για να μπορεί το ρομπότ να περιστρέφεται στη θέση του, και τα δύο σέτ τροχών πρέπει να στρίβουν με ίση ταχύτητα σε αντίθετες κατευθύνσεις. Το κέντρο περιστροφής βρίσκεται στο κέντρο του τον άξονα ( $R = 0$ ).

- Διαφορετικά, ανάλογα με την ταχύτητα περιστροφής και την κατεύθυνση της, το κέντρο περιστροφής μπορεί να πέσει οπουδήποτε στη γραμμή που ορίζεται από τα δύο σημεία επαφής των ελαστικών

Όσον αφορά το ευθύ κινηματικό μοντέλο ενός διαφορικού ρομπότ αναλύεται ως εξής. Έστω ένα διαφορικός οδηγούμενο όχημα. Η κατάσταση του ρομποτικού οχήματος, συνήθως περιγράφεται από έξι μεταβλητές, τις καρτεσιανές συντεταγμένες του  $x$ ,  $y$ ,  $z$ , ως προς ένα αυθαίρετο εξωτερικό σύστημα συντεταγμένων και τις γωνίες Euler yaw, pitch, roll. Προκύπτει λοιπόν ότι οι συνολικοί βαθμοί ελευθερίας για το συγκεκριμένο όχημα να είναι έξι. Ο λόγος διότι το συγκεκριμένο όχημα μπορεί να κινηθεί στο δισδιάστατο καρτεσιανό επίπεδο που ορίζουν οι άξονες  $x$  και  $y$  και μπορεί να μεταβάλλει τον προσανατολισμό του  $\theta$ .

Αλλάζοντας τις ταχύτητες του δεξιά και αριστερού τροχού  $V_l$ ,  $V_r$ , μπορούμε να κάνουμε το ρομπότ να μετακινηθεί σε διαφορετικές θέσεις αλλά προσανατολισμό. Γνωρίζοντας τις ταχύτητες  $V_l$ ,  $V_r$  και χρησιμοποιώντας την **Εξίσωση 2.3**, μπορούμε να βρούμε τη θέση ICC [15]:

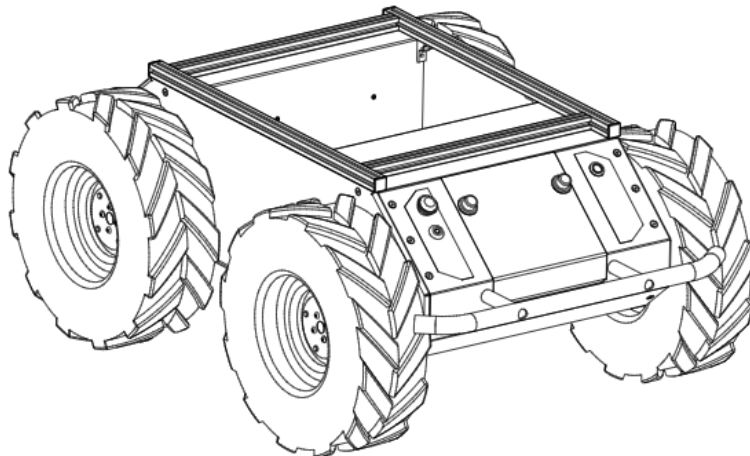
$$ICC = [x - R * \sin(\theta), y + R * \cos * (\theta)] \quad (2.4)$$

Σε μορφή πίνακα για κάποια χρονική στιγμή  $t + \delta t$  προκύπτει ότι η το ευθύ κινηματικό μοντέλο είναι:

$$\begin{bmatrix} x' \\ y' \\ \theta' \end{bmatrix} = \begin{bmatrix} \cos(\omega \delta t) & -\sin(\omega \delta t) & 0 \\ \sin(\omega \delta t) & \cos(\omega \delta t) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x - ICC_x \\ y - ICC_y \\ \theta \end{bmatrix} + \begin{bmatrix} ICC_x \\ ICC_y \\ \omega \delta t \end{bmatrix} \quad (2.5)$$

## Ρομποτική Πλατφόρμα Clearpath Husky-a200

Στα πλαίσια αυτής της διπλωματικής χρησιμοποιήθηκε η ρομποτική πλατφόρμα Husky-a200 [16] η οποία σχεδιάστηκε και κατασκευάστηκε από την εταιρεία Clearpath. Η συγκεκριμένη πλατφόρμα ανήκει ρομποτικές πλατφόρμες μεσαίου μεγέθους η οποία έχει μέγεθος 990 x 670 x 390 mm, πράγμα που την κάνει αρκετά καλή επιλογή για πειραματισμούς σε αμπελώνα λόγο του μικρού μεγέθους της. Το βάρος της είναι στα 50 κιλά, το μέγιστο βάρος που μπορεί να σηκώσει είναι στα 75 κιλά καθώς επίσης η μέγιστη ταχύτητα που μπορεί να αναπτύξει είναι 1 μέτρο/δευτερόλεπτο.



Εικόνα 3. Το ρομπότ Husky από την εταιρεία Clearpath. Πηγή [16]

Η συγκεκριμένη πλατφόρμα αποτελείται από δύο σετ κινούμενων τροχών τοποθετημένων εκατέρωθεν πάνω στο σασί του ρομπότ και με αυτό τον τρόπο να ακολουθεί την κίνηση μίας διαφορική οδήγησης όπως περιγράφηκε στην προηγούμενη Παράγραφο. Αυτό σημαίνει ότι για οποιαδήποτε περιστροφή ως προς τον κάθετο άξονα της πλατφόρμας, μεταβάλλονται σε αντίστοιχο ρυθμό περιστροφής οι τροχοί της πλατφόρμας. Αυτή η ικανότητα επιτρέπει στην πλατφόρμα Husky-a200, να έχει αρκετά λιγότερους κινηματικούς περιορισμούς έναντι σε άλλα μοντέλα κίνησης (όπως για παράδειγμα στροφή κατά Ackerman), πράγμα που την κάνει ευέλικτη για στροφές μεταξύ των γραμμών σε έναν αμπελώνα (U-turns). Αξίζει επίσης να αναφερθεί ότι η πλατφόρμα έχει εγκατεστημένο υπολογιστή με ROS kinetic [17], το οποίο και χρησιμοποιείται σαν κύριο μέσο επικοινωνίας του πειράματος.

## 2.3. Αισθητήρες

Για να μπορεί μία πλατφόρμα να λειτουργεί αυτόνομα θα πρέπει να έχει την ικανότητα να αντιλαμβάνεται το εξωτερικό περιβάλλον στο οποίο και λειτουργεί. Ως εκ τούτου, η χρήση διαφόρων αισθητήρων αλλά και η σωστή διαχείριση των διαφόρων πληροφοριών από καθ' ένα εξ αυτών αποτελεί επιτακτική ανάγκη η οποία επιτρέπει την αντίληψη εξωτερικών παρατηρήσεων κατά την διάρκεια λειτουργίας. Σύμφωνα με τους Siegwart και άλλοι [12], οι αισθητήρες ταξινομούνται ανάλογα με την λειτουργικότητα τους σε:

### 1) Ιδιοδεκτικούς/Εξωδεκτικούς (Proprioceptive/Exteroceptive)

Όπου Ιδιοδεκτικοί είναι οι αισθητήρες οι οποίοι δίνουν πληροφορία σχετική με το εσωτερικό σύστημα μίας ρομποτικής πλατφόρμας όπως η ταχύτητα των μοτέρ, η τάση της μπαταρίας, η θέση των γωνιών των αρθρώσεων ενός ρομποτικού βραχίονα κ.α. Στην δεύτερη κατηγορία ανήκουν οι Εξωδεκτικοί αισθητήρες οι οποίοι παρέχουν πληροφορία για το εξωτερικό περιβάλλον όπως μετρήσεις απόστασης από κάποιο αντικείμενο στο εξωτερικό περιβάλλον, η ένταση του φωτός, το εύρος ήχου κ.α.

## 2) Ενεργούς/Παθητικούς (Active/Passive)

Στην κατηγορία των ενεργών αισθητήρων περιλαμβάνονται οι διατάξεις οι οποίες εκπέμπουν ενέργεια στο εξωτερικό περιβάλλον και εν συνεχεία μετράνε την απόκριση που προκλήθηκε από αυτήν την εκπομπή. Αυτού του τύπου οι αισθητήρες θεωρούνται ανώτεροι, λόγω του ότι επιτρέπουν την άμεση αλληλεπίδραση με το περιβάλλον. Οι οπτικοί κωδικοποιητές, οι αισθητήρες υπερήχων και οι ανιχνευτές λέιζερ αποτελούν γνωστά παραδείγματα αυτής της κατηγορίας αισθητήρων. Τέλος στους παθητικούς αισθητήρες ανήκουν οι αισθητήρες οι οποίοι δεν εκπέμπουν κάποια μορφή ενέργειας αλλά μετρούν την ενέργεια που προέρχεται από κάποια φυσική πηγή. Ως παράδειγμα αποτελούν οι ανιχνευτές θερμοκρασίας, τα μικρόφωνα και οι κάμερες CCD ή CMOS. Στις επόμενες παραγράφους θα παρουσιαστούν εν συντομία όλοι οι αισθητήρες που εγκαταστάθηκαν αλλά χρησιμοποιήθηκαν στα πλαίσια αυτής της διπλωματικής εργασίας όπως επίσης και η βασική λειτουργική δομή καθενός.

### 2.3.1. Οπτικοί Κωδικοποιητές

Οι οπτικοί κωδικοποιητές είναι διατάξεις με τις οποίες μπορεί να μετρηθεί η γραμμική και η περιστροφική μετατόπιση. Η βασική αρχή λειτουργίας αυτών των αισθητήρων στηρίζεται στην ανίχνευση της φωτεινής ακτινοβολίας, η οποία διαπερνά διαμέσου μίας κατάλληλα διαμορφωμένης επιφάνειας, αλλά και η μετατροπή αυτής της ανίχνευσης σε ηλεκτρικούς παλμούς [18].

Η γραμμική ή περιστροφική κίνηση ενός οπτικού κωδικοποιητή, προκαλεί την μετατόπιση μιας κατάλληλα διαμορφωμένης επιφάνειας η οποία ονομάζεται **οπτική επιφάνεια**, και αποτελείται από διαφανείς ή μη διαφανείς περιοχές οι οποίες επιτρέπουν ή μη την διέλευση του φωτός. Οι διαφανείς περιοχές αυτής της επιφάνειας ονομάζονται **οπτικά παράθυρα**. Σε έναν τυπικό περιστροφικό οπτικό κωδικοποιητή τα οπτικά παράθυρα τοποθετούνται με τέτοιο τρόπο πάνω στην οπτική επιφάνεια έτσι ώστε κατά την περιστροφή του μηχανισμού να υπάρχουν χρονικά σημεία στα οποία επιτρέπεται η διέλευση του φωτός ή όχι, με αποτέλεσμα να δημιουργείται μία ακολουθία παλμών που αντιστοιχίζεται με την προς την μέτρηση μετατόπιση. Σε γενική περίπτωση, ένας οπτικός κωδικοποιητής για να μπορέσει να ανιχνεύσει την διέλευση του φωτός χρησιμοποιεί δύο φωτοδιακόπτες.

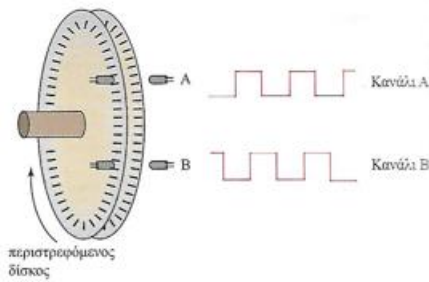
Οι οπτικοί κωδικοποιητές χωρίζονται σε δύο βασικές κατηγορίες:

- Τους αυξητικούς οπτικούς κωδικοποιητές
- Τους απόλυτους οπτικούς κωδικοποιητές

Η κύρια διαφορά μεταξύ των δύο διατάξεων είναι ότι πρώτοι παράγουν πληροφορία για την σχετική γωνιακή/γραμμική ταχύτητα, ενώ οι δεύτεροι για την απόλυτη γωνιακή/γραμμική ταχύτητα ως προς ένα σημείο αναφοράς.

Οι αυξητικοί οπτικοί κωδικοποιητές αποτελούν μία δημοφιλής συσκευή για τη μέτρηση της γωνιακής ταχύτητας και της θέσης ενός μοτέρ κίνησης το οποίο τοποθετείται πάνω στον άξονα ενός τροχού. Αυτό το χαρακτηριστικό των κωδικοποιητών, επιτρέπει τον έλεγχο της θέσης ή της ταχύτητας των τροχών σε μία ρομποτικό όχημα, το οποίο χρησιμοποιεί αυτούς τους κινητήρες για την οποιαδήποτε κίνηση του στον χώρο. Ένας τύπος της κατηγορίας των αυξητικών οπτικών κωδικοποιητών αποτελεί ο **“Τετραγωνικός Κωδικοποιητής” (Quadrature Encoder)**. Σε αυτήν την κατηγορία τοποθετούνται δύο φωτοδιακόπτες οι οποίοι είναι τοποθετημένοι με τέτοιο τρόπο στην περιφέρεια του οπτικού δίσκου έτσι

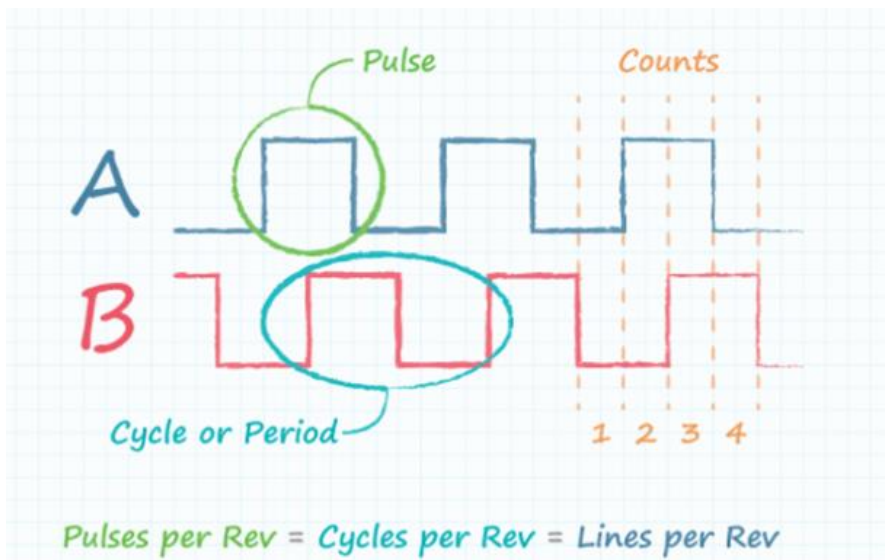
ώστε όταν ο δίσκος περιστραφεί με κάποια γωνιακή ταχύτητα, η διαφορά φάση των δύο παλμών να είναι 90°.



Εικόνα 4. Κωδικοποίηση δίσκου από διαφανείς και μη διαφανείς επιφάνειες και παραγόμενες ακολουθίες. Πηγή [2]

Η ταχύτητα της μετατόπισης δίνεται μετρώντας τον αριθμό των παλμών ενός εκ των δύο καναλιών και την αναγωγή τους σε γραμμική ή γωνιακή μετατόπιση με βάση τα γεωμετρικά χαρακτηριστικά του κατασκευαστή (γραμμική ή γωνιακή απόσταση ανάμεσα στα οπτικά παράθυρα). Όσον αφορά την διεύθυνση της μετατόπισης, εξάγεται με βάση την διαφορά φάσης που εντοπίζεται στα δύο κανάλια όπως φαίνεται και στην **Εικόνα 4**.

Η ακρίβεια ενός τέτοιου μηχανισμού μετριέται σε “**Κύκλος ανά Περιστροφή**” (**Cycles per Revolution** ή **CPR**), και ρυθμίζει το πόσο ακριβείς θέλουμε να είναι η μέτρηση μας. Η μονάδα CPR ορίζει τέσσερις διαφορετικές καταστάσεις με βάση τις εξόδους των δύο καναλιών. Χρησιμοποιώντας αυτήν τη μέθοδο, ο κωδικοποιητής υπολογίζει κάθε ακμή του σήματος σε γραφήματα δύο τετραγωνικών κυμάτων (ακμή A(1), πτώση A(0), ακμή B(1), πτώση B(0)).



Εικόνα 5. Έξοδος καναλιών για τις τέσσερις διαφορετικές καταστάσεις για την μέτρηση CPR. Πηγή [19]

Τυπικοί κωδικοποιητές που χρησιμοποιούνται στις ρομποτικές πλατφόρμες είναι της τάξης των 2.000 CPR, ενώ στην βιομηχανία χρησιμοποιούνται κωδικοποιητές της τάξης των 10.000 CPR [12].

### 2.3.2. Αισθητήρες Θέσης/Προσανατολισμού

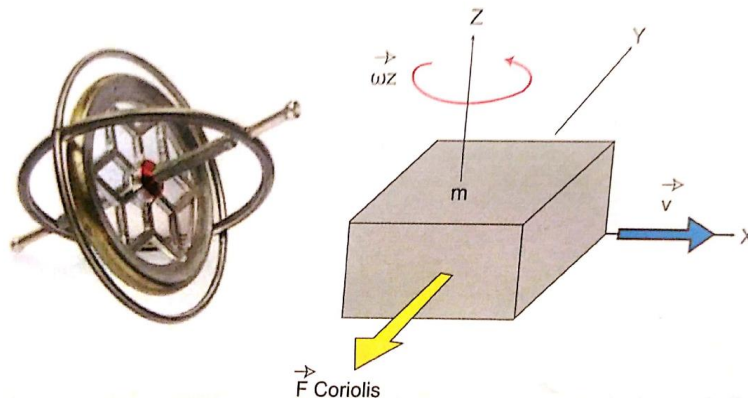
Κρίσιμο κομμάτι για την ορθή εκτίμηση της πόζας μίας ρομποτικής πλατφόρμας, αποτελούν οι αισθητήρες προσανατολισμού. Ως αισθητήρες προσανατολισμού θεωρούμε τις διατάξεις οι οποίες χρησιμοποιούνται για τον προσδιορισμό του προσανατολισμού και της κλίσης της πλατφόρμας αυτής καθ' αυτής. Σε αυτήν κατηγορία περιέχονται αισθητήρες οι οποίοι μπορεί να είναι Ιδιοδεκτικοί (όπως το γυροσκόπιο, επιταχυνσιόμετρο) και Εξωδεκτικοί (όπως η πυξίδα). Σκοπός αυτής της ενότητας είναι να παρουσιάσει τις δύο κεντρικές διατάξεις οι οποίες όταν συνδυαστούν δημιουργούν μία αισθητήρια διάταξη με όνομα **Μονάδα Αδρανειακής Μέτρησης** ή όπως συναντάται συχνά στην βιβλιογραφία με ακρωνύμιο IMU (Inertial Measurement Unit).

#### Γυροσκόπιο

Τα γυροσκόπια αποτελούν διατάξεις οι οποίες μετράνε τον γωνιακό ρυθμό της περιστροφής ως προς ένα σταθερό άξονα [20]. Η κατασκευή ενός γυροσκοπίου αποτελείται από έναν αδρανειακό δίσκο ο οποίος περιστρέφεται γύρω από ένα σύνολο δακτυλίων. Κατά την κίνηση ενός γυροσκοπίου η μάζα αντισταθμίζει αυτή την κίνηση με την αντίστοιχη αλλαγή του προσανατολισμού της. Ο πιο γνωστός τύπος γυροσκοπίου είναι το γυροσκόπιο Coriolis (Coriolis Vibratory Gyroscopes ή CVG), στο οποίο η αδρανειακή μάζα ταλαντώνεται με μεγάλη συχνότητα στην διεύθυνση του άξονα κίνησης και κατά την περίπτωση αλλαγής προσανατολισμού η δύναμη Coriolis ωθεί την αδρανειακή μάζα σε κατεύθυνση κάθετη σε αυτήν την περιστροφή [20]. Η δύναμη Coriolis δίνεται από τον τύπο:

$$\vec{F}_{Coriolis} = -2m(\vec{\omega} * \vec{v}) \quad (2.6)$$

και το αρνητικό πρόσημο της δηλώνει την αντιστάθμιση των πλευρικών επιταχύνσεων σε ένα αδρανές σύστημα.



Εικόνα 6. Μηχανικό Γυροσκόπιο και δύναμη Coriolis. Πηγή [20]

Συγκεκριμένη κατηγορία γυροσκοπίων αποτελούν τα ηλεκτρομηχανικά γυροσκόπια MEMS, στα οποία τοποθετούνται πιεζοηλεκτρικοί αισθητήρες οι οποίοι είναι σε θέση να ανιχνεύσουν την συχνότητα των δονήσεων που είναι ανάλογοι του ρυθμού περιστροφής ως προς κάποιο σταθερό άξονα. Από κατασκευαστικής άποψης, ένα ηλεκτρομηχανικό γυροσκόπιο αποτελείται από μία μάζα η οποία τοποθετείται στο κέντρο της συσκευής και συγκρατείται από κατάλληλα διαμορφωμένα στοιχεία ελατηρίων. Σε περίπτωση που προκληθεί οποιαδήποτε περιστροφή σε αυτό το σύστημα οι χωρητικοί αισθητήρες της συσκευής θα μετρήσουν την δύναμη Coriolis η οποία θα προκληθεί από την περιστροφή αυτή καθ' αυτή.



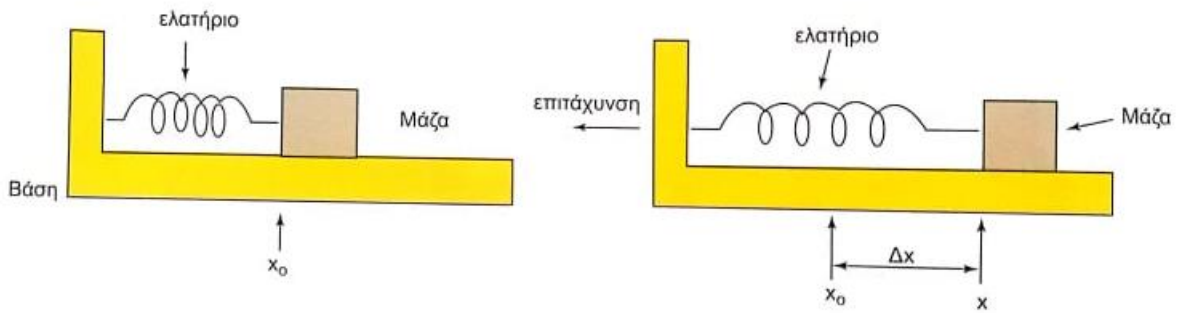
Εικόνα 7. Διάταξη ηλεκτρομηχανικού γυροσκοπίου. Πηγή [20]

Ένα κοινό πρόβλημα σε όλα τα γυροσκόπια είναι αυτό της απόκλισης (drift) των μετρήσεων και βασίζεται σε σφάλματα μετρήσεων τα οποία σχετίζονται με τις μεμονωμένες τεχνολογίες γυροσκοπίου. Οπότε κάθε μεμονωμένη μέτρηση αλλοιώνεται από αυτά τα σφάλματα τα οποία συσσωρεύονται με την πάροδο του χρόνου [21].

## Επιταχυνσιόμετρο

Εκτός από την μέτρηση του προσανατολισμού, σημαντική πληροφορία αποτελεί και η εκτίμηση της θέσης. Γι' αυτόν τον σκοπό χρησιμοποιούνται μία κατηγορία αδρανειακών αισθητήρων γνωστοί ως επιταχυνσιόμετρα, τα οποία έχουν την δυνατότητα να μετρήσουν τις εξωτερικές δυνάμεις οι οποίες ασκούνται σε μία ρομποτική πλατφόρμα και να την αποτυπώσουν σε μία ψηφιακή μορφή.

Η αρχή λειτουργίας ενός μηχανικού επιταχυνσιόμετρου αποτελείται από ένα σύστημα ενός ελατηρίου, μίας μάζας και ενός αποσβεστήρα, καθώς επίσης και ενός μηχανισμού ικανού να παρακολουθεί οποιαδήποτε μεταβολή ασκείται σε αυτό το σύστημα.



Εικόνα 8. Μηχανικό μοντέλο επιταχυνσιόμετρου. Πηγή [20]

Κάθε φορά που ασκείται κάποια εξωτερική δύναμη στο σύστημα, η μάζα  $m$  μετατοπίζεται κατά  $\Delta x$  με φορά αντίθετη στην κατεύθυνση παρασύροντας έτσι και ένα ελατήριο σταθεράς  $k$  που την συγκρατεί. Η επιτάχυνση του συστήματος δίνεται από την εφαρμογή των νόμων Newton και Hook μέσα από τον τύπο:

$$\alpha = \frac{k}{m} \Delta x \quad (2.7)$$

Όπου  $k$  είναι η σταθερά του ελατηρίου (N/m)

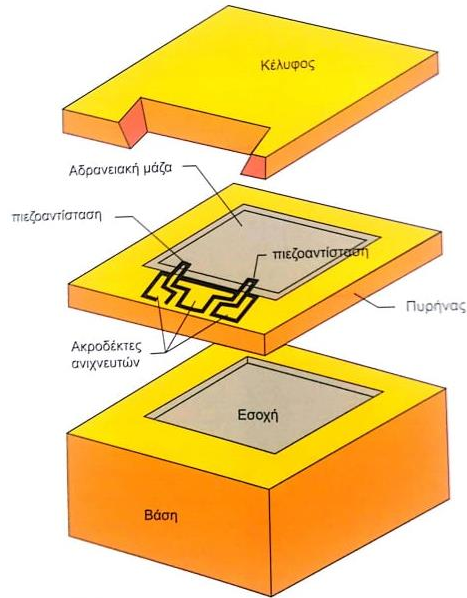
$\Delta x$  είναι η μετατόπιση του ελατηρίου από το σημείο ηρεμίας

$m$  είναι η τιμή της μάζας

$\alpha$  είναι η ζητούμενη επιτάχυνση.

Τέλος για την μέτρηση της επιτάχυνσης αλλά και την αποτύπωση της σε μία αριθμητική τιμή χρησιμοποιείται κάποιος αισθητήρας κίνησης όπως για παράδειγμα ένα ποτενσιόμετρο. Μία πολύ γνωστή κατηγορία επιταχυνσιόμετρων, αποτελούν τα πιεζοηλεκτρικά επιταχυνσιόμετρα. Αντί να στηρίζονται στην άμεση μηχανική μέτρηση των εξωτερικών δυνάμεων, τα πιεζοηλεκτρικά επιταχυνσιόμετρα βασίζονται σε μια ιδιότητα που εμφανίζεται σε ορισμένους κρυστάλλους, στους οποίους δημιουργείται ηλεκτρική τάση σε περίπτωση καταπόνησης τους. Από κατασκευαστικής άποψης ο πιεζοηλεκτρικός κρύσταλλος τοποθετείται ανάμεσα στο περίβλημα της συσκευής και μίας μάζας και σε περίπτωση πρόκλησης κάποια εξωτερικής δύναμης ασκείται καταπόνηση κρύσταλλο.





Εικόνα 9. Μοντέλο Πιεζοηλεκτρικού Επιταχυνσιόμετρου. Πηγή [20]

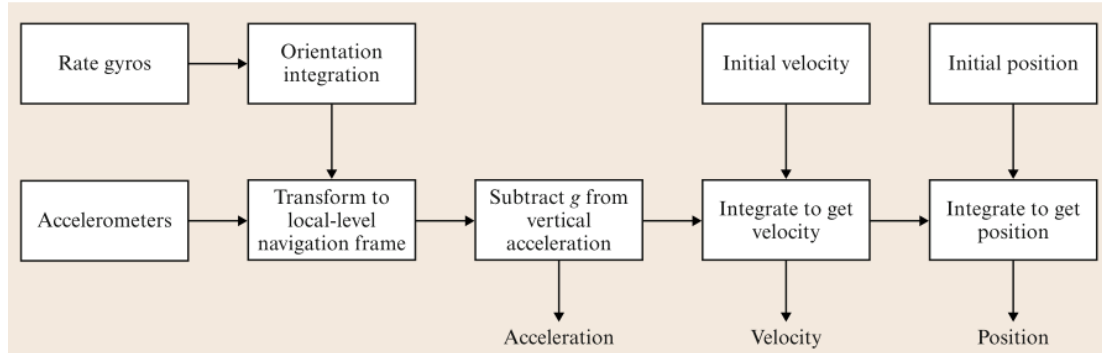
## Μονάδα Αδρανειακής Μέτρησης

Μια “**Μονάδα Αδρανειακής Μέτρησης**” (IMU) είναι μια συσκευή που χρησιμοποιεί συστήματα μέτρησης όπως γυροσκόπια και επιταχυνσιόμετρα για να εκτιμήσει τη σχετική θέση, την ταχύτητα και την επιτάχυνση ενός οχήματος το οποίο κινείται [21]. Ένα σύστημα πλοήγησης το οποίο χρησιμοποιεί μόνο τους ιδιοδεκτικούς αισθητήρες του ή/και συνδυασμό αυτών είναι γνωστό ως “**Σύστημα Αδρανειακής Πλοήγησης**” (Inertial Navigation System ή INS). Μια IMU κάνει μία εκτίμηση του χώρου με 6 βαθμούς ελευθερίας (DOF) που αυτό περιλαμβάνει την εκτίμηση της πόζας του οχήματος: θέση (x, y, z) και προσανατολισμός (yaw, pitch, roll).

Οι IMU χωρίζονται σε δύο είδη με ονομασίες “**Συστήματα Σταθερής Βάσης**” (Gimbaled Systems) και “**Συστήματα Δεμένα σε Ιμάντα**”(Strap-down Systems) [21]. Στην πρώτη κατηγορία ανήκουν οι συσκευές οι οποίες είναι τοποθετημένες σε πολύπλοκες σταθερές δομές προκειμένου να παρέχουν μια σταθερή πλατφόρμα στην οποία και γίνονται οι διάφορες μετρήσεις. Ο τελικός προσανατολισμός που προκύπτει χρησιμοποιεί ένα μετασχηματισμό από την σταθερή δομή ως προς το πλαίσιο αναφοράς της ρομποτικής πλατφόρμας. Στην δεύτερη και πιο δημοφιλής κατηγορία η IMU τοποθετείται πάνω στο όχημα με μία άκαμπτη σύνδεση, πράγμα που σημαίνει ότι δεν απαιτείται ο μετασχηματισμός της πρώτης κατηγορίας.

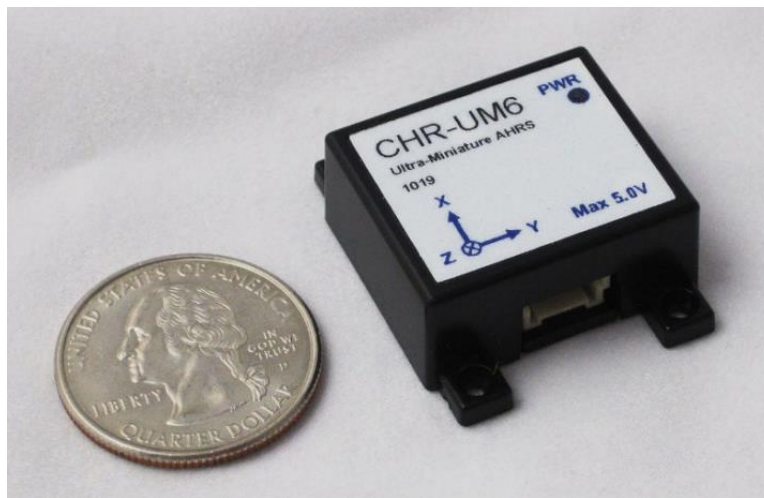
Στην **Εικόνα 10** παρουσιάζεται η βασική λειτουργική δομή μίας IMU όπως περιγράφεται από τους Dudek και άλλοι [21]. Σε αυτό το μοντέλο χρησιμοποιούνται τρία επιταχυνσιόμετρα για την εκτίμηση της στιγμιαίας επιτάχυνσης του οχήματος  $a$  και τρία γυροσκόπια για την εκτίμηση του προσανατολισμού  $\theta$ . Ο λόγος που χρησιμοποιούνται τρεις αισθητήρες ανά περίπτωση είναι διότι κάθε μία μεμονωμένη διάταξη είναι υπεύθυνη για την μέτρηση για έναν και μόνο άξονα. Ομαδοποιώντας λοιπόν σε τριάδες μία IMU είναι σε θέση να μετρήσει την θέση και τον προσανατολισμό στο τρισδιάστατο επίπεδο. Στην συνέχεια τα δεδομένα του προσανατολισμού χρησιμοποιούνται έτσι ώστε το διάνυσμα της βαρύτητας να

μπορεί να εκτιμηθεί και να αφαιρεθεί από τη τελική μέτρηση. Τέλος οι τιμές που προκύπτουν ολοκληρώνονται μία ή δύο φορές σε περίπτωση που θέλουμε την ταχύτητα ή την θέση του οχήματος αντίστοιχα.



Εικόνα 10. Σχηματικό που περιγράφει την γενική λειτουργική δομή μίας IMU. Πηγή [21]

Όσον αφορά τον υπολογισμό του προσανατολισμού της ρομποτικής πλατφόρμας χρησιμοποιήθηκε ένας αισθητήρας IMU από την εταιρεία CH Robotics. Το μοντέλο UM6 αποτελεί έναν εξαιρετικά μικρό αισθητήρα τεχνολογίας MEMS ο οποίος συνδυάζει μετρήσεις επιταχυνσιόμετρων, γυροσκοπίων και μαγνητικούς αισθητήρες για να μετρήσει τον προσανατολισμό με συχνότητα έως και 500 Hz. [22] Ο προσανατολισμός του αισθητήρα αναφέρεται είτε με τεταρτημόρια είτε με γωνίες Euler σε σειριακή διεπαφή TTL με προσαρμόσιμες τιμές από τον χρήστη.

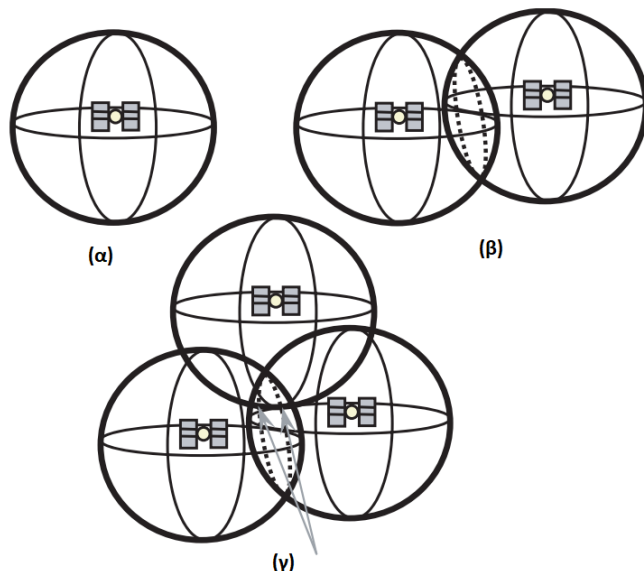


Εικόνα 11. Μονάδα Αδρανειακής Μέτρησης με ονομασία UM6 από την εταιρεία CH Robotics. Πηγή [22]

### 2.3.3. GPS-RTK

Μία διαδεδομένη προσέγγιση για την επίλυση του προβλήματος Αυτοεντοπισμού βασίζεται στην χρήση του συστήματος GPS (Global Positioning System), για προσδιορίσει την θέση της με μεγάλη ακρίβεια στον παγκόσμιο χάρτη. Το Παγκόσμιο Σύστημα Προσδιορισμού Γεωγραφικής Θέσης ή GPS (Global Positioning System) είναι ένα σύστημα ραδιοπλοήγησης, το οποίο αποτελείται από ένα δίκτυο 24 δορυφόρων και από επίγειους δέκτες καταναμημένους σε όλο τον κόσμο. Οι δορυφόροι βρίσκονται σε ύψος περίπου 20 χιλιομέτρων από την επιφάνεια της γης και περιφέρονται γύρω από αυτή ακολουθώντας έξι διαφορετικές τροχιές. Κάθε δορυφόρος μεταδίδει συνεχώς δεδομένα που υποδεικνύουν τη θέση του και την τρέχουσα ώρα. Σημαντικό γεγονός αποτελεί ότι, όλοι οι δορυφόροι GPS συγχρονίζουν τις μεταδόσεις τους έτσι ώστε τα σήματα τους να αποστέλλονται ταυτόχρονα. Όταν ένας δέκτης GPS δέχεται ένα σήμα από δύο ή περισσότερους δορυφόρους, υπολογίζει την διαφορά του χρόνου άφιξης του εκάστοτε σήματος και κατ' επέκταση την απόσταση του από αυτόν για μία δεδομένη στιγμή. Συνδυάζοντας την πληροφορία για τον χρόνο άφιξης αλλά και την στιγμιαία θέση τεσσάρων δορυφόρων, ο δέκτης είναι σε θέση να συμπεράνει την δική του θέση χρησιμοποιώντας μία τεχνική που ονομάζεται τριγωνισμός. Ο τριγωνισμός είναι μία μέθοδος προσδιορισμού της θέσης με βάση την απόσταση από άλλα σημεία ή αντικείμενα των οποίων η απόσταση θεωρείται γνωστή. [23]

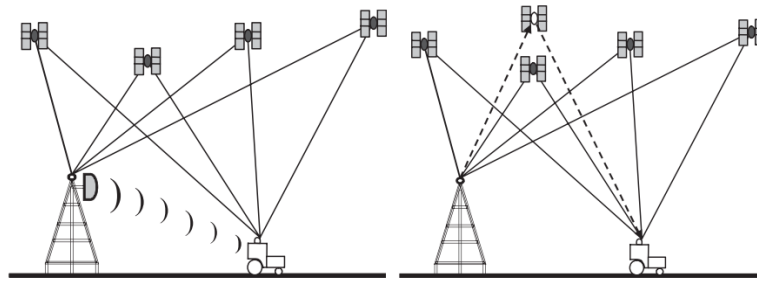
Όπως φαίνεται και από την **Εικόνα 12(α)**, όταν ένας δέκτης δεχτεί κάποιο σήμα από έναν δορυφόρο, είναι σε θέση να καθορίσει την θέση εντός μίας νοητής σφαίρας. Για να μπορέσει να περιορίσει περαιτέρω τη θέση του, ο δέκτης θα πρέπει να μετρήσει την απόσταση του και από έναν δεύτερο δορυφόρο. Αυτή η δεύτερη μέτρηση απόστασης θα τοποθετήσει τον δέκτη στην τομή των δύο νοητών σφαιρών όπως φαίνεται στην **Εικόνα 12(β)**. Τέλος η χρήση ενός τρίτου δορυφόρου θα μειώσει τον αριθμό των πιθανών τοποθεσιών όπως φαίνεται στην **Εικόνα 12(γ)**, εκ' των οποίων συνήθως το ένα από τα δύο σημεία μπορεί να εξαλειφθεί διότι δεν βρίσκεται πάνω στην επιφάνεια της Γης. [23]



Εικόνα 12. Προσδιορισμός της θέσης με χρήση ενός δορυφόρου (α), με χρήση δύο δορυφόρων (β) και τρεις δορυφόρους (γ). Πηγή [23].

Αποτελεί γεγονός ότι, ένα σύστημα GPS ενέχει πολλά σφάλματα κατά τον υπολογισμό της θέσης, πράγμα που επηρεάζει και περιορίζει την ακρίβεια της μέτρησης. Τα πιο συνηθισμένα σφάλματα περιλαμβάνουν σφάλματα ρολογιού (clock errors), σφάλματα ημερολογίου (ephemeris errors), δορυφορική διαμόρφωση (satellite configurations), Ατμοσφαιρική Παρεμβολή (Atmospheric Interference), σφάλματα πολλαπλών διαδρομών (Multipath Errors) [23], τα οποία επηρεάζουν την μέτρηση και υποβαθμίζουν την ακρίβεια του καθορισμού της θέσης.

Για την μέτρηση της θέσης με μεγαλύτερη ακρίβεια, οι περισσότεροι δέκτες GPS χρησιμοποιούν το Διαφορικό GPS (Differential GPS - DGPS). Ένας δέκτης DGPS χρησιμοποιεί πληροφορίες από έναν ή περισσότερους δέκτες GPS οι οποίοι ονομάζονται “**Σταθμοί Βάσης**” (Base Station) οι οποίοι βρίσκονται σε σταθερές και γνωστές θέσεις. Υπάρχουν δύο τύποι για την δημιουργία ενός Σταθμού Βάσης ανάλογα με τον τρόπο εκπομπής του σήματος [23]. Στην πρώτη κατηγορία οι διαφορικές διορθώσεις μεταδίδονται από έναν σταθμό βάσης τοπικής περιοχής, δηλαδή από χερσαίους ραδιοφωνικούς πύργους οι οποίοι συλλέγουν πληροφορίες και στην συνέχεια τις μεταδίδουν το διορθωτικό σήμα. Στην δεύτερη κατηγορία το διορθωτικό σήμα μεταδίδεται από ένα δίκτυο συνδεδεμένων τοπικών σταθμών οι οποίοι συλλέγουν πληροφορίες για την θέση τους και της στέλνουν σε ένα κεντρικό δέκτη. Ο κεντρικός δέκτης συλλέγει αυτές τις πληροφορίες και στέλνει διορθωτικό σήμα με σε ένα ευρύ φάσμα τοπικής κάλυψης [24].



Εικόνα 13. Τοπική διόρθωση (αριστερά) και διόρθωση ευρέως κάλυψης (δεξιά) σε συστήματα DGPS. Πηγή [23]

Ένας δέκτης που παίζει τον ρόλο του Σταθμού Βάσεως, αρχικά υπολογίζει μία θέση από τα διάφορα δορυφορικά σήματα τα οποία δέχεται. Γνωρίζοντας στην πραγματικότητα την απόλυτη θέση του, ο σταθμός βάσης, είναι σε θέση να υπολογίσει το σφάλμα μεταξύ της πραγματικής (απόλυτης) θέσης σε σχέση με την μετρούμενη που δημιουργήθηκε από τα διάφορα δορυφορικά σήματα. Αυτό το σφάλμα μεταδίδεται στην συνέχεια, σε ένα δεύτερο δέκτη με ονομασία “**Κινητός Δέκτης**” ή **GPS Rover Receiver**. Έχοντας λάβει αυτήν την διόρθωση ο Κινητός Δέκτης, έχει την δυνατότητα να υπολογίσει με μεγαλύτερη ακρίβεια την θέση του στον παγκόσμιο χάρτη.

Μία κατηγορία δεκτών που χρησιμοποιούν την τεχνική του διαφορικού GPS (DGPS) είναι αυτοί με ονομασία “**Real Time Kinematic**”(RTK) και συναντώνται πολύ συχνά στον τομέα της ρομποτικής λόγω της ακρίβειας των μετρήσεων που προσδίδουν αλλά και της δυνατότητας τους να δουλεύουν σε πραγματικό χρόνο [25]. Αυτή η κατηγορία των δεκτών μπορεί να δημιουργήσει ένα τοπικό σύστημα βάσης και κινητού δέκτη χρησιμοποιώντας δύο ίδιους δέκτες. Ο κινητός δέκτης τοποθετείται πάνω στην γεωργικό ρομποτικό μηχανισμό και ο σταθερός δέκτης παρέχει διορθωτικό σήμα σε πραγματικό χρόνο, παρέχοντας έτσι μεγάλη ακρίβεια στην μέτρηση. Σαν μία γενική παραδοχή σε ένα σύστημα RTK υποθέτουμε ότι το σφάλμα εντοπισμού που εμφανίζεται δέκτη βάσης επηρεάζει τον κινητό δέκτη με τον ίδιο τρόπο. Σε ένα RTK, για να αποκτηθεί μια “σταθερή” ή “fixed” λύση, με ακρίβεια επιπέδου εκατοστόμετρου, τουλάχιστον πέντε δορυφόροι πρέπει να είναι προσβάσιμοι στον σταθμό βάσης αλλά και στο κινητό δέκτη ταυτόχρονα.

Σε αυτήν την διπλωματική εργασία χρησιμοποιήθηκαν δύο δέκτες GPS για την δημιουργία ενός συστήματος RTK. Η ονομασία των δεκτών είναι “Reach RS2” [26] και δημιουργήθηκαν από την εταιρεία Emlid. Κάθε μονάδα αποτελείται από δύο πανομοιότυπες μονάδες επεξεργασίας οι οποίες μπορούν να ρυθμιστούν για να έχουν τον ρόλο της βάσης και του κινητού δέκτη αντίστοιχα. Κάθε δέκτης σε επίπεδο αποτελείται από ένα GNSS τσιπ “GPS NEO-M8T” [27] από την εταιρεία U-Blox συνδεδεμένο με μια μονάδα επεξεργασίας “Intel Edison” [28] οι οποίες εσωκλείονται από ένα ειδικά σχεδιασμένο περίβλημα. Όσον αφορά το λογισμικό εντοπισμού και διόρθωσης το οποίο εκτελείται σε πραγματικό χρόνο, το σύστημα χρησιμοποιεί ένα πακέτο ανοιχτού κώδικα με ονομασία “RTKLIB” [29]. Το RTKLIB αποτελείται από πολλές βιβλιοθήκες γραμμένες σε γλώσσα C που παρέχουν τις κύριες λειτουργίες πλοήγησης, αλγόριθμους εκτίμησης σχετικής θέσης, υπολογισμός θέσης με RTK με ακέραιη ανάλυση σφάλματος κ.α. Με αυτήν την βιβλιοθήκη το συγκεκριμένο σύστημα επιτρέπει να γίνουν διορθώσεις σε πραγματικό χρόνο. Τέλος αξίζει να σημειωθεί ότι οι δορυφόροι που υποστηρίζονται είναι οι: GPS, GLONASS, GALILEU, QZSS, BeiDou και SBAS.



Εικόνα 14. Δέκτης GPS με ονομασία Reach RS2 από την εταιρεία Emlid. Πηγή [21]

### 2.3.4. LiDAR

Οι αισθητήρες ενεργού εύρους (active range) αποτελούν μία από τις πιο δημοφιλείς κατηγορίες αισθητήρων στον τομέα της ρομποτικής. Αυτού του τύπου οι αισθητήρες αποτελούν ιδανικές λύσεις για την ανίχνευση και την αποφυγή εμποδίων όπως επίσης και για την επίλυση του προβλήματος του Αυτοεντοπισμού. Οι αισθητήρες ενεργού εύρους χρησιμοποιούν συνήθως την τεχνική που ονομάζεται “Χρόνος Πτήσης” ή “Time-of-flight” (ToF) για την ανίχνευση αντικειμένων στο χώρο. Η αρχή ToF είναι μια μέθοδος μέτρησης της απόστασης που μετράει την χρονική διαφορά μεταξύ της εκπομπής ενός σήματος και της επιστροφής του όταν αυτό ανακλάται σε ένα αντικείμενο. Μπορούν να χρησιμοποιηθούν διάφοροι τύποι με τους πιο δημοφιλείς να είναι ο ήχος και το φως. Εν γένει, η απόσταση που χρειάζεται, από την εκπομπή έως και την επιστροφή, ενός σήματος ήχου ή ηλεκτρομαγνητικού κύματος δίνεται από τον τύπο:

$$d = c * t \quad (2.8)$$

Όπου  $d$ , διανυθείσα απόσταση (συνήθως μετ' επιστροφής)

$c$ , ταχύτητα διάδοσης κυμάτων  $t$

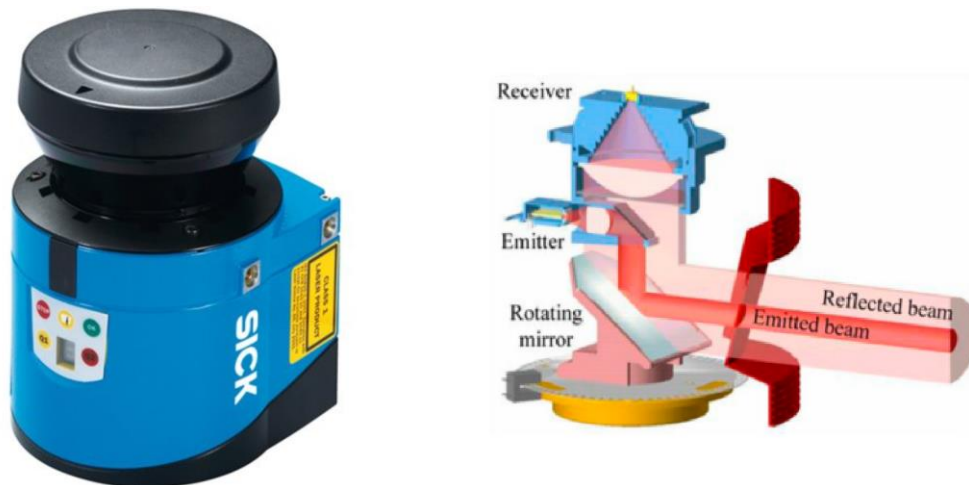
$t$ , χρόνος πτήσης

Είναι σημαντικό να επισημανθεί ότι η ταχύτητα διάδοσης  $v$  του ήχου είναι περίπου  $0,3 \frac{m}{ms}$ , ενώ η ταχύτητα των ηλεκτρομαγνητικών σημάτων είναι  $0,3 \frac{m}{ns}$ , η οποία όπως βλέπουμε είναι ένα εκατομμύριο φορές ταχύτερη. Ο χρόνος πτήσης για μια τυπική απόσταση, ας πούμε 3 μέτρα, είναι  $10 ms$  για ένα σύστημα υπερήχων, αλλά μόνο  $10 ns$  για έναν ανιχνευτή λέιζερ [12]. Στην παρούσα διπλωματική θα παρουσιάσουμε μόνο τους αισθητήρες που χρησιμοποιούν το φως για την ανίχνευση αντικειμένων.

Οι ανιχνευτές λέιζερ, οι οποίοι αναφέρονται και στην βιβλιογραφία και ως “**Οπτικό Ραντάρ**” ή “**LiDAR**” (Light Detection and Ranging), αποτελούν τυπικό παράδειγμα αισθητήρων που λειτουργούν με την μέθοδο ToF και η δομή τους συνίσταται από έναν πομπό ο οποίος φωτίζει έναν στόχο με μια συσσωρευμένη δέσμη (π.χ. λέιζερ), και έναν δέκτη ικανό να ανιχνεύει αυτή την δέσμη όταν ανακληθεί από τον στόχο. Η ταξινόμηση των οργάνων LiDAR μπορεί να είναι ευρεία και υποκειμενική, και εξαρτάται με το πλαίσιο εφαρμογής. Παρ' όλα αυτά, αυτό το όργανο ταξινομείται συνήθως σε τρεις κατηγορίες ανάλογα με τον τύπο πληροφορίας τον οποίο προσφέρει με ονομασίες “χωρική” (spatial), “φασματική” (spectral) και “χρονική” (temporal) [30]. Στην χωρική κατηγορία, ανήκουν οι αισθητήρες LiDAR οι οποίοι προσφέρουν πληροφορία σχετικά με την ανίχνευση αντικειμένων σε μία (1D), δύο (2D) ή τρεις (3D) διαστάσεις, χρησιμοποιώντας την τεχνική ToF. Οι χωρικές πληροφορίες είναι απαραίτητες για την κατασκευή ενός ακριβούς 3D χάρτη ενός περιβάλλοντος. Ωστόσο, οι χωρικές πληροφορίες δεν επαρκούν για εφαρμογές απαιτείται ανίχνευση συγκεκριμένων αντικειμένων. Στην δεύτερη κατηγορία οργάνων LiDAR ανήκουν οι αισθητήρες οι οποίοι έχουν την ικανότητα να μετρήσουν τις φασματικές πληροφορίες ενός υλικού όπως η ένταση επιστροφής λέιζερ (laser return intensity ή LRI). Το LRI αναφέρεται στην ανακλαστική ικανότητα που έχει το μήκος κύματος του μεταδιδόμενου παλμού από το αισθητήρα έως υλικό-στόχος [30]. Δεδομένου ότι το LRI είναι χαρακτηριστικό για έναν συγκεκριμένο τύπο υλικού, είναι δυνατά χρήσιμο για τον προσδιορισμό των επιφανειακών ιδιοτήτων ενός υλικού στόχου. Ωστόσο, για να αποφευχθούν αμφιβολίες στις αναγνώσεις LRI, απαιτούνται τουλάχιστον δύο μήκη κύματος λέιζερ. Επιπλέον, σε ορισμένες εφαρμογές απαιτείται λειτουργικότητα συλλογής χρονικών πληροφοριών εκτός από χωρικών και φασματικών πληροφοριών. Αυτό μπορεί να επιτευχθεί χρησιμοποιώντας την επαναλαμβανόμενη τεχνική LiDAR. Το επαναλαμβανόμενο LiDAR είναι μια διαδικασία συλλογής χρονικών δεδομένων ενός περιβάλλοντος στόχου σε πεπερασμένο χρονικό διάστημα [31]. Οι χρονικές πληροφορίες είναι απαραίτητες για την κατανόηση δυναμικών διεργασιών όπως η ανάπτυξη των φυτών και η διάβρωση του εδάφους.

Στα πλαίσια αυτής της διπλωματικής χρησιμοποιήθηκε ένας διδιάστατος αισθητήρας LiDAR με ονομασία LMS 10-1000 κατασκευασμένος από την εταιρεία SICK. Ο συγκεκριμένος αισθητήρας LiDAR ανήκει στην πρώτη και χωρική αλλά και στην φασματική κατηγορία, καθώς παρέχει πληροφορίες σχετικά με την απόσταση ενός αντικειμένου όπως επίσης και πληροφορίες σχετικά με την ανακλαστική ικανότητα του αντικειμένου αυτού καθ' αυτού. Σε αυτό το όργανο, ένας καθρέφτης τοποθετημένος σε  $45^\circ$  χρησιμοποιείται για την εκτροπή της δέσμης λέιζερ. Αυτός ο καθρέφτης κινείται κυκλικά από έναν ηλεκτροκινητήρα εγκατεστημένο στο κάτω μέρος της συσκευής. Η πηγή και ο δέκτης λέιζερ χρησιμοποιούν τον ίδιο καθρέφτη για να μεταδίδουν και να λαμβάνουν παλμούς λέιζερ διαδοχικά. Κάποια από τα βασικά τεχνικά χαρακτηριστικά του LiDAR LMS 10-1000 αποτελούν το μέγιστο μήκος σάρωσης

στα 20 μέτρα, γωνία σάρωσης οριζόντια στις 270 μοίρες, γωνιακή ανάλυση(γωνία μεταξύ δύο διαδοχικών δεσμών) 0,05 μοίρες και μέγιστη συχνότητα σάρωσης στα 50 Hz [32].

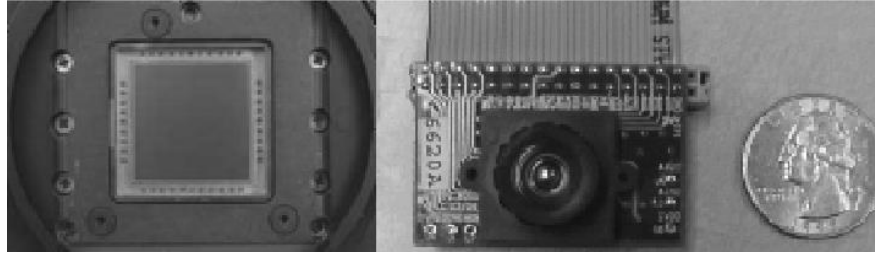


Εικόνα 15. LiDAR LMS 10-1000 από την εταιρεία Sick (αριστερά), Σχηματικό λειτουργίας (δεξιά).

### 2.3.5. Μονοσκοπική Όραση

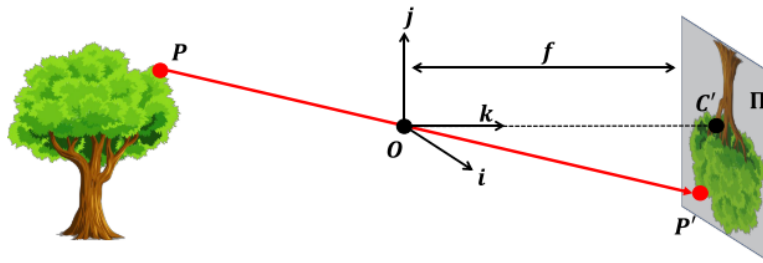
Στον κόσμο της ρομποτικής και πιο συγκεκριμένα της αντίληψης (perception), η όραση θεωρείται μία από τις πιο “ισχυρές” αισθήσεις. Ο λόγος είναι γιατί η όραση παρέχει μία πολύ μεγάλη ποσότητα πληροφοριών για το εξωτερικό περιβάλλον και επιτρέπει μία έξυπνη αλληλεπίδραση σε ένα δυναμικό περιβάλλον. Δεν αποτελεί έκπληξη λοιπόν, πως ένα μεγάλο μέρος της επιστημονικής κοινότητας έχει καταβάλει μία μεγάλη προσπάθεια για την δημιουργία αισθητήρων ικανών να μιμούνται τις δυνατότητες του ανθρώπινου συστήματος όρασης. Δύο από τους πιο συνηθισμένοι αισθητήρες όρασης αποτελούν δύο τεχνολογίες με ονομασία “**συσσκευή συζευγμένου φορτίου**” ή **charged-coupled device (CCD)** και **CMOS (Complementary metal-oxide-semiconductor)**.

Στην πρώτη κατηγορία CCD έχουμε μία συσκευή η οποία αποτελείται από μία σειρά από εικονοστοιχεία ή pixels, της τάξης από 20.000 μέχρι αρκετά εκατομμύρια τα οποία παρουσιάζουν μία ευαισθησία στο φως. Κάθε εικονοστοιχείο είναι στην πραγματικότητα ένας πυκνωτής ο οποίος μετατρέπει τα εισερχόμενα φωτόνια σε φορτία ηλεκτρονίων. Με την πάροδο του χρόνου, κάθε εικονοστοιχείο συσσωρεύει ένα διαφορετικό επίπεδο φόρτισης με βάση τον συνολικό αριθμό των φωτονίων που έχει συγκεντρώσει, όπου και καθορίζει και την φωτεινότητα του αντίστοιχου pixel στην συνολική εικόνα [12]. Στην δεύτερη κατηγορία, οι αισθητήρες CMOS αποτελούνται επίσης από μια σειρά pixel, αλλά για κάθε pixel υπάρχει ένα αντίστοιχο τρανζίστορ που στόχο έχει την αποθήκευση της τιμής του φορτίου για κάθε ένα pixel ξεχωριστά. Καθώς η τεχνολογία έχει αρκετά πλεονεκτήματα έναντι της πρώτης όπως λιγότερη κατανάλωση ενέργειας, οι αισθητήρες CCD συναντώνται συχνότερα λόγω της τεχνολογικής ωριμότητας λόγω της παλαιότητας τους.



Εικόνα 16 Ένας αισθητήρας CCD μεγέθους 2048 x 2048 (Αριστερά), Κάμερα CMOS με φακό (Δεξιά). Πηγή [12]

Πολύ σημαντικό κομμάτι της όρασης αποτελεί και η μοντελοποίηση μίας κάμερας όσον αφορά τον τρόπο λειτουργίας της. Η βασική αρχή λειτουργίας μία φωτογραφικής μηχανής βασίζεται στηρίζεται σε ένα μοντέλο με όνομα “**Κάμερα Σημειακής οπής**” (Pinhole Camera) ή “**Σκοτεινός Θάλαμος**” (Camera Obscura). Σε αυτό το μοντέλο σκοπός είναι η αναπαράσταση ενός τρισδιάστατου αντικειμένου σε μία δισδιάστατη απεικόνιση. Για να μπορέσει να επιτευχθεί αυτή η μετατροπή, στο συγκεκριμένο μοντέλο τοποθετείται ένα φράγμα με ένα μικρό άνοιγμα ανάμεσα στον αισθητήρα και στο τρισδιάστατο αντικείμενο το οποίο πρόκειται να καταγραφεί. Όπως φαίνεται και στην **Εικόνα 17**, παρότι κάθε σημείο του τρισδιάστατου αντικειμένου εκπέμπει πολλαπλές ακτίνες φωτός το φράγμα το οποίο τοποθετείτε επιτρέπει μόνο μία(ή μερικές) ακτίνες να περάσουν και να ανιχνευτούν από το φωτογραφικό φιλμ. Με αυτόν τον τρόπο μπορούμε και δημιουργούμε μία αντιστοίχιση 1 προς 1 μεταξύ των διαφόρων σημείων του αντικειμένου πάνω στο φωτογραφικό φιλμ [33]. Όπως βλέπουμε και στην παρακάτω εικόνα το φωτογραφικό φιλμ που αποτυπώνεται το τρισδιάστατο αντικείμενο ονομάζεται “**Εικόνα**” (Image) ή “**Επίπεδο Αμφιβληστροειδούς**” (Retinal Plane). Το διάφραγμα που αναφέραμε βρίσκεται στο σημείο  $O$  και βρίσκεται στο κέντρο της κάμερας, ενώ η απόσταση μεταξύ εικόνας και του σημείου  $O$  ονομάζεται “**Εστιακή Απόσταση**” (Focal Length) και συμβολίζεται με το γράμμα  $F$ .



Εικόνα 17. Μοντελοποίηση ενός φωτογραφικού συστήματος με ονομασία “Κάμερα Μικροσκοπικής Οπής”. Πηγή [33]

Σκοπός του μοντέλου μικροσκοπικής οπής όπως αναφέραμε είναι η προβολή ενός σημείου το οποίο τοποθετείται στον τρισδιάστατο χώρο με συντεταγμένες  $P = [x \ y \ z]^T$  πάνω στο δισδιάστατο επίπεδο της εικόνας με συντεταγμένες  $P' = [x' \ y']^T$  που ορίζεται από το σύστημα συντεταγμένων της κάμερας με αρχή το σημείο  $O$ . Βλέπουμε λοιπόν ότι από τα σημεία  $P' C' O$  δημιουργείται ένα τρίγωνο όμοιο με το τρίγωνο  $P, O$  και τον άξονα  $(0,0, z)$  και χρησιμοποιώντας τον νόμο όμοιων τριγώνων προκύπτει η εξίσωση:

$$P' = [x' \ y']^T = \left[ f \frac{x}{z} \quad f \frac{y}{z} \right]^T \quad (2.9)$$



που μας δίνει τις συντεταγμένες ενός τρισδιάστατου αντικειμένου πάνω σε μία δισδιάστατη εικόνα. Γράφοντας τον μετασχηματισμό σε μορφή Πίνακα προκύπτει ότι:

$$\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} f * x' \\ f * y' \\ z \\ 1 \end{bmatrix} \quad (2.10)$$

$$\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x' \\ y' \\ z \\ 1 \end{bmatrix} \quad (2.11)$$

$$x = P X \quad (2.12)$$

Όπου πίνακας μετασχηματισμού P, γνωστός και ως μήτρα προβολής κάμερας ο οποίος καθορίζεται πλήρως από τις παραμέτρους της κάμερας.

Σαν τελικό βήμα θα πρέπει αυτός ο μετασχηματισμός να αποτυπωθεί σε μία ψηφιακή δισδιάστατη εικόνα η οποία αποτελείται από διακριτά ψηφιακά κελιά τα οποία ονομάζονται **Εικονοστοιχεία (picture element ή pixel)**. Αυτός ο μετασχηματισμός δεν μπορεί να γίνει άμεσα καθώς το επίπεδο εικόνας δεν αντιστοιχεί άμεσα σε αυτό που βλέπουμε στην ψηφιακή εικόνα για διάφορους λόγους. Πρώτον, τα σημεία στις ψηφιακές εικόνες είναι, σε γενικές γραμμές, σε ένα διαφορετικό σύστημα συντεταγμένων από αυτά στο επίπεδο εικόνας. Δεύτερον, οι ψηφιακές εικόνες χωρίζονται σε διακριτά εικονοστοιχεία, ενώ τα σημεία στο επίπεδο εικόνας είναι συνεχή. Για να ληφθούν υπόψη αυτές οι διαφορές, θα εισαγάγουμε μια σειρά μετασχηματισμών που μας επιτρέπουν να χαρτογραφήσουμε οποιοδήποτε σημείο από τον κόσμο του 3D σε συντεταγμένες pixel.

Οι πρώτες παράμετροι που θα πρέπει να συμπεριληφθούν για τον μετασχηματισμό σε ψηφιακή απεικόνιση αποτελούν οι μεταβλητές  $c_x$  και  $c_y$  οι οποίες μεταφέρουν το κέντρο του συστήματος συντεταγμένων του επιπέδου εικόνας το οποίο βρίσκεται σε διαφορετικό σημείο σε σχέση με το σύστημα συντεταγμένων της ψηφιακής εικόνας το οποίο συνηθίζεται να βρίσκεται στο κάτω μέρος της εικόνας. Η δεύτερη παράμετρος που θα πρέπει να λάβουμε υπόψη είναι τα σημεία στις ψηφιακές εικόνες που όπως είπαμε εκφράζονται σε εικονοστοιχεία, ενώ τα σημεία στο επίπεδο εικόνας αντιπροσωπεύονται σε φυσικές μετρήσεις (π.χ. εκατοστά). Για να προσαρμόσουμε αυτήν την αλλαγή μονάδων, πρέπει να εισαγάγουμε δύο νέες παραμέτρους  $k$  και  $l$ . Σε ορισμένες περιπτώσεις μία εικόνα μπορεί να περιέχει μια επιπρόσθετη παράμετρο  $s$ , η οποία περιγράφει τη μη ορθογωνιότητα των αξόνων  $x$ ,  $y$ . Το φαινόμενο αυτό ονομάζεται στρέβλωση (skewness) και προκαλεί παραμόρφωση στην προβαλλόμενη εικόνα. Αυτές οι παράμετροι, των οποίων οι μονάδες είναι εικονοστοιχεία/cm και αντιστοιχούν στην αλλαγή των μονάδων μέτρησης στους δύο άξονες του επιπέδου εικόνας. Σύμφωνα με τα παραπάνω η **Εξίσωση 2.12** γίνεται:

$$\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} a & s & x_o & 0 \\ 0 & b & y_o & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x' \\ y' \\ z \\ 1 \end{bmatrix} \quad (2.13)$$

όπου  $a = f * k, b = f * l, x_0 = k * p_x$  και  $y_0 = l * p_y$ , με  $k$  και  $l$  τον αριθμό των εικονοστοιχείων ανά μονάδα απόστασης για τις διευθύνσεις  $x$  και  $y$ , αντίστοιχα

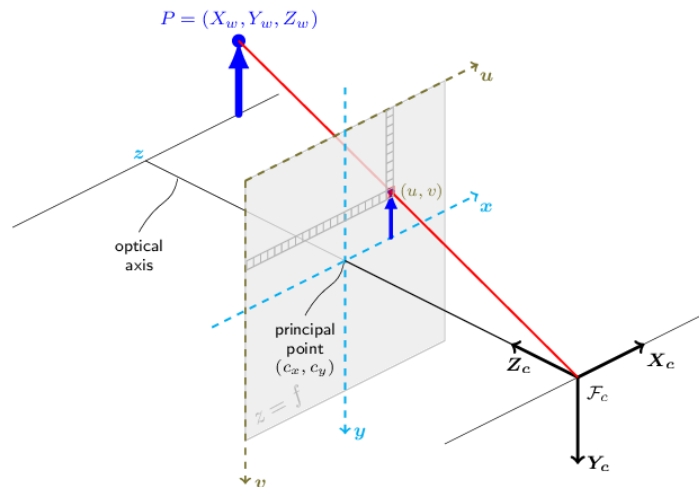
$$P' = \begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} a & s & x_0 \\ 0 & b & y_0 \\ 0 & 0 & 1 \end{bmatrix} [I \quad 0] \begin{bmatrix} x \\ y \\ z \end{bmatrix} = K [I \quad 0] P \quad (2.14)$$

$$P' = K [I \quad 0] P \quad (2.15)$$

Η μήτρα  $K$  αναφέρεται συχνά ως “**Πίνακας Κάμερας**”(Camera Matrix) και περιέχει όλες τις μοναδικές παραμέτρους που σχετίζονται με τις βασικές ιδιότητες της κάμερας που ονομάζονται “**Ενδογενείς Παράμετροι**”(Intrinsic Parameters).

Εκτός από τις εσωτερικές παραμέτρους, μία κάμερα χαρακτηρίζεται από ένα σύνολο παραμέτρων που ονομάζονται “**Εξωγενείς Παράμετροι**” (Extrinsic Parameters). Αυτές στην πραγματικότητα περιγράφουν ένα μετασχηματισμό που συνδέει ένα αντικείμενο που αποτυπώνουμε το οποίο βρίσκεται στο παγκόσμιο σύστημα αναφοράς, στο σύστημα αναφοράς της κάμερας. Αυτός ο μετασχηματισμός αποτυπώνεται από έναν πίνακα περιστροφής  $R$  και διάνυσμα μεταφοράς  $T$ . Εισάγοντας τις εξωτερικές παραμέτρους στην **Εξίσωση 2.15**, προκύπτει η τελική **Εξίσωση 2.16**, η οποία περιγράφει την χαρτογράφηση από ένα τρισδιάστατο σημείο  $P$  σε έναν αυθαίρετο παγκόσμιο σύστημα αναφοράς, στο σύστημα αναφοράς της κάμερας.

$$P' = K [R \quad T] P \quad (3) \quad (2.16)$$



Εικόνα 18. Μοντέλο Κάμερας Σημειακής Οπής.

### 2.3.6. Στερεοσκοπική Όραση

Στην προηγούμενη ενότητα παρουσιάστηκε ο τρόπος υπολογισμού τρισδιάστατων συντεταγμένων σε μία δισδιάστατη απεικόνιση χρησιμοποιώντας μία μεθοδολογία μονής απεικόνισης. Παρότι μπορούμε

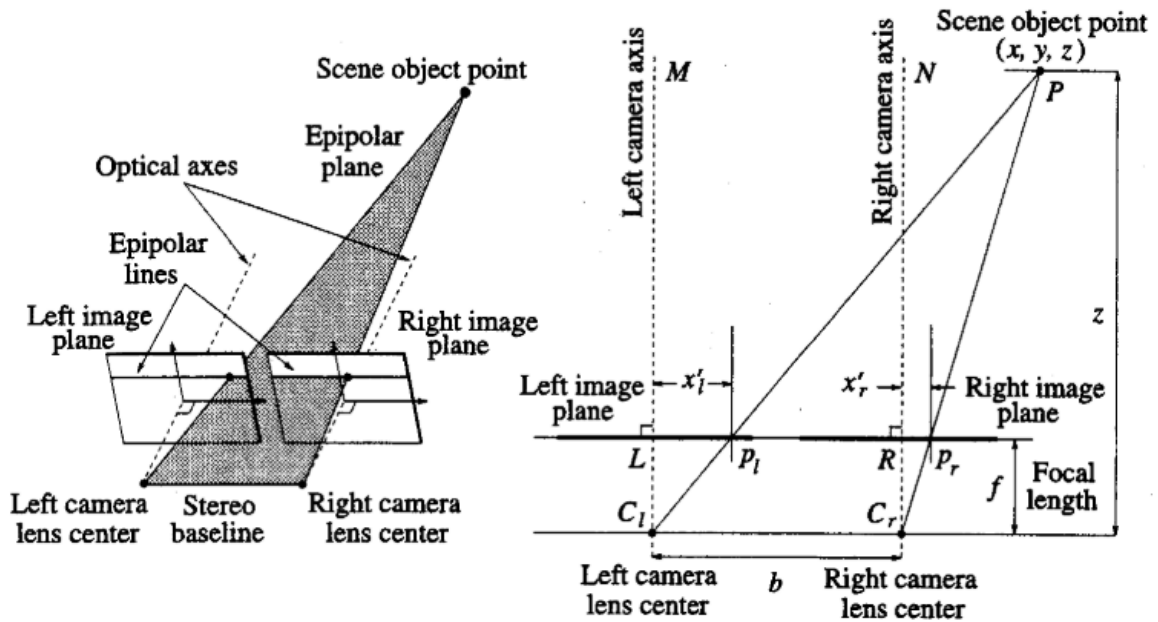
με αυτήν την διαδικασία να δημιουργήσουμε μία απεικόνιση του τρισδιάστατου κόσμου σε μία εικόνα, αποτελεί σύνθετο πρόβλημα η έλλειψη αίσθησης του πραγματικού βάθους με χρήση μίας λήψης. Αυτό οφείλεται κυρίως στην έλλειψη πληροφορίας μίας διάστασης όταν μεταφερόμαστε από τον τρισδιάστατο στο δισδιάστατο χώρο. Αξίζει να αναφέρουμε ότι υπάρχει και τεχνική λύση η οποία χρησιμοποιεί ένα φωτογραφικό καρέ και μέσω κατάλληλης ρύθμισης της γεωμετρίας της κάμερα, όπως θέση της εστίασης ή της ίριδας του φακού, μας δίνει πληροφορία απόστασης ενός αντικειμένου. Αυτή είναι η θεμελιώδης ιδέα πίσω από μία τεχνική που ονομάζεται “**Βάθος από την Εστίαση**”(Depth from focus) και το “**Βάθος από την Αποεστίαση**”(Depth from defocus). [12]



*Εικόνα 19. Μια μεμονωμένη λήψη όπως αυτή της εικόνας ενός ανθρώπου που κρατά ψηλά τον Πύργο της Πίζα οδηγεί σε διαφορετικά σενάρια όσον αφορά την αίσθηση του βάθους. Πηγή [34]*

Παρ' όλα αυτά, στα πλαίσια αυτής της διπλωματικής χρησιμοποιήθηκε διαφορετική τεχνική για την απόκτηση της αίσθησης του βάθους. Ένας πολύ δημοφιλής τρόπος επίλυσης αυτού του προβλήματος, αποτελεί η χρήση της στερεοσκοπικής όρασης στην οποία εν γένει χρησιμοποιούνται πολλαπλές εικόνες του ίδιου του αντικειμένου από διαφορετική οπτική, μέσω διαδοχικών λήψεων με μία φωτογραφική μηχανή ή χρησιμοποιώντας δύο διαφορετικές φωτογραφικές μηχανές.

Στην **Εικόνα 20** βλέπουμε το γενικευμένο γεωμετρικό μοντέλο μίας διοπτρικής κάμερας όπως περιγράφεται από τους Jain R. και άλλοι [35] και για την ανάλυση και επίλυση αυτού χρησιμοποιούμε όρους οι οποίοι πηγάζουν από μία γεωμετρία με όνομα “**Επιπολική Γεωμετρία**” (Epipolar Geometry).



Εικόνα 20. Γενικευμένο γεωμετρικό μοντέλο διοπτικής κάμερας. [35]

Σε αυτό το απλό μοντέλο έχουμε δύο πανομοιότυπες κάμερες που απέχουν μεταξύ τους κατά μήκος του άξονα  $x$ , μία απόσταση  $b$  η οποία ονομάζεται “**Γραμμή Βάσεων**” (**Baseline**). Αξίζει να αναφέρουμε ότι σε αυτό το μοντέλο, το αριστερό και δεξί επίπεδο εικόνας είναι ομοεπίπεδα. Επίσης το επίπεδο το οποίο ορίζεται από ένα σημείο  $P$  αλλά και τα κέντρα των δύο φωτογραφικών μηχανών ονομάζεται “**Επιπολικό Επίπεδο**” (**Epipolar Plane**) και οι γραμμές που ορίζονται από τη τομή του Επιπολικού Επιπέδου και των δύο Επιπέδων Εικόνας είναι γνωστές ως “**Επιπολικές Γραμμές**” (**Epipolar Lines**). Στην στερεοσκοπική όραση κάθε ένα σημείο  $P$  στον χώρο το οποίο μπορεί να εντοπιστεί από τις δύο κάμερες ταυτόχρονα, προβάλλεται σε διαφορετικές θέσεις σε κάθε επίπεδο εικόνας στην αριστερή και δεξιά φωτογραφική μηχανή. Αυτή η μετατόπιση μεταξύ των δύο θέσεων που βρίσκεται στο επίπεδο εικόνας σε κάθε φωτογραφική μηχανή ονομάζεται “**Ανομοιότητα**” ή στον αγγλικό όρο ως “**Disparity**”.

Στην Εικόνα 4, το σημείο  $P$  τον τρισδιάστατο χώρο παρατηρείται στα σημεία  $P_l$  και  $P_r$  στο αριστερό και δεξί επίπεδο εικόνας, αντίστοιχα. Συγκρίνοντας τα όμοια τρίγωνα τα οποία δημιουργούνται από τα σημεία  $PM C_l$  και  $p_r L C_l$ , παίρνουμε ότι:

$$\frac{x}{z} = \frac{x'_l}{f} \quad (2.17)$$

όπου  $f$  είναι η εστιακή απόσταση όπως περιγράφηκε στην προηγούμενη ενότητα.

Παρόμοια έχουμε για τα τρίγωνα  $PNC_r$  και  $p_r R C_r$ :

$$\frac{(x - b)}{z} = \frac{x'_r}{f} \quad (2.18)$$

όπου  $b$  είναι η Γραμμή Βάσεων, δηλαδή η απόσταση μεταξύ των δύο καμερών.

Οπότε συνδυάζοντας τις **Εξισώσεις 2.17** και **2.18** παίρνουμε την εξίσωση:

$$z = \frac{b * f}{(x'_l - x'_r)} \quad (2.19)$$

όπου η διαφορά  $(x'_l - x'_r)$  είναι η ανομοιότητα.

Όπως μπορούμε να δούμε έχοντας την ανομοιότητα (disparity), την γραμμή βάσεων (baseline) και την εστιακή απόσταση, είμαστε σε θέση να ανακτήσουμε πληροφορία σχετική με το βάθος του αντικειμένου. Συνοψίζοντας και τις τρεις εξισώσεις οι οποίες περιγράφουν την τρισδιάστατη θέση  $(x,y,z)$  ενός σημείου P στον χώρο έχουμε:

$$x = b \frac{(x'_l + x'_r)/2}{(x'_l - x'_r)}, y = b \frac{(y'_l - y'_r)/2}{(x'_l - x'_r)}, z = b \frac{f}{(x'_l - x'_r)} \quad (2.20)$$

Κάποιες σημαντικές παρατηρήσεις που πηγάζουν από αυτές τις τρεις εξισώσεις είναι :

1. Η απόσταση είναι αντιστρόφως ανάλογη της ανομοιότητας, πράγμα που σημαίνει ότι τα κοντινότερα αντικείμενα μπορούν να μετρηθούν με μεγαλύτερη ακρίβεια έναντι των μακρινότερων.
2. Η ανομοιότητα είναι ανάλογη της τιμής που παίρνει η γραμμή βάσεων  $b$  πράγμα που σημαίνει ότι όσον αυξάνει η μεταβλητή  $b$  τόσο πιο ακριβείς γίνεται η εκτίμηση του βάθους.
3. Με την μεγάλη αύξηση του  $b$  λόγο της μεγάλης απόστασης που μπορούν να πάρουν οι δύο κάμερες κάποια σημεία στο χώρο μπορεί να μην είναι φαίνονται ταυτόχρονα και στις δύο κάμερες.
4. Τέλος ένα τρισδιάστατο σημείο P το οποίο είναι ορατό και στις δύο κάμερες, παράγει ένα ζεύγος σημείων στο επίπεδο εικόνας (ένα σε κάθε φωτογραφικό φακό) γνωστό ως “**Συζευγμένο Ζεύγος**” (**Conjugate Pair**) .

Ένα εργαλείο που πηγάζει από την στερεοσκοπική όραση είναι η δημιουργία εικόνων οι οποίες ονομάζονται “**Εικόνες Βάθους**”(Depth Images). Οι εικόνες βάθους αποτελούνται από εικόνες στις οποίες κάθε pixel περιέχει πληροφορίες με την απόσταση των επιφανειών των αντικειμένων σε ένα φωτογραφικό καρέ.



Εικόνα 21. Εικόνα με Χρώμα (αριστερά), Εικόνα Βάθους (δεξιά).

Στον τομέα της ρομποτικής η στερεοσκοπική όραση είναι ένα πολύ σημαντικό εργαλείο για τον εντοπισμό και αποφυγή εμποδίων αλλά στην δημιουργία μίας τεχνικής που ονομάζεται οπτική οδομετρία την οποία αναλύεται σε επόμενο κεφάλαιο. Σε αυτήν την διπλωματική εργασία χρησιμοποιήθηκε μία κάμερα της οικογένειας Realsense D400 από την εταιρεία Intel [36]. Το μοντέλο της κάμερας είναι D435i και αποτελείται από μία RGB κάμερα και δύο ακόμα κάμερες που χρησιμοποιούνται για την δημιουργία μίας εικόνας βάθους. Ο τρόπος λειτουργίας όπως περιγράφεται στο τεχνικό εγχειρίδιο της κάμερας αποτελείται από βήματα τα οποία περιγράφονται στο συγκεκριμένο κεφάλαιο [37]. Ένα ακόμα χαρακτηριστικό του συγκεκριμένου μοντέλου είναι ότι έχει εγκατεστημένη μία IMU έτσι ώστε να εξάγει πληροφορίες σχετικές με το προσανατολισμό της κάμερας. Παρ' όλ' αυτά αυτή η δυνατότητα δεν χρησιμοποιήθηκε διότι η ρομποτική πλατφόρμα είχε ήδη εγκατεστημένη μία IMU όπως περιγράψαμε σε προηγούμενο κεφάλαιο.



Εικόνα 22 Διάφορες όψεις της κάμερας Intel D435i. Πηγή [37]

Η παρακάτω Εικόνα παρουσιάζει κάποια από τα βασικά τεχνικά χαρακτηριστικά όπως παρουσιάζονται από τον κατασκευαστή.

Use Environment	Indoor/Outdoor
Image Sensor Technolog	Global Shutter
Ideal Range	.3 m to 3 m
Depth Technology	Stereoscopic
Minimum Depth Distance (Min-Z) at Max Resolution	~28 cm
Depth Accuracy	<2% at 2 m
Depth Field of View (FOV)	87° x 58°
Depth Stream Output Resolution	Up to 1280 x 720
Depth Stream Output Frame Rate	Up to 90 fps
RGB Frame Resolution	1920 x 1080
RGB Frame Rate	30 fps
RGB Sensor Technology	Rolling Shutter
RGB Sensor FOV	69° x 42°
RGB Sensor Resolution	2 MP
Camera Module	Intel RealSense Module D430 + RGB Camera

*Εικόνα 23. Τεχνικά Χαρακτηριστικά κάμερα Intel Realsense D435i. Πηγή [36]*

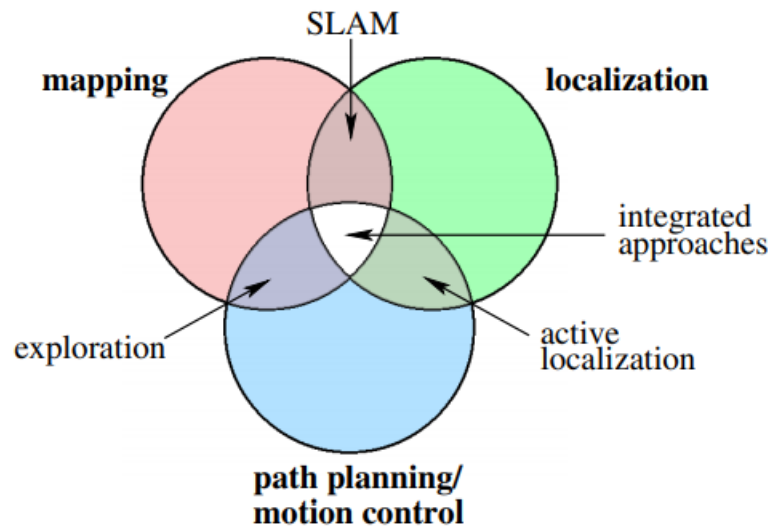
## Κεφάλαιο 3

### Αυτόνομη Πλοήγηση σε άγνωστο Περιβάλλον

#### 3.1 Αυτοεντοπισμός, Χαρτογράφηση και Σχεδιασμός Δρόμου

Σε αυτό το κεφάλαιο θα γίνει μία περιγραφή τριών εννοιών: “Αυτοεντοπισμός” (Localization), “Χαρτογράφηση” (Mapping) και “Σχεδιασμός Δρόμου” (Path Planning), εννοιών που αποτελούν ακρογωνιαίο λίθο για την ανάπτυξη ενός αυτόνομου ρομποτικού συστήματος. Ως επί το πλείστον, η ικανότητα της Χαρτογράφησης η οποία περιγράφεται στην βιβλιογραφία μπορεί να συμπυχθεί στην ερώτηση “Πώς φαίνεται ο κόσμος;”. Με άλλα λόγια, η διαδικασία της χαρτογράφησης αποτελεί τον τρόπο αποτύπωσης των διαφόρων αισθητήρων με τέτοιο τρόπο ο οποίος είναι κατανοητός και επεξεργάσιμος από ένα ρομπότ. Εν συνεχεία, η ικανότητα του “Αυτοεντοπισμού” μπορεί να απαντηθεί στην ερώτηση “Που βρίσκομαι;” και περιέχει όλες τις αλγοριθμικές δομές που επιτρέπουν μία ρομποτική πλατφόρμα να εκτιμά την θέση της κατά την λειτουργία της. Και τέλος, ο Σχεδιασμός Τροχιάς εμπεριέχει τις τεχνικές κατασκευής μίας τροχιάς από ένα αρχικό σε ένα επιθυμητό σημείο και βασίζονται στο ερώτημα “Πώς μπορώ να φτάσω σε μία επιθυμητή τοποθεσία” [38].

Θα αποτελούσε παράλειψη να μην επισημάνουμε ότι, κάθε μία από τις τρεις έννοιες οι οποίες και αναφέρθηκαν παραπάνω είναι άρρηκτα συνδεδεμένες μεταξύ τους και για την επίλυσή τους θα πρέπει να γίνει συνδυασμός τουλάχιστον δύο εκ’ των τριών. Όπως φαίνεται και στην **Εικόνα 24**, η εκάστοτε ονομασία προέρχεται από τον συνδυασμό των διαφόρων προβλημάτων που επρόκειτο να λύσουμε.



Εικόνα 24. Ονομασία κάθε προβλήματος ανάλογα με τον συνδυασμό των προβλημάτων “Σχεδίασης” “Εντοπισμού” και “Χαρτογράφησης”. Πηγή [38]

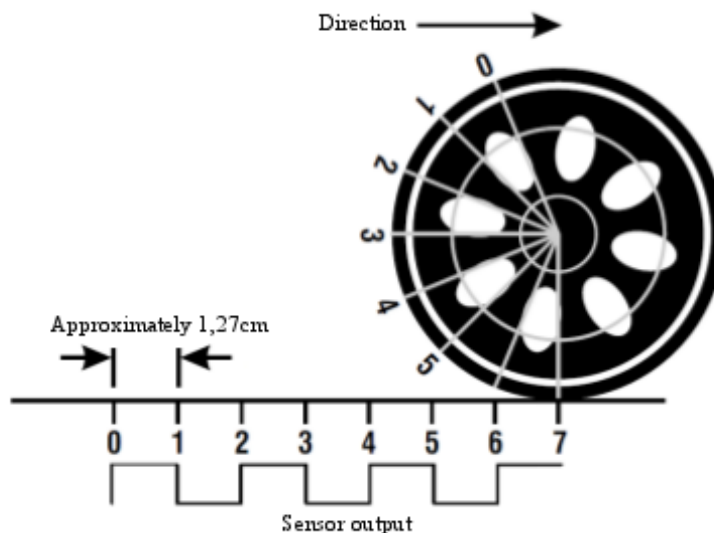
Ο Ταυτόχρονος Εντοπισμός και Χαρτογράφηση (Simultaneous Localization and Mapping - SLAM ) αποτελεί το πρόβλημα της δημιουργίας ενός χάρτη ενώ ταυτόχρονα εντοπίζει την θέση το ρομπότ μέσα σε αυτόν, αποτελεί ίσως και το δημοφιλέστερο πρόβλημα σε επίπεδο βιβλιογραφίας αλλά και απτών υλοποιήσεων [38]. Ενδεικτικά αναφέρουμε πως ο Ενεργός Εντοπισμός (Active Localization) επιδιώκει να καθοδηγήσει το ρομπότ σε τοποθεσίες εντός του χάρτη με στόχο την βελτίωση της εκτίμησης της θέσης



του αυτού καθ' αυτού. Εν αντιθέσει με την προσέγγιση της Εξερεύνησης (Exploration), οι οποίες προϋποθέτουν την ακριβή θέση με στόχο την αποτελεσματική καθοδήγηση του ρομπότ για τη δημιουργία έγκυρου χάρτη. Τέλος, ο συνδυασμός και των τριών προβλημάτων ο οποίος έχει ως στόχο την ταυτόχρονη και των τριών ονομάζεται Ταυτοχρονή Σχεδίαση, Εντοπισμός και Χαρτογράφηση (Simultaneous Planning Localization and Mapping –SPLAM). Η παρούσα διπλωματική εργασία, εστιάζει στον τρόπο επίλυσης του Αυτοεντοπισμού χρησιμοποιώντας δύο υλοποιήσεις οι οποίες διαφέρουν ως προς τις αισθητήριες διατάξεις που χρησιμοποιήθηκαν.

## 3.2.Οδομετρία

Η Οδομετρία αποτελεί μία μέθοδο υπολογισμού της μεταβολής της θέσης αλλά και του προσανατολισμού ενός ρομποτικού οχήματος. Μία οδομετρία μπορεί να δημιουργηθεί χρησιμοποιώντας Ιδιοδεκτικούς ή Εξωδεκτικούς αισθητήρες ανάλογα με τις απαιτήσεις της εκάστοτε ρομποτικής πλατφόρμας. Οι πιο δημοφιλείς μέθοδοι οδομετρίας, σε ρομποτικές εφαρμογές, είναι η οδομετρία τροχών (Wheel Odometry) και η οπτική οδομετρία (Visual Odometry). Η οδομετρία τροχών χρησιμοποιεί τις μετρήσεις των ταχυτήτων των τροχών χρησιμοποιώντας οπτικούς κωδικοποιητές (**Παράγραφος 2.3.1**) ενός ρομποτικού οχήματος για την εξαγωγή της ταχύτητας αυτού και μέσω ολοκλήρωσης της ταχύτητας, υπολογίζει την μεταβολή της πόζας του οχήματος.



Εικόνα 25 Μέτρηση ταχύτητας τροχών για τον προσδιορισμό της θέσης.

Ωστόσο αποτελεί γεγονός ότι η οδομετρία εισάγει αβεβαιότητα ως προς τον υπολογισμό της θέσης μίας ρομποτικής πλατφόρμας. Κύρια προβλήματα μίας οδομετρίας των τροχών αποτελούν η τραχύτητα του εδάφους αλλά και η ολίσθηση των τροχών πρόβλημα το οποίο είναι συνηθέστερο σε ρομποτικές πλατφόρμες που λειτουργούν σε εξωτερικό έδαφος για αγρονομικές λειτουργίες. Για να οριοθετηθεί η αβεβαιότητα της θέσης μίας κινητής ρομποτικής πλατφόρμας, συνήθης πρακτική αποτελεί η χρήση των Επεκταμένων Φίλτρων Κάλμαν το οποία και παρουσιάζονται στην παρούσα διπλωματική στην **Παράγραφο 3.4.**

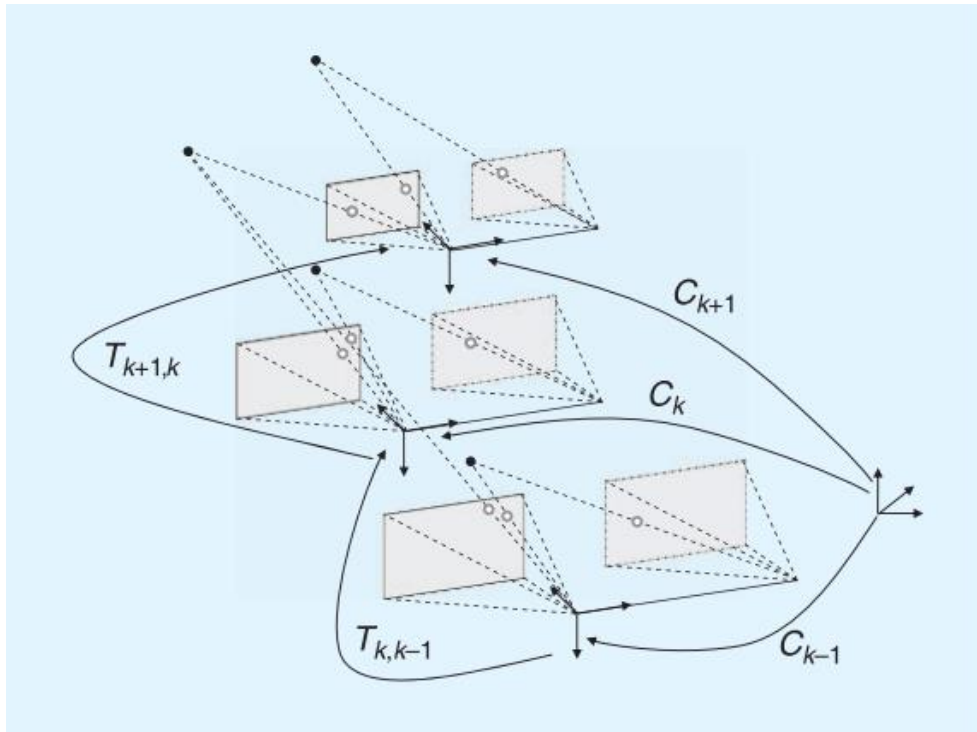
### 3.3. Οπτική Οδομετρία

Μία διαφορετική προσέγγιση για την δημιουργία μίας οδομετρίας αποτελεί η “**Οπτική Οδομετρία**” ή **Visual Odometry (VO)** η οποία είναι μία μέθοδος για τον υπολογισμό της θέσης αλλά και του προσανατολισμού μίας ρομποτικής πλατφόρμας υπολογίζοντας τις μεταβολές μεταξύ διαδοχικών εικόνων χρησιμοποιώντας μία κάμερα. Για να λειτουργήσει αποτελεσματικά η VO, θα πρέπει να υπάρχει επαρκής φωτισμός στο περιβάλλον, όπως επίσης το περιβάλλον λειτουργίας της πλατφόρμας να έχει πλούσια οπτική πληροφορία ικανή για να εξάγουμε αποτελέσματα για την κίνηση της κάμερας. Ένα από τα σημαντικότερα πλεονεκτήματα μίας οπτικής οδομετρίας αποτελεί το γεγονός ότι μπορεί να χρησιμοποιηθεί σε περιβάλλοντα τα οποία δεν έχουν κάλυψη του σήματος GPS, πράγμα πολύ σύνηθες σε ένα αγροτικό περιβάλλον όπως αναφέρθηκε και νωρίτερα.

Το πρόβλημα της οπτικής οδομετρίας περιγράφεται από μία ρομποτική πλατφόρμα η οποία κινείται σε ένα άγνωστο περιβάλλον και λαμβάνει συνεχόμενα διαδοχικές εικόνες, από μία κάμερα η οποία είναι προσαρτημένη πάνω της, σε διακριτές χρονικές στιγμές  $k$  οι οποίες συμβολίζονται ως  $I_{0:k} = \{I_0, \dots, I_k\}$ . Για κάθε νέα εικόνα  $I_k$  τα δύο πρώτα βήματα μίας οπτικής οδομετρίας αποτελούνται από την ανίχνευση αλλά και την αντιστοίχιση διαφόρων χαρακτηριστικών με αυτά των προηγούμενων καρέ. Σαν επόμενο βήμα σκοπός της οπτικής οδομετρίας είναι να υπολογίσει το μητρώο μετασχηματισμού

$$T_k = \begin{bmatrix} R_{k,k-1} & t_{k,k-1} \\ 0 & 1 \end{bmatrix} \quad (3.1)$$

όπου  $R_{k,k-1} \in SO(3)$  το οποίο είναι το μητρώο στροφής και όπου  $t_{k,k-1} \in R^{3 \times 1}$  είναι το διάνυσμα μετατόπισης, χρησιμοποιώντας της εικόνες  $I_k$  και  $I_{k-1}$  και σχηματίζοντας έτσι ένα δρόμο συνδέοντας τις διαφορετικές πόζες της κάμερας μέσω τον μετασχηματισμών  $T$ .



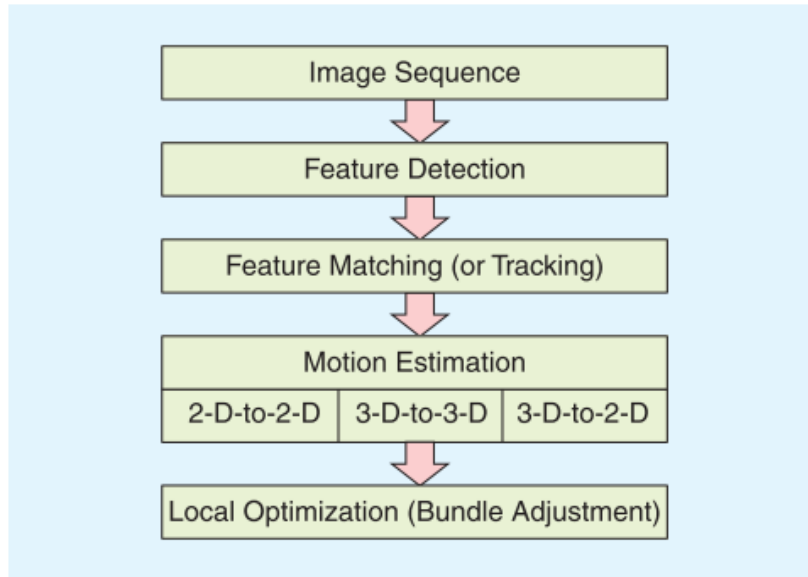
Εικόνα 26. Απεικόνιση του προβλήματος της Οπτικής Οδομετρίας. Πηγή [39]

Η οπτική οδομετρία αποτελεί ένα από τα δημοφιλέστερα προβλήματα στο τομέα της ρομποτικής, το οποίο απασχολεί ένα μεγάλο μέρος της επιστημονικής κοινότητας. Δεν αποτελεί έκπληξη λοιπόν ότι, οι προσεγγίσεις επίλυσης του συγκεκριμένου προβλήματος είναι πολλές και μπορούν να κατηγοριοποιηθούν σε αυτές που χρησιμοποιούν μονοσκοπική όραση αλλά και αυτές που χρησιμοποιούν στερεοσκοπική όραση όπως είδαμε και στη **Παράγραφο 2.3.5** και **Παράγραφο 2.3.6** αντίστοιχα.

Αυτές οι δύο κατηγορίες μπορούν να κατηγοριοποιηθούν περαιτέρω σε δύο βασικές προσεγγίσεις: την “Αντιστοίχιση Χαρακτηριστικών” (Feature Matching) δηλαδή αντιστοίχιση χαρακτηριστικών μεταξύ ενός προκαθορισμένου αριθμού φωτογραφικών εικόνων, την “Παρακολούθηση Χαρακτηριστικών” (Feature Tracking) δηλαδή αντιστοίχιση χαρακτηριστικών μεταξύ μόνο διαδοχικών φωτογραφικών εικόνων και της “Οπτικής Ροής” (Optical Flow) δηλαδή μίας τεχνικής που βασίζεται στην “ένταση” όλων των Εικονοστοιχείων ή μίας συγκεκριμένης περιοχής σε διαδοχικές εικόνες [39].

Παρ’ όλ’ αυτά μία οπτική οδομετρία ακολουθεί ένα σύνολο βημάτων όπου κάθε ένα μπορεί να υλοποιηθεί με διαφορετικές προσεγγίσεις. Τα βήματα που ακολουθούνται αναφέρονται ονομαστικά και φαίνονται στην **Εικόνα 27**:

1. **Εξαγωγή Χαρακτηριστικών Εικόνας (Extract Visual Features)**
2. **Εντοπισμός Χαρακτηριστικών (Feature Detection)**
3. **Ταυτοποίηση Χαρακτηριστικών (Feature Matching)**
4. **Εκτίμηση Κίνησης (Motion Estimation)**
5. **Τοπική Βελτιστοποίηση (Local Optimization)**



Εικόνα 27. Διάγραμμα Ροής που περιγράφει τα δομικά στοιχεία μίας Οπτικής Οδομετρίας. Πηγή [40]

### 3.3.1. Χαρακτηριστικά Εικόνας

Το πρώτο βήμα που ακολουθούν όλες οι προσεγγίσεις για την δημιουργία μίας οπτικής οδομετρίας, είναι η εξαγωγή διαφόρων **Χαρακτηριστικών (Image Features)** από κάθε φωτογραφία και η ταυτοποίηση αυτών στα επόμενα φωτογραφικά καρέ. Ως χαρακτηριστικό εννοούμε ένα μοτίβο εικόνας το οποίο διαφέρει από τα άμεσα γειτονικά χαρακτηριστικά του ως προς την ένταση, το χρώμα, τα χαρακτηριστικά ή την υφή του. Οι γωνίες αποτελούν σύνθετες παράδειγμα ενός χαρακτηριστικού, το οποίο καθορίζεται ως ένα σημείο του οποίου η τομή αποτελείται από δύο ή περισσότερες πλευρές. Ένα ακόμα γνωστό παράδειγμα ενός χαρακτηριστικού μίας εικόνας αποτελούν οι κηλίδες, οι οποίες είναι μοτίβα τα οποία διαφέρουν από τα γειτονικά σημεία τους ως προς την ένταση, το χρώμα ή την υφή [41]. Η βιβλιογραφία βρίθει προσεγγίσεων για την εύρεση χαρακτηριστικών σε μία εικόνα όπως είναι οι γωνίες (Frostner [42], Harris [43], Shi-Tomasi [44], FAST [45]) και οι κηλίδες (SIFT [46], SURF [47] και CENSURE [48]). Παρότι υπάρχουν διάφορες προσεγγίσεις για την εύρεση χαρακτηριστικών μέσα σε μία εικόνα, όλες ακολουθούν δύο κεντρικά βήματα όπως αναφέρεται στο [41]. Το πρώτο βήμα αποτελείται από την εφαρμογή μίας συνάρτησης μεταφοράς σε όλη την εικόνα, όπως εφαρμόζεται στον αλγόριθμο Harris ή στον αλγόριθμο SIFT που χρησιμοποιείται η διαφοράς Γκάους. Στην συνέχεια χρησιμοποιείται η τεχνική της μη-μέγιστης σύμμειξης (non-maxima suppression) η οποία έχει ως στόχο την επιλογή ενός μοναδικού χαρακτηριστικού σε μία γειτονιά πάνω στην εικόνα [49]. Τέλος η εύρεση τοπικών ελαχίστων (ή μεγίστων) στην εικόνα που έχει προκύψει στο τέλος του δεύτερου βήματος υποδηλώνει την ύπαρξη ενός χαρακτηριστικού [41].

Στην συνέχεια θα δώσουμε μία περιγραφή ενός πολύ γνωστού αλγορίθμου ο οποίος χρησιμοποιείται για την εξαγωγή γωνιών από ένα φωτογραφικό καρέ. Ο εν λόγω αλγόριθμος αναπτύχθηκε από τους J. Lie και C. Tomashi [44] και αποτελεί μία βελτιστοποίηση του αλγορίθμου από τους C G Harris and J M Pike [43]. Όσον αφορά την βασική λειτουργική δομή ο αλγόριθμος ψάχνει για γωνίες μέσα σε μικρά κομμάτια εικόνας τα οποία ονομάζονται "παράθυρα". Απομονώνοντας λοιπόν την εικόνα σε αυτό το μικρότερο

τετραγωνικό υποσύνολο κρατάμε στην ουσία μόνο τα pixel τα οποία βρίσκονται σε αυτό. Ακολουθώντας, μετακινούμε αυτό το χωρίο σε μία δεδομένη απόσταση πάνω στον άξονα και x και y και βλέπουμε την μεταβολή της έντασης των εικονοστοιχείων σε σχέση με αυτήν την μεταβολή. Η μαθηματική διατύπωση αυτής της μεταβολής φαίνεται στον παρακάτω τύπο:

$$E(u, v) = \sum_{x,y} w(x, y) [I(x + u, y + v) - I(x, y)]^2 \quad (3.2)$$

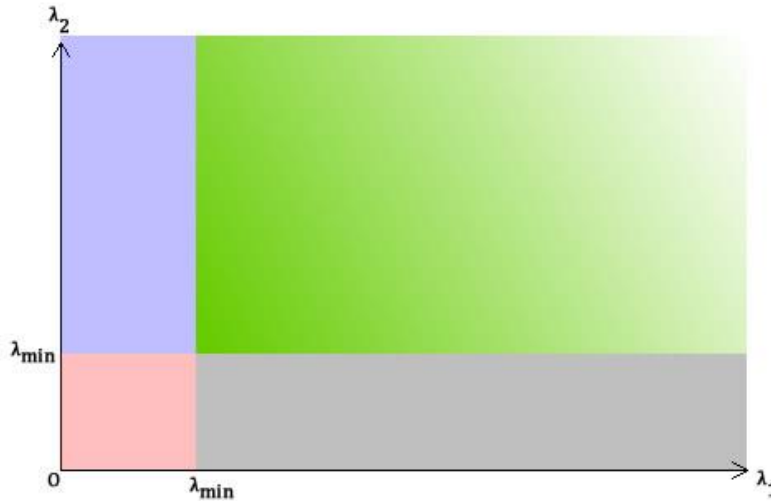
- Όπου E είναι η διαφορά μεταξύ του αρχικού από το τελικό παράθυρο
- u είναι η μετακίνηση παραθύρου πάνω στον x άξονα
- v είναι η μετακίνηση παραθύρου πάνω στον y άξονα
- $w(x, y)$  είναι το παράθυρο στη θέση (x, y). Αυτό λειτουργεί σαν μάσκα, βεβαιώνοντας ότι χρησιμοποιείται μόνο το επιθυμητό παράθυρο.
- I είναι η ένταση της εικόνας σε μια θέση (x, y)
- $I(x + u, y + v)$  είναι η ένταση του μετακινούμενου παραθύρου
- $I(x, y)$  είναι η ένταση του αρχικού παραθύρου

Χρησιμοποιώντας την σειρά Taylor για την γραμμικοποίηση του μη γραμμικού όρου  $[I(x + u, y + v) - I(x, y)]^2$  παίρνοντας μόνο τους τρεις πρώτους όρους αλλά και σε μορφή πίνακα η παραπάνω εξίσωση παίρνει την παρακάτω μορφή:

$$E \approx [u \quad v] \left( \sum \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \right) \begin{bmatrix} u \\ v \end{bmatrix} \quad (3.3)$$

$$E \approx [u \quad v] M \begin{bmatrix} u \\ v \end{bmatrix} \quad (3.4)$$

Για να μπορέσει να θεωρηθεί ένα καλό σημείο παρακολούθησης, θα πρέπει να αναλυθεί ο συμμετρικός πίνακας M. Από αυτόν μπορούμε να δούμε δύο βασικές προϋποθέσεις οι οποίες θα πρέπει να πληρούνται: να έχει χαμηλά επίπεδα θορύβου αλλά και να είναι καλώς ορισμένος [44]. Αυτό στην πραγματικότητα μεταφράζεται στο να έχει μεγάλες ιδιοτιμές ο Πίνακας M όπως επίσης δεν μπορούν να διαφέρουν στην τάξη μεγέθους. Αυτό σημαίνει ότι ανάλογα με την τιμή που παίρνουν οι τιμές των ιδιοτιμών μπορούμε να δούμε αν ένα σημείο θεωρείται καλό χαρακτηριστικό το οποίο θα πρέπει να λάβουμε υπόψιν. Στην **Εικόνα 28** τις διάφορες περιοχές των ιδιοτιμών οι οποίες πληρούν αυτήν την προϋπόθεση.



Εικόνα 28 Περιοχές ιδιοτιμών στις οποίες χαρακτηρίζεται ένα σημείο ως επιθυμητό χαρακτηριστικό. Πηγή [50]

Η μόνη περιοχή στην οποία οι ιδιοτιμές έχουν τιμές έχουν αποδεκτές είναι η πράσινη και άρα μπορεί να χαρακτηριστεί η περιοχή με τα pixel ως γωνία.

### 3.3.2. Περιγραφείς Χαρακτηριστικών

Στην συνέχεια για την ταυτοποίηση του εκάστοτε σημείου στα επόμενα φωτογραφικά καρέ χρησιμοποιείται μία τεχνική με ονομασία “**Περιγραφή Χαρακτηριστικού**” ή **Feature Descriptor**. Για να μπορεί να αποκτήσει μία μοναδική “ταυτότητα” κάθε χαρακτηριστικό, χρησιμοποιούνται πληροφορίες των γειτονικών σημείων του οι οποίες δίνουν μία μοναδικότητα σε καθένα με στόχο την ταυτοποίηση τους στην επόμενη φωτογραφία. Το πιο απλό παράδειγμα της δημιουργίας μιας περιγραφής ενός χαρακτηριστικού, αποτελούν οι γειτονικές τιμές που παίρνουν τα Εικονοστοιχεία αυτού καθ’ αυτού. Μία τέτοια πληροφορία δεν αποτελεί πολύ έγκυρη μοναδική περιγραφή ενός χαρακτηριστικού καθώς εξαρτάται σε μεγάλο βαθμό από τον προσανατολισμό, την κλίμακα αλλά και από το οπτικό πεδίο της λήψης της φωτογραφίας. Οι πιο γνωστές τεχνικές για την δημιουργία μιας περιγραφής ενός χαρακτηριστικού αποτελούν οι τεχνικές BRIEF [51], ORB [52] και BRISK [53] οι οποίες δίνουν μία γρήγορη και αρκετά αποδοτική λύση [41].

Εκτός από την ορθότητα που θα πρέπει να έχει ένας αλγόριθμος για τον υπολογισμό έγκυρων περιγραφέων, πολύ σημαντικός παράγοντας αποτελεί και η ταχύτητα του υπολογισμού αυτής. Σε μία κινητή ρομποτική πλατφόρμα η ταχύτητα αλλά και η υπολογιστική ισχύ ενός τέτοιου αλγορίθμου, αποτελεί ύψιστη προτεραιότητα λόγω της έλλειψης υπολογιστικών πόρων που συνήθως χρησιμοποιούνται σε τέτοιες μηχανές. Γι’ αυτούς τους λόγους χρησιμοποιούνται πολύ συχνά μία κατηγορία περιγραφέων με ονομασία “**Δυαδικοί Περιγραφείς**” (**Binary Descriptors**), οι οποίοι χρησιμοποιούν δυαδικούς αριθμούς για την απεικόνιση αυτής της περιγραφής. Η πιο γνωστή υλοποίηση αυτής της τεχνικής αποτελεί ο αλγόριθμος BRIEF [51] και στην συνέχεια θα δώσουμε μία σύντομη περιγραφή αυτού.

Ένας δυαδικός περιγραφέας όπως αναφέραμε αποτελεί μία μοναδική ταυτότητα ενός χαρακτηριστικού σε μία εικόνα. Για να μπορέσουμε να δημιουργήσουμε αυτήν την ταυτότητα λοιπόν θα πρέπει να χρησιμοποιήσουμε πληροφορίες που πηγάζουν από τα γειτονικά pixel αυτού του

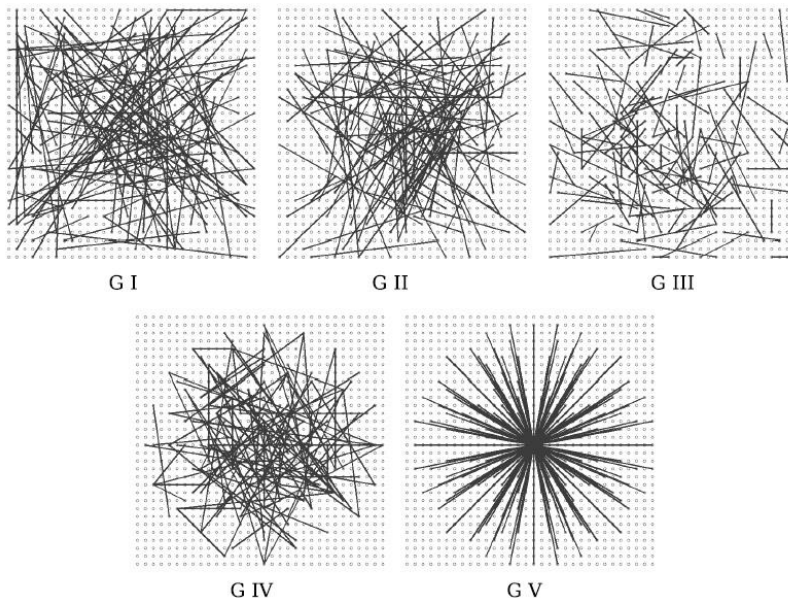
χαρακτηριστικού. Αυτή η τοπική περιοχή του χαρακτηριστικού συχνά αποκαλείται ως “Κομμάτια” (**Patches**) και στην πραγματικότητα αποτελείται από ένα σύνολο Εικονοστοιχείων σε μία εικόνα. Φυσικά κάθε ένα pixel αυτής της περιοχής αποτελείται από μία τιμή η οποία ανάλογα με το πόσο ανοιχτόχρωμο ή μη είναι αυτό το Εικονοστοιχείο αυξάνεται ή μειώνεται. Παίρνοντας λοιπόν μία γειτονιά από pixels μπορούμε να πούμε ότι παίρνουμε και έναν πίνακα συγκεκριμένων διαστάσεων όπου κάθε τιμή του περιέχει την τιμή της έντασης του κάθε pixel. Σαν μία πάγια τακτική θα μπορούσαμε να δημιουργήσουμε μία συνάρτηση όπου κάθε φορά που έρχεται ένα ζεύγος από δύο τιμές του πίνακα, η συνάρτηση να έχει έξοδο “1” ή “0” ανάλογα με το αν η πρώτη τιμή έχει μεγαλύτερη ένταση ή όχι αντίστοιχα. Σε μαθηματική μορφή αυτό αποτυπώνεται ως:

$$b = \begin{cases} 1, & \text{εάν } I(\text{pixel1}) < I(\text{pixel2}) \\ 0, & \text{αλλιώς} \end{cases} \quad (3.5)$$

Οπότε καταλαβαίνει κανείς ότι, αν χρησιμοποιήσουμε αυτήν την συνάρτηση παίρνοντας πολλά και διαφορετικά ζεύγη αυτού του Πίνακα καταλήγουμε να έχουμε σαν έξοδο έναν δυαδικό αριθμό. Αυτός ο δυαδικός αριθμός είναι και στην πραγματικότητα και ο δυαδικός περιγραφέας ο οποίος αναφέραμε και νωρίτερα. Κρίσιμο κομμάτι όπως καταλαβαίνουμε αποτελεί και η επιλογή της στρατηγικής για την δημιουργία ζευγών που θα εισαχθούν στην συνάρτηση. Όπως παρουσιάζεται από του συγγραφείς του BRIEF, εξετάστηκαν διαφορετικές στρατηγικές για την επιλογή αυτών των ζευγών οι οποίες συνθέτουν ένα δυαδικό αριθμό με μέγεθος 256 bits. Ενδεικτικά αναφέρουμε ότι οι στρατηγικές που δοκιμάστηκαν ως προς την επιλογή των σημείων ήταν [51]:

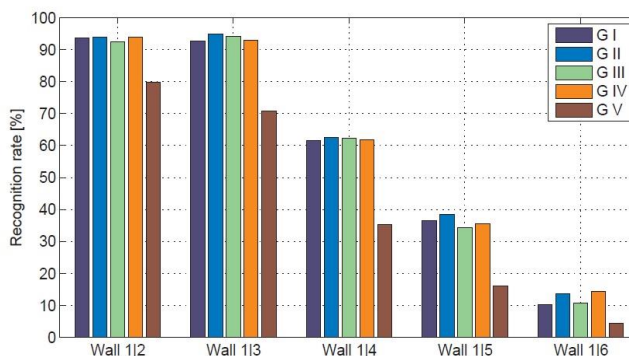
- i. τυχαία δειγματοληψία
- ii. δειγματοληψία με βάση την Γκαουσιανή κατανομή
- iii. δειγματοληψία με βάση την Γκαουσιανή κατανομή και διαφορετική επιλογή τιμής διακύμανσης για κάθε ζεύγος
- iv. δειγματοληψία που προκύπτει από την δημιουργία διακριτού πολικού πλέγματος,
- v. επιλογή όλων των πιθανών τιμών σε πολικό πλέγμα

Τα ακόλουθα σχήματα, απεικονίζουν τα αποτελέσματα από την ταυτοποίηση των χαρακτηριστικών με βάση την εκάστοτε στρατηγική.



Εικόνα 29 Αποτελέσματα από την ταυτοποίηση των χαρακτηριστικών με βάση την εκάστοτε στρατηγική. Πηγή [51]

Το παρακάτω σχήμα παρουσιάζει ποσοστά αναγνώρισης χρησιμοποιώντας και τις πέντε στρατηγικές δειγματοληψίας. Μπορούμε να δούμε ότι τα ποσοστά αναγνώρισης είναι περίπου τα ίδια, εκτός από την πέμπτη στρατηγική δειγματοληψίας που δείχνει χειρότερες επιδόσεις. Όπως συμβαίνει με όλους τους δυαδικούς περιγραφείς, το μέτρο απόστασης του BRIEF είναι ο αριθμός των διαφορετικών δυαδικών ψηφίων μεταξύ δύο δυαδικών συμβολοσειρών που μπορούν επίσης να υπολογιστούν ως άθροισμα της πράξης XOR μεταξύ των συμβολοσειρών.



Εικόνα 30. Ποσοστά αναγνώρισης χρησιμοποιώντας και τις πέντε στρατηγικές δειγματοληψίας. Πηγή [51]

### 3.3.3. Αντιστοίχιση Χαρακτηριστικών

Σε αυτό το βήμα με τελικό προορισμό την δημιουργία μίας οπτικής οδομετρίας, έχοντας τα χαρακτηριστικά αλλά και τους αντίστοιχους περιγραφείς τους θα πρέπει να γίνει μία αντιστοίχιση των χαρακτηριστικών σε διαδοχικά καρέ. Το σύνολο των εφαρμογών και προσεγγίσεων οι οποίες δίνουν λύση στο συγκεκριμένο πρόβλημα συναντάται πολύ συχνά στην βιβλιογραφία ως “**Αντιστοίχιση Δεδομένων**”(Data Association ή DA)[21].



Η πλειοψηφία των προσεγγίσεων, για την στρατηγική επίλυσης του DA, χρησιμοποιεί ως βασική προϋπόθεση ότι οι περιγραφείς που ανακτήθηκαν στο προηγούμενο βήμα έχουν σχεδιαστεί με τέτοιο τρόπο έτσι ώστε να μπορεί να χρησιμοποιηθεί η Ευκλείδεια απόσταση στο χωρικό επίπεδο των χαρακτηριστικών [54]. Αυτή η απόσταση αποτελεί κεντρικό εργαλείο για τον χαρακτηρισμό αλλά και την κατάταξη των πιθανών αντιστοιχιών που μπορεί να δημιουργηθούν από κάθε υλοποίηση.

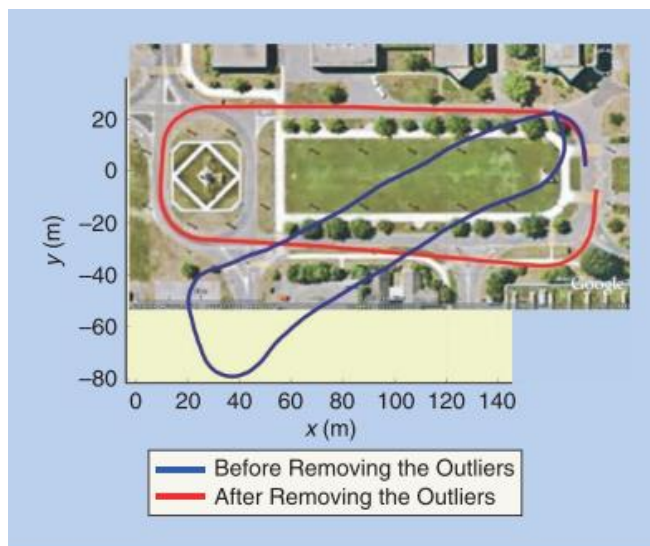
Δεδομένης της Ευκλείδειας μέτρησης της απόστασης, η απλούστερη στρατηγική DA αποτελεί η ύπαρξη ενός ορίου που εκφράζει την μέγιστη απόσταση στην οποία μπορούν να βρεθούν αντιστοιχίσεις των χαρακτηριστικών. Ο καθορισμός αυτού του ορίου εάν οριστεί με υψηλή τιμή είναι πιθανό να οδηγήσει σε πολλές ψευδής θετικές αντιστοιχίσεις ενώ αντίθετα η χαμηλή τιμή μπορεί να οδηγήσει στο να χαθούν πολλές θετικές εκτιμήσεις. Ιδανικά λοιπόν, αυτό το όριο θα πρέπει να προσαρμόζεται με τιμή ανάλογα με την περιοχή του χώρου χαρακτηριστικών. Για αυτόν τον σκοπό η χρήση ενός ευρετικού αλγορίθμου με ονομασία “**Λόγος Πλησιέστερου Γείτονα**” (**Nearest Neighbor Ratio**). Σε αυτήν κάθε όριο ρυθμίζεται με βάση τον λόγο της απόστασης του πλησιέστερου γείτονα με αυτή του δεύτερου πλησιέστερου γείτονα ως προς ένα συγκεκριμένο περιγραφέα, δηλαδή έχουμε [55]:

$$NNDR = \frac{d_1}{d_2} = \frac{|D_A - D_B|}{|D_A - D_C|} \quad (3.6)$$

όπου  $d_1, d_2$  είναι οι αποστάσεις του πρώτου και του δεύτερου πλησιέστερου γείτονα,  $D_A$  είναι ο περιγραφέας και  $D_B, D_C$  είναι οι δύο κοντινότεροι γείτονες.

## RANSAC

Όπως καταλαβαίνουμε σε αυτό το βήμα οι αντιστοιχίσεις των διαφόρων σημείων θα είναι είτε αληθείς (σημεία τα οποία ορθώς έχουν ταυτιστεί) ή ψευδείς (σημεία τα οποία λανθασμένα έχουν ταυτιστεί). Αποτελεί πολύ συνηθισμένο πρόβλημα ότι, ο αριθμός των ψευδών ταυτίσεων να αλλοιώνει και να διαφοροποιεί την συνολική επίδοση μίας οπτικής οδομετρίας. Για να μπορέσει να αντιμετωπιστεί αυτό το πρόβλημα χρησιμοποιείται μία αλγοριθμική δομή με όνομα “Συναίνεση Τυχαίου Δείγματος” (Random Sample Consensus ή πιο γνωστός με το ακρωνύμιο RANSAC), η οποία έχει ως στόχο την αφαίρεση των ψευδών αντιστοιχίσεων (outliers). Στην παρακάτω Εικόνα μπορούμε να δούμε ένα παράδειγμα για την βελτίωση της απόδοσης της οπτικής οδομετρίας σε ένα σενάριο αφαιρώντας ή μη τις ψευδείς αντιστοιχίσεις.



Εικόνα 31. Μπλε διαδρομή χωρίς την χρήση του RANSAC, κόκκινη διαδρομή με χρήση RANSAC

Ο αλγόριθμος RANSAC λόγω της απλότητας αλλά και της καλής επίδοσης που προσδίδει, χρησιμοποιείται κατά κόρον στις διάφορες εφαρμογές μίας οπτικής οδομετρίας. Η μέθοδος RANSAC αποτελείται από τα ακόλουθα βήματα [39]:

- 1) Επιλογή τυχαίου ελάχιστου αριθμού αντιστοιχισμένων χαρακτηριστικών που απαιτούνται για την εκτίμηση του μετασχηματισμού (π.χ. τρία σημεία για ένα σύστημα 6DOF χρησιμοποιώντας τη μέθοδο εκτίμησης σφαλμάτων επαναπροβολής 3D σε 2D)
- 2) Εκτίμηση του μετασχηματισμού (παράμετροι περιστροφής και μεταφοράς) χρησιμοποιώντας τα επιλεγμένα σημεία.
- 3) Εφαρμογή του συγκεκριμένου μετασχηματισμού στα υπόλοιπα σημεία.
- 4) Εύρεση της απόστασης (χρησιμοποιείται συνήθως η απόσταση  $l_2$ )  $D$  μεταξύ των μετασχηματισμένων σημείων και των αντίστοιχων αντιστοιχιών τους.
- 5) Ζεύγη με  $D$  μικρότερο από ένα προκαθορισμένο όριο  $\tau$  θεωρούνται αληθείς ταυτίσεις.
- 6) Μέτρηση του συνολικού αριθμού των αληθών ταυτίσεων που προκύπτουν από αυτόν τον μετασχηματισμό
- 7) Επανάληψη προηγούμενων βημάτων  $n$  φορές.
- 8) Ο μετασχηματισμός με τον μεγαλύτερο αριθμό αληθών ταυτίσεων θεωρείται ότι είναι ο σωστός μετασχηματισμός.
- 9) Επανεκτίμηση του μετασχηματισμού χρησιμοποιώντας την προσέγγιση ελαχίστων τετραγώνων χρησιμοποιώντας τις εκτιμήσεις που προέκυψαν από τον αλγόριθμο.

### 3.3.4. Εκτίμηση Κίνησης

Σαν επόμενο βήμα μία οπτική οδομετρία χρησιμοποιεί τα σημεία που βρέθηκαν στις διαφορετικές λήψεις σε μία διαδικασία που ονομάζεται “**Εκτίμηση Κίνησης**” ή “**Motion Estimation**”. Η Εκτίμηση Κίνησης αποτελεί ένα βασικό βήμα υπολογισμού που πραγματοποιείται για κάθε εικόνα σε ένα σύστημα VO. Πιο συγκεκριμένα, σε αυτό το βήμα, υπολογίζεται η κίνηση της κάμερας μεταξύ της τρέχουσας εικόνας και της προηγούμενης εικόνας. Συνδυάζοντας όλες αυτές τις μεμονωμένες κινήσεις, μπορεί να

ανακτηθεί η πλήρης τροχιά της κάμερας και κατ' επέκταση η τροχιά της ρομποτικής πλατφόρμας. Ανάλογα με το αν τα χαρακτηριστικά που βρέθηκαν στα προηγούμενα βήματα προσδιορίζονται σε δύο ή τρεις διαστάσεις, η εκτίμηση κίνησης μπορεί να κατηγοριοποιηθεί από τρεις βασικές μεθόδους τις οποίες αναφέρουμε ενδεικτικά [41]:

Έστω  $T_k$  το μητρώο μετασχηματισμού, μεταξύ δύο εικόνων  $I_{k-1}$  και  $I_k$ , και τα σύνολα των χαρακτηριστικών  $f_{k-1}$ ,  $f_k$  για τις χρονικές στιγμές  $k - 1$  και  $k$  αντίστοιχα.

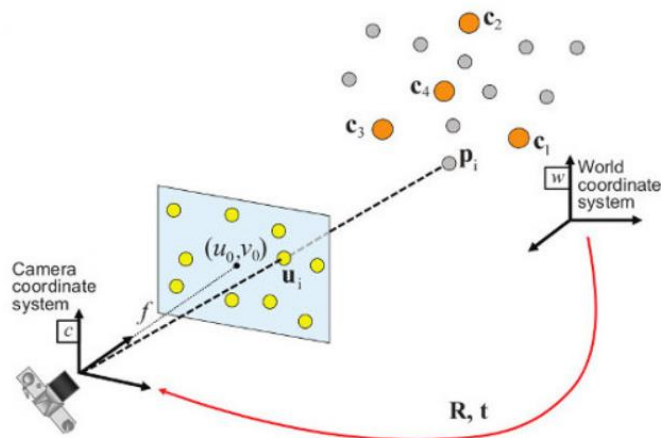
- **2D-2D**: Σε αυτήν την περίπτωση κάθε σύνολο  $f_{k-1}$ ,  $f_k$  προσδιορίζεται στο δισδιάστατο επίπεδο
- **3D-3D**: Όπως μπορεί να γίνει αντιληπτό από τον αναγνώστη, κάθε σύνολο προσδιορίζεται  $f_{k-1}$ ,  $f_k$  στο τρισδιάστατο επίπεδο.
- **3D-2D**: Τέλος, σε αυτή την περίπτωση το σύνολο  $f_{k-1}$  προσδιορίζεται στο τρισδιάστατο επίπεδο και το σύνολο  $f_k$  αποτελεί την αντίστοιχη δισδιάστατη προβολή στην εικόνα  $I_k$ .

Κάθε μία από τις τρεις μεθόδους λειτουργεί με διαφορετική μαθηματική προσέγγιση και σύμφωνα με τον Nister και άλλοι [56], η δεύτερη μέθοδος παρουσιάζει μεγαλύτερη ακρίβεια στα αποτελέσματα της σε σχέση με τις άλλες δύο μεθόδους.

Ενδεικτικά αναφέρουμε την τρίτη υπολογιστική μέθοδο για τον υπολογισμό εκτίμησης θέσης με βάση τις 3D-2D αντιστοιχίσεις των χαρακτηριστικών σε μία εικόνα. Όπως περιγράφεται από τους Fraundorfer και Scaramuzza [41], η γενική διατύπωση αυτής της προσέγγισης στηρίζεται στην εύρεση ενός μητρώου μετασχηματισμού  $T_k$  για την ελαχιστοποίηση του σφάλματος επαναπροβολής όπως φαίνεται παρακάτω:

$$\arg \min_{T_k} \sum_i \|p_k^i - \hat{p}_{k-1}^i\|^2 \quad (3.7)$$

Όπου  $\hat{p}_{k-1}^i$  είναι η επαναπροβολή του τρισδιάστατου σημείου  $X_{k-1}^i$  στην εικόνα  $I_k$  σύμφωνα με τον μετασχηματισμό  $T_k$ . Αυτό το πρόβλημα ανάγεται στην εκτίμηση θέσης μίας βαθμονομημένης κάμερας, δεδομένου ενός συνόλου  $n$  τρισδιάστατων σημείων με τις αντίστοιχες δισδιάστατες απεικονίσεις τους πάνω στο επίπεδο της εικόνας της κάμερας. Όπως είδαμε και στην **Παράγραφο 2.3.5**, μία φωτογραφική μηχανή εκφράζεται συνολικά από 11 βαθμούς ελευθερίας στους οποίους: οι 6 προκύπτουν από τις Εξωγενείς Παραμέτρους (Extrinsics Parameters) και 5 από τις Ενδογενείς Παραμέτρους (Intrinsics Parameters). Με γνωστά βαθμονομημένη κάμερα, οι άγνωστοι περιορίζονται μόνο στις Εξωγενείς, οι οποίες είναι το μητρώο περιστροφής (yaw, pitch, roll) και το διάνυσμα μετατόπισης (x, y, z). Στην βιβλιογραφία αυτό το πρόβλημα ονομάζεται “**Perspective from n Points**” ή “**PnP**” και υπάρχουν διαφορετικές προσεγγίσεις επίλυσης του.



Εικόνα 32. Εκτίμηση Θέσης κάμερας με βάση την 3D-2D αντιστοίχιση χαρακτηριστικών.

Αυτό το πρόβλημα ανάλογα με το αν θεωρούμε τις εξωγενείς και ενδογενείς άγνωστες παραμέτρους λύνεται με διαφορετικό τρόπο. Ενδεικτικά αναφέρουμε λοιπόν ότι, στην περίπτωση που δεν γνωρίζουμε και τα δύο σέτ παραμέτρων χρησιμοποιείται ο “**Άμεσος Γραμμικός Μετασχηματισμός**” (**Direct Linear Transformation**), ο οποίος επιτάσσει για την επίλυση του προβλήματος την ύπαρξη το λιγότερο 6 αντιστοιχίσεων χαρακτηριστικών σε μία εικόνα [57].

Στην περίπτωση που είναι γνωστές οι ενδογενείς παράμετροι της κάμερας, πράγμα που σημαίνει ότι η κάμερα είναι βαθμονομημένη, ο ελάχιστος αριθμός των αντιστοιχιών για την επίλυση του προβλήματος είναι τρεις, το οποίο ονομάζεται “**Perspective 3 Points**”(P3P). Το συγκεκριμένο πρόβλημα εμφανίζει πολλές παραλλαγές για την επίλυση του με κεντρικότερες να αποτελούν ο αλγόριθμος P3P [58], το οποίο είναι η αρχική μορφή του προβλήματος, ο αλγόριθμος EPnP [59] ο οποίος εκφράζει κάθε ένα από τα  $n$  χαρακτηριστικά ως ένα σταθμισμένο άθροισμα τεσσάρων εικονικών σημείων ελέγχου και ο αλγόριθμος UPnP [60], ο οποίος στηρίζεται στον EPnP αλλά δίνει την δυνατότητα εκτίμησης της θέσης για την περίπτωση που οι ενδογενείς παράμετροι θεωρούνται άγνωστες.

## Βελτιστοποίηση εκτίμησης θέσης

Από την εκτίμηση της κίνηση, αναφέραμε πραγματοποιείται μία εκτίμηση της θέσης της κάμερας μέσω της ενσωμάτωσης των διαδοχικών εκτιμήσεων του μητρώου μετασχηματισμού από διαδοχικά φωτογραφικά καρέ. Όμως, κάθε ένας από αυτούς τους μετασχηματισμούς αποτελεί μία εκτίμηση και έχει μία αβεβαιότητα στον υπολογισμό, γεγονός που σημαίνει σε κάθε επανάληψη το συνολικό σφάλμα αθροίζεται από τις αβεβαιότητες των προγενέστερων μετασχηματισμών. Με κύριο στόχο την ελαχιστοποίηση αυτού του σφάλματος χρησιμοποιούνται δύο μεθοδολογίες, οι οποίες βελτιστοποιούν την θέση της κάμερας, με ονομασία “**Συνόρθωση Δέσμης**” (**Bundle Adjustment**) και “**Βελτιστοποίηση μέσω Συνόρθωσης Γράφου**” (**Pose-Graph Optimization**)

Η Συνόρθωση Δέσμης χρησιμοποιεί μία προσέγγιση ελαχίστων τετραγώνων για την προσέγγιση της πόζας της κάμερας χρησιμοποιώντας χαρακτηριστικά στον τρισδιάστατο χώρο. Ισχύει για περιπτώσεις

όπου τα χαρακτηριστικά εικόνες τα οποία εντοπίζονται σε περισσότερα από δύο καρέ. Έχοντας δεδομένο αυτά τα τρισδιάστατα χαρακτηριστικά, υπολογίζουμε σε ποιο pixel θα πρέπει να βρίσκονται σε ένα εικονικό δισδιάστατο επίπεδο εικόνας. Αφού κάνουμε αυτόν τον υπολογισμό, στην συνέχεια βλέπουμε στην πραγματικότητα που βρίσκονται στην πραγματική εικόνα της κάμερας. Αυτή η διαφορά μεταξύ της μετρούμενης και πραγματικής τιμής ονομάζεται σφάλμα επαναπροβολής (reprojection error) και κύριος στόχος αυτής της υπολογιστικής δομής είναι η ελαχιστοποίηση του.

Η μαθηματική έκφραση της Συνόρθωσης Δέσμης, η οποία έχει ως στόχο την ελαχιστοποίηση του σφάλματος επαναπροβολής (reprojection error) δίνεται από τον τύπο:

$$\arg_{x^i, C_k} \min \sum \|p_k^i - g(X^i, C_k)\|^2 \quad (3.8)$$

Θεωρώντας το  $X^i$  το  $i$ -οστό τρισδιάστατο σημείο του μοντέλου,  $p_k^i$  το αντίστοιχο δισδιάστατο χαρακτηριστικό στην εικόνα  $k$ , και  $g(X^i, C_k)$  την επαναπροβολή του σύμφωνα με τη θέση της κάμερας  $C_k$ .

Μία περαιτέρω κατηγοριοποίηση που αξίζει να αναφερθεί, συμβαίνει στην περίπτωση που κατά την Συνόρθωση Δέσμης λαμβάνονται όλα τα φωτογραφικά καρέ η οποία ονομάζεται ολική (Global Bundle Adjustment), ενώ στην περίπτωση που λαμβάνονται  $n$  αριθμοί φωτογραφικών καρέ ονομάζεται τοπική (Local Bundle Adjustment).

### 3.4. Φίλτρα Κάλμαν

Το φίλτρο Kalman προτάθηκε τη δεκαετία του 1950 από τον Rudolph Emil Kalman, ως τεχνική φιλτραρίσματος και πρόβλεψης σε γραμμικά συστήματα. Πρόκειται στην πραγματικότητα για μία βέλτιστη προσέγγιση σύμμειξης διαφόρων αισθητηριακών πληροφοριών για την παρακολούθηση της κατάστασης και πρόβλεψη ενός δυναμικού συστήματος. Το φίλτρο Kalman εφαρμόζει έναν υπολογισμό πεποιθήσεων για συνεχείς καταστάσεις και δεν ισχύει για διακριτούς ή υβριδικούς χώρους κατάστασης. Βασική αρχή ενός τέτοιου φίλτρου αποτελεί ότι τα μοντέλα πρόβλεψης ακολουθούν μία Γκαουσιανή κατανομή πολλαπλών μεταβλητών. Αυτό σημαίνει ότι κάθε πεποίθηση θα είναι της μορφής:

$$p(x) = \frac{1}{\sqrt{2 * \pi * \sigma}} e^{-\frac{(x-\mu)^2}{2 * \sigma^2}} \quad (3.9)$$

Η πεποίθηση  $x$  χαρακτηρίζεται από δύο σύνολα παραμέτρων: την μέση τιμή  $\mu$  και η συνδιακύμανση  $\Sigma$ . Ο μέση τιμή  $\mu$  είναι ένα διάνυσμα που έχει την ίδια διάσταση με την κατάσταση  $x$ . Η συνδιακύμανση είναι ένας τετραγωνικός πίνακας που είναι συμμετρικός και θετικά ημιορισμένος. Η διάστασή του Πίνακα είναι ίση με την διάσταση της κατάστασης  $x$  στο τετράγωνο. Έτσι, ο αριθμός των στοιχείων στον πίνακα συνδιακύμανσης εξαρτάται τετραγωνικά από τον αριθμό των στοιχείων του διανύσματος κατάστασης [61].

Η βασική αρχή ενός τέτοιου φίλτρου αποτελείται από δύο βασικά βήματα τα οποία επαναλαμβάνονται σε έναν ατέρμων βρόχο [62] :

- 1) Βήμα πρόβλεψης ή εκτέλεσης
- 2) Βήμα μέτρησης ή διόρθωσης

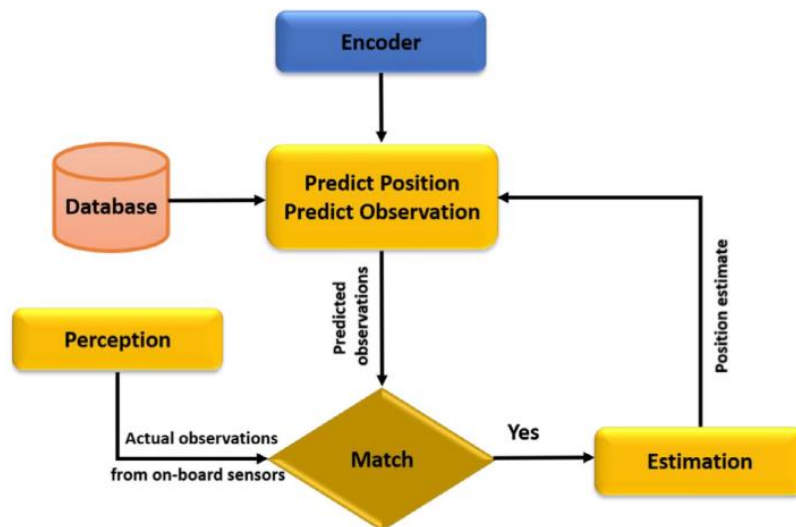
Κατά τη διάρκεια του πρώτου βήματος μία ρομποτική πλατφόρμα χρησιμοποιεί τους ιδιοδεκτικούς της αισθητήρες (π.χ. αισθητήρας τροχού/κινητήρα, αισθητήρας επιτάχυνσης) για την πρόβλεψη της θέσης για ένα πεπερασμένο χρονικό διάστημα. Αυτό εκφράζεται από την παρακάτω εξίσωση ως [61] :

$$x_t = A_t x_{t-1} + B_t u_t + \varepsilon_t \quad (3.10)$$

Όπου  $x_t$  και  $x_{t-1}$  αποτελούν τα διανύσματα κατάστασης ενώ οι  $A$  και  $B$  είναι πίνακες που δημιουργούν μία γραμμική σχέση σε όλο το δυναμικό σύστημα. Η μεταβλητή  $u_t$  αποτελεί την είσοδο ελέγχου του συστήματος ενώ η μεταβλητή  $\varepsilon_t$  αποτελεί ένα τυχαίο διάνυσμα Gauss που μοντελοποιεί τον θόρυβο κατά την μετάβαση κατάστασης. Ο μέσος όρος του διανύσματος είναι μηδέν και η συνδιακύμανση του θα συμβολίζεται  $R_t$ . Όπως συζητήσαμε και σε προηγούμενες παραγράφους, η οδομετρική μέτρηση προσθέτει αρκετή αβεβαιότητα στην πρόβλεψή μας σύμφωνα λόγω των διαφόρων θορύβων που υπάρχουν στο σύστημα. Με στόχο την μείωση αυτού του σφάλματος ένα φίλτρο Κάλμαν χρησιμοποιεί ένα δεύτερο βήμα με στόχο την διόρθωση της προβλεπόμενης τιμής χρησιμοποιώντας τους εξωδεκτικούς αισθητήρες της και παίρνοντας έτσι μία δεύτερη μέτρηση. Για να γίνει η μέτρηση όσον δυνατόν πιο ακριβής χρησιμοποιείται η σύμμιξη διαφόρων πληροφοριών οι οποίες πηγάζουν από διάφορους αισθητήρες (**Sensor Fusion**). Αυτό σημαίνει ότι, ένα ρομπότ συνδυάζει όσον δυνατόν περισσότερους αισθητήρες με στόχο να πάρει μία καλύτερη εκτίμηση για το πού βρίσκεται στον χώρο. Σε μορφή εξίσωσης αυτό μεταφράζεται ως:

$$z_t = C_t x_t + \delta_t \quad (3.11)$$

Εδώ  $C_t$  είναι ένας πίνακας μεγέθους  $k \times n$ , όπου  $k$  είναι η διάσταση του διανύσματος των διαφόρων μετρήσεων  $z_t$ . Το διάνυσμα  $\delta_t$  περιγράφει τον θόρυβο μέτρησης και όπως και στην προηγούμενη περίπτωση μοντελοποιείται από μία Γκαουσιανή κατανομή με μηδενικό μέσο όρο και συνδιακύμανση  $Q_t$ . Υπολογίζοντας έτσι τη διαφορά μεταξύ της προβλεπόμενης τιμής και της μετρούμενης τιμής αποφασίζουμε ποια τιμή θα κρατήσουμε υπολογίζοντας έτσι το “**Κέρδος Κάλμαν**”(Kalman Gain). Στην πραγματικότητα το Κέρδος Κάλμαν καθορίζει τον βαθμό στον οποίο η μέτρηση ενσωματώνεται στη νέα εκτίμηση κατάστασης και δείχνει κατά πόσο εμπιστευόμαστε τις προβλεπόμενες μετρήσεις από το δεύτερο βήμα.



Εικόνα 33. Διάγραμμα Ροής του αλγορίθμου ενός φίλτρου Κάλμαν. Πηγή [62]

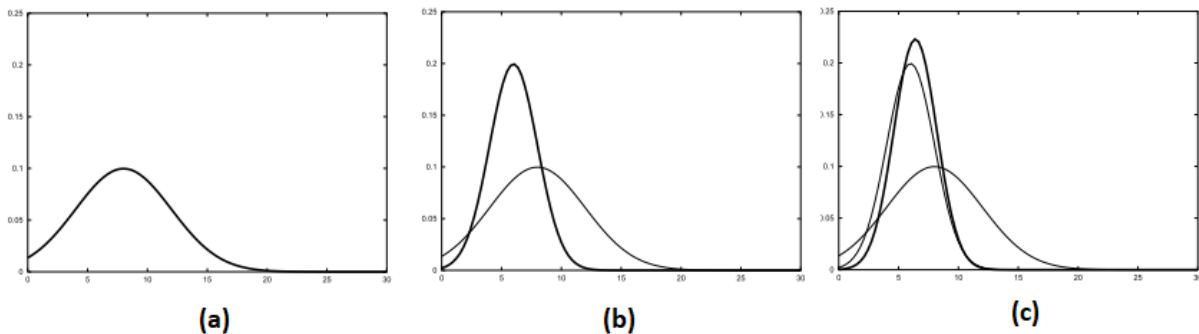
Έχοντας λοιπόν περιγράψει την βασική λειτουργική δομή του φίλτρου, μπορούμε να δούμε και τον ψευδοκώδικα του αλγορίθμου όπως φαίνεται στην **Εικόνα 34**. Για να μπορέσει να λειτουργήσει το φίλτρο θα πρέπει να δοθεί ως είσοδος η μέση τιμή και η συνδιακύμανση για την χρονική στιγμή  $t - 1$  καθώς επίσης η είσοδος ελέγχου αλλά και το διάνυσμα μετρήσεων για την χρονική στιγμή  $t$ . Στις γραμμές 2 και 3 υπολογίζεται η μέση τιμή αλλά και συνδιακύμανση χρησιμοποιώντας μόνο το μοντέλο πρόβλεψης. Έχοντας λοιπόν αυτήν την τιμή στην γραμμή 4 υπολογίζουν το κέρδος Kalman το οποίο όπως αναφέραμε καθορίζει τον βαθμό στον οποίο θα ενσωματώσουμε την τιμή που πήραμε από το προηγούμενο βήμα στη νέα εκτίμηση κατάστασης. Στην γραμμή 5 και 6, υπολογίζεται η νέα τιμή της μέσης τιμής αλλά και η συνδιακύμανση με βάση την τιμή που έχουμε πάρει από το κέρδος Κάλμαν.

**Algorithm Kalman filter**( $\mu_{t-1}, \Sigma_{t-1}, u_t, z_t$ ):

$$\begin{aligned} \bar{\mu}_t &= A_t \mu_{t-1} + B_t u_t \\ \bar{\Sigma}_t &= A_t \Sigma_{t-1} A_t^T + R_t \\ K_t &= \bar{\Sigma}_t C_t^T (C_t \bar{\Sigma}_t C_t^T + Q_t)^{-1} \\ \mu_t &= \bar{\mu}_t + K_t (z_t - C_t \bar{\mu}_t) \\ \Sigma_t &= (I - K_t C_t) \bar{\Sigma}_t \\ \text{return } &\mu_t, \Sigma_t \end{aligned}$$

Εικόνα 34. Ψευδοκώδικας που υλοποιεί τα βήματα ενός Φίλτρου Κάλμαν. Πηγή [61]

Τέλος στην **Εικόνα 35**, παρουσιάζουμε ένα ποιοτικό παράδειγμα στο οποίο γίνεται πιο αντιληπτή η εφαρμογή ενός φίλτρου Κάλμαν για ένα απλοϊκό σενάριο μονοδιάστατου εντοπισμού. Ας υποθέσουμε ότι ένα ρομπότ κινείται κατά μήκος του οριζόντιου άξονα σε κάθε διάγραμμα στην **Εικόνα 35**. Στην **Εικόνα 35(a)** φαίνεται η πεποίθηση της θέσης της ρομποτικής πλατφόρμας η οποία δίνεται από την κανονική κατανομή του σχήματος. Το ρομπότ χρησιμοποιώντας τους αισθητήρες του για τη εντοπισμό της θέσης του (π.χ. ένα σύστημα GPS, κάμερα, LiDAR) παίρνει μια μέτρηση που βρίσκεται στο επίκεντρο της κορυφής της Γκαουσιανής κατανομής όπως φαίνεται στο **Εικόνα 35(b)** με το κέντρο της να δείχνει την τιμή πρόβλεψης ενώ το πλάτος της κατανομής να αντιστοιχεί στην αβεβαιότητα της μέτρησης. Ο συνδυασμός των δύο προηγούμενων μετρήσεων δίνει την μέτρηση που φαίνεται στην **Εικόνα 35(c)** και αποτελεί των μέσο όρο των δύο προηγούμενων μετρήσεων.



Εικόνα 35. Απεικόνιση φίλτρων Κάλμαν: (α) αρχική πεποίθηση, (β) μέτρηση (με έντονα γράμματα) με τη σχετική αβεβαιότητα, (γ) πεποίθηση μετά την ενσωμάτωση της μέτρησης χρησιμοποιώντας μετά την χρήση του φίλτρου. Πηγή [61]

### 3.5.Επεκταμένο Φίλτρο Κάλμαν

Όπως αναφέραμε στην προηγούμενη παράγραφο για την χρήση ενός φίλτρου Κάλμαν η αρχική θέση του ρομπότ θα πρέπει να είναι γνωστή και στην περίπτωση που η ρομποτική πλατφόρμα χαθεί, δεν θα μπορέσει να ανακτήσει τη θέση του. Όπως επίσης αναφέραμε, οι βασικές υποθέσεις που έγιναν είναι ότι το σύστημα θα πρέπει να είναι γραμμικό καθώς επίσης ο θόρυβος μοντελοποιείται από μία κανονική κατανομή. Ωστόσο, πολλά ρομποτικά συστήματα απαιτούν το σύστημα να είναι μη γραμμικό. Αυτό δημιουργεί μια άλλη προσέγγιση εντοπισμού, γνωστή και ως “**Επεκταμένο φίλτρο Κάλμαν**” (**Extended Kalman Filter** ή **EKF**). Το EKF είναι η επέκταση του φίλτρου Κάλμαν για μη γραμμικά συστήματα. Εδώ, η κύρια υπόθεση είναι οι εξισώσεις που περιγράφουν το πρώτο και το δεύτερο βήμα να αποτελούνται από μη γραμμικές συναρτήσεις  $g$  και  $h$ , αντίστοιχα [63]. Αυτό πρακτικά σημαίνει ότι:

$$x_t = g(u_t, x_{t-1}) + \varepsilon_t \quad (3.11)$$

$$z_t = h(x_t) + \delta_t \quad (3.12)$$

Έτσι, το EKF κληρονομεί από το βασικό φίλτρο Κάλμαν τη βασική αναπαράσταση πεποιθήσεων, αλλά διαφέρει στο ότι αυτή η πεποίθηση είναι μόνο κατά προσέγγιση, όχι ακριβής όπως συνέβαινε στο αρχικό φίλτρο. Τα EKF χρησιμοποιούν σαν μέθοδο γραμμικοποίησης μη γραμμικών όρων, που ονομάζεται επέκταση Taylor (πρώτης τάξης). Η επέκταση Taylor κατασκευάζει μια γραμμική προσέγγιση σε μια συνάρτηση  $g$ , χρησιμοποιώντας την συνάρτηση  $g$  αυτή καθ’ αυτήν αλλά επίσης και την κλίση της (μερική παράγωγος), την χρονική στιγμή  $t - 1$ .

Προκύπτει λοιπόν ότι [63]:

$$g(u_t, x_{t-1}) \approx g(u_t, \mu_{t-1}) + g'(u_t, \mu_{t-1})(x_{t-1} - \mu_{t-1}) \quad (3.13)$$

$$g(u_t, x_{t-1}) \approx g(u_t, \mu_{t-1}) + G_t (x_{t-1} - \mu_{t-1}) \quad (3.14)$$

όπου  $G$  είναι ο Ιακωβιανός πίνακας της συνάρτησης  $G$  εξαρτώμενος από τις μεταβλητές  $u_t, \mu_{t-1}$  ο οποίος αλλάζει τιμές ανάλογα με την χρονική σειρά.

Στην παρακάτω εικόνα συνοψίζουμε τα βήματα που ακολουθεί ο αλγόριθμος EKF αλλά και τις διαφορές σε σχέση με το απλό φίλτρο Κάλμαν [63].

```

1:   Algorithm Extended_Kalman_filter( $\mu_{t-1}, \Sigma_{t-1}, u_t, z_t$ ):
2:      $\bar{\mu}_t = g(u_t, \mu_{t-1})$ 
3:      $\bar{\Sigma}_t = G_t \Sigma_{t-1} G_t^T + R_t$ 
4:      $K_t = \bar{\Sigma}_t H_t^T (H_t \bar{\Sigma}_t H_t^T + Q_t)^{-1}$ 
5:      $\mu_t = \bar{\mu}_t + K_t (z_t - h(\bar{\mu}_t))$ 
6:      $\Sigma_t = (I - K_t H_t) \bar{\Sigma}_t$ 
7:     return  $\mu_t, \Sigma_t$ 

```

Εικόνα 36. Ψευδοκώδικας που υλοποιεί τα βήματα ενός Επεκταμένου Φίλτρου Κάλμαν. Πηγή [63]



Οι διαφορές που προκύπτουν σε σχέση με ένα απλό φίλτρο Κάλμαν είναι οι γραμμικές συσχετίσεις που αντικαθίσταται από μη γραμμικές σε ένα ΕΚΦ. Επιπρόσθετα, σε μία τέτοια προσέγγιση χρησιμοποιούνται οι Ιακωβιανοί πίνακες  $G_t$  και  $H_t$  αντί των αντίστοιχων γραμμικών πινάκων  $A_t$ ,  $B_t$  και  $C_t$  που είδαμε στην προηγούμενη παράγραφο.

## Κεφάλαιο 4

### Αρχιτεκτονική Συστήματος σε επίπεδο υλοποίησης

Σε αυτό το κεφάλαιο θα παρουσιαστεί η αρχιτεκτονική του συστήματος που υλοποιήθηκε, το λογισμικό αλλά και τα επιμέρους εργαλεία τα οποία χρησιμοποιήθηκαν με κύριο στόχο την αυτόνομη λειτουργία της ρομποτικής πλατφόρμας Clearpath Husky-a200. Επίσης, παρουσιάζονται και το περιβάλλον προσομοίωσης του ρομπότ το οποίο χρησιμοποιήθηκε κατά την υλοποίηση και τον έλεγχο της γενικής λειτουργικότητας του πειραματικού σχεδιασμού.

#### 4.1 ROS - Robot Operating System

Οι περισσότερες ρομποτικές εφαρμογές απαιτούν ένα εξαιρετικά σύνθετο λογισμικό για να μπορούν να εκτελέσουν βασικές λειτουργίες, λογισμικό που μεγεθύνεται όσο εξελίσσεται το αντικείμενο της ρομποτικής και αυξάνουν οι απαιτήσεις της. Παράλληλα, κάθε ρομποτικό σύστημα παρουσιάζει ένα βαθμό ιδιαιτερότητας, λόγω διαφορετικού hardware, λογισμικού και σκοπού λειτουργίας, γεγονός που καθιστά την επαναχρησιμοποίηση κώδικα εξαιρετικά δύσκολο εγχείρημα. Επίσης, λόγω του μεγάλου μεγέθους του λογισμικού και των απαιτήσεων ταυτόχρονης λειτουργίας και επικοινωνίας μεταξύ των επιμέρους τμημάτων, από το επίπεδο των οδηγών (drivers) των επιμέρους συσκευών του ρομποτικού συστήματος μέχρι τα επίπεδα της αντίληψης, επεξεργασίας και λήψης αποφάσεων, καθίσταται απαραίτητη η ανάγκη για μία μέθοδο υψηλής κλίμακας ενσωμάτωσης λογισμικού.

Για την επίλυση των παραπάνω προβλημάτων, την ανάπτυξη των διαφόρων αλγορίθμων αλλά και την εγκατάσταση των διαφόρων αισθητήρων χρησιμοποιήθηκε το meta-λειτουργικό σύστημα ROS (Robot Operating System) [64]. Το ROS αποτελεί μία ευέλικτη δομή η οποία προσφέρει ένα σύνολο εργαλείων αλλά και βιβλιοθηκών λογισμικού το οποίο αναπτύχθηκε από την εταιρεία Willow Garage και αποτελεί ένα σημαντικό εργαλείο μεταξύ ερευνητών αλλά εμπορικών ρομποτικών επιχειρήσεων, από την αρχική του κυκλοφορία το 2010.

Κάποια από τα κύρια χαρακτηριστικά του ROS αποτελούν:

**Peer-to-Peer:** Ένα σύστημα το οποίο είναι σχεδιασμένο σε ROS αποτελείται από έναν αριθμό προγραμμάτων τα οποία μπορούν να λειτουργούν σε διαφορετικά μηχανήματα σε μία τοπολογία Peer to Peer και να επικοινωνούν μεταξύ τους διαμέσω μίας διεργασίας η οποία ονομάζεται ROS Master.

**Πολύγλωσσο,** η δομή του ROS επιτρέπει την ανάπτυξη διαφόρων προγραμμάτων τα οποία μπορούν να είναι γραμμένα σε διάφορες γλώσσες προγραμματισμού με τις πιο διαδεδομένες να αποτελούν τις Python, C++, XML, YAML.

**Βασισμένο σε εργαλεία:** πολύ σημαντικό στοιχείο αποτελεί ότι το ROS έχει ενσωματωμένα συστήματα που διευκολύνουν την ανάπτυξη και αποσφαλμάτωση των διαφόρων προγραμμάτων μίας ρομποτικής πλατφόρμας.

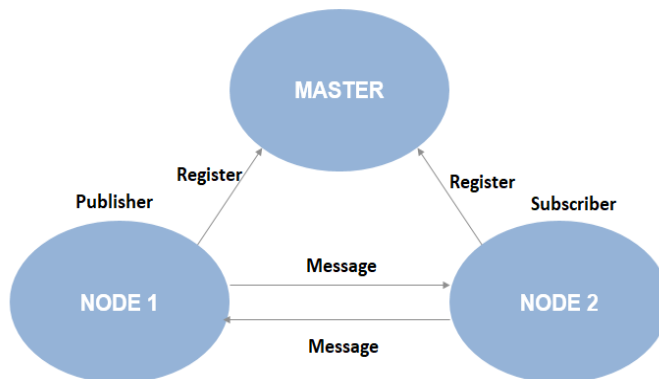
**Ελαφρύ:** όσον αφορά την πλήρη λειτουργικότητα του το ROS αποτελεί μία δομή η οποία δεν χρειάζεται πολύ επεξεργαστική ισχύ αλλά και μνήμη για να λειτουργήσει. Επιπρόσθετα, ένα πολύ σημαντικό στοιχείο

αποτελεί ότι επιτρέπει ευκολότερη εξαγωγή κώδικα αλλά και την επαναχρησιμοποίηση παλαιότερων λειτουργικών προγραμμάτων ρομποτικών εφαρμογών.

**Πρόγραμμα ανοιχτού κώδικα:** ίσως να αποτελεί και το σημαντικότερο όλων των χαρακτηριστικών καθώς το γεγονός ότι αποτελεί ένα εργαλείο με τόσες δυνατότητες σε συνδυασμό με το γεγονός ότι αποτελεί πρόγραμμα ανοιχτού κώδικά, δύναται να χρησιμοποιηθεί από την όλη την επαγγελματική αλλά και πανεπιστημιακή ρομποτική κοινότητα.

Τις θεμελιώδεις λειτουργίες του ROS οι οποίες επιτρέπουν την ανάπτυξη μίας ολόκληρης ρομποτικής εφαρμογής απαρτίζουν οι **κόμβοι** (nodes), τα **θέματα** (topics) , τα **μηνύματα** (messages), οι **υπηρεσίες** (services) και οι **ενέργειες** (actions).

Την κύρια λειτουργική δομή αποτελούν οι κόμβοι, οι οποίοι στην πραγματικότητα πρόκειται για διεργασίες οι οποίες έχουν στόχο την εκτέλεση κάποιου υπολογισμού. Η κατανεμημένη ιδιότητα που έχουν αυτά τα προγράμματα επιτρέπει την ανταλλαγή και άμεση επικοινωνία τους, η οποία οργανώνεται από έναν κεντρικό κόμβο ο οποίος ονομάζεται ROS Master. Κάθε κόμβος μπορεί να μεταφέρει ή να πάρει πληροφορίες από οποιαδήποτε άλλο κόμβο/ους, πράγμα που εξαρτάται από την αρχιτεκτονική της εκάστοτε υλοποίησης. Οι κόμβοι έχουν την δυνατότητα να μεταφέρουν κάποια πληροφορία ονομάζονται publishers, ενώ οι κόμβοι οι οποίοι παίρνουν πληροφορία από κάποιον άλλο κόμβο ονομάζονται subscribers. Είναι γεγονός ότι ένας κόμβος μπορεί να δίνει αλλά και να διαμοιράζει ταυτόχρονα κάποια πληροφορία σε οποιαδήποτε άλλο κόμβο. Για την επικοινωνία των κόμβων χρησιμοποιούνται κάποιοι δίαυλοι οι οποίοι ονομάζονται Θέματα. Ένας κόμβος επικοινωνεί ανταλλάσσει διάφορες πληροφορίες, όπως "οδομετρία" ή "χάρτης" , μέσω μίας δομής που ονομάζεται Μηνύματα. Τα Μηνύματα αποτελούν μία δομή δεδομένων η οποία υποστηρίζει διάφορους τύπους όπως ακέραιοι, αριθμοί μονής ακρίβειας, πίνακες, συμβολοσειρές κλπ., όπως επίσης και συνδυασμούς αυτών.



Εικόνα 37. Γράφος κόμβων που δείχνει την κατεύθυνση των δεδομένων με βάση τον "ρόλο" του κάθε κόμβου.

Στην παρακάτω Εικόνα, μπορούμε να δούμε την μορφή της δομής δεδομένων που ακολουθείται σε περίπτωση που θέλουμε να δημιουργήσουμε ένα μήνυμα μίας οδομετρίας στο ROS.

## nav\_msgs/Odometry Message

---

File: `nav_msgs/Odometry.msg`

### Raw Message Definition

```
# This represents an estimate of a position and velocity in free space.
# The pose in this message should be specified in the coordinate frame given by header.frame_id.
# The twist in this message should be specified in the coordinate frame given by the child_frame_id
Header header
string child_frame_id
geometry_msgs/PoseWithCovariance pose
geometry_msgs/TwistWithCovariance twist
```

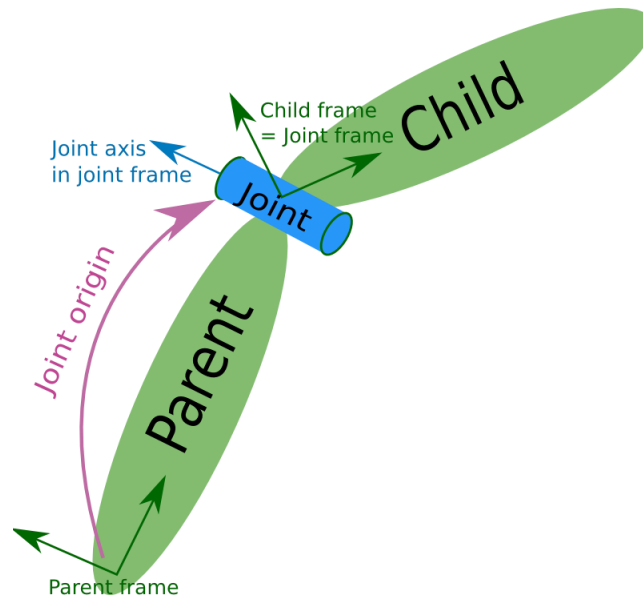
### Compact Message Definition

```
std_msgs/Header header
string child_frame_id
geometry_msgs/PoseWithCovariance pose
geometry_msgs/TwistWithCovariance twist
```

*Εικόνα 38. Δομή δεδομένων για την δημιουργία ενός μηνύματος οδομετρίας στο ROS. [65]*

Ένα πολύ σημαντικό και γνωστό εργαλείο στο ROS, αποτελούν οι “**Σάκοι**” (**bags** ή **Rosbags**), όπου πρόκειται για αρχείο που περιέχουν, ένα ή περισσότερα, topics τα οποία έχουν καταγραφεί σε πραγματικό χρόνο. Αυτά τα μηνύματα μπορούν να επαναχρησιμοποιηθούν δίνοντας μας έτσι την δυνατότητα για την μετέπειτα επεξεργασία.

Για την σωστή αναπαράσταση των ιδιοτήτων μίας ρομποτικής πλατφόρμας σε ψηφιακό επίπεδο το σύστημα ROS χρησιμοποιεί μία γλώσσα μοντελοποίησης η οποία ονομάζεται “**Ενιαία Μορφή Ρομποτικής Περιγραφής**” ή “**Unified Robotic Description Format**” (**URDF**). Ένα αρχείο URDF αποτελεί μία περιγραφή γραμμένη εκφρασμένη σε γλώσσα προγραμματισμού XML (Extensible Markup Language [66]) και περιέχει όλα τα στοιχεία τα οποία συνθέτουν μία περιγραφή της ρομποτικής διάταξης. Για την διάρθρωση αυτής της περιγραφής κάθε ένα καινούργιο στοιχείο θα πρέπει να αποτελείται από τα βασικά στοιχεία όπως σύνδεσμοι και αρθρώσεις όπως επίσης και αρχεία τα οποία χρησιμοποιούνται για την οπτικοποίηση αυτών. Οι σύνδεσμοι χρησιμοποιούνται για να περιγράψουν ένα άκαμπτο σώμα με χαρακτηριστικά όπως αδράνεια, οπτικά χαρακτηριστικά και ιδιότητες σύγκρουσης. Κάθε ένας σύνδεσμος, συνδέεται με οποιαδήποτε άλλον χρησιμοποιώντας ένα στοιχείο που ονομάζεται άρθρωση όπως φαίνεται στην παρακάτω Εικόνα. Αξίζει να σημειωθεί ότι ένα αρχείο URDF, όλοι οι σύνδεσμοι συνδέονται μεταξύ τους σε μία μορφή συνεκτικού γράφου χωρίς κύκλους, δηλαδή σαν δέντρα. Αυτό πρακτικά σημαίνει ότι, κάθε γιός σύνδεσμος μπορεί να έχει έως ένα πατέρα, κάθε πατέρας μπορεί να έχει όσα παιδιά θέλει και κάθε σύνδεσμος πατέρα-γιου συνδέεται μέσω μίας άρθρωσης.



Εικόνα 39. Σχέδιο που περιγράφει την δενδροειδή ένωση δύο συνδέσμων μέσω μίας άρθρωσης. Πηγή [67]

Όπως μπορούμε να καταλάβουμε κάθε ένας από του συνδέσμων συνοδεύεται με το δικό του σύστημα συντεταγμένων, το οποίο είναι στην πραγματικότητα η “οπτική” από άποψη της θέσης και του προσανατολισμού εκφρασμένη για κάθε ξεχωριστό σύνδεσμο. Μία κινητή πλατφόρμα για να μπορέσει να προσανατολιστεί σε έναν χώρο, είναι σύνηθες να χρησιμοποιεί κάποιο εξωτερικό, διάφορο του τοπικού της, σύστημα συντεταγμένων. Με σκοπό την ύπαρξης μίας κοινής σύμβασης σχετικά με τις ονομασίες αλλά και την λειτουργία των συστημάτων συντεταγμένων, το ROS έχει δημιουργήσει έναν κανονισμό με όνομα “REP 105: Συστήματα Συντεταγμένων για Κινητές Ρομποτικές Πλατφόρμες” [68]. Σε αυτόν τον κανονισμό ορίζονται τρεις ονομασίες συστημάτων συντεταγμένων ανάλογα με τον ρόλο που έχουν στον σχεδιασμό ενός συστήματος αυτοενοτοπισμού μίας κινητής ρομποτικής πλατφόρμας. Πιο συγκεκριμένα θα πρέπει να οριστεί τα παρακάτω:

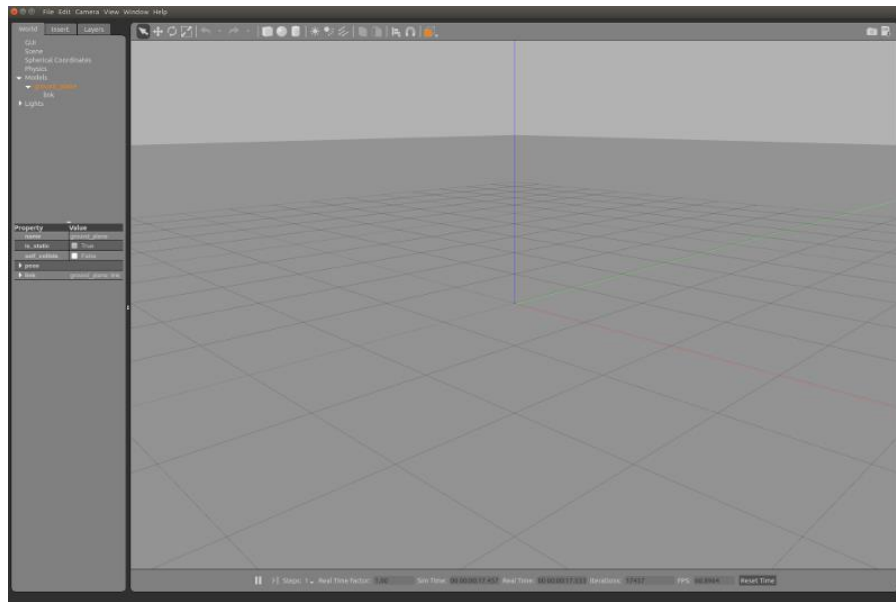
- Το **base\_link** αφορά το σύστημα συντεταγμένων το οποίο είναι προσαρτημένο στη βάση της πλατφόρμας
- Το **odom** αφορά ένα σταθερό παγκόσμιο σύστημα συντεταγμένων. Σε αυτό το πλαίσιο αναφοράς η εκτίμηση της πόζας της ρομποτικής πλατφόρμας αποκλίνει στο πέρασμα του χρόνου δημιουργώντας έτσι ένα μη φραγμένο σφάλμα. Παρ’ όλ’ αυτά, αυτό το πλαίσιο είναι συνεχές, πράγμα που σημαίνει ότι μπορεί να χρησιμοποιηθεί για κινήσεις σε βραχυπρόθεσμη τοπική αναφορά.
- Τέλος το **map** είναι και αυτό ένα παγκόσμιο σύστημα συντεταγμένων το οποίο συντεταγμένων, το οποίο διαφέρει ως προς την συνέχεια και την εκτίμηση της θέσης της πλατφόρμας. Εδώ η εκτίμηση της θέσης δεν είναι συνεχής πράγμα που σημαίνει ότι μπορεί να εμφανιστούν διακριτές εναλλαγές των τιμών εκτίμησης με βάση τις πληροφορίες που παρέχουν οι αισθητήρες του ρομπότ εξαλείφοντας έτσι οποιαδήποτε απόκλιση από την πραγματική τιμή. Αυτό το σύστημα συντεταγμένων χρησιμοποιείται κυρίως για μακροπρόθεσμη παγκόσμια αναφορά.

## 4.2. Μηχανή Προσομοίωσης Gazebo

Η ανάπτυξη αλγορίθμων για ρομποτικά συστήματα αποτελεί μία αρκετά σύνθετη, πολυπαραγοντική και χρονοβόρα διαδικασία κατά την οποία θα πρέπει να δοκιμαστεί αλλά και βελτιστοποιηθεί πάρα πολλές φορές ένας αλγόριθμος μέχρι την τελική του έκδοση. Όπως καταλαβαίνει κανείς, η ανάπτυξη και η εκτέλεση ενός αλγορίθμου αλλά και ενός πειράματος πριν δοκιμαστεί σε πραγματικές συνθήκες θα πρέπει να εξεταστεί σε έναν προσομοιωτή για να επιβεβαιώσουμε ότι η εφαρμογή μας μπορεί να λειτουργήσει. Επιπρόσθετα, η ανάπτυξη μίας ρομποτικής εφαρμογής μπορεί να δημιουργήσει διαφορετικά αποτελέσματα από αυτά που περιμένουμε, πράγμα επικίνδυνο και για το φυσικό περιβάλλον αλλά και για τον ίδιο τον ρομποτικό μηχανισμό. Τέλος, η εν λόγω εφαρμογή μπορεί συνήθως να απαιτεί συνεχείς δοκιμές πάνω στο ρομπότ και επομένως προϋποθέτει την διαρκή φυσική παρουσία και πλήρη ετοιμότητα αυτού, πράγμα που δεν είναι πάντα εφικτό.

Κάτω από αυτό το πρίσμα, έχουν αναπτυχθεί ανά τα χρόνια προσομοιωτές ρομποτικών συστημάτων, μέσω των οποίων είναι δυνατή η μοντελοποίηση ενός ή περισσότερων ρομποτικών συστημάτων και η προσομοίωση αυτών, σε ένα ψηφιακό περιβάλλον ικανό να προσεγγίσει σε έναν βαθμό την πραγματικότητα. Προφανώς και ένας προσομοιωτής δεν αποτελεί μία τέλεια προσέγγιση της πραγματικότητας, αλλά μπορούμε να δούμε σε ένα ελεγχόμενο, γρήγορο και επεξεργάσιμο περιβάλλον την βασική λειτουργικότητα μίας ρομποτικής εφαρμογής.

Ο μηχανή προσομοίωσης **Gazebo** αποτελεί ένα από τα πιο γνωστά εργαλεία, τα οποία προσφέρουν ένα τρισδιάστατο προσομοιωτή ρομποτικών συστημάτων, χρησιμοποιώντας σε εσωτερικά και εξωτερικά δυναμικά περιβάλλοντα μέσω ανεπτυγμένων μηχανών φυσικής.



Εικόνα 40. Το αρχικό GUI της μηχανής προσομοίωσης Gazebo. Πηγή [69]

### 4.3. Robot Localization

Πολύ σημαντικό κομμάτι στη τεχνική υλοποίηση του πειράματος αποτέλεσε το πακέτο “**robot\_localization**”, το οποίο αποτελεί μία συλλογή από κόμβους εκτίμησης κατάστασης καθένας από τους οποίους αποτελεί μία υλοποίηση ενός μη γραμμικού φίλτρου Κάλμαν (EKF) το οποίο χρησιμοποιείται κυρίως σε ρομποτικές πλατφόρμες οι οποίες κινούνται στο τρισδιάστατο επίπεδο [70]. Το robot\_localization πραγματοποιεί μία εκτίμηση της θέσης αλλά και του προσανατολισμού στον τρισδιάστατο χώρο, χρησιμοποιώντας την σύμμειξη πληροφοριών που πηγάζουν από διάφορους αισθητήρες, οι οποίες ακολουθούν ως μέσο επικοινωνίας αλλά και διαχείρισης το λειτουργικό σύστημα ROS.

Αξίζει να σημειωθεί ότι αυτή η υλοποίηση χρησιμοποιεί για τον υπολογισμό του βήματος μέτρησης το τρισδιάστατο κινηματικό μοντέλο της Νευτώνειας μηχανικής. Επιπρόσθετα, για τον υπολογισμό του βήματος πρόβλεψης, ο κάθε αισθητήρας για να εισαχθεί στο φίλτρο θα πρέπει να παράγει μετρήσεις πρόβλεψης των μεταβλητών κατάστασης που εκτιμούμε (θέση και προσανατολισμό). Τέλος θα αποτελούσε παράλειψη να μην αναφέρουμε ότι, οι πίνακες συνδιακύμανσης της αρχικής κατάστασης αλλά και του θορύβου που εισάγεται στο βήμα μέτρησης θα πρέπει να εισαχθούν από τον χρήστη.

Όπως μπορούμε να καταλάβουμε στην παρακάτω εικόνα, ο ψευδοκώδικας της οποίας αναφέρεται στην υλοποίηση ενός Επεκταμένου Φίλτρου Κάλμαν (**Παράγραφος 3.5**), τα κόκκινα κουτιά υποδηλώνουν τις μεταβλητές τις οποίες θα πρέπει να εισάγει ο χρήστης για την σωστή λειτουργία του φίλτρου.

**Preparing Your Sensor Data: Covariance** CLEARPATH ROBOTICS

initial\_estimate\_covariance ( $P_0$ )      process\_noise\_covariance

$$\hat{x}_{k|k-1} = f(\hat{x}_{k-1|k-1})$$

$$P_{k|k-1} = P_{k-1|k-1} + Q$$

$$K_k = P_{k|k-1} H_k^T (H_k^T P_{k|k-1} H_k^T + R_k)^{-1}$$

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k (z_k - H_k \hat{x}_{k|k-1})$$

$$P_{k|k} = (I - K_k H_k) P_{k|k-1} (I - K_k H_k)^T + K_k R_k K_k^T$$

nav\_msgs/Odometry  
 geometry\_msgs/PoseWithCovarianceStamped  
 geometry\_msgs/TwistWithCovarianceStamped  
 sensor\_msgs/Imu

Εικόνα 41. Στιγμιότυπο από την παρουσίαση του πακέτου robot\_localization στο συνέδριο ROSCon. [71]

Από την προηγούμενη εικόνα μπορούμε να δούμε ότι ο χρήστης του συγκεκριμένου πακέτου θα πρέπει να εισάγει τρεις βασικές μεταβλητές. Όπως αναφέρεται στην περιγραφή του πακέτου [72], οι Πίνακες  $P$  και  $Q$  υποδεικνύουν τους πίνακες συνδιακύμανσης στη τρέχουσα κατάσταση και στο στάδιο πρόβλεψης αντίστοιχα, ενώ η μεταβλητή  $R$  δηλώνει το μήνυμα οδομετρίας με συγκεκριμένη δομή όπως εκφράζεται στο ROS (**Παράγραφος 4**). Η κύρια παραμετροποίηση του πακέτου στηρίζεται στην κατάλληλη ρύθμιση των Πινάκων  $P$  και  $Q$  η οποία κατά βάση γίνεται με επαναλαμβανόμενη δοκιμή. Κάθε Πίνακας

συνδιακύμανσης περιέχει το σύνολο των εξεταζόμενων μεταβλητών του δυναμικού συστήματος και στην ουσία περιγράφει τον τρόπο συσχέτισης κάθε ζεύγους μεταβλητών. Οι μεταβλητές που εξετάζονται στην περίπτωση στην περίπτωση ενός υπολογισμού μίας οδομετρίας είναι συνολικά 15 και αφορούν την θέση, την ταχύτητα, την επιτάχυνση, την περιστροφή, την γωνιακή ταχύτητα στο επίπεδο. Σημαντικές ιδιότητες ενός τέτοιου Πίνακα είναι ότι είναι συμμετρικός, θετικά ορισμένος, κάθε στοιχείο του θα πρέπει να περιέχει έναν μη αρνητικό πραγματικό αριθμό και τα στοιχεία της μη κύριας διαγωνίου εκφράζουν την συσχέτιση των δύο εξεταζόμενων μεταβλητών που εκφράζει η γραμμή και η στήλη. Ένας τέτοιος Πίνακας θα είναι της μορφής:

	x	y	z	Roll	Pitch	Yaw	$V_x$	$V_y$	$V_z$	$V_{roll}$	$V_{pitch}$	$V_{yaw}$	$accel_x$	$accel_y$	$accel_z$
x	d	0	0	0	0	0	0	0	0	0	0	0	0	0	0
y	0	d	0	0	0	0	0	0	0	0	0	0	0	0	0
z	0	0	d	0	0	0	0	0	0	0	0	0	0	0	0
Roll	0	0	0	d	0	0	0	0	0	0	0	0	0	0	0
Pitch	0	0	0	0	d	0	0	0	0	0	0	0	0	0	0
Yaw	0	0	0	0	0	d	0	0	0	0	0	0	0	0	0
$V_x$	0	0	0	0	0	0	d	0	0	0	0	0	0	0	0
$V_y$	0	0	0	0	0	0	0	d	0	0	0	0	0	0	0
$V_z$	0	0	0	0	0	0	0	0	d	0	0	0	0	0	0
$V_{roll}$	0	0	0	0	0	0	0	0	0	d	0	0	0	0	0
$V_{pitch}$	0	0	0	0	0	0	0	0	0	0	d	0	0	0	0
$V_{yaw}$	0	0	0	0	0	0	0	0	0	0	0	d	0	0	0
$accel_x$	0	0	0	0	0	0	0	0	0	0	0	0	d	0	0
$accel_y$	0	0	0	0	0	0	0	0	0	0	0	0	0	d	0
$accel_z$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	d

Πίνακας 1. Μορφή πίνακα Συνδιακύμανσης στο πακέτο `robot_localization`.

Με  $0 \leq d \leq 1, d \in R$

Μεγαλύτερη τιμή της διαγωνίου  $d$ , θα προκαλέσει και ένα μεγαλύτερο προβλεπόμενο σφάλμα για το συγκεκριμένο ζεύγος μεταβλητών.

#### 4.4. RTAB-Map

Σε αυτήν την παράγραφο θα παρουσιάσουμε το πακέτο RTAB-Map (Real-Time Appearance-Based Mapping), το οποίο χρησιμοποιήθηκε για την δημιουργία μίας οπτικής οδομετρίας η οποία παρουσιάστηκε στην **Παράγραφο 3.3** και αποτελεί ένα από τα πιο γνωστά πακέτα ανοιχτού κώδικα για την υλοποίηση μίας οπτικής οδομετρίας στον τομέα της ρομποτικής. Το RTAB-Map είναι μια προσέγγιση επίλυσης του προβλήματος SLAM που έχει ενσωματωθεί στο ROS ως πακέτο με ονομασία `rtabmap_ros` από το 2013. Σαν κύριες αισθητήριες διατάξεις, το συγκεκριμένο πακέτο μπορεί να χρησιμοποιήσει RGB-D και στερεοσκοπικές κάμερες, αισθητήρες LiDAR και συνδυασμούς των προηγούμενων. Εκτός από μία ολοκληρωμένη προσέγγιση του προβλήματος της ταυτόχρονης χαρτογράφησης αλλά και αυτοεντοπισμού, το συγκεκριμένο πακέτο μπορεί να συμβάλει μόνο στην επίλυση του αυτοεντοπισμού, δημιουργώντας μία οδομετρία, ή μόνο στην χαρτογράφηση, χρησιμοποιώντας εξωτερική οδομετρία. Στα πλαίσια αυτής της διπλωματικής χρησιμοποιήθηκε η οπτική οδομετρία που παρέχει το πακέτο RTAB-Map χρησιμοποιώντας



ως κύρια αισθητήρια διάταξη την κάμερα που έχουμε αναφέρει σε προηγούμενο κεφάλαιο. Όσον αφορά τις βασικές αλγοριθμικές δομές, το πακέτο RTAB-Map στηρίζεται κυρίως στις διάφορες υλοποιήσεις που προσφέρει η ανοιχτή βιβλιοθήκη OpenCV.

Όπως μπορούμε να δούμε στο [73], τα βήματα που ακολουθεί το πακέτο για να δημιουργήσει μία οπτική οδομετρία είναι τα εξής:

1. **Ανίχνευση Χαρακτηριστικών:** Για κάθε ένα φωτογραφικό καρέ, εφαρμόζεται ο αλγόριθμος των Shi-Tomashi όπως περιγράψαμε στην **Παράγραφο 3.3.1** ανιχνεύοντας έτσι έναν αριθμό χαρακτηριστικών σε μία εικόνα (στην προκειμένη περίπτωση γωνίες). Το RTAB-Map υποστηρίζει όλους τους τύπους εξαγωγής χαρακτηριστικών οι οποίοι είναι διαθέσιμοι στην βιβλιοθήκη OpenCV, αλλά ο αλγόριθμος των Shi-Tomashi έχει επιλεγεί για να διευκολύνει στην παραμετροποίηση και να λαμβάνει ομοιόμορφα ανιχνευμένες λειτουργίες σε διαφορετικό μέγεθος εικόνας και ένταση φωτός.
2. **Αντιστοίχιση Χαρακτηριστικών:** Η αντιστοίχιση πραγματοποιείται με την βοήθεια του αλγορίθμου (**Παράγραφος 3.3.3**), χρησιμοποιώντας Περιγραφείς BRIEF (**Παράγραφος 3.3.2**) από τα Χαρακτηριστικά που προέκυψαν από το προηγούμενο βήμα.
3. **Πρόβλεψη κίνησης:** Ένα μοντέλο κίνησης χρησιμοποιείται για να προβλέψει στο τρέχον καρέ, με βάση τον προηγούμενο μετασχηματισμό κίνησης. Αυτό περιορίζει το παράθυρο αναζήτησης για Αντιστοίχιση χαρακτηριστικών και παρέχει καλύτερες αντιστοιχίσεις, ιδιαίτερα σε περιβάλλοντα με δυναμικά αντικείμενα και επαναλαμβανόμενες υφές.
4. **Εκτίμηση κίνησης:** Μετά τον υπολογισμό των αντιστοιχιών χρησιμοποιείται ο αλγόριθμος RANSAC (**Παράγραφος 3.3.3**), καθώς επίσης και το μοντέλο 3D-2D (**Παράγραφος 3.3.4**) το οποίο είναι ήδη υλοποιημένο από την βιβλιοθήκη OpenCV με όνομα RANSAC Perspective-n-Point (PnP) [74].
5. **Τοπική Συνόρθωση Δέσμης:** Ο μετασχηματισμός που προκύπτει βελτιστοποιείται χρησιμοποιώντας Τοπική Συνόρθωση Δέσμης (Local Bundle Adjustment) (**Παράγραφος 3.3.4**).
6. **Ενημέρωση θέσης:** Με τον εκτιμώμενο μετασχηματισμό, δημιουργείται μία οπτική οδομετρία η οποία στέλνεται σαν μετασχηματισμός, μέσω του ROS, από το σύστημα συντεταγμένων της οδομετρίας (odom frame) στο σύστημα συντεταγμένων της βάσης της ρομποτικής πλατφόρμας (base\_link frame) (**Παράγραφος 4**).

Στις παρακάτω δύο Εικόνες μπορούμε να δούμε δύο στιγμιότυπα δημιουργία της οπτικής οδομετρίας κατά την διάρκεια της πειραματικής διαδικασίας. Στην **Εικόνα 42** φαίνονται με κόκκινα σημεία τα σημεία τα οποία έχουν εντοπιστεί όλα τα υποψήφια χαρακτηριστικά εικόνας τα οποία θα πρέπει να εντοπιστούν στο επόμενο φωτογραφικό καρέ.



Εικόνα 42. Φωτογραφικό καρέ στο οποίο φαίνονται με κόκκινο χρώμα τα σημεία που εξήχθησαν για το πρώτο βήμα για την δημιουργία μιας οπτικής οδομετρίας με χρήση του πακέτου RTAB-Map.

Στην **Εικόνα 43** φαίνονται τα ταυτισμένα σημεία μετά το φιλτράρισμα με την χρήση του αλγορίθμου RANSAC. Με κόκκινο χρώμα φαίνονται τα σημεία που θεωρούνται outliers, ενώ με μπλε τα σημεία που χρησιμοποιήθηκαν για την εκτίμηση θέσης μεταξύ αυτών των δύο διαδοχικών καρέ τα οποία ενώνονται με μία μπλε γραμμή.

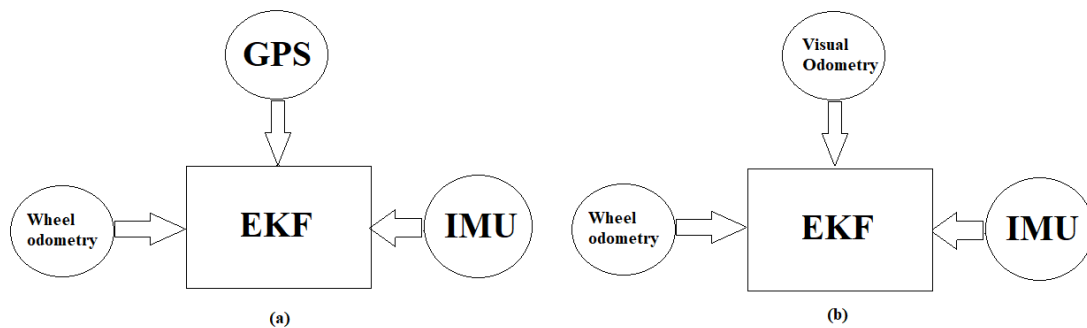


Εικόνα 43. Ταύτιση χαρακτηριστικών εικόνας σε δύο διαδοχικά καρέ, με κόκκινο χρώμα φαίνονται τα σημεία τα οποία θεωρούνται outliers.

## Κεφάλαιο 5

### Πείραμα και Αποτελέσματα

Σκοπός της παρούσας διπλωματικής είναι η μελέτη, η υλοποίηση αλλά και η σύγκριση δύο εφαρμογών στο κομμάτι του Αυτοεντοπισμού (Localization). Στα πλαίσια αυτής της μελέτης δημιουργήθηκαν δύο διαφορετικά Επεκταμένα φίλτρα Κάλμαν με την βοήθεια του πακέτου robot\_localization (**Παράγραφος 4.3**) όπου στο καθένα κάνουμε μία διαφορετική σύμμειξη αισθητήριων πληροφοριών. Η κύρια διαφορά αυτών των δύο φίλτρων προέρχεται στο γεγονός ότι στο πρώτο φίλτρο εισάγεται μία οδομετρία η οποία πηγάζει χρησιμοποιώντας πληροφορίες από GPS-RTK ενώ στην δεύτερη περίπτωση η οδομετρία δημιουργείται με χρήση κάμερας (οπτική οδομετρία). Αξίζει να σημειωθεί ότι και στις δύο περιπτώσεις προστέθηκε για την εκτίμηση της θέσης η οδομετρία που παράχθηκε από τους τροχούς και οι πληροφορίες της IMU όσον αφορά τον προσδιορισμό του προσανατολισμού της πλατφόρμας. Τέλος η σύγκριση των δύο οδομετριών γίνεται μόνο ως προς την θέση της πλατφόρμας και όχι ως προς τον προσανατολισμό αυτής. Στην **Εικόνα 44** μπορούμε να δούμε τι αισθητήρια δεδομένα εισάγονται στο φίλτρο Κάλμαν ανά περίπτωση.



Εικόνα 44. Αισθητήρια δεδομένα που χρησιμοποιήθηκαν ανά περίπτωση.

Για να μπορεί να γίνει μία τέτοια σύγκριση, θα πρέπει να εξασφαλιστούν δύο σημαντικά χαρακτηριστικά τα οποία διασφαλίζουν την εγκυρότητα των μετρήσεων.

Το πρώτο εκ των δύο είναι η εξασφάλιση ίδιας διαδρομής για την εκτίμηση θέσης με χρήση GPS-RTK αλλά και με την χρήση κάμερας. Για να επιτευχθεί αυτό το χαρακτηριστικό, η ρομποτική πλατφόρμα οδηγήθηκε χειροκίνητα κατά μήκος μίας ευθείας τροχιάς σε μία τυχαία γραμμή αμπελώνα. Κατά την διάρκεια εκτέλεσης αυτής της κίνησης, καταγράφηκαν σε ένα αρχείο rosbag (**Παράγραφος 4**) όλες οι πληροφορίες που πηγάζουν από τους διάφορους αισθητήρες της πλατφόρμας (GPS, IMU, οδομετρία τροχών, Κάμερα). Έχοντας λοιπόν αυτό το αρχείο δημιουργήθηκαν δύο διαφορετικές οδομετρίες για κάθε μία από τις δύο περιπτώσεις χρησιμοποιώντας μόνο τα δεδομένα από τους αισθητήρες που χρειαζόμαστε, εξασφαλίζοντας έτσι την εγκυρότητα της ίδιας διαδρομής ανά διαφορετική περίπτωση.

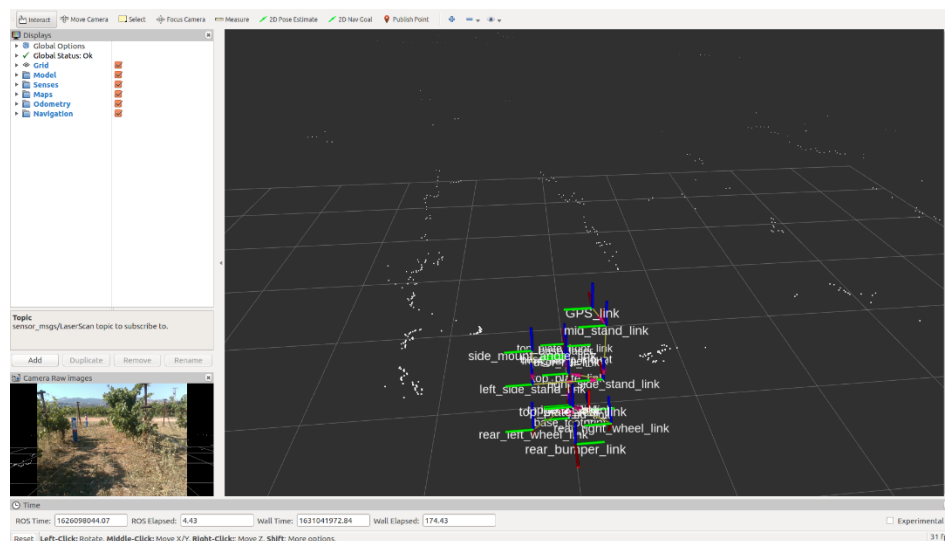
Όσον αφορά το δεύτερο χαρακτηριστικό το οποίο μας δίνει την δυνατότητα μίας επιτυχής σύγκρισης, είναι η δημιουργία μίας εκτίμησης για το πού βρίσκεται το ρομπότ στην πραγματικότητα μέσα στο επίπεδο

στο οποίο λειτουργεί. Στην επόμενη παράγραφο παρουσιάζουμε την μεθοδολογία που ακολουθήθηκε για την δημιουργία αυτής της εκτίμησης.

## 5.1 Δημιουργία Groundtruth

Ο όρος “**Groundtruth**”, αποτελεί μία ονομασία που χρησιμοποιείται σε διάφορα πεδία και αναφέρεται σε μία πληροφορία ή ένα δεδομένο το οποίο υποδηλώνει την πραγματικότητα και χρησιμοποιείται ως μία αληθής μέτρηση για σύγκριση διαφόρων εκτιμήσεων. Η δημιουργία μίας τέτοιας πληροφορίας προκύπτει από δεδομένα τα οποία είναι αληθής και παρέχονται από άμεση παρατήρηση και μέτρηση (δηλαδή εμπειρικά στοιχεία). Για την σύγκριση δύο εκτιμήσεων λοιπόν, θα πρέπει να γίνει ως προς μία τρίτη εκτίμηση η οποία γνωρίζουμε ότι είναι αληθής.

Στην παρούσα διπλωματική, αποφασίστηκε να χρησιμοποιηθεί κάποιο από τα διαθέσιμα αισθητήρια όργανα, το οποίο με την κατάλληλη επεξεργασία θα μας δείχνει το πού βρίσκεται στην πραγματικότητα η ρομποτική πλατφόρμα με έγκυρο και άμεσο τρόπο. Γι’ αυτόν τον λόγο επιλέχθηκε η χρήση ενός αισθητήρα LiDAR ο οποίος παρέχει μία πληροφορία με γρήγορο και ελάχιστο σφάλμα μέτρησης (μέγιστο συστηματικό σφάλμα μέτρησης 50 mm [75]). Όπως είδαμε και στο **Κεφάλαιο 2.3.4**, ένας τέτοιος αισθητήρας σαρώνει περιοδικά μία δισδιάστατη περιοχή χρησιμοποιώντας δέσμες φωτός. Σε περίπτωση που υπάρχει κάποιο αντικείμενο μέσα στο εύρος σάρωσης, ένας αισθητήρας LiDAR θα μας δώσει την ακριβή τοποθεσία του στο δισδιάστατο επίπεδο καθώς επίσης και μία τιμή η οποία μας δείχνει την ένταση της προσπίπτουσας-επιστρεφόμενης ακτίνας. Ο αισθητήρας που χρησιμοποιήθηκε για το πείραμα αποτελεί έναν δισδιάστατο αισθητήρα και στην **Εικόνα 45** φαίνονται με άσπρες τελείες τα αντικείμενα που έχει εντοπίσει εντός εύρους σάρωσης κατά την διάρκεια του πειράματος.



Εικόνα 45. Σύνολο αντικειμένων τα οποία εντοπίζονται από τον αισθητήρα LiDAR.

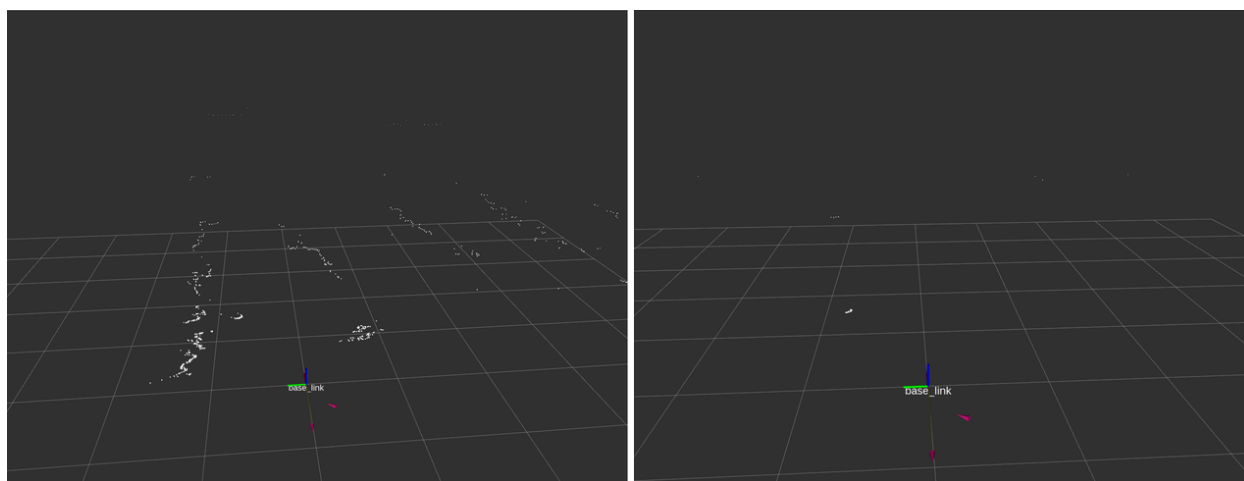
Η κύρια ιδέα για να γνωρίζουμε καθ’ όλη την διάρκεια την πραγματική θέση της ρομποτικής πλατφόρμας στηρίζεται στο να εντοπίζουμε και να παρακολουθούμε ένα στατικό γνωστό σημείο αναφοράς (landmark). Καθώς κινείται στον χώρο η πλατφόρμα, σε κάθε χρονική στιγμή ελέγχει το πόσο μακριά ή κοντά βρίσκεται από αυτό το σημείο και κατ’ επέκταση να μπορεί με ακρίβεια να γνωρίζει ανά πάσα ώρα και στιγμή το πού βρίσκεται σε ένα τοπικό σύστημα συντεταγμένων. Γι’ αυτό το λόγο δημιουργήθηκαν

κάποια τεχνητά κυλινδρικά αντικείμενα τα οποία τοποθετήθηκαν σε ευθεία απόσταση μεταξύ τους κατά **τρία μέτρα**. Η συγκεκριμένη απόσταση επιλέχθηκε διότι μετά από επαναλαμβανόμενες δοκιμές, ο αισθητήρας LiDAR εντόπιζε με πλήρη σιγουριά όσο το δυνατόν περισσότερες δέσμες το αντικείμενο. Για να μπορέσουμε να απομονώσουμε τις δέσμες οι οποίες αφορούσαν το αντικείμενο-ορόσημο, τοποθετήθηκαν περιμετρικά κάποιες στρώσεις ανακλαστικής ταινίας



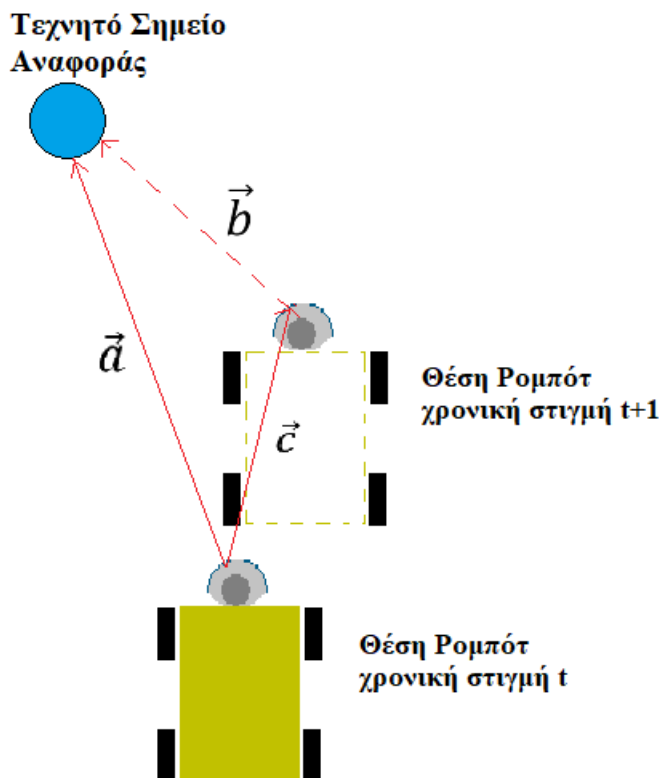
Εικόνα 46. Τεχνητά σημεία αναφοράς (landmarks) το οποίο είναι τυλιγμένο με ανακλαστική ταινία (άσπρη και κόκκινη στρώση).

Η επιλογή της ανακλαστικής ταινίας βασίζεται στην υψηλή ανακλαστική ιδιότητα που θα δίνει κατευθείαν στο τεχνητό σημείο έναντι των φυσικών υλικών του περιβάλλοντος. Με αυτό τον τρόπο μπορούμε να φιλτράρουμε τα σημεία τα οποία παρουσιάζουν χαμηλή ανακλαστική τιμή και να κρατήσουμε μόνο τις υψηλές τιμές οι οποίες αντιπροσωπεύουν στην πραγματικότητα τα τεχνητά σημεία τα οποία και θέλουμε να παρακολουθεί η πλατφόρμα σε κάθε χρονική στιγμή χρησιμοποιώντας τον αισθητήρα LiDAR. Μετά από επαναλαμβανόμενες δοκιμές επιλέχθηκε η αριθμητική τιμή **1500** η οποία εκφράζει την υψηλή ανακλαστική ιδιότητα του ορόσημου. Στην **Εικόνα 47** (αριστερά) μπορούμε να δούμε το σύνολο των σημείων τα οποία είναι εντοπίσιμα από τον αισθητήρα LiDAR. Στην **Εικόνα 47** (δεξιά) φιλτράρουμε το σύνολο των σημείων κρατώντας μόνο τα σημεία που παρουσιάζουν υψηλή ανακλαστική τιμή (τεχνητά σημεία).



Εικόνα 47. Σύνολο αντικειμένων (λευκά σημεία) τα οποία εντοπίζονται από το LiDAR(αριστερά), φιλτράρισμα με βάση την ανακλαστική ιδιότητα των αντικειμένων (δεξιά).

Στην συνέχεια, έχοντας επιτυχημένα απομονώσει τις δέσμες που αφορούν μόνο το σημείο αναφοράς θα πρέπει να υπολογίσουμε την θέση της πλατφόρμας με βάση την εκάστοτε απόσταση που απέχουμε από αυτό. Αξίζει να αναφέρουμε σε κάθε χρονική στιγμή χρησιμοποιούμε **μία μόνο δέσμη** για να υπολογίσουμε την απόσταση. Η δέσμη που χρησιμοποιείτε κάθε φορά είναι η δέσμη που εντοπίζει την κοντινότερη απόσταση από το τεχνητό σημείο αναφοράς. Τέλος, για να βρούμε την απόσταση που έχουμε διανύσει σε δύο διαδοχικές χρονικές στιγμές θα πρέπει να γίνει η ανάλυση όπως περιγράφεται παρακάτω.



Εικόνα 48. Κάτοψη που περιέχει την θέση της ρομποτικής πλατφόρμας σε δύο διαδοχικές στιγμές ως προς το σημείο αναφοράς.

Βλέποντα **Εικόνα 48**, τα διανύσματα  $\vec{a}$  και  $\vec{b}$  αποτελούν τα διανύσματα της απόστασης εντοπίστηκε το σημείο αναφοράς την χρονική στιγμή  $t$  και  $t + 1$  αντίστοιχα. Αυτή η πληροφορία αποκτήθηκε χρησιμοποιώντας τον αισθητήρα LiDAR, ο οποίος μας δίνει την απόσταση ενός αντικειμένου σε πολικές συντεταγμένες. Έχοντας αυτά ως δεδομένο, θα πρέπει να βρούμε το διάνυσμα  $\vec{c}$  το οποίο αποτελεί το διάνυσμα της απόστασης για την χρονική στιγμή από  $t$  σε  $t + 1$ . Γνωρίζοντας ότι η πλατφόρμα έχει τρεις βαθμούς ελευθερίας (κίνηση  $x, y$  και  $roll$ ), μπορούμε να εκφράσουμε αυτήν την κίνηση με έναν μητρώο περιστροφής και ένα διάνυσμα μετατόπισης. Στην συγκεκριμένη διπλωματική εργασία χρησιμοποιήθηκε η υπόθεση ότι το διάνυσμα περιστροφής είναι πάνω στον άξονα  $z$  και το διάνυσμα μετατόπισης πάνω στον άξονα  $y$ , δηλαδή:

$$A_{t+1}^t = Rot(z, \theta)Tra(y, |\vec{c}|) \quad (5.1)$$

Προκύπτει λοιπόν ότι:

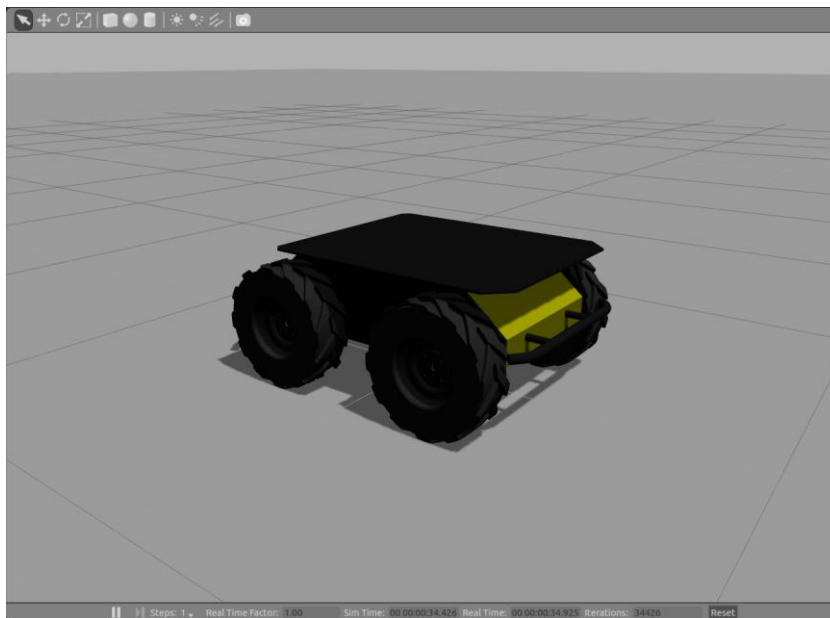
$$\vec{b} = A_{t+1}^t * \vec{a} \quad (5.2)$$

$$\begin{bmatrix} b_x \\ b_y \\ b_z \\ 1 \end{bmatrix} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 & 0 \\ \sin(\theta) & \cos(\theta) & 0 & |\vec{c}| \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_x \\ a_y \\ a_z \\ 1 \end{bmatrix} \quad (5.3)$$

Από την **Εξίσωση 5.3** καταλήγουμε σε ένα σύστημα δύο εξισώσεων και δύο μεταβλητών ( $\theta$  και  $|\vec{c}|$ ), πράγμα που μας δίνει την **θέση** της ρομποτικής πλατφόρμας για κάθε χρονική στιγμή. Η συγκεκριμένη μεθοδολογία αναπτύχθηκε με την βοήθεια της γλώσσας προγραμματισμού Python, λόγω της ευκολίας και της πληρότητας όσον αφορά τα εργαλεία ανάπτυξης και σχεδίασης που δίνει.

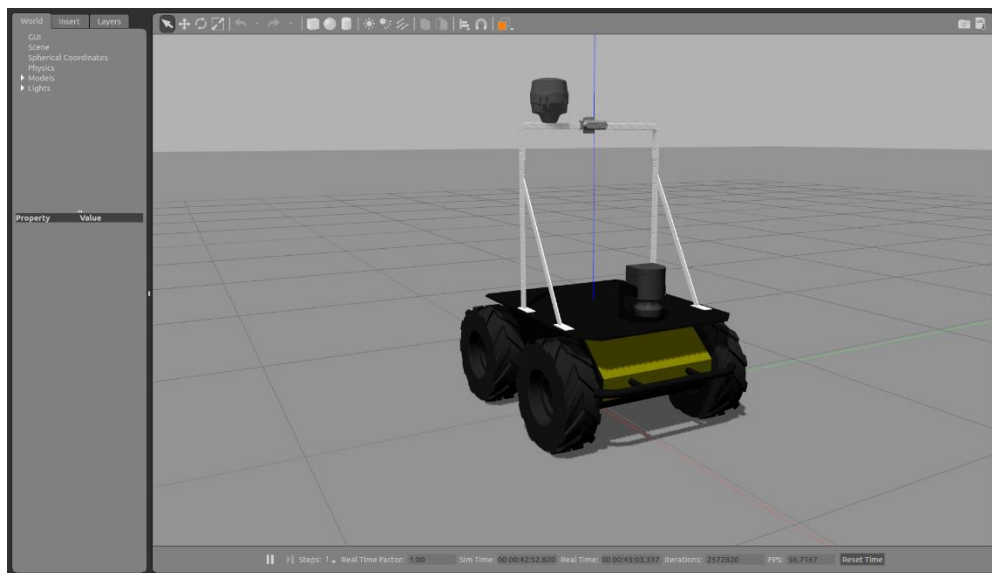
## 5.2 Προσομοίωση Πειράματος στο Gazebo

Όπως αναφέρθηκε και νωρίτερα, το κομμάτι της προσομοίωσης μίας ρομποτικής εφαρμογής αποτελεί ένα σημαντικό τμήμα. Γι' αυτό τον λόγο στην παρούσα διπλωματική εργασία, χρησιμοποιήθηκε το τρισδιάστατο μοντέλο που δημιουργήθηκε από την εταιρεία Clearpath, για την ρομποτική πλατφόρμα Husky-a200 [76].



Εικόνα 49. Ρομποτική πλατφόρμα στην μηχανή προσομοίωσης Gazebo.

Έχοντας λοιπόν την βασική λειτουργική δομή, κύριος στόχος ήταν να εγκατασταθούν τα αισθητήρια όργανα σε πραγματικό αλλά και ψηφιακό επίπεδο πάνω στην ρομποτική πλατφόρμα. Τα αισθητήρια όργανα τα οποία εγκαταστάθηκαν πάνω στο ρομπότ αποτελούνταν από μία RGB-D κάμερα (**Παράγραφος 2.3.6**), ένα 2D LiDAR (**Παράγραφος 2.3.4**) και ένα GPS (**Παράγραφος 2.3.3**). Έχοντας το τρισδιάστατο ψηφιακό μοντέλο καθενός αισθητήρα από τον κατασκευαστή, σχεδιάστηκε μία βάση με την βοήθεια του προγράμματος Solidworks η οποία και εγκαταστάθηκε στο πραγματικό αλλά και το ψηφιακό μοντέλο το οποίο αναπαραστάθηκε μέσω της μηχανής προσομοίωσης Gazebo. Αξίζει να αναφέρουμε ότι όλοι οι αισθητήρες αλλά και η βάση που προστέθηκε, γράφτηκε σε μορφή XML στο αρχείο URDF (**Παράγραφος 4**) του μοντέλου του ρομπότ.



Εικόνα 50. Ρομποτική πλατφόρμα με εγκατεστημένα τις βάσεις και τους αισθητήρες στην μηχανή προσομοίωσης Gazebo.

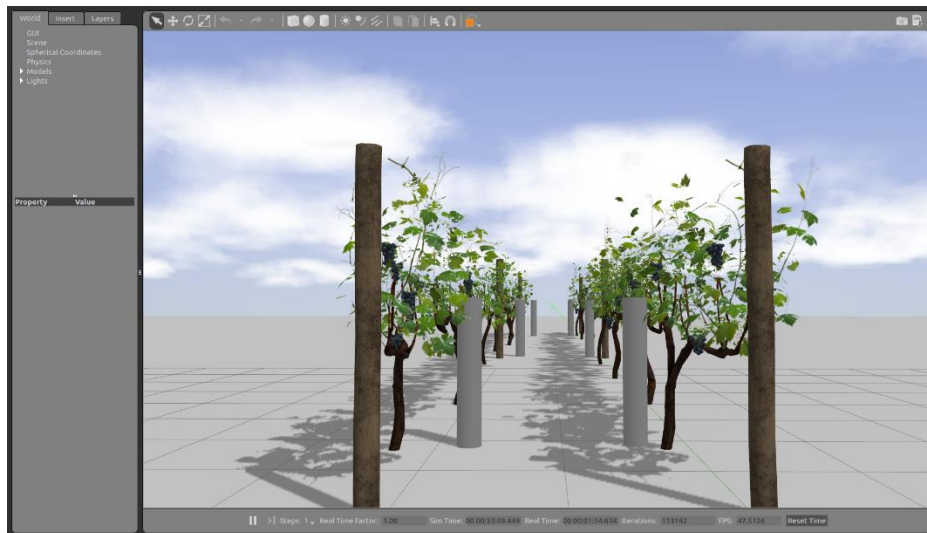
Έχοντας το τρισδιάστατο μοντέλο του ρομπότ, σαν επόμενο βήμα είναι να δημιουργηθούν τα ψηφιακά δεδομένα τα οποία πηγάζουν από τους διάφορους αισθητήρες. Η δημιουργία αυτών των δεδομένων γίνεται με την βοήθεια εφαρμογών οι οποίες ονομάζονται “**Επεκτάσεις**” (**Plugins**). Οι Επεκτάσεις είναι στην πραγματικότητα κομμάτια κώδικα τα οποία μεταγλωττίζονται και χρησιμοποιούνται ως κοινόχρηστες βιβλιοθήκες οι οποίες εισάγονται στην προσομοίωση. Ένα plugin έχει άμεση πρόσβαση σε όλες τις λειτουργίες του Gazebo, για την δημιουργία τους χρησιμοποιείται η γλώσσα προγραμματισμού C++ και η μεταφορά των πληροφοριών γίνεται μέσω ROS. [77]

Στην παρούσα διπλωματική τα plugins που προσομοιώνουν το GPS και την IMU ήταν ήδη εγκατεστημένα και λειτουργικά από την εταιρεία Clearpath. Χρειάστηκε να χρησιμοποιηθεί λοιπόν ένα plugin το οποίο προσομοιώνει έναν αισθητήρα 2D LiDAR το οποίο παρέχεται ήδη από το gazebo [78] το οποίο παραμετροποιήθηκε με τρόπο έτσι ώστε να παρουσιάζει τα ίδια χαρακτηριστικά με το πραγματικό. Τέτοια χαρακτηριστικά αποτελούν το εύρος σάρωσης, ο αριθμός των δεσμών, η συχνότητα αναμετάδοσης της πληροφορίας του αισθητήρα. Επιπρόσθετα, χρησιμοποιήθηκε ένα plugin το οποίο παρέχει πληροφορίες ίδιες με αυτές που δίνει μία κάμερα RGB-D. Η παραμετροποίηση αυτού του plugin συσχετίζεται με το μέγιστο βάθος που δίνει η κάμερα, τις εσωτερικές και εξωτερικές παραμέτρους, την απόσταση μεταξύ των φακών (baseline), την ανάλυση της χρωματικής κάμερα κ.α.

Εκτός από την ψηφιακή αναπαράσταση της πλατφόρμας και του ορόσημου, θα πρέπει να φτιαχτεί και το εξωτερικό περιβάλλον στο οποίο θα λειτουργήσει η προσομοίωση του πειράματος. Γι’ αυτό το λόγο



δημιουργήθηκε ένα μοντέλο ψηφιακού αμπελιού το οποίο χρησιμοποιήθηκε με τέτοιο τρόπο ορίζοντας μία γραμμή αμπελώνων.



Εικόνα 51. Περιβάλλον προσομοίωσης αμπελώνα με σημεία αναφοράς στο περιβάλλον Gazebo.

Σαν επόμενο βήμα σχεδιάστηκε ένα τεχνητό σημείο αναφοράς για να μπορούμε να παρακολουθούμε την θέση της πλατφόρμας σε κάθε χρονική στιγμή. Αξίζει να σημειωθεί ότι στην μηχανή προσομοίωσης Gazebo μπορούμε να παραμετροποιήσουμε κάποια επιθυμητά χαρακτηριστικά τα οποία θέλουμε να έχουν τα αντικείμενα μας. Τέτοια μπορεί να είναι το βάρος, την φωτεινότητα όπως και την ανακλαστική ιδιότητα την οποία μπορεί να έχει αυτό το υλικό. Έχοντας αυτήν την δυνατότητα δόθηκε μία τιμή μεγαλύτερη στα σημεία αναφοράς, πράγμα που σημαίνει ότι είμαστε σε θέση να φιλτράρουμε οποιαδήποτε τιμή χαμηλότερη και κατ' επέκταση να κρατήσουμε μόνο τις δέσμες που αφορούσαν τα τεχνητά σημεία.

Έχοντας λοιπόν δημιουργήσει το ψηφιακό περιβάλλον, το μοντέλο της ρομποτικής πλατφόρμας αλλά και των τεχνητών σημείων αναφοράς, ακολουθήθηκε η ίδια ακριβώς διαδικασία που θα χρησιμοποιούνταν και στο πείραμα. Αυτό σημαίνει ότι, η πλατφόρμα οδηγήθηκε σε μία ευθεία γραμμή και ταυτόχρονα όλες οι αισθητήριες πληροφορίες καταγράφηκαν σε ένα αρχείο rosbag. Στην συνέχεια, έχοντας αυτό το αρχείο δημιουργήθηκαν δύο οδομετρίες η οποίες συγκρίθηκαν σε σχέση με την πραγματικότητα (Groundtruth) την οποία περιγράψαμε στην προηγούμενη ενότητα.

### 5.3 Πείραμα σε Αμπελώνα

Το πείραμα για την σύγκριση και αξιολόγηση των δύο διαφορετικών EKF με GPS και κάμερα αντίστοιχα, διεξήχθη σε πειραματικό αμπελώνα τον Ιούλιο 2021. Ο αμπελώνας, ο οποίος βρίσκεται στην περιοχή Σπάτα (Αθήνα, Ελλάδα), έχει αποστάσεις φύτευσης στα 2,5 μέτρα και 1,25 με 1,35 μέτρα μεταξύ των φυτών πάνω στη σειρά. Σε μία γραμμή αμπελώνα τοποθετήθηκαν τρία τεχνητά στατικά σημεία αναφοράς με μεταξύ τους απόσταση τα 3 μέτρα. Η ρομποτική πλατφόρμα όπως φαίνεται και στην παρακάτω εικόνα τοποθετήθηκε σε 3 μέτρα απόσταση από το πρώτο σημείο αναφοράς με φορά κατεύθυνσης τα σημεία αναφοράς.



Εικόνα 52. Αρχική θέση ρομποτικής πλατφόρμας.

Στο πείραμα αξιολόγησης της εκτίμησης κατά την αυτόνομη οδήγηση του ρομπότ στον αμπελώνα, η ρομποτική πλατφόρμα οδηγήθηκε χειροκίνητα σε μία ευθεία απόσταση 9 μέτρων με σταθερή ταχύτητα  $0,4 \frac{m}{s}$ . Κατά την διάρκεια της κίνησης η RGB-D κάμερα είχε ενεργοποιηθεί καθώς επίσης ο αισθητήρας GPS-RTK ήταν σε κατάσταση “fix” πράγμα που σημαίνει ότι το μέγιστο σφάλμα μέτρησης είναι στα 0,02 μέτρα. Επίσης οι αισθητήρες LiDAR, IMU όπως επίσης και οι οπτικοί κωδικοποιητές παρείχαν σχετικές πληροφορίες κατά την διάρκεια της κίνησης. Όλες αυτές οι πληροφορίες αποθηκεύτηκαν στον υπολογιστή του ρομπότ για μετέπειτα επεξεργασία.

Έχοντας όλα αυτά δεδομένα σε μορφή rosbag, δημιουργήθηκαν δύο οδομετρικές εκτιμήσεις που βασίζονται στο Επεκταμένο Φίλτρο Κάλμαν (**Παράγραφος 3.5**) για τον υπολογισμό της θέσης και του προσανατολισμού της ρομποτικής πλατφόρμας χρησιμοποιώντας το πακέτο robot\_localization (**Παράγραφος 4.3**). Σημαντικό κομμάτι για την ορθή χρήση του Φίλτρου Κάλμαν η οποία υλοποιείται από



```
process_noise_covariance: [0.5, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0.5, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 1e-3, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0.3, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0.3, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0.01, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0.5, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0.5, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0.1, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0.3, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0.3, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0.3, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0.3, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0.3, 0.3]
```

```
initial_estimate_covariance: [1.0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 1.0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 1e-9, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 1.0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 1.0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 1e-9, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 1.0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 1.0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 1.0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 1.0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1.0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1.0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1.0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1.0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1.0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1.0]
```

Εικόνα 54. Πίνακες Συνδιακύμανσης για P και Q για την παραμετροποίηση του δεύτερου Επεκταμένου Φίλτρου Κάλμαν.

Στη πρώτη εκτίμηση χρησιμοποιήθηκαν τα δεδομένα που παρήχθησαν από τους αισθητήρες GPS, IMU και τους οπτικούς κωδικοποιητές και στην δεύτερη η κάμερα, η IMU και οι οπτικοί κωδικοποιητές. Για την σύγκριση των δύο χρησιμοποιήθηκε μία εκτίμηση με την χρήση του LiDAR με την μεθοδολογία όπως παρουσιάστηκε στην **Παράγραφο 5.1**. Αξίζει να αναφέρουμε ότι, κάθε εκτίμηση της ρομποτικής πλατφόρμας καταγράφηκε σε ένα αρχείο σύμφωνα με το πρότυπο το οποίο αναπτύχθηκε από το Πολυτεχνείο του Μονάχου (TUM file format) [79].

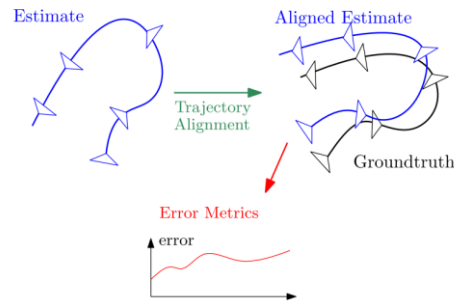
Τέλος, για την σύγκριση και εμφάνιση των αποτελεσμάτων χρησιμοποιήθηκε το πρόγραμμα eno, το οποίο είναι πακέτο γραμμένο σε γλώσσα προγραμματισμού Python και παρέχει μία βιβλιοθήκη για την αξιολόγηση και σύγκριση της οδομετρίας και διαφόρων υλοποιήσεων SLAM [80].

## 5.4 Αποτελέσματα

Σε αυτήν την ενότητα θα παρουσιάσουμε τα αποτελέσματα που παρήχθησαν, από τις δυο εκτιμήσεις (GPS και κάμερα) σε σχέση με την πραγματική θέση της πλατφόρμας και δημιουργήθηκαν από τον πειραματικό σχεδιασμό όπως περιγράφηκε στην προηγούμενη παράγραφο. Για την ανάλυση αλλά και την απεικόνιση τους, όπως αναφέρθηκε και νωρίτερα χρησιμοποιήθηκε το πρόγραμμα eno.

Κάθε οδομετρία που δημιουργήθηκε από τις δύο εκτιμήσεις αλλά και η οδομετρία που δημιουργήθηκε χρησιμοποιώντας τα σημεία αναφοράς (Groundtruth), αποτελείται από διαδοχικές χρονικές στιγμές οι οποίες συνοδεύονται από μία θέση στο δισδιάστατο επίπεδο (x και y) καθώς επίσης και έναν προσανατολισμό εκφρασμένο συνήθως σε γωνίες Euler. Συνθέτοντας όλες αυτές τις πόξες τις ρομποτικής πλατφόρμας σχεδιάζεται ένας δρόμος (trajectory) ο οποίος εκφράζει την συνολική κίνηση της πλατφόρμας. Για να μπορέσει να γίνει μία ποσοτική σύγκριση των εκτιμώμενων δρόμων σε σχέση με την πραγματικότητα θα πρέπει να λάβουμε υπόψη δύο σημαντικά προβλήματα. Το πρώτο είναι ότι κάθε τροχιά

είναι εκφρασμένη σε διαφορετικό σύστημα συντεταγμένων οπότε η άμεση σύγκρισή τους αποτελεί σημαντικό λάθος [81]. Το δεύτερο πρόβλημα είναι ότι ένας δρόμος αποτελείται από πόζες σε διαφορετικές χρονικές στιγμές και ως εκ τούτου αποτελεί πολυδιάστατο δεδομένο [81]. Για την αντιμετώπιση του πρώτου προβλήματος το οποίο αναφέρεται στην βιβλιογραφία και ως “**Ευθυγράμμιση Δρόμου**” (**Trajectory Alignment**) θα πρέπει να χρησιμοποιηθεί μία αλγοριθμική δομή με κατάλληλους μετασχηματισμούς ο οποίος θα μεταφέρει τους δύο δρόμους στο ίδιο σύστημα συντεταγμένων. Όσον αφορά την λύση του δεύτερου προβλήματος θα πρέπει να χρησιμοποιηθούν ουσιαστικές μετρήσεις σφάλματος των οποίων οι ιδιότητες θα πρέπει να είναι καλά κατανοητές.



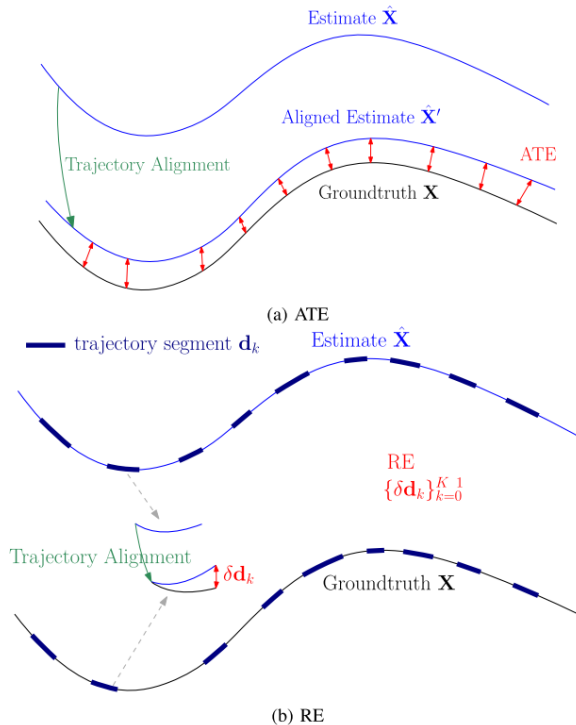
Εικόνα 55. Η διαδικασία της αξιολόγησης ενός δρόμου ξεκινάει με την ευθυγράμμιση με την πραγματική εκτίμηση και στην συνέχεια υπολογίζεται το σφάλμα με συγκεκριμένη μεθοδολογία μέτρησης σφάλματος. Πηγή [81]

Σύμφωνα με τους Zichao και Scaramuzza [81], το πρόβλημα της οπτικής οδομετρίας το οποίο ανάγεται στην εύρεση ενός μετασχηματισμού από δύο διαδοχικές πόζες στον χώρο μπορεί να εκφραστεί με πολλούς διαφορετικούς μετασχηματισμούς οι οποίοι είναι ισοδύναμοι μεταξύ τους δημιουργώντας έτσι μία ασάφεια. Αυτή η ασάφεια περιπλέκει την διαδικασία αξιολόγησης ενός δρόμου καθώς δεν μπορούμε να πάρουμε την άμεση διαφορά από την πραγματική τιμή χρησιμοποιώντας για παράδειγμα την Ευκλείδεια απόσταση μεταξύ μίας εκτίμησης θέσης  $X_{est}$  και της πραγματικής τιμής  $X_{groundtruth}$ . Αντ’ αυτού, θα πρέπει να υπολογίσουμε πρώτα τον διανυσματικό υποχώρο όλων των διανυσμάτων που εμφανίζουν ισοδύναμους μετασχηματισμούς και στην συνέχεια να τον συγκρίνουμε με την πραγματική τιμή. Η συνήθης πρακτική είναι με την εύρεση μίας ισοδύναμης εκτίμησης  $X_{est}$  κοντινής σε απόσταση από την πραγματική  $X_{groundtruth}$  όσον αφορά την απόσταση. Αυτή η διαδικασία όπως αναφέραμε και προηγουμένως ονομάζεται Ευθυγράμμιση Δρόμου και η πιο δημοφιλής αλγοριθμική δομή που χρησιμοποιείται για την επίλυσή της ονομάζεται μέθοδος Umeyama’s [82]. Αυτή η μέθοδος χρησιμοποιείται για την ευθυγράμμιση δύο συνόλων που αποτελούνται από σημεία, με την εύρεση ενός μετασχηματισμού ο οποίος λαμβάνει υπόψη την ελαχιστοποίηση του μέσου τετραγωνικού σφάλματος (root mean square error – RMSE).

Έχοντας λοιπόν δύο δρόμους (εκτίμηση και πραγματική τιμή) οι οποίοι είναι ευθυγραμμισμένοι από το προηγούμενο βήμα, σκοπός είναι η σύγκριση αυτών καθ’ αυτών χρησιμοποιώντας έναν κατάλληλο μετρητή σφάλματος. Σε αυτό το βήμα χρησιμοποιούνται δύο γνωστοί μετρητές με ονομασία “**Απόλυτο Σφάλμα Δρόμου**” (**Absolute Trajectory Error** ή **ATE**) και “**Σχετικό Σφάλμα**” (**Relative Error** ή **RE**), οι οποίοι συνοδεύονται από κάποια πλεονεκτήματα και μειονεκτήματα αντίστοιχα.

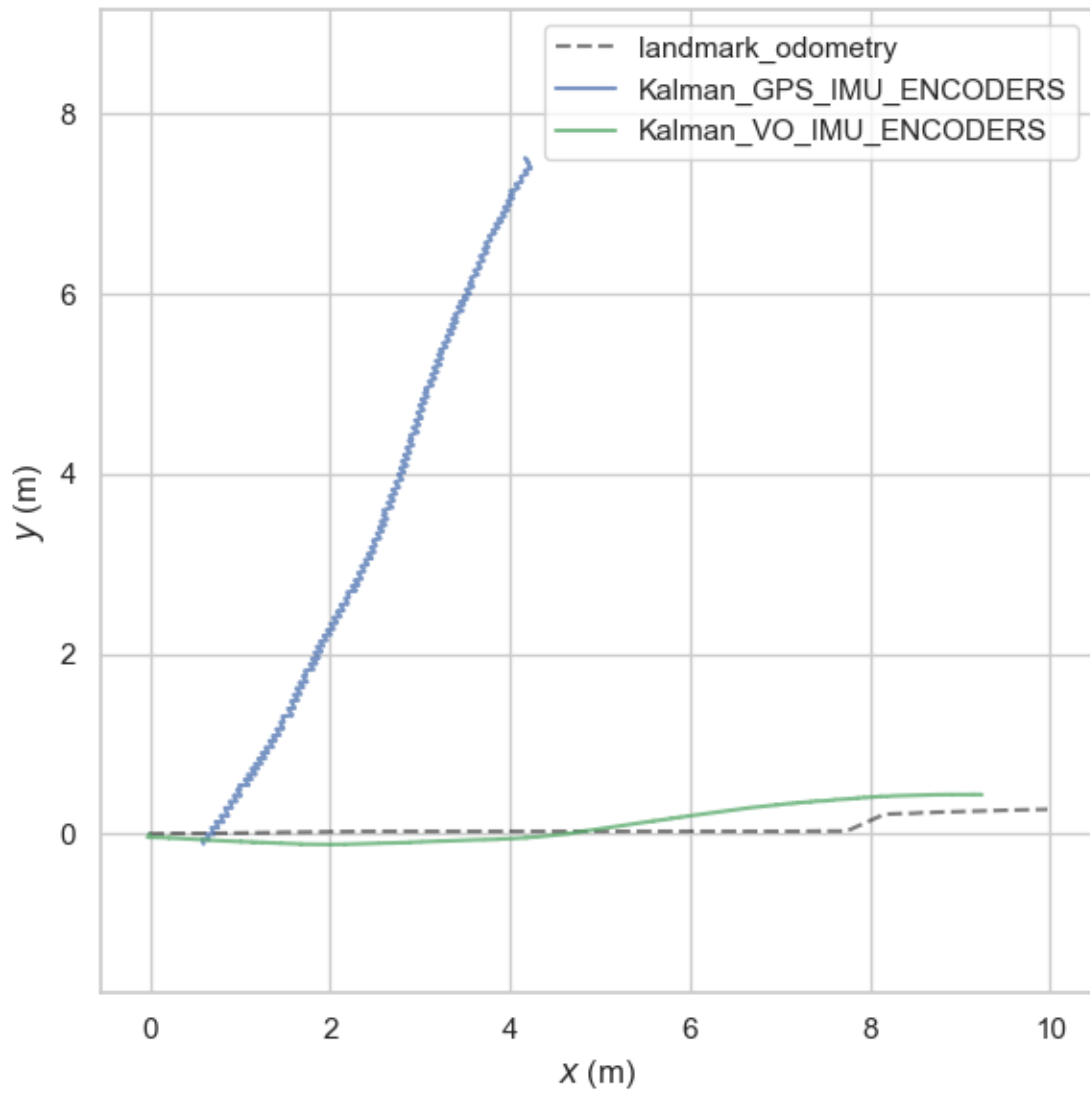
Για την σύγκριση των θέσεων αλλά και προσανατολισμού διαδοχικών σημείων σε έναν δρόμο χρησιμοποιείται συνήθως το μέσο τετραγωνικό σφάλμα (RMSE). Η απλότητα αυτής της μεθόδου για την σύγκριση αποτελεί σημαντικό πλεονέκτημα. Ωστόσο πολλοί ερευνητές παρατήρησαν ότι αυτή η μέθοδος δεν παρουσιάζει ευαισθησία την χρονική στιγμή του σφάλματος [83]. Μία δεύτερη εκτίμηση του σφάλματος παρέχει το Σχετικό Σφάλμα, όπου ο δρόμος χωρίζεται σε μικρότερα κομμάτια σύμφωνα με

κάποια παράμετρο (π.χ. απόσταση). Έχοντας ένα μικρότερο τμήμα, μπορούμε στην συνέχεια να χρησιμοποιήσουμε την ευθυγράμμιση δρόμου και να υπολογίσουμε το σφάλμα με τον ίδιο τρόπο όπως το Απόλυτο Σφάλμα. Δεδομένου ότι το σχετικό σφάλμα δεν αντιπροσωπεύεται από έναν συγκεκριμένο αριθμό αλλά μια συλλογή σφαλμάτων για όλα τα τμήματα του δρόμου, μπορούν να υπολογιστούν στατιστικά στοιχεία όπως ο διάμεσος, ο μέσος όρος και ποσοστό σφάλματος, τα οποία δίνουν περισσότερες πληροφορίες από την μέθοδο ATE. Το μειονέκτημα του RE είναι ότι είναι μία σχετικά περίπλοκη μέθοδος λιγότερο προφανής μέθοδος για να κατατάξουμε την ποιότητα εκτίμησης έναντι στην μέθοδο ATE. Τέλος αξίζει να αναφερθεί ότι οι συγκεκριμένες μετρήσεις έχουν μία μεροληπτική βάση (bias), της οποίας το σφάλμα είναι η Ευκλείδεια απόσταση της εκτίμησης με την πραγματικότητα (Groundtruth).



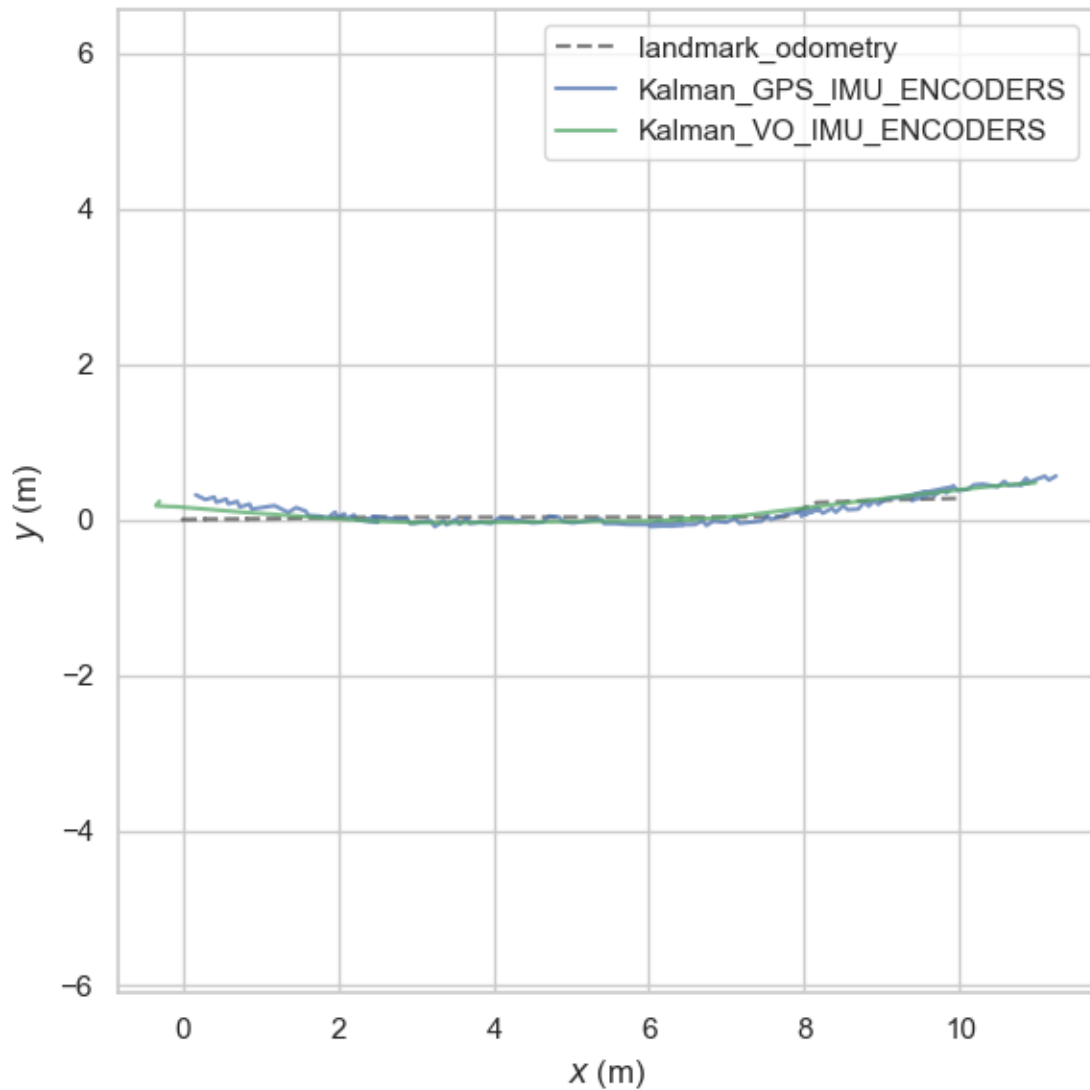
Εικόνα 56. Απεικονίσεις του Απόλυτου Σφάλματος Δρόμου και Σχετικού Σφάλματος. Πηγή [81]

Στην **Εικόνα 57** φαίνονται οι τρεις διαφορετικές εκτιμήσεις της θέσης χρησιμοποιώντας κάμερα, (μπλε χρώμα) χρησιμοποιώντας GPS (πράσινο χρώμα) ενώ η διακεκομμένη μαύρη γραμμή αποτελεί την πραγματική θέση στον χώρο (Groundtruth). Όπως φαίνεται στην εικόνα οι δρόμοι δεν μπορούν να συγκριθούν άμεσα διότι κάθε ένας από αυτούς είναι εκφρασμένος σε διαφορετικό σύστημα συντεταγμένων και ως αποτέλεσμα έχει διαφορετική διεύθυνση και κλίμακα.



Εικόνα 57 Γράφημα που περιέχει τους δρόμους που δημιουργήθηκαν από τις διάφορες εκτιμήσεις αλλά και της πραγματικής θέσης στο χώρο.

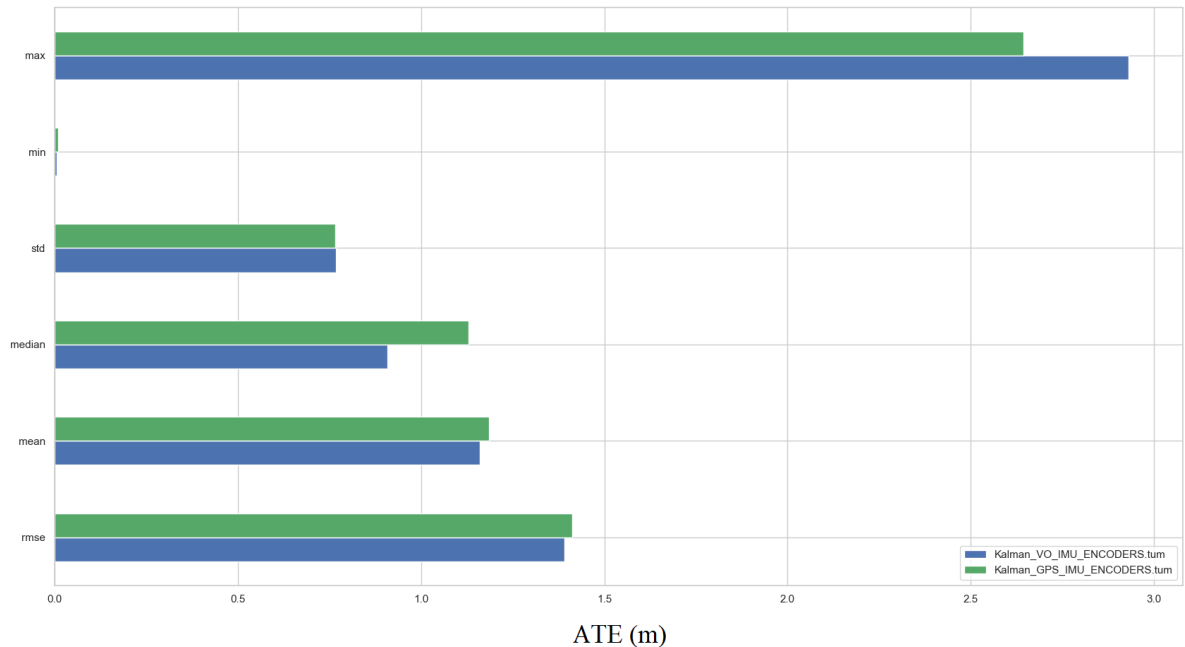
Χρησιμοποιώντας την μέθοδο Umeyama για την ευθυγράμμιση των δύο δρόμων σε σχέση με τον δρόμο που δημιουργήθηκε από τα σημεία αναφοράς, παίρνουμε τα αποτελέσματα όπως φαίνονται στην **Εικόνα 58**.



Εικόνα 58. Γράφημα που περιέχει τους ευθυγραμμισμένους δρόμους που δημιουργήθηκαν από τις διάφορες εκτιμήσεις αλλά και της πραγματικής θέσης στο χώρο με τον αλγόριθμο Uteyama.

Για την μέτρηση του σφάλματος για κάθε μία από τις δύο εκτιμήσεις χρησιμοποιήθηκαν οι δύο μετρητές ATE (Absolute Trajectory Error) και RE (Relative Error). Ο πρώτος μετρητής χρησιμοποιήθηκε για να εξεταστεί η **συνολική επίδοση** της εκάστοτε εκτίμησης σε σχέση με την πραγματική τιμή, ενώ ο δεύτερος για να χαρακτηριστεί σε η ακρίβεια σε **τοπικό επίπεδο**. Όπως μπορούμε να δούμε από την **Εικόνα 59**, ο **μέσος όρος του σφάλματος** της οπτικής οδομετρίας έχει **μικρότερα σφάλματα** σε σχέση με την μέτρηση του GPS, παρότι οι μετρήσεις της οδομετρίας με χρήση GPS έχουν μικρότερο μέγιστο σφάλμα σε σχέση με την χρήση της κάμερας.





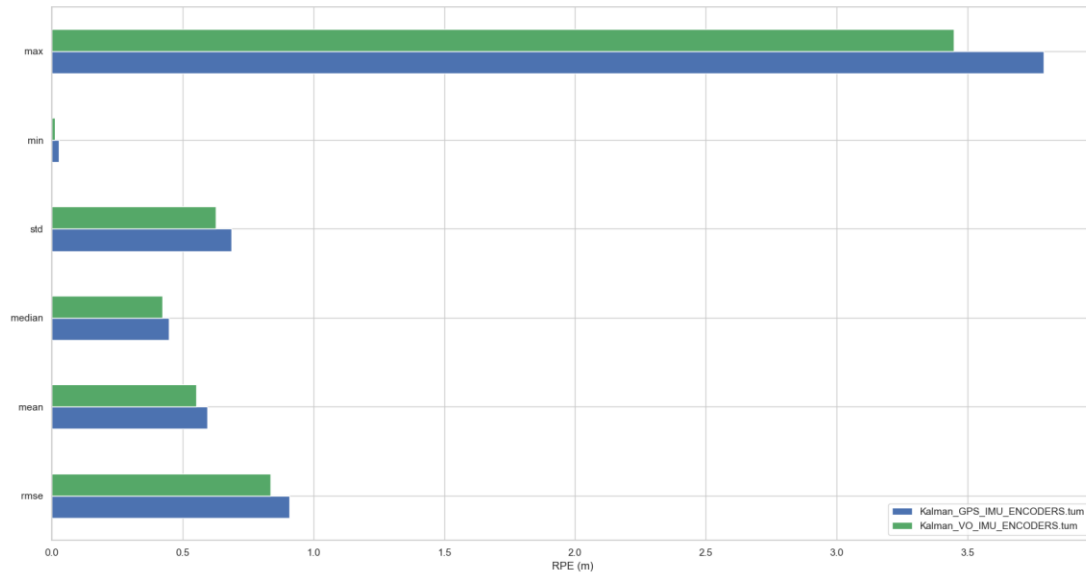
Εικόνα 59. Μέγιστη και ελάχιστη τιμή, μέσος όρος και διάμεσος τιμών, διακύμανση και μέσου τετραγωνικού σφάλματος του μετρητή ATE (Absolute Trajectory Error) για την εκτίμηση με χρήση GPS (πράσινο χρώμα) και για την εκτίμηση με χρήση Κάμερας (μπλε χρώμα).

Και σε μορφή Πίνακα

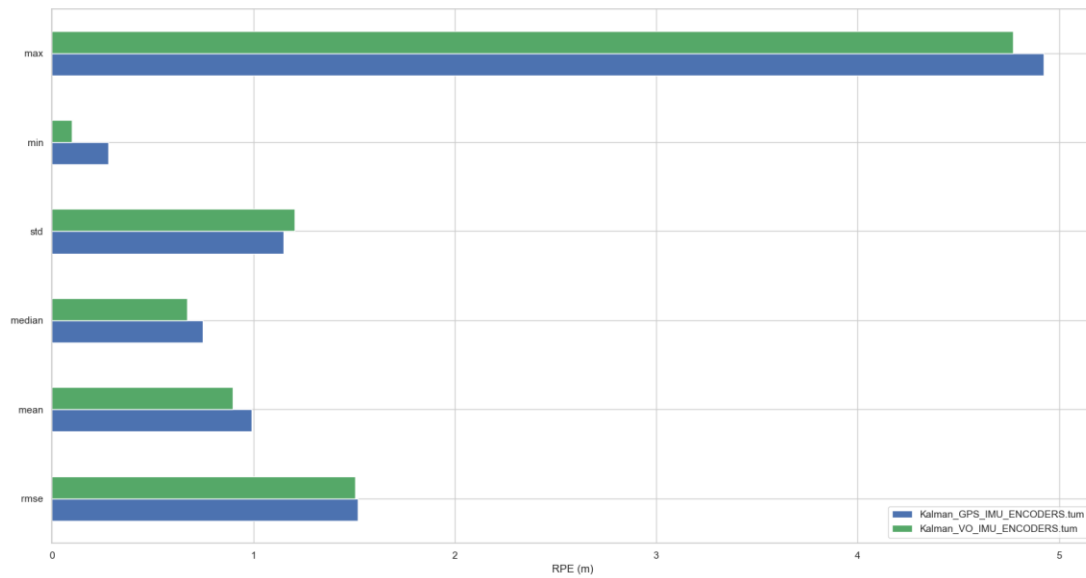
	Rmse(m)	Mean(m)	Median(m)	Std(m)	Min(m)	Max(m)
<b>GPS</b>	1.411699	1.18604	1.129886	0.765639	0.00831	2.645071
<b>VO</b>	1.390569	1.160025	0.907171	0.766827	0.005895	2.931806

Πίνακας 2. Πίνακας που περιέχει τις τιμές της μέγιστης και ελάχιστης τιμής, τον μέσο όρος και την διάμεσο τιμών, την διακύμανση και το μέσο τετραγωνικό σφάλμα του μετρητή ATE (Absolute Trajectory Error) για την εκτίμηση με χρήση GPS και για την εκτίμηση της οπτικής οδομετρίας.

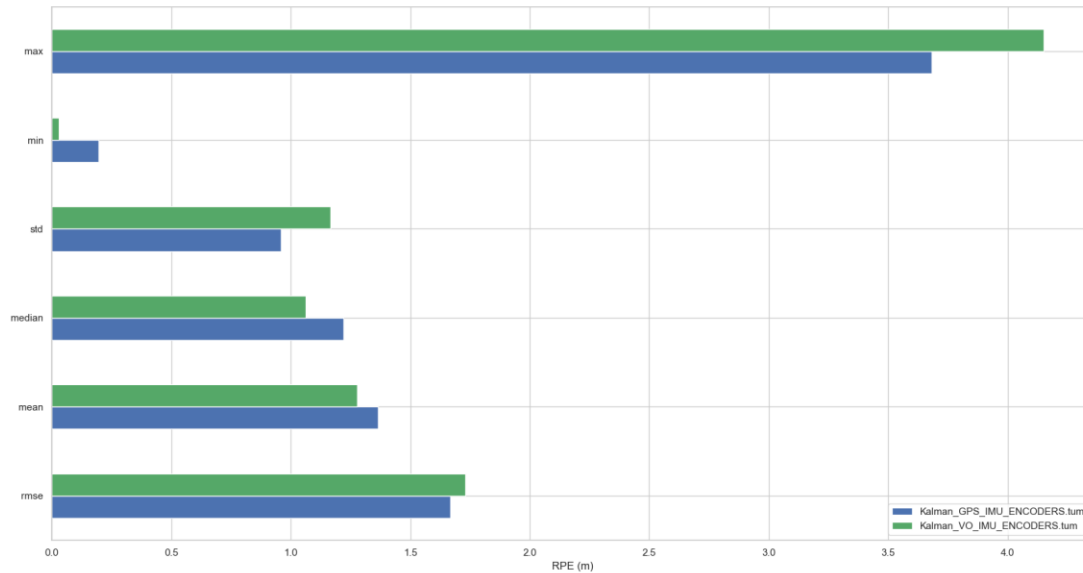
Για την εξέταση του τοπικού σφάλματος χρησιμοποιήθηκε ο μετρητής RE (Relative Error) με βάση την δημιουργία διαδοχικών τμημάτων στην εκάστοτε εκτίμηση εκφρασμένη σε απόσταση (μέτρα). Για την δημιουργία διαδοχικών τμημάτων επιλέχθηκαν οι τιμές **1.5 μέτρο**, **1.0 μέτρο**, **0.5 μέτρα** οι οποίες έδωσαν μία μεγαλύτερη εικόνα για την τοπικό σφάλμα κάθε μίας εκτίμησης σε σχέση με την πραγματικότητα. Στις επόμενες τρεις εικόνες φαίνονται τα αποτελέσματα για κάθε διαφορετική τιμή της παραμέτρου για τον μετρητή RE για τις δύο εκτιμήσεις.



Εικόνα 60. Μέγιστη και ελάχιστη τιμή, μέσος όρος και διάμεσος τιμών, διακύμανση και μέσο τετραγωνικό σφάλμα για το RE με βάση την απόσταση 0.5 μέτρων.



Εικόνα 61. Μέγιστη και ελάχιστη τιμή, μέσος όρος και διάμεσος τιμών, διακύμανση και μέσο τετραγωνικό σφάλμα για το RE με βάση την απόσταση 1 μέτρον.



Εικόνα 62. Μέγιστη και ελάχιστη τιμή, μέσος όρος και διάμεσος τιμών, διακύμανση και μέσο τετραγωνικό σφάλμα για το RE με βάση την απόσταση 1.5 μέτρον.

Στον παρακάτω Πίνακα συνοψίζουμε τις τιμές της μέγιστης, ελάχιστης, μέσης τιμής, της διακύμανση του τοπικού σφάλματος για τις δύο εκτιμήσεις ανά διαφορετική παράμετρο.

	<b>Rmse(m)</b>	<b>Mean(m)</b>	<b>Median(m)</b>	<b>Std(m)</b>	<b>Min(m)</b>	<b>Max(m)</b>
GPS με παράμετρο 1.5 μέτρο	1.668263	1.365958	1.219342	0.957736	0.19558	3.681922
VO με παράμετρο 1.5 μέτρο	1.729689	1.276643	1.063465	1.16705	0.02933	4.151513
GPS με παράμετρο 1 μέτρο	1.517633	0.991048	0.747935	1.149363	0.278803	4.92492
VO με παράμετρο 1 μέτρο	1.504759	0.898958	0.671792	1.206721	0.100229	4.770477
GPS με παράμετρο 0.5 μέτρα	0.90823	0.595154	0.446935	0.686056	0.027145	3.793027
VO με παράμετρο 0.5 μέτρο	0.835691	0.553271	0.423303	0.626314	0.010553	3.44751

Πίνακας 3. Μέση τιμή και την διακύμανση για τον μετρητή RE, ανά διαφορετική παράμετρο και οδομετρική εκτίμηση.

## Κεφάλαιο 6

### Συμπεράσματα

Στην παρούσα διπλωματική εργασία παρουσιάστηκε μία μεθοδολογία για την σύγκριση δύο διαφορετικών εκτιμήσεων θέσης με χρήση GPS και κάμερας για το πρόβλημα του αυτοεντοπισμού μίας ρομποτικής πλατφόρμας σε ένα αγροτικό περιβάλλον. Όπως αναφέραμε και νωρίτερα, το κομμάτι του αυτοεντοπισμού αποτελεί μείζον ζήτημα για την αυτοματοποίηση μίας ρομποτικής κινητής πλατφόρμας η οποία λειτουργεί σε ένα δεδομένο περιβάλλον. Επίσης αξίζει να αναφερθεί πως ένα αγροτικό περιβάλλον χαρακτηρίζεται ως μία δύσκολη περίπτωση για σωστή εκτίμηση έναντι των εσωτερικών χώρων που συνήθως χρησιμοποιούνται βιομηχανικά ρομπότ. Ένα αγροτικό περιβάλλον λόγω του ανώμαλου εδάφους, την απρόβλεπτη γεωμετρία των φυτών αλλά και την ανάπτυξη υψηλών θερμοκρασιών δυσχεραίνει κατά πολύ την σωστή πρόβλεψη και μελέτη του προβλήματος.

Είδαμε λοιπόν ότι αποτελεί συνήθη πρακτική για την επίλυση του προβλήματος του αυτοεντοπισμού η χρήση ενός αισθητήρα GPS η οποία χρησιμοποιείται και σε εμπορικές πλέον λύσεις. Καθώς ένας αισθητήρας GPS χαρακτηρίζεται από πολλούς ως μία λειτουργική λύση για την εκτίμηση της θέσης, ένας τέτοιος αισθητήρας δεν χαρακτηρίζεται από την ευρωστία του κατά την διάρκεια λειτουργίας. Αυτό στην πραγματικότητα σημαίνει ότι είναι πολύ σύνθηες να χάνεται το σήμα GPS λόγω της πυκνής βλάστησης που υπάρχει σε έναν αμπελώνα πράγμα που εισάγει πολύ μεγάλο θόρυβο στην εκτίμηση θέσης. Ως εκ τούτου δοκιμάστηκε η χρήση της οπτικής οδομετρίας για την εκτίμηση της θέσης της ρομποτικής πλατφόρμας σαν άμεση αντικατάσταση ενός αισθητήρα GPS για την λειτουργία σε έναν πραγματικό αμπελώνα.

Στο **Κεφάλαιο 5.1**, παρουσιάστηκε μία μεθοδολογία για την σύγκριση των δύο εκτιμήσεων (κάμερα και GPS) χρησιμοποιώντας έναν διαφορετικό αισθητήρα με σκοπό την δημιουργία μίας εκτίμησης πραγματικής θέσης η οποία δοκιμάστηκε σε μηχανή προσομοίωσης πριν της πραγματοποίηση του πραγματικού πειράματος σε αμπελώνα.

Τέλος στο **Κεφάλαιο 5.4** είδαμε ότι σύγκριση των δύο εκτιμήσεων δεν αποτελεί μία απλή διαδικασία και για μπορέσει να επιτευχθεί με επιτυχία θα πρέπει να χρησιμοποιηθούν οι σωστές μεθοδολογίες και μετρήσεις οι οποίες δίνουν μία πιο έγκυρη εικόνα των αποτελεσμάτων. Βλέποντας το διάγραμμα στην **Εικόνα 58** μπορούμε να δούμε πως ο δρόμος που δημιουργήθηκε χρησιμοποιώντας την πληροφορία από το GPS εμφανίζει κορυφές και γρήγορες αλλαγές από σημείο σε σημείο. Αυτό οφείλεται στην φύση της πληροφορίας που εισάγει ένας τέτοιος αισθητήρας, ο οποίος στην ουσία υπολογίζει μέσα σε ένα νοητό κύκλο 2 εκατοστών (σε περίπτωση που το σύστημα είναι σε κατάσταση fix) την θέση της πλατφόρμας. Αυτό σημαίνει αυτόματα ότι θα υπάρχουν διακυμάνσεις κάθε φορά που ο αισθητήρας GPS δίνει αυτή την πληροφορία στο φίλτρο Κάλμαν. Αντίθετα η οπτική οδομετρία παρουσιάζει έναν δρόμο χωρίς μεγάλες διακυμάνσεις και κορυφές ανάμεσα σε δύο διαδοχικές χρονικές στιγμές. Αυτό οφείλεται στο γεγονός ότι κάθε επόμενη θέση στηρίζεται στην προηγούμενη για να υπολογιστεί η νέα θέση της πλατφόρμας. Για κάθε νέα εκτίμηση θέσης όμως, εισάγεται ένα μικρό σφάλμα το οποίο συσσωρεύεται σε μεγάλο χρονικό διάστημα.

Παρατηρώντας τον **Πίνακα 2** μπορούμε να δούμε ότι το σφάλμα ATE δεν μας δίνει και πολλές πληροφορίες καθώς η μέση τιμή, η διακύμανση και η ελάχιστη και μέγιστη τιμή έχουν πολύ μικρές αποκλίσεις. Η μόνη πληροφορία που μπορούμε να πάρουμε από αυτήν την μέτρηση είναι η διάμεσος.

Παρατηρούμε λοιπόν ότι η διάμεσος της εκτίμησης με χρήση GPS είναι μεγαλύτερη, πράγμα που σημαίνει ότι οι τιμές αυτής της εκτίμησης “επικεντρώνονται” γύρω από μία μεγαλύτερη τιμή σφάλματος.

Αντίθετα κοιτάζοντας τον μετρητή APE, μπορούμε να πάρουμε μία περισσότερη πληροφορία σχετικά με την τοπική ακρίβεια των δύο εκτιμήσεων. Από την **Εικόνα 61** μπορούμε να δούμε ότι παρ’ ότι το σφάλμα σε κάθε διαφορετική παραμετροποίηση έχει υψηλότερη τιμή για την εκτίμηση θέσης με την χρήση του GPS σε σχέση με την οπτική οδομετρία. Το συγκεκριμένο δικαιολογείται με δεδομένο ότι ο αισθητήρας GPS δίνει πληροφορία η οποία είναι διακριτή και μη συνεχείς, σε σχέση με την οπτική οδομετρία η οποία δημιουργεί ένα συνεχές δρόμο. Αυτό το σφάλμα που εισάγεται κατά την διάρκεια μέτρησης του αισθητήρα GPS εισάγει μεγαλύτερο “τοπικό” σφάλμα σε σχέση με την οπτική οδομετρία και άρα η διάμεσος αλλά και η μέση τιμή θα έχει μεγαλύτερο τοπικό σφάλμα. Παρ’ ότι λοιπόν ένας αισθητήρας GPS θεωρείτε ένας πιο ακριβής αισθητήρας για την μέτρηση σε εξωτερικό περιβάλλον, στο συγκεκριμένο πειραματικό σχεδιασμό μπορούμε να πούμε ότι η εκτίμηση της οπτικής οδομετρίας σε ολικό επίπεδο έχει παρόμοια συμπεριφορά με την χρήση του GPS καθώς επίσης σε τοπικό επίπεδο παρουσιάζει καλύτερα αποτελέσματα.

## Κεφάλαιο 7

### Μελλοντική Εργασία

Στην παρούσα διπλωματική παρουσιάστηκε μία μεθοδολογία σύγκρισης δύο εκτιμήσεων θέσης με χρήση GPS και κάμερας σε σχέση με μία πραγματική τιμή η οποία δημιουργήθηκε με χρήση του αισθητήρα LiDAR. Σαν περαιτέρω βελτιστοποιήσεις και επόμενους στόχους προτείνεται:

- Επέκταση της παρούσας μεθοδολογίας για περισσότερα μέτρα όπως επίσης και κάλυψη στροφής για να εξεταστεί η οπτική οδομετρία σε στροφές  $180^\circ$ , όπως συνηθίζεται για αλλαγή σειράς αμπελώνα (U-turn).
- Εξέταση μόνο της οπτικής οδομετρίας χρησιμοποιώντας ως πραγματική εκτίμηση την θέση από αισθητήρα GPS.
- Χρήση διαφόρων οπτικών Χαρακτηριστικών όπως SURF,SIFT κ.α. καθώς επίσης και σύγκριση μεταξύ της επίδοσης τους.
- Σύγκριση επίδοσης οπτικής οδομετρίας με βάση την ταχύτητα του οχήματος.
- Σύγκριση διαφόρων άλλων πακέτων οπτικής οδομετρίας όπως Kimera [84], OrbSlam3 [85], RGBDSLAMv2 [86].

## Βιβλιογραφία

- [1] European Commition, "Agrobofood Project," 2019. [Electronic Source]. Available: <https://agrobofood.eu/>. [Access 2 September 2021].
- [2] European Commition, "Robots For Protecting the Crops," February 2021. [Electronic Source]. Available: <https://robs4crops.eu/>. [Access 02 September 2021].
- [3] Agreenculture, "AgreenCulture's Ceol Robot," Agreenculture, [Electronic Source]. Available: <https://www.agreenculture.net/ceol>. [Access 2 September 2021].
- [4] Agrointelli, "Robboti Robot," Agrointelli, [Electronic Source]. Available: <https://www.agrointelli.com/robboti/>. [Access 2 September 2021].
- [5] "Farmdroid," Farmdroid, [Electronic Source]. Available: <https://farmdroid.dk/>. [Access 2 September 2021].
- [6] M. Perez-Ruiz and S. K. Upadhyaya, "GNSS in Precision Agricultural Operation," *New Approach of Indoor and Outdoor Localization Systems*, London, InTech, 2012. DOI: 10.13140/2.1.4162.3362.
- [7] J. Guo, X. Li, Z. Li, L. Hu, G. Yang, C. Zhao, D. Fairbairn, D. Watson and M. Ge, "Multi-GNSS precise point positioning for precision agriculture," *Precision Agriculture*, vol. 19, pp. 895-911, 2018. DOI: <https://doi.org/10.1007/s11119-018-9563-8>.
- [8] S. Zaman, L. Comba, A. Biglia, D. Aimonino, P. Barge and P. Gay, "Cost-effective visual odometry system for vehicle motion control in agricultural environments," *Computers and Electronics in Agriculture*, vol. 162, pp. 82-94, 2019. DOI: <https://doi.org/10.1016/j.compag.2019.03.037>.
- [9] N. Chebrolu, P. Lottes, T. Labe and C. Stachniss, "Robot Localization Based on Aerial Images for Precision Agriculture Tasks in Crop Fields," *In: Proceedings of the 2019 International Conference on Robotics and Automation (ICRA)*, Montreal, 2019. DOI: 10.1109/ICRA.2019.8794030.
- [10] L. Piyathilaka and R. Munasinghe, "Vision-only outdoor localization of two-wheel tractor for autonomous operation in agricultural fields," *In: Proceedings of the 2011 6th International Conference on Industrial and Information Systems*, Kandy, Sri Lanka, 2011. DOI: 10.1109/ICIINFS.2011.6038094.
- [11] L. Santos, A. Aguiar, F. Santos, A. Valente, J. Ventura and A. Sousa, "Navigation Stack for Robots Working in Steep Slope Vineyard," *Advances in Intelligent Systems and Computing*, New York, NY, USA, Springer International Publishing, 2020, p. 264–285.
- [12] R. Siegwart, I. R. Nourbakhsh and D. Scaramuzza, "Introduction to Autonomous Mobile Robots, second edition (Intelligent Robotics and Autonomous Agents series)," 2011, p. 80.

- [13] R. N. Jazar, "Chapter 7: Steering Dynamics," In *Vehicle Dynamics: Theory and Application*, Springer Science and Business Media, LLC, 2009.
- [14] K. Spentzas, I. Alkhazali and M. Demic, "Kinematics of four-wheel-steering vehicles," *Forschung im Ingenieurwesen*, vol. 66, pp. 211-216, 2001.
- [15] G. Dudek and M. Jenkin, *Computational Principles of Mobile Robotics 2nd Edition*, Cambridge University Press, 2010.
- [16] Clearpath, "HUSKY Unmanned Ground Vehicle", [Electronic Source]. Available: <https://clearpathrobotics.com/husky-unmanned-ground-vehicle-robot/>. [Access 5 September 2021].
- [17] ROS, "ROS:Kinetic Kame," ROS, [Electronic Source]. Available: <http://wiki.ros.org/kinetic>. [Access 5 September 2021].
- [18] K. Kalovrentis and K. Katevas, "Μετρήσεις Θέσης και Μετατόπισης," *Αισθητήρες Μέτρησης και Ελέγχου*, Θεσσαλονίκη, Εκδόσεις Τζιόλα, 2015, pp. 89-90.
- [19] R. Smoot, "What's the Difference Between an Incremental Encoder's PPR, CPR, and LPR?," CUI Devices, [Electronic Source]. Available: <https://www.cuidevices.com/blog/what-is-encoder-ppr-cpr-and-lpr>. [Access 30 August 2021].
- [20] K. Kalovrentis and N. Katevas, "Μετρήσεις ταχύτητας και επιτάχυνσης," *Αισθητήρες Μέτρησης και Ελέγχου*, Θεσσαλονίκη, Εκδόσεις Τζιόλα, 2015, pp. 270-271.
- [21] G. Dudek and M. Jenkin, "Inertial Sensors, GPS, and Odometry," *Springer Handbook of Robotics*, Berlin, Heidelberg, Springer, 2008, pp. 477-490. DOI: [https://doi.org/10.1007/978-3-540-30301-5\\_21](https://doi.org/10.1007/978-3-540-30301-5_21).
- [22] CH-Robotics, "UM6 Ultra-Miniature Orientation Sensor Datasheet," [Electronic Source]. Available: <http://www.robotcombat.com/products/images/0-PL1255.pdf>. [Access 07 September 2021].
- [23] J. L. Awange, "The global positioning system," *Environmental Science and Engineering (Subseries: Environmental Science)*, pp. 23-39, 2012.
- [24] C. KEE, W. B. W. PARKINSON and P. AXELRAD, "Wide Area Differential GPS," *NAVIGATION: Journal of The Institute of Navigation*, vol. 38, no. 2, pp. 123-144, 1991. DOI: <https://doi.org/10.1002/j.2161-4296.1991.tb01720.x>.
- [25] D. S. M. Valente, A. Momin, T. Grift and A. Hansen, "Accuracy and precision evaluation of two low-cost RTK global navigation satellite systems," *Computers and Electronics in Agriculture*, vol. 168, 2020. DOI: <https://doi.org/10.1016/j.compag.2019.105142>.
- [26] Emlid, "Reach RS2," Emlid, [Electronic Source]. Available: <https://emlid.com/reachrs2/>. [Access 07 September 2021].
- [27] U-Blox, "NEO/LEA-M8T series," U-Blox, [Electronic Source]. Available: <https://www.u-blox.com/en/product/neolea-m8t-series>. [Access 07 September 2021].



- [28] Intel, "Intel Edison Development Platform," Intel, 2015. [Electronic Source]. Available: [https://www.intel.com/content/dam/support/us/en/documents/edison/sb/edison\\_pb\\_331179002.pdf](https://www.intel.com/content/dam/support/us/en/documents/edison/sb/edison_pb_331179002.pdf). [Access 07 September 2021].
- [29] T. Takasu, N. Kubo and A. Yasuda, "Development, evaluation and application of rtklib:A program library for rtk-gps," In: *Proceedings of the International Symposium on GPS/GNSS*, 2007.
- [30] T. Raj, F. H. Hashim, A. B. Huddin, M. F. Ibrahim and A. Hussain, "A survey on LiDAR scanning mechanisms," *Electronics (Switzerland)*, vol. 9, no. 5, 2020. DOI: <https://doi.org/10.3390/electronics9050741>
- [31] J. R. Grove, J. Croke and C. Thompson, "Making a difference: examples of the use of repeat LiDAR datasets to guide river management decisions following extreme floods," In: *Proceedings of the 7th Australian Stream Management Conference*, pp. 232-239,2014 .
- [32] Sick, "LMS1xx Laser Measurement Sensors," 02 December 2019. [Electronic Source]. Available: [https://cdn.sick.com/media/docs/1/31/331/operating\\_instructions\\_lms1xx\\_laser\\_measurement\\_sensors\\_en\\_im0031331.pdf](https://cdn.sick.com/media/docs/1/31/331/operating_instructions_lms1xx_laser_measurement_sensors_en_im0031331.pdf). [Access 07 September 2021].
- [33] H. Kenji and S. Silvio, *CS231A Course Notes 1: Camera Models*, Stanford University , 2017.
- [34] Kenji and S. S. Hata, *CS231A Course Notes 3: Epipolar Geometry*, Stanford, California: University of Stanford, 2017.
- [35] R. Jain, R. Jaina, R. Kasturi, B. Schunck and B. Schunck, "Chapter 11: Depth," In: *Machine Vision*, McGraw-Hill, 1995, pp. 289-291.
- [36] Intel®, "Intel® RealSense™ Technology," Intel®, [Electronic Source]. Available: <https://www.intel.com/content/www/us/en/architecture-and-technology/realsense-overview.html>. [Access 29 August 2021].
- [37] Intel ® , *Intel® RealSense™*, Intel ® , June 2020.
- [38] C. Stachniss, *Robotic Mapping and Exploration*, Springer-Verlag Berlin Heidelberg, 2009. DOI: <https://doi.org/10.1007/978-3-642-01097-2>
- [39] K. Yousif, A. Bab-Hadiashar and R. Hoseinnezhad, "An Overview to Visual Odometry and Visual SLAM: Applications to Mobile Robotics," *Intelligent Industrial Systems*, vol. 1, no. 4, pp. 289-311, 2015. DOI: <https://doi.org/10.1007/s40903-015-0032-7>
- [40] D. Scaramuzza and F. Fraundorfe, "Visual Odometry Part I: The First 30 Years and Fundamentals," *IEEE ROBOTICS & AUTOMATION MAGAZINE*, vol. 18, no. 5, pp. 80-92, 2011.
- [41] F. Fraundorfer and D. Scaramuzza, "Visual Odometry: Part II - Matching , Robustness, and Applications," *IEEE ROBOTICS & AUTOMATION MAGAZINE*, vol. 19, no. 1, pp. 78-90, 2012.
- [42] W. Forstner, *A feature based correspondence algorithm for image matching*, Int. Arch. of Photogrammetry, 1986.

- [43] C. G. Harris and J. M. Pike, "3D positional integration from image sequences," *Image and Vision Computing*, vol. 6, no. 2, pp. 87-90, 1988. DOI: [https://doi.org/10.1016/0262-8856\(88\)90003-0](https://doi.org/10.1016/0262-8856(88)90003-0).
- [44] J. Shi and Tomasi, "Good features to track," In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Seattle, WA, USA, 1994. DOI: 10.1109/CVPR.1994.323794.
- [45] E. Rosten and T. Drummond, "Machine Learning for High-Speed Corner Detection," In: *European Conference on Computer Vision*, Graz, Austria, 2006. DOI: 10.1007/11744023\_34.
- [46] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," *International Journal of Computer Vision*, vol. 60, pp. 90-110, 2004. DOI: <https://doi.org/10.1023/B:VISI.0000029664.99615.94>
- [47] H. Bay, T. Tuytelaars and L. V. Gool, "Speeded-Up Robust Features (SURF)" *Computer Vision and Image Understanding*, vol. 110, no. 3, 2008. DOI: <https://doi.org/10.1016/j.cviu.2007.09.014>.
- [48] M. Agrawal, K. Konolige and M. Blas, "CenSurE: Center Surround Extremas for Realtime Feature Detection and Matching," In: *Computer Vision - ECCV 2008, 10th European Conference on Computer Vision*, Marseille France, pp. 102-105, 2008. DOI: 10.1007/978-3-540-88693-8\_8.
- [49] J. Sanchez, N. Monzon and A. Salgado, "An Analysis and Implementation of the Harris Corner Detector," *Image Processing on Line*, vol. 8, pp. 305-328, 2018. DOI: <https://doi.org/10.5201/ipol.2018.229>
- [50] U. Sinha, "The Shi-Tomasi Corner Detector," AI Shach, 2010. [Electronic Source]. Available: <https://aishack.in/tutorials/shitomasi-corner-detector/>. [Access 3 September 2021].
- [51] M. Calonder, V. Lepetit, C. Strecha and P. Fua, "BRIEF: Binary robust independent elementary features," In: *Proc. European Conf. Computer Vision*, pp. 778-792, 2010. DOI: 10.1007/978-3-642-15561-1\_56.
- [52] E. Rublee, V. Rabaud, K. Konolige and G. Bradski, "Orb: An efficient alternative to sift or surf," In: *Proc. IEEE Int. Conf. Computer Vision (ICCV)*, Barcelona, 2011. DOI: 10.1109/ICCV.2011.6126544.
- [53] S. Leutenegger, M. Chli and R. R. Siegwart, "Binary robust invariant scalable keypoints," In: *Proc. Int. Conf. Computer Vision*, 2011. DOI: 10.1109/ICCV.2011.6126542.
- [54] R. Szeliski, "Feature detection and matching," *Computer Vision: Algorithms and Applications*, Springer, 2010, pp. 225-230. DOI: <https://doi.org/10.1007/978-1-84882-935-0>.
- [55] K. Mikolajczyk and C. Schmid, "Scale & Affine Invariant Interest Point Detectors," *International Journal of Computer Vision*, vol. 1, no. 60, pp. 63-86, 2004. DOI: 10.1023/B:VISI.0000027790.02288.f2
- [56] D. D. Nister, O. Naroditsky and J. Bergen, "Visual odometry," In: *Proc. Int. Conf. Computer Vision and Pattern Recognition*, vol. 1, pp. I-I, 2004.

- [57] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981. DOI: <https://doi.org/10.1145/358669.358692>.
- [58] X.-S. Gao, X.-R. Hou, J. Tang and H.-F. Cheng, "Complete Solution Classification for the Perspective-Three-Point Problem," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 8, pp. 930- 943, 2003.
- [59] V. Lepetit, F. Moreno-Noguer and P. Fua, "EPnP: An accurate  $O(n)$  solution to the PnP problem," *International Journal of Computer Vision*, vol. 81, no. 2, 2009.
- [60] A. Penate-Sanchez, J. Andrade-Cetto and M.-N. Francesc, "Exhaustive Linearization for Robust Camera Pose and Focal Length Estimation.," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 10, pp. 2387–2400, 2013.
- [61] S. Thrun, D. Fox and W. Burgard, "Chapter 3: Gaussian Filters," *Probabilistic robotics*, pp. 33-36, MIT Press, 2005.
- [62] P. K. Panigrahi and S. K. Bisoy, "Localization strategies for autonomous mobile robots: A review," *Journal of King Saud University – Computer and Information Sciences*, 2021. DOI: <https://doi.org/10.1016/j.jksuci.2021.02.015>.
- [63] S. Thrun, W. Burgard and D. Fox, "The Extended Kalman Filter," *Probabilistic Robotics*, pp. 48-51, MIT Press, 2002.
- [64] M. Quigley, B. Gerkey, K. Conley, J. Faust, T. Foote, J. Leibs, E. Berger, R. Wheeler and A. Ng, "ROS: an open-source Robot Operating System," In: *ICRA - IEEE Robotics and Automation Society*, 2009.
- [65] ROS, "nav\_msgs/Odometry Message," ROS Documentations, 13 Jan 2021. [Electronic Source]. Available: [http://docs.ros.org/en/noetic/api/nav\\_msgs/html/msg/Odometry.html](http://docs.ros.org/en/noetic/api/nav_msgs/html/msg/Odometry.html). [Access 25 August 2021].
- [66] liam, "Extensible Markup Language (XML)," W3C, 11 10 2016. [Electronic Source]. Available: <https://www.w3.org/XML/>. [Access 1 September 2021].
- [67] R. Angeli, "urdf/XML/joint," ROS, 05 11 2018. [Electronic Source]. Available: <http://wiki.ros.org/urdf/XML/joint>. [Access 1 September 2021].
- [68] W. Meeussen, "REP: 105 Title: Coordinate Frames for Mobile Platforms," Robot Operating System, 27 October 2010. [Electronic Source]. Available: <https://github.com/ros-infrastructure/rep/blob/master/rep-0105.rst#id8>. [Access 1 September 2021].
- [69] Gazebo, "Beginner: GUI," Willow Garage, [Electronic Source]. Available: [http://gazebo.org/tutorials?tut=guided\\_b2](http://gazebo.org/tutorials?tut=guided_b2). [Access 1 September 2021].
- [70] T. Moore and D. Stouch, "A Generalized Extended Kalman Filter Implementation for the Robot Operating System," *Springer International Publishing*, Springer, 2016, pp. 335-348.

- [71] Moore and Tom, In: *ROSCon* , Hamburg, 2015.
- [72] T. Moore, "robot\_localization 2.6.11 documentation," 2016. [Electronic Source]. Available: [http://docs.ros.org/en/melodic/api/robot\\_localization/html/state\\_estimation\\_nodes.html](http://docs.ros.org/en/melodic/api/robot_localization/html/state_estimation_nodes.html). [Access 14 October 2021].
- [73] M. Labbe and F. Michaud, "RTAB-Map as an Open-Source Lidar and Visual SLAM Library for Large-Scale and Long-Term Online Operation," *Journal of Field Robotics*, vol. 36, no. 2, pp. 416-446, 2018. DOI: 10.1002/rob.21831.
- [74] G. Bradski and A. Kaehler, *Learning OpenCV: Computer vision with the OpenCV library*, O'Reilly Media, 2008.
- [75] SiCK, "Laser Measurement Sensors LMS1xx," 02 12 2019. [Electronic Source]. Available: [https://cdn.sick.com/media/docs/1/31/331/operating\\_instructions\\_lms1xx\\_laser\\_measurement\\_sensors\\_en\\_im0031331.pdf](https://cdn.sick.com/media/docs/1/31/331/operating_instructions_lms1xx_laser_measurement_sensors_en_im0031331.pdf). [Access 13 10 2021].
- [76] T. Baltovski, "husky\_simualtor," Clearpath Robotics, January 2014. [Electronic Source]. Available: <https://github.com/husky/husky>. [Access 25 August 2021].
- [77] Open Source Robotics Foundation, "Overview of Gazebo Plugins," Open Source Robotics Foundation, 2014. [Electronic Source]. Available: [http://gazebosim.org/tutorials?tut=plugins\\_hello\\_world&cat=write\\_plugin](http://gazebosim.org/tutorials?tut=plugins_hello_world&cat=write_plugin). [Access 27 August 2021].
- [78] J. L. Rivero, "gazebo\_ros\_pkgs," Open Robotics, 2016. [Electronic Source]. Available: [https://github.com/ros-simulation/gazebo\\_ros\\_pkgs/tree/noetic-devel/gazebo\\_plugins](https://github.com/ros-simulation/gazebo_ros_pkgs/tree/noetic-devel/gazebo_plugins). [Access 27 August 2021].
- [79] Technical University of Munich, "File Formats," [Electronic Source]. Available: [https://vision.in.tum.de/data/datasets/rgbd-dataset/file\\_formats](https://vision.in.tum.de/data/datasets/rgbd-dataset/file_formats). [Access 2 September 2021].
- [80] M. Grupp, "evo: Python package for the evaluation of odometry and SLAM.," 2017. [Electronic Source]. Available: <https://github.com/MichaelGrupp/evo>. [Access 02 September 2021].
- [81] Z. Zhang and D. Scaramuzza, "A Tutorial on Quantitative Trajectory Evaluation for Visual(-Inertial) Odometry," In: *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Madrid, Spain, 2018. DOI: 10.1109/IROS.2018.8593941.
- [82] S. Umeyama, "Least-squares estimation of transformation parameters between two point patterns," In: *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 4, pp. 376 - 380, 1991. DOI: 10.1109/34.88573.
- [83] W. Burgard, C. Stachniss, G. Grisetti, B. Steder, R. Kummerle, C. Dornhege, M. Ruhnke, A. Kleiner and J. D. Tardós, "A Comparison of SLAM Algorithms Based on a Graph of Relations," In: *IEEE/RSJ Int. Conf. Intell. Robot. Syst. (IROS)*, pp. 2089-2095, USA, 2009. DOI: 10.1109/IROS.2009.5354691.

- [84] A. Rosinol, M. Abate, Y. Chang and L. Carlone, "Kimera: an Open-Source Library for Real-Time Metric-Semantic Localization and Mapping," In: *IEEE Int. Conf. Robot. Autom. (ICRA)*, pp. 1689-1696, 2020.
- [85] C. Campos, R. Elvira, J. J. G. Rodríguez, J. M. M. Montiel and J. D. Tardós, «ORB-SLAM3: An Accurate Open-Source Library for Visual, Visual–Inertial, and Multimap SLAM,» *IEEE Transactions and Robotics*, pp. 1-17, 2021. DOI: 10.1109/TRO.2021.3075644.
- [86] F. Endres, J. Hess, J. Sturm, D. Cremers and W. Burgard, "3D Mapping with an RGB-D Camera," *IEEE TRANSACTIONS ON ROBOTICS*, vol. 30, no. 1, pp. 177-187, 2014. DOI: 10.1109/TRO.2013.2279412.