



LATVIJAS
UNIVERSITĀTE

Promocijas darba
kopsavilkums

Pēteris Račinskis

UZTVERE UN VADĪBA
AUTONOMAJĀ ROBOTIKĀ,
IZMANTOJOT
REDZES-VALODAS
IEGULTOS VEKTORUS

Rīga 2026



LATVIJAS
UNIVERSITĀTE

EKSAKTO ZINĀTŅU UN TEHNOLOĢIJU FAKULTĀTE

Pēteris Račinskis

**UZTVERE UN VADĪBA AUTONOMAJĀ
ROBOTIKĀ, IZMANTOJOT
REDZES-VALODAS IEGULTOS
VEKTORUS**

PROMOCIJAS DARBA KOPSAVILKUMS

zinātniskā doktora grāda (*Ph. D.*) iegūšanai
datorzinātnes un informātikas nozarē
Apakšnozare: intelektuālo sistēmu teorija

Rīga 2026

Promocijas darbs izstrādāts robotikas un mašīnuztveres laboratorijā, Elektronikas un datorzinātņu institūtā (EDI), no 2023. gada līdz 2025. gadam.

Darbs sastāv no ievada, 4 nodaļām, kopsavilkuma, galvenajiem rezultātiem un secinājumiem, literatūras saraksta.

Darba forma: disertācija.

Darba zinātniskais vadītājs: *Dr. sc. comp.* **Modris Greitāns**, vadošais pētnieks, Elektronikas un datorzinātņu institūts, Latvija.

Darba recenzenti:

1. profesore, *Dr. sc. comp.* **Inguna Skadiņa**, Latvijas Universitāte, Latvija;
2. profesors, *Ph. D.* **Vytautas Bučinskas**, Viļņas Ģedimina tehniskā universitāte, Lietuva;
3. profesors, *Ph. D.* **Po Ling Tin**, Taivānas Nacionālā zinātnes un tehnoloģiju universitāte, Taivāna.

Promocijas darba aizstāvēšana notiks Latvijas Universitātes Datorzinātnes un informātikas un Elektrotehnikas, elektronikas un komunikāciju tehnoloģiju nozares promocijas padomes atklātajā sēdē 2026. gada 23. janvārī.

Ar promocijas darbu un tā kopsavilkumu var iepazīties Latvijas Universitātes Bibliotēkā Rīgā, Kalpaka bulvārī 4.

LU Datorzinātnes un informātikas un
Elektrotehnikas, elektronikas un komunikāciju
tehnoloģiju nozares

promocijas padomes priekšsēdētājs:

Dr. sc. comp. Guntis Bārzdīņš

promocijas padomes sekretāre:

Sintija Siliņa

© Latvijas Universitāte, 2026

© Pēteris Račinskis, 2026

ISBN 978-9934-36-479-2

ISBN 978-9934-36-480-8 (PDF)

ANOTĀCIJA

Šajā disertācijā aprakstīti pētnieciskie rezultāti uztveres un kontroles jomā autonomajā robotikā, integrējot redzes-valodas semantiku, lai risinātu problēmas nestrukturētā vidē. Pētījums koncentrējas uz semantisko vienlaicīgās lokalizācijas un kartēšanas (SLAM) sistēmu izstrādi. Galvenie ieguldījumi ietver vizuālās atvērtās kopas semantikas pielietošanu liela mēroga LiDAR kartēs apvidus segmentēšanai autonomu robotu navigācijā un valodas modeļu-plānotāju integrāciju ar šādu semantisko karti galda objektu kraušanas scenārijā. Pētījuma mērķis ir apvienot dažādus sensoru datus – LiDAR, dziļuma kameru, krāsu attēlus un GNSS – vienotā redzes-valodas iegulto vektoru reprezentācijā, veicinot cilvēka un robota saskarni, izmantojot balsi vai tekstu. Darba ietvaros arī izvirzīta un eksperimentāli validēta sensora atskaites pozas metode, kas izmanto vizuālos marķierus, lai nodrošinātu neatkarību no GNSS signāla un augstāku precizitāti vidēs ar signāla degradāciju. Šī darba rezultāti ir publicēti trīs galvenā autora rakstos indeksētos žurnālos, divās starptautiskās konferencēs un tehnoloģijas dokumentācijas pakotnē projekta gala nodevuma ietvaros. Šis darbs iepazīstina ar diviem publiski pieejamiem resursiem – EDI-SLAM datu kopu un atvērtā koda programmatūras pakotni SLAMVDB. Divus disertācijas ietvaros izstrādātus tehnoloģiju demonstratorus — semantiska uztveres sistēma un statiskai manipulācijai paredzēta robota kontroles sistēma — pozitīvi novērtējuši industrijas pārstāvji.

SATURS

Apzīmējumi	5
Ievads	6
1 Teorētiskais pamatojums	12
1.1 Uztveres un kartēšanas sistēmas	12
1.2 Augsta līmeņa plānošana un vadība	13
1.3 Izaicinājumi, mērķi un ieguldījums nozares attīstībā	14
2 Datu kopa un atskaites pozas	15
2.1 Sensoru komplekts	15
2.2 Atskaites pozu mērīšana	16
2.3 Datu kopa	17
2.4 Rezultāti	19
3 Stacionārais objektu kraušanas demonstrators	21
3.1 Vairāku līmeņu vadības sistēma	21
3.2 Eksperimentālie novērtējumi	23
3.3 Rezultāti	24
4 Semantiskās uztveres sistēmas	26
4.1 Izstrādātās versijas	26
4.2 Eksperimentālie novērtējumi	29
4.3 Rezultāti	29
Kopsavilkums un secinājumi	31

APZĪMĒJUMI

EDI - Elektronikas un datorzinātņu institūts, Latvija

SLAM - Vienlaicīgā lokalizācija un kartēšana

LiDAR - Lāzera attāluma skeneris

VLM - Redzes-valodas (ieguļšanas) modelis

LLM - Lielais valodu modelis

NLP - Dabiskās valodas apstrāde

GNSS - Globālā navigācijas satelītu sistēma

ECEF - Globālā Dekarta koordinātu sistēma

ROS, ROS2 - *Robot Operating System, Robot Operating System 2*

RGB - Krāsu attēls, tiek lietots gan sensora tipam, gan datu modalitātei

RGB-D - Dziļuma-krāsu attēls, tiek lietots gan sensora tipam, gan datu modalitātei

RTK - Reāllaika kinematiskā GNSS

ATE - Absolūtā trajektorijas kļūda

RPE - Relatīvā pozas kļūda

ICP - Iteratīvais tuvākā punkta algoritms

HLP - Augsta līmeņa plānošanas apakšsistēma

LLP - Zema līmeņa plānošanas apakšsistēma

SotA, BSotA - *state of the art, beyond state of the art*

PnP - Perspektīvas-n-punktu problēma vai tās risināšanas algoritms

IEVADS

Tādās pamatnozārēs kā ražošana, lauksaimniecība un mežsaimniecība demogrāfisko pārmaiņu, mainīgo darba prasību un ekonomiskā spiediena dēļ rodas bažas par draudošu darbaspēka trūkumu un prasmju krīzes saasināšanos. Nesen veikti pētījumi izceļ šos izaicinājumus, norādot, ka uzņēmumiem visā pasaulē ir grūtības atrast un noturēt darbiniekus, īpaši tos, kuriem ir specializētas prasmes, kas nepieciešamas mūsdienīgā darba vietā [1, 2]. Robotika jau sen tiek uzskatīta par potenciālu risinājumu darbaspēka iztrūkumiem tirgū. Tipveida uzdevumos fiksētās darba stacijās, ar kādiem saskaras liela mēroga ražošanas uzņēmumi, industriālie manipulatori ir sevi apliecinājuši. Tomēr, lai gan šādas robotrokas ir plaši izplatītas, lielākoties tās joprojām tiek programmētas manuāli un izpilda iepriekš ieprogrammētas trajektorijas [3], kas ierobežo to pielāgojamību un līdz ar to arī potenciālos pielietojumus. Ražošanas līnijas joprojām uzstāda augsti kvalificēti tehniķi, bet starp robotu sistēmu un ražošanas līnijas darbinieku pastāv minimāla atgriezeniskā saite. Ārpus lielsēriju ražošanas, lielākie panākumi autonomas robotikas pielietojumā ir gūti salīdzinoši strukturētās vidēs, piemēram, ielas apstākļos, ar kādiem saskaras pabraucošie automobiļi [4]. Skatot pielietojuma jomas sarežģītākās un mazāk paredzamās vidēs, uztveres un kontroles tehnoloģiju gatavības līmeņi ir ievērojami zemāki daļēji strukturētās (piemēram, lauksaimniecības [5]) un gandrīz pilnīgi nestrukturētās (piemēram, mežsaimniecības [6]) vidēs. Šo likumsakarību var izskaidrot ar diviem galvenajiem izaicinājumiem:

1. Nestrukturētās vidēs nevar paļauties uz konstruktīviem pieņēmumiem vai augsta līmeņa abstrakcijām apkārtnes modelēšanā uztveres procesiem;
2. Plašāks, mazāk atkārtots uzdevumu klāsts sarežģī robota vadības sistēmas izveidi un programmēšanu.

Abu šo izaicinājumu risināšana ir nepieciešama, lai veicinātu plašāku robotu izmantošanu tirgus nozarēs, kas nav lielsēriju rūpnieciskā ražošana. Potenciālie pielietojumi ir ļoti plaši — sākot ar industriālo manipulatoru ieviešanu mazu sēriju ražošanā vai pakalpojumu nozarē un beidzot ar autonomu mobilo robotu izmantošanu sarežģītās, nestrukturētās ārtelpu vidēs, piemēram, lauksaimniecībā, mežsaimniecībā, infrastruktūras uzturēšanā un aizsardzībā.

Darba zinātniskā novitāte

Šī disertācija aptver pētniecības un inženierijas darbu kopumu, kas galvenokārt vērsts uz semantisko SLAM sistēmu izstrādi, to izmantošanu industriālo robotu vadībā un šiem mērķiem paredzētas aparatūras, programmatūras un metodoloģijas izstrādi — no datu vākšanas līdz atvērtā pirmkoda programmatūras publiskai izlaišanai. Darba galvenie zinātniskie jaunpienesumi ir:

- Vizuālās atvērtās kopas semantikas izmantošana liela mēroga LiDAR kartē apvidus segmentācijai, kas paredzēta uz virszemes autonomu robotu navigācijai;
- Lielajos valodas modeļos balstītas augsta līmeņa plānošanas sistēmas un atvērtās kopas semantiskās kartes integrācija;
- Metodoloģija SLAM sistēmu testēšanai un novērtēšanai neatkarīgi no GNSS, izmantojot vizuālus marķierus un mērmiecības aprīkojumu.

Darba mērķis un uzdevumi

Darba **mērķis** un galvenā motivācija ir sekmēt robotu autonomiju un interaktīvu sadarbību ar cilvēkiem caur dažādu sensoru mērījumu — LiDAR vai dziļuma kameras punktu mākoņu, krāsu attēlu, GNSS — apvienošanu vienotā, valodas jēdzienos balstītā apkārtējās valodas modelī. Šī darba rezultātiem būtu jāpalīdz nākotnes autonomo robotu vadības sistēmām labāk orientēties sarežģītās, nestrukturētās vidēs (piemēram, mežos), kā arī jāveicina intuitīvu saskarni starp robotu un tā operatoru.

Darba konkrētie pētnieciskie **uzdevumi** ir:

1. Izstrādāt atvērtās kopas semantiskās kartēšanas sistēmu darbam garās trajektorijās caur nestrukturētām vidēm;
2. Demonstrēt kartēšanas sistēmas spēju veikt dažādas uzmeklēšanas operācijas, kas ir būtiskas industriālās manipulācijas un āra visurgājēju kontekstā — objektu satveršanas pozu noteikšanu, apvidus segmentāciju;
3. Izveidot rīkus un metodes, kas nepieciešamas, lai testētu, verificētu un validētu sistēmas veiktspēju atbilstošos apstākļos.

Aizstāvēšanai izvirzītās tēzes

Disertācijas ietvaros veiktās pētnieciskās darbības motivācijai un pamatošanai tika izvirzītas šādas trīs tēzes, kas pētījumu gaitā tika eksperimentāli pārbaudītas:

- **1. Vizuālie marķieri SLAM datu kopu anotācijā** — vizuāli marķieri ar stūru pozīcijām, kas uzmērītas izmantojot mērnietu aprīkojumu, kalpo kā no GNSS neatkarīga atskaites pozas noteikšanas metode, kura turklāt ir uzticamāka par RTK koriģētu daudzjoslu GNSS vidēs ar signāla degradāciju.
- **2. Atvērtās kopas semantiskās kartes industriālo robotu vadībai manipulācijās** — atvērtās kopas semantiskās vokseļu un dziļuma kartes saņemšana objektu pārklājuma rādītājus virs 90% galda virsmas vidē, un šāda uztveres metode, integrēta divlīmeņu dabiskās valodas komandu izpildes sistēmā, nodrošina veiksmīgu vienkāršu kraušanas instrukciju izpildi;

- **3. Atvērtās kopas semantiskās kartes liela mēroga āra apvidus segmentācijai** — atvērtās kopas vokseļu karte, kas satur redzes-valodas modeļa iegultos vektorus, sasniedz vokseļu klasifikācijas precizitāti apvidus segmentācijas uzdevumā līdz 87%, ja tiek ieviesta heuristika tukšo novērojumu atmešanai.

Īss metožu raksturojums

Lai motivētu un vadītu turpmāko pētniecību, autors ir veicis literatūras apskatu autonomās robotikas uztveres, vides modelēšanas, lokalizācijas un karšu veidošanas jomā [7]. Balstoties uz tajā identificētajām galvenajām atziņām un nozares attīstības trūkumiem, autors izstrādāja rokās pārnēsājamu sensoru komplektu, lai vāktu sinhronizētus LiDAR, IMU, GNSS un attēlu datus, kā arī nepieciešamo programmatūru tā darbināšanai un rīkus saistīto datu apstrādei. Būtisks jaunpienesums ir vizuālu marķieru izmantošana atskaites pozu uzmērīšanā, kas aprakstīta [8], un šis darbs noslēdzās ar *EDI-SLAM* datu kopas [9] izveidi. Lai validētu atvērtās kopas semantiskās kartēšanas piemērotību robotu kontroles uzdevumiem, tika izstrādāts iekštelpu demonstrators objektu satveršanas un kraušanas uzdevumu veikšanai ar industriālo manipulatoru [10]. Tika izstrādāti rīki galda virsmas ainu manuālai anotācijai un daļēji automātiskai patieso darbību plānu izveidei, lai pārbaudītu instrukciju izpratni un kvantitatīvi novērtētu sistēmas veiktspēju. Visbeidzot, autors izstrādāja otru demonstratoru — atvērtā pirmkoda programmatūras pakotnes — lai veiktu atvērtās kopas semantisko kartēšanu āra vidēs un veidotu liela mēroga vokseļu režģa kartes ar saskarnēm apvidus segmentācijai un objektu meklēšanai. Tā publicēta ar nosaukumu *SLAMVDB* [11]. Tā kā liela daļa šī darba tika veikta Latvijas valsts pētījumu programmas VPP-EM-FOTONIKA-2022/1-0001 ietvaros, autors arī bija galvenais autors tehniskās dokumentācijas pakotnei, kas publiski pieejama [12].

Promocijas darba rezultātu aprobācija

Ar šo promocijas darbu saistītās publikācijas un konferenču tēzes:

- (I) Racinskis, Peteris, Janis Arents, and Modris Greitans. "Constructing maps for autonomous robotics: An introductory conceptual overview." *Electronics* 12, no. 13 (2023): 2925.
- (II) Racinskis, P., Arents, J., Greitans, M. Towards a Multi-modal, Multi-layer Mapping Framework for Autonomous Robotics—an Outline, International Workshop on Embedded Digital Intelligence (IWoEDI'2023), Riga, Latvia.
- (III) Racinskis, Peteris, Janis Arents, and Modris Greitans. "Annotating SLAM data sets with Apriltag markers." In 2024 10th International Conference on Automation, Robotics and Applications (ICARA), pp. 438-442. IEEE, 2024.

- (IV) Racinskis, Peteris, Oskars Višmanis, Toms Eduards Zinars, Janis Arents, and Modris Greitans. "Towards Open-Set NLP-Based Multi-Level Planning for Robotic Tasks." *Applied Sciences* 14, no. 22 (2024): 10717.
- (V) Racinskis, Peteris, Gustavs Krasnikovs, Janis Arents, and Modris Greitans. "The EDI Multi-Modal Simultaneous Localization and Mapping Dataset (EDI-SLAM)." *Data* 10, no. 1 (2025): 5.

Šajā disertācijā detalizētie rezultāti tikuši prezentēti šādās starptautiskās zinātniskās konferencēs un pētniecības un industrijas forumos:

- (a) Divu lappušu tēzes: 2023 International Workshop on Embedded Digital Intelligence (IWoEDI'2023). Rīga, Latvija, 2023;
- (b) Konferences raksts: 2024 IEEE 10th International Conference on Automation, Robotics and Applications (ICARA). Atēnas, Grieķija, 2024;
- (c) Stenda referāts: 2024 IEEE 14th International Conference Nanomaterials: Applications & Properties (NAP). Rīga, Latvija, 2024;
- (d) Stenda referāts un fizisks demonstrators: *5g Techritory*. Rīga, Latvija, 2024;
- (e) Vienas lappuses tēzes, stenda referāts, fizisks demonstrators: European Robotics Forum 2025. Štutgarte, Vācija, 2025.

Pētījumi šajā disertācijā galvenokārt tika finansēti divos projektos:

- Latvijas valsts pētījumu programma Nr. VPP-EM-FOTONIKA-2022/1-0001 "Viedo materiālu, fotonikas, tehnoloģiju un inženierijas ekosistēma"
- EdgeAI "Edge AI Technologies for Optimised Performance Embedded Processing" projekts, kas saņēmis finansējumu no CHIPS JU saskaņā ar granta līgumu Nr. 101097300.

Autora ieguldījums

Visās publikācijās autors bija galvenais autors un sagatavoja vai nu visu pamattekstu, vai lielāko tā daļu. (I), (II), (III) autors veica visu pētniecības un izstrādes darbību. Publikācijā (IV) autors piedāvāja ideju, izstrādāja sistēmas arhitektūru, pārraudzīja demonstratora izstrādi, integrāciju un eksperimentālo novērtēšanu, kā arī tieši implementēja šādas apakšsistēmas: atvērtais kopas semantisko karti, attēlu datu starpprogrammatūru, eksperimentālo datu vākšanu, datu kopu anotācijas rīkus, novērtēšanas statistikas aprēķināšanu. Kolēģis Oskars Višmanis izstrādāja zema līmeņa robotu darbību primitīvu apakšsistēmu. Kolēģis T.E. Zinars izstrādāja augsta līmeņa plānošanas apakšsistēmu. Abi kolēģi palīdzēja

manuskripta sagatavošanā, uzrakstot sākotnējās melnraksta versijas sadaļām, kas atbilst viņu apakšsistēmām. Viņu izstrādāto komponentu apraksti ir izmantoti arī piekrišanu, kur tas ir būtiski. (V) autors veica visu pētniecības un izstrādes darbību, taču kolēģis Gustavs Krasņikovs ierakstīja lielāko daļu sensoru trajektoriju, kas tika iekļautas datu kopas publiskajā versijā, kā arī veica sensoru komplekta kalibrēšanas, uzstādīšanas un uzturēšanas darbus. Visās publikācijās vadošais pētnieks, EDI robotikas grupas vadītājs Ph.D. Jānis Ārents un vadošais pētnieks, EDI direktors, promocijas darba vadītājs Dr.sc.comp. Modris Greitāns palīdzēja padomdevēja lomā un projektu administrācijā.

Promocijas darba apjoms un struktūra

Ņemot vērā to, ka autors jau ir publicējis ievērojamu apjomu materiāla par šajā disertācijā aprakstīto pētniecisko darbību, liela daļa promocijas darba satūra ir sniegta saīsināta kopsavilkuma veidā, papildinot iepriekš rakstīto ar dažām jaunākām atsaucēm, rezultātiem un atziņām. Šī disertācija satur 30 attēlus, 10 tabulas un 2 algoritmu aprakstus. Tā sastāv no 4 nodaļām:

- **1. nodaļa — Teorētiskais pamatojums** ir apskats par galvenajiem jēdzieniem, pētījumiem un pašreizējo stāvokli nozarēs, kas aplūkotas šajā disertācijā. Tā lielā mērā ir apvienots kopsavilkums no rakstu (I), (IV), (V) un projekta ziņojuma [12] teorētiskajām sadaļām, ar dažiem jaunākiem rezultātiem, kas iekļauti, kur tas ir būtiski, lai atjauninātu informāciju atbilstoši jaunākajām nozares attīstības tendencēm;
- **2. nodaļa — Sensoru komplekts un datu kopa** apraksta sensoru komplektu, kas izstrādāts datu vākšanai ārpus telpām; jauno atskaites pozas mērīšanas metodi, tās precizitātes eksperimentālos novērtējumus, salīdzinājumu ar sensoru komplekta GNSS-INS trajektoriju; rezultātā iegūto *EDI-SLAM* datu kopu;
- **3. nodaļa — Galda virsmas kraušanas demonstrators** detalizēti aplūko robotu vadības sistēmas arhitektūru, kas izstrādāta, lai izmantotu atvērtās kopas semantiku statistiskā instalācijā; lielā valodas modeļa instrukciju izpratnes sistēmas eksperimentālu novērtējumu; kā arī rezultātus pilna cikla instrukciju izpildē uz reāla robota;
- **4. nodaļa — Semantiskās uztveres sistēmas** detalizēti apraksta šī pētījuma ietvaros izstrādātās atvērtās kopas semantiskās kartēšanas sistēmas; to dizaina aspektus un katrā iterācijā ieviesto papildu funkcionalitāti; verificācijas metodes un rezultātus trīs aspektos — objektu atpazīšanā, lokalizācijas precizitātē un apvidus segmentācijā.

Galveno rezultātu kopsavilkums

Tika izgatavota pārnēsājams sensoru komplekts ar kuru tika ierakstīta apjomīga publiski pieejama datu kopa — *EDI-SLAM*. Tā satur vairākus kilometrus

trajektoriju ierakstu no dažādiem (rokās nestiem un uz transportlīdzekļa fiksētiem) rakursiem. Procesā tika pielietota jaunieviestā patiesās pozas noteikšanas metode, un šīs metodes precizitāte tika eksperimentāli novērtēta — ar maksimālo pozīcijas kļūdu zem 5 cm un orientācijas kļūdu zem 0.5° , izmantojot 4 marķierus katrā pozicionēšanas stacijā. Līdztekus šiem darbiem tika izstrādāta vairāku līmeņu robota kontroles sistēma, kas izmanto atvērtās kopas semantiku objektu satveršanas pozu noteikšanai. Šī sistēma veiksmīgi rekonstruē virs 90% vienkāršu objektu kraušanas instrukciju, un izpilda divas trešdaļas no tām bez kļūdām ar industriālu robotrokas. Tika izstrādāta arī virkne atvērtās kopas semantiskās kartēšanas sistēmu, kas sekmīgi paveic objektu satveršanas pozas noteikšanu (līdz 94.69% objektu atrašana, ar oktokoka karti un ternāri kvantizētiem semantikas vektoriem) kā arī zemes virsmas segmentāciju (līdz 87.35% vokseļu klasifikācijas precizitāte uz *RELLIS-3D* trajektorijas ar indeksu *00000*).

1. TEORĒTISKAIS PAMATOJUMS

Lai autonomais robots varētu veikt jēgpilnus uzdevumus, tam ir jāspēj uzvert apkārtējo vidi, novērtēt savu pozīciju un orientāciju, plānot un izpildīt darbības. Lielākajā daļā pašlaik pastāvošo robotu kontroles ietvaru zema līmeņa kontrolieri darbina atsevišķas locītavas vai riteņus, un plānotāji ģenerē robota konfigurāciju trajektorijas, kas sasniedz mērķi telpā [13, 14]. Taču šajā abstrakcijas līmenī joprojām ir nepieciešami precīzi uzdoti galamērķi robota konfigurāciju vai Dekarta telpā. Lai šādus kustības mērķus ģenerētu dinamiski, būtisks posms starp plānošanu un darbību ir robota uztveres sistēma. Tā nodrošina vides modeli, kas ļauj robotam veikt patstāvīgu navigāciju un izvairīties no šķēršļiem [15]. Savukārt sarežģītākai uzvedībai un dabiskās valodas saskarnēm nepieciešama kontrole augstākos abstrakcijas līmeņos.

1.1. Uztveres un kartēšanas sistēmas

Ja robots darbojas bez iepriekšēja vides modeļa, tam ir jāatrisina vienlaicīgās lokalizācijas un kartēšanas (SLAM) problēma — vienlaikus veidojot apkārtnes karti un sekojot līdzi savai pozai (pozīcijai un orientācijai) šajā kartē [16, 17]. Formāli tā ir Beiesa secināšanas problēma, kuras mērķis ir novērtēt robota trajektoriju un kartes struktūru, pamatojoties uz sensoru mērījumu laikrindu [16]. SLAM problēmas tipiski risina vienā no diviem veidiem. Varbūtiskās filtrēšanas metodes, piemēram, paplašinātais Kalmana filtrs (EKF), novērtē tikai nesenāko robota stāvokli, pakāpeniski integrējot pagātnes informāciju, lai saglabātu fiksētu algoritmisko sarežģītību [18—20]. Lai gan šī pieeja ir efektīva odometrijas uzdevumā — relatīvā pārvietojuma noteikšanai reālā laikā — filtrēšanas pieejas cieš no dreifa garās trajektorijās [16]. Savukārt izlīdzinošās pieejas saglabā visu robotu pozu un sensoru mērījumu vēsturi, periodiski atrisinot optimizācijas uzdevumu, lai atrastu pašreizējā stāvokļa novērtējumu ar visaugstāko varbūtību [17]. Šādas metodes pieļauj ciklu slēgšanu — iepriekš apmeklētu vietu atpazīšanu un “cilpu” atsaīšu pievienošanu optimizācijas uzdevumā [21—23]. Izlīdzināšanas SLAM problēmas bieži tiek formulētas, izmantojot faktoru grafus ar divu veidu virsotnēm — robotu pozas un kartes orientieri ir mainīgie, bet sensoru mērījumu atsaītes ir faktori [24]. Šāds grafs definē nelineāru mazāko kvadrātu optimizācijas problēmu [17].

SLAM sistēmas rezultāts parasti ir tīri telpiska karte, piemēram, punktu mākonis vai režģakarte [25], kas iekodē vides ģeometriju. Lai gan telpiskas kartes ir nepieciešamas robota lokalizācijai un navigācijai, tās neietver pietiekamu informāciju sarežģītāku plānošanas uzdevumu risināšanai [26]. Tāpēc jau izsenis veikti pētījumi vērsti uz karšu papildināšanu ar semantiskiem slāņiem, kas atspoguļo objektu un telpas apgabalu nozīmi [27, 28]. Semantiskā informācija šādām kartēm pārsvarā tiek iegūta ar datorredzes modeļiem, kas apstrādā kameru attē-

lus. Tādi uzdevumi kā objektu atpazīšana un pikseļu **semantiskā segmentācija** sašķiro attēla daļas diskrētās kategorijās (piemēram, "krēsls", "galds") [29]. Nesens attīstības virziens šajā nozarē ir **atvērtās kopas semantika**, kas atbrīvojas no nepieciešamības iepriekš definēt diskrētu kategoriju kopu. Izmantojot redzesvalodas iegulšanas modeļu (VLM), piemēram, CLIP [30], vektoru telpas, šīs sistēmas var identificēt un meklēt iepriekš neredzētus objektus vai jēdzienus, kas aprakstīti dabiskā valodā, bez nepieciešamības veikt atkārtotu apmācību [31, 32]. Šāda spēja ir nepieciešama, lai radītu elastīgas un cilvēkam intuitīvi saprotamas robotu sistēmas.

Īpaši sarežģīts semantiskās kartēšanas pielietojums ir **apvidus segmentācija** navigācijai nestrukturētās ārtelpu vidēs. Atšķirībā no tipiskas objektu atpazīšanas, šis uzdevums prasa atšķirt izbraucamas un neizbraucamas virsmas (piemēram, "zāle", "grants", "dublī"). Pastāvošās datu kopas šai jomai ir ievērojami mazākas nekā tās, kas pieejamas plaši izplatītiem objektiem [33, 34], un modeļu arhitektūras, kas izcili veic objektu segmentāciju, ne vienmēr ir labi vispārināmas apvidus analīzei [35]. Šajā disertācijā tiek izvirzīta hipotēze, ka, sapludinot lokālus atvērtās kopas semantiskos novērojumus no daudziem attēliem vienotā 3-dimensionālā kartē, vispārīgākam pielietojumam apmācītu redzes modeli var efektīvi pielāgot šim šauri specializētajam uzdevumam.

1.2. Augsta līmeņa plānošana un vadība

Robotikas kontekstā uztveres sistēmas galvenais mērķis ir sniegt informāciju darbību realizēšanai. Nesenos pētījumos iezīmējusies tendence izmantot **Lielos valodu modeļus (LLM)** uzdevumu plānošanai augstā līmenī [36, 37]. Šādas sistēmas spēj izprast dabiskā valodā uzdotas komandas un ģenerēt soļu secību, ko robots var izpildīt. Dažas pieejas izmanto LLM, lai vienā piegājienā ģenerētu veselu plānu, kas pēc tam tiek validēts un izpildīts [38, 39], savukārt citas iteratīvi izsauc modeli, lai tas izvēlētos vispiemērotāko no pieejamajām darbībām pašreizējā kontekstā [40].

Būtisks daudzu pašreizējo valodas modeļos balstīto plānotāju ierobežojums ir tas, ka to "pasaules modelis" bieži ir sastāv tikai no pašreizējiem sensoru datiem vai īsas novērojumu vēstures [40, 41]. Tās lielākoties neizmanto strukturētas un pastāvīgas apkārtējās vides reprezentācijas, kādas var nodrošināt SLAM sistēma. Tas rada plaisu: no vienas puses, pastāv sarežģītas SLAM sistēmas, kas spēj izveidot detalizētas semantiskās kartes; no otras puses, pastāv intelektuālas augsta līmeņa sistēmas, kas nevar šādas kartes efektīvi izmantot. Šīs plaisas pārvarēšana, izstrādājot liela mēroga, pārmeklējamu semantisko karti, kas piemērota saskarnei ar mūsdienu plānošanas sistēmām, ir viens no galvenajiem šajā disertācijā prezentētā darba motivatoriem.

1.3. Izaicinājumi, mērķi un ieguldījums nozares attīstībā

Lai motivētu šajā promocijas darbā aprakstīto pētniecisko darbu, ir svarīgi apkopot galvenās atziņas un identificēt dažas svarīgas problēmas, kas šobrīd nav atrisinātas, kā arī funkcionālos un veikspējas aspektus, kuros pašreizējais nozares stāvoklis (SotA) ir nepilnīgs. Praktiski lauka izmēģinājumi, piemēram, [42], kurā tika novērtētas vairākas sistēmas [18, 19, 21, 22, 43], liecina, ka liela mēroga robusta ārtelpu vides kartēšana joprojām ir problēma. Tāpēc ir nepieciešams nodrošināt, ka jebkuras pētījuma gaitā radītās uztveres sistēmas pamatā tiek izvietota robusta, uzticama un modulāra lokalizācijas sistēma. Lai to nodrošinātu, ir ļoti svarīgi spēt ievākt jaunus sensoru datus, kas adekvāti atspoguļotu paredzamos ekspluatācijas apstākļus, un izvairītos no pārāk ciešas pielāgošanās ierobežotam gatavo datu kopu klāstam. Ir jāapkopo un jāpārbauda retāk pārstāvēti vides veidi — daļēji vai pilnībā nestrukturēti — un garas trajektorijas (simtiem metru, kilometri). Pārvietojoties starp atklātām vietām ar skaidriem, netraucētiem GNSS signāliem un iekštelpām vai ļoti apbūvētām teritorijām, ir jānodrošina arī no GNSS signāla neatkarīga patiesās pozas anotācijas metode. Ar šādu, no GNSS neatkarīgu atsauces pozas mērījumu ir iespējams objektīvi novērtēt tādu lokalizācijas sistēmas veikspēju, kas tieši iepļudina GNSS datus savos aprēķinos. Ņemot vērā, kā SotA pieejas LLM vai multimodālo modeļu plānošanai apstrādā vai nu tikai pašlaik redzamo ainu [40], vai nesēn redzētu attēlu secību [41], bet neizmanto SLAM tehnoloģijas, ir jāizpēta integrācija starp valodas modeļos balstītu augsta līmeņa plānošanu un strukturētiem semantiskiem vides modeļiem. Visbeidzot, apvidus segmentācijā, kur iepriekšējās pieejas galvenokārt šauri specializējas uz uzlabotiem rezultātiem tikai konkrētajā uzdevumā attēlos [35] vai punktu mākoņos [44], šī disertācija pēta iespēju pielietot universālu atvērtās kopas attēlu segmentācijas modeli un 3-dimensionālu karti, lai atrastu izbraucamus virsmas ceļus robotu navigācijai. Ņemot vērā visu iepriekš minēto, bija nepieciešama vairāku tehnoloģiju izstrādes un eksperimentu programmu īstenošana:

- **Datu nodrošināšana** — lai nodrošinātu, ka ir pieejami atbilstoši testa, verifikācijas un validācijas dati, jāizstrādā metodoloģija un instrumenti datu kopu ievākšanai, patiesās pozas marķēšanai un izmantošanai skaitliskos novērtējumos (1. tēze).
- **Vadības integrācijas** — lai novērtētu uztveres sistēmas pielietojamību reālos robotikas vadības uzdevumos (2. tēze), jāizstrādā robota vadības sistēmas arhitektūra, kas izmanto šādu karti; jāizveido nepieciešamie moduļi augsta un zema līmeņa plānošanai un vadībai, jāsteno saskarnes starp tiem; jāvalidē arhitektūra reāla robota kontrolē;
- **Uztveres sistēma** — atvērtās kopas semantiskās kartēšanas sistēmas izstrāde, kas paredzēta robotu uztverei. Bija jāidentificē pieeja, kas spēj veikt gan objektu noteikšanu un pozas novērtēšanu manipulācijas uzdevumiem (2. tēze), gan apvidus segmentāciju lielā mērogā (3. tēze);

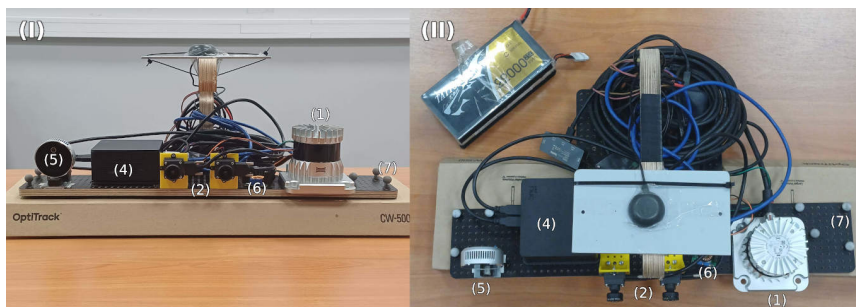
2. DATU KOPA UN ATSKAITES POZAS

Lai ievāktu datus, kas izmantoti šajā disertācijā aprakstīto uztveres sistēmu izstrādē, testēšanā un novērtēšanā, bija nepieciešama izstrādāt sensoru instalāciju un datu vākšanas metodoloģiju. Uzsākot šajā disertācijā aprakstīto pētniecisko darbību, EDI robotikas laboratorijā jau bija ierīkota divu robotu darbstacija, kas aprīkota ar vairākām dziļuma kamerām. Taču nebija pieejama aparatūra, lai ierakstītu LiDAR vai GNSS datus ārtelpu trajektorijās patvaļīgās, nestrukturētās vidēs. Tāpēc tika izstrādāts pārnēsājams sensoru komplekts. Turklāt, lai ļautu veikt lokalizācijas precizitātes novērtēšanu GNSS liegtās vidēs un sistēmām, kas izmanto GNSS, tika realizēta neatkarīgu atskaites pozu mērīšanas metode un tās precizitāte tika novērtēta, izmantojot fiksētu kustības uztveres instalāciju. Izveidotais sensoru komplekts un jaunā atskaites pozu mērīšanas metode pēc tam tika izmantoti, ierakstot *EDI-SLAM* [9] datu kopu. Šī datu kopa tika izmantota borta GNSS-INS sensora lokalizācijas kļūdas mērīšanai, kā arī dažos no semantiskajās uztveres sistēmās izmantoto SLAM sistēmu veikspējas novērtējumiem.

2.1. Sensoru komplekts

Iedvesmojoties no sensoru komplekta, kas izmantots *TUM-VI* [45] vizuāli-inerciālās odometrijas datu kopas ierakstīšanai, tika izveidots pārnēsājams, regulējams rāmis, kam piestiprināt sensorus, skaitļošanas moduli, GNSS antenu un barošanas shēmas, parādīts Attēlā 2.1. LiDAR tika izvēlēts pār dziļuma kameru pamatojoties ar tā tālāko darbības rādiusu un augstāko noturību pret āra apgaismojuma apstākļiem. Kameras tika uzstādītas galvenokārt, lai ļautu izmantot datorredzi semantiskajā segmentācijā; divas ir nepieciešamas, lai būtu iespējams testēt arī stereo un stereo-inerciālās SLAM sistēmas un saglabātu iespēju nākotnē izmantot tīri vizuālas punktu mākoņu ieguves metodes. GNSS-INS navigācijas sistēma kalpo kā galvenā blīvu pozu anotācijas (augstfrekvences un augstas pieejamības) ieguves metode vidēs, kur sagaidāms zems GNSS dreifs, piemēram, atklāta lauka trajektorijās.

Sensoru komplektā ir uzstādīts *Ouster OS1 rev7* [46] 32 līniju mehāniskais LiDAR (1), kas publicē punktu mākoņus ar 10 Hz frekvenci un IMU datus ar 100 Hz frekvenci; $2 \times$ *Basler Dart 1920-160uc* [47] globālā aizvara USB3.0 kameras (2), kas aprīkotas ar fiksēta fokusa attāluma $4mm$ $1/1.8''$ objektīviem; *Xsens MTi-680g* [48] inerciālās navigācijas ierīce (INS) (attēlā nav redzama) ar daudzjoslu *u-blox ZED F9 RTK GNSS* uztvērēju (3); un *Intel nuc* kompaktais dators, kas darbojas ar *Ubuntu Linux 20.04 LTS* un *ROS1 Noetic*, un veic datu ierakstīšanu (4). Testēšanas un kalibrācijas nolūkos iespējams pie rāmja piestiprināt atstarojošos marķierus (7), ko izmanto *Optitrack* kustības uztveres sistēma. Akumulators tiek turēts atsevišķi — mugursomā, ja ierakstīšanu veic, sensoru komplektu turot rokās, salonā, ierakstot ar automašīnu, vai atsevišķi piestiprināts pie UGV plat-



2.1. Attēls: (I) priekšskats un (II) augšskats uz sensoru komplektu, norādot dažus no sensoriem. (III) viens no attēliem, kas apstrādāts atsaucēs pozas mērīšanai *courtyard_gt* trasē, parādot izsekošanas vārtus ar noteiktiem marķieru indeksiem un to stūriem, kas iezīmēti melnā krāsā.

formas — un ir savienots ar rāmi izmantojot noņemamu pagarinātāju. LiDAR un kameras attēlu sinhronizācijai tika radīts specializēts *Pylon* kameras ROS draiveris, kas ieraksta attēlu no vienas vai vairākām kamerām ik reizi, kad ticis publicēts LiDAR punktu mākonis. Kameru projektīvos parametrus un kalibrācijas matricas starp dažādiem sensoriem iegūst izmantojot *kalibr* programmatūras rīkkopu [49].

2.2. Atskaites pozu mērīšana

Ņemot vērā nepieciešamību novērtēt SLAM sistēmas, kas pludina GNSS datus savos lokalizācijas aprēķinos, kā arī vajadzību ierakstīt datus vidēs bez kvalitatīva GNSS signāla, jānodrošina neatkarīga atskaites pozas mērīšanas metode. Viena pieeja, ko izmanto citās datu kopās, ir kustības uztveres sistēmu pielietošana — piemēram, *Optitrack* [45, 50]. Lai gan šādas instalācijas nodrošina pilnīgi no GNSS neatkarīgus mērījumus un milimetra daļās mērāmu precizitāti, tām ir ļoti ierobežota darba zonā, jo nepieciešams izvietot vairākas (specializētas un dārgas) kameras fiksētā konfigurācijā ap interesējošo objektu, lai atstarojošie marķieri būtu redzami no vairākiem rakursiem.

Šajā darbā autors piedāvā alternatīvu atskaites pozas mērīšanas sistēmu SLAM datu kopām, kurās pieejami kameras attēli [51]. Vispirms gar paredzamo kameras trajektoriju tiek izvietoti *Apriltag2* [8] marķieri, pēc tam ar tradicionāliem mērniecības instrumentiem — tahimetru — tiek uzmērīti marķieru stūru punkti vai nu pret brīvi fiksētu atskaites punktu, vai, ja iespējams, ģeoreferencētā ECEF koordinātu sistēmā (izmantojot ar mērnieku aprīkojumu iegūtus neatkarīgus GNSS-RTK mērījumus). Ar zināmiem kameras kalibrācijas parametriem un šo punktu koordinātēm, kameras poza tiek neatkarīgi novērtēta katram kadram, kurā redzami marķieri, izmantojot *perspektīvas-n-punktu* (PnP) algoritmu

[52]. Tipiskā izvietojuma konfigurācijā vizuālie marķieri ir sakārtoti “izsekošanas vārtos”, katrs no kuriem sastāv no divām pamatnes plāksnēm, uz kurām ir uzstādīti divi marķieri, tādējādi nodrošinot kopumā 16 stūrus PnP algoritmam.

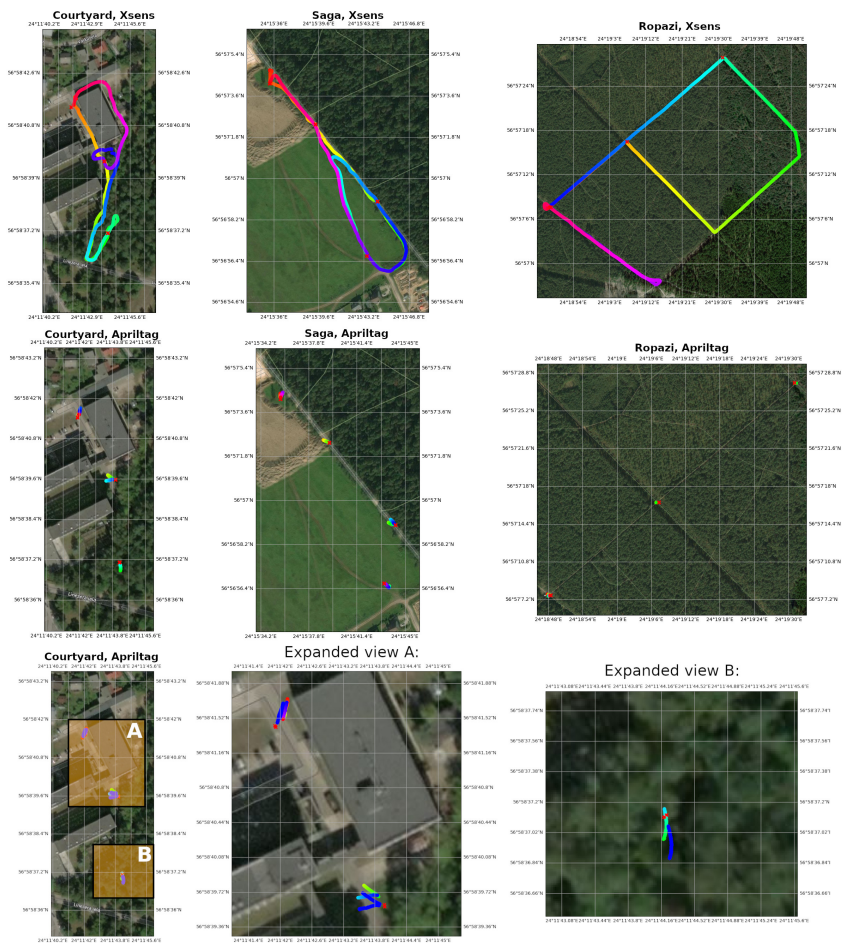
Galvenā šīs sistēmas priekšrocība salīdzinājumā ar *Optitrack* un līdzīgām sistēmām ir tā, ka to var izvietot praktiski jebkur, jo vizuālo izsekošanas vārtu izvietojuma ir daudz vienkāršāka un lētāka nekā kustības uztveres instalācija. Turklāt jebkuri pozu aprēķini pēc konstrukcijas ir laika ziņā sinhronizēti ar kameras kadriem, neveicot papildu kalibrācijas procedūras. Salīdzinot ar GNSS-INS galvenās priekšrocības ir iepriekš minētā statistiskā neatkarība, kā arī spēja izvietot šādu sistēmu iekštelpās vai citās vietās, kur satelīta signāls nav pieejams vai nav uzticams. Ievērojamākie trūkumi ir samazināts pārklājums — pozu aprēķini ir pieejami tikai tuvu uzmērītajiem izsekošanas vārtiem — un fakts, ka pirms jebkādas ierakstīšanas ir jāveic marķieru izvietojuma un uzmērīšana, kas palielina laiku un izmaksas trajektoriju vākšanai jaunās vietās.

2.3. Datu kopa

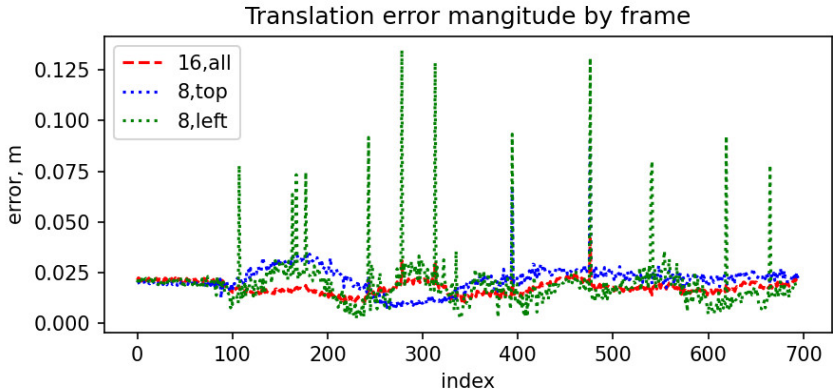
Izmantojot šo sensoru komplektu, draiveru programmatūru un kalibrēšanas metodoloģiju, tika ierakstīta datu kopa. Šobrīd datu kopa ir tās pirmajā publiskās relīzes versijā, un pieejam to aprakstoša publikācija akadēmiskā žurnālā [9]. Relīzes versijā pieejami ieraksti ar un bez atskaites pozu mērījumiem, kas savākti piecās sesijās trijās dažādās vidēs:

- EDI teritorijā, daļā tā apkārtnes — pilsētvide ar dažiem blīvas veģētācijas un krūmāju posmiem. Augstās ēkas pagalmā pasliktina GNSS signāla kvalitāti (2 trases);
- līdzena pļava — ar augstas kvalitātes GNSS datiem, bet plakanu, neizteiksmīgu ainavu, kas apgrūtina relatīvo pārvietojumu noteikšanu ar LiDAR vai kameru (2 trases);
- meža ceļu tīkls — taisni, gari grants ceļi vai sliedes zem meža nojumes, ar vienmuļu ainavu (kas palielina nepatiesu ciklu slēgšanas risku) un pārtraukumiem GNSS signāla pieejamībā (1 trase).

Veicot ierakstus ar atskaites pozas anotāciju, *Apriltag* izsekošanas vārti tika izvietoti dažādās vietās gar plānoto maršrutu. Trajektorija vienmēr sākas ar pirmajiem vārtiem kameras redzes lokā. Pārējie vārti tiek izvietoti aptuveni vienmērīgi gar paredzēto trasi, lai nodrošinātu, ka sistemātiska lokalizācijas kļūda ir manāma jebkurā trajektorijas posmā. EDI pagalma trasē izsekošanas vārti prioritāri tika novietoti vietās ar aizsegtu skatu uz debesīm, kur tika sagaidīta būtiska GNSS degradācija. Sensoru komplekts tika pārvietots divos dažādos veidos vākšanas laikā — pagalmā un vienā no ierakstiem, kas veikti atklātā laukā, komplektu nesa rokās. Daudz garākajai meža ceļa trasei, kā arī otrai lauka trasei, iekārta tika



2.2. Attēls: Satelīta kartes ar atskaites pozīcijām — iegūtām ar GNSS-INS (augšā) vai marķieriem (vidū). Izplesti skati (apakšā) ar marķieru pozām (zilie punkti), pašiem marķieriem (sarkanā) un atbilstošām GNSS-INS pozām (gradienta krāsa).



2.3. Attēls: Vizuālo marķieru pozas novirze no *Optitrack* attiecībā pret novērojuma indeksa numuru, pie dažādām marķieru konfigurācijām.

2.1. Tabula: Atskaites pozu mērīšanas precizitāte.

Marker count and arrangement		Evaluation metric			
		mean ϵ_θ	max ϵ_θ	mean ϵ_t	max ϵ_t
all 4 markers, 16 points		0.43°	2.04°	0.018m	0.042m
top 2 markers, 8 points		0.42°	2.39°	0.022m	0.090m
left 2 markers, 8 points		0.59°	4.51°	0.019m	0.135m
Track	ϵ_{rmse}	$\epsilon_{rmse,XY}$	ϵ_{ate}	$\epsilon_{ate,XY}$	ϵ_{rpe}
<i>courtyard_gt</i>	77.291	6.873	4.000	3.193	0.008
<i>saga_gt</i>	1.305	0.452	0.114	0.099	0.024
<i>ropazi_gt</i>	8.223	2.165	3.059	2.405	0.047

uzstādīta uz automašīnas jumta, lai ļautu veikt garākus ierakstus. Trim no trasēm ir atskaites pozu anotācijas — pa vienai katrā no vidēm. Šo trašu satelītattēli ir parādīti Attēlā 2.2. Datu kopa pieejama zem *Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International* licences [53] EDI domēnā [54].

2.4. Rezultāti

Lai nodrošinātu *EDI-SLAM* datu kopā [9] iekļauto atskaites pozu ticamību, tika veikti gan autora ieviestās uzmērīto *Apriltag* lokalizācijas metodes, gan *Xsens* GNSS-INS trajektorijas novērtējumi, kuru rezultāti apkopoti tabulā 2.1. Vizuālo marķieru metode tika novērtēta attiecībā pret fiksētu *Optitrack* kustības uztveres sistēmu. Tika pārbaudītas vairākas marķieru izvietojuma konfigurācijas, un, kā redzams attēlā 2.3, lai samazinātu izlēcēju kļūdu risku, ir svarīgi izman-

tot vairāk stūra punktu un izvietot tos maksimālu tālu vienu no otra attēla telpā. Izmantojot 4 marķieru plāksnes, tika iegūta vidējā orientācijas kļūda 0.43° un pārvietojuma kļūda aptuveni 1.8 cm, neņemot vērā kalibrācijas sistemātisko kļūdu, ar maksimālo kļūdu zem 5 cm. GNSS-INS sistēmas kļūdas novērtējums ir veikts arī uz *EDI-SLAM* datu kopas, kas apkopots tabulas 2.1 apakšējā daļā. Šie rezultāti liecina, ka GNSS-INS aprēķini lielā mērā sakrīt ar marķieru pozicionēšanu atklātā laukā *saga_gt* trasē — ar vidējo kvadrātisko kļūdu (RMSE) mazāku par 50 cm XY plaknē, neveicot nekādu koordinātu sistēmu savietošānu pirms kļūdas novērtējuma —, bet novērojama būtiska novirze *courtyard_gt* trasē, kur debesis bieži ir aizsegta un augstu ēku klātbūtne izraisa grūti novērtējumu signāla atstarošanas. Lielākā kļūdas komponente ir sistemātiska nobīde GNSS augstumā, bet pastāv arī pārrāvumi, kas ietekmē absolūtās trajektorijas kļūdas (ATE) novērtējumu. Kompromisi starp jauno atskaites pozas mērīšanas metodi un GNSS-INS ir skaidri ilustrēti Attēlā 2.2. Augšējā rindā redzams daudz blīvākais pārklājumu, ko nodrošina augstas *Xsens* GNSS-INS ierīce, savukārt vidējā rindā redzami izklaidētie novērtējumi ar vizuālajiem marķieriem, kas pieejami tikai tuvu izsekošanas vārtiem. Tomēr, aplūkojot *courtyard_gt* trasi, kur konstatēta skaitliski vislielākā GNSS-INS nobīde no *Apriltag* pozām, arī attēlā redzama skaidra trajektorijas novirze — lai gan lielākā kļūdas komponente ir vertikālajā koordinātē, ko nevar redzēt satelītattēlos.

3. STACIONĀRAIS OBJEKTU KRAUŠANAS DEMONSTRATORS

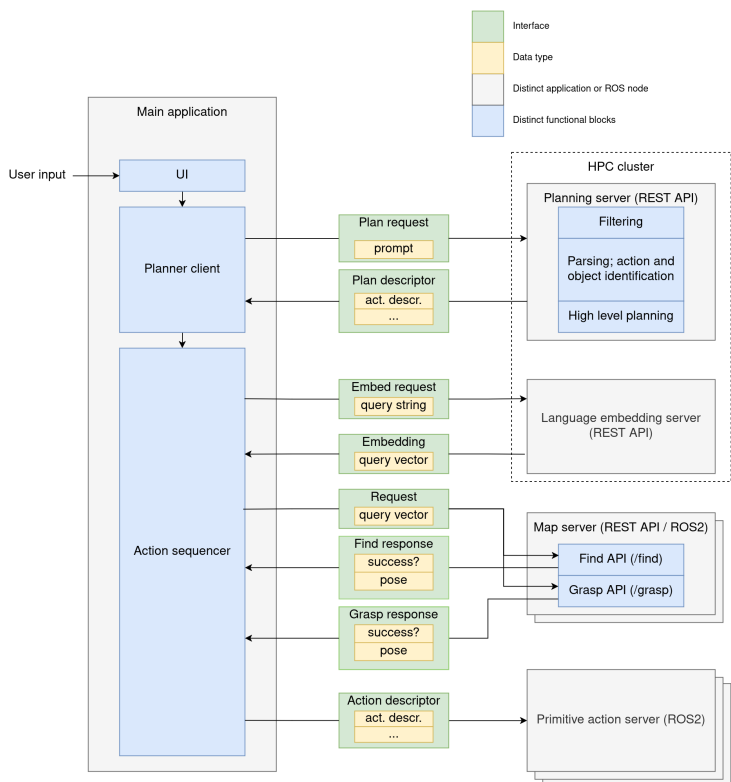
Galvenais ar robota vadības sistēmām saistītās pētnieciskās darbības mērķis ir piedāvāt un realizēt arhitektūru patvaļīga formāta dabiskās valodas komandu izpratnei un izpildei. Saskaņā ar ievadā definētajiem disertācijas mērķiem un uzdevumiem, lēmums par strukturētas apkārtējās vides kartes izmantošanu uztveres sistēmā tika pieņemts jau agrīnā darba stadijā. Lai nodrošinātu modularitāti un pārnesamību, zema līmeņa kontrolei tika nolemts izmantot darbību primitīvus — robota “prasmes”. Augsta līmeņa plānošanas apakšsistēma, kuras pamatā ir valodas modelis, tiek izmantota, lai nestrukturētas komandas dabiskajā valodā pārveidotu robotam izpildāmos darbību plānos. Piedāvātā arhitektūra tika realizēta uz industriālā manipulatora un pārbaudīta gan atsevišķu komponentu līmenī, gan pilna cikla izpildē.

3.1. Vairāku līmeņu vadības sistēma

Lai efektīvi izmantotu mūsdienīgu LLM un multimodālo modeļu piedāvātās iespējas un savietotu tās ar strukturētu vides modeli, plānošanai, vadībai un uztverei tika izstrādāta sekojošā trīs bloku arhitektūra:

- **Augsta līmeņa plānošanas (HLP)** apakšsistēma — nodrošina lietotāja saskarni un dabiskās valodas komandu izpratni, apstrādājot operatora vaicājumus balss komandu vai teksta veidā. Šī moduļa galvenais uzdevums ir ģenerēt kartei uzdotās meklēšanas operācijas un robotam veicamās darbības;
- **Zema līmeņa plānošanas (LLP)** apakšsistēma — šī tehnoloģiju bloka funkcija ir HLP komandēto darbību secības fizisku izpilde, abstrahējot robota konstruktīvos aspektus un tā vadības programmatisko izpildījumu;
- **Semantiskās kartes** apakšsistēma — atvērtās kopas semantiskā karte, kas veido vienotu, strukturētu robota apkārtējās vides (piemēram, statistiskas darba zonas vai ārtelpu apkārtnes) modeli, sapludinot sensoru mērījumus.

Testēšanai un validācijai bija nepieciešams definēt pielietojuma scenāriju un operatora saskarnes modeli. Cilvēks, kas pārrauga robota kontroles sistēmu, uzdod balss vai teksta komandas brīvā valodā. Uz datora, kas kalpo kā robota darbstacijas kontroles mezgls, uzstādīta LLP programmatūra, caur kuru notiek sistēmas vadība, jo tai nepieciešams cieši integrēties ar robota un sensoru draiveriem. Darbību primitīvi realizēti kā ROS darbību protokola serveri — tos izsaucot, tiek izpildīta pārtraucama, parametrizēta kustība. HLP un karte var tikt uzstādīti uz tā paša datora vai attālinātām ierīcēm, un informācija no tiem tiek



3.1. Attēls: Industriālā robota galdas virsmas manipulācijas demonstratora arhitektūra, kas tika izveidota kartes-HLP-LLP integrācijai un kartes validācijai objektu noteikšanas un satveršanas pozas novērtēšanas uzdevumos [10, 12].

Iegūta caur klienta-servera izsaukumiem. Attēlā 3.1 attēlota demonstratora faktiskās realizācijas blokshēma saskaņā ar iepriekš publicētajiem rezultātiem [10, 12]. Reālu sensoru datu ievākšana objektu atpazīšanas testiem un pilna cikla izpilde tika veikta EDI robotikas grupas industriālo robotu darbstatijā. Tā sastāv no divām robotrokām — *Universal Robots UR5e* un *UR5*, no kurām šajā demonstratorā tiek izmantota pirmā — *UR5e*, kas aprīkota ar pirkstu tipa *RobotiQ* satvērēju; virs darba zonas statiski ir uzstādīta strukturētās gaismas tipa *Zivid One+* RGB-D kamera; uz pašas robota rokas ir uzstādīta *Intel Realsense D435* kamera. Darba zona ir plakana tērauda virsma ar komplektēšanas un šķirošanas uzdevumu demonstrācijās izmantotām kastēm paredzētiem stiprinājuma punktiem. Tā kā viena no kamerām ir statiska, bet otra ir uzstādīta uz robota rokas — un robota ROS draiveris aprēķina instrumenta centra punkta koordinātas ar augstu

frekvenci, risinot tiešo kinemātiku — semantisko karti iespējams pārbaudīt bez nepieciešamības izmantot papildus lokalizācijas metodes, piemēram, SLAM.

Augsta līmeņa plānošanas (HLP) apakšsistēma pārveido ienākošos lietotāja pieprasījumus — balsī vai tekstā — par izpildāmu plānu. Tas tiek panākts, veicot secīgus valodas modeļa izsaukumus, kas pakāpeniski aizpilda strukturētas teksta sagataves, papildinot tās ar detaļām no uzdotās komandas. HLP izmanto *LLama3 8B Instruct* [55] C++ implementāciju *llama.cpp* ar 4 bitu GGUF kvantizāciju un *Q_4_K_M* svariem no [56]. Papildus kvantizācijai, ko izmanto, lai samazinātu atmiņas prasības, tiek izmantota uzvedņu ķēdēšana [57], lai samazinātu nepieciešamo konteksta garumu. HLP apakšsistēma īsteno trīs ievades apstrādes posmus: filtrēšanu, parsēšanu un plānošanu. Filtrēšanas posms klasificē ienākošos pieprasījumus vienā no iepriekš definētajām kategorijām — *”mobilā manipulācija”*, kas norāda uz nepieciešamību pāriet uz tālāko apstrādes posmu izpildi, un *”jautājumu atbildēšana”*, *”ikdienišķa saruna”* vai *”jebkāds cits vaicājums”*. Nākamais posms, pieņemot, ka tikusi pieprasīta plānošana, ir instrukciju parsēšana. Šis process sadala uzdevumu smalkākos soļos, katram no kuriem tiek ģenerēts savs plāns, kā Least-to-Most [58] pieejā. Pēdējā posmā — plānošanā — katrs no soļiem tiek interpretēts kā LLP pieejamo darbību secība un papildināts ar nepieciešamajām detaļām. Saliekot visas darbību secības, tiek iegūts plāna deskriptors, ko nodot tālāk LLP sistēmai.

Zema līmeņa plānošanas (LLP) apakšsistēma apstrādā plānu deskriptorus, ko ģenerē HLP. Tā komunicē ar robotu kontrolleriem, plānotājiem un cita veida aparatūras draiveru programmatūru, izmantojot *ROS 2* ietvaru. Kustību plānošanai tiek izmantota *ROS2 MoveIt* [13] versija, kas pārvērš kustības mērķus Eiklīda telpā par robota konfigurācijas telpas iestatījuma punktu secībām. Pētījumu gaitā tiktāl tikuši realizēti pieci darbību primitīvi: **embed**, **find.get_grasp**, **activate_gripper** un **move**. Tie visi tiek implementēti, izmantojot *ROS 2* pakalpojumu izsaukuma saskarni, un tiek komponēti apakšrutīnās, kas veido augstāka līmeņa abstrakciju — **darbības**. Šobrīd LLP realizācija nodrošina divas darbības — **pick** (paņemt) un **place** (novietot).

3.2. Eksperimentālie novērtējumi

Lai novērtētu HLP apakšsistēmas veiktspēju, šajā darbā izmantotā pieeja ir sākt ar iepriekš zināmu plānu kopu, pārveidot tos par komandām dabiskā valodā, un pēc tam pārbaudīt, vai HLP, saņemot pārveidoto komandu, spēj atgūt sākotnējā plāna struktūru — darbību secību un ar tām saistītos pareizos objektus. Ņemot vērā ļoti vienkāršo darbību kopu, kas pieejama pārbaudāmajai demonstratora versijai — *”paņemt”* un *”novietot”* —, tika izstrādāti trīs komandu veidi, lai konstatētu dažādus kļūmju paveidus, kas var rasties plānošanas procesā: *“simple”* — tiešas paņemšanas-novietošanas secības; *“forward”* — atkārtotas paņemšanas-novietošanas komandas parādīšanās secībā; *“reverse”* — divas paņemšanas-novietošanas komandas, kas jāizpilda apgrieztā secībā. Lai ģene-



3.2. Attēls: Vienas komandas izpildes izpilde, uzņemta vienā no pilna cikla izpildes testiem. Uzdotā komanda “*Rest the electrical tape on the multimeter, but only after positioning the computer mouse over the pink foam*” tika veiksmīgi izpildīta. Starp darbībām robota roka pamet darba zonu, lai neaizklātu objektus.

rētu formulējumu permutācijas divos stilistiskos veidos, tika izmantota *OpenAI GPT-4o* [59] interaktīvā asistenta saskarne. Šādi iegūtajās komandu veidnēs var ievietot konkrētu objektu aprakstus, lai iegūtu uzdevumus robotam.

Pilna cikla integrācijas testēšanai un validācijai tika izstrādāts protokols nejausi izvēlētas komandu apakškopas izpildei cilvēka uzraudzībā. Vienas šādas izpildes epizodes laika intervāla attēls ir redzams Attēlā 3.2. Lai iegūtu izpildāmo apakškopu, tiek nejausi izvēlēta grupa ar veidnēm un permutācija ar objektiem no pieejamā klāsta, ko tajās ievietot. Ņemot vērā faktu, ka veicot instrukciju izpildi uz fiziska robota pastāv ļoti daudzi nezināmi faktori, ko iepriekš paredzēt nav iespējams, rezultāti tiek iegūti novērotājam manuāli reģistrējot katrā sistēmas posmā novērotās kļūmes — HLP, uztveres sistēmā, zema līmeņa robotu vadībā — un piešķirot atbilstošu posma izpildes vērtējumu — veiksmīgi, daļēji veiksmīgi vai neveiksmīgi. Kopējais epizodes vērtējums atbilst minimālajam, kas iegūts atsevišķajos izpildes posmos — viens daļēji veiksmīgs rezultāts nozīmē, ka kopējais vērtējums ir ne labāks par daļēju veiksmi, un viena neveiksme nozīmē, ka visa epizode ir neveiksmīga.

3.3. Rezultāti

Abu veidu eksperimentālo novērtējumu būtiskākie rezultāti ir apkopoti Tabulā 3.1. Tabulas augšējā daļa atspoguļo HLP novērtējuma rezultātus. Kreisā kolonna norāda komandas strukturālo tipu (neizšķirot starp stilistiskajiem tipiem).

3.1. Tabula: Robota vadības sistēmas novērtējumu rezultāti.

HLP Verification		
Command type	Count in data set	Correct %
simple	394	99.49
forward	400	92.50
reverse	400	96.40

End-to-end Execution (15 episodes)	
System stage	Success %
HLP	93.33
Perception	86.67
Robot control	73.33
Total	66.67

Vidējā kolonna norāda, cik reizu šāda veida komanda pārstāvēta testa datu kopā. Atšķirības komandu skaitā ir saistītas ar dublēšanos — “*simple*” komandā ir tikai divas vietas objektiem, kas nozīmē, ka virknes, kuras atšķiras tikai ar pēdējiem diviem elementiem, atbilst vienādām aizpildītām veidnēm. Labējā kolonna norāda pareizi rekonstruēto plānu procentuālo daudzumu katram komandas veidam. Aplūkojot HLP veikspējas novērtējuma rezultātus, kļūst skaidrs, ka izmantotā pieeja spēj rekonstruēt komandas no ierobežotās telpas, kas pārstāvēta novērtējuma datos, ar sakritības rādītājiem virs 90% visiem plānu paveidiem. Tomēr novērota neintuitīva sakarība — apgrieztiem saliktajiem plāniem ir zemāki kļūdu rādītāji, nekā tiešiem. Pārskatot konkrētas nepareizi rekonstruētas komandas redzams, ka lielākā daļa neveiksmju ir saistītas vai nu ar halucinētām liekām, atkārtotām darbībām, vai arī ar objektu aprakstu dublēšanos — abas šīs parādības ir biežāk sastopamas tieši strukturētajās komandās. Vēl viena novērota kļūme ir pārlietu pieklājīgu pieprasījumu nepareiza klasifikācija kā “*ikdienišķa saruna*” vai “*jebkurš cits vaicājums*”, kas noved pie šādu komandu ignorēšanas. Tabulas apakšējā daļa apkopo rezultātus, kas iegūti nesenākajā pilna cikla izpildes testā. Kreisā kolonna norāda sistēmas posmu, bet labā kolonna norāda to epizožu procentuālo daudzumu, kurās šis posms veiksmīgi izpildīja savu uzdevumu, ar kopējo veiksmes rādītāju (kad visi posmi bija veiksmīgi un uzdevums tika izpildīts) apakšā. Lai gan ir skaidrs, ka sistēma būtībā spēj veikt vienkāršus kraušanas uzdevumus, ir nepieciešams turpmāks darbs, lai uzlabotu katra posma veikspēju un padarītu to gatavu rūpnieciskai izvietošanai — piemēram, robustāka satveršanas pozas novērtēšana.

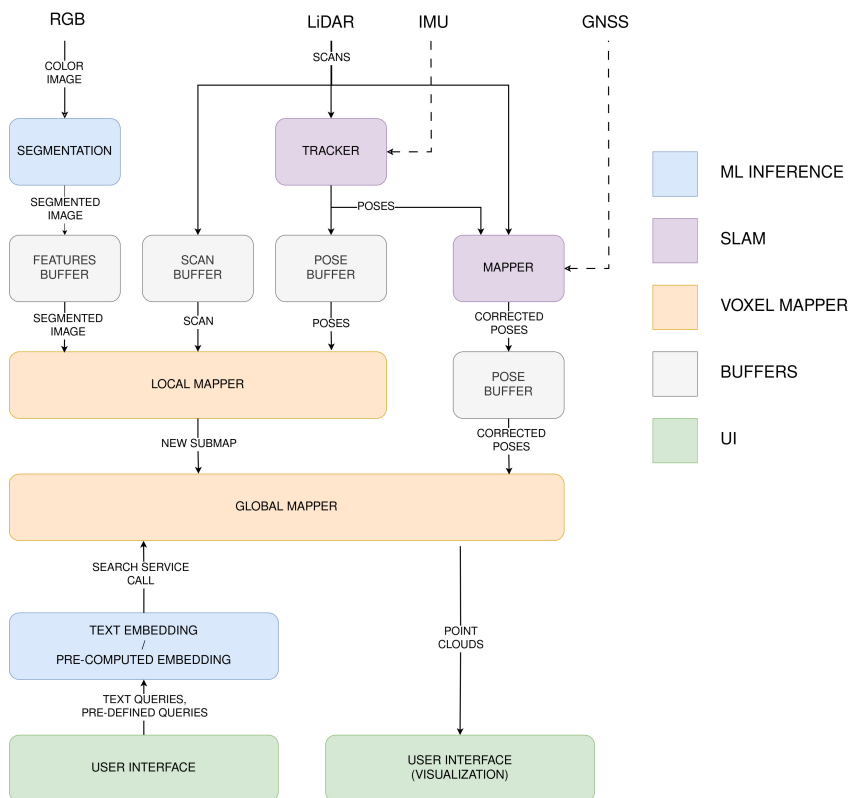
4. SEMANTISKĀS UZTVERES SISTĒMAS

Šīs disertācijas pētnieciskās darbības pamatā bija atvērtās kopas semantisko uztveres sistēmu izstrāde, verifikācija un validācija — procesi, kas īsi aprakstīti šajā nodaļā. Visas tēzes galu galā ir saistītas ar šo tēmu — 1. tēze attiecas uz eksperimentālo metodoloģiju, ko izmanto uztveres sistēmu novērtēšanai; 2. tēze apgalvo, ka atvērtas kopas segmentācijas modeļi var veiksmīgi izmantot manipulācijas uzdevumu realizācijā ar robotu; 3. tēze uzsver, ka ar atbilstoši izvēlētiem vaicājumiem redzes-valodas iegultos vektorus var izmantot apvidus paveidu atšķiršanā, kas ir būtiska problēma autonomu virszemes robotu navigācijā. Tika veikti trīs dažādi eksperimentālo novērtējumu veidi, kas saistīti ar katru no tēzēm — piedāvātā atskaites pozu mērīšanas metode tika izmantota, lai novērtētu SLAM sistēmas lokalizācijas precizitāti; stacionārā demonstratora uztveres sistēmas tika pārbaudītas objektu atpazīšanā; nesenākā ārtelpu kartēšanas sistēmas versija — *SLAMVDB* — tika novērtēta apvidus segmentācijas precizitātē.

4.1. Izstrādātās versijas

Iteratīva izstrādes procesa ietvaros tika izstrādātas četras uztveres sistēmas, augošā secībā pēc sarežģītības, katrā no tām ieviešot iepriekšējos posmos gūtās atziņas:

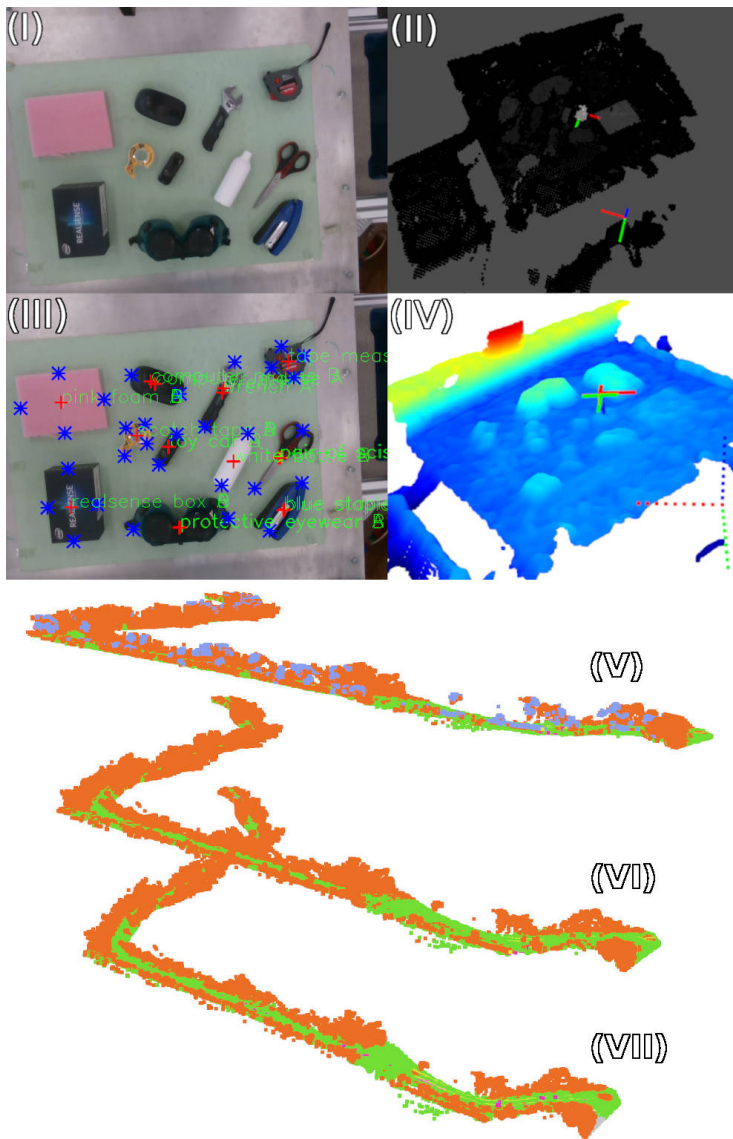
1. **Dziļuma karte** — ”2.5-dimensionāla” karte, iegūta no viena krāsu un dziļuma attēla pāra, izmantota sākotnējai darbības pamatprincipu pārbaudei objektu atrašanas un to satveršanas pozu noteikšanas metodēm;
2. **Vektoru oktantkoks** — vienota, 3-dimensionāla oktantkoka datu struktūra, kas režģa šūnās glabā redzes-valodas iegultos vektorus. Izstrādāta galvenokārt, lai pārbaudītu metodes vairāku semantisko novērojumu integrācijai vienā un tajā pašā režģa šūnā, kā arī lai noteiktu objektu noteikšanas metožu pārnesamību no viena dziļuma attēla uz pilnībā 3-dimensionālu karti;
3. **Sekvenciālā** semantiskā SLAM sistēma — paredzēta liela mēroga karšu veidošanai; pēta veidus, kā pielāgot vokseļu režģa karti ārtelpu uztveres izaicinājumiem — proti, sensoru dreifa uzkrāšanos, jau integrētu režģa šūnu apvienošanu ciklu slēgšanas ietvaros un nepieciešamību veidot režģa kartes, kas pārsniedz maksimāli indeksējamus izmērus. Izmanto sekvenciālu **apakškaršu** datu struktūru, kurā katra apakškarte realizēta, izmantojot reinātas matricas;
4. Laikā indeksētā semantiskā SLAM sistēma (**SLAMVDB**) [11] — nesenākā uztveres sistēmas versija, kas pilnībā nošķir lokalizācijas un vokseļu kartēšanas posmus, lai nodrošinātu modularitāti, apvienojot Vektoru oktantkoka



4.1. Attēls: Komponentes un datu plūsmas nesenākajā atvērtās kopas semantiskās kartēšanas sistēmas iterācijā [11]. Pārtrauktās līnijas apzīmē datu plūsmas, kas nav obligāti nepieciešamas — IMU un GNSS.

datu struktūru ar Sekvenciālās sistēmas apakškaršu nodalījumu. Šīs sistēmas blokshēma ir attēlota attēlā 4.1.

Dziļuma kartes un Vektoru oktantkoka sistēmas tika uzstādītas un pārbaudītas stacionārā objektu kraušanas scenārijā [10], kas detalizēti aprakstīts iepriekšējā nodaļā. Objektu atpazīšanas precizitātes novērtējumi [10] tika veikti, izmantojot šīs sistēmas, un algoritmi, ar kuriem tika pārbaudīts Vektoru oktantkoks, ir pārnesti uz abām turpmākajām kartes versijām. Sekvenciālā sistēma tika izmantota sākotnējos āra testos un tika izmantota lokalizācijas precizitātes novērtējumā [12] — tai skaitā validējot SLAM sistēmā izmantoto GNSS mērījumu integrācijas metodi. *SLAMVDB* apkopo iepriekšējo iterāciju izstrādē gūtās atziņas, kā arī demonstrē ārēju lokalizācijas sistēmu pielietojumu.



4.2. Attēls: (I-IV) Aina no pārbaudes datu kopas. (V-VII) Kategorijas “void” atmešana ārtelpu kartēšanas sistēmā.

4.2. Eksperimentālie novērtējumi

Lai novērtētu Dziļuma kartes un Vektoru oktantkoka spēju atrast vēlamo objektu pēc teksta apraksta, tika savākta pārbaudes ainu datu kopa, kas redzama Attēlā 4.2 (I-IV). Ievācot datus, uz darba virsmas tiek izkārtoti objekti (I), un virsmas attēls, ielādētu īpaši šim mērķim izstrādātā lietojumprogrammā (II), manuāli anotē cilvēks, norādot objektu centroīdas (sarkanie marķieri), satveršanas virzienus (zilie marķieri) un teksta aprakstus (zaļā krāsā). Tiek ierakstīta arī dziļuma attēlu sērija, ko var rekonstruēt kā vokseļu karti (III). Objektu noteikšanu šādā kartē veic, izmantojot līdzības novērtējuma sliekšni (gaišā krāsa attēlā norāda līdzību ar vaicājumu "līmleņķi"). Novērtēšanā kartē tiek meklēts katrs no pieejamajiem objektiem, izmantojot tā teksta aprakstu, un kartes atgrieztā objekta centroīda vai satveršanas poza tiek salīdzināta ar tuvāko objektam piešķirto atskaites pozu (IV). Lai noteiktu, vai var izmantot iegulto vektoru kvantizāciju kā atmiņas prasību samazināšanas metodi, tika veikti testi ar skalārām vērtībām ternārā, baitu un 32 bitu izšķirtspējā.

Lai pārbaudītu āra kartēs izmantoto lokalizācijas novērtējumu kvalitāti, Sekvenciālās kartēšanas sistēmas SLAM precizitāte tika mērīta attiecībā pret *Apriltag* marķieriem EDI pagalmā (cilpveida trajektorijā) un atklātā laukā (atvērtā trajektorijā, ko koriģē GNSS). Šis novērtējums izdarīts pēc tāda paša principa kā GNSS-INS pozas kļūdas novērtēšana, kas aprakstīta sadaļā 2.2 — lai gan šajos rezultātos tiek ņemta vērā tikai absolūtā trajektorijas kļūda (ATE) [60]. *lio2* [20] relatīvās pozas noteikšanas sistēma, ko atbalsta *SLAMVDB*, tika novērtēta uz *RELLIS-3D*. Visbeidzot, arī *SLAMVDB* apvidus segmentācijas precizitāte tika izvērtēta uz *RELLIS-3D* datu kopas, kas satur divu veidu patiesās semantikas anotāciju — attēlos un LiDAR punktu mākoņos. Abas šīs reprezentācijas ir sensoram lokālas un jāprojicē kartes koordināšu sistēmā, lai veiktu jebkādu salīdzinājumu. Tā kā pat neliels lokalizācijas sistēmas dreifs ātri noslēpj jebkādu semantiskās segmentācijas kļūdu citu iemeslu dēļ, lokalizācijas precizitāte un kartes veidošana tiek uzskatītas par atsevišķām problēmām. Tāpēc, konstruējot karti semantiskajam novērtējumam, izmanto patiesās sensora pozas, un tā pati trajektorija tiek izmantota arī, projicējot sensoram lokālos patiesās semantiskās kategorijas marķējumus kartes fiksētajā koordināšu sistēmā.

4.3. Rezultāti

Tabula 4.1 ievērojamākos rezultātus no trim minētajiem eksperimentālajiem novērtējumiem. Augšējais bloks apkopo objektu atpazīšanas precizitāti, kas, izņemot vienu sistēmas konfigurāciju, abām pārbaudītajām kartes versijām ir virs 90% un neuzrāda skaidras likumsakarības ar skalāru kvantizācijas pakāpi. Tomēr atsauces ainu datu kopa ir neliela, tāpēc rezultāti ir tikai indikatīvi. Vairākos gadījumos pareizi un nepareizi atgūto objektu kopas ir identiskas, bet skaitlisko atšķirību pamatā ir neliels skaits ar šķietami nejaušām izmaiņām.

Vidējais bloks apkopo rezultātus no lokalizācijas precizitātes testiem. *EDI-*

4.1. Tabula: Uztveres sistēmas precizitātes novērtējuma rezultāti.

Object Recall for Grasping				
System		Quantization	Recall, %	
Depth Map		ternary	91.04%	
		byte	92.29%	
		float32	92.29%	
Vector Octree		ternary	94.69%	
		byte	87.60%	
		float32	93.85%	

Localization				
System	Data set	Track	l, m	ATE, m
Sequential, loops	<i>EDI-SLAM</i>	<i>courtyard_gt</i>	696.09	0.057
Sequential, GNSS	<i>EDI-SLAM</i>	<i>saga_gt</i>	421.86	0.485
<i>SLAMVDB, lio2</i>	<i>RELLIS-3D</i>	<i>00000</i>	329.76	2.12
<i>SLAMVDB, lio2</i>	<i>RELLIS-3D</i>	<i>00003</i>	256.58	6.08
<i>SLAMVDB, lio2</i>	<i>RELLIS-3D</i>	<i>00004</i>	228.41	0.91

Terrain Segmentation Accuracy				
Track	with <i>void</i>		without <i>void</i>	
	ρ_{hit} %	acc. %	ρ_{hit} %	acc. %
<i>00000</i>	80.28	84.69	81.26	87.35
<i>00003</i>	79.57	79.19	80.39	81.37
<i>00004</i>	73.17	74.22	73.90	76.17

SLAM datu kopā LiDAR-ICP *SLAM* sistēma skaidri demonstrē spēju neitralizēt dreifu pat salīdzinoši garās trajektorijās — ar ciklu slēgšanu garajā, cilpveida pagalma trasē un izmantojot GNSS ierobežojumus atklātā laukā bez cilpām. *RELLIS-3D* datu kopā pat augstākas tehniskas sarežģītības pakāpes LiDAR-inercialā *lio2* sistēma uzkrāj ievērojamu dreifu īsākās trasēs, jo daudzos posmos trūkst pietiekami uzkrītošu pazīmju efektīvai relatīvā pārvietojuma noteikšanai, un ir novērojamas straujas robota platformas rotācijas — skaidri norādot uz nepieciešamību atdalīt lokalizācijas un semantiskās precizitātes novērtējumus.

Apakšējais bloks parāda LiDAR punktu klasifikācijas precizitāti ρ_{hit} un vokseļu klasifikācijas precizitāti $acc.\%$, ko sasniedz *SLAMVDB*. Konstruktīvi neiespējamo “*void*” šūnu novērojumu likvidēšana kartē uzlabo precizitāti, sasniedzot pat 87.35% garākajā, bet vismazāk sarežģītajā *00000* trasē. Tas ir redzams arī attēlā 4.2 — daļēji caurspīdīgas koku galotnes (V) sākotnēji tiek kļūdaini uzskatītas par debesīm, bet pēc “*void*” izņemšanas no ontoloģijas tiek atgūta to patiesā (VII) “*obstacle*” klase (VI).

KOPSAVILKUMS UN SECINĀJUMI

Promocijas darba ietvaros tika izvirzītas un validētas trīs galvenās tēzes — saistītas ar SLAM sistēmas lokalizācijas precizitātes novērtēšanas metodoloģiju, atvērtās kopas semantikas izmantošanu robotu manipulācijā, paklausot nespēcīta uzdotām komandām dabiskā valodā, un apvidus segmentācijas veikšanu virszemes robotu navigācijai, izmantojot vispārīga pielietojuma redzes-valodas iegulšanas modeli. Visos trīs gadījumos tēzēs izteiktie apgalvojumi ir pamatoti ar eksperimentāliem rezultātiem. Darbā piedāvātā apvidus segmentācijas metode, ciktāl autoram zināms, ir unikāla savā pieejā, pirmoreiz ierosināta kā viena daļa no plašākas atvērtās kopas uztveres un cilvēka saskarnes tehnoloģijas, kas izstrādāta MOTE projekta ietvaros [12]. Turklāt no tēzēm izdarītie secinājumi tālāk izmantoti tajā pašā pētnieciskajā darbā — piedāvātā atskaites pozu mērīšanas metode, kuras priekšrocības pār GNSS-RTK vidēs ar degradētu signāla kvalitāti ir eksperimentāli pārbaudīta, pēc tam tika izmantota izstrādāto lokalizācijas metožu pārbaudē.

Darba gaitā iegūtie rezultāti ir iezīmējuši jaunus pētniecības virzienus un jautājumus. Diskusijās ar nozares pārstāvjiem ir identificēta skaidra nepieciešamība industriālo robotu kontrolē spēt izdarīt vairāk, nekā izpildīt balss komandu vienu reizi. Tā vietā, lai izmantotu šajā disertācijā prezentēto komandu-darbību arhitektūru, tirgū ir nepieciešama spēja valodu pārvērst par tipveida, atkārtojamām robota programmām, ko tas spēj atkārtot daudzas reizes. Koncentrējoties uz uztveres aspektiem, skaidri iezīmējas nepieciešamība integrēt spējīgākus objekta satveršanas un novietošanas pozu noteikšanas paņēmienus semantiskajā kartē, jo pēdējais solis pilna cikla izpildes testos — fiziskas manipulācijas ar objektiem — visbiežāk noved pie kļūmēm. Apvidus segmentācijā nepieciešams veikt pētījumus ar jaunākām redzes iegulšanas modeļu arhitektūrām, lai mazinātu šajā darbā sastapto kompromisu starp attēla apstrādes laika prasībām un pieļaujamo attēla izšķirtspēju. Turklāt ir nepieciešami pētījumi par veidiem, kā modelēt mobilu objektu klātbūtni vokseļu kartēs — kas pašlaik spēj modelēt tikai nekustīgas ainas, un dzēs momentānus novērojumus kā troksni. Visbeidzot, nepieciešamība integrēt semantisko informāciju ģeometriskās režģa kartes veidošanas un lokalizācijas posmos kļūst arvien manāmāka — paredzams, ka šāda pieeja samazinās tādu problēmu incidenci kā semantiski “tukšu” vērtību eksistence aizņemtos kartes vokseļos un uzlabos precizitāti SLAM cilpu slēgšanā.

BIBLIOGRĀFIJA

1. Forum, W. E. *The Future of Jobs Report 2025* tehn. ziņ. (World Economic Forum, Geneva, 2025). https://reports.weforum.org/docs/WEF_Future_of_Jobs_Report_2025.pdf (2025).
2. Ryan, M. *Labour and skills shortages in the agro-food sector* OECD Food, Agriculture and Fisheries Papers 189 (Organisation for Economic Co-operation and Development (OECD), Paris, 2023). <https://doi.org/10.1787/ed758aab-en> (2025).
3. Arents, J. & Greitans, M. *Smart Industrial Robot Control Trends, Challenges and Opportunities within Manufacturing* (2022).
4. Badue, C. S. u. c. Self-Driving Cars: A Survey. *ArXiv* abs/1901.04407 sējums (2019).
5. Oliveira, L. F., Moreira, A. P. & Silva, M. F. Advances in agriculture robotics: A state-of-the-art review and challenges ahead. *Robotics* 10. sējums, 52. lpp. (2021).
6. Oliveira, L. F., Moreira, A. P. & Silva, M. F. Advances in forest robotics: A state-of-the-art survey. *Robotics* 10. sējums, 53. lpp. (2021).
7. Racinskis, P., Arents, J. & Greitans, M. Constructing maps for autonomous robotics: An introductory conceptual overview. *Electronics* 12. sējums, 2925. lpp. (2023).
8. Wang, J. & Olson, E. *AprilTag 2: Efficient and robust fiducial detection Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (2016).
9. Racinskis, P., Krasnikovs, G., Arents, J. & Greitans, M. The EDI Multi-Modal Simultaneous Localization and Mapping Dataset (EDI-SLAM). *Data (2306-5729)* 10. sējums (2025).
10. Racinskis, P., Vismanis, O., Zinars, T. E., Arents, J. & Greitans, M. Towards Open-Set NLP-Based Multi-Level Planning for Robotic Tasks. *Applied Sciences* 14. sējums, 10717. lpp. (2024).
11. Peteris Racinskis. *SLAMVDB - the EDI SLAM Vector Data Base (pre-release version)* <https://github.com/edi-administrator/SLAMVDB>. Accessed: 2025-05-03.
12. *RoLISe T4.1 Project Report* https://www.edi.lv/RoLISe_T4_1. (Last accessed: 30.04.2025).
13. Open Robotics. *moveit - Package Summary* <http://wiki.ros.org/moveit>. Accessed: 2025-05-02.

14. Open Robotics. *move_base - Package Summary* http://wiki.ros.org/move_base. Accessed: 2025-05-04.
15. *ROS Wiki: Movebase Global Planner* [Online; accessed on 03-June-23]. http://wiki.ros.org/global_planner.
16. Thrun, S., Burgard, W. & Fox, D. *Probabilistic robotics* ISBN: 9780262201629 (MIT Press, Cambridge, Mass., 2005).
17. Dellaert, F. & Kaess, M. Factor Graphs for Robot Perception. *Foundations and Trends® in Robotics* 6. sējums, 1.—139. lpp. ISSN: 1935-8253. <http://dx.doi.org/10.1561/2300000043> (2017).
18. Sun, K. *u. c.* Robust Stereo Visual Inertial Odometry for Fast Autonomous Flight. *IEEE Robotics and Automation Letters* 3. sējums, 965.—972. lpp. (2018).
19. Bloesch, M., Omari, S., Hutter, M. & Siegwart, R. *Robust visual inertial odometry using a direct EKF-based approach 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (2015), 298.—304. lpp.
20. Xu, W., Cai, Y., He, D., Lin, J. & Zhang, F. FAST-LIO2: Fast Direct LiDAR-Inertial Odometry. *IEEE Transactions on Robotics* 38. sējums, 2053.—2073. lpp. (2022).
21. Campos, C., Elvira, R., Rodr'iguez, J. J. G., Montiel, J. M. M. & Tardós, J. D. ORB-SLAM3: An Accurate Open-Source Library for Visual, Visual-Inertial, and Multimap SLAM. *IEEE Transactions on Robotics* 37. sējums, 1874.—1890. lpp. (2020).
22. Qin, T., Li, P. & Shen, S. VINS-Mono: A Robust and Versatile Monocular Visual-Inertial State Estimator. *IEEE Transactions on Robotics* 34. sējums, 1004.—1020. lpp. (2018).
23. Lu, G., Yang, H., Li, J., Kuang, Z. & Yang, R. A Lightweight Real-Time 3D LiDAR SLAM for Autonomous Vehicles in Large-Scale Urban Environment. *IEEE Access* 11. sējums, 12594.—12606. lpp. (2023).
24. Dellaert, F. & Contributors. *borglab/gtsam* 4.2a8 versija. [Online; accessed 01-June-2023]. <https://github.com/borglab/gtsam>.
25. Hornung, A., Wurm, K. M., Bennewitz, M., Stachniss, C. & Burgard, W. OctoMap: an efficient probabilistic 3D mapping framework based on octrees. *Autonomous Robots* 34. sējums, 189.—206. lpp. <https://api.semanticscholar.org/CorpusID:8655888> (2013).
26. Chatila, R. & Laumond, J.-P. Position referencing and consistent world modeling for mobile robots. *Proceedings. 1985 IEEE International Conference on Robotics and Automation* 2. sējums, 138.—145. lpp. (1985).

27. Kuipers, B. *Modeling Spatial Knowledge International Joint Conference on Artificial Intelligence* (1978).
28. Kuipers, B. The Spatial Semantic Hierarchy. *Artif. Intell.* 119. sējums, 191.—233. lpp. (2000).
29. Kirillov, A., He, K., Girshick, R. B., Rother, C. & Dollár, P. Panoptic Segmentation. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 9396.—9405. lpp. (2018).
30. Radford, A. *u. c. Learning Transferable Visual Models From Natural Language Supervision International Conference on Machine Learning* (2021).
31. Jatavallabhula, K. M. *u. c. ConceptFusion: Open-set Multimodal 3D Mapping. ArXiv abs/2302.07241 sējums* (2023).
32. Li, B., Weinberger, K. Q., Belongie, S. J., Koltun, V. & Ranftl, R. Language-driven Semantic Segmentation. *ArXiv abs/2201.03546 sējums*. <https://api.semanticscholar.org/CorpusID:245836975> (2022).
33. Jiang, P., Osteen, P. R., Wigness, M. B. & Saripalli, S. RELIS-3D Dataset: Data, Benchmarks and Analysis. *2021 IEEE International Conference on Robotics and Automation (ICRA)*, 1110.—1116. lpp. (2020).
34. Wigness, M., Eum, S., Rogers, J. G., Han, D. & Kwon, H. *A RUGD Dataset for Autonomous Navigation and Visual Perception in Unstructured Outdoor Environments International Conference on Intelligent Robots and Systems (IROS)* (2019).
35. Guan, T., Kothandaraman, D., Chandra, R. & Manocha, D. GANav: Group-wise Attention Network for Classifying Navigable Regions in Unstructured Outdoor Environments. *ArXiv abs/2103.04233 sējums* (2021).
36. Hu, Y. *u. c. Toward General-Purpose Robots via Foundation Models: A Survey and Meta-Analysis. arXiv preprint: arXiv:2312.08782* (2023).
37. Wang, L. *u. c. A survey on large language model based autonomous agents. Frontiers of Computer Science* 18. sējums, 186345. lpp. (2024).
38. Song, C. H. *u. c. LLM-Planner: Few-Shot Grounded Planning for Embodied Agents with Large Language Models Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)* (2023).
39. Lin, K., Agia, C., Migimatsu, T., Pavone, M. & Bohg, J. Text2Motion: from natural language instructions to feasible plans. *Autonomous Robots*. ISSN: 1573-7527. <https://doi.org/10.1007/s10514-023-10131-7> (2023).
40. Ahn, M. *u. c. Do As I Can, Not As I Say: Grounding Language in Robotic Affordances Conference on Robot Learning* (2022).

41. Brohan, A. *u. c.* RT-2: Vision-Language-Action Models Transfer Web Knowledge to Robotic Control. *ArXiv* abs/2307.15818 sējums. <https://api.semanticscholar.org/CorpusID:260293142> (2023).
42. Cremona, J., Comelli, R. & Pire, T. Experimental evaluation of Visual-Inertial Odometry systems for arable farming. *Journal of Field Robotics* 39. sējums, 1123.—1137. lpp. (2022).
43. Rosinol, A., Abate, M., Chang, Y. & Carlone, L. *Kimera: an Open-Source Library for Real-Time Metric-Semantic Localization and Mapping* 2020 *IEEE International Conference on Robotics and Automation (ICRA)* (2020), 1689.—1696. lpp.
44. Liu, Y. *u. c.* Accurate 3-D Semantic Segmentation of Point Clouds for Intelligent Vehicles Based on Multiview Edge Guidance and Fusion. *IEEE Sensors Journal* 24. sējums, 26853.—26865. lpp. <https://api.semanticscholar.org/CorpusID:270807942> (2024).
45. Schubert, D. *u. c.* The TUM VI Benchmark for Evaluating Visual-Inertial Odometry. *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 1680.—1687. lpp. (2018).
46. Ouster. *OS1 Hardware User Manual* <https://data.ouster.io/downloads/hardware-user-manual/hardware-user-manual-revd-os1.pdf>. Accessed: 2025-05-08.
47. Basler. *daa1920-160uc* <https://docs.baslerweb.com/daa1920-160uc>. Accessed: 2025-05-08.
48. Movella. *MTi-680G* <https://www.xsens.com/hubfs/Downloads/Leaflets/MTi-680G.pdf>. Accessed: 2025-05-03.
49. Furgale, P. T., Rehder, J. & Siegwart, R. Y. Unified temporal and spatial calibration for multi-sensor systems. *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 1280.—1286. lpp. <https://api.semanticscholar.org/CorpusID:15778738> (2013).
50. Burri, M. *u. c.* The EuRoC micro aerial vehicle datasets. *The International Journal of Robotics Research*. eprint: <http://ijr.sagepub.com/content/early/2016/01/21/0278364915620033.full.pdf+html>. <http://ijr.sagepub.com/content/early/2016/01/21/0278364915620033.abstract> (2016).
51. Racinskis, P., Arents, J. & Greitans, M. *Annotating SLAM data sets with Apriltag markers* 2024 *10th International Conference on Automation, Robotics and Applications (ICARA)* (2024), 438.—442. lpp.
52. Terzakis, G. & Lourakis, M. I. A. *A Consistently Fast and Globally Optimal Solution to the Perspective-n-Point Problem* *European Conference on Computer Vision* (2020). <https://api.semanticscholar.org/CorpusID:226239551>.

53. Creative Commons. *Attribution-NonCommercial-ShareAlike 4.0 International* <https://creativecommons.org/licenses/by-nc-sa/4.0/>. Accessed: 2024-11-01.
54. EDI. *EDI-SLAM data* http://edi.lv/EDI-SLAM_dataset. 2024.
55. Dubey, A. u. c. *The Llama 3 Herd of Models* 2024. arXiv: 2407.21783 [cs.AI]. <https://arxiv.org/abs/2407.21783>.
56. Bartowski. *Llamacpp imatrix Quantizations of Meta-Llama-3-8B-Instruct* <https://huggingface.co/bartowski/Meta-Llama-3-8B-Instruct-GGUF>. Accessed: 2024-10-22.
57. Wu, T., Terry, M. & Cai, C. J. *AI Chains: Transparent and Controllable Human-AI Interaction by Chaining Large Language Model Prompts* 2022. arXiv: 2110.01691 [cs.HC]. <https://arxiv.org/abs/2110.01691>.
58. Zhou, D. u. c. *Least-to-Most Prompting Enables Complex Reasoning in Large Language Models* 2023. arXiv: 2205.10625 [cs.AI]. <https://arxiv.org/abs/2205.10625>.
59. OpenAI. *GPT-4o System Card* <https://cdn.openai.com/gpt-4o-system-card.pdf>. Accessed: 2024-11-14.
60. Sturm, J., Engelhard, N., Endres, F., Burgard, W. & Cremers, D. A benchmark for the evaluation of RGB-D SLAM systems. *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 573.—580. lpp. (2012).