



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

ÚSTAV AUTOMATIZACE A MĚŘICÍ TECHNIKY

DEPARTMENT OF CONTROL AND INSTRUMENTATION

DETEKCE ZMĚN TRASY

ROUTE CONTROL

DIPLOMOVÁ PRÁCE

MASTER'S THESIS

AUTOR PRÁCE

AUTHOR

Bc. Dávid Dolobáč

VEDOUCÍ PRÁCE

SUPERVISOR

Ing. Miloslav Richter, Ph.D.

BRNO 2024



Diplomová práce

magisterský navazující studijní program **Kybernetika, automatizace a měření**

Ústav automatizace a měřicí techniky

Student: Bc. Dávid Dolobáč

ID: 211141

Ročník: 2

Akademický rok: 2023/24

NÁZEV TÉMATU:

Detekce změn trasy

POKYNY PRO VYPRACOVÁNÍ:

Úkolem je navrhnout aplikaci, která zajistí asistenci při průchodu známou trasou s upozorněním na významné body trasy a na změny (geometrické rozdíly) v dané trase.

- 1) Nastudujte možná řešení úkolu a srovnejte je. Pro vybrané řešení nastudujte základní vlastnosti detektorů, zpracování a uchování dat.
- 2) Popište vlastnosti scény mající vliv na řešení. Zvolte úlohu (typ scény a její vlastnosti), pro kterou budete realizovat řešení. Nasnímejte data pro testování algoritmů.
- 3) Navrhňte a realizujte metody a algoritmy pro zpracování dat a zvolte vhodnou datovou reprezentaci pro manipulaci s daty a uložení vzorové předlohy pro porovnání.
- 4) Navrhňte a realizujte algoritmy pro nalezení aktuální polohy v uložených vzorových datech a stanovení rozdílů. Pro testovací i vzorová data použijte videosekvence z bodu 2.
- 5) Zhodnoťte výsledky realizovaného řešení.

DOPORUČENÁ LITERATURA:

Gonzalez R.C., Woods R.E.: Digital Image Processing, 4th edition, Pearson, 2017, ISBN 978-0133356724

Termín zadání: 5.2.2024

Termín odevzdání: 15.5.2024

Vedoucí práce: Ing. Miloslav Richter, Ph.D.

doc. Ing. Petr Fiedler, Ph.D.

předseda rady studijního programu

UPOZORNĚNÍ:

Autor diplomové práce nesmí při vytváření diplomové práce porušit autorská práva třetích osob, zejména nesmí zasahovat nedovoleným způsobem do cizích autorských práv osobnostních a musí si být plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č.40/2009 Sb.

ABSTRAKT

Táto práca si kladie za úlohu detegovať geometrické zmeny, ktoré sa nachádzajú na vopred známej trase, ktorá je braná ako predloha alebo vzor. Geometrické zmeny predstavujú rigídne transformácie objektov v scéne, konkrétnie ich rotáciu, posunutie a ich kombinácie. Ďalším typom zmien je pridanie nových alebo odstránenie pôvodných objektov. Pôvodná trasa je snímaná monokulárnu kamerou a z videosekvencie je metódami fotogrammetrie zrekonštruovaný 3D model scény v bodovej reprezentácii. Rovnakým spôsobom je vytvorený 3D model scény, ktorá obsahuje zmeny, avšak tento model zachytáva iba lokálny úsek pôvodnej trasy. Toto lokálne mračno bodov je registrované v 3D modeli pôvodnej trasy tak, aby sa 3D modely mohli medzi sebou priamo porovnať. Vyšetrovaný priestor je rozdelený na kubické voxely rovnakej veľkosti, ktoré sú postupne prechádzané a pre každý z nich sa vyšetruje miera podobnosti lokálnych povrchov. Výsledkom riešenia sú 3D body z oboch mračien označené ako zmenené alebo nezmenené. Uvedené metódy sú testované na reálnych dátach v scenároch s rôznymi geometrickými úpravami scény a vyhodnotenie detekcie zmeny je realizované ako v prípade binárneho klasifikátoru.

KĽÚČOVÉ SLOVÁ

detekcia zmien, voxel, povrchové normály, histogram orientácií, fotogrammetria, kalibrácia kamery, 3D rekonštrukcia, mračno bodov, 3D model, registrácia, lokalizácia, MATLAB

ABSTRACT

This thesis sets itself the task to detect geometric changes, located on a previously known route, which is considered as a template or pattern. Geometric changes represent rigid transformations of objects in a scene, specifically their rotation, displacement and their combinations. Another type of change is the addition of new or removal of original objects. The original route is captured with a monocular camera and by using methods of photogrammetry, a 3D model of a scene in a point representation is reconstructed from the video sequence. In the same manner, a 3D model of the scene containing changes is created, but this model captures only the local section of the original route. This local point cloud is registered in the 3D model of the original route, so that the 3D models can be directly compared with each other. The investigated space is divided into cubic voxels of the same size which are successively traversed and the degree of similarity of local surfaces is investigated for each of them. The solution results in 3D points from both clouds marked as changed or unchanged. The mentioned methods are tested on real data in scenarios with different geometric modifications of the scene and the evaluation of change detection is implemented as in the case of a binary classifier.

KEYWORDS

change detection, voxel, surface normals, orientation histogram, photogrammetry, camera calibration, 3D reconstruction, point cloud, 3D model, registration, localization, MATLAB

DOLOBÁČ, Dávid. *Detekce změn trasy*. Diplomová práca. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav automatizace a měřicí techniky, 2024. Vedúci práce: Ing. Miloslav Richter, Ph.D.

Vyhľásenie autora o pôvodnosti diela

Meno a priezvisko autora: Bc. Dávid Dolobáč

VUT ID autora: 211141

Typ práce: Diplomová práca

Akademický rok: 2023/24

Téma záverečnej práce: Detekce změn trasy

Vyhlasujem, že svoju záverečnú prácu som vypracoval samostatne pod vedením vedúcej/cého záverečnej práce, s využitím odbornej literatúry a ďalších informačných zdrojov, ktoré sú všetky citované v práci a uvedené v zozname literatúry na konci práce.

Ako autor uvedenej záverečnej práce ďalej vyhlasujem, že v súvislosti s vytvorením tejto záverečnej práce som neporušil autorské práva tretích osôb, najmä som nezasiahol nedovoleným spôsobom do cudzích autorských práv osobnostných a/alebo majetkových a som si plne vedomý následkov porušenia ustanovenia § 11 a nasledujúcich autorského zákona Českej republiky č. 121/2000 Sb., o práve autorskom, o právach súvisiacich s právom autorským a o zmene niektorých zákonov (autorský zákon), v znení neskorších predpisov, vrátane možných trestnoprávnych dôsledkov vyplývajúcich z ustanovenia časti druhej, hlavy VI. diel 4 Trestného zákonného Českej republiky č. 40/2009 Sb.

Brno
.....
podpis autora*

*Autor podpisuje iba v tlačenej verzii.

POĎAKOVANIE

Rád bych poděkoval vedoucímu diplomové práce panu Ing. Miloslavovi Richterovi, Ph.D. za odborné vedení, konzultace, trpělivost a podnětné návrhy k práci.

Obsah

Úvod	13
1 Prehľad existujúcich riešení	15
1.1 2D detekcia zmien	15
1.1.1 Diferencia obrazov	15
1.1.2 Podiel obrazov	15
1.2 3D detekcia zmien	16
1.2.1 Zhrnutie	18
2 Teoretický prehľad	23
2.1 Homogénne súradnice a transformácie	23
2.2 Model kamery	27
2.3 Kalibrácia kamery	34
2.4 Projective 3-Point	37
2.5 Relatívna orientácia	40
2.6 Fundamentálna a esenciálna matica	42
2.7 Epipolárna geometria	45
2.8 Výpočet esenciálnej matice z bodových korešpondencií	47
2.9 RANSAC	52
2.10 Triangulácia 3D bodov	53
2.11 Bundle Adjustment	57
2.12 Registrácia mračien bodov	58
2.13 Povrchové normály	60
2.14 Voxelová reprezentácia mračna bodov	61
3 Návrh scény a testovacích dát	64
3.1 Vlastnosti scény	66
4 Návrh systému detektie zmien	70
4.1 3D rekonštrukcia	70
4.1.1 Výber snímača	70
4.1.2 Kalibrácia kamery	71
4.1.3 Pohyb kamery	71
4.1.4 3D štruktúra	75
4.1.5 Odstránenie šumu z mračna bodov	77
4.2 Registrácia mračien bodov	77
4.3 Detekcia zmien	79

5 Realizácia systému detektie zmien	82
5.1 Výber snímača	82
5.2 Kalibrácia kamery	83
5.3 Ukážka riešenia	84
5.3.1 Pohyb kamery	84
5.3.2 3D štruktúra	87
5.3.3 Odstránenie šumu z mračna bodov	88
5.3.4 Registrácia mračien bodov	89
5.3.5 Detekcia zmien	91
5.3.6 Vyhodnotenie	94
Závěr	96
Literatúra	98
Zoznam príloh	103
A Vyhodnotenie testovacích trás	104
A.1 Trasa č. 95 a 99_1	104
A.2 Trasa č. 100 a 101	106
A.3 Trasa č. 102 a 103	108
A.4 Trasa č. 104 a 105	109
A.5 Trasa č. 106 a 107	111
A.6 Trasa č. 111 a 112	113
A.7 Tabuľka parametrov matice zámen	115
B Demonstračná aplikácia	116
C Obsah prílohy na CD	117
D Obsah elektronickej prílohy	119

Zoznam obrázkov

1.1	Detekcia zmeny využitím metódy späťnej projekcie snímky I_1 do 3D modelu [6]	17
1.2	Vlastnosti 3D mračien bodov a) Nerovnomernosť: zelená šípka ukazuje na hustú oblasť, červená šípka na riedku oblasť. b) Neštruktúrovanosť: vzdialenosť medzi susednými bodmi sú rôzne. c) Neusporiadanosť: množina bodov je nezávislá na poradí [3]	20
2.1	Bod (1/3,2/3,1) vyjadrený v homogénnych súradničach [14]	24
2.2	Koľaje sa pomyselne pretínajú v bode v nekonečnej vzdialosti [14] . .	25
2.3	Podobnostná transformácia	27
2.4	Dierkový model kamery [17]	28
2.5	Súradnicové systémy uvažované pri projekcii 3D bodu [18]	29
2.6	Centrálna projekcia [18]	30
2.7	Prevod medzi obrazovou rovinou a pixelovou mriežkou [18]	31
2.8	Súdkovité skreslenie v tyrkysovom štvorci	33
2.9	Kalibračný vzor šachovnice	34
2.10	Kompenzácia súdkovitého skreslenia	36
2.11	<i>Projective 3-Point</i> metóda [23]	37
2.12	Projekčné lúče ${}^k\mathbf{x}_i^s$ [23]	38
2.13	4 možné riešenia P3P [23]	39
2.14	Nestabilné riešenie P3P [23]	40
2.15	Relatívna poloha - neznáme merítko	41
2.16	Koplanarita [25]	43
2.17	Parametrizácia relatívnej orientácie [25]	45
2.18	Epipolárna geometria dvoch snímok	46
2.19	Normalizované súradnice 8-bodového algoritmu [29]	49
2.20	4 možné riešenia relatívnej orientácie [29]	50
2.21	Triangulácia [34]	54
2.22	Uhlová paralaxa [34]	56
2.23	Chyba reprojekcie [38]	57
2.24	Registrácia dvoch bodových množín [42]	58
2.25	Registrácia dvoch bodových množín [42]	59
2.26	Reprezentácia normály vo sférických súradničach [45]	61
2.27	Rôzne definície voxelového okolia [47]	62
3.1	Scéna pri pohľade zhora	64
3.2	Pridanie objektu	65
3.3	Posunutie objektov	65
3.4	Rotácia objektu	66

3.5	Zámena objektov	66
3.6	Oklúzia	67
3.7	Nesprávne určené korešpondencie kvôli vzhľadovo podobným vzorom	68
4.1	Systém 3D detekcie zmien	70
4.2	Relatívna orientácia prvých dvoch pohľadov	72
4.3	Určenie pózy zvyšných pohľadov pomocou <i>P3P</i> algoritmu	75
4.4	Postup tvorby mračna bodov	76
4.5	Návrh registrácie mračien bodov	78
4.6	Návrh detekcie zmien medzi registrovanými mračnami bodov	79
5.1	Použitá monokulárna kamera SENCOR 3CAM	82
5.2	Rôzne zachytené kalibračné vzory	83
5.3	Pôvodná testovacia trasa č. 109	84
5.4	Aktuálna testovacia trasa č. 110	84
5.5	Detekované SURF významné body na prvej snímke	85
5.6	Prvotné korešpondencie medzi prvou a druhou snímkou	85
5.7	Filtrované korešpondencie medzi prvou a druhou snímkou	86
5.8	Počiatočná dvojica klúčových snímok - snímka 1 a 32	86
5.9	Pózy klúčových snímok trasy č. 109	86
5.10	Bodové korešpondencie medzi dvojicou klúčových snímok: a) Prvotné korešpondencie, b) Epipolárne korešpondencie, c) Korešpondencie filtrované podľa pixelovej vzdialenosťi, d) Korešpondencie s dobrou uhlovou paralaxou	88
5.11	Počiatočné mračno bodov pôvodnej trasy č. 109	89
5.12	Odšumenie mračna bodov	89
5.13	Mračno bodov zbavené šumu reprezentujúce pôvodnú trasu č. 109	90
5.14	Mračno bodov zbavené šumu reprezentujúce lokálnu trasu č. 110	90
5.15	3D korešpondencie medzi mračnami bodov	90
5.16	Registrované mračná bodov trás 109 a 110	91
5.17	a) Vymedzenie vyšetrovaného priestoru, b) Voxelizácia	92
5.18	Normály pre mračno bodov aktuálnej trasy č. 110	93
5.19	a) Vybraný voxel, b) Lokálny povrch, c) Histogram orientácie daného povrchu	93
5.20	a) Zmeny v pôvodnej trase č. 109, b) Zmeny v aktuálnej trase č. 110	94
5.21	a) Matica zámen pre trasy 109 a 110, b) ROC krivka pre trasy 109 a 110	95
A.1	Pôvodná testovacia trasa č. 95	104
A.2	Aktuálna testovacia trasa č. 99_1	104
A.3	a) Pôvodná testovacia trasa č. 95, b) Aktuálna testovacia trasa č. 99_1	104
A.4	a) Zmeny v pôvodnej trase č. 95, b) Zmeny v aktuálnej trase č. 99_1	105

A.5 a) Matica zámen pre trasy 95 a 99_1, b) ROC krivka pre trasy 95 a 99_1	105
A.6 Pôvodná testovacia trasa č. 100	106
A.7 Aktuálna testovacia trasa č. 101	106
A.8 a) Pôvodná testovacia trasa č. 100, b) Aktuálna testovacia trasa č. 101	106
A.9 a) Zmeny v pôvodnej trase č. 100, b) Zmeny v aktuálnej trase č. 101 .	107
A.10 a) Matica zámen pre trasy 100 a 101, b) ROC krivka pre trasy 100 a 101	107
A.11 Pôvodná testovacia trasa č. 102	108
A.12 Aktuálna testovacia trasa č. 103	108
A.13 a) Pôvodná testovacia trasa č. 102, b) Aktuálna testovacia trasa č. 103	108
A.14 a) Zmeny v pôvodnej trase č. 102, b) Zmeny v aktuálnej trase č. 103 .	109
A.15 a) Matica zámen pre trasy 102 a 103, b) ROC krivka pre trasy 102 a 103	109
A.16 Pôvodná testovacia trasa č. 104	109
A.17 Aktuálna testovacia trasa č. 105	110
A.18 a) Pôvodná testovacia trasa č. 104, b) Aktuálna testovacia trasa č. 105	110
A.19 a) Zmeny v pôvodnej trase č. 104, b) Zmeny v aktuálnej trase č. 105 .	110
A.20 a) Matica zámen pre trasy 104 a 105, b) ROC krivka pre trasy 104 a 105	111
A.21 Pôvodná testovacia trasa č. 106	111
A.22 Aktuálna testovacia trasa č. 107	111
A.23 a) Pôvodná testovacia trasa č. 106, b) Aktuálna testovacia trasa č. 107	112
A.24 a) Zmeny v pôvodnej trase č. 106, b) Zmeny v aktuálnej trase č. 107 .	112
A.25 a) Matica zámen pre trasy 106 a 107, b) ROC krivka pre trasy 106 a 107	113
A.26 Pôvodná testovacia trasa č. 111	113
A.27 Aktuálna testovacia trasa č. 112	113
A.28 a) Pôvodná testovacia trasa č. 111, b) Aktuálna testovacia trasa č. 112	114
A.29 a) Zmeny v pôvodnej trase č. 111, b) Zmeny v aktuálnej trase č. 112 .	114
A.30 a) Matica zámen pre trasy 111 a 112, b) ROC krivka pre trasy 111 a 112	115
B.1 Aplikácia demonštrujúca vytvorené algoritmy	116

Zoznam tabuliek

5.1	Parametre kamery	82
5.2	Vyhodnotenie trasy 109 a 110	95
A.1	Tabuľka parametrov matice zámen testovacích trás	115

Úvod

Detekcia zmien spadá k úlohám počítačového videnia a jej cieľom je nájsť rozdiely v scéne, ktorá je snímaná v rôznych časoch. Detekcia zmien sa uplatňuje v rozličných aplikáciách ako napríklad monitorovanie odlesňovania, sledovanie rastu vegetácie, detekcia a monitorovanie zosuvov pôdy, detekcia nelegálnych stavieb, monitorovanie vývoja stavby budovy, dohľadové aplikácie, udržiavanie aktuálnej 3D mapy mesta apod., preto má význam sa tomuto problému venovať. Úloha detekcie zmien je teda hlavným zameraním tejto práce.

Cieľom tejto práce je nájdenie aktuálnej polohy a detekcia geometrických zmien na vopred známej trase. Geometrické zmeny predstavujú situácie, kedy sa na trase objaví nový objekt alebo zmizne pôvodný alebo dôjde k posunu či otočeniu pôvodných objektov. Táto úloha je riešená nasledujúcim spôsobom: trasa je nasnímaná kamerou a zo snímok je rekonštruovaný jej 3D model, ktorý je braný ako predloha. Následne sa nasníma a rekonštruuje menší úsek trasy, ktorý už obsahuje dané geometrické zmeny. Tento menší/lokálny 3D model predstavuje aktuálnu polohu alebo významný bod/miesto v známej trase. Cieľom je oba 3D modely medzi sebou porovnať a určiť rozdiely medzi týmito dvoma trasami. Aby to bolo možné, je treba umiestniť lokálny 3D model do globálneho 3D modelu predlohy/pôvodnej trasy - týmto je nájdená aktuálna poloha vo vzorových dátach. Po tomto kroku, kedy už sú 3D modely zarovnané, ich môžeme priamo porovnať za účelom nájdenia rozdielov. Priestor, v ktorom sa nachádzajú zarovnané 3D modely je rozdelený do rovnako veľkých častí tzv. voxelov. Tieto časti sú postupne prechádzané a pre každú z nich je rozhodnuté o tom, či predstavuje zmenu. O zmene je rozhodnuté na základe porovnania orientácie lokálnych povrchov. Vstupom sú v tejto úlohe videosekvencie dvoch trás a výstupom sú vyznačené časti v 3D modeloch oboch trás, ktoré predstavujú geometrické rozdiely medzi týmito trasami.

Zadanie je zložené z piatich bodov. V prvom bode zadania vykonám prieskum literatúry - vyhľadám, naštudujem a porovnám práce/odborné články, ktoré sa zaoberali detekciou zmeny a na základe nich vyberiem jeden vhodný prístup k riešeniu - kapitola 1. Pre vybrané riešenie naštudujem a popíšem princípy metód a algoritmov, ktoré budú využité, čomu v texte odpovedá kapitola 2. Konkrétnie sa jedná o princípy 3D rekonštrukcie (popísané v podkapitolách 2.1 až 2.11), lokalizácie 3D modelu vo vzorových dátach (podkapitola 2.12) a detekcie zmien (podkapitoly 2.13 a 2.14). V druhom bode zadania navrhnu scenáre a podobu trás, ktoré budú snímané - určím, aké bude mať scéna vlastnosti a aké budú jej limity. Tým definujem podmienky za akých bude scéna snímaná a taktiež definujem aké zmeny v nej budem hľadať. Druhému bodu zadania odpovedá v texte práce kapitola 3. Tretí bod zadania sa venuje spracovaniu dát/snímok nazbieraných z predošlého bodu a ich prevodu do

vhodnej podoby/reprezentácie - ktorou bude v tejto práci 3D model. V tomto bode sú teda navrhnuté a realizované algoritmy 3D rekonštrukcie, čomu odpovedajú kapitoly 4.1 a 5.3.1 až 5.3.3. Štvrtý bod zadania sa venuje návrhu a realizácii algoritmov, ktoré umiestnia daný lokálny 3D model do globálnej mapy (3D modelu predlohy) - nájdenie aktuálnej polohy vo vzorových dátach, čo odpovedá kapitolám 4.2 a 5.3.4 a ktoré určia zmeny medzi týmito dvoma trasami, čo odpovedá kapitolám 4.3 a 5.3.5. V piatom bode zadania sú číselne a slovne ohodnotené výstupy navrhnutého systému detekcie zmien, čomu odpovedá kapitola 5.3.6.

1 Prehľad existujúcich riešení

V tejto kapitole je uvedený prehľad možných prístupov, ktorými sa môže úloha detektie zmien vyriešiť. Prístupy v riešení detektie zmien sa dajú rozdeliť podľa typu vstupných dát na 2D a 3D metódy. Oba prístupy sú v tejto kapitole popísané a porovnané a následne je vybrané jedno riešenie, ktoré sa implementuje.

1.1 2D detekcia zmien

Tento prístup využíva k detekcii zmien digitálne 2D obrazy, kde jeden je označený ako referencia, zachycujúci scénu v čase t_1 a druhý, obsahujúci zmeny a zachycujúci scénu v čase t_2 . 2D detekcia zmien zahŕňa viacero rôzne pokročilých metód, z ktorých sú, pre ukážku, vybrané a stručne popísané základné metódy založené na aritmetike jasových hodnôt pixelov.

1.1.1 Diferencia obrazov

Jednou zo základných metód je rozdiel obrazov:

$$I_d(x, y) = |I_1(x, y) - I_2(x, y)|, \quad (1.1)$$

kde I_1 a I_2 sú obrazy zachytené v časoch t_1 a t_2 , (x, y) sú súradnice pixelov a I_d je výsledný obraz, ktorý predstavuje rozdiel v jasových hodnotách medzi I_1 a I_2 . Zmena je detegovaná prahovaním obrazu $I_d(x, y)$ s často empiricky zvolenou hodnotou prahu τ :

$$T(x, y) = \begin{cases} 1, & \text{if } I_d(x, y) \geq \tau. \\ 0, & \text{inak,} \end{cases} \quad (1.2)$$

kde $T(x, y)$ je finálna binárna maska. Táto technika vyžaduje, aby boli vstupné obrazy zarovnané - odpovedajúce si pixely majú rovnaké pozície.[2]

1.1.2 Podiel obrazov

Podobne ako v predošej technike musia byť vstupné obrazy zarovnané. Potom je podiel vstupných obrazov definovaný ako:

$$I_r(x, y) = \frac{I_1(x, y)}{I_2(x, y)}, \quad (1.3)$$

kde $I_r(x, y)$ je hodnota v rozsahu $[0, \infty)$. Ak je $I_r(x, y)$ rovná 1, jasové hodnoty sú v danom mieste rovnaké, opak znamená v danom mieste zmenu.[2]

1.2 3D detekcia zmien

Metódy 3D detektie zmien rozhodujú o zmene na základe 3D vstupných dát alebo v ich kombinácii s 2D dátami. 3D dáta môžu mať rôznu podobu, napr. mračná bodov, hranové alebo ploškové (*mesh*) 3D modely, digitálny model terénu alebo RGB-D dáta, záleží od použitého snímača a od metód spracovania dát. Pred započatím detektie zmien musia byť vstupné dáta registrované resp. zosúhlásené, čo predstavuje ďalšiu problematiku, ktorá bude rozobraná kapitole 2.12. Táto kapitola je venovaná prieskumu a stručnému popisu článkov, ktoré riešili 3D detekciu zmeny.

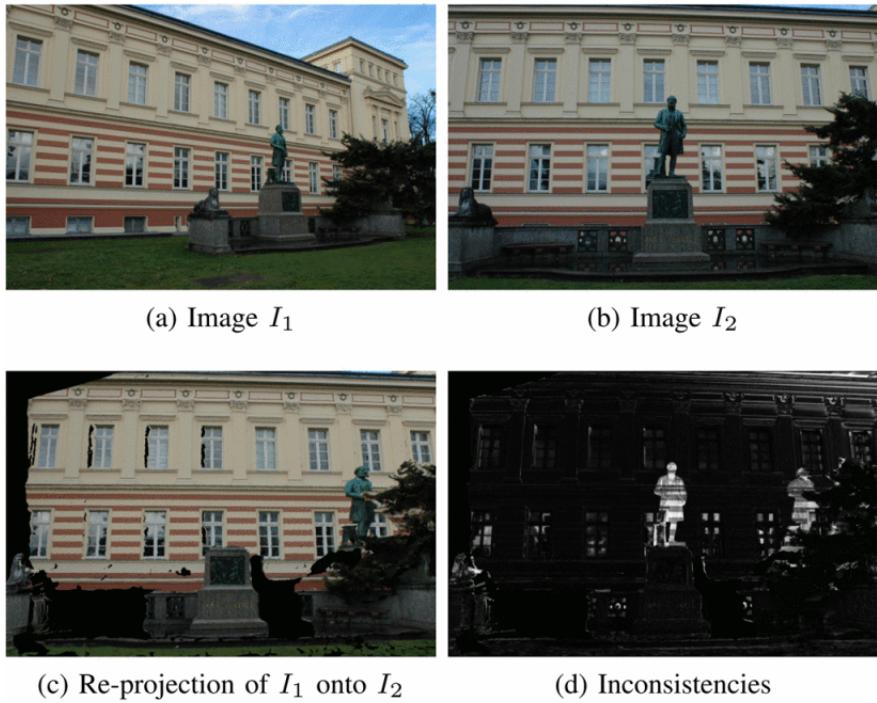
V [4] autori detegujú geometrické zmeny v mestskom prostredí medzi referenčným 3D skenom zachyteným mobilným laserom a 2D snímkami z mobilného mapovacieho systému zachytenými o nejaký čas neskôr. 2D snímky sú registrované v mračne bodov, ktoré je premietnuté na každú snímkou pomocou algoritmu z-buffer. V ďalšom kroku sú stereo páry (referenčná snímka s premietnutým 3D modelom a nasledujúca susediaca snímka) rektifikované za účelom porovnania rozdielov snímkov na pixelovej úrovni.[4]

Podobný princíp je použitý v [6], kde sa porovnáva referenčný 3D model a krátká sekvencia 2D snímkov, ktoré obsahujú zmeny. Uplatňujú tu spätnú projekciu 2D snímky I_1 do 3D modelu, resp. sa nájdú 3D body, ktoré sú pretnuté lúčmi vychádzajúcimi z projekčného stredu kamery skrz jednotlivé pixely. Tieto body sa premietnu do nového obrázku $I_{1 \rightarrow 2}$ pomocou projekčnej matice susednej snímky I_2 , čím sa dosiahne to, že $I_{1 \rightarrow 2}$ je snímaná z rovnakého uhla ako I_2 a to dovoľuje ich porovnanie na pixelovej úrovni, čo zachytáva obrázok 1.1. Zmenou je v tomto prípade socha, ktorá sa nenachádza v pôvodnom 3D modeli, ale v aktuálnych snímkach. Spätnou projekciu I_1 do 3D modelu a následnou reprojekciu do súradnicového systému I_2 , sa zmena objaví v porovnávanom obrazze dvakrát, ako je vidieť na obrázku 1.1 d). Zmeny potom môžu byť detegované pomocou absolútneho rozdielu snímkov. Autori tu ako referenciu využívajú ploškový 3D model a majú k dispozícii polohy snímkov získané fúziou systémov GPS, IMU a vizuálnej odometrie.[6]

Práca [9] je veľmi podobná predošej s tým rozdielom, že o zmene je rozhodnuté pomocou pravdepodobnostného modelu a identifikované zmeny sú zaznačené vo voxelizovanom 3D modeli. [9] [6]

[8] využíva metódu projekcie referenčného 3D modelu do porovnávaných 2D snímkov. Autori pracujú s 3D modelom vo voxelovej reprezentácii a detegujú pravdepodobnosť zmeny pomocou porovnania farby voxelu a farby pixelu, do ktorého sa daný voxel premietol. [8] [6]

V [5] autori detegujú štrukturálne zmeny vo veľkomeste medzi dvoma riedkymi mračnami bodov rekonštruovanými technikou *Structure-from-Motion* z 2D snímkov dvoch kamier, kde každá kamera je pripojená na jednej bočnej strane vozidla



Obr. 1.1: Detekcia zmeny využitím metódy späťnej projekcie snímky I_1 do 3D modelu [6]

prechádzajúceho trasou. Mračná sú registrované nájdením nerigidnej transformácie s použitím metód hlbokého učenia. Rozdiely sú detegované priamym porovnaním mračien bodov a to konkrétnie výpočtom euklidovskej vzdialenosťi každého bodu referenčného mračna bodov od jeho najbližšieho suseda z porovnávaného mračna.[5]

V [7] je prezentovaná metóda priameho porovnania dvoch mračien bodov, ktoré sú zackytené pomocou systému LIDAR. Po ich registrácii je vyšetrovaný priestor rozdelený na kubické voxely implementované pomocou oktalového stromu. Medzi referenčnými bodmi a porovnávanými bodmi, ktoré sa nachádzajú v rovnakých voxeloch je vypočítaná Hausdorffova vzdialenosť, ktorá je kritériom pre detekciu zmeny. [4] [7]

Článok [10] sa venuje monitorovaniu 3D zmien na stavenisku pri výstavbe budovy. Ako referenciu používajú autori BIM model budovy, ktorý porovnávajú s mračnom bodov reprezentujúcim rozostavanú budovu. Mračno bodov rekonštruujú z neu-sporiadanej a vopred neskalibrovanej sady snímok, ktoré sú pravidelne zhotovované. Následne je BIM model registrovaný s rekonštruovaným 3D modelom pomocou nájdienia podobnostnej transformácie. Po registrácii sú porovnávané modely diskretizované pomocou rovnostranných voxelov, ktoré sú prechádzané a označované podľa kritéria obsadenosti, viditeľnosti a oklúzie. Na záver je o zmene pre každý voxel rozhodnuté pomocou SVM klasifikátoru.[10]

V [11] autori riešia úlohu detekcie 3D objektov ako detekciu zmien medzi opakoványm navštívením rovnakej scény. Objekty definujú intuitívne ako všetky prvky v scéne, ktoré môžu byť presúvané, teda dochádza u nich k rigidným transformáciám narozenod od statického pozadia. Zmeny sú počiatočne určené z rozdielu registrovaných hlbkových máp. Následne sú v registrovaných 3D skenoch detegované geometrické 3D príznaky a ich príznakové vektory, pomocou ktorých sá nájdú bodové korešpondencie medzi dvoma 3D skenmi. Z týchto korešpondencií sa snažia určiť dominantné rigidné transformácie, ktoré predstavujú geometrické zmeny objektov.[11]

Autori práce [12] cielia na detekciu štrukturálnych zmien skrz metódu priameho porovnania dvoch mračien bodov. 3D dátá sú vyhotovené z kolekcie 2D snímok: pomocou techniky *Structure-from-Motion* je vytvorený riedky mrak bodov, ktorý je následne zhustený technikou viac-pohľadového stereja (*Multi-View Stereo*). Mračná bodov sú registrované pomocou rigidnej transformácie, vypočítanej na základe korešpondencií medzi 3D bodmi. Každý 3D bod má svoju projekciu v 2D, čo je v tomto kroku využité - korešpondencie sú určené medzi 2D bodmi pomocou ich príznakových vektorov a povýšené do 3D domény. Posledným krokom pred detekciou zmien je doladenie transformácie algoritmom *ICP*. Detekcia zmien prebieha tak, že cez ohraničený priestor zosúhlásených mračien sa posúva kubický voxel a porovnáva sa v ňom početnosť, tvar a rozmiestnenie 3D bodov. Tieto kritériá tvoria hlasovací systém, podľa ktorého sa jednotlivým voxelom prideľuje tzv. skóre zmeny, ktoré je zaznamenané v teplotnej mape. [12] [13]

1.2.1 Zhrnutie

2D detekcia zmien

Jednou z výhod 2D metód je fakt, že ide o dlhodobejšie riešený problém, ktorý je dobre preskúmaný a existuje množstvo dostupných datasetov a implementácií. Z toho plynie jednoduchšia implementácia než u 3D metód. [1] [3]

Pixelové porovnávanie obrazov vyžaduje presnú registráciu, ktorá je limitovaná, ak sú obrazy snímané z rôznych uhlov a pozícii - perspektívne skreslenie. 2D detekcia zmien je taktiež silno ovplyvnená zmenou osvetlenia a počasia. [1] [3]

Tento prístup nájde uplatnenie v aplikáciách diaľkového prieskumu, kde sa pracuje so satelitnými snímkami a na scénu sa nazerá ako na rovinu. Ďalšie uplatnenie nachádza pri použití stacionárnych kamier, kde medzi snímkami nedochádza k perspektívному skresleniu. Pri oboch aplikáciách môže stále dochádzať k chybám kvôli zmenám v osvetlení napríklad pri zmene počasia.[1] [3]

3D detekcia zmien

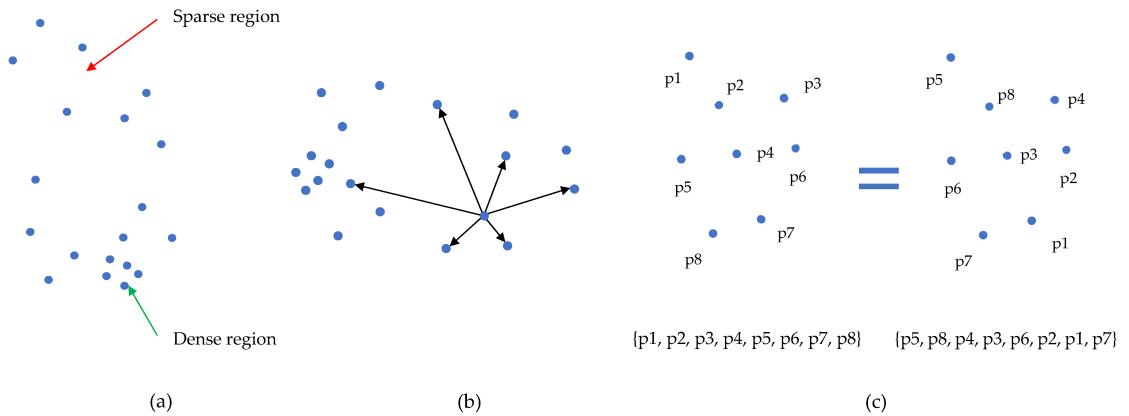
Pridaním tretieho rozmeru - hĺbky sa odkrýva znalosť priestorovej geometrie scény. Z toho plynú viaceré výhody oproti 2D detekcii zmien, kvôli ktorým sa oplatí týmto prístupom zaoberať. Týmito výhodami sú: 1) *Invariantnosť voči zmenám osvetlenia*. Ako už bolo spomenuté, mračná bodov reprezentujú priestorové merania 3D objektov, tým pádom porovnávanie geometrie 3D dát je neovplyvnené zmenami v osvetlení; 2) *Invariantnosť voči perspektívnym skresleniam*. Porovnávanie geometrie mračien bodov prebieha v trojdimenzionálnom priestore, preto nie je ovplyvnené zmenou uhlu pohľadu, ako to býva v 2D dátach, kde je efekt perspektívneho skreslenia badateľný. Mračná bodov teda nemusia byť nasnímané z rovnakého miesta; 3) *Objemová informácia*. Detekcia zmien v 3D poskytuje informáciu o volumetrických zmenách scény, čo je užitočné v aplikáciách, ktoré sledujú odlesňovanie, topografické zmeny terénu, priebeh konštrukcie na stavenisku atď. [1][3]

Pridanie ďalšieho rozmeru so sebou prináša rôzne výzvy v oblasti zberu a spracovania týchto dát. V tejto práci sú 3D dáta reprezentované v podobe 3D mračien bodov a tie, rovnako ako aj 2D dáta, majú svoje špecifické vlastnosti, s ktorými sa pri ich zbore a spracovaní musi počítať:[1][3]

1) Artefakty. Miera šumu snímača, typ snímača, vzdialenosť objektov od snímača, pozadie, počasie a algoritmy spracovania dát vnášajú do výsledného 3D modelu nepresnosti, resp. artefakty [1]. Priehľadné a „lesknúce sa“ povrchy [3] sú vlastnosti prostredia, ktoré v 3D rekonštrukcii nie sú žiadané a môžu v 3D modeli spôsobiť nepresnosti. V častiach scény bez výraznej textúry [1] nemusí byť zachytený dostačujúci počet významných bodov, čo znepresní výpočet pózy kamery a následne aj presnosť 3D modelu. Šum môže byť občas spôsobený aj prítomnosťou častíc vo vzduchu [3]. V prípade dynamickej scény sa musia uvažovať a ignorovať napr. pohybujúce sa autá, chodci, vtáctvo atď.

2) Oklúzia. Nastáva, ak sa nemenný objekt v scéne nachádza v jej prvom 3D modeli, ale v ďalšom už nie. Deje sa to vtedy, ak objekt nie je pre kameru viditeľný - napr. sa nedostane do záberu, alebo sa doň dostane, ale je zablokovaný iným nepriehľadným objektom/povrchom. Takto nenasnímaný objekt môže byť nesprávne označený ako zmena.[3] [1]

3) Nerovnomernosť. Body v mračne nie sú rovnomerne rozmiestnené, t.j. mračná bodov zachytávajúce rovnakú scénu majú rozdielnú hustotu, ako je vidieť na obrázku 1.2. Niektoré oblasti v mračne bodov zachytávajúce objekt alebo časť scény sú zastúpené bodov hustejšie alebo redšie. Je to spôsobené rôznymi faktormi ako napr. vzdialenosť objektov od snímača alebo typ použitého snímača. Nerovnomernosť môže byť do istej miery potlačená vo fáze spracovania dát pomocou techník podvzorkovania, ale nikdy nie je úplne odstránená. [3]



Obr. 1.2: Vlastnosti 3D mračien bodov a) Nerovnomernosť: zelená šípka ukazuje na hustú oblasť, červená šípka na riedku oblasť. b) Neštruktúrovanosť: vzdialenosť medzi susednými bodmi sú rôzne. c) Neusporiadanosť: množina bodov je nezávislá na poradí [3]

4) Neštruktúrovanosť. Mračná bodov, v porovnaní s 2D obrazom, nie sú vo svojej počiatočnej podobe umiestnené v pravidelnej mriežke. 3D body sú zachytené nezávisle a ich vzdialenosť od susedných bodov môže byť rôzna, čo znázorňuje obrázok 1.2. Mračná bodov sú preto v niektorých prípadoch transformované do voxelovej podoby, kde vzdialenosť medzi jednotlivými voxelmi sú vždy konštantné, v analógií s pixelmi umiestnenými v 2D mriežke. Táto technika sa používa pre uľahčenie ďalšieho spracovania 3D dát.[3]

5) Neusporiadanosť. Mračno bodov predstavuje množinu bodov v XYZ súradničiach (prípadne spolu s ďalšími atribútmi ako vektory normál, farba bodov atď.), ktoré zvyčajne zachytávajú povrch nejakého objektu. Táto množina bodov je vyjadená ako matica o rozmeroch $N \times 3$, kde N je počet 3D bodov a preusporiadáním jej riadkov, resp. zmenou poradia 3D bodov sa výsledný 3D model nezmení. Takáto množina bodov je potom nezávislá na poradí jej prvkov, čo zachytáva obrázok 1.2.[3]

Podľa preskúmanej literatúry, sa autori pri riešení 3D detekcie zmien uberali dvoma smermi: 1) *Reprojekcia 2D snímok pomocou 3D geometrie*. Ide o práce [4], [6], [9] a [8]. Ich spoločným znakom je to, že porovnávajú 3D model pôvodnej scény so sekvenciou 2D snímok, ktoré zachytávajú aktuálny stav scény. Jedna zo snímok pomocou späťnej projekcie „prejde cez“ 3D model a je následne premietnutá do súradnicového systému nasledujúcej snímky. Týmto spôsobom sú do danej snímky zaznamenané zmeny v geometrii scény a jej perspektíva je zhodná s nasledujúcou snímkou, čo umožňuje ich porovnanie po pixeloch. 2) *Priame porovnanie geometrie 3D mračien bodov*. Jedná sa o práce [5], [7], [10], [11], [12]. Autori vyšetrujú zmeny medzi pôvodným a aktuálnym stavom scény pomocou priameho porovna-

nia 3D mračien bodov, ktoré sú registrované do spoločného súradnicového systému vypočítaním vhodnej transformácie. Algoritmus detekcie zmien vo väčšine preskúmaných článkov zahŕňa techniku voxelizácie a o zmene je rozhodnuté lokálne pre každý voxel, ktorý je označený príznakom *zmena* alebo *žiadna zmena*, prípadne ďalšími. Medzi dvoma množinami bodov vo voxelu môže byť vypočítaná Hausdorffova vzdialenosť [7] alebo iná vzdialostná metrika na porovnanie podobnosti. Voxely môžu byť vyhodnocované podľa ich bodovej hustoty, obsadenosti alebo viditeľnosti. Tieto kritériá sa potom môžu využiť ako príznaky pre model strojového učenia [10]. Podobnosť množín bodov vo voxelu môže byť vyhodnotená aj pomocou porovnania orientácie povrchov, ktorý dané body tvoria. Orientácia povrchu sa získa výpočtom povrchových normál a ich ďalším spracovaním do formy, ktorá je dobre porovnatelná, napríklad histogramy.

Táto práca bude smerovať k využitiu metód 3D detekcie zmien. Tento prístup je oproti 2D metódam invariantný voči perspektívnym skresleniam a zmenám v osvetlení a takisto sa geometrické rozdiely detegujú jednoduchšie, ak je k dipozícii priestorový model scény. 3D detekcia zmien predstavuje na druhú stranu komplexnejší problém a jej implementácia je zložitejšia. Plynie to hlavne z charakteru 3D dát a z ich problematickejšieho zberu a spracovania. Metódy 3D detekcie zmien sú rozdelené do dvoch spomínaných kategórií: 1) určuje zmenu medzi aktuálnymi 2D snímkami, ktoré sú registrované podľa geometrie pôvodného 3D modelu. Registrácia musí byť pixelovo perfektná, aby sa snímky mohli správne porovnať. K tomu sa vyžaduje čo najpresnejší 3D model pôvodnej scény, ktorý musí byť v hustej alebo ploškovej reprezentácii, riedke mračno bodov na tento účel nestačí. Takisto musia byť pózy snímkov voči 3D modelu určené s čo najvyššou presnosťou. Autori v [6] majú kvalitné pózy snímkov zaistené fúziou viacerých systémov ako GPS, inerciálny senzor a vizuálna odometria. V tejto práci je 3D model scény vytvorený vlastnou implementáciou a nie je natol'ko presný, aby vyhovoval tomuto prístupu. Presnosť 3D modelu by bolo možné zvýšiť použitím nástroja alebo softvéru tretej strany pre 3D rekonštrukciu, ako napríklad *Bundler*, *VisualSfM* alebo *COLMAP*. Pózy snímkov sú v tejto práci určené iba pomocou implementovaného systému vizuálnej odometrie a sám o sebe by pravdepodobne pre tento prístup nestačil; V 2) sú 3D mračná bodov porovnávané priamo medzi sebou. Tomuto porovnaniu predchádza registrácia, ktorej cieľom je určiť transformáciu, ktorá transformuje mračná bodov do rovnakého súradnicového systému tak, aby vzdialosti medzi odpovedajúcimi si 3D bodmi mračien boli minimálne. Táto metóda nutne nevyžaduje 3D model v hustej alebo ploškovej reprezentácii, ani jeho vysokú presnosť a presnosť odometrie, ako metóda 1). Samozrejme, čím presnejší 3D model bude (s tým priamo súvisí aj presnosť pôz kamier), tým presnejšia bude aj detekcia zmien, no cieľom tejto práce nie je vytvoriť čo najrealistickejší 3D model scény. Dôležité je, aby sa pomocou neho

dalo rozhodnúť o geometrických zmenách a aby do istej miery hodnoverne zobrazoval snímanú scénu. Z týchto dôvodov je pre 3D detekciu zmien v tejto práci vybraná metóda 2).

2 Teoretický prehľad

2.1 Homogénne súradnice a transformácie

Homogénne súradnice predstavujú projektívny priestor, ktorý vznikne rozšírením euklidovského priestoru pomocou pridania ďalšieho rozmeru, ktorého súradnica je zvyčajne rovná 1. Takto rozšírený môže byť ľubovoľný n-dimenzionálny priestor; ukázaný je prípad pre bod \mathbf{x}_E v 2D:

$$\mathbf{x}_E = \begin{bmatrix} x \\ y \end{bmatrix} \Rightarrow \mathbf{x}_H = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (2.1)$$

Bod \mathbf{x}_H v homogénnych súradničiach môže byť násobený ľubovoľným skalárom $w \neq 0$:

$$\mathbf{x}_H = \begin{bmatrix} x \cdot w \\ y \cdot w \\ w \end{bmatrix} = \begin{bmatrix} u \\ v \\ w \end{bmatrix} \quad (2.2)$$

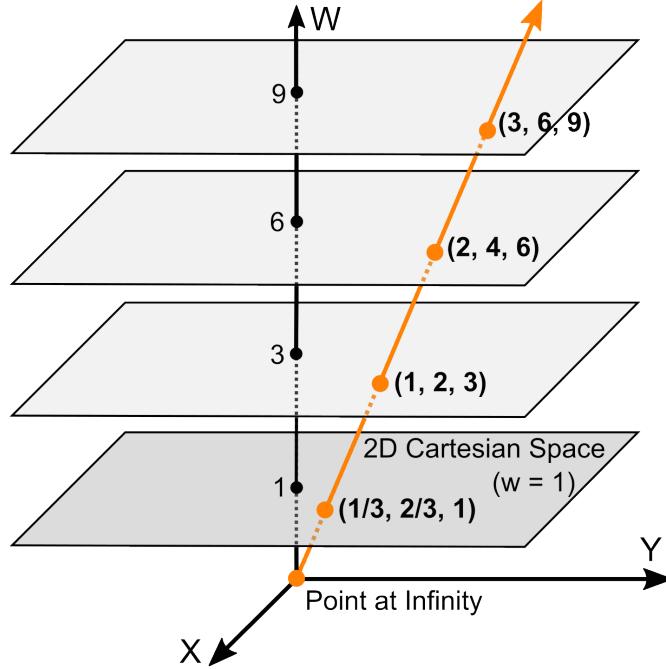
a prevedený naspäť do euklidovského priestoru normalizáciou podľa w :

$$\mathbf{x}_H = \begin{bmatrix} u/w \\ v/w \\ w/w \end{bmatrix} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \Rightarrow \mathbf{x}_E = \begin{bmatrix} x \\ y \end{bmatrix} \quad (2.3)$$

Pre rôzne hodnoty násobiaceho faktoru w sa stále jedná o rovnaký euklidovský bod \mathbf{x}_E a teda ho môžeme vyjadriť viacerými bodmi \mathbf{x}_H , ktoré ležia na spoločnej priamke, viď obrázok 2.1. [14] [15] [16]

Počiatok euklidovského súradnicového systému sa v homogénnych súradničiach nachádza v bode $(0, 0, 1)$ a každej hodnote w odpovedá pomyselná rovina (pre $w = 1$ ide o euklidovské zobrazenie). Homogénna reprezentácia euklidovského bodu je preto definovaná až na ľubovoľné skalárne merítko w a umožňuje nám vykonávať operácie, na ktoré by samotný euklidovský priestor nestačil. [15] [16]

Jednou z týchto operácií je možnosť vyjadriť body, ktoré sa nachádzajú v nekonečne. V našom prípade, projekciou 3D scény do 2D obrazovej roviny cez dierkový model kamery dochádza popri strate hĺbky aj k strate rovnobežnosti, teda rovnobežky v reálnom svete nie sú rovnobežkami v projektívnom zobrazení kamery a môžu sa pretnúť. K pretnutiu dochádza v „miznúcom bode“ (*vanishing point*), ktorý je v nekonečnej vzdialosti od kamery (viď obrázok 2.2) a v euklidovskom systéme je vyjadrený ako $[\infty, \infty]^T$, čo nie je prakticky využiteľné. V homogénnych súradničiach



Obr. 2.1: Bod $(1/3, 2/3, 1)$ vyjadrený v homogénnych súradničach [14]

je takýto bod definovaný pre $w = 0$. Ak by sa bod \mathbf{x}_H presunul do nekonečna:

$$\mathbf{x}_H = \begin{bmatrix} u \\ v \\ 0 \end{bmatrix} \quad (2.4)$$

a chceli by sme ho previesť do euklidovských súradníc, tak by sme dostali:

$$\mathbf{x}_E = \begin{bmatrix} u/w \\ v/w \end{bmatrix} = \begin{bmatrix} u/0 \\ v/0 \end{bmatrix} = \begin{bmatrix} \infty \\ \infty \end{bmatrix} \quad (2.5)$$

čo je bod v nekonečne. Homogénne súradnice umožňujú explicitne vyjadriť bod v nekonečne: u a v predstavujú lúč, na ktorom tento bod leží. [14] [15] [16]

Ďalším využitím homogénnych súradníc je vyjadrenie nelineárnych transformácií pomocou maticového násobenia. Typickou nelineárnu transformáciou je translácia \mathbf{t} , ktorá sa v euklidovskom zobrazení dá zapísat iba pomocou sčítania vektorov (bod \mathbf{x}_E je vyjadrený v 3D):

$$\mathbf{x}'_E = \mathbf{x}_E + \mathbf{t} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} = \begin{bmatrix} x + t_x \\ y + t_y \\ z + t_z \end{bmatrix} \quad (2.6)$$

Pridanie homogénnych súradníc nám umožňuje vyjadriť transláciu a všetky ostatné projektívne transformácie ako napr. rotáciu, zmenu merítka, skosenie, zrkadlové otočenie atď. maticovým zápisom a taktiež umožňuje ich reťazenie. [15] [16]



Obr. 2.2: Koľaje sa pomyselne pretínajú v bode v nekonečnej vzdialosti [14]

Matica \mathbf{H} vyjadruje všeobecnú projektívnu transformáciu, ktorá transformuje bod \mathbf{x}_H do nového bodu \mathbf{x}'_H :

$$\mathbf{x}'_H = \mathbf{H}\mathbf{x}_H \quad (2.7)$$

Transláciu teda môžeme vyjadriť pomocou matice \mathbf{H} ako

$$\mathbf{H} = w \begin{bmatrix} \mathbf{I} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix} \quad (2.8)$$

kde w je ľubovoľný skalár (vlastnosť homogénnych súradníc), \mathbf{I} je jednotková matica:

$$\mathbf{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2.9)$$

a $\mathbf{0}$:

$$\mathbf{0} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad (2.10)$$

Posledný riadok matice \mathbf{H} predstavuje pridanie homogénnej súradnice a výsledný zápis posunutia bude podľa vzťahu 2.7 pre $w = 1$ vyzerat nasledovne:

$$\mathbf{x}'_H = \mathbf{H}\mathbf{x}_H = \begin{bmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} x + t_x \\ y + t_y \\ z + t_z \\ 1 \end{bmatrix} \Rightarrow \mathbf{x}'_E = \begin{bmatrix} x + t_x \\ y + t_y \\ z + t_z \\ 1 \end{bmatrix} \quad (2.11)$$

s troma stupňami voľnosti: t_x , t_y a t_z . [15] [16]

Pre zápis rotácie bude matica \mathbf{H} vyzerat nasledovne:

$$\mathbf{H} = w \begin{bmatrix} \mathbf{R} & \mathbf{0} \\ \mathbf{0}^T & 1 \end{bmatrix} \quad (2.12)$$

\mathbf{R} je matica rotácie o rozmeroch 3×3 všeobecne vyjadrená ako

$$\mathbf{R} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \quad (2.13)$$

Dá sa tiež vyjadriť postupným otáčaním okolo osí x , y a z o uhly α , β a γ :

$$\begin{aligned} \mathbf{R}_x(\alpha) &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & \sin\alpha \\ 0 & -\sin\alpha & \cos\alpha \end{bmatrix} & \mathbf{R}_y(\beta) &= \begin{bmatrix} \cos\beta & 0 & -\sin\beta \\ 0 & 1 & 0 \\ \sin\beta & 0 & \cos\beta \end{bmatrix} \\ \mathbf{R}_z(\gamma) &= \begin{bmatrix} \cos\gamma & \sin\gamma & 0 \\ -\sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{aligned} \quad (2.14)$$

Postupným násobením týchto matíc dostaneme výslednú maticu R , ktorá vyjadruje súčasné otočenie objektu o dané uhly:

$$\mathbf{R}(\alpha, \beta, \gamma) = \mathbf{R}_z(\gamma)\mathbf{R}_y(\beta)\mathbf{R}_x(\alpha) \quad (2.15)$$

Pri skladaní transformácií závisí na ich poradí, pretože násobenie matíc nie je komutatívna operácia. [15] [16] [46]

Pre maticu \mathbf{R} dalej platí, že:

$$\det(\mathbf{R}) = 1 \quad (2.16)$$

a

$$\mathbf{R}^{-1} = \mathbf{R}^T. \quad (2.17)$$

Rotácia má v 3D priestore 3 stupne voľnosti - uhly α , β a γ . Zložením rotácie a translácie do matice \mathbf{H} :

$$\mathbf{H} = w \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix} \quad (2.18)$$

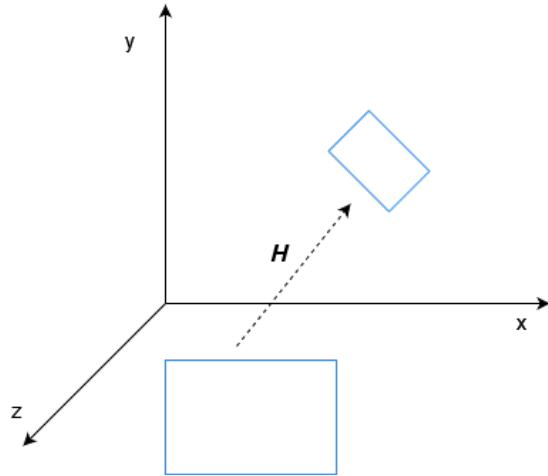
dostávame rigidnú transformáciu, ktorá zachováva linearitu (rovné čiary zostávajú rovné), rovnobežnosť (rovnobežné čiary zostávajú rovnobežné), veľkosť uhlov a dĺžky čiar a má 6 stupňov voľnosti (3 pre transláciu a 3 pre rotáciu). [15] [16]

Transformácia typu zmena merítka znamená vynásobenie bodov skalárnym faktorom s . Pre túto prácu sa uvažuje uniformná zmena merítka vo všetkých smeroch.

Pre $s > 1$ ide o zväčšenie a pre $0 < s < 1$ o zmenšenie. Pridaním skalárneho merítka do matice \mathbf{H} :

$$\mathbf{H} = w \begin{bmatrix} s\mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix} \quad (2.19)$$

dostaneme zloženie translácie, rotácie a merítka. Táto transformácia sa nazýva podobnostná a v 3D má 7 stupňov voľnosti: 3 pre transláciu, 3 pre rotáciu a 1 pre merítko. Podobnostná transformácia zachováva linearitu, paralelizmus, veľkosti uhlov, ale nezachováva pôvodné dĺžky čiar, viď obrázok 2.3. [15] [16] [46]



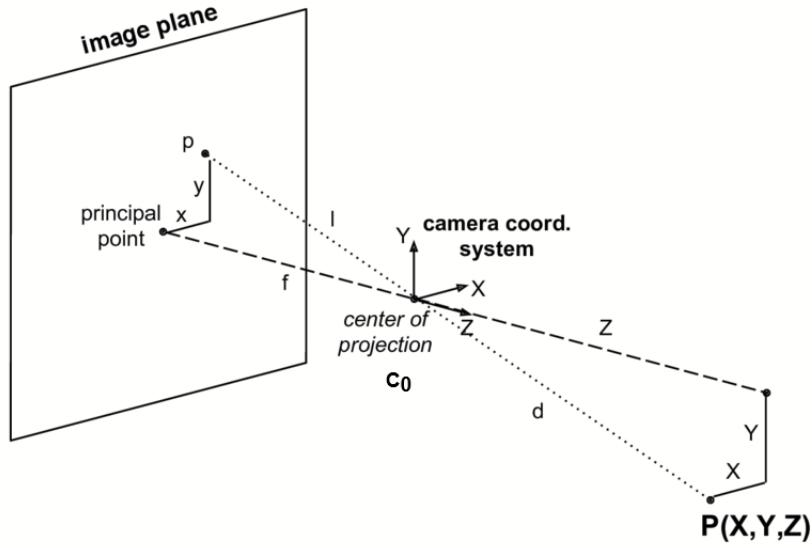
Obr. 2.3: Podobnostná transformácia

2.2 Model kamery

Ako zjednodušená reprezentácia kamery sa v tejto práci uvažuje dierkový model kamery, viď obrázok 2.4, kde všetky lúče svetla vychádzajúce z 3D bodov scény dopadajú a prechádzajú jedným bodom \mathbf{C}_O , ktorý sa označuje ako projekčný stred kamery. Za týmto bodom sa v ohniskovej vzdialosti f od bodu \mathbf{C}_O pozdĺž osi z nachádza obrazová rovina, ktorá je na osu z kolmá. Lúč vychádzajúci z 3D bodu \mathbf{P} pretína obrazovú rovinu v bode \mathbf{p} , na ktorý je bod \mathbf{P} premietnutý.

Matematický model kamery, ktorý popisuje transformáciu 3D bodu \mathbf{X} z vonkajšieho (svetového) súradnicového systému do 2D bodu \mathbf{x} v pixelových súradničiach senzoru sa vyjadri ako:

$$\mathbf{x} = \mathbf{P}\mathbf{X} \quad (2.20)$$

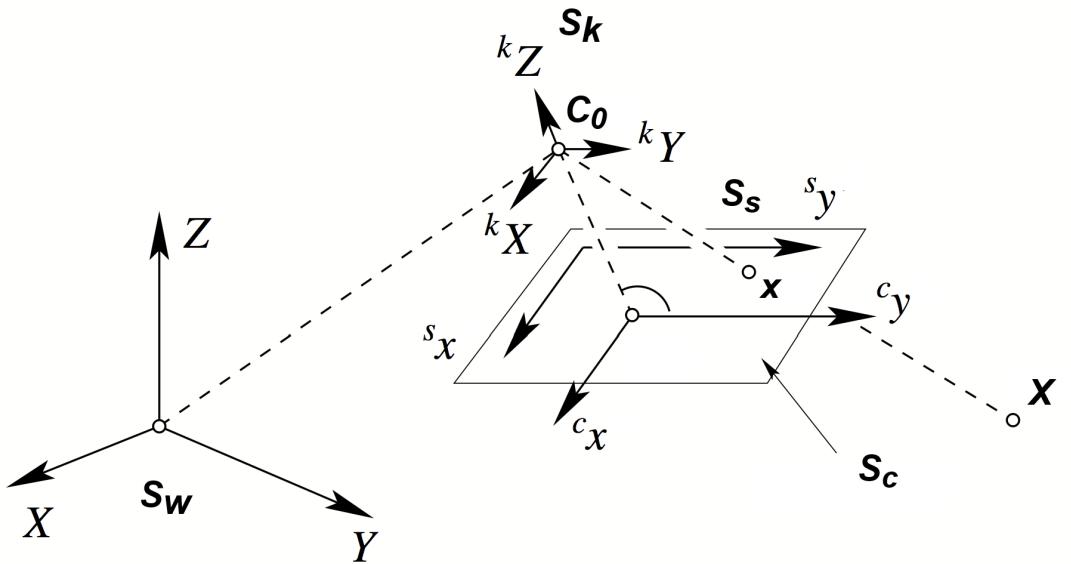


Obr. 2.4: Dierkový model kamery [17]

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \mathbf{P} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (2.21)$$

kde \mathbf{P} predstavuje projektívnu transformáciu. Matica \mathbf{P} je zložená z viacerých transformácií, pretože uvažujeme viacero súradnicových systémov, medzi ktorými musíme vstupný 3D bod postupne transformovať, kým ho dostaneme do systému pixelových súradníčkov senzora. Tieto súradnicové systémy sú 4, vidieť obrázok 2.5 (ich poradie odpovedá smeru transformácií):

- 1) **Svetový súradnicový systém.** Tiež vonkajší súradnicový systém, v ktorom sa nachádzajú všetky 3D body a má svoj vlastný počiatok. Je označený ako S_w a bod, ktorý sa v ňom nachádza ako $[X \ Y \ Z]^T$
- 2) **Súradnicový systém kamery.** Projekčný stred kamery C_O je jeho počiatkom a všetky body sú uvažované vzhľadom k nemu. Je označený ako S_k a bod, ktorý sa v ňom nachádza ako $[^kX \ ^kY \ ^kZ]^T$
- 3) **Súradnicový systém obrazovej roviny.** Je označený ako S_c a bod, ktorý sa v ňom nachádza ako $[^cX \ ^cY]^T$. Počiatok je zhodný s predošlým súradnicovým systémom, je to bod C_O , posunutý o ohniskovú vzdialenosť v smere osi Z a má o jeden rozmer menej, pretože transformáciou do tohto súradnicového systému dochádza k strate hĺbky.
- 4) **Súradnicový systém senzoru.** Je označený ako S_s a bod, ktorý sa v ňom nachádza ako $[^sX \ ^sY]^T$. Predstavuje súradnicový systém čipu, teda transformovaný



Obr. 2.5: Súradnicové systémy uvažované pri projekcii 3D bodu [18]

bod je umiestnený v pixelovej mriežke, ktorej počiatok je oproti počiatku predošlého súradnicového systému posunutý do ľavého horného rohu čipu.

Aby sa bod 3D bod \mathbf{X} postupne dostal z S_w do S_s , viď obrázok 2.5, musí prejsť troma transformáciami:

$$\begin{bmatrix} {}^s x \\ {}^s y \\ 1 \end{bmatrix} = {}^s \mathbf{H}_c {}^c \mathbf{P}_k {}^k \mathbf{H}_w \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (2.22)$$

${}^k \mathbf{H}_w$ je transformácia bodu \mathbf{X} zo svetového súradnicového systému S_w (3D) do súradnicového systému kamery S_k (3D). Na to je potrebné určiť pózu kamery v S_w , ktorá má 6 parametrov: 3 udávajú polohu jej projekčného stredu C_O a 3 jej orientáciu. Tieto parametre sú označované ako *extrinsicke* a popisujú kameru zvonku, teda kde sa vzhľadom k svetovému súradnicovému systému nachádza a ako je vzhľadom k nemu otočená. Tieto parametre tvoria rigidnú transformáciu zloženú z transformácií:

$$\mathbf{X}_O = [X_O \ Y_O \ Z_O]^T \quad (2.23)$$

ktorá vyjadruje posunutie projekčného stredu kamery (počiatok S_k) od počiatku S_w resp. polohu kamery v S_w . Matica rotácia $\mathbf{R}^{3 \times 3}$ určuje otočenie kamery z S_w do S_k . Bod \mathbf{X} sa teda do systému S_k transformuje nasledovne:

$${}^k \mathbf{X} = \mathbf{R}(\mathbf{X} - \mathbf{X}_O) \quad (2.24)$$

tým pádom bude bod ${}^k \mathbf{X}$ vyjadrený v S_k vzhľadom k počiatku, ktorý predstavuje projekčný stred kamery C_O . V homogénnych súradniciach je tento vzťah vyjadrený

ako:

$$\begin{bmatrix} {}^k \mathbf{X} \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R} & -\mathbf{R} \mathbf{X}_O \\ \mathbf{0}^T & 1 \end{bmatrix} \begin{bmatrix} \mathbf{X} \\ 1 \end{bmatrix} \quad (2.25)$$

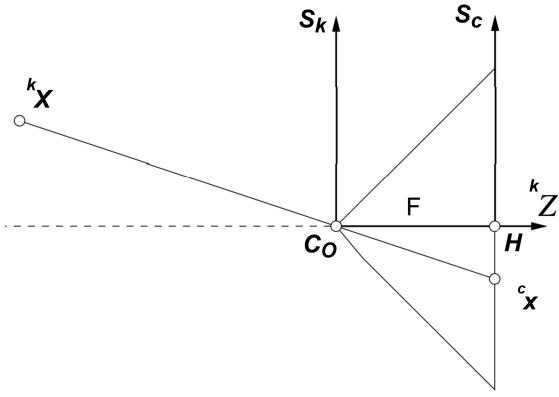
a transformácia ${}^k \mathbf{H}_w$:

$${}^k \mathbf{H}_w = \begin{bmatrix} \mathbf{R} & -\mathbf{R} \mathbf{X}_O \\ \mathbf{0}^T & 1 \end{bmatrix} \quad (2.26)$$

Nasledujúca transformácia ${}^c \mathbf{P}_k$ zo vzťahu 2.22 predstavuje transformáciu 3D bodu ${}^k \mathbf{X}$ zo súradnicového systému kamery S_k do 2D bodu ${}^c \mathbf{x}$ obrazovej roviny S_c :

$${}^c \mathbf{x} = {}^c \mathbf{P}_k {}^k \mathbf{X} \quad (2.27)$$

Táto transformácia sa týka vnútorných parametrov kamery, označených ***intrinsicské*** parametre, ktoré popisujú kameru vnútorný model kamery, ktorý je nezávislý od pózy kamery vo vonkajšom svete. Táto transformácia predstavuje ideálnu centrálnu projekciu 3D bodov do 2D roviny, kde sa uvažujú všetky lúče ako rovné čiary, ktoré prechádzajú projekčným stredom \mathbf{C}_O a premietajú sa na obrazovú rovinu S_c s počiatkom \mathbf{H} (*principal point*), ktorý leží na optickej osi (označenej ${}^k \mathbf{Z}$), vid. obrázok 2.6. Počiatok \mathbf{H} je v konštantnej vzdialosti F od \mathbf{C}_O , ktorá značí ohniskovú vzdialenosť. Je jedným z intrinsických parametrov kamery a v tejto práci je určený kalibráciou. Táto transformácia je neinvertovateľná, teda nemôžeme sa dostať späťne z 2D bodu v rovine do 3D bodu v priestore, pretože sme projekciou prišli o jeden rozmer. Čo ale poznáme je smer/lúč z akého sa 3D bod premietol. V konfigurácii



Obr. 2.6: Centrálna projekcia [18]

na obrázku 2.6 je obraz premietnutý naruby voči snímanej scéne. To sa kompenzuje otočením obrazovej roviny o 180° , tým pádom bude $F < 0$ a obrazová rovina sa bude nachádzať pred projekčným centrom \mathbf{C}_0 . Z geometrie trojuholníkov na obrázku 2.6

môžeme odvodiť polohu bodu ${}^c\mathbf{x} = [{}^c x \quad {}^c y]^T$ v obrazovej rovine:

$$\frac{{}^c x}{F} = \frac{k X}{k Z} \quad \frac{{}^c y}{F} = \frac{k Y}{k Z}$$

$${}^c x = F \frac{k X}{k Z} \quad {}^c y = F \frac{k Y}{k Z}$$

Tieto vzťahy vyjadrujme maticovo pomocou homogénnych súradníc:

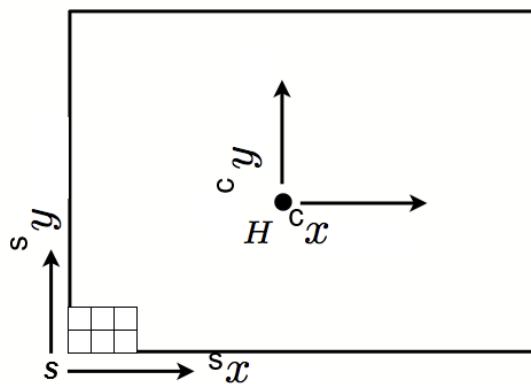
$$\begin{bmatrix} {}^c x \\ {}^c y \\ 1 \end{bmatrix} = \begin{bmatrix} F & 0 & 0 & 0 \\ 0 & F & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} {}^k X \\ {}^k Y \\ {}^k Z \\ 1 \end{bmatrix} = \begin{bmatrix} F {}^k X \\ F {}^k Y \\ {}^k Z \\ 1 \end{bmatrix} = \begin{bmatrix} F \frac{{}^k X}{k Z} \\ F \frac{{}^k Y}{k Z} \\ 1 \end{bmatrix} \quad (2.28)$$

Podľa rovnice 2.27 je matica ${}^c \mathbf{P}_k = \begin{bmatrix} F & 0 & 0 & 0 \\ 0 & F & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$ a vynechaním posledného stĺpca dostaneme kalibračnú maticu ${}^c \mathbf{K}$. [18] [19]

${}^s \mathbf{H}_c$ je poslednou transformáciou rovnice 2.22 a transformuje 2D bod ${}^c \mathbf{x}$ z obrazovej roviny S_c na 2D bod ${}^s \mathbf{x}$ umiestnený v pixelovej mriežke čipu S_s :

$${}^s \mathbf{x} = {}^s \mathbf{H}_c {}^c \mathbf{x} \quad (2.29)$$

Počiatok \mathbf{H} obrazovej roviny nie je zhodný s počiatkom \mathbf{S} pixelovej mriežky, preto sa musí určiť poloha \mathbf{H} v súradnicovom systéme čipu, aby bolo možné bod ${}^c \mathbf{x}$ správne posunúť. Situácia je zachytená na obrázku 2.7. Toto posunutie o $[x_H \quad y_H]^T$



Obr. 2.7: Prevod medzi obrazovou rovinou a pixelovou mriežkou [18]

je maticovo v homogénnych súradničiach podľa vzťahu 2.29 vyjadrené nasledovne:

$$\begin{bmatrix} {}^s x \\ {}^s y \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & x_H \\ 0 & 1 & y_H \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} {}^c x \\ {}^c y \\ 1 \end{bmatrix} = \begin{bmatrix} {}^c x + x_H \\ {}^c y + y_H \\ 1 \end{bmatrix} \quad (2.30)$$

Súradnice x_H a y_H vyjadrujú polohu stredu obrazovej roviny v pixelových súradničach. Zložením matíc ${}^c\mathbf{K}$ a ${}^s\mathbf{H}_c$ vznikne matica kalibrácie $\mathbf{K} = \begin{bmatrix} F & 0 & x_H \\ 0 & F & y_H \\ 0 & 0 & 1 \end{bmatrix}$. [18]

Výsledná zreteľaná transformácia, podľa rovnice 2.22 bude vyzerat nasledovne:

$$\begin{bmatrix} {}^s x \\ {}^s y \\ 1 \end{bmatrix} = \mathbf{K} \begin{bmatrix} \mathbf{R} & -\mathbf{R}\mathbf{X}_O \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (2.31)$$

$$\begin{bmatrix} {}^s x \\ {}^s y \\ 1 \end{bmatrix} = \mathbf{K}\mathbf{R} \begin{bmatrix} \mathbf{I}_{3 \times 3} & -\mathbf{X}_O \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (2.32)$$

$${}^s \mathbf{x} = \mathbf{P} \mathbf{X} \quad (2.33)$$

kde \mathbf{P} sa označuje ako projekčná matica, ktorá má v tomto prípade dokopy 9 parametrov: 6 extrinsických a 3 intrinsické.[18]

Kalibračná matica obsahuje formálne ešte parameter skosenia s a rozdiel v merítku m , ktoré sa v tejto práci neuplatňujú.

Situácia, kedy chceme invertovať projekciu danú vzťahom 2.33, teda chceme zisťiť, kde sa nachádza 2D bod ${}^s \mathbf{x}$ z pixelovej mriežky v priestore, je predmetom 3D rekonštrukcie. Inverznou transformáciou bodu ${}^s \mathbf{x}$ nie sme schopní získať polohu 3D bodu \mathbf{X} , pretože sme projekciou stratili jeden rozmer, ktorý nejde vrátiť späť (pomocou jednej snímky). Ak by sme sa o inverziu projekcie pokúsili, tak by to vyzeralo nasledovne:

$$\begin{aligned} w {}^s \mathbf{x} &= \mathbf{P} \mathbf{X} = \mathbf{K}\mathbf{R} \begin{bmatrix} \mathbf{I}_{3 \times 3} & -\mathbf{X}_O \end{bmatrix} \mathbf{X} = \begin{bmatrix} \mathbf{K}\mathbf{R} & -\mathbf{K}\mathbf{R}\mathbf{X}_O \end{bmatrix} \mathbf{X} = \\ &= \mathbf{K}\mathbf{R}\mathbf{X} - \mathbf{K}\mathbf{R}\mathbf{X}_O \end{aligned} \quad (2.34)$$

kde w je ľubovoľný skalár plynúci z použitia homogénnych súradníc. Začínajúc z:

$$w {}^s \mathbf{x} = \mathbf{K}\mathbf{R}\mathbf{X} - \mathbf{K}\mathbf{R}\mathbf{X}_O \quad (2.35)$$

dostávame:

$$\mathbf{X} = (\mathbf{K}\mathbf{R})^{-1} \mathbf{K}\mathbf{R}\mathbf{X}_O + w(\mathbf{K}\mathbf{R})^{-1} {}^s \mathbf{x} = \mathbf{X}_O + w(\mathbf{K}\mathbf{R})^{-1} {}^s \mathbf{x} \quad (2.36)$$

Tento vzťah sa dá interpretovať tak, že člen $(\mathbf{K}\mathbf{R})^{-1}$ určuje smer priamky vychádzajúcej z projekčného stredu kamery \mathbf{X}_O skrz bod ${}^s \mathbf{x}$ a na tejto priamke leží bod \mathbf{X} . Kde presne na priamke leží, resp. ako je daleko od kamery udáva homogenizačný faktor w a ten môže byť ľubovoľný. w predstavuje stupeň voľnosti, o ktorý

sme projekciou prišli. Znalosť priamky, na ktorej hľadaný 3D bod leží je kľúčová pri 3D rekonštrukcii. [18] [19]

Doteraz bola uvažovaná ideálna lineárna projekcia 3D bodu \mathbf{X} na 2D bod ${}^s\mathbf{x}$ v pixelovej mriežke obrázku bez nelineárnych transformácií. V reálnych prípadoch nie je poloha ${}^s\mathbf{x}$ 2D bodu konečná, ale je ďalej ovplyvnená nelineárnymi transformáciami, ktoré pramenia z rôznych nedokonalostí. Najčastejšou z nich je nedokonalosť šošovky, ktorá spôsobuje súdkovité skreslenie. Súdkovité skreslenie je nelineárna transformácia, kvôli ktorej prestáva platiť podmienka linearity - rovné čiary nezosstávajú rovné, ale oblúkovito sa zakrivujú. Tento efekt je najvýraznejší na okrajoch obrazu a znižuje sa smerom ku stredu obrazovej roviny (*principal point*), v ktorom je najmenší, viď obrázok 2.8, tyrkysový štvorec - hrana stola, ktorá je v realite rovná je v obraze zakrivená. Súdkovité skreslenie je tým väčšie, čím väčší je zorný uhol kamery. Môžeme ho vyjadriť ako posunutie bodu ${}^s\mathbf{x}$ do bodu ${}^a\mathbf{x}$ v rámci rovnakej pixelovej mriežky:

$${}^a\mathbf{x} = {}^s\mathbf{x} + \Delta\mathbf{x}(\mathbf{x}, \mathbf{q}) \quad (2.37)$$

$${}^a\mathbf{y} = {}^s\mathbf{y} + \Delta\mathbf{y}(\mathbf{y}, \mathbf{q}) \quad (2.38)$$

$${}^a\mathbf{x} = {}^s\mathbf{x}(1 + q_1r^2 + q_2r^4) \quad (2.39)$$

$${}^a\mathbf{y} = {}^s\mathbf{y}(1 + q_1r^2 + q_2r^4) \quad (2.40)$$

kde členy $\Delta\mathbf{x}(\mathbf{x}, \mathbf{q})$ a $\Delta\mathbf{y}(\mathbf{y}, \mathbf{q})$ predstavujú posunutie pixelov dané súdkovitým



Obr. 2.8: Súdkovité skreslenie v tyrkysovom štvorci

skreslením. Toto posunutie je závislé na polohe \mathbf{x} a \mathbf{y} vstupného pixelu (pre každý pixel má inú hodnotu) a na parametroch \mathbf{q} , ktorých môže byť ľubovoľný počet. Nelineárny parameter r predstavuje vzdialenosť pixelu od stredu obrazovej roviny.

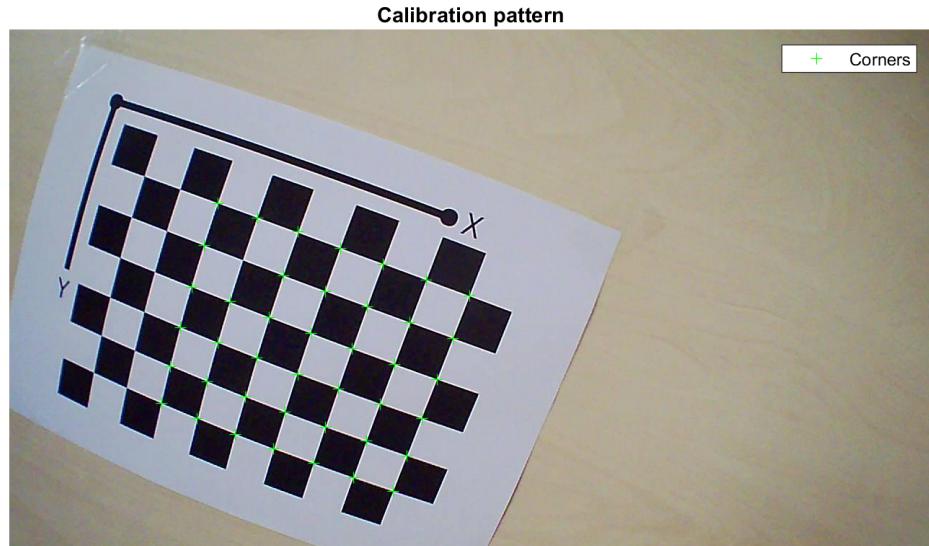
Rozšírená kalibračná matica o posunutie spôsobené súdkovitým skreslením je nasledovná:

$$\mathbf{K}(\mathbf{x}, \mathbf{q}) = \begin{bmatrix} F & 0 & x_H + \Delta x(\mathbf{x}, \mathbf{q}) \\ 0 & F & y_H + \Delta y(\mathbf{x}, \mathbf{q}) \\ 0 & 0 & 1 \end{bmatrix} \quad (2.41)$$

Je vidieť, že už nie je konštantná, ale závisí na polohe pixelu \mathbf{x} a parametroch \mathbf{q} modelujúcich súdkovité skreslenie. [18] [19]

2.3 Kalibrácia kamery

Kalibrácia kamery predstavuje meranie, ktorým dokážeme určiť extrinsické a intrinsicské parametre kamery. V tejto práci sa kalibráciou kamery rozumie určenie jej intrinsicských parametrov obsiahnutých v kalibračnej matici $\mathbf{K}(\mathbf{x}, \mathbf{q})$. Ide o ohniskovú vzdialenosť F , polohu stredu obrazovej roviny (*principal point*) v pixelových súradničiach $[x_H \ y_H]^T$ a koeficienty $[q_1 \ \dots \ q_N]$ modelujúce súdkovité skreslenie. Kalibrácia kamery je založená na Zhangovej metóde [20], kde sa ako kalibračný vzor využíva šachovnica umiestnená na rovnom povrchu, viď obrázok 5.2. Skladá sa z dvoch krokov: v prvom kroku sa určujú lineárne intrinsicské parametre a v druhom kroku sa určujú nelineárne parametre, týkajúce sa súdkovitého skreslenia. Táto me-



Obr. 2.9: Kalibračný vzor šachovnice

toda predpokladá, že detekované body na šachovnici, viď obrázok 5.2 ležia v XY rovine, kde zložka Z je na ňu kolmá a rovná 0. Pre každý obrázok so šachovnicou je

definovaný globálny (svetový) súradnicový systém s počiatkom v ľavom hornom rohu šachovnice a zo znalosti rozmeru políčok dokážeme určiť XY súradnice detekovaných bodov v súradnicovom systéme šachovnice. [21] [22]

Pri určovaní intrinsických parametrov využijeme rovnici projekcie 3D bodu \mathbf{X} na šachovnici v jej svetovom súradnicovom systéme do 2D bodu \mathbf{x} v súradnicovom systéme obrazu:

$$\mathbf{x} = \mathbf{P}\mathbf{X} \quad (2.42)$$

$$\mathbf{x} = \mathbf{K}\mathbf{R} \left[\mathbf{I}_{3 \times 3} \mid -\mathbf{X}_O \right] \mathbf{X} \quad (2.43)$$

Body \mathbf{x} a \mathbf{X} sú známe z pozorovania a matica \mathbf{P} je neznáma - pre náš prípad sledujeme časť \mathbf{K} . Vzťah 2.43 môžeme prepísat do tvaru:

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} F & 0 & x_H \\ 0 & F & y_H \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (2.44)$$

v ktorom zatial neuvažujeme súdkovité skreslenie. Ak položíme $Z=0$, tak nám z extrinsickej matice $\mathbf{R} \left[\mathbf{I}_{3 \times 3} \mid -\mathbf{X}_O \right]$ vypadne stĺpec $[r_{13} \ r_{22} \ r_{33}]^T$:

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} F & 0 & x_H \\ 0 & F & y_H \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & t_1 \\ r_{21} & r_{22} & t_2 \\ r_{31} & r_{32} & t_3 \end{bmatrix} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \quad (2.45)$$

Každý bod na šachovnici generuje tento vzťah a pre všetky body zo všetkých obrázkov šachovnice je kalibračná matica \mathbf{K} rovnaká. Matica extrinsickej parametrov $\mathbf{R} \left[\mathbf{I}_{3 \times 3} \mid -\mathbf{X}_O \right]$ je pre každý obrázok iná, lebo šachovnica môže byť snímaná z inej pozície alebo uhla, tým pádom sa zmení jej svetový súradnicový systém. Pre N pozorovaných bodov na šachovnici (v jednom obrázku) upravíme vzťah vyššie do tvaru:

$$\begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} = \mathbf{H}^{3 \times 3} \begin{bmatrix} X_i \\ Y_i \\ 1 \end{bmatrix} \quad i = 1, \dots, N \quad (2.46)$$

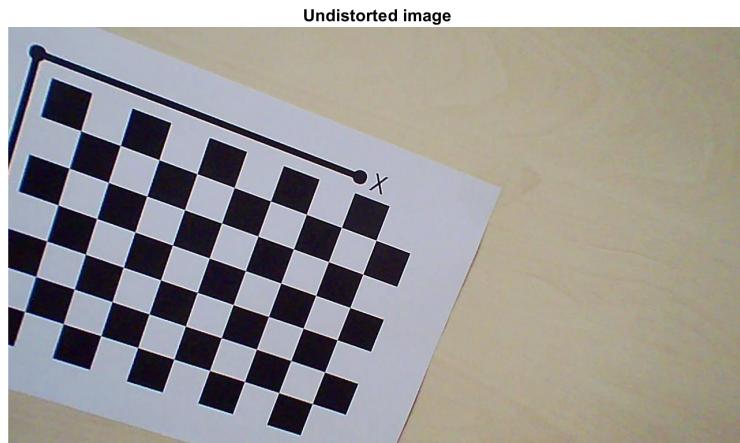
kde matica $\mathbf{H}^{3 \times 3}$ je zložená z kalibračnej matice a matice extrinsickej parametrov. Tento vzťah vyjadruje systém lineárnych rovníc, kde matica $\mathbf{H}^{3 \times 3}$ je neznáma. Má 9 neznámych parametrov, z ktorých jeden predstavuje homogenizačný faktor, takže máme 8 stupňov volnosti. Každý pozorovaný bod na šachovnici generuje 2 rovnice (pre x a pre y) a aby sme dokázali určiť $\mathbf{H}^{3 \times 3}$ potrebujeme 8 rovníc, resp. 4 body. Vyriešením systému lineárnych rovníc dostaneme maticu $\mathbf{H}^{3 \times 3}$, ktorá je pre každý obrázok šachovnice iná, pretože súradnicový systém šachovnice závisí na polohe šachovnice v obrázku. Následne sa z matice $\mathbf{H}^{3 \times 3}$ jej rozkladom extrahuje kalibračná

matica \mathbf{K} . Toto odvodenie už v tejto práci popísané nie je, ale vedie k zisteniu, že potrebujeme minimálne 3 rôzne pohľady na šachovnicu, aby sme odvodili \mathbf{K} . [21] [22]

Po prvom kroku kalibrácie, v ktorom sme určili lineárne parametre kamery pre ideálny prípad bez skreslenia, nasleduje určenie nelineárnych parametrov, v tomto prípade $\mathbf{q} = [q_1 \dots q_N]$, ktoré modelujú súdkovité skreslenie. Jeho efekt je vyjadrený v rovniciach 2.39 a 2.40: prejaví sa posunom pôvodných pixelov na nové miesta v obrazu. Tieto nelineárne parametre sú určené iteratívne metódou najmenších štvorcov resp. minimalizáciou výrazu:

$$\min_{(\mathbf{K}, \mathbf{q}, \mathbf{R}_n, \mathbf{t}_n)} \sum_n \sum_i \|\mathbf{x}_{ni} - \hat{\mathbf{x}}(\mathbf{K}, \mathbf{q}, \mathbf{R}_n, \mathbf{t}_n, \mathbf{X}_{ni})\|^2 \quad (2.47)$$

kde \mathbf{x}_{ni} je pozorovaná poloha i -tého bodu v n -tom obrázku, $\hat{\mathbf{x}}$ je očakávaná poloha bodu šachovnice premietnutého do obrázku, \mathbf{X}_{ni} je poloha i -tého bodu na šachovnici v n -tom pozorovaní. Snažíme sa minimalizovať reprojekčnú chybu medzi bodmi \mathbf{x}_{ni} a $\hat{\mathbf{x}}$, nájdením optimálnych parametrov kalibračnej matice, medzi ktorými sú aj nelineárne parametre \mathbf{q} . Tento krok je v procese kalibrácie iteratívny, narozdiel od prvého kroku, kde sme určovali lineárne parametre kamery (išlo o priame riešenie). Iteratívna metóda potrebuje počiatočný odhad parametrov, ktorý je poskytnutý na základe určených intrinsických parametrov z prvého kroku kalibrácie. Určením nelineárnych parametrov súdkovitého skreslenia dostávame aktualizovanú kalibračnú maticu kamery zo vzťahu 2.41 a môžeme pomocou nej kompenzovať efekt súdkovitého skreslenia z obrázku 5.2, vidieť obrázok 2.10, na ktorom je vidieť, že rovné čiary zostali rovné. [21] [22]



Obr. 2.10: Kompenzácia súdkovitého skreslenia

2.4 Projective 3-Point

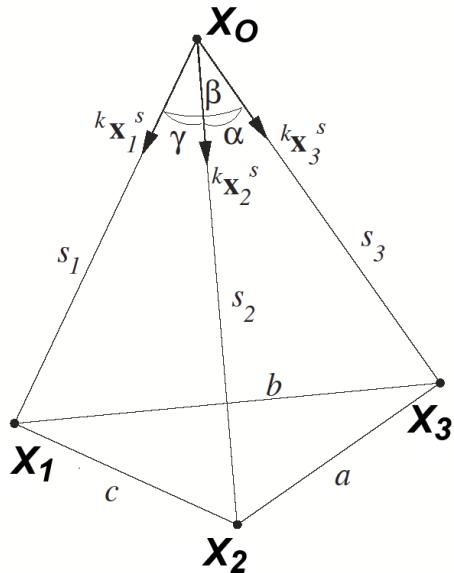
Cieľom metódy *Projective 3-Point* (*P3P*) je lokalizácia kamery na základe pozorovanej scény - ak poznáme súradnice minimálne troch 3D bodov a pozorujeme ich 2D súradnice v obrázku, tak dokážeme touto metódou určiť pózu kamery - pozíciu jej projekčného centra \mathbf{X}_O a jej orientáciu \mathbf{R} na základe iba jednej snímky. Táto metóda predpokladá skalibrovanú kameru, teda poznáme jej vnútorné parametre. [23] [24]

Cieľom je určiť, kde sa kamera nachádza, ak **poznáme** 3D pozície bodov \mathbf{X}_i v scéne, **pozorujeme** ich 2D súradnice v obrázku - body \mathbf{x}_i a **hľadáme** vonkajšie parametre kamery \mathbf{X}_O a \mathbf{R} . Taktiež poznáme korešpondencie i medzi bodmi - vieme ktorý bod \mathbf{X}_i odpovedá ktorému bodu \mathbf{x}_i v obrázku, ktorého pózu chceme určiť. Neznámych parametrov je 6 (3 pre \mathbf{X}_O a 3 pre \mathbf{R}) a na ich určenie potrebujeme vedieť súradnice aspoň 3 bodov. Ide o priame riešenie, teda pózu kamery sme schopní určiť v jednom kroku bez počiatočného odhadu. Metóda pozostáva z dvoch krov:

1) Určenie dĺžky projekčných lúčov

2) Určenie pózy

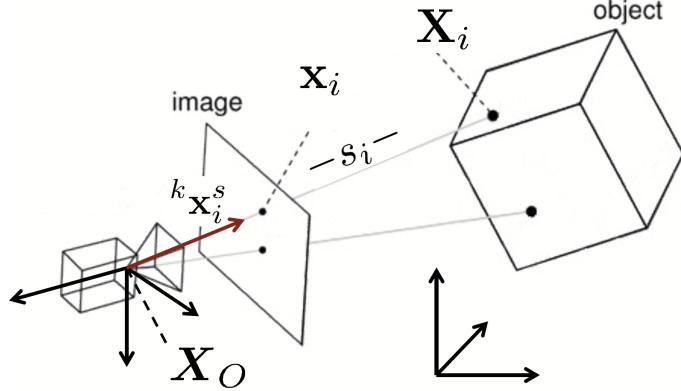
a je zobrazená na obrázku 2.11: [23] [24]



Obr. 2.11: *Projective 3-Point* metóda [23]

V prvom kroku musíme určiť dĺžky projekčných lúčov s_1 , s_2 a s_3 . Vektory ${}^k\mathbf{x}_i^s$ predstavujú smerové vektory projekčných lúčov vyjadrené v súradnicovom systéme kamery, na ktorých ležia 3D body \mathbf{X}_i , ktoré sú zobrazené ako 2D body \mathbf{x}_i v obrázku. Dĺžky s_i sú neznáme, projekciou sme túto informáciu stratili a zo samotnej snímky

nejde bez ďalších informácií odvodiť. ${}^k\mathbf{x}_i^s$ vychádzajú z projekčného centra kamery \mathbf{X}_O , prechádzajú cez 2D body \mathbf{x}_i a končia v 3D bodoch \mathbf{X}_i , vid obrázok 2.12. Tým,



Obr. 2.12: Projekčné lúče ${}^k\mathbf{x}_i^s$ [23]

že poznáme 2D body \mathbf{x}_i v obrázku a kalibračnú maticu \mathbf{K} , sme schopní smerové vektoru ${}^k\mathbf{x}_i^s$ vyjadriť ako:

$${}^k\mathbf{x}_i^s = -sign(F)N(\mathbf{K}^{-1}\mathbf{x}_i) \quad (2.48)$$

$-sign(F)$ otáča smer ${}^k\mathbf{x}_i^s$ (pri projekcii uvažujeme ohniskovú vzdialenosť F zápornú, aby sa nám obrázok nezobrazil naruby, čo nám umiestni obrazovú rovinu pred \mathbf{X}_O , ktorá v skutočnosti leží za \mathbf{X}_O). N je normalizačný faktor, ktorý zaistí, že ${}^k\mathbf{x}_i^s$ má jednotkovú veľkosť, aby sme po jeho vynásobení vzdialenosťou s_i dostali polohu bodu \mathbf{X}_i . Zo znalosti vektorov ${}^k\mathbf{x}_i^s$ môžeme priamo vypočítať uhly α, β, γ medzi nimi, vid. obrázok 2.11:

$$\alpha = arccos\left(\frac{{}^k\mathbf{x}_2^s \cdot {}^k\mathbf{x}_3^s}{\| {}^k\mathbf{x}_2^s \| \| {}^k\mathbf{x}_3^s \|}\right) \quad (2.49)$$

$$\beta = arccos\left(\frac{{}^k\mathbf{x}_1^s \cdot {}^k\mathbf{x}_3^s}{\| {}^k\mathbf{x}_1^s \| \| {}^k\mathbf{x}_3^s \|}\right) \quad (2.50)$$

$$\gamma = arccos\left(\frac{{}^k\mathbf{x}_1^s \cdot {}^k\mathbf{x}_2^s}{\| {}^k\mathbf{x}_1^s \| \| {}^k\mathbf{x}_2^s \|}\right) \quad (2.51)$$

Ak by sme nemali skalibrovanú kameru, tieto uhly by nebolo možné vypočítať. Zo znalosti bodov $\mathbf{X}_1, \mathbf{X}_2$ a \mathbf{X}_3 môžeme vypočítať dlžky čiar a, b a c viď obrázok 2.11:

$$a = \|\mathbf{X}_3 - \mathbf{X}_2\| \quad (2.52)$$

$$b = \|\mathbf{X}_3 - \mathbf{X}_1\| \quad (2.53)$$

$$c = \|\mathbf{X}_2 - \mathbf{X}_1\| \quad (2.54)$$

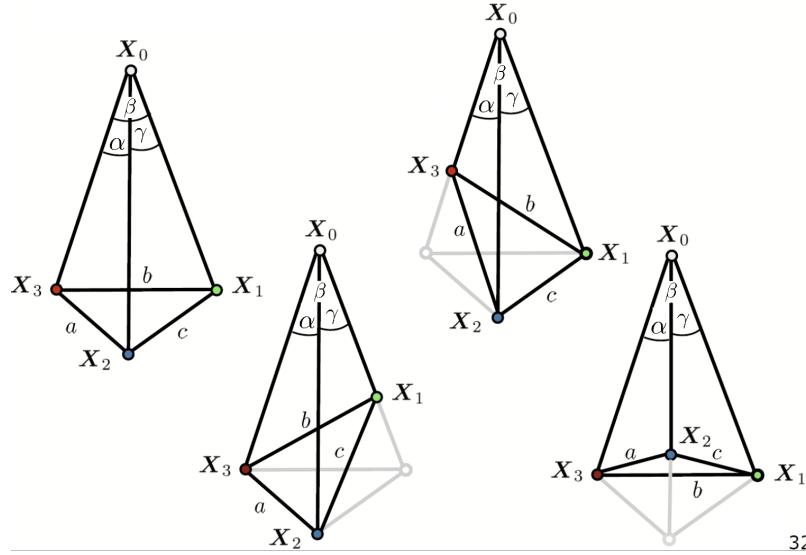
V obrázku 2.11 máme 3 trojuholníky: $\triangle \mathbf{X}_O \mathbf{X}_2 \mathbf{X}_3$, $\triangle \mathbf{X}_O \mathbf{X}_1 \mathbf{X}_3$, $\triangle \mathbf{X}_O \mathbf{X}_1 \mathbf{X}_2$. V týchto trojuholníkoch môžeme pomocou kosínusovej vety vyjadriť neznáme vzdialenosťi s_1, s_2 a s_3 :

$$a^2 = s_2^2 + s_3^2 - 2s_2s_3\cos\alpha \quad (2.55)$$

$$b^2 = s_1^2 + s_3^2 - 2s_1s_3\cos\beta \quad (2.56)$$

$$c^2 = s_1^2 + s_2^2 - 2s_1s_2\cos\gamma \quad (2.57)$$

Dostávame systém 3 rovníc o 3 neznámych s_1, s_2 a s_3 , ktoré sa nachádzajú v kvadraticej a nekvadratickej podobe. Riešenie tohto systému rovníc viedie na 4 možné varianty riešení: $\{s_1, s_2, s_3\}_{1\dots 4}$, kde každé riešenie má rovnaké dĺžky strán a, b, c a rovnaké uhly α, β, γ , no jedná sa o rozdielne geometrické konfigurácie, viď. obrázok 2.13, kde práve jedno riešenie (vľavo) je správne. Variantné riešenie dostaneme rotá-



Obr. 2.13: 4 možné riešenia P3P [23]

ciu trojuholníka $\triangle abc$ okolo jednej z jeho osí, čo neovplyvní uhly α, β, γ , ani dĺžky jeho strán a, b, c , no zmenia sa vzdialenosťi s_1, s_2 a s_3 . Tento problém sa dá vyriešiť pridaním ďalšieho kontrolného bodu \mathbf{X}_4 , pomocou ktorého sa vyberie riešenie, ktoré je s \mathbf{X}_4 kompatibilné. [23] [24]

V druhom kroku, kedy už máme vypočítané smerové vektory ${}^k\mathbf{x}_i^s$ a ich dĺžky s_i , môžeme pomocou nich vypočítať, kde sa budú nachádzať 3D body ${}^k\mathbf{X}_i$ v súradnicovom systéme kamery S_k :

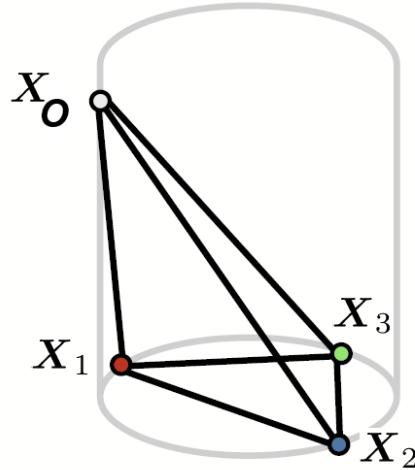
$${}^k\mathbf{X}_i = {}^k\mathbf{x}_i^s s_i \quad (2.58)$$

Potom určíme rigidnú transformáciu medzi súradnicovým systémom kamery a svetovým súradnicovým systémom pomocou bodov ${}^k\mathbf{X}_i$ a \mathbf{X}_i :

$${}^k\mathbf{X}_i = \mathbf{R}(\mathbf{X}_i - \mathbf{X}_O) \quad (2.59)$$

kde \mathbf{R} a \mathbf{X}_O je hľadaná póza. [23] [24]

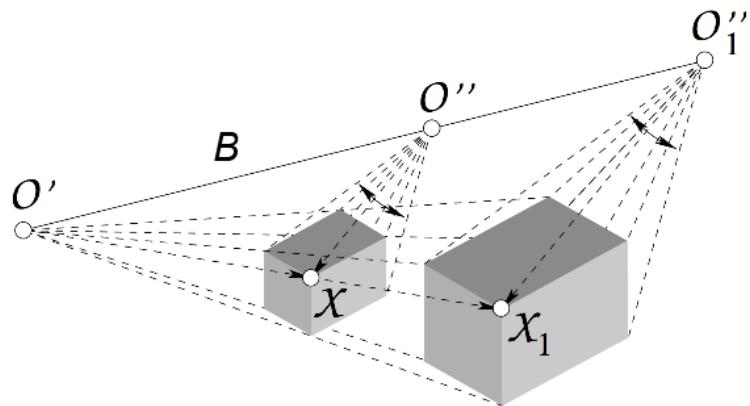
Existujú určité prípady kedy môže metóda P3P zlyhať. Ide o geometrickú konfiguráciu, kde projekčné centrum kamery \mathbf{X}_O leží na povrchu tubusu valca, s bodmi \mathbf{X}_i ležiacimi na jeho základni, viď obrázok 2.14. Malé zmeny uhlov spôsobia veľké zmeny v súradniciach, čo vedie na nestabilné riešenie. [23] [24]



Obr. 2.14: Nestabilné riešenie P3P [23]

2.5 Relatívna orientácia

V predošej kapitole sme určovali orientáciu/pózu kamery na základe jednej snímky a znalosti pozície 3D bodov. V tejto kapitole budeme určovať orientáciu dvoch kamer/pohľadov iba pomocou dvojice 2D snímok. Táto konfigurácia odpovedá prípadu, kedy máme jednu kameru, ktorá sa pohybuje scénou a zaznamená ju z dvoch rôznych bodov. Cieľom je určiť orientáciu prvého pohľadu a relatívnu orientáciu druhého pohľadu voči prvému, iba na základe dvoch 2D snímok (na ktorých detektujeme 2D body a poznáme ich korešpondencie), bez ďalších informácií o scéne s tým, že máme kameru skalibrovanú. Orientácia kamery je teda vyjadrená jej extrinsickými parametrami: 1) 3 parametre pre natočenie, resp. matica rotácie $\mathbf{R}^{3 \times 3}$ (kam sa kamera pozerá) a 2) 3 parametre pre pozíciu jej projekčného stredu \mathbf{X}_O . Pre dva pohľady je to spolu 12 neznámych parametrov, ktoré chceme určiť. Otázkou je, kolko z týchto parametrov je možné určiť iba na základe dvoch obrázkov. Platí, že z jedného obrázku nie sme schopní (bez ďalších informácií o scéne) určiť skutočnú veľkosť objektu, pretože nepoznáme ako ďaleko je objekt od kamery. Merítko (*scale*) scény teda ostáva nedefinované. Podobná úvaha platí aj pre prípad dvoch pohľadov - sú isté parametre, ktoré nie je možné odvodiť iba na základe 2D snímok. [25] [26]



Obr. 2.15: Relatívna poloha - neznáme merítko

Na obrázku 2.15 je znázornená dvojica pohľadov \mathbf{O}' a \mathbf{O}'' skalibrovanej kamery, ktoré zachytávajú 3D bod \mathbf{X} v scéne, o ktorej nemáme žiadne informácie. Počiatok súradnicového systému je umiestnený v bode \mathbf{O}' , teda všetko je vyjadrené vzhľadom k prvej kamere. Ak by sme druhý pohľad/kameru \mathbf{O}'' posunuli po spojnici \mathbf{B} do bodu \mathbf{O}_1'' a zväčšili by sme objekt v scéne, pohľad \mathbf{O}_1'' by generoval rovnaký obrázok ako pohľad \mathbf{O}'' . Tým, že nepoznáme skutočnú veľkosť snímaného objektu (alebo iné externé informácie), nedokážeme určiť veľkosť vektoru \mathbf{B} (merítko - 1 parameter), teda nevieme, kde na priamke \mathbf{B} leží druhý pohľad \mathbf{O}'' . Nedostaneme teda plnú transláciu medzi týmito dvoma pohľadmi, no poznáme smer, daný vektorom \mathbf{B} , na ktorom leží druhý pohľad. Tým, že sme pohľad \mathbf{O}' zaviedli ako počiatok súradnicového systému, nepoznáme jeho globálnu polohu v rámci svetového súradnicového systému - tým sme prišli o znalosť rotácie a posunutia (6 parametrov) kamery \mathbf{O}' . Z dvoch snímok pre skalibrovaný prípad teda nevieme určiť veľkosť vektoru ich relatívneho posunutia (merítko - 1 parameter) a globálnej rotáciu a transláciu prvej kamery, ktorá predstavuje počiatok zavedeného súradnicového systému (6 parametrov). Týchto 7 parametrov predstavuje podobnostnú transformáciu. Z pôvodných 12 parametrov vieme vypočítať 5 parametrov, ktoré majú relatívnu povahu: 1) Rotácia druhej kamery voči prvej (3 parametre), 2) Smerový vektor \mathbf{B} spájajúci projekčné stredy dvoch kamier (2 parametre) - jeho veľkosť je neznáma. Tieto parametre nám udávajú relatívnu polohu/orientáciu druhej kamery voči prvej, ktorá je umiestnená v počiatku súradnicového systému. [25] [26]

Z dvoch pohľadov v tejto konfigurácii môžeme rekonštruovať takzvaný fotogrammetrický 3D model scény, ktorý je definovaný až na spomínanú podobnostnú transformáciu (rotácia, translácia a merítko). Nepoznáme teda jeho globálnu rotáciu a transláciu vo svetovom/externom súradnicovom systéme a nepoznáme veľkosť objektov v 3D modeli. Ak ho chceme registrovať alebo umiestniť do externého súrad-

nicového systému (napríklad do väčšieho 3D modelu scény) musíme nájsť parametre tejto podobnostnej transformácie pomocou minimálne troch 3D bodov. Ide o úlohu, v ktorej chceme určiť absolútну orientáciu 3D modelu scény - *absolute orientation problem*, ktorá je rozobraná v kapitole 2.12. [25] [26]

Metóda relatívnej orientácie má tú výhodu, že dokážeme do určitej miery vypočítať orientáciu kamery iba na základe dvoch snímok, bez žiadnej dodatočnej informácie o scéne. Musí sa ale brať do úvahy to, že nepoznáme absolútnu veľkosť posunutia medzi dvoma pohľadmi, ale iba smer priamky, na ktorej druhý pohľad leží. *Projective 3-Point* metóda na druhej strane síce vyžaduje informáciu o scéne (pozície 3D bodov), ale takto vypočítaná póza, na základe 3D bodov v scéne, má definovanú veľkosť vektoru translácie (merítka), narozenie od prípadu relatívnej orientácie. V tejto práci sú skombinované obe metódy v časti 3D rekonštrukcie, ktorá sa zaoberá výpočtom pohybu kamery. Prvý pohľad je umiestnený do počiatku s.s. a relatívna poloha druhého pohľadu je vypočítaná z esenciálnej matice. Z prvých dvoch pohľadov sú triangulované 3D body, ktoré slúžia na výpočet pózy tretieho pohľadu pomocou algoritmu *Projective 3-Point*. Póza každej ďalšej pridanej snímky je určená pomocou P3P algoritmu ná základe 3D bodov triangulovaných z predošej dvojice snímok. [25] [26]

2.6 Fundamentálna a esenciálna matica

Relatívna orientácia kamery z predošej kapitoly sa vypočíta z fundamentálnej matice (neskalibrovaná kamera, všeobecný prípad) alebo z esenciálnej matice (skalibrovaná kamera, špeciálny prípad fundamentálnej matice). Tieto matice popisujú geometriu medzi dvoch pohľadov a sú v nich zakódované parametre relatívnej orientácie, ku ktorým sa chceme dostať. Fundamentálna matica pracuje s neskalibrovanou kamerou, je to všeobecnejšia reprezentácia geometrie dvoch pohľadov a je z nej odvodena esenciálna matica, ktorá sa používa ak máme skalibrovanú kameru. [25] [26]

Odvodenie fundamentálnej matice vychádza z podmienky koplanarity pre kamery, ktoré zachovávajú podmienku linearity (rovné čiary sú zobrazené ako rovné): Na obrázku 2.16 sú zobrazené dve kamery/pohľady s projekčnými stredmi \mathbf{O}' a \mathbf{O}'' a 3D bodom \mathbf{X} . Predpokladáme, že projekčné lúče resp. smerové vektory $\mathbf{O}'\mathbf{X}'$ a $\mathbf{O}''\mathbf{X}''$ sa pretínajú v bode \mathbf{X} . Ďalej uvažujeme, že vektory $\mathbf{O}'\mathbf{X}'$, $\mathbf{O}''\mathbf{X}''$ a $\mathbf{O}'\mathbf{O}''$ ležia v jednej rovine, čo je vyjadrené vzťahom:

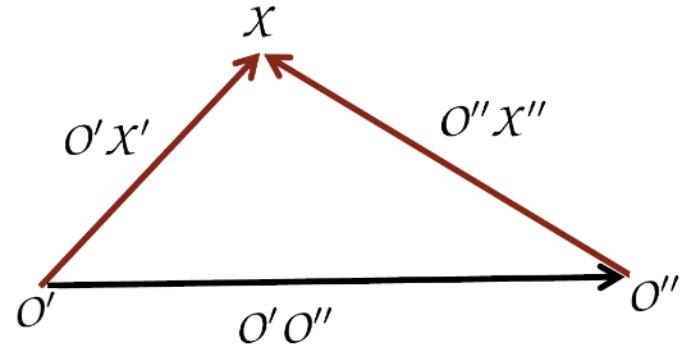
$$[\mathbf{O}'\mathbf{X}' \ \mathbf{O}'\mathbf{O}'' \ \mathbf{O}''\mathbf{X}''] = 0 \quad (2.60)$$

kde \mathbf{X}' a \mathbf{X}'' predstavujú rovnaký bod \mathbf{X} v priestore. Vzťah 2.60 používa operátor $[\cdot, \cdot, \cdot]$, ktorý predstavuje trojity súčin vektorov \mathbf{A}, \mathbf{B} a \mathbf{C} ako skalárny súčin jedného

z vektorov s vektorovým súčinom zvyšných dvoch:

$$[\mathbf{A}, \mathbf{B}, \mathbf{C}] = (\mathbf{A} \times \mathbf{B}) \cdot \mathbf{C} \quad (2.61)$$

Ak je $[\mathbf{A}, \mathbf{B}, \mathbf{C}] = 0$, znamená to, že vektory ležia v jednej rovine. Smerové vek-



Obr. 2.16: Koplanarita [25]

tory $\mathbf{O}'\mathbf{X}'$ a $\mathbf{O}''\mathbf{X}''$, môžeme odvodiť zo znalosti 2D bodov \mathbf{x}' a \mathbf{x}'' v obrazových súradničach, ktoré pozorujú 3D bod \mathbf{X} (inverzia projekčnej matice, vzťah 2.36):

$${}^n\mathbf{x}' = (\mathbf{R}')^{-1}(\mathbf{K}')^{-1}\mathbf{x}' \quad (2.62)$$

Vektor ${}^n\mathbf{x}'$ je vďaka členu $(\mathbf{R}')^{-1}$ vyjadrený vo svetových súradničach a udáva smer projekčného lúča z centra kamery, do 3D bodu \mathbf{X} . Analogicky vyjadrieme aj vektor ${}^n\mathbf{x}''$ druhého pohľadu. Bázový vektor $\mathbf{O}'\mathbf{O}''$ plynie priamo zo súradníč projekčných stredov $\mathbf{X}_{O'}$ a $\mathbf{X}_{O''}$:

$$\mathbf{b} = \mathbf{X}_{O'} - \mathbf{X}_{O''} \quad (2.63)$$

Prepísaním vzťahu 2.60 dostávame:

$$\begin{aligned} \begin{bmatrix} {}^n\mathbf{x}' & \mathbf{b} & {}^n\mathbf{x}'' \end{bmatrix} &= 0 \\ {}^n\mathbf{x}' \cdot (\mathbf{b} \times {}^n\mathbf{x}'') &= 0 \\ {}^n\mathbf{x}'^T \mathbf{S}_b {}^n\mathbf{x}'' &= 0 \end{aligned} \quad (2.64)$$

\mathbf{S}_b je antisymetrická matica:

$$\mathbf{S}_b = \begin{bmatrix} 0 & -b_3 & b_2 \\ b_3 & 0 & -b_1 \\ -b_2 & b_1 & 0 \end{bmatrix} \quad (2.65)$$

Dosadením do vzťahu 2.64 za ${}^n\mathbf{x}'$ a ${}^n\mathbf{x}''$ dostaneme:

$$\mathbf{x}'^T (\mathbf{K}')^{-T} (\mathbf{R}')^{-T} \mathbf{S}_b (\mathbf{R}'')^{-1} (\mathbf{K}'')^{-1} \mathbf{x}'' = 0 \quad (2.66)$$

kde člen

$$(\mathbf{K}')^{-T}(\mathbf{R}')^{-T}\mathbf{S}_b(\mathbf{R}'')^{-1}(\mathbf{K}'')^{-1} = (\mathbf{K}')^{-T}\mathbf{R}'\mathbf{S}_b\mathbf{R}''^T(\mathbf{K}'')^{-1} \quad (2.67)$$

predstavuje fundamentálnu maticu:

$$\mathbf{F} = (\mathbf{K}')^{-T}\mathbf{R}'\mathbf{S}_b\mathbf{R}''^T(\mathbf{K}'')^{-1} \quad (2.68)$$

ktorá obsahuje všetky parametre relatívnej orientácie dvoch pohľadov, ktoré chceme získať a dá sa pomocou nej vyjadriť podmienka koplanarity ako:

$$\mathbf{x}'^T\mathbf{F}\mathbf{x}'' = 0 \quad (2.69)$$

ktorá platí pre korešpondujúce 2D body \mathbf{x}' a \mathbf{x}'' v obrazových súradničach z oboch pohľadov. [25] [26]

Esenciálna matica je špeciálnym prípadom fundamentálnej matice, kedy poznáme kalibračnú maticu \mathbf{K} . Väčšina fotogrametrických systémov používa skalibrované kamery (ak je kalibrácia možná), pretože to uľahčuje estimáciu relatívnej orientácie. V tejto práci je zavedené, že oba pohľady, resp. všetky pohľady majú rovnaké kalibračné matice a v rámci tohto odvodenia neuvažujeme skreslenia v obrazu ani iné neistoty. Pre skalibrované kamery sa podmienka koplanarity oproti prípadu neskalibrovaných kamier zjednoduší nasledovne: smerový vektor zo vzťahu 2.62 sa zjednoduší na:

$${}^k\mathbf{x}' = \mathbf{K}'^{-1}\mathbf{x}' \quad (2.70)$$

Kalibračná matica je známa, preto môžeme 2D bod \mathbf{x}' z obrazových súradníc vyjadriť ako smerový vektor ${}^k\mathbf{x}'$ v súradnicovom systéme kamery. Analogicky vyjadrimo aj vektor ${}^k\mathbf{x}''$ druhého pohľadu kamery. Ak dosadíme do vzťahu 2.66 smerové vektory ${}^k\mathbf{x}'$ a ${}^k\mathbf{x}''$ tak sa daný vzťah zjednoduší na:

$${}^k\mathbf{x}'^T\mathbf{R}'\mathbf{S}_b\mathbf{R}''^T{}^k\mathbf{x}'' = 0 \quad (2.71)$$

kde esenciálna matica je:

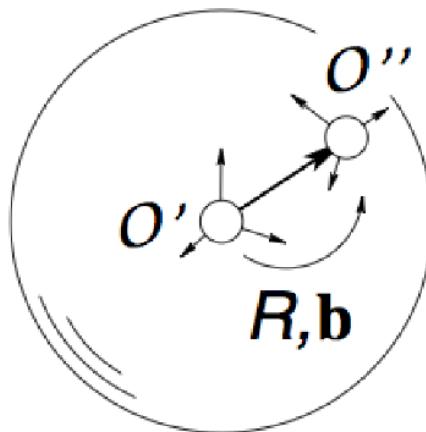
$$\mathbf{E} = \mathbf{R}'\mathbf{S}_b\mathbf{R}''^T \quad (2.72)$$

Vidíme, že obsahuje iba extrinsické parametre kamery, konkrétnie 5 hľadaných parametrov relatívnej orientácie. Podmienka koplanarity sa pomocou nej dá vyjadriť ako:

$${}^k\mathbf{x}'^T\mathbf{E}{}^k\mathbf{x}'' = 0 \quad (2.73)$$

pre korešpondujúce body ${}^k\mathbf{x}'$ a ${}^k\mathbf{x}''$. Tieto body sú vyjadrené v súradnicovom systéme kamery resp. v normalizovaných obrazových súradničach, narozené od prípadu fundamentálnej matice, kde sú korešpondujúce body v obrazových (pixelových) súradničach. [25] [26]

Esenciálna matica má teda 5 stupňov voľnosti (3 parametre pre rotáciu a 2 pre transláciu), ktoré môžeme určiť na základe relatívneho pohybu dvoch kamier/pohľadov. Existuje viacero spôsobov, akými môže byť vyjadrená relatívna orientácia dvoch pohľadov resp. 5 stupňov voľnosti esenciálnej matice. V tejto práci je zvolený spôsob na obrázku 2.17, kedy je prvý pohľad O' umiestnený do počiatku súradnicového systému, poloha jeho projekčného stredu je v bode $[0 \ 0 \ 0]^T$ a jeho matica rotácie je jednotková matica. Druhý pohľad O'' je vyjadrený relatívne voči prvému pomocou jeho matice rotácie $R = R''$, ktorá tvorí 3 stupne voľnosti esenciálnej matice a pomocou smerového vektoru b , ktorý je normalizovaný na jednotkovú veľkosť - predstavuje zvyšné 2 stupne voľnosti esenciálnej matice. [25] [26]

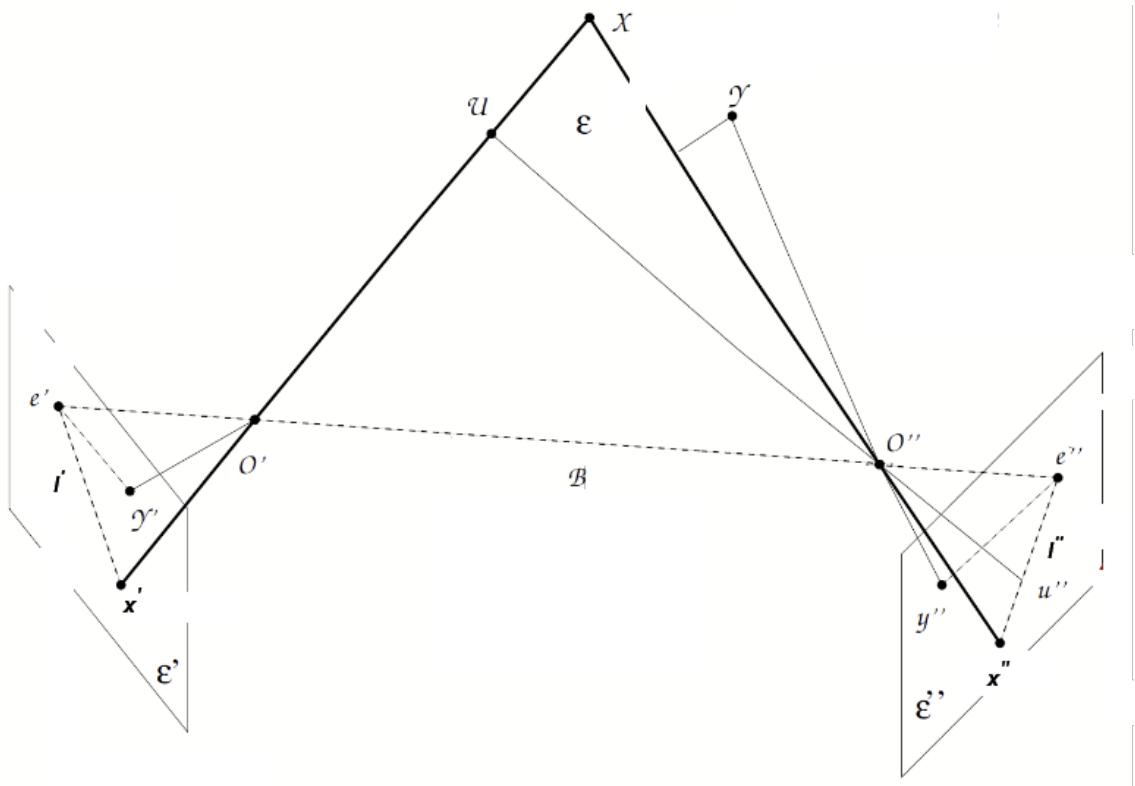


Obr. 2.17: Parametrizácia relatívnej orientácie [25]

2.7 Epipolárna geometria

Epipolárna geometria popisuje geometrické vzťahy medzi párom dvoch kamier/pohľadov, ktoré snímajú rovnakú scénu. Opisuje to, ako sa bod na jednej snímke zobrazí v susednej snímke. Cieľom je pre daný bod x' z prvej snímky nájsť korešpondujúci bod x'' v druhej snímke. Epipolárna geometria zjednoduší hľadanie bodových korešpondencií tým, že redukuje prehľadávanú oblasť celého 2D obrazu na 1D priamku, na ktorej leží hľadaný bod. Tým je stanovenie korešpondencií medzi snímkami rýchlejšie a takisto nám to umožňuje predpovedať, kde sa bude hľadaný bod nachádzať v druhej snímke. To umožňuje odhalovanie chybne stanovených korešpondencií napríklad medzi viacerými podobnými objektmi v scéne. [27] [28]

Epipolárna geometria dvoch pohľadov je zobrazená na obrázku 2.18: 3D bod X je pozorovaný dvoma pohľadmi s projekčnými stredmi O' a O'' . Priamka B spájajúca projekčné stredy sa nazýva epipolárna os. Body O' , O'' a X tvoria epipolárnu rovinu



Obr. 2.18: Epipolárna geometria dvoch snímok

ϵ , ktorá závisí na polohe projekčných stredov a bodu \mathbf{X} . Pozorované 3D body generujú rôzne epipolárne roviny, ktoré vznikajú rotáciou okolo epipolárnej osi, vidť pre bod \mathbf{Y} . Body e' a e'' sa nazývajú epipóly. Epipól je bod v obrazovej rovine, do ktorého je premietnutý projekčný stred susednej snímky: $e' = (\mathbf{O}'')'$ pre epipól prvého pohľadu a $e'' = (\mathbf{O}')''$ pre epipól druhého pohľadu. Epipóly tiež predstavujú body, kde epipolárna os pretína obrazovú rovinu. Priamky l' a l'' sú epipolárne priamky, ktoré predstavujú prienik epipolárnej a obrazovej roviny. Epipolárna priamka tiež vyjadruje zobrazenie projekčného lúča v obraze susednej snímky. Všetky epipolárne priamky na danej snímke sa pretínajú v jej epipóle. Projekčný lúč/vektor $\mathbf{O}'\mathbf{X}$ je zobrazený v susednej snímke ako epipolárna priamka: $l''(\mathbf{X}) = (\mathbf{O}'\mathbf{X})''$. Ak máme daný bod \mathbf{x}' a hľadáme jeho korešpondujúci bod \mathbf{x}'' v druhej snímke, tak tento bod bude ležať na epipolárnej priamke l'' . Tým pádom je vyhľadávanie korešpondencií obmedzené na 1D priamku, čo je rýchlejšie než prehľadávanie celého obrazu a tiež môžeme podľa vzdialenosť korešpondujúceho bodu od epipolárnej priamky rozhodovať o správnosti stanovených bodových korešpondencií, ktoré získame napríklad z porovnania príznakov významných bodov. [27] [28]

Výpočet epipolárnej priamky sa vyjadri pomocou vzťahu:

$${}^k \mathbf{x}'^T \mathbf{l}' = 0 \quad (2.74)$$

ktorý hovorí, že bod ${}^k\mathbf{x}'$ leží na priamke \mathbf{l}' . Bod ${}^k\mathbf{x}'$ je vyjadrený v normalizovaných obrazových súradniciach resp. je vyjadrený v súradnicovom systéme kamery ako:

$${}^k\mathbf{x}' = \mathbf{K}'^{-1} \mathbf{x}' \quad (2.75)$$

kde bod \mathbf{x}' je v obrazových resp. pixelových súradniciach. Ďalej môžeme využiť podmienku koplanarity z predošej kapitoly vyjadrenú pomocou esenciálnej matice:

$${}^k\mathbf{x}'^T \mathbf{E} {}^k\mathbf{x}'' = 0 \quad (2.76)$$

Porovnaním vzťahov 2.74 a 2.76 môžeme epipolárnu priamku \mathbf{l}' vyjadriť ako:

$$\mathbf{l}' = \mathbf{E} {}^k\mathbf{x}'' \quad (2.77)$$

Pre epipolárnu priamku \mathbf{l}'' na susednej snímke potom platí:

$$\mathbf{l}'' = \mathbf{E}^T {}^k\mathbf{x}' \quad (2.78)$$

Na výpočet epipolárnej priamky teda potrebujeme vedieť pozíciu bodu v susednej snímke a esenciálnu maticu. Zo znalosti epipolárnej priamky môžeme predpovedať, kde sa bude korešpondujúci bod nachádzať - prehľadávaný priestor je obmedzený na 1D priamku. Taktiež môžeme pomocou epipolárnej priamky testovať vierohodnosť stanovených korešpondencií a odhalovať falošné asociácie medzi bodmi, napríklad pomocou vzdialenosť korešpondujúceho bodu od epipolárnej priamky. [27] [28]

2.8 Výpočet esenciálnej matice z bodových korešpondencií

V predošlých kapitolách bolo ukázané odvodenie esenciálnej matice z podmienky koplanarity dvoch pohľadov a jej uplatnenie v epipolárnej geometrii. Esenciálna matica bola definovaná vzťahom:

$$\mathbf{E} = \mathbf{R}' \mathbf{S}_b \mathbf{R}''^T \quad (2.79)$$

kde \mathbf{R}' je matica rotácie prvého pohľadu, uvažovaná ako jednotková, \mathbf{S}_b je antisymetrická matica, ktorá má v sebe zahrnutý smerový vektor bázy medzi dvoma pohľadmi (jeho absolútна velkosť nie je známa) a \mathbf{R}'' je matica rotácie druhého pohľadu. Parametrizáciou esenciálnej matice z kapitoly 2.6 uvažujeme iba rotáciu druhého pohľadu a smerový vektor bázy:

$$\mathbf{E} = \mathbf{S}_b \mathbf{R}^T \quad (2.80)$$

V tejto kapitole bude esenciálna matica vypočítaná na základe bodových korešpondencií medzi dvoma pohľadmi a potom z nej bude odvodena relatívna orientácia druhého pohľadu. [29] [30]

Máme dané korešpondujúce body v dvoch pohľadoch: ${}^k\mathbf{x}'_n$ a ${}^k\mathbf{x}''_n$ pre $n = 1, \dots, N$ a chceme určiť maticu \mathbf{E} . Index k znamená, že sa dané body nachádzajú v súradnicovom systéme kamery, do ktorého sú prepočítané body \mathbf{x}'_n z pixelových súradníc:

$${}^k\mathbf{x}'_n = \mathbf{K}'^{-1}\mathbf{x}'_n \quad (2.81)$$

Využijeme podmienku koplanarity z predošlých kapitol:

$${}^k\mathbf{x}'_n \mathbf{E} {}^k\mathbf{x}''_n = 0 \quad (2.82)$$

ktorú prepíšeme do jednotlivých zložiek:

$$\begin{bmatrix} {}^k\mathbf{x}'_n & {}^k\mathbf{y}'_n & F' \end{bmatrix} \begin{bmatrix} E_{11} & E_{12} & E_{13} \\ E_{21} & E_{22} & E_{23} \\ E_{31} & E_{32} & E_{33} \end{bmatrix} \begin{bmatrix} {}^k\mathbf{x}''_n \\ {}^k\mathbf{y}''_n \\ F'' \end{bmatrix} = 0 \quad (2.83)$$

kde F je ohnisková vzdialenosť kamery. Matica prvkov E_{ij} je neznáma a vektoru ${}^k\mathbf{x}'_n$ a ${}^k\mathbf{x}''_n$ poznáme. Roznásobením tohto vzťahu dostávame lineárnu funkciu v tvare:

$${}^k\mathbf{x}'_n E_{11} {}^k\mathbf{x}''_n + {}^k\mathbf{y}'_n E_{21} {}^k\mathbf{x}''_n + \dots = 0 \quad (2.84)$$

s neznámymi E_{ij} , ktoré sú násobené známymi koeficientmi. Tento vzťah môžeme vyjadriť ako násobenie vektoru známych koeficientov \mathbf{a}_n^T s vektorom neznámych \mathbf{e} :

$$\mathbf{a}_n^T = \begin{bmatrix} {}^k\mathbf{x}'_n {}^k\mathbf{x}''_n & {}^k\mathbf{y}'_n {}^k\mathbf{x}''_n & F' {}^k\mathbf{x}''_n & {}^k\mathbf{x}'_n {}^k\mathbf{y}''_n & {}^k\mathbf{y}'_n {}^k\mathbf{y}''_n & F' {}^k\mathbf{y}''_n & {}^k\mathbf{x}'_n F' & {}^k\mathbf{y}'_n c' & F' F'' \end{bmatrix} \quad (2.85)$$

$$\mathbf{e} = \begin{bmatrix} E_{11} & E_{12} & E_{13} & E_{21} & E_{22} & E_{23} & E_{31} & E_{32} & E_{33} \end{bmatrix}^T \quad (2.86)$$

$$\mathbf{a}_n^T \mathbf{e} = 0 \quad (2.87)$$

Každej dvojici bodov odpovedá jeden vektor \mathbf{a}_n^T a pre N dvojíc bodov zložíme tieto vektoru do matice \mathbf{A} :

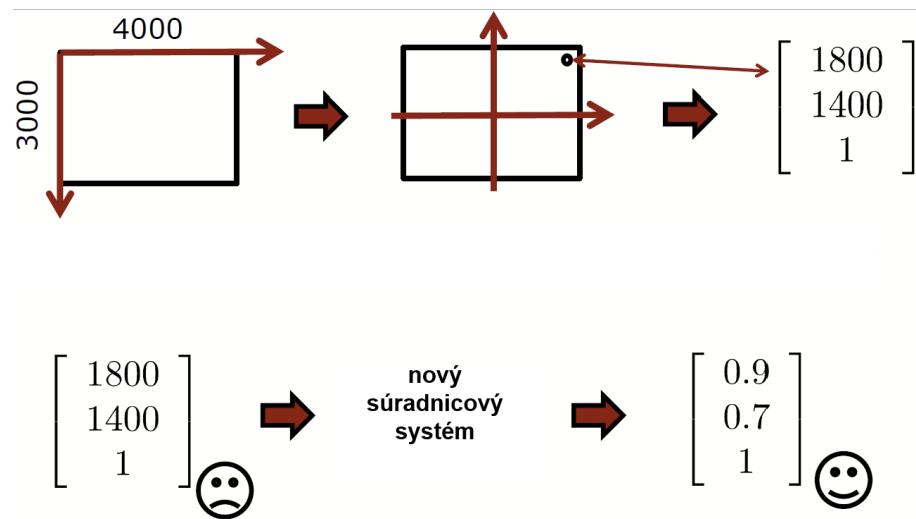
$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_n^T \\ \vdots \\ \mathbf{a}_N^T \end{bmatrix} \quad (2.88)$$

Následne dostaneme systém lineárnych rovníc daný vzťahom:

$$\mathbf{A}\mathbf{e} = 0 \quad (2.89)$$

Vektor e má 9 neznámych a aby sme ich mohli určiť musí mať matica A maximálnu hodnosť 8, čo znamená 8 bodových korešpondencií. Tento systém rovníc sa vyrieší metódou singulárneho rozkladu, z ktorého určíme maticu E . [29] [30]

Táto metóda výpočtu esenciálnej matice sa nazýva 8-bodový algoritmus a vyžaduje, aby boli pixelové súradnice vstupných bodov transformované tak, že počiatok $(0, 0)$ nového súradnicového systému bude v strede pixelovej mriežky a súradnice bodov budú prepočítané na stupnici od $[-1, 1]$, čo je znázornené na obrázku 2.19. Táto normalizácia je potrebná, pretože medzi homogénou súradnicou a pixelovými súradnicami môžu nastať veľké rozdiely, kvôli čomu by bola metóda numericky nestabilná. Takto získaná esenciálna matica sa musí na záver spätnou transformáciou vrátiť späť do pôvodných pixelových súradníc. [29] [30]



Obr. 2.19: Normalizované súradnice 8-bodového algoritmu [29]

Existujú prípady, v ktorých 8-bodový algoritmus zlyhá. Jedným z nich je situácia, kedy všetky vstupné korešpondujúce body ležia v rovine. Z toho plynie hodnosť matice A menšia než 8. Ďalšou prípadom je, ak majú oba pohľady zhodné pozície projekčných stredov, teda medzi pohľadmi nie je žiaden posun, iba vzájomná rotácia.[29] [30]

Pre esenciálnu maticu E platí, že je singulárna (jej determinant je rovný nule) a nemá plnú hodnosť - jej hodnosť je rovná 2. Ďalej platí, že jedna jej singulárna hodnota je nulová a zvyšné 2 nenulové singulárne hodnoty sú identické. Kvôli použitiu homogénnych súradníc sa rovnajú 1:

$$E = U \begin{bmatrix} d & 0 & 0 \\ 0 & d & 0 \\ 0 & 0 & 0 \end{bmatrix} V^T = U \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} V^T \quad (2.90)$$

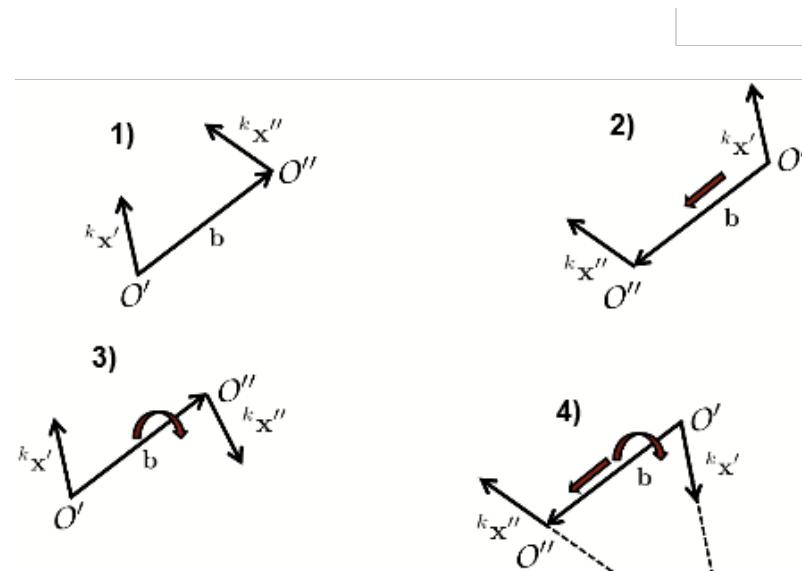
Na výpočet esenciálnej matice existuje viacero variant algoritmov, ktoré sa medzi sebou líšia v počte vstupných bodových dvojíc. Výpočet \mathbf{E} bol ukázaný na 8-bodovom algoritme, no existuje aj 5-bodový algoritmus [31], ktorý využíva minimálny počet vstupných bodových dvojíc, no jeho výpočet je zložitejší. Je variantou 8-bodového algoritmu a vedie na riešenie polynómu 10. stupňa, s 10 možnými riešeniami. Dnes je štandardne využívaný v knižnicových funkciách, ktoré slúžia na výpočet \mathbf{E} a následne relatívnej orientácie. Používa sa v kombinácii s iteratívnou metódou *RANSAC* (popísaná v nasledujúcej kapitole) pri odhalovaní nesprávne stanovených korešpondencií. Menší počet požadovaných bodov (v tomto prípade 5) znamená menší počet iterácií v *RANSAC* schéme. [29] [30]

5-bodový algoritmus v kombinácii s *RANSAC* metódou je v tejto práci použitý v rámci výpočtu relatívnej orientácie kamery.

Výpočítanú esenciálnu maticu chceme ďalej rozložiť na parametre relatívnej orientácie - maticu rotácie \mathbf{R} a antisymetrickú maticu \mathbf{S}_b :

$$\mathbf{E} = \mathbf{S}_b \mathbf{R}^T \quad (2.91)$$

v ktorej sa nachádza smerový vektor posunutia medzi dvoma pohľadmi. Z matematického hľadiska sú v esenciálnej matici zakódované 4 možné riešenia relatívnej orientácie resp. 4 konfigurácie dvoch pohľadov, z ktorých iba 1 je fyzikálne možné, vidieť obrázok 2.20. Toto riešenie znamená takú konfiguráciu, v ktorej sa oba projekčné lúče zbiehajú v bode pred kamerami, teda obe kamery sa dívajú pred seba a body sú zobrazené pred kamerami - 1) konfigurácia v obrázku 2.20. 2) varianta vznikne



Obr. 2.20: 4 možné riešenia relatívnej orientácie [29]

vynásobením (-1) a vektoru \mathbf{b} (vektor \mathbf{b} je známy až na jeho veľkosť). 3) variantu

dostaneme rotáciou druhého pohľadu o 180° . 4) varianta vznikne kombináciou predošlých operácií. Tieto usporiadania nie sú fyzikálne možné a je treba zaviesť spôsob akým z rozkladu esenciálnej matice vybrať iba jedno správne riešenie. Singulárny rozklad matice je:

$$\mathbf{E} = \mathbf{U} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \mathbf{V}^T \quad (2.92)$$

kde \mathbf{U} a \mathbf{V}^T sú rotačné matice. Ďalej zavedieme antisymetrickú maticu \mathbf{Z} a rotačnú maticu \mathbf{W} :

$$\mathbf{Z} = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \mathbf{W} = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

tak aby:

$$\mathbf{ZW} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (2.93)$$

\mathbf{ZW} dosadíme do vzťahu 2.92:

$$\mathbf{E} = \mathbf{U} \mathbf{ZWV}^T \quad (2.94)$$

Pretože je \mathbf{U} matica rotácie, je ortogonálna a platí $\mathbf{I} = \mathbf{U}^T \mathbf{U}$, čo môžeme umiestniť do vzťahu vyššie:

$$\mathbf{E} = \mathbf{U} \mathbf{Z} \mathbf{U}^T \mathbf{U} \mathbf{W} \mathbf{V}^T \quad (2.95)$$

Rozložením vzťahu hore na 2 časti dostávame výraz $\mathbf{U} \mathbf{Z} \mathbf{U}^T$, ktorý predstavuje hľadanú antisymetrickú maticu \mathbf{S}_b , pretože \mathbf{U} je matica rotácie a ak je násobená jej transpozíciou \mathbf{U}^T dostávame jednotkovú maticu, takže celý výraz $\mathbf{U} \mathbf{Z} \mathbf{U}^T$ zostáva stále antisymetrickou maticou. Druhá časť vzťahu vyššie, výraz $\mathbf{U} \mathbf{W} \mathbf{V}^T$ je súčinom 3 rotačných matíc, preto tento výraz predstavuje hľadanú maticu rotácie \mathbf{R}^T . [29] [30]

4 možné riešenia plynú z toho, že výraz \mathbf{ZW} môžeme vyjadriť dvojako:

$$\mathbf{ZW} = \mathbf{Z}^T \mathbf{W}^T \quad (2.96)$$

čo genereje 2 varianty pre \mathbf{S}_b a pre \mathbf{R} :

$$\mathbf{S}_b^1 = \mathbf{U} \mathbf{Z} \mathbf{U}^T \quad \mathbf{S}_b^2 = \mathbf{U} \mathbf{Z}^T \mathbf{U}^T \quad \mathbf{R}_1^T = \mathbf{U} \mathbf{W} \mathbf{V}^T \quad \mathbf{R}_2^T = \mathbf{U} \mathbf{W}^T \mathbf{V}^T$$

a ich kombináciou vzniknú 4 rôzne esenciálne matice. Pre každú esenciálnu maticu resp. relatívnu orientáciu potom testujeme, či sa všetky body zobrazili pred oboma kamerami, čo nám odkryje jedno fyzikálne prijatelné riešenie relatívnej orientácie. [29] [30]

2.9 RANSAC

Algoritmus *RANSAC* (*RA*ndom *S*ample *C*onsensus) slúži na detekciu nesprávnych asociácií medzi dátovými bodmi vo vstupnej množine dát. V tejto práci sa používa na detekciu chybné určených bodových korešpondencií medzi dvoma snímkami. Prvotne získané korešpondencie medzi významnými bodmi dvoch snímok môžu obsahovať chybné asociácie, ktoré znemožnia ďalšie operácie, ako je napríklad výpočet relatívnej orientácie dvoch pohľadov. Výstupom metódy *RANSAC* je rozdelená vstupná množina bodových dvojíc na správne a nesprávne korešpondencie a v nasledujúcich operáciách už teda môžeme pracovať len s množinou správnych bodových dvojíc. Ide o iteratívnu metódu typu pokus-omyl, ktorá si dokáže poradiť s vysokým podielom nesprávnych korešpondencií vo vstupných dátach. Dá sa rozdeliť do troch krokov:

- 1) Náhodný výber s bodových dvojíc. V tomto prípade je $s = 5$, pretože používame 5-bodový algoritmus. Tieto body sa uvažujú ako správne korešpondencie.
- 2) Výpočet modelu pomocou týchto vybraných bodov. Modelom je v tomto prípade esenciálna matica \mathbf{E} .
- 3) Ohodnotenie modelu na základe počtu zvyšných bodov v dátovej sade, ktoré podporia vypočítaný model. O tom, či zvyšné body podporia vypočítaný model rozhoduje kritérium, ktorým je v tomto prípade podmienka koplanarity pre dvojicu bodov \mathbf{x}'_n a \mathbf{x}''_n :

$$\mathbf{x}'_n \mathbf{E} \mathbf{x}''_n = 0 \quad (2.97)$$

a stanovená prahová hodnota δ :

$$\mathbf{x}'_n \mathbf{E} \mathbf{x}''_n < \delta \quad (2.98)$$

Počet zvyšných bodov, ktoré spĺňajú túto podmienku udáva skóre modelu a tieto body, spolu s počiatocnými s bodmi sú označené ako správne korešpondencie.

[32] [33]

Tento postup sa zopakuje T -krát a následne sa vyberie model resp. matica \mathbf{E} s najvyšším skóre. Body, ktoré tento model podporili sú označené ako správne korešpondencie. Z tejto matice \mathbf{E} môže byť následne odvodena relatívna orientácia kamery. [32] [33]

Otázkou je, aký počet pokusov T je potrebný. T závisí od počtu bodov s požadovaných na výpočet modelu. O týchto náhodne vybraných bodoch sa uvažuje, že sú správne a preto menší počet bodov znamená vyššiu šancu, že vytiahneme skutočne správne korešpondencie. Ďalším parametrom je podiel nesprávnych korešpondencií v dátach e , ktorý môže, ale nemusí byť vopred známy. Pravdepodobnosť, že vytiahneme správnu korešpondenciu z celého súboru s -krát je $(1 - e)^s$. Pravdepodobnosť úspechu p - náhodný výber je bez nesprávnych korešpondencií je p . Opačný jav,

teda kedy bude daný výber zatažený chybou je $1 - p$. Skombinovaním týchto dvoch vzťahov dostávame:

$$1 - p = 1 - (1 - e)^s \quad (2.99)$$

Tento vzťah predstavuje zlyhanie pri náhodnom výbere v jednej iterácii. Zlyhanie vo všetkých T iteráciách - prípad, kedy ani v jednej iterácii nevyberieme s bezchybných dvojíc je:

$$1 - p = (1 - (1 - e)^s)^T \quad (2.100)$$

z tejto rovnice vyjadríme T :

$$\begin{aligned} \log(1 - p) &= T \log((1 - (1 - e)^s)^s) \\ T &= \frac{\log(1 - p)}{\log((1 - (1 - e)^s)^s)} \end{aligned} \quad (2.101)$$

T sa volí tak, aby bol s pravdepodobnosťou p aspoň jeden náhodný výber s dvojíc bezchybný. Počet pokusov teda závisí na tom aký je podiel nesprávnych korešpondencií e vo vstupných dátach a koľko parametrov s je potrebných na výpočet modelu. Čím menej parametrov potrebujeme, tým je počet pokusov T menší. To platí aj o počte chybných korešpondencií. Výhodou metódy *RANSAC* je robustnosť pri určovaní falošných korešpondencií pri nižšom počte parametrov s (zhruba 1 až 10). Nevýhodou je, že časová náročnosť algoritmu rastie s podielom falošných korešpondencií v dátach a s počtom parametrov s . [32] [33]

V tejto práci je metóda *RANSAC* použitá v kombinácii s 5-bodovým algoritmom pri určovaní relatívnej orientácie kamery skrz esenciálnu maticu a pri určovaní pózy kamery pomocou metódy *Projective 3-Point*.

2.10 Triangulácia 3D bodov

3D body sú rekonštruované pomocou triangulácie zo známych orientácií/póz dvoch pohľadov a korešpondujúcich 2D bodov na snímkach. Existuje viacero metód, ktorými môžeme odvodiť pozíciu 3D bodov, v tejto kapitole bude ukázaná metóda využívajúca geometrický prístup a metóda lineárnej triangulácie. Na obrázku 2.21 je znázornená geometria dvoch pohľadov. \mathbf{P} a \mathbf{Q} predstavujú prvý a druhý pohľad s projekčnými stredmi \mathbf{X}'_O a \mathbf{X}''_O . \mathbf{r} a \mathbf{s} sú smerové vektory alebo projekčné lúče, ktoré predstavujú priamku, resp. smer, v ktorom leží daný 3D bod zobrazený ako 2D bod v obrázku a sú definované ako:

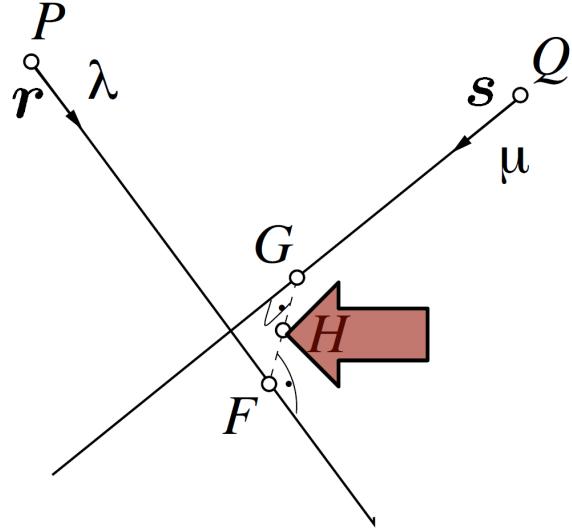
$$\mathbf{r} = \mathbf{R}'^T \mathbf{x}' \qquad \mathbf{s} = \mathbf{R}''^T \mathbf{x}''$$

${}^k\mathbf{x}'$ a ${}^k\mathbf{x}''$ sú taktiež smerové vektory, na ktorých leží daný 3D bod, ale sú vyjadrené v súradnicovom systéme kamery, kam boli prepočítané z pixelových súradníc bodov

\mathbf{x}' a \mathbf{x}'' pomocou známej kalibračnej matice \mathbf{K} , podľa vzťahu:

$${}^k\mathbf{x}' = \mathbf{K}'^{-1}\mathbf{x}' \quad (2.102)$$

Aby sme ${}^k\mathbf{x}'$ a ${}^k\mathbf{x}''$ vyjadrili v svetovom súradnicovom systéme, v ktorom sa nachádza hľadaný 3D bod, musíme do výpočtu zahrnúť matice rotácie \mathbf{R}' a \mathbf{R}'' . λ a μ sú



Obr. 2.21: Triangulácia [34]

neznáme skaláry, ktoré vyjadrujú veľkosť projekčných lúčov \mathbf{r} a \mathbf{s} a určujú bod na danej priamke, kde sa nachádza hľadaný 3D bod. [34] [35]

Bod, v ktorom sa stretávajú lúče \mathbf{r} a \mathbf{s} je hľadaný 3D bod. Lúče sa ale reálne nikdy nestretnú, preto sa musí zaviesť kompromis - určia sa body \mathbf{F} a \mathbf{G} , v ktorých je vzdialenosť medzi lúčmi najkratšia a hľadaný 3D bod \mathbf{H} sa bude nachádzať v jej strede. Body \mathbf{F} a \mathbf{G} sú pre nás neznáme a plynú z neznámych skalárov λ a μ . Situáciu na obrázku 2.21 môžeme vyjadriť ako rovnice dvoch priamok v 3D:

$$\mathbf{f} = \mathbf{X}_O' + \lambda\mathbf{r} \quad \mathbf{g} = \mathbf{X}_O'' + \mu\mathbf{s}$$

kde \mathbf{f} a \mathbf{g} tiež vyjadrujú polohu neznámych bodov \mathbf{F} a \mathbf{G} . Pre priamku, spojnicu bodov \mathbf{F} a \mathbf{G} , ktorá predstavuje najkratšiu vzdialenosť medzi projekčnými lúčmi \mathbf{r} a \mathbf{s} platí, že je na dané projekčné lúče kolmá. Z toho plynie vzťah:

$$(\mathbf{f} - \mathbf{g}) \cdot \mathbf{r} = 0 \quad (\mathbf{f} - \mathbf{g}) \cdot \mathbf{s} = 0$$

čo vedie na systém dvoch rovníc o dvoch neznámych λ a μ :

$$\begin{aligned} (\mathbf{X}_O' + \lambda\mathbf{r} - \mathbf{X}_O'' - \mu\mathbf{s}) \cdot \mathbf{r} &= 0 \\ (\mathbf{X}_O' + \lambda\mathbf{r} - \mathbf{X}_O'' - \mu\mathbf{s}) \cdot \mathbf{s} &= 0 \end{aligned} \quad (2.103)$$

Z obdržaných λ a μ vypočítame body \mathbf{F} a \mathbf{G} a hľadaný 3D bod \mathbf{H} určíme ako:

$$\mathbf{H} = \frac{\mathbf{F} + \mathbf{G}}{2} \quad (2.104)$$

Tento prístup využil iba základnú geometriu vektorov a teda aj takýto prístup vedie k riešeniu. [34] [35]

Ďalšou metódou je lineárna triangulácia, ktorá vychádza zo vzťahu projekcie 3D bodu \mathbf{X} do 2D bodu \mathbf{x} pomocou projekčnej matice \mathbf{P} :

$$\mathbf{x} = \mathbf{P}\mathbf{X} \quad (2.105)$$

Z tohto vzťahu je vidieť, že sa vektory \mathbf{x} a $\mathbf{P}\mathbf{X}$ rovnajú, ide teda o vektory s rovnakým smerom a pre ich vektorový súčin platí:

$$\mathbf{x} \times \mathbf{P}\mathbf{X} = 0 \quad (2.106)$$

Rozpísaním tohto vzťahu dostávame:

$$\begin{aligned} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \times \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} &= \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \\ \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \times \begin{bmatrix} \mathbf{p}_1^T \\ \mathbf{p}_2^T \\ \mathbf{p}_3^T \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} &= \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \\ \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \times \begin{bmatrix} \mathbf{p}_1^T \mathbf{X} \\ \mathbf{p}_2^T \mathbf{X} \\ \mathbf{p}_3^T \mathbf{X} \end{bmatrix} &= \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \\ \begin{bmatrix} y\mathbf{p}_3^T \mathbf{X} - \mathbf{p}_2^T \mathbf{X} \\ \mathbf{p}_1^T \mathbf{X} - x\mathbf{p}_3^T \mathbf{X} \\ x\mathbf{p}_2^T \mathbf{X} - y\mathbf{p}_1^T \mathbf{X} \end{bmatrix} &= \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \end{aligned} \quad (2.107)$$

Tretí riadok v poslednom vzťahu je lineárnnou kombináciou prvého a druhého riadku, takže ho môžeme vynechať:

$$\begin{bmatrix} y\mathbf{p}_3^T \mathbf{X} - \mathbf{p}_2^T \mathbf{X} \\ \mathbf{p}_1^T \mathbf{X} - x\mathbf{p}_3^T \mathbf{X} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (2.108)$$

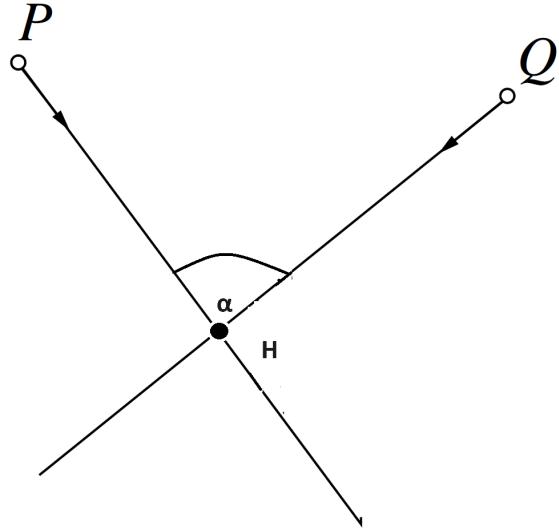
To isté zopakujeme pre bod \mathbf{x}' a projekčnú maticu \mathbf{P}' druhej kamery a pridáme ich rovnice k rovniciam vo vzťahu vyššie:

$$\begin{bmatrix} y\mathbf{p}_3^T \mathbf{X} - \mathbf{p}_2^T \mathbf{X} \\ \mathbf{p}_1^T \mathbf{X} - x\mathbf{p}_3^T \mathbf{X} \\ y'\mathbf{p}'_3^T \mathbf{X} - \mathbf{p}'_2^T \mathbf{X} \\ \mathbf{p}'_1^T \mathbf{X} - x'\mathbf{p}'_3^T \mathbf{X} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad (2.109)$$

$$\begin{bmatrix} y\mathbf{p}_3^T - \mathbf{p}_2^T \\ \mathbf{p}_1^T - x\mathbf{p}_3^T \\ y'\mathbf{p}'_3^T - \mathbf{p}'_2^T \\ \mathbf{p}'_1^T - x'\mathbf{p}'_3^T \end{bmatrix} \mathbf{X} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Týmto sme dostali systém 4 lineárnych rovnic a 3 neznámych: $\mathbf{AX} = \mathbf{0}$, ktorý vyriešime singulárnym rozkladom a obdržíme polohu 3D bodu \mathbf{X} . [36] [37]

Kvalita triangulovaného bodu závisí na horizontálnej vzdialosti medzi dvoma pohľadmi, ktorá je označovaná ako *baseline* - b . Pre kvalitnú trianguláciu chceme, aby bola b čo najvyššia. Kvalita triangulácie je tiež ovplyvnená tým, ako ďaleko sú objekty od kamery a môžeme ju určiť pomocou uhlovej paralaxy, ktorá je zachytená na obrázku 2.22. \mathbf{P} a \mathbf{Q} sú projekčné stredy dvoch pohľadov, \mathbf{H} je triangulovaný



Obr. 2.22: Uhlová paralaxa [34]

3D bod a α je uhol, ktorý zvierajú vektory \vec{PH} a \vec{QH} . Vypočíta sa ako:

$$\alpha = \cos^{-1} \frac{\vec{PH} \cdot \vec{QH}}{\|\vec{PH}\| \|\vec{QH}\|} \quad (2.110)$$

Nazýva sa uhlová paralaxa a čím je väčší, tým presnejšia je triangulácia. Veľkosť uhlu paralaxy teda závisí jednak na horizontálnej vzdialosti dvoch pohľadov a na vzdialosti objektu od kamery. Veľkosť uhlu paralaxy v tejto práci rozhoduje o tom, či

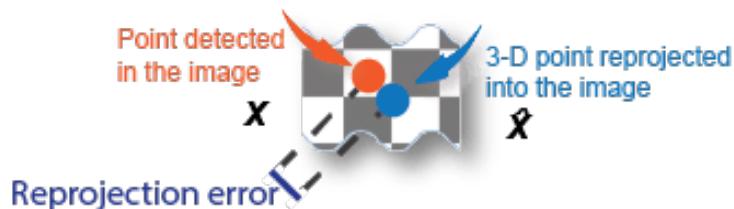
medzi dvomi po sebe idúcimi pohľadmi došlo k dostatočnému posunu (teda či je daný pohľad vhodný na trianguláciu) a tiež rozhoduje o tom, ktoré body budú triangulované a ktoré nie.

2.11 Bundle Adjustment

Bundle Adjustment je iteratívna optimalizačná metóda, ktorá sa používa na estimáciu/optimalizáciu 3D bodov \mathbf{X} a projekčných matíc \mathbf{P} na základe počiatočného odhadu - vieme polohy 3D bodov \mathbf{X} , projekčné matice pohľadov \mathbf{P} a 2D body \mathbf{x} v daných pohľadoch, ktoré zachytávajú 3D body \mathbf{X} scény. Chceme minimalizovať reprojekčnú chybu medzi bodom \mathbf{x} , ktorý vyjadruje správnu resp. očakávanú hodnotu a bodom $\hat{\mathbf{x}}$:

$$\hat{\mathbf{x}} = \mathbf{P}\mathbf{X} \quad (2.111)$$

ktorý vyjadruje reprojekciu známeho bodu \mathbf{X} späť do snímky, obrázok 2.23. Slep-



Obr. 2.23: Chyba reprojekcie [38]

dujeme euklidovskú vzdialenosť medzi týmito dvoma bodmi v snímke a chceme ju minimalizovať nájdením optimálnych parametrov \mathbf{P} a \mathbf{X} :

$$\min_{(\mathbf{X}_j, \mathbf{P}_i)} \sum_j \sum_i \|\mathbf{x}_{ij} - \mathbf{P}_i \mathbf{X}_j\|^2 \quad (2.112)$$

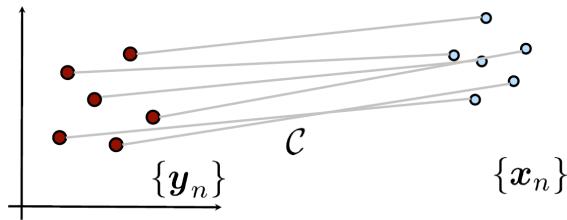
Index i znamená i -tý pohľad a index j zas j -tý bod v danom pohľade. Aby tento algoritmus fungoval, musíme poznať asociácie medzi 2D bodmi \mathbf{x} na snímke a 3D bodmi \mathbf{X} v scéne. [39] [40] [41]

V tejto práci je algoritmus *Bundle Adjustment* použitý na zlepšenie vypočítanej orientácie/pózy kamery a polohy triangulovaných 3D bodov. Každá vypočítaná póza kamery je optimalizovaná na základe predošlých pôz a na základe 3D bodov triangulovaných z predošlých pohľadov s tým, že poznáme korešpondencie medzi 2D a 3D bodmi.

2.12 Registrácia mračien bodov

V tejto kapitole je ukázané, akým spôsobom sa vypočíta podobnostná transformácia medzi dvoma 3D mračnami bodov, ak medzi nimi poznáme bodové korešpondencie. Táto transformácia prevedie mračná bodov do spoločného súradnicového systému tak, že vzdialenosť medzi odpovedajúcimi bodmi sú minimálne - takto je definovaná registrácia mračien bodov. Tento problém, v ktorom hľadáme podobnostnú transformáciu medzi dvoma množinami bodov, je v literatúre tiež označovaný ako *Absolute orientation problem*. [42] [43]

Majme dve množiny 3D bodov $\{\mathbf{y}_n\}$ a $\{\mathbf{x}_n\}$ a známe korešpondencie \mathcal{C} medzi týmito bodmi, vidieť obrázok 2.24. Obe množiny bodov predstavujú rovnaké body v



Obr. 2.24: Registrácia dvoch bodových množín [42]

3D svete, no sú umiestnené v rôznych súradnicových systémoch. Môže ísť napríklad o dva 3D modely scény rekonštruované v rôznych časoch. Ak uvažujeme, že boli rekonštruované pomocou monokulárnej kamery znamená to, že je medzi nimi podobnostná transformácia - translácia, rotácia a merítka. Množinu $\{\mathbf{y}_n\}$ zavedieme ako referenčný/globálny súradnicový systém, do ktorého chceme transformovať body $\{\mathbf{x}_n\}$ z lokálneho súradnicového systému:

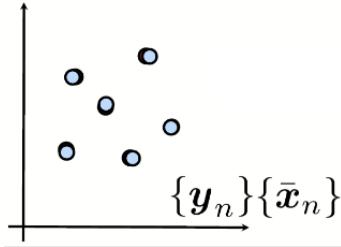
$$\bar{\mathbf{x}}_n = \lambda \mathbf{R} \mathbf{x}_n + \mathbf{t} \quad (2.113)$$

tak, aby vzdialosť medzi novými transformovanými bodmi $\bar{\mathbf{x}}_n$ a \mathbf{y}_n bola minimálna:

$$\sum \|\mathbf{y}_n - \bar{\mathbf{x}}_n\|^2 \rightarrow \min \quad (2.114)$$

vidieť obrázok 2.25. Neznámymi sú v tomto prípade merítka λ (1 parameter), rotácia \mathbf{R} (3 parametre) a translácia \mathbf{t} (3 parametre). Máme teda 7 stupňov voľnosti a na ich určenie potrebujeme minimálne 3 bodové korešpondencie. [42] [43]

Tento prípad odpovedá úlohe, ktorá je riešená v tejto práci: chceme registrovať lokálne mračno bodov s iným referenčným 3D modelom/mračnom bodov. Ak boli tieto modely rekonštruované zo snímok monokulárnej kamery, jedná sa o fotogrametrické modely, takže ako neznámy parameter nám pri translatii a rotácii pribudne skalárne merítko λ . Ak dokážeme medzi týmito dvoma mračnami určiť dostatočný počet bodových korešpondencií, môžeme určiť hľadanú transformáciu. Stanovovanie



Obr. 2.25: Registrácia dvoch bodových množín [42]

korešpondencií je dôležitá úloha v rámci problému registrácie, pretože na základe nich budeme počítať výslednú transformáciu a predpokladáme, že sú správne. V reálnych dátach sa ale nevyhneme prítomnosti chybných korešpondencií tzv. *outliers*, ktoré musíme vhodnou metódou (popísaná v kapitole návrhu algoritmov 4.2) rozlísiť od správnych korešpondencií tzv. *inliers*.

Ked' je stanovený dostatočný počet korešpondencií, vypočítame pre dané body ich tažiská:

$$\mathbf{x}_0 = \frac{\sum \mathbf{x}_n}{N} \quad \mathbf{y}_0 = \frac{\sum \mathbf{y}_n}{N}$$

kde N je počet korešpondencií. Následne vypočítame kovariančnú maticu \mathbf{H} :

$$\mathbf{H} = \sum (\mathbf{x}_n - \mathbf{x}_0)(\mathbf{y}_n - \mathbf{y}_0)^T \quad (2.115)$$

z ktorej určíme singulárnym rozkladom maticu rotácie \mathbf{R} :

$$\begin{aligned} svd(\mathbf{H}) &= \mathbf{U} \mathbf{D} \mathbf{V}^T \\ \mathbf{R} &= \mathbf{V} \mathbf{U}^T \end{aligned} \quad (2.116)$$

Merítko λ určíme ako:

$$\lambda = \sqrt{\frac{\sum (\mathbf{y}_n - \mathbf{y}_0)^T (\mathbf{y}_n - \mathbf{y}_0)}{\sum (\mathbf{x}_n - \mathbf{x}_0)^T (\mathbf{x}_n - \mathbf{x}_0)}} \quad (2.117)$$

Posunutím \mathbf{t} zabezpečíme, že tažiská budú v rovnakom bode:

$$\mathbf{t} = \mathbf{y}_0 - \lambda \mathbf{R} \mathbf{x}_0 \quad (2.118)$$

Tým sme obdržali parametre podobnostnej transformácie a môžeme na základe tej transformovať všetky body v lokálnom mračne bodov podľa vzťahu 2.113. [42] [43]

Ide o priame riešenie, ktoré sa vypočíta v jednom kroku, no predpokladom je znalosť bodových korešpondencií. Toto priame riešenie je základom metódy ICP (*Iterative Closest Point*), ktorá sa skladá z týchto krokov: 1) stanovenie bodových korešpondencií na základe najbližších susedných bodov, 2) výpočet a aplikácia transformácie na dané body a 3) opakuj predošlé dva body, kým nie je dosiahnutá minimálna vzdialenosť medzi bodmi. Metóda popísaná v tejto kapitole je druhým krokom

algoritmu ICP. ICP metóda nevyžaduje vopred známe korešpondencie, no vyžaduje aby boli mračná bodov aspoň približne zarovnané. ICP sa používa skôr na finálne doladenie registrácie a jej nevýhodou je, že dokáže určiť iba rigidnú transformáciu. [42] [43]

V tejto práci je využité to, že bodové korešpondencie medzi dvomi mračnami bodov sa môžu vopred stanoviť a na základe nich sa vypočítá hľadaná podobnostná transformácia v jednom kroku. Na základe tejto transformované lokálne mračno bodov do súradnicového systému referenčného mračna. Vzdialosti medzi odpovedajúcimi si bodmi z oboch mračien sú minimálne - mračná bodov sú zarovnané / registrované alebo inak povedané 'do seba zapadnú' a v tejto podobe ich môžeme medzi sebou priamo porovnať za účelom nájdenia zmien.

2.13 Povrchové normály

Povrchové normály patria medzi základné charakteristiky 3D mračien bodov. Vyjadrujú orientáciu daného povrchu, ktorý je reprezentovaný prepojenou množinou 3D bodov. Povrchové normály predstavujú jednotkové vektory kolmé na daný povrch a sú určené pre každý bod povrchu. Na určenie normály \vec{n} v bode \mathbf{p} sa používa množina bodov \mathbf{P}^k , ktorá predstavuje k najbližších bodov v euklidovskej vzdialosti od bodu \mathbf{p} . Určenie normály pre daný bod na povrchu je prevedené na určenie normály pre rovinu dotyčnicovú k danému povrchu. Táto rovina je reprezentovaná bodom $\bar{\mathbf{p}}$:

$$\bar{\mathbf{p}} = \frac{1}{k} \cdot \sum_{i=1}^k \mathbf{p}_i \quad (2.119)$$

ktorý predstavuje ľažisko bodov $\mathbf{p}_i \in \mathbf{P}^k$ a normálovým vektorom \vec{n} , ktorý sa určí z vlastných čísel a vlastných vektorov kovariančnej matice $\mathbf{C}^{3 \times 3}$:

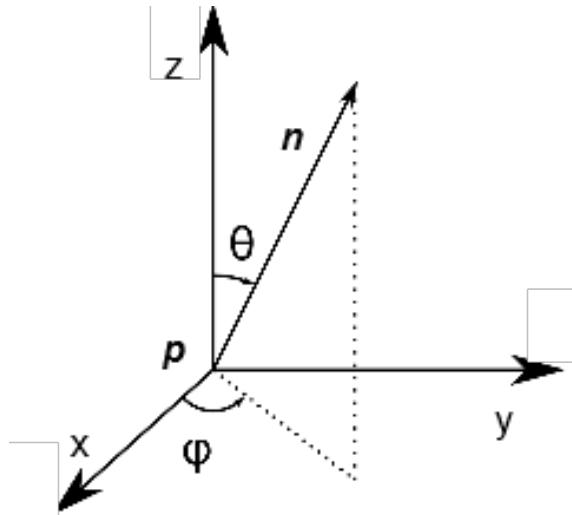
$$\mathbf{C} = \sum_{i=1}^k (\mathbf{p}_i - \bar{\mathbf{p}}) \cdot (\mathbf{p}_i - \bar{\mathbf{p}})^T. \quad (2.120)$$

Ak pre vlastné čísla $\{\lambda_i\}_{i=1..3}$ matice \mathbf{C} platí $0 \leq \lambda_1 \leq \lambda_2 \leq \lambda_3$, tak jednotkový vektor \mathbf{v}_1 odpovedajúci najmenšiemu vlastnému číslu λ_1 je potom aproximáciou normálového vektoru \vec{n} alebo $-\vec{n}$. Pre znamienko vektoru normály neexistuje žiadne matematické riešenie, preto je orientácia normálového vektoru dvojaká - môže smerovať dovnútra alebo von z povrchu. Orientácia každej normály musí byť určená tak, aby všetky normály daného povrchu mali konzistentný smer. Problém určenia smeru \vec{n} môže byť vyriešený, ak poznáme pozíciu kamery/pohľadu \mathbf{X}_O , pomocou ktorého bol daný povrch nasnímaný a rekonštruovaný. Na základe uhlu medzi normálou \vec{n} a vektorom $(\mathbf{X}_O - \mathbf{p})$ spájajúcim daný bod s pozíciou kamery môžeme nastaviť správny smer normály \vec{n} vzhľadom k miestu snímania. [44]

Povrchová normála $\vec{n} = \{n_x, n_y, n_z\}$ môže byť tiež vyjadrená v sférických súradničach pomocou uhlov ϕ a θ , obrázok 2.26:

$$\phi = \tan^{-1} \frac{n_y}{n_x} \quad \theta = \cos^{-1} \frac{n_z}{\sqrt{n_x^2 + n_y^2 + n_z^2}}$$

Uhol ϕ sa označuje ako azimut a θ ako uhol sklonu a sú v rozsahu $0 \leq \phi \leq 2\pi$ a



Obr. 2.26: Reprezentácia normály vo sférických súradničach [45]

$$0 \leq \theta \leq \pi. [44]$$

Tieto uhly potom vyjadrujú orientáciu daného povrchu, čo je v tejto práci využité.

2.14 Voxelová reprezentácia mračna bodov

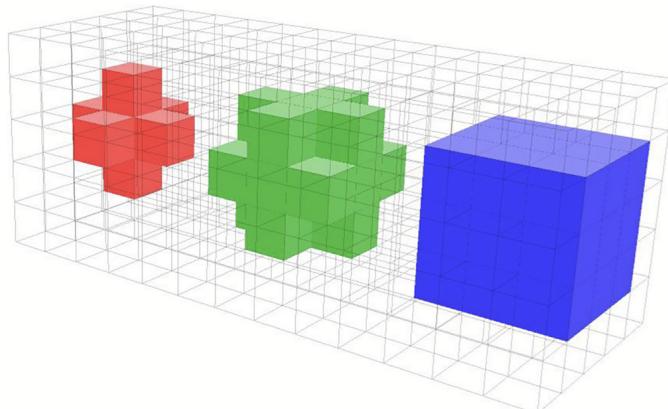
Voxely sú kubické elementy v 3D priestore, ktorých funkcia je analogická k pixelom v obrázku - tzv. volumetrický pixel. Zmysel voxelov spočíva v segmentácii priestoru (ktorý napríklad obsahuje mračno bodov) do pravidelnej mriežky/siete zloženej z týchto objemových elementov. Voxely mávajú zvyčajne tvar kocky alebo kvádru a nesú so sebou informáciu o svojej obsadenosti v podobe dvojstavovej hodnoty. Každý voxel je definovaný svojou polohou vo voxelovej mriežke, rozmermi (rozsílenie segmentácie) a obsadenostou. V prípade, že je voxel obsadený 3D bodmi obsahuje indexy týchto bodov alebo priamo ich súradnice, farbu, normály, početnosť a ďalšie. [46] [47]

Takáto segmentácia mračna bodov sa nazýva voxelizácia a jej význam spočíva v tom, že mračná bodov sú vo svojej prvotnej podobe neštruktúrované - 3D body sú zachytené nezávisle a môžu mať medzi sebou rôzne vzdialenosť. Rôzne oblasti v mračne bodov zas môžu mať rôzne bodové hustoty. Mračno bodov je väčšinou

uložené ako zoznam 3D bodov a ich poradie nevyjadruje to, ako sú body rozložené v priestore alebo ktoré body sú susedné. Mračná bodov teda nemajú jasne danú topológiu/štruktúru. Voxelizáciou teda stanovíme pre dané mračno bodov jasne definovanú štruktúru, v ktorej sú jednotlivé prvky (voxely) umiestnené v pravidelnej mriežke - poznáme polohu a rozmer každého voxelu a vzdialenosť medzi jednotlivými voxelmi, ktoré sú konštantné. V takto zavedenej dátovej štruktúre môžeme jednoduchšie pristupovať k jednotlivým prvkom (voxelom a následne k samotným 3D bodom), tiež máme informáciu o tom, kde sa daný voxel v priestore nachádza a poznáme jeho susedné voxely. Štruktúra, do ktorej organizujeme mračno bodov (v tomto prípade voxelová mriežka) závisí na type operácií, ktoré budeme s mračnom bodov vykonávať a musí byť zvolená tak, aby sa tieto operácie vykonávali čo najefektívnejšie. [47]

Voxelizácia mračna bodov prebieha v nasledujúcich krokoch. V prvom kroku sa určí priestor, ktorý ohraničuje dané mračno bodov. V druhom kroku je tento vyšetrovaný priestor rozdelený do kubických elementov (napríklad kociek) o zvolenej veľkosti, ktoré budú predstavovať jednotlivé voxely. Zvolená veľkosť/rozmer voxelov je pre každý voxel rovnaká a tiež určuje rozlíšenie, ktorým chceme mračno bodov vzorkovať - čím je menšia veľkosť voxelov, tým viac detailov je možné zachytit a opačne. V poslednom kroku sú každému voxelu pridelené 3D body z mračna bodov, ktoré sa nachádzajú vo vnútri daného voxelu. Každý voxel je teda definovaný polohou v rámci voxelovej mriežky, veľkosťou a bodmi, ktoré sa v ňom nachádzajú. [47]

Voxelová mriežka predstavuje 3D maticu, kde každý jej prvak predstavuje daný voxel a táto matica odpovedá tomu, ako sú voxely umiestnené v priestore. Výhodou je, že iba pomocou indexov môžeme pristupovať ku konkrétnym voxelom a taktiež k ich susedným voxelom, bez potreby prepočítavania priestorových súradníc. Voxelové okolie resp. susedstvo, viď obrázok 2.27, je pre daný voxel definované ako a) 6-okolie,



Obr. 2.27: Rôzne definície voxelového okolia [47]

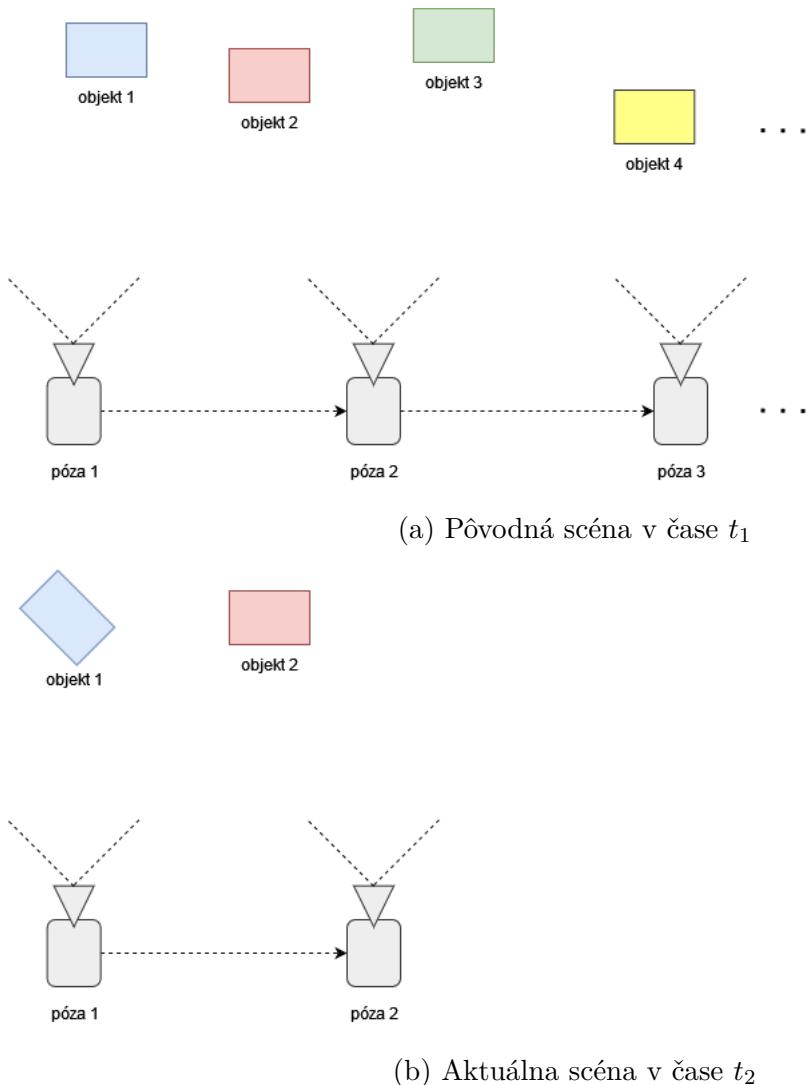
daný voxel má so susednými voxelmi spoločné steny; b) 18-okolie, spoločné steny

a hrany; c) 26-okolie, spoločné steny, hrany a rohy a voľba typu okolia závisí na danej aplikácii. Pomocou susedstva voxelov môžeme viaceré voxely zoskupovať do zhľukov, tzv. supervoxelov a analyzovať ich ako samostatné elementy. [47]

V tejto práci je koncept voxelizácie využitý v časti samotnej detekcie zmien. Mračná bodov sú z predoších krokov registrované - vzdialenosť medzi odpovedajúcimi bodmi sa uvažuje ako nulová. Vyšetrovaný priestor ohraničujúci obe mračná bodov je voxelizovaný a postupne sú všetky voxely prechádzané. V každom voxelovi sa porovnávajú body resp. povrchy, ktoré sa v ňom nachádzajú a na základe toho je každý voxel (spolu s danými bodmi) prvotne klasifikovaný ako 'zmena' alebo 'žiadna zmena'. Finálna klasifikácia voxelov berie do úvahy aj hodnoty voxelov z lokálneho 26-okolia.

3 Návrh scény a testovacích dát

Úlohou tejto práce je detekcia geometrických rozdielov medzi scénou nasnímanou v čase t_1 a scénou nasnímanou neskôr v čase t_2 . Obe scény sú reprezentované vo forme 3D mračien bodov, ktoré sú rekonštruované z videosekvencie monokulárnej kamery. Scéna predstavuje trasu, na ktorej sú premiestnené objekty a cez ktorú sa kamera pohybuje a zaznamenáva ju. Typický scenár je znázornený na obrázku 3.1.



Obr. 3.1: Scéna pri pohľade zhora

Prvým prechodom scénou v čase t_1 sa nasníma a rekonštruuje vzorová predloha danej scény v podobe 3D mračna bodov. Potom sú vybrané objekty v scéne rôzne premiestnené, otočené, odobrané alebo pridané, čím dojde k zmene geometrie scény. Následne je nasnímaný a rekonštruovaný menší úsek trasy/scény v čase t_2 , ktorý obsahuje dané zmeny a je braný ako aktuálna resp. porovnávaná trasa, vid obrázok 3.1,

ktorá má tiež podobu 3D mračna bodov. Tento lokálny úsek resp. lokálne mračno bodov tiež predstavuje aktuálnu polohu v rámci globálneho 3D modelu/mapy predlohy. Toto menšie mračno bodov, nasnímané v čase t_2 je registrované v globálnom 3D modeli trasy (predlohe) zachytenom v čase t_1 . Registrácia v tomto zmysle znamená aj lokalizáciu, teda nájdenie aktuálnej polohy vo vzorových dátach. Registrované mračná bodov sú pripravené na ich priame porovnanie a detekciu zmien.

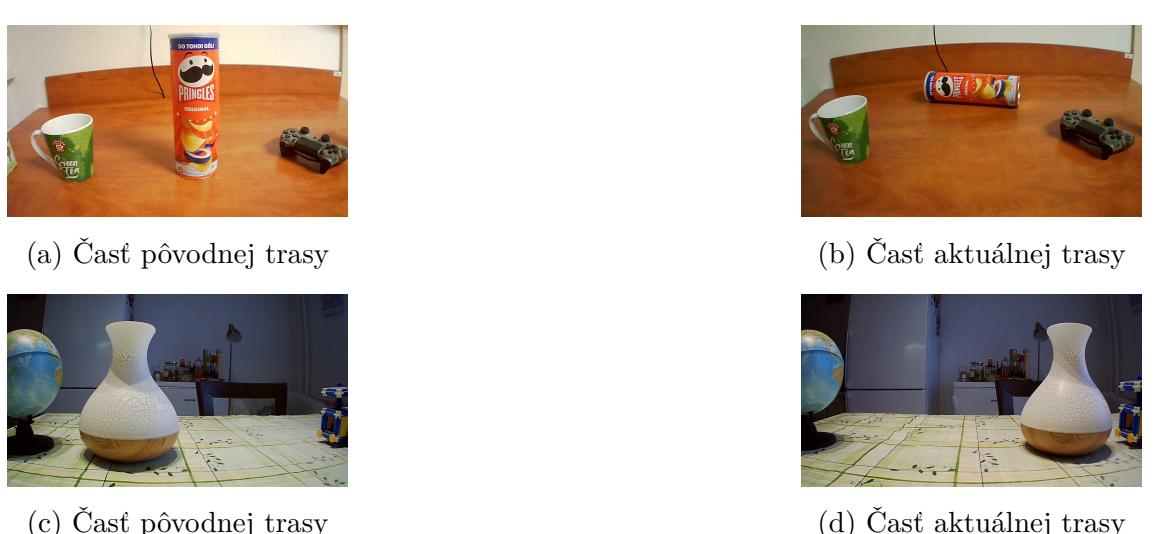
Detegované geometrické zmeny sú uvedené v nasledujúcim odstavci a zoradené od najjednoduchšieho prípadu po najnáročnejší. Z testovacích dát sú vybrané dvojice pôvodnej a aktuálnej trasy a v nich zobrazené konkrétné miesta, kde sa nachádzajú zmeny:

1) Pridanie alebo odstránenie objektov.



Obr. 3.2: Pridanie objektu

2) Posunutie objektov.



Obr. 3.3: Posunutie objektov

3) Rotácia objektov.



(a) Časť pôvodnej trasy

(b) Časť aktuálnej trasy

Obr. 3.4: Rotácia objektu

4) Zámena objektov. Kombináciou vyššie uvedených operácií je možné vytvoriť rôzne usporiadania scény.



(a) Časť pôvodnej trasy

(b) Časť aktuálnej trasy



(c) Časť pôvodnej trasy

(d) Časť aktuálnej trasy

Obr. 3.5: Zámena objektov

3.1 Vlastnosti scény

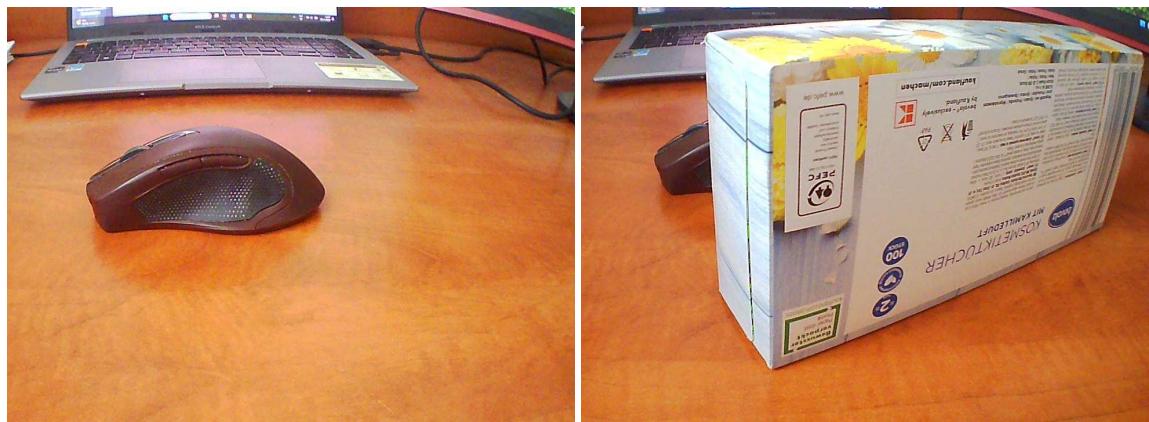
Vlastnosti snínamej scény, ktoré majú vplyv na riešenie a ktoré určujú limitujúce podmienky sú:

1) *Interiér a exteriér.* Testovacie dáta sú nasnímané predovšetkým v interiéri, v kontrolovanom prostredí, v ktorom sa dajú nasimulovať rôzne scenáre a typy zmien bez toho, aby na výsledok vplývali vonkajšie vplyvy ako je nestále osvetlenie, dynamika scény, oklúzia, počasie, opakujúce sa štruktúry atď.

2) *Pohyb.* Z hľadiska pohybu objektov, môže byť scéna: a) *Statická*. Objekty a pozadie v scéne sú počas snímania nehybné; b) *Dynamická*. Objekty alebo pozadie sa počas snímania hýbu, typicky autá, chodci, vtáctvo, zver. V tejto práci je scéna počas snímania považovaná za statickú. K pohybu objektov dochádza len medzi jednotlivými snímaniami scény, kvôli simulácií geometrických zmien.

3) *Zmeny osvetlenia.* Počas snímania scény aj medzi jednotlivými snímaniami sú uvažované budžiadne alebo minimálne zmeny v osvetlení. V kontrolovanom interiéri takmer nedochádza k výrazným zmenám v osvetlení (ak sa tam nedostane svetlo z vonku, napr. cez okno). Komplikovanejšie sú scény v exteriéri, kde je osvetlenie dané časom a aktuálnym počasím a môže sa častejšie a vo väčšej mieri meniť. 3D detekcia zmien by podľa prieskumu literatúry v kapitole 1.2.1 nemala byť, s výnimkou extrémnych prípadov, ovplyvnená zmenami v osvetlení scény (čo je slabinou 2D prístupu). Miera odolnosti 3D detekcie zmeny voči zmenám v osvetlení nie je v tejto práci overovaná.

4) *Oklúzia.* Snímané scény sú usporiadane tak, aby k oklúzii nedochádzalo; tento prípad teda riešený nie je. Pre testovacie dáta sa predpokladá, že to čo kamera „nevidí“, tak to sa v scéne nenachádza. Príklad oklúzie, kde jeden objekt blokuje druhý, je vidieť na obrázku 3.6.



(a) Scéna v čase t_1

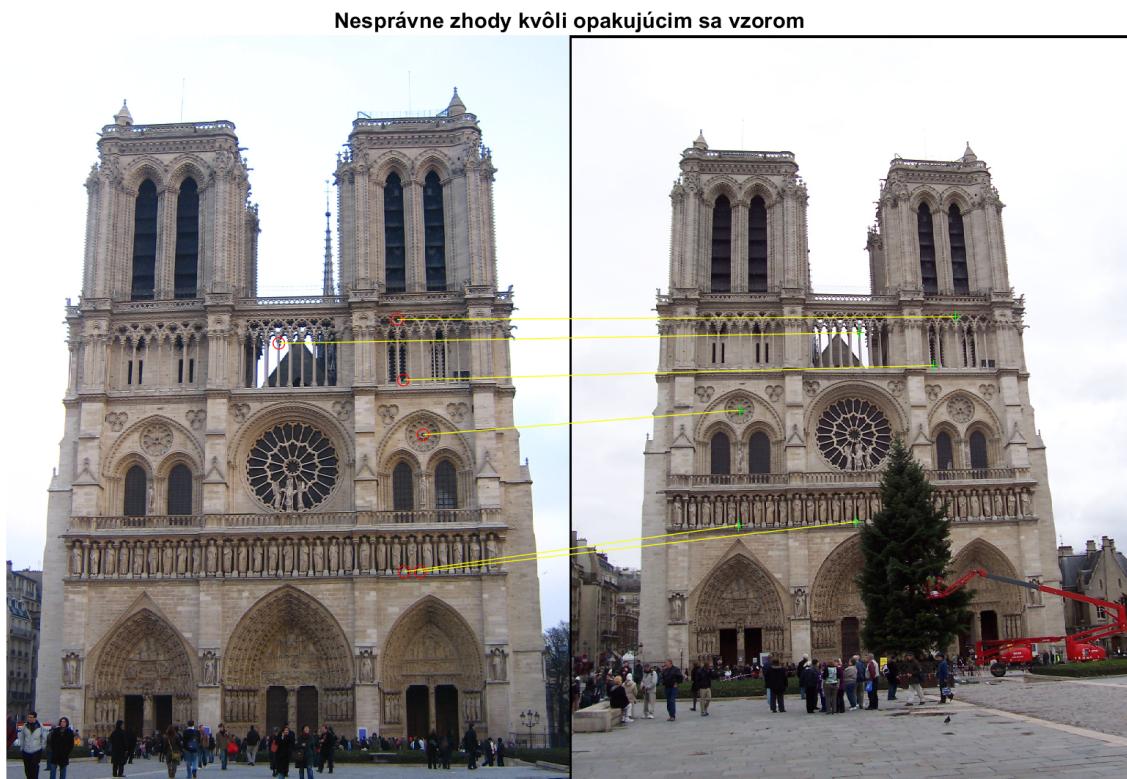
(b) Scéna v čase t_2

Obr. 3.6: Oklúzia

5) *Počasie.* Testovacie dáta sú snímané prevažne v interiéri, ak by ale zber prebiehal vonku, tak ideálny čas je okolo poludnia, kedy je slnko najvyššie a ideálne počasie je zamračené. Zamračené počasie a čas poludnia zaručujú rovnomerné a dostatočné

osvetlenie bez výrazných tieňov a príliš lesknúcich sa povrchov. Práve odrazivé povrhy, ktoré sa zvyknú za slnečného počasia lesknúť alebo ktoré reflektujú scénu ovplyvňujú 3D rekonštrukciu a dokážu v 3D modeli produkovať artefakty.

6) *Opakujúce sa vzory*. Ide o problém hlavne v mestskom prostredí, kde sa vyskytujú opakujúce sa prvky ako napríklad semafory, dopravné značky, okná, fasády budov. Kvôli tomu môže v snímkach rovnakej scény dochádzať k chybným korešpondenciám medzi detegovanými bodmi, ktoré majú rovnaké fotometrické vlastnosti (napríklad rohy viacerých okien), ale nejedná sa o rovnaké body v reálnom svete, viď obrázok 3.7. Chybne určené korešpondencie znepresňujú výpočet pózy kamery a následne aj presnosť 3D modelu, preto je nutné ich filtrovať.



Obr. 3.7: Nesprávne určené korešpondencie kvôli vzhľadovo podobným vzorom

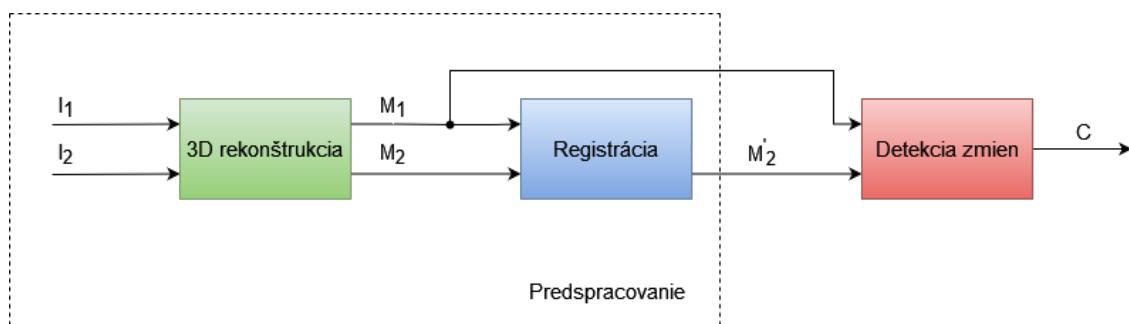
7) Homogénne oblasti. Snímanie homogénnych oblastí bez výraznej textúry znamená v daných snímkach nízky počet významých bodov, tým pádom aj málo korešpondencií medzi snímkami, čo má zá následok zhoršený alebo chybný výpočet pózy kamery. Prechody cez homogénne oblasti sa dajú softvérovo ošetriť napríklad detekciou náhleho poklesu v počte korešpondencií a ignorovaním takýchto snímok. Testovacie dátá sú zaznamenané v kontrolovanom prostredí, takže sa dá zaistiť, aby bola scéna dostatočne rozmanitá a bohatá na textúru.

Tieto vlastnosti predstavujú limity systému detektie zmien a týkajú sa hlavne robustnosti 3D rekonštrukcie. Ošetrenie všetkých stavov by bolo časovo náročné, preto je snímaná scéna vhodne prispôsobená.

4 Návrh systému detekcie zmien

Úloha detekcie zmien sa formálne delí na 3 časti: 1) *Zber a spracovanie dát*; 2) *Registrácia dát*; 3) *Analýza zmien*. Prvé dva kroky sú považované za kroky predspracovania, v ktorých sa generujú a transformujú dátá do takej podoby, aby boli pripravené na následnú analýzu zmien. [1]

Tomuto konceptu odpovedá aj systém detekcie zmien navrhovaný v tejto práci. Jeho bloková schéma je na obrázku 4.1, kde vstupmi sú sekvenčne snímok I_1 a I_2 zachytené v časoch t_1 a t_2 . M_1 je 3D model pôvodnej trasy, rekonštruovaný zo snímok I_1 a slúži ako predloha. M_2 je 3D model aktuálnej trasy, rekonštruovaný zo snímok I_2 . Registrované mračná bodov M_1 a transformované M'_2 sú následne podrobene analýze zmien a výstupom celého systému je binárna maska C , ktorá každému 3D bodu z M_1 a M'_2 prideľuje príznaky: „zmena“ alebo „žiadna zmena“.



Obr. 4.1: Systém 3D detekcie zmien

4.1 3D rekonštrukcia

Vstupom pre tento modul je videosekvencia zachytávajúca scénu a výstupom je jej 3D reprezentácia vo forme mračna bodov. Táto kapitola sa zaobrá návrhom algoritmov 3D rekonštrukcie. 3D rekonštrukcia sa v tejto práci skladá z troch častí. V prvej časti sú vypočítané pózy/vonkajšie parametre vybraných snímok zo vstupnej videosekvencie. V druhej časti je na základe týchto vybraných snímok rekonštruované 3D mračno bodov. V poslednej časti sa z tohto mračna bodov vyfiltrujú/odstránia chybné body, ktoré predstavujú šum.

4.1.1 Výber snímača

Vstupom pre navrhovaný systém detekcie zmien sú videosekvencie 2D snímok, čo zužuje výber možných snímačov na kamerové systémy. Do úvahy prichádzajú dve

varianty a to monokulárna a stereo kamera. V prípade monokulárnej kamery je k dispozícii jeden objektív resp. jeden pohľad. Aby bolo možné scénu zrekonštruovať, musí byť nasnímaná z viacerých pohľadov (minimálne z dvoch), ktoré sa dostatočne prekrývajú a zároveň je medzi nimi dostatočná paralaxa. Výhodou oproti stereo kamere je nižšia cena a širšia dostupnosť. Ak sú snímky z videosekvencie spracovávané sekvenčne, čo je prípad tejto práce, tak musí byť orientácia a pohyb kamery prispôsobený tak, aby medzi po sebe idúcimi snímkami bola dostatočná paralaxa. Stereo kamera sa skladá z dvoch objektívov (pohľadov), medzi ktorými je konštantná horizontálna vzdialenosť tzv. *baseline*. Vďaka tomu sú k dispozícii vždy dva pohľady s dostatočnou paralaxou, čo umožňuje volný pohyb kamery scénou bez obmedzení. Tým, že je známa skutočná vzdialenosť medzi dvoma pohľadmi kamier v reálnych jednotkách, výsledný 3D model zodpovedá skutočným rozmerom scény - má teda definované merítko, ktoré je v prípade 3D modelu monokulárnej kamery neznáme. Nevýhodou je vyššia cena v porovnaní s monokulárnom kamerou.

V tejto práci je vybraná monokulárna kamera z dôvodu cenovej dostupnosti a tiež preto, že riešenie tejto úlohy nepožaduje znalosť skutočných rozmerov objektov v scéne. Pohyb kamery bude musieť byť prispôsobený tak, aby bola medzi dvoma pohľadmi vždy dostatočná disparita/paralaxa a je znázornený na obrázku 3.1.

4.1.2 Kalibrácia kamery

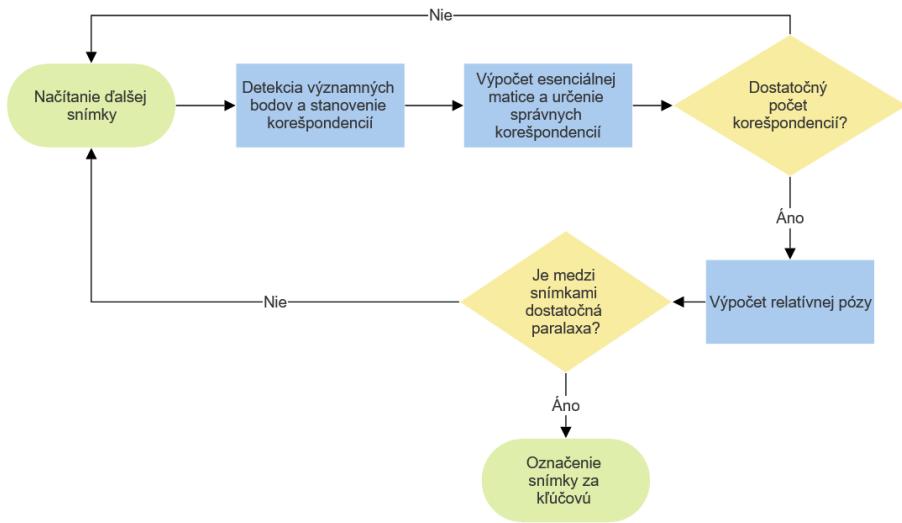
V tejto práci je pred započatím rekonštrukcie kamera skalibrovaná - sú určené jej vnútorné parametre ako ohnisková vzdialenosť, stred obrazovej roviny v pixelových súradničiach a koeficienty, kompenzujúce súdkovité skreslenie, ktoré kamera generuje. Tieto parametre potom tvoria kalibračnú maticu \mathbf{K} . Kalibrácia kamery prebieha podľa metódy opísanej v kapitole 2.3. Znalosť vnútorných parametrov kamery zjednoduší výpočet pózy, ktorú môžeme určiť z esenciálnej matice. Tá vyžaduje 5 bodových korešpondencií, oproti fundamentálnej matici pre neskalibrovaný prípad, ktorá vyžaduje 8 bodov. Čím menej bodov je potrebných, tým je väčšia šanca, že sa medzi nimi nebudú nachádzať chybné korešpondencie a algoritmus RANSAC bude potrebovať menší počet iterácií.

4.1.3 Pohyb kamery

V tejto časti sú určované pózy jednotlivých pohľadov pri pohybe kamery scénou. Vstupná množina snímok je spracovávaná sekvenčne a sú z nej vyberané iba určité tzv. klúčové snímky, s ktorými sa bude v 3D rekonštrukcii ďalej pracovať, ostatné sa zahodia. Zachytenie scény pomocou videa má tú výhodu, že je k dispozícii veľké množstvo snímok, ktoré sa prekrývajú a sú redundantné - ak je napríklad daná

snímka v zlej kvalite, tak sa za ňou nachádzajú ďalšie snímky, ktoré ju môžu rovnomenne nahradíť.

Určenie pózy/orientácie prvých dvoch snímok/pohľadov je založené na princípe relatívnej orientácie z kapitoly 2.5, teda póza druhého pohľadu je určená relatívne k prvému pohľadu, ktorý je umiestnený do bodu $(0, 0, 0)^T$ v počiatku súradnicového systému s jednotkovou maticou rotácie a táto snímka je označená ako klúčová. Určenie relatívnej orientácie druhého pohľadu je zachytené ná diagrame v obrázku 4.2.



Obr. 4.2: Relatívna orientácia prvých dvoch pohľadov

Detekcia významných bodov a stanovenie korešpondencií. Cieľom je určiť body v snímke, ktoré majú výrazné fotometrické vlastnosti a sú dobre odlíšiteľné. Môže ísť o body s vysokým kontrastom alebo v najzákladnejšom prípade to môžu byť rohy objektov v obraze. Významné body, označované aj ako lokálne príznaky, musia byť detekované a popísané tak, aby ich bolo možné vo viacerých snímkach nezávisle na geometrických a fotometrických deformáciách znova objaviť a označiť ich za rovnaké body. Lokálne príznaky by teda mali splňať podmienku opakovateľnosti a mali by byť invariantné voči posunutiu, rotácii, zmene merítka a do istej miery voči zmene perspektívy a osvetlenia. Určenie lokálnych príznakov v obraze je zložené z dvoch krokov: detekcie a popisu. V prvom kroku je analyzovaný celý obraz a sú v ňom nájdené body, ktoré majú výrazné fotometrické vlastnosti. V druhom kroku je okolie stanovených významných bodov popísané príznakovým vektorom tzv. *descriptor*, ktorý je invariantný voči rigidným transformáciám. Každý významný bod je teda reprezentovaný svojou polohou v obraze a príznakovým vektorom. Stanovovanie korešpondencií medzi bodmi z rôznych snímok sa vykonáva pomocou porovnania

ich príznakových vektorov pomocou zvolenej metriky, napríklad euklidovská vzdialenosť. Existuje viacero detektorov lokálnych príznakov, najznámejšími sú SIFT, SURF, Harrisov detektor rohov, ORB, MSER, FREAK atď. Líšia sa medzi sebou v stupni invariantnosti voči rôznym transformáciám, v type príznakov (môže íst o rohy alebo kvapky tzv. *blobs*), v rýchlosťi detekcie alebo vo veľkosti príznakového vektoru. Pre detekciu lokálnych príznakov je v tejto práci vybraná metóda SURF - ako detektor významných bodov a rovnako aj ako príznakový vektor.

SURF príznaky sú invariantné voči posunutiu, zmene merítka, rotácií a do istej miery voči zmenám miesta snímania a zmenám v osvetlení, čo je v tejto práci potrebné. Oproti SIFT príznakom sú menej robustnejšie, no detekcia príznakov a stanovovanie korešpondencií je oproti SIFT oveľa rýchlejšie, čo bol hlavný dôvod pre výber SURF metódy. Počet detekovaných významných bodov je u metódy SURF väčší než u SIFT. [48]

Detekcia významných bodov prebieha na snímkach, v ktorých je kompenzované súdkovité skreslenie. Detekované príznaky/body na aktuálnej snímke sú porovnané s príznakmi na predošlej a sú medzi nimi stanovené korešpondencie pomocou porovnania ich príznakových vektorov. V týchto prvotných korešpondenciách sa väčšinou nachádzajú aj nesprávne určené dvojice, ktoré je potreba odfiltrovať.

Výpočet esenciálnej matice a určenie správnych korešpondencií. V tomto kroku sa na základe vstupných bodových korešpondencií vypočíta esenciálna matica. Na výpočet esenciálnej matice je použitý 5-bodový algoritmus v kombinácii s metódou RANSAC, ktorej výstupom je esenciálna matica a množina správnych korešpondencií, ktoré podporili výpočet modelu (esenciálnej matice). Zo vstupných bodových korešpondencií sme v tomto kroku odstránili chybné korešpondencie a vypočítali esenciálnu maticu.

Výpočet relatívnej pózy. Relatívna póza druhého pohľadu, matica rotácie (3 parametre) a smerový vektor posunutia (2 parametre), je odvodená z esenciálnej matice, podľa princípu popísanom v kapitole 2.8. Pomocou množiny správnych bodových dvojíc, ktoré slúžia ako kontrolné body, je vybraná zo 4 možných konfigurácií relatívnej pózy tá, v ktorej sa najväčší počet bodov zobrazí pred oboma kamerami.

Uhol paralaxy medzi dvoma pohľadmi Ako už bolo spomenuté v kapitole 2.10, veľkosť uhlu paralaxy určuje ako kvalitná je triangulácia. V tomto kroku je vypočítaný uhol paralaxy pre každý 3D bod rekonštruovaný z korešpondujúcich 2D bodových dvojíc medzi aktuálnou a predošlou snímkou. Ak je uhol paralaxy dostačne veľký, znamená to, že medzi pohľadmi je dostatočná horizontálna vzdialenosť a aktuálna snímka sa označí ako klúčová - bude teda vo výbere snímok, ktoré budú reprezentovať scénu a pomocou ktorých sa bude rekonštruovať 3D model. Ak aktuálna snímka nevyhovuje minimálnemu uhlu paralaxy, zahodí sa a prejde sa vo vstupnej sekvencii na ďalší pohľad. To, či je medzi dvoma pohľadmi dostatočné ho-

rizontálne posunutie, sa tiež môže určiť zo vzdialenosťi korešpondujúcich 2D bodov v pixelových súradničiach medzi dvoma pohľadmi. Pixelová vzdialosť medzi bodmi je preto použitá ako doplňujúca podmienka.

Takto je určená orientácia/póza prvých dvoch pohľadov - na základe 2 snímok a ich korešpondujúcich 2D bodov bez znalosti 3D scény. Je treba ale bráť ohľad na to, že vektor translácie druhého pohľadu vyjadruje iba smer, kde daná kamera leží a merítka/veľkosť tohto vektoru je neznáma, ktorú nie je možné odvodiť iba na základe 2D informácie, kapitola 2.5. Problém neurčitého merítka translácie sme schopní do istej miery vyriešiť, ak máme k dispozícii polohu 3D bodov scény. Tieto 3D body si môžeme triangulovať z predošlého páru snímok. Potom môžeme orientáciu daného pohľadu určiť pomocou metódy *Projective 3-Point*, kapitola 2.4, na základe korešpondencií 3D bodov a 2D bodov v danom pohľade, ktorý scénu pozoruje. Vektor translácie už teda nie je iba smerový vektor, ale má aj definovanú veľkosť, ktorá je odvodnená z triangulovaných 3D bodov. Skutočné merítko scény a translačných vektorov jednotlivých pohľadov (napríklad v metroch), ale známe nie je, pretože je použitá monokulárna kamera bez externých informácií o scéne (vzdialosť medzi dvoma bodmi v metroch alebo poloha kamery zo systému GPS) 2.4.

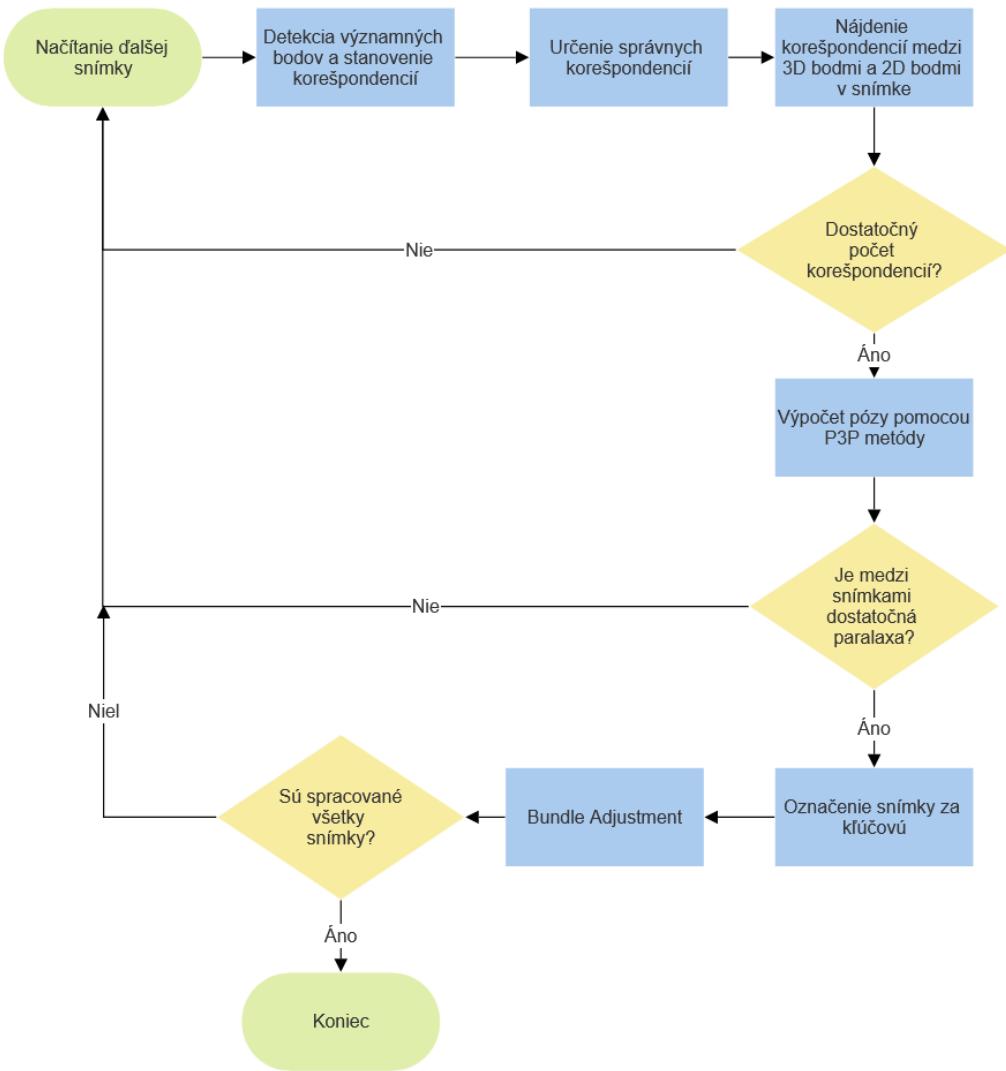
Z prvých dvoch pohľadov sú teda triangulované 3D body, pomocou ktorých sa algoritmom *Projective 3-Point* určí póza tretieho pohľadu. Výpočet pózy tretej a zvyšných klúčových snímok je zobrazený na diagrame v obrázku 4.3. Postup spracovania snímok je podobný ako v prípade prvých dvoch pohľadov, odlišné časti sú popísané v nasledujúcich odstavcoch.

Určenie správnych korešpondencií. Prebieha podobne ako v prípade relatívnej orientácie - množinu správnych korešpondencií určujeme pomocou algoritmu RANSAC, kde je ako model použitý esenciálna matica. Zaujíma nás ale už len množina správnych korešpondencií a esenciálnu matiku už ďalej nevyužívame.

Nájdenie korešpondencií medzi 3D bodmi a 2D bodmi v snímke. Majme trojicu po sebe idúcich snímok I_{i-2} , I_{i-1} a I_i , kde I_i je aktuálna snímka. Z dvojice snímok I_{i-2} a I_{i-1} sú triangulované 3D body a následne sú nájdené korešpondencie medzi týmito 3D bodmi a 2D bodmi detektovanými na snímke I_i pomocou 2D korešpondencií medzi pohľadmi I_{i-1} a I_i z predošlého kroku.

Výpočet pózy pomocou P3P metódy. Prebieha podľa princípu popísanom v kapitole 2.4. V rámci metódy *Projective 3-Point* je implementovaný algoritmus RANSAC, ktorý ďalej filtriuje predložené 3D-2D korešpondencie.

Bundle adjustment. Po pridaní pózy každého nového klúčového pohľadu/snímky je daná aktuálna póza a všetky doterajšie pózy optimalizované technikou *Bundle adjustment* popísanou v kapitole 2.11, na základe 3D bodov triangulovaných zo všetkých predošlých klúčových snímok a korešpondencií medzi nimi a 2D bodmi v daných snímkach.

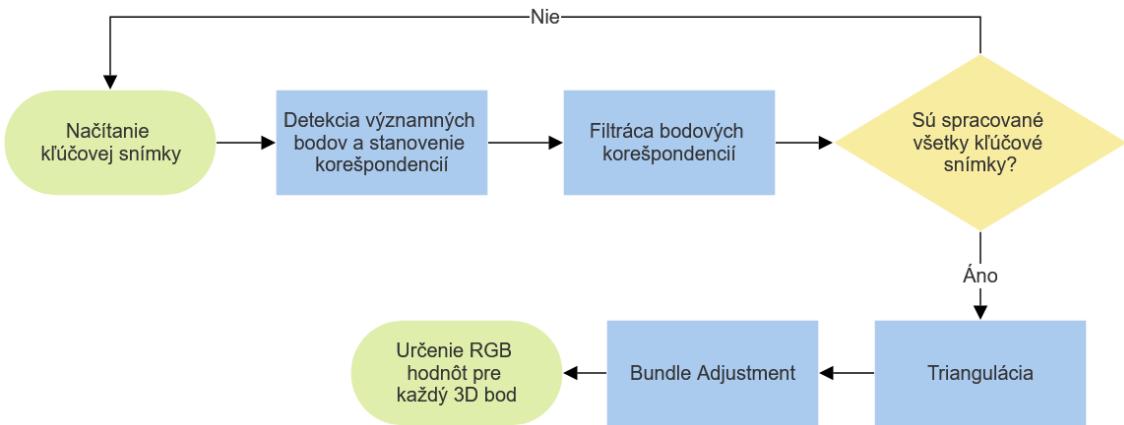


Obr. 4.3: Určenie pózy zvyšných pohľadov pomocou $P3P$ algoritmu

4.1.4 3D štruktúra

Vstupmi pre túto časť 3D rekonštrukcie sú klúčové snímky a ich odpovedajúce pózy. Na základe nich rekonštruujeme/triangulujeme výslednú 3D štruktúru scény - výstupom tejto časti je 3D mračno bodov. Postup rekonštrukcie 3D mračna bodov vyjadruje diagram na obrázku 4.4.

Detektia významných bodov a stanovenie korešpondencií. V tomto bode sú znova detekované významné body na klúčových snímkach, no s tým rozdielom, že parametre SURF detektora sú nastavené tak, aby sme obdržali maximálny počet detekovaných bodov na obrázku. To isté platí aj v stanovovaní korešpondencií medzi snímkami - parametre porovnávania príznakových vektorov sú nastavené tak, aby sme dostali čo najviac počiatočných bodových korešpondencií. Je to preto, aby



Obr. 4.4: Postup tvorby mračna bodov

výsledné mračno bodov bolo čo najhustejsie.

Mračná bodov, ktoré sú rekonštruované na základe významných bodov resp. príznakov ako SURF, SIFT apod. (čo je prípad tejto práce) sa označujú ako riedke mračná, pretože vznikajú iba z určitej podmnožiny bodov v obrazu. Naproti tomu sú husté mračná bodov, ktoré vznikajú za použitia techniky viac-pohľadového stereo (*multi-view stereo*), väčšinou s použitím kamery v stereo konfigurácii. Vtedy je mračno bodov rekonštruované zo stereo párov snímok a výsledné 3D mračno bodov vzniká pomocou všetkých pixelov v obrázku, čím dostaneme hustejšiu reprezentáciu 3D scény.

Korešpondencie sú stanovované medzi aktuálnou klúčovou snímkou a n nasledujúcimi klúčovými snímkami. Čím je interval týchto snímok väčší, t.j. daný bod je pozorovaný z viacerých pohľadov, tým je triangulácia presnejšia, no zároveň narastá aj čas potrebný na rekonštrukciu.

Filtrácia bodových korešpondencií. Filtrácia počiatočných korešpondencií zabezpečuje, že vyberieme iba najkvalitnejšie body resp. bodové dvojice, z ktorých sa bude triangulovať výsledné mračno bodov. Filtrácia prebieha v troch krokoch. V prvom kroku je použitý RANSAC algoritmus v kombinácii s 5-bodovým algoritmom a esenciálou maticou. Zaujíma nás iba množina vyfiltrovaných správnych bodových dvojíc. V druhom kroku sú prebrané korešpondencie z predošlého kroku a medzi odpovedajúcimi bodmi z dvoch snímok sa počíta ich euklidovská vzdialenosť v pixelových súradničiach. Tie bodové dvojice, ktorých vzdialenosť sa príliš líši od priemernej vzdialnosti všetkých korešpondencií, sú zahodené. V poslednom kroku sú zvyšné bodové dvojice filtrované podľa veľkosti uhlu paralaxy. Korešpondencie s malým uhlom paralaxy sú zahodené.

Triangulácia. Na základe pôz a bodových korešpondencií medzi klúčovými snímkami je vypočítaná pozícia daného bodu v 3D prostredí.

kami sú triangulované 3D body, ktoré sú následne filtrované na základe veľkosti chyby reprojekcie (kapitola 2.11) a znamienka ich z súradnice (body zobrazené za pohľadom kamery).

Bundle adjustment. Pomocou tohto algoritmu sú optimalizované polohy 3D bodov v mračne. Optimalizácia pôz daných pohľadov nás už nezaujíma.

Určenie RGB hodnôt pre každý 3D bod. Každý 3D bod má náváznosť na konkrétnu snímku a konkrétny bod v snímke, z ktorého bol rekonštruovaný. Pomocou toho určíme farbu každého bodu v 3D modeli.

4.1.5 Odstránenie šumu z mračna bodov

V počiatokom mračne bodov sa môžu stále vyskytovať chybné alebo nepresné dátá, ktoré môžeme odhaliť štatistickou analýzou. Chybné body identifikujeme ako body, ktoré sa nachádzajú vo veľkej vzdialosti od svojich najbližších susedov v porovnaní s priemerom pre celé mračno bodov [49]. Zvolíme veľkosť najbližšieho okolia n , z ktorého sa pre každý bod mračna vypočíta priemerná vzdialenosť \bar{l} od jeho n najbližších bodov:

$$\bar{l} = \frac{1}{n} \sum_{i=1}^n l_i \quad (4.1)$$

Priemer pre celé mračno bodov s početnosťou o potom bude:

$$\bar{l}_{mean} = \frac{1}{o} \sum_{i=1}^o \bar{l}_i \quad (4.2)$$

Pomocou \bar{l}_{mean} je určená hodnota σ a pre každý bod je testované či platí:

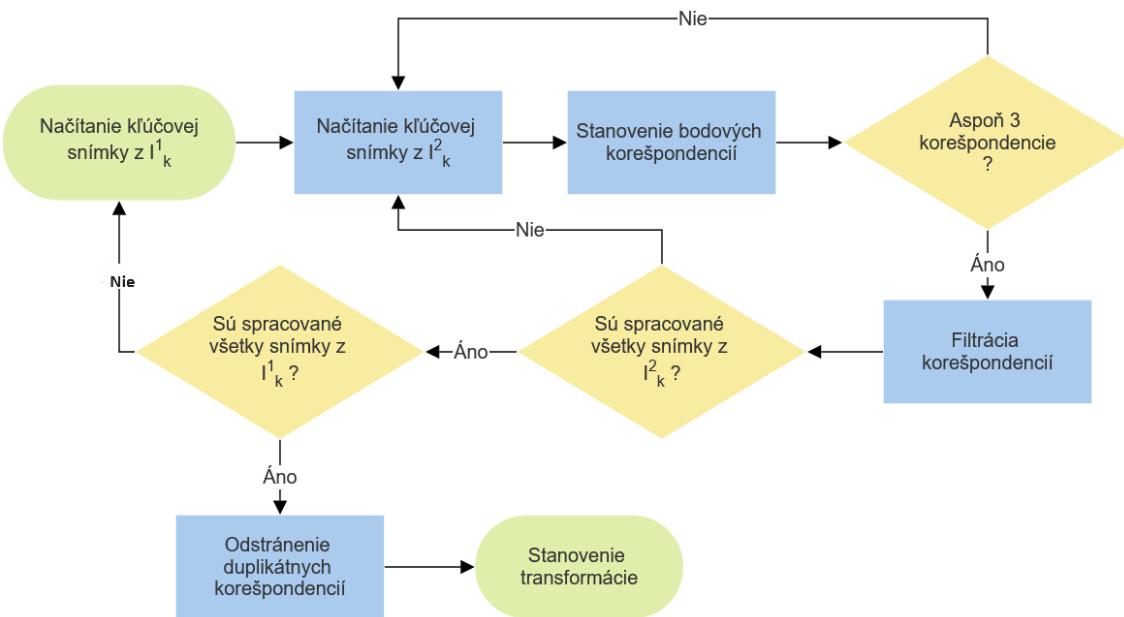
$$\bar{l} < \sigma \quad (4.3)$$

Bod, ktorý túto podmienku nesplní sa označí ako chybný a z mračna bodov sa odstráni. Čím väčšia je hodnota okolia n , tým väčšia je robustnosť filtra voči šumu. Čím menšia je hodnota prahu σ , tým bude filtrovanie bodov agresívnejšie [49].

4.2 Registrácia mračien bodov

Vstupom pre tento modul sú dve mračná bodov M_1 a M_2 . M_1 označme ako referenčný 3D model - tento v našom prípade reprezentuje pôvodnú trasu a M_2 reprezentuje lokálny úsek pôvodnej trasy obsahujúci geometrické zmeny. Ďalším vstupom je množina klúčových snímok I_k^1 (z ktorých vzniklo mračno M_1) a I_k^2 (z ktorých vzniklo mračno M_2) a 2D body, z ktorých boli obe mračná rekonštruované. Cieľom tejto kapitoly je popísť spôsob, akým sa lokálne mračno M_2 lokalizuje/umiestni v rámci globálneho mračna bodov pôvodnej trasy M_1 . Výstupom je teda podobnostná transformácia, ktorá prevedie mračno M_2 do súradnicového systému M_1 .

Hľadaná podobnostná transformácia sa počíta z korešpondencií 3D bodov medzi M_1 a M_2 - kapitola 2.12. Nájdenie týchto korešpondencií je pre úspešnú registráciu mračien klúčové, pretože ak už máme určené spoľahlivé bodové korešpondencie, výpočet hľadanej transformácie je jednoduchý. Pri určovaní 3D bodových korešpondencií využijeme to, že každý 3D bod odpovedá 2D SURF príznaku/bodu v klúčovej snímke a informácia o náväznosti 3D bodov na 2D príznaky v klúčových snímkach je známa. Označme P_1 a P_2 ako množiny 2D SURF bodov, ktoré sa nachádzajú v klúčových snímkach I_k^1 a I_k^2 a odpovedajú 3D bodom mračien M_1 a M_2 . Stanovením korešpondencií medzi množinami 2D bodov P_1 a P_2 , napríklad porovnaním ich príznakových vektorov, zároveň obdržíme aj korešpondencie medzi 3D bodmi mračien M_1 a M_2 , na základe ktorých potom môžeme vypočítať hľadanú podobnostnú transformáciu. Postup stanovenia korešpondencií a výpočtu transformácie je zobrazený na diagrame v obrázku 4.5.



Obr. 4.5: Návrh registrácie mračien bodov

Stanovenie bodových korešpondencií. Každá klúčová snímka z množiny I_k^2 je porovnaná s každou klúčovou snímkou množiny I_k^1 - v každej iterácii sú stanovené korešpondencie medzi SURF príznakmi z P_1 a z P_2 (pomocou porovnania ich príznakových vektorov), ktoré sú pozorované daným párom snímkov z I_k^1 a z I_k^2 . Korešpondencie sú teda stanovované iteratívne po jednotlivých snímkach a na konci každej iterácie sú zhromaždené do jednej matice korešpondencií C .

Filtrácia korešpondencií. V predložených korešpondenciách z predošlého bodu sa môžu stále vyskytovať falosné bodové dvojice, ktoré je treba filtrovať. V každej

iterácií je teda na základe predložených korešpondencií vypočítaná podobnostná transformácia T_{ij} medzi danými 3D bodmi, tak ako je ukázané v kapitole 2.12. Aktuálne skúmané body z M_2 sú transformované pomocou T_{ij} a ich vzdialenosť od bodov z M_1 by mala byť minimálna. Tie bodové dvojice, ktorých vzdialenosť je vyššia než určitá zvolená hodnota sa zahodia.

Odstránenie duplikátnych korešpondencií. Po spracovaní všetkých kľúčových snímok a zhromaždení všetkých čiastkových korešpondencií do matice \mathbf{C} , sa z nej v tomto kroku odstránia duplikátne bodové dvojice.

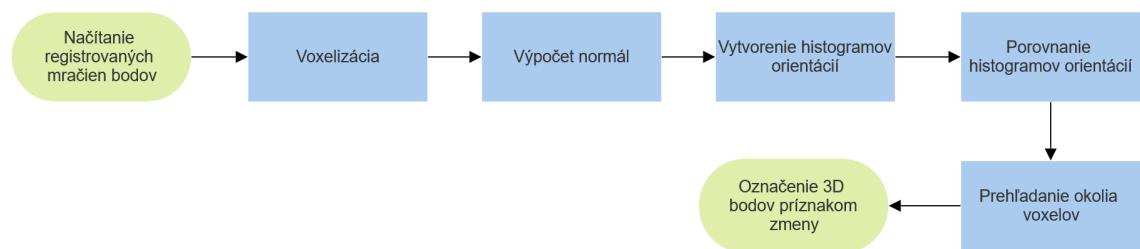
Stanovenie transformácie. Na záver je z daných korešpondencií vypočítaná finálna podobnostná transformácia T , ktorá sa aplikuje na mračno M_2 :

$$M'_2 = T(M_2) \quad (4.4)$$

ktoré je týmto prevedené do súradnicového systému mračna M_1 resp. úsek trasy reprezentovaný mračnom M_2 je lokalizovaný v pôvodnej trase reprezentovanej mračnom bodov M_1 .

4.3 Detekcia zmien

V tejto kapitole je popísaný spôsob, akým sú detektované zmeny medzi dvoma registrovanými mračnami bodov, ktoré sú vstupmi pre túto časť systému. Výstupom tejto časti je binárny vektor, ktorý každému bodu z oboch mračien priradí príznak 'zmena' alebo 'žiadna zmena'. Postup je zobrazený na diagrame v obrázku 4.6.



Obr. 4.6: Návrh detekcie zmien medzi registrovanými mračnami bodov

Voxelizácia. Priestor, v ktorom sa nachádzajú registrované mračná bodov je segmentovaný do voxelovej mriežky podľa kapitoly 2.14. Ohraničenie vyšetrovaného priestoru je dané hranicami mračna bodov M_2 , pretože M_2 je výsekom z mračna bodov M_1 , ktoré odpovedá pôvodnej trase. Okraje voxelovej mriežky sú vyplnené prázdnymi voxelmi z dôvodu prehľadávania 26-okolia každého voxelu, ktoré prebehne v jednom z nasledujúcich krovov. Následne sa zo všetkých voxelov vyberú tie, ktoré sú obsadené 3D bodmi a prázdne voxel sa v ďalších operáciách už neuvažujú.

Výpočet normál. Pre každý 3D bod z oboch mračien bodov je vypočítaná povrchová normála. Znamienko každej normály je určené na základe pózy pohľadu, z ktorého bol 3D povrch nasnímaný. Tento pohľad je vybraný ako prvý pohľad zo sekvencie kľúčových snímok/pohľadov a nachádza sa v počiatku súradnicového systému.

Vytvorenie histogramov orientácií. Vypočítané povrchové normály z predošlého kroku sú prevedené do sférických súradníc. Každá normála je teda vyjadrená dvojicou uhlov θ (sklon) a ϕ (azimut). Uhol θ je v rozsahu $[-90^\circ, 90^\circ]$ a uhol ϕ v rozsahu $[-180^\circ, 180^\circ]$. Pre skupinu normál, ktoré sa nachádzajú vo voxelu je vytvorený 2D histogram, ktorý má 8 košov pre uhol θ a 4 koše pre uhol ϕ - obe intervaly uhlov sú teda rozdelené po 45 stupňoch a každá normála podľa svojej orientácie spadá do jedného koša 2D histogramu. Vrcholy histogramu odpovedajú počtu normál/bodov, ktoré majú danú orientáciu. Takýto 2D histogram sa v tejto práci nazýva histogram orientácie a vyjadruje tvar/orientáciu povrchu zloženého z 3D bodov, ktoré sa nachádzajú v konkrétnom voxelu. Pre každý voxel sú určené 2 histogramy orientácií - jeden pre povrch odpovedajúci mračnu M_1 a druhý pre povrch odpovedajúci M_2 s tým, že energia histogramov sa rovná počtu bodov z daného mračna, ktoré sa vo voxelu nachádzajú.

Porovnanie histogramov orientácií. Z predošlého kroku máme pre každý voxel určenú dvojicu histogramov orientácií, ktorá reprezentuje dvojicu povrchov z mračien M_1 a M_2 v danom voxelu. Porovnaním tejto dvojice histogramov zistíme v každom voxelu mieru podobnosti medzi danými dvoma povrchmi. Ak je miera podobnosti vysoká, ide o rovnaké resp. rovnako orientované povrhy a môžeme usúdiť, že v danom voxelu k zmene nedošlo a naopak. Porovanie histogramov je v tomto prípade špecifické v tom, že histogramy nemajú rovnakú energiu (súčet vrcholov), pretože hustota bodov sa môže medzi mračnami lísiť. Ak by sme napríklad porovnávali histogramy pomocou absolútneho rozdielu metódou diel po dieli (*bin-by-bin*) a vyšetrovali by sme dva povrhy, ktoré majú rovnakú orientáciu/tvar, ale jeden z nich je zložený z väčšieho počtu bodov, mohli by sa tieto povrhy označiť ako rozdielne. Rozdielna hustota je jednou z hlavných nevýhod mračien bodov, čo je uvedené v kapitole 1.2.1 a preto by metóda ich porovnania mala byť nezávislá na počte bodov. Histogramy sú preto normalizované tak, aby ich energia bola rovná 1, čo znamená, že jednotlivé vrcholy majú výšku od 0 po 1:

$$h_n(i) = \frac{h(i)}{\sum_{i=1}^k h(i)} \quad (4.5)$$

$h(i)$ predstavuje pôvodný histogram a $h_n(i)$ normalizovaný histogram a oba majú rovnaký počet košov k . Tým sme dosiahli to, že histogramy majú rovnakú energiu a môžeme ich medzi sebou porovnať. Normalizované histogramy p a q sú porovnávané

metódou prieniku histogramov, ktorá patrí k metódam typu diel po dieli (*bin-by-bin*):

$$s(p, q) = \sum_{i=1}^k \min(p(i), q(i)) \quad (4.6)$$

kde $s(p, q)$ vyjadruje mieru podobnosti daných histogramov/povrchov. Metóda diel po dieli je použitá preto, lebo chceme porovnávať iba tie vrcholy v histograame, ktoré si priamo odpovedajú. Vrcholy s rozdielnymi indexami zrovnávať nechceme (metódy kríženia dielov - *cross bin*), pretože vrcholy na rôznych indexoch znamenajú rôzne orientácie povrchu. Histogramy preto môžu mať rovnaký geometrický tvar, ale ak sú ich vrcholy na iných pozíciah znamená to, že dané povrhy majú úplne odlišnú orientáciu. $s(p, q)$ teda vyjadruje mieru podobnosti povrchov a je v intervale $[0, 1]$. $s(p, q) = 0$ znamená, že povrhy sú maximálne rozdielne a $s(p, q) = 1$ značí, že povrhy sú identické. Ďalej je zvolená hodnota s_{min} , ktorá značí minimálnu podobnosť a je prahom tejto klasifikácie. Ak je $s(p, q) < s_{min}$, znamená to, že dané povrhy sú rozdielne a voxel je označený ako 'zmena', v opačnom prípade je voxel označený ako 'žiadna zmena'. Ak sa v danom voxelí nachádzajú body iba z jedného mračna bodov (čo znamená, že v scéne nejaký povrch zmizol alebo pribudol) automaticky nastavujeme $s(p, q) = 0$ a voxel je tým pádom označený príznakom 'zmena'. V prípadoch, kedy sa vo voxeloch stretávajú dva povrhy resp. nachádzajú sa tam body z oboch mračien, už musíme porovnať ich histogramy.

Prehľadanie okolia voxelov. Z predchádzajúceho kroku sme všetky vyšetrované voxelé prešli a označili ich príznakom 'zmena' alebo 'žiadna zmena' na základe porovnania histogramov orientácií. V tomto kroku prechádzame každý voxel znova a zaujíma nás jeho 26-okolie, kapitola 2.14. V tomto okolí sledujeme, aké majú susedné voxelé hodnoty príznakov zmeny. Zhluk týchto 27 voxelov označíme ako supervoxel. Ďalej si vytvoríme 3D maticu, ktorá odpovedá rozmerom zavedenej voxelovej mriežky. Ak je supervoxel zložený iba z voxelov označených ako 'zmena', prepíšeme supervoxel s danými príznakmi do odpovedajúceho miesta v novej 3D matici. Po prejdení všetkých voxelov sú zvyšné voxelé, ktoré neboli prepísané do novej 3D matice (neboli súčasťou predošlých supervoxelov 'zmena'), označené ako 'žiadna zmena'. Týmto spôsobom nájdeme prepojené zhluky voxelov, ktoré sa zhodujú s príznakom 'zmena'. Voxely, ktorých okolie má zmiešané hodnoty príznakov zmeny, sú teda definitívne označené ako 'žiadna zmena'. Toto prehľadávanie lokálneho okolia voxelov má zmysel v tom, že klasifikácia voxelov z predošlého kroku nie je dostatočne presná. Avšak spoliehame sa na to, že klasifikátor v predošлом bode je dostatočne presný pri odhalovaní skutočných zmien, resp. že skutočne rozdielne povrhy označí ako 'zmenu' tzv. *true positive* a v povrchoch, ktoré sú skutočne identické a má ich označiť príznakom 'žiadna zmena' tzv. *true negative*, má priestor na to sa zmýliť.

5 Realizácia systému detektie zmien

V tejto kapitole je opísaná realizácia systému detektie zmien na základe návrhu z predosnej kapitoly vrátane praktickej ukážky, v ktorej budú postupne demonštrované všetky časti systému na vybranej dvojici porovnávaných trás.

Systém detektie zmien bol implementovaný v jazyku *MATLAB* vo verzii *R2022b* za použitia knižníc *Computer Vision Toolbox*, *Image Processing Toolbox* a *Statistics and Machine Learning Toolbox*.

5.1 Výber snímača

Vybraná monokulárna kamera je SENCOR 3CAM 2002 - obrázok 5.1 a jej parametre, s ktorými boli nasnímané dáta sú v tabuľke 5.1. Pozorovací uhol 120°



Obr. 5.1: Použitá monokulárna kamera SENCOR 3CAM

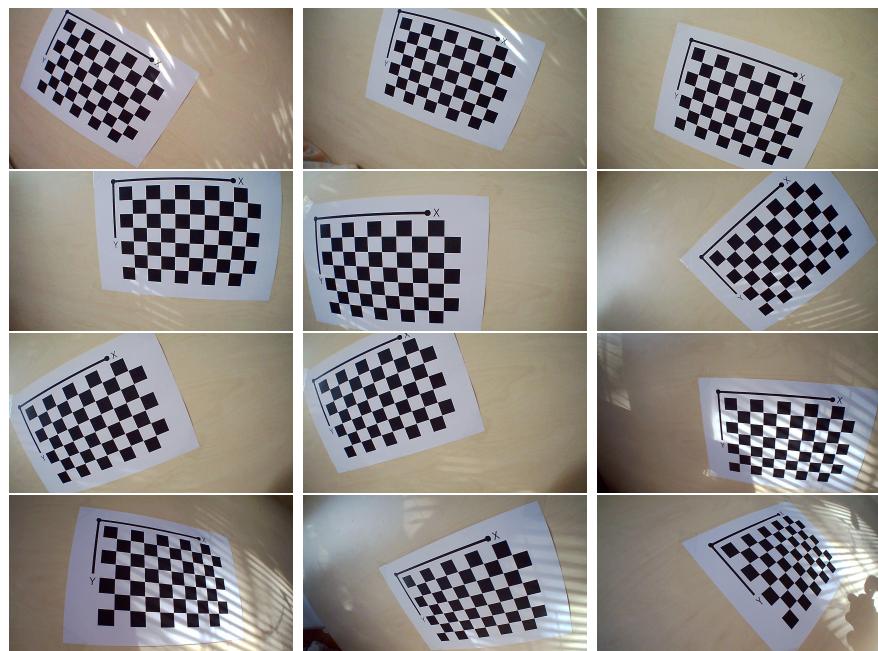
Tab. 5.1: Parametre kamery

Rozlíšenie	FPS	Zorný uhol	Formát videa
1280x720 pixelov	30	120°	avi

generuje v obraze súdkovité skreslenie, zobrazené na obrázkoch 2.8 a 2.9 a je nutné ho kompenzovať, čomu sa venuje kalibrácia kamery.

5.2 Kalibrácia kamery

Kamera je skalibrovaná pomocou *MATLAB* aplikácie *Camera Calibrator* metódou, ktorej princíp je popísaný v kapitole 2.3. Kalibračný vzor šachovnice bol umiestnený na rovný pevný povrch a nasnímaný dostatočným počtom snímok (aspoň 20), obrázok 5.2. Snímky boli zaznamenané tak, aby sa šachovnica nachádzala v rôznych častiach obrázku. Je to preto, aby sa zachytilo súdkovité skreslenie ktoré sa zvyšuje od stredu k okrajom snímky a v rámci obrázku nie je rovnomerne rozložené. Ďalej boli snímky zaznamenané tak, aby šachovnica zaberala aspoň 20% z obrázku a aby bola nasnímaná z rôznych uhlov, nie však viac než 45° medzi obrazovou rovinou kamery a šachovnicou, podľa odporúčaní v *MATLAB* dokumentácii [50]. Po



Obr. 5.2: Rôzne zachytené kalibračné vzory

pridaní snímok sú nastavené parametre kalibrácie: model kamery je nastavený na štandardný model kamery a počet parametrov súdkovitého skrenenia (kapitola 2.3) je nastavený na 3. Postupom popísaným v kapitole 2.3 sú určené súradnice stredu obrazovej roviny, ohnisková vzdialenosť a parametre súdkovitého skreslenia. Parameter skosenia a tangenciálne skreslenie v našom prípade neuvažujeme. Úspešnosť kalibrácie môže byť vyhodnotená pomocou priemernej chyby reprojekcie (kapitola 2.11) v každej snímke a snímky s veľkou chybou reprojekcie môžu byť z procesu kalibrácie vyradené. V tomto prípade bola veľkosť priemernej chyby reprojekcie 0,23 pixelov, čo je akceptovateľné.

5.3 Ukážka riešenia

Funkcionalita systému detektie zmien (jeho bloková schéma je na obrázku 4.1) je implementovaná v piatich moduloch, ktoré odpovedajú návrhu z predošej kapitoly 4: 1) Pohyb kamery - kapitola 4.1.3, 2) 3D štruktúra - kapitola 4.1.4, 3) Odstránenie šumu z mračna bodov - kapitola 4.1.5, 4) Registrácia mračien bodov - kapitola 4.2 a 5) Detekcia zmien - kapitola 4.3. V tejto kapitole bude demonštrovaný prechod všetkými piatimi modulmi pre vybranú dvojicu testovacích trás.

Na obrázku 5.3 je nasnímaná pôvodná globálna trasa č. 109 a na obrázku 5.4 je nasnímaná aktuálna lokálna trasa č. 110, ktorá tvorí úsek pôvodnej trasy a obsahuje zmeny, konkrétnie ide o zámennu ovládača a červenej tuby. Moduly 1), 2) a 3) sú pre obe trasy rovnaké, preto sa prejdú iba pre jednu z nich, konkrétnie pre pôvodnú trasu, ktorá je na obrázku 5.3.



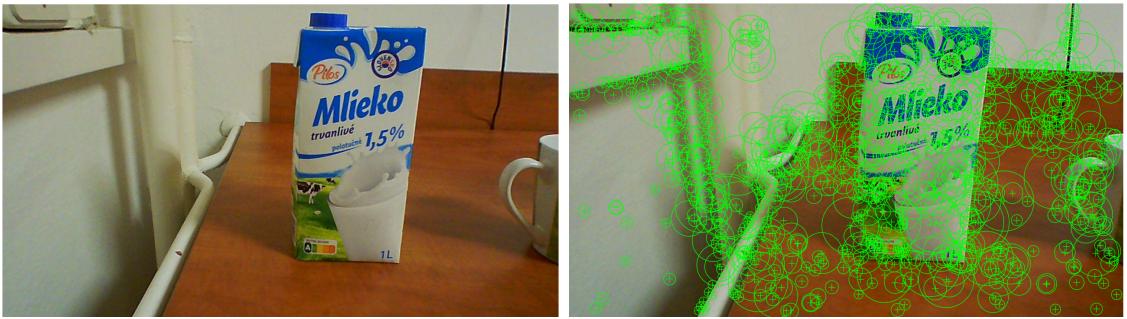
Obr. 5.3: Pôvodná testovacia trasa č. 109



Obr. 5.4: Aktuálna testovacia trasa č. 110

5.3.1 Pohyb kamery

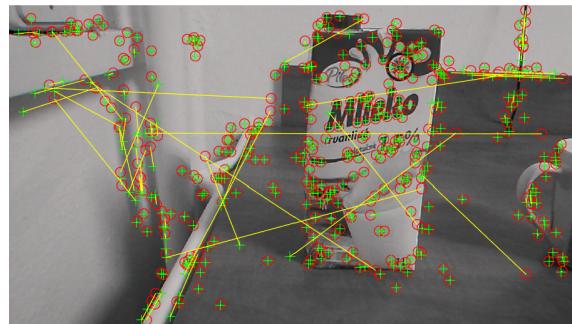
Tento modul odpovedá v rámci kapitoly návrhu algoritmov kapitole 4.1.3 a bude demonštrovaný pre pôvodnú globálnu trasu, ktorá je na obrázku 5.3.



Obr. 5.5: Detekované SURF významné body na prvej snímke

Detekcia SURF významných bodov je zachytená na obrázku 5.5. Počet detektovaných SURF bodov môžeme ovplyvniť hodnotou prahu, označme ju ako S . Čím je S menšie, tým viac bodov v obrazu detekujeme a naopak. Z týchto bodov je následne vybraná podmnožina bodov (nastavená na tretinu), ktoré sú rovnomerne rozložené v obrazu. Čím viac rovnomerne rozložených bodov je na snímke detektovaných, tým je výpočet pózy robustnejší výmenou za časovú náročnosť.

Prvotné korešpondencie medzi aktuálnou a predošlou snímkou sú stanovené pomocou porovnania príznakových vektorov daných SURF bodov. Tieto korešpondencie zachytáva obrázok 5.6, z ktorého je vidieť, že niektoré korešpondencie sú chybné. Chybné korešpondencie sú filtrované pomocou epipolárnej geometrie, resp. výpoč-



Obr. 5.6: Prvotné korešpondencie medzi prvou a druhou snímkou

tom esenciálnej matice v kombinácii s *RANSAC* algoritmom. Správne korešpondencie bez falošných asociácií sú zobrazené na obrázku 5.7.

Na základe týchto korešpondencií je vypočítaná relatívna póza druhého pohľadu a ďalej je testované, či spĺňa podmienku minimálnej uhlovej paralaxy. Minimálny uhol paralaxy je nastavený na hodnotu 1° a pixelová vzdialenosť na 15 pixelov. Ak všetky bodové dvojice majú uhol paralaxy a pixelovú vzdialenosť väčšiu než nastavený prah, daná snímka je označená ako kľúčová. Čím menšie sú prahové hodnoty



Obr. 5.7: Filtrované korešpondencie medzi prvou a druhou snímkou

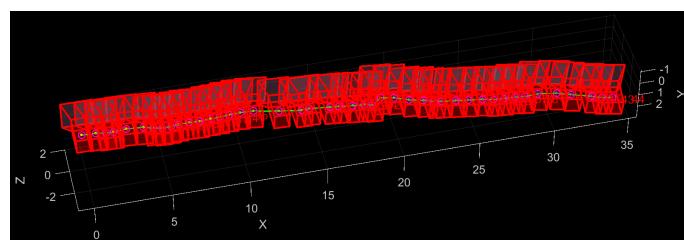
pre minimálny posuv/paralaxu medzi dvoma po sebe idúcimi snímkami, tým viac snímok sa označí za klúčové. Počet klúčových snímok alebo frekvencia, s ktorou sa vyberajú klúčové snímky taktiež závisí na vzdialosti objektov/významných bodov od kamery a rýchlosťi akou sa kamera pohybuje.

Počiatočná dvojica klúčových snímok je zobrazená na obrázku 5.8.



Obr. 5.8: Počiatočná dvojica klúčových snímok - snímka 1 a 32

Pózy nasledujúcich snímok sú už určené algoritmom *Projective 3-Point* a po každej pridanej klúčovej snímke sa vykoná *Bundle adjustment*. Vypočítané pózy pre vybrané klúčové snímky reprezentujúce danú trasu sú zobrazené na obrázku 5.9.



Obr. 5.9: Pózy klúčových snímok trasy č. 109

Pri výpočte pohybu kamery je dôležité, aby v bodových korešpondenciách, na základe ktorých sa určuje póza kamery, boli identifikované a odstránené falošné asociácie tzv. *outliers*. Tieto falošné korešpondencie spôsobujú chybné a nepresné určenie pózy, čo má vplyv na fungovanie ďalších modulov systému detekcie zmien. Ich úspešná filtračia je preto kritická a musí prebehnúť v každej časti tohto modulu, kde sa pracuje s bodovými korešpondenciami, t.j. pri výpočte relatívnej pózy pomocou esenciálnej matice, pri výpočte pózy pomocou algoritmu *Projective 3-Point* a pri optimalizácii pôz pomocou algoritmu *Bundle adjustment*.

Dĺžka samotnej trasy nie je definovaná v reálnych jednotkách (napríklad v metrech), pretože je použitá monokulárna kamera a bez externých informácií o scéne (napríklad znalosť rozmerov objektov alebo informácia zo senzoru) nevieme odvodiť skutočné rozmery 3D modelu a teda ani skutočné vzdialenosť medzi pôzami jednotlivých pohľadov. Limitom pre tento modul a vlastne aj pre celkové riešenie teda nie je fyzická dĺžka trasy ale počet klúčových snímok, pretože s pribúdajúcimi klúčovými snímkami rastie aj chyba vo výpočte pózy a pre "dlhšie" trasy môže výpočet polohy postupne degradovať, ide o tzv. *drift*, ktorý sa do istej miery dokáže potlačiť optimálnym algoritmom *Bundle adjustment*, ktorý je v tomto module použitý. Počet klúčových snímok zaznamenannej trasy závisí na viacerých faktoroch, ktorými sú minimálny posuv/uhlová paralaxa medzi dvoma po sebe idúcimi pohľadmi, vzdialenosť objektov/významných bodov od kamery, rýchlosť pohybu kamery a tiež fyzická dĺžka trasy.

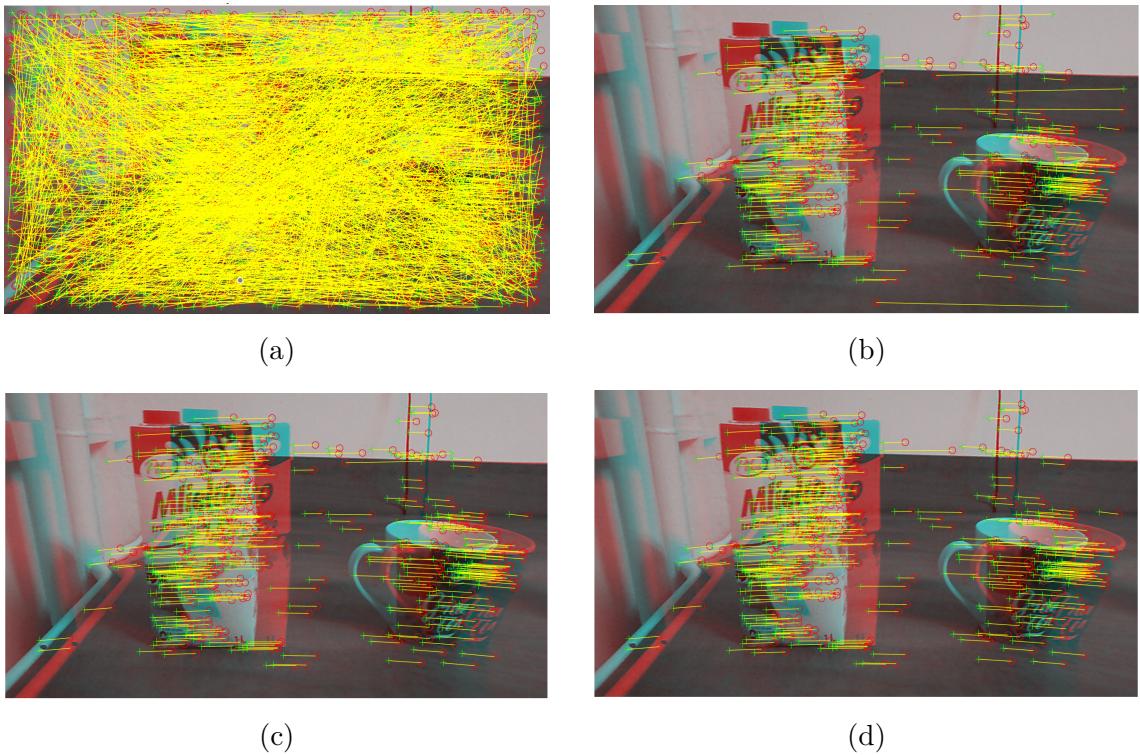
5.3.2 3D štruktúra

Tento modul odpovedá v rámci kapitoly návrhu algoritmov kapitole 4.1.4.

Stanovovanie hustých bodových korešpondencií prebieha medzi aktuálnou klúčovou snímkou a n nasledujúcimi klúčovými snímkami. Hodnota n bola pre všetky trasy nastavená na 2 nasledujúce snímky. Čím väčší je interval n snímok, tým presnejšie sú triangulované 3D body, výmenou za zvýšenú časovú náročnosť.

Filtrácia bodových korešpondencií medzi dvoma klúčovými snímkami je zobrazená na obrázku 5.10. Na prvej snímke a) sú zobrazené prvotné husté korešpondencie (vznikli porovnaním príznakových vektorov SURF významných bodov) a chceme ich dostať čo najviac (kvôli hustejšiemu mračnu bodov). Z týchto korešpondencií budú v ďalších krokoch prísnejšie vyberané čo najkvalitnejšie bodové dvojice.

Filtrácia týchto korešpondencií prebieha kaskádovo pomocou 3 filtrov. V prvom filtrovi sú odstránené bodové dvojice, ktoré nevyhovujú epipolárnej geometrii (algoritmus *RANSAC* v kombinácii s výpočtom esenciálnej matice) a odozva tohto filtrovi je na obrázku b). V druhom filtrovi sú odstránené bodové dvojice, ktorých euklidovská vzdialenosť v pixeloch je p -krát väčšia než priemerná vzdialenosť všetkých bodo-



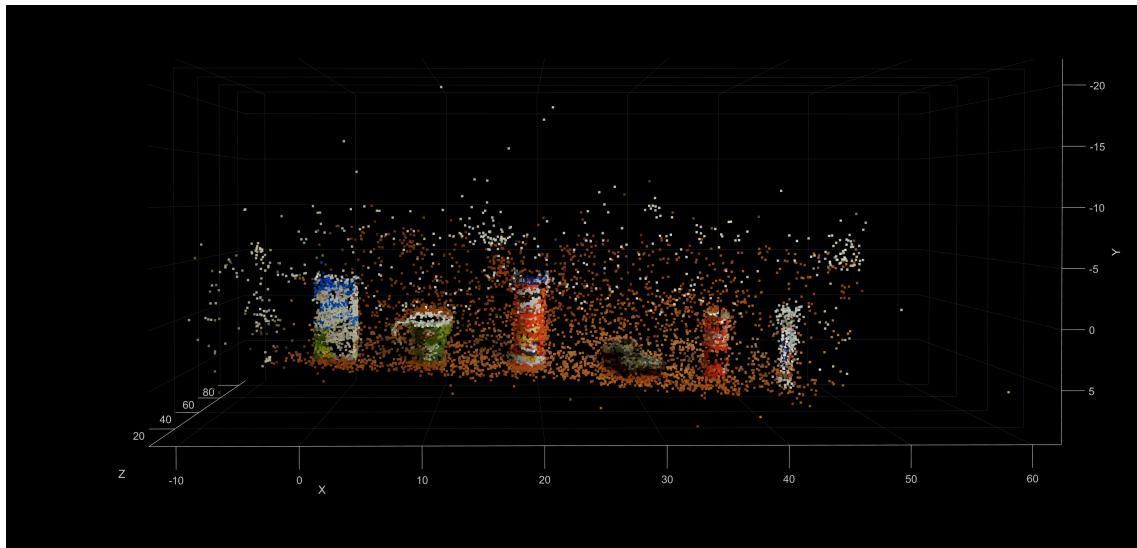
Obr. 5.10: Bodové korešpondencie medzi dvojicou klúčových snímok: a) Prvotné korešpondencie, b) Epipolárne korešpondencie, c) Korešpondencie filtrované podľa pixelovej vzdialenosť, d) Korešpondencie s dobrou uhlovou paralaxou

vých dvojíc. p je nastavené na hodnotu 2,5 a odozva tohto filtra je na obrázku c). V treťom filteri sú odstránené bodové dvojice, ktorých uhlová paralaxa je menšia než 1° .

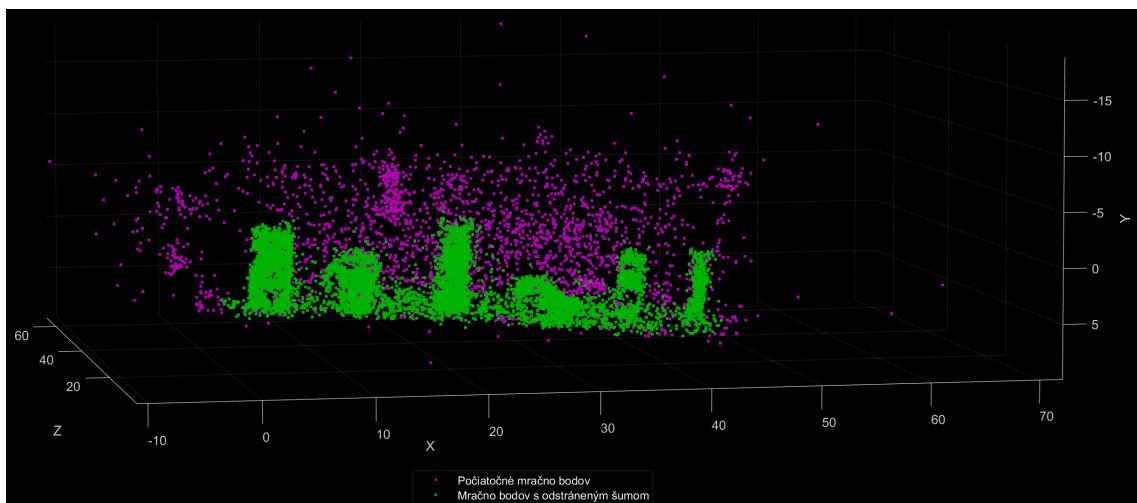
Po prejdení všetkých klúčových snímok sa na základe určených korešpondencií triangulujú 3D body, ktorých poloha je následne optimalizovaná pomocou algoritmu *Bundle adjustment* a pre dané 3D body sa určia ich RGB hodnoty z obrazu. Výsledné 3D mračno bodov je zobrazené na obrázku 5.11. Toto počiatocné mračno bodov môže byť stále zašumené chybnými/nepresnými bodmi, ktoré sú v nasledujúcej kapitole odstránené.

5.3.3 Odstránenie šumu z mračna bodov

Tento modul odpovedá v rámci kapitoly návrhu algoritmov kapitole 4.1.5. Veľkosť okolia n je zvolená pre každú trasu individuálne, pre túto konkrétnu trasu je použitá hodnota 300 a hodnota prahu σ je pre každú trasu rovnaká, nastavená na hodnotu $1e^{-5}$. Odšumené mračno bodov pôvodnej trasy je na obrázkoch 5.12 a 5.13 a v tejto podobe sa s ním pracuje v ďalších moduloch.



Obr. 5.11: Počiatočné mračno bodov pôvodnej trasy č. 109

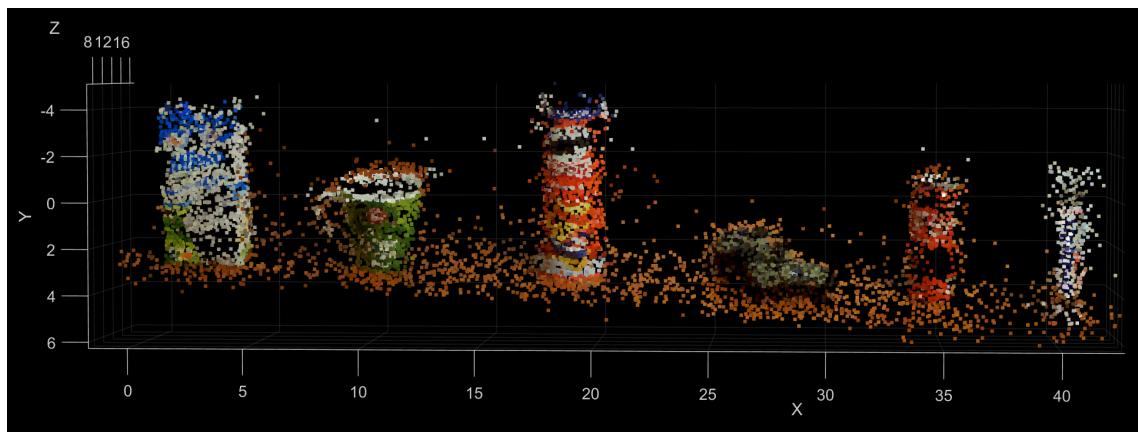


Obr. 5.12: Odšumenie mračna bodov

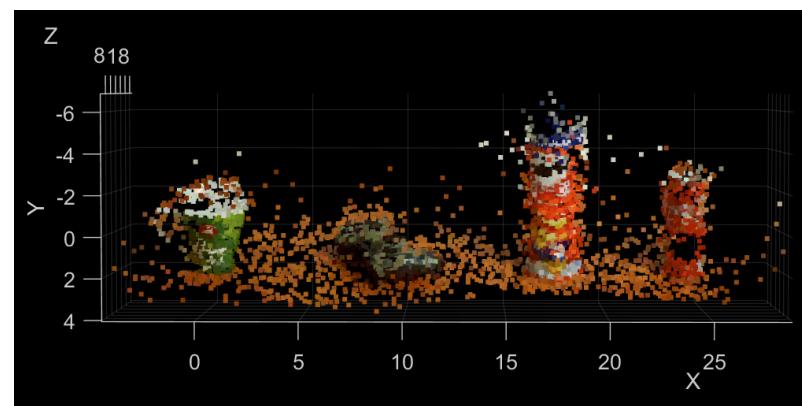
5.3.4 Registrácia mračien bodov

Tento modul odpovedá v rámci kapitoly návrhu algoritmov kapitole 4.2. V tomto module prebieha registrácia lokálneho mračna bodov, ktoré je zobrazené na obrázku 5.14 a reprezentuje trasu č. 110 (obrázok 5.4) v globálnom mračne bodov, ktoré je na obrázku 5.13 a ktoré zas reprezentuje originálnu trasu č. 109 na obrázku 5.3.

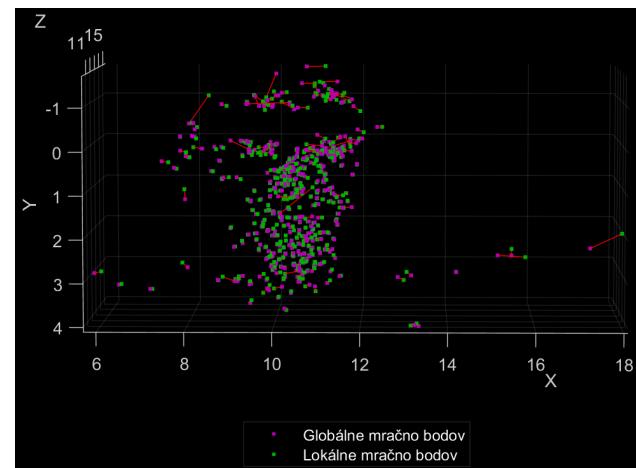
Po prejdení a porovnaní klúčových snímok oboch trás sú stanovené bodové ko-rešpondencie medzi 3D bodmi oboch mračien bodov, zobrazené na obrázku 5.15. Ide hlavne o body patriace zelenému hrnčeku, ktorý predstavuje objekt, ktorý ostal medzi trasami geometricky nezmenený a na jeho bodoch sa môže ukotviť výpočet



Obr. 5.13: Mračno bodov zbavené šumu reprezentujúce pôvodnú trasu č. 109

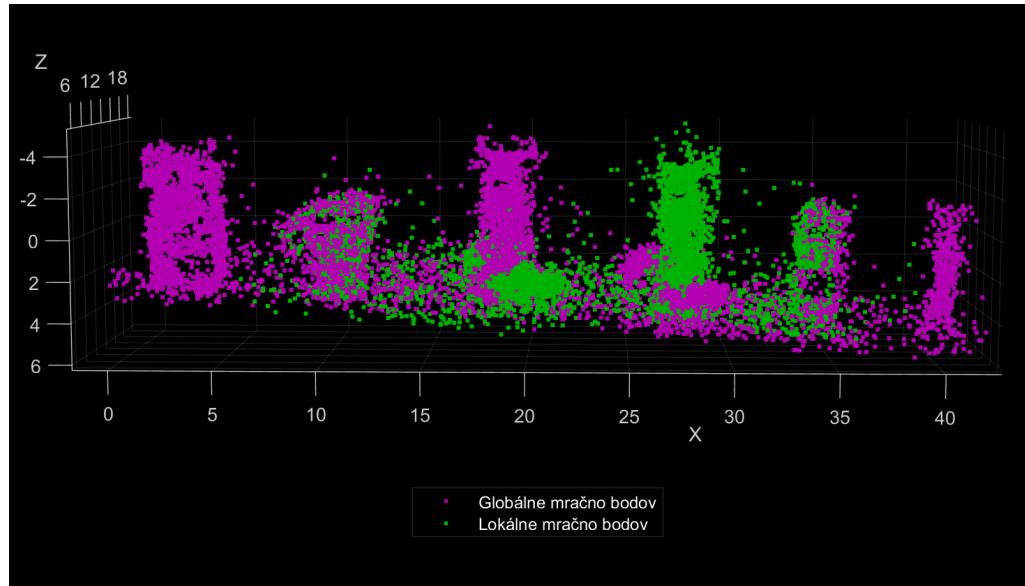


Obr. 5.14: Mračno bodov zbavené šumu reprezentujúce lokálnu trasu č. 110



Obr. 5.15: 3D korešpondencie medzi mračnami bodov

podobnostnej transformácie medzi týmito mračnami bodov. Pre úspešnú registráciu je teda dôležité aby obe mračná bodov mali nejaké spoločné body, ktoré ostávajú v scéne nezmenené. Lokálne mračno bodov transformované/umiestnené do globálneho mračna bodov pôvodnej trasy je zobrazené v obrázku 5.16. V takejto konfigurácii



Obr. 5.16: Registrované mračná bodov trás 109 a 110

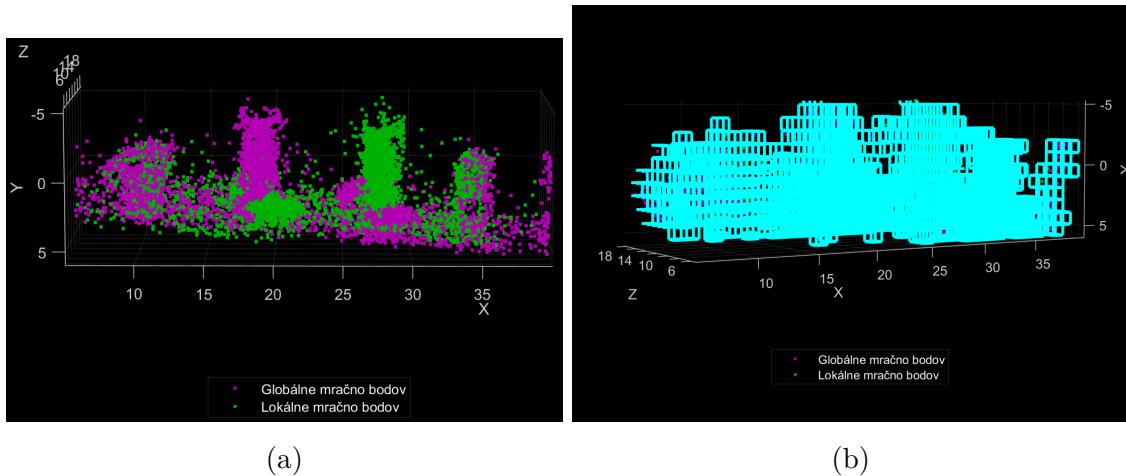
môžeme mračná bodov medzi sebou priamo porovnať a preskúmať možné zmeny, čomu sa venuje nasledujúci modul detekcie zmien.

5.3.5 Detekcia zmien

Tento modul odpovedá v rámci kapitoly návrhu algoritmov kapitole 4.3.

Voxelizácia začína vymedzením vyšetrovaného priestoru registrovaných mračien z obrázku 5.16. Jeho hranice sú určené lokálnym mračnom bodov, ktoré je menšie. Tento priestor je následne rozdelený na kubické voxely o veľkosti g , čo je znázorené na obrázku 5.17 a ďalej sú vyšetrované iba neprázdné voxely. Veľkosť voxelovej mriežky g je pre každú dvojicu trás iná a nastavuje sa individuálne. Veľkosť rozlíšenia g sa odvíja od veľkosti/merítka registrovaných mračien bodov, ktoré je dané merítkom mračna bodov pôvodnej trasy, keďže tento 3D model ostáva zafixovaný a podobnostná transformácia (translácia, rotácia a merítko) sa aplikuje na lokálny-/menší 3D model aktuálnej trasy. Pôvodná trasa je vopred známa, preto sa môže podľa nej nastaviť vhodné voxelové rozlíšenie resp. veľkosť mriežky.

Následne sa pre všetky 3D body vypočítajú ich povrchové normály. Veľkosť okolia k pre výpočet normál je nastaviteľným parametrom a pre každú dvojicu trás je



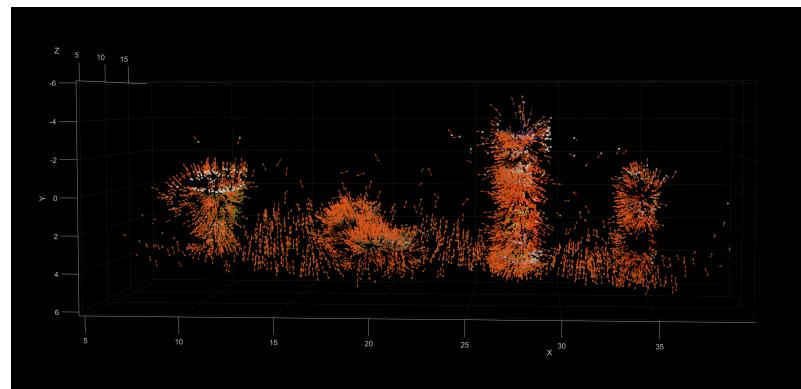
Obr. 5.17: a) Vymedzenie vyšetrovaného priestoru, b) Voxelizácia

jeho hodnota iná. Ak je okolie k príliš veľké, môžu sa doň zahrnúť aj body, ktoré nie sú súčasťou lokálneho povrchu (patria inému objektu) a aproximácia povrchu môže byť týmto skreslená. Pre nižšie hodnoty k sa viac prejavia jemné detaily, no môže sa zvýrazniť aj šum. Veľkosť okolia pre výpočet normál je teda analogická k filtrovaniu šumu a hranovým detektorom v 2D obraze. Veľkost okolia znamená potlačenie jemných detailov/vysokých frekvencií, čo je podobné napríklad filtru s Gaussovým rozložením. Malé okolie zas vyzdvihne vysoké frekvencie, čo je podobné hranovým detektorom. Ideálne by sa veľkosť okolia mala určiť individuálne pre každý 3D bod pomocou nejakého optimálneho kritéria založeného napríklad na variácii povrchu. Táto technika sa využíva v renderovaní, kde je dôležité, aby bol povrch approximovaný čo najpresnejšie aj za cenu vysokej časovej náročnosti. Veľkosť okolia k sa teda musí zvoliť tak, aby orientácia normál odpovedala orientácii povrchu objektov v danom mračne bodov. Tým, že je pôvodná trasa vopred známa, môže sa podľa jej 3D modelu vopred nastaviť veľkosť okolia k . Approximácia povrchu mračna bodov je zobrazená na obrázku 5.18.

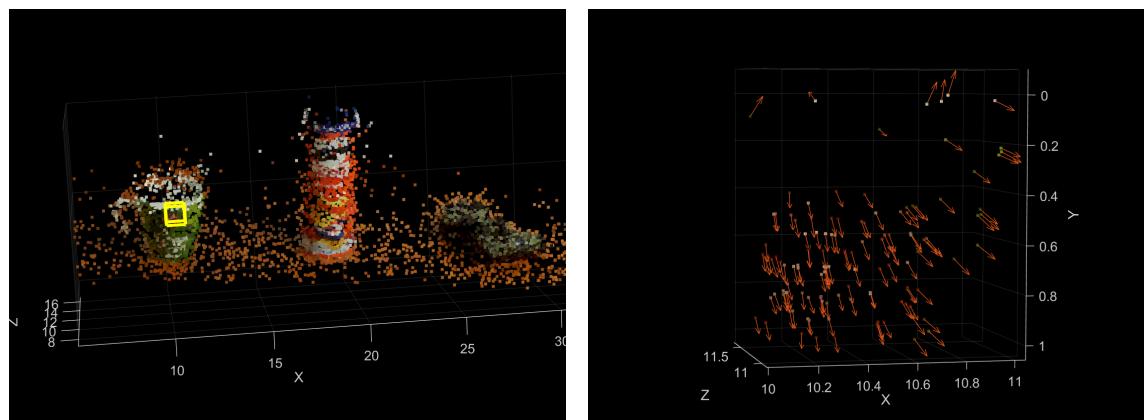
Pre každý voxel sa na základe povrchových normál vytvorí dvojica histogramov orientácií. Exemplárny histogram orientácií pre povrch vo vybranom voxelí je na obrázku 5.19.

Porovnaním takejto dvojice histogramov (ktoré sú predtým ešte normalizované) sa rozhodne o zmene v danom voxelí. Hodnota prahu klasifikácie, t.j. minimálna povrchová podobnosť s_{min} (kapitola 4.3) je nastavená na hodnotu 0,5. Po prehľadaní 26-okolia každého voxelu sú 3D body oboch mračien klasifikované do dvoch tried - 'zmena' a 'žiadna zmena', čo je zobrazené na obrázku 5.20. Môžeme vidieť, že boli správne klasifikované body objektov, ktoré boli zamenené. Časť bodov v okrajoch mračien sa tiež označila ako zmena, pretože tam už lokálny/menší 3D model neob-

sahoval žiadne body a pôvodný 3D model áno. Pôvodný 3D model trasy môže byť

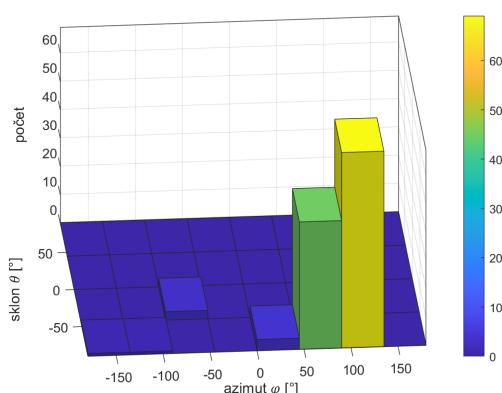


Obr. 5.18: Normály pre mračno bodov aktuálnej trasy č. 110



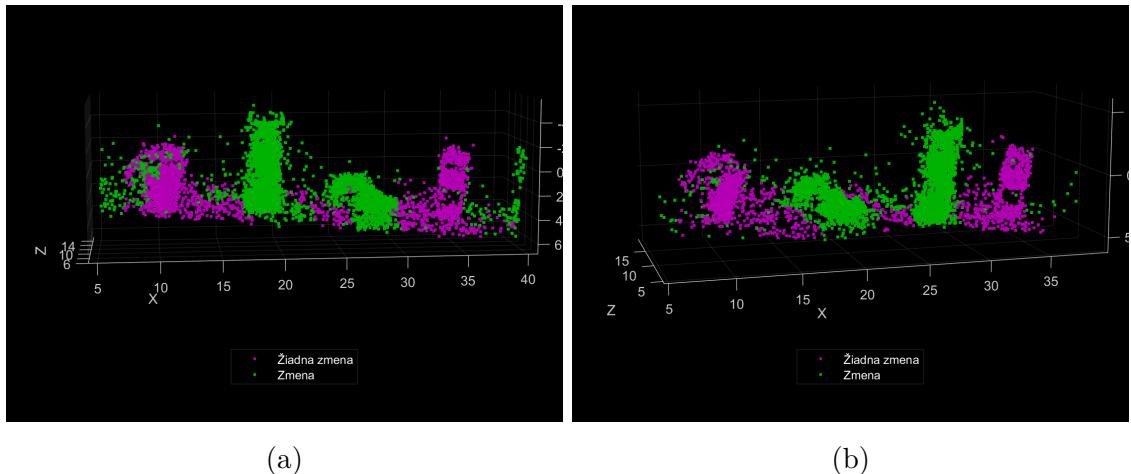
(a)

(b)



(c)

Obr. 5.19: a) Vybraný voxel, b) Lokálny povrch, c) Histogram orientácie daného povrchu



Obr. 5.20: a) Zmeny v pôvodnej trase č. 109, b) Zmeny v aktuálnej trase č. 110

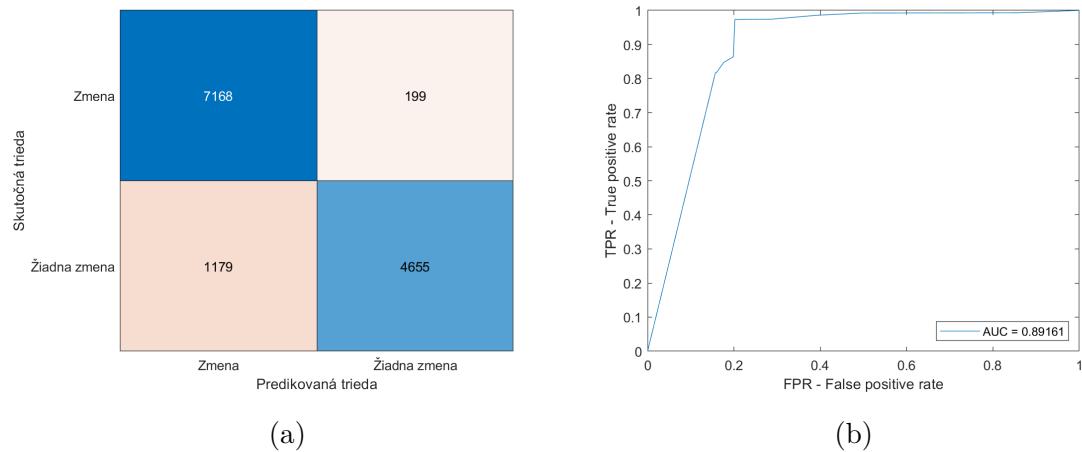
následne aktualizovaný o dané detekované zmeny.

5.3.6 Vyhodnotenie

V tejto práci sú vyhodnotené výstupy modulu detektie zmien, ktorý sumarizuje funkčnosť celého systému. Vyhodnotenie môže prebiehať aj po častiach, teda môže sa vyhodnotiť každý modul zvlášť, ale v rámci celého systému je merítko/veľkosť scény neznáme, takže hodnota/parameter presnosti výpočtu pohybu kamery a presnosti registrácie by bol bezrozmerný t.j. neboli by vyjadrený v reálnych jednotkách - metroch. Pri určovaní presnosti pohybu kamery by bolo tiež potrebné mať k dispozícii referenčné/správne hodnoty a v rámci 3D rekonštrukcie by bol potrebný referenčný 3D model. Vyhodnotenie celého systému detektie zmien bude teda prebiehať vyhodnotením jeho posledného modulu. Za týmto účelom boli v každom mračne bodov ručne vyznačené 3D body, ktoré predstavujú zmeny a zvyšné body sú uvažované ako nezmenené. Systém detektie zmien funguje ako binárny klasifikátor, ktorý rozdeľuje vstupnú množinu 3D bodov do dvoch tried - 'zmena' a 'žiadna zmena'. Vyhodnotenie tohto klasifikátoru zahŕňa výpočet matice zámen a jej parametrov ako senzitivita (TPR - *true positive rate*), špecificita (TNR - *true negative rate*), FPR - *false positive rate* a presnosť. Na základe hodnôt TPR a FPR je pre každú dvojicu trás vykreslená ROC krivka pre rôzne hodnoty prahu klasifikátoru, ktorým je minimálna povrchová podobnosť s_{min} (kapitola 4.3). Ako pozitívna trieda je braná trieda 'zmena' a ako negatívna je braná trieda 'žiadna zmena'. Dôležitým parametrom je senzitivita, teda úspešnosť detektie skutočných zmien a je dôležitejšia než špecificita, teda detekcia nezmenených/identických povrchov. Ak sa nesprávne vyhodnotia dva identické povrhy ako 'zmena', tak to nie je až tak kritické, pretože

pri aktualizácii (doplnení detekovaných zmien) do pôvodného 3D modelu je povrch z pôvodnej trasy odstránený a nahradený jeho identickým náprotivkom z aktuálnej trasy. V opačnom prípade, ak je skutočná zmena (napríklad povrch, ktorý v scéne pribudol) klasifikovaný ako 'žiadna zmena', tak o daný povrch prichádzame a táto zmena sa nepretaví do aktualizovanej mapy/modelu pôvodnej trasy.

Matica zámen a ROC krivka pre demonštrovanú dvojicu trás 109 a 110 je na obrázku 5.21. Parametre matice zámen sú v tabuľke 5.2. Z obrázku 5.20 a z grafu a tabuľky nižšie je vidieť, že geometrické zmeny sú detekované správne a chyba klasifikátora je akceptovateľná. Grafy a tabuľky s vyhodnotením zvyšných tras sa nachádzajú v prílohe X spolu s popisom aplikácie, ktorá bola vytvorená na otestovanie a prechod jednotlivými časťami systému detektie zmien.



Obr. 5.21: a) Matica zámen pre trasy 109 a 110, b) ROC krivka pre trasy 109 a 110

Tab. 5.2: Vyhodnotenie trasy 109 a 110

Pôvodná trasa	Aktuálna trasa	Senzitivita (TPR)	Špecificita (TNR)	Presnosť	AUC
109	110	0,973	0,798	0,896	0,892

Záver

V prvom bode zadania boli preskúmané práce/články, ktoré sa zaoberali detekciou zmeny. Z rôznych prístupov k riešeniu detekcie zmeny bola vybraná metóda s využitím 3D dát, konkrétnie metóda priameho porovnania dvoch 3D mračien bodov. Tomuto sa venuje kapitola 1. Pre vybrané riešenie boli naštudované princípy metód a algoritmov, ktoré budú použité, čomu odpovedá kapitola 2.

V druhom bode zadania sú navrhnuté testovacie dáta a definované podmienky a limity v nasnímaní dát. Ďalej boli definované rôzne prípady zmien, ktoré dokáže systém detektovať a na základe toho boli nasnímané rôzne scenáre pomocou monokulárnej kamery. Tomuto bodu sa venuje kapitola 3.

V treťom bode zadania boli navrhnuté a realizované algoritmy 3D rekonštrukcie, ktorými sa zo vstupných videosekvencií vytvorí 3D mračno bodov, ktoré reprezentuje danú trasu. 3D rekonštrukcia je rozdelená do 3 častí. V prvej je zo vstupných snímok vybraná časť klúčových snímok/pohľadov, pre ktoré sa určí ich poloha na základe SURF významých bodov sledovaných v obrazu. V druhej časti sú na základe klúčových snímok triangulované 3D body, ktoré tvoria mračno bodov. V tretej časti je toto mračno bodov zbavené šumu. Tento bod zadania je popísaný v kapitolách 4.1 a 5.3.1 až 5.3.3.

V štvrtom bode zadania boli navrhnuté a realizované algoritmy, ktorými sa nájde aktuálna poloha vo vzorových dátach a stanovia sa geometrické rozdiely. Aktuálna poloha resp. úsek pôvodnej trasy je reprezentovaný mračnom bodov, ktoré je lokalizované/registrované v globálnom 3D mračne bodov pôvodnej trasy. Registrácia znamená výpočet podobnostnej transformácie (rotácia, translácia a merítka), ktorá prevedie lokálne mračno bodov do súradnicového systému globálneho mračna bodov na základe korešpondencí medzi 3D bodmi oboch mračien. Tieto 3D korešpondencie sú určené skrz korešpondencie medzi 2D SURF významnými bodmi, ktoré odpovedajú 3D bodom. Návrh a realizácie registrácie mračien bodov je obsiahnutá v kapitolách 4.2 a 5.3.4. Nad registrovanými mračnami bodov následne priebieha detekcia zmien. Vyšetrovaný priestor je rozdelený do rovnako veľkých kubických voxelov. Pre každý voxel sa určí dvojica histogramov orientácií, ktoré popisujú tvar/orientáciu lokálneho povrchu zloženého z 3D bodov v danom voxelovi. Orientácia povrchu je daná výpočtom povrchových normál pre každý 3D bod a ich prevodom do sférických súradníc, na uhly ϕ a θ . Porovnaním dvojice histogramov (prienik histogramov - metóda diel po dieli) je rozhodnuté o tom, či dvojica povrchov obsiahnutých v danom voxelovi je identická resp. má rovnakú orientáciu alebo nie. Takto je prejdený každý voxel, ktorý je klasifikovaný ako 'zmena' alebo 'žiadna zmena'. Následne je skúmané lokálne 26-okolie každého voxelu, pomocou čoho sú nájdené prepojené zhľuky voxelov s rovnakým príznakom zmeny. Výstupom detekcie zmien

je teda rozdelená vstupná množina 3D bodov do tried 'zmena' a 'žiadna zmena'. Návrh a realizácia algoritmov detekcie zmien je v kapitolách 4.3 a 5.3.5.

V piatom bode zadania boli vyhodnetené výstupy systému detekcie zmien. Ako metrika bola použitá matica zámen a jej parametre spolu s určením ROC krivky a plochy pod ňou. Vyhodnotenie teda prebiehalo ako v prípade binárneho klasifikátoru. Z výsledkov je vidieť, že detekcia geometrických zmien je funkčná a chyba klasifikácie je akceptovateľná. Najlepšie výsledky má klasifikátor v prípade zmien typu pridanie/odobranie objektu, ktoré sú pri prechode ľubovoľným prostredím časté. Zmeny typu zámena objektov alebo rotácia objektov na mieste, dosahujú v porovnaní s predošlou situáciou o trochu menej presné výsledky, pretože sa tam naroziel od predošlého prípadu stretávajú dva ľubovoľné povrchy, čo je zložitejší prípad, ale detekcia týchto zmien je stále dostatočne presná. Vyhodnotenie na rôznych testovacích dátach ukázalo, že navrhnutý a realizovaný systém detekcie zmien je funkčný, čiže dokáže detektovať geometrické zmeny v konkrétnom mieste v pôvodnej trase. Ďalším prínosom tejto práce je rekonštrukcia 3D modelu trasy, ktorý môže byť po každom jej prechode aktualizovaný o nájdené zmeny, čo môže mať praktické využitie v rôznych aplikáciách, napr. pri udržovaní aktuálnej 3D mapy mesta. Vyhodnoteniu sa venuje kapitola 5.3.6.

Navrhnuté a realizované algoritmy môžu slúžiť ako základ pre aplikáciu, ktorá zaistí asistenciu pri priechode známou trasou - teda bude upozorňovať na významné miesta a geometrické zmeny v danej trase. V súčasnej podobe pracuje navrhnutý systém offline, teda zmeny sú detektované až po priechode oboma trasami, čo má stále praktické využitie, ale ak by bolo cieľom práce realizovať detekciu zmien online resp. počas priechodu trasou v reálnom čase museli by sa viaceré časti systému zrýchliť. Najviac časovo náročný je modul 3D rekonštrukcie. Nasnímanie a 3D rekonštrukcia pôvodnej trasy prebieha vopred, čiže offline a teda žiadny časový limit nemá. Časová náročnosť nás zaujíma až pri rekonštrukcii lokálneho 3D modelu, ktorý reprezentuje aktuálnu polohu v pôvodnej trase. Tu je potrebné, aby bolo 3D mračno bodov rekonštruované čo najrýchlejšie, ale aby malo stále výpovednú hodnotu. Zrýchlenie by sa dalo zaistiť v module pre výpočet pohybu kamery za použitia rýchlejších významných bodov a príznakových vektorov, napr. ORB v kombinácii s binárnymi príznakovými vektormi. Ďalej by sa mohol znížiť počet detektovaných bodov v jednotlivých snímkach, okno optimalizačného algoritmu *Bundle adjustment* na menší interval snímkov a mohla by sa zvýšiť minmálna paralaxa/posuv medzi snímkami. Ďalej by sa mohla redukovať vstupná sekvencia snímkov na každú druhú alebo treťiu snímku, keďže väčšina snímkov je redundantných. Taktiež by sa mohli algoritmy prepísat do rýchlejšieho jazyka, napr. C++. Tieto opatrenia by zabezpečili zrýchlenie algoritmov, no zároveň by znížili robustnosť celého systému (okrem posledného opatrenia).

Literatúra

- [1] QIN, Rongjun, Peter REINARTZ a Jiaojiao TIAN. 3D change detection – Approaches and applications. *ISPRS Journal of Photogrammetry and Remote Sensing*. 2016, **122**, 41-56. ISSN 0924-2716.
- [2] İLSEVER, Murat a Cem ÜNSALAN. *Two-Dimensional Change Detection Methods: Remote Sensing Applications* [online]. London: Springer London, 2012 [cit. 2024-02-27]. SpringerBriefs in Computer Science. ISBN 978-1-4471-4255-3.
- [3] KHARROUBI, Abderrazzaq, Florent POUX, Zouhair BALLOUCH, Rafika HAJJI a Roland BILLEN. Three Dimensional Change Detection Using Point Clouds: A Review. *Geomatics*. 2022, **2**(4), 457–486.
- [4] QIN, Rongjun a Armin GRUEN. 3D change detection at street level using mobile laser scanning point clouds and terrestrial images. *ISPRS Journal of Photogrammetry and Remote Sensing*. 2014, **90**, 23-35.
- [5] YEW, Zi Jian a Gim Hee LEE. City-scale Scene Change Detection using Point Clouds. *2021 IEEE International Conference on Robotics and Automation (ICRA)*. 2021, 13362-13369.
- [6] PALAZZOLO, Emanuele a Cyrill STACHNISS. Fast Image-Based Geometric Change Detection Given a 3D Model. *2018 IEEE International Conference on Robotics and Automation (ICRA)*. 2018, 6308-6315.
- [7] GIRARDEAU-MONTAUT, D., Michel ROUX, Raphaël MARC a Guillaume THIBAULT. Change detection on points cloud data acquired with a ground laser scanner. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. 2005, **36**, 30-35. ISSN 16821750.
- [8] POLLARD, Thomas a Joseph L. MUNDY. Change Detection in a 3-d World. *2007 IEEE Conference on Computer Vision and Pattern Recognition*. 2007, 1-6.
- [9] TANEJA, Aparna, Luca BALLAN a Marc POLLEFEYS. Image Based Detection of Geometric Changes in Urban Environments. *2011 International Conference on Computer Vision*. 2011, 2336-2343.
- [10] GOLPARVAR-FARD VIRGINIA, Mani, Peña-Mora FENIOSKY a Silvio SA-VARESE. Monitoring Changes of 3D Building Elements from Unordered Photo Collections. *2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops)*. 2011, 249-256.

- [11] ADAM, Aikaterini, Torsten SATTLER, Konstantinos KARANTZALOS a Tomas PAJDLA. Objects Can Move: 3D Change Detection by Geometric Transformation Consistency. *Computer Vision – ECCV 2022*. 2022, 108-124.
- [12] FANFANI, Marco Fanfani a Carlo COLOMBO. Structural Change Detection by Direct 3D Model Comparison. In: *Proceedings of the 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2019) - Volume 5: VISAPP*. SciTePress, 2019, s. 760-767. ISBN 978-989-758-354-4. Dostupné z: doi:10.5220/0007260607600767
- [13] GÖNCZ, Levente a András László MAJDIK. Object-Based Change Detection Algorithm with a Spatial AI Stereo Camera. *Sensors*. 2022, 22(6342). Dostupné z: doi:10.3390/s22176342
- [14] Homogeneous Coordinates. Songho.ca [online]. ©2005-2024 [cit. 2024-03-20]. Dostupné z: <http://www.songho.ca/math/homogeneous/homogeneous.html>
- [15] Homogeneous Coordinates (Cyrill Stachniss, 2020). In: Youtube [online]. 20.08.2020 [cit. 2024-03-23]. Dostupné z: https://www.youtube.com/watch?v=MQdm0Z_gNcw. Kanál uživateľa Cyril Stachniss.
- [16] STACHNISS, Cyril. Homogeneous Coordinates [prezentácia]. In: StachnissLab Photogrammetry & Robotics, Bonn [online]. [cit. 2024-03-25]. Dostupné z: <http://www.ipb.uni-bonn.de/photo12-2021/index.html>
- [17] VAN OOSTERWYCK, Nick. Real Time Human Robot Interactions and Speed Control of a Robotic Arm for Collaborative Operations. 2018.
- [18] STACHNISS, Cyril. Camera Parameters: Extrinsics and Intrinsics [prezentácia]. In: StachnissLab Photogrammetry & Robotics, Bonn [online]. [cit. 2024-03-25]. Dostupné z: <https://www.ipb.uni-bonn.de/photo12-2021/index.html#2>
- [19] Camera Parameters - Extrinsics and Intrinsics (Cyrill Stachniss). In: Youtube [online]. 15.04.2020 [cit. 2024-03-25]. Dostupné z: <https://www.youtube.com/watch?v=uHApDqH-8UE>. Kanál uživateľa Cyril Stachniss.
- [20] ZHANG, Z. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2000, **22**(11), 1330-1334.
- [21] STACHNISS, Cyril. Camera Calibration: Zhang's Method [prezentácia]. In: StachnissLab Photogrammetry & Robotics, Bonn [online]. [cit. 2024-03-26]. Dostupné z: <https://www.ipb.uni-bonn.de/photo12-2021/index.html#2>

- [22] Camera Calibration using Zhang's Method (Cyrill Stachniss). In: Youtube [online]. 20.04.2020 [cit. 2024-03-26]. Dostupné z: <https://www.youtube.com/watch?v=-9He7Nu3u8s>. Kanál užívateľa Cyrill Stachniss.
- [23] STACHNISS, Cyril. Projective 3-Point (P3P) Algorithm / Spatial Resection [prezentácia]. In: StachnissLab Photogrammetry & Robotics, Bonn [online]. [cit. 2024-03-27]. Dostupné z: <https://www.ipb.uni-bonn.de/photo12-2021/index.html#2>
- [24] Projective 3-Point Algorithm using Grunert's Method (Cyrill Stachniss). In: Youtube [online]. 20.04.2020 [cit. 2024-03-27]. Dostupné z: <https://www.youtube.com/watch?v=N1aCvzF116Q>. Kanál užívateľa Cyrill Stachniss.
- [25] STACHNISS, Cyril. Relative Orientation, Fundamental and Essential Matrix [prezentácia]. In: StachnissLab Photogrammetry & Robotics, Bonn [online]. [cit. 2024-04-05]. Dostupné z: <https://www.ipb.uni-bonn.de/photo12-2021/index.html#2>
- [26] Relative Orientation, Fundamental and Essential Matrix (Cyrill Stachniss). In: Youtube [online]. 20.04.2020 [cit. 2024-04-05]. Dostupné z: <https://www.youtube.com/watch?v=gYYumFSDsvA>. Kanál užívateľa Cyrill Stachniss.
- [27] STACHNISS, Cyril. Epipolar Geometry [prezentácia]. In: StachnissLab Photogrammetry & Robotics, Bonn [online]. [cit. 2024-04-06]. Dostupné z: <https://www.ipb.uni-bonn.de/photo12-2021/index.html#2>
- [28] Epipolar Geometry Basics (Cyrill Stachniss). In: Youtube [online]. 20.04.2020 [cit. 2024-04-06]. Dostupné z: <https://www.youtube.com/watch?v=cLeF-KNHgwU>. Kanál užívateľa Cyrill Stachniss.
- [29] STACHNISS, Cyril. Direct Solutions for Computing Fundamental and Essential Matrix [prezentácia]. In: StachnissLab Photogrammetry & Robotics, Bonn [online]. [cit. 2024-04-07]. Dostupné z: <https://www.ipb.uni-bonn.de/photo12-2021/index.html#2>
- [30] Direct Solution for Estimating the Fundamental and Essential Matrix (Cyrill Stachniss). In: Youtube [online]. 21.04.2020 [cit. 2024-04-07]. Dostupné z: <https://www.youtube.com/watch?v=zX5NeY-GT00>. Kanál užívateľa Cyrill Stachniss.
- [31] NISTÉR, David. An Efficient Solution to the Five-Point Relative Pose Problem. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2004, **26**(6), 756-770.

- [32] STACHNISS, Cyril. RANSAC – Random Sample Consensus [prezentácia]. In: StachnissLab Photogrammetry & Robotics, Bonn [online]. [cit. 2024-04-07]. Dostupné z: <https://www.ipb.uni-bonn.de/photo12-2021/index.html#2>
- [33] RANSAC - Random Sample Consensus (Cyrill Stachniss). In: Youtube [online]. 14.04.2020 [cit. 2024-04-07]. Dostupné z: <https://www.youtube.com/watch?v=Cu1f6vpEilg>. Kanál užívateľa Cyrill Stachniss.
- [34] STACHNISS, Cyril. Triangulation and Absolute Orientation [prezentácia]. In: StachnissLab Photogrammetry & Robotics, Bonn [online]. [cit. 2024-04-08]. Dostupné z: <https://www.ipb.uni-bonn.de/photo12-2021/index.html#2>
- [35] Triangulation for Image Pairs (Cyrill Stachniss). In: Youtube [online]. 21.09.2020 [cit. 2024-04-08]. Dostupné z: <https://www.youtube.com/watch?v=UZlRhEUWSas>. Kanál užívateľa Cyrill Stachniss.
- [36] MRKVIČKA, D. Rekonstrukce 3D objektů z více pohledů. Brno: Vysoké učení technické v Brně, Fakulta strojního inženýrství, 2019. 74 s. Vedoucí Mgr. Jana Procházková, Ph.D.
- [37] KITANI, Kris. Triangulation [prezentácia]. In: CMU School of Computer Science [online]. [cit. 2024-04-08]. Dostupné z: <https://www.cs.cmu.edu/~16385/s17/Slides/>
- [38] Adjust collection of 3-D points and camera poses - MATLAB bundleAdjustment. MathWorks - Makers of MATLAB and Simulink [online]. © 1994-2024 [cit. 2024-04-08]. Dostupné z: <https://www.mathworks.com/help/vision/ref/bundleadjustment.html>
- [39] STACHNISS, Cyril. Bundle Adjustment – Part I Introduction & Application [prezentácia]. In: StachnissLab Photogrammetry & Robotics, Bonn [online]. [cit. 2024-04-08]. Dostupné z: <https://www.ipb.uni-bonn.de/photo12-2021/index.html#2>
- [40] The Basics about Bundle Adjustment (Cyrill Stachniss). In: Youtube [online]. 26.09.2020 [cit. 2024-04-08]. Dostupné z: <https://www.youtube.com/watch?v=sobyKHwgB0Y>. Kanál užívateľa Cyrill Stachniss.
- [41] HORÁK, Karel. Vícesnímková rekonstrukce scény [prezentácia]. In: Machine Vision Group [online]. [cit. 2024-04-08]. Dostupné z: <http://vision.uamt.feec.vutbr.cz/?course=R0Z>

- [42] STACHNISS, Cyril. Absolute Orientation Problem: Derivation of the Solution [prezentácia]. In: StachnissLab Photogrammetry & Robotics, Bonn [online]. [cit. 2024-04-10]. Dostupné z: <https://www.ipb.uni-bonn.de/photo12-2021/index.html#2>
- [43] Absolute Orientation Problem: Similarity Transformations Between Point Sets (Cyrill Stachniss, 2021). In: Youtube [online]. 09.06.2020 [cit. 2024-04-10]. Dostupné z: <https://www.youtube.com/watch?v=K5WG40gMlz8>. Kanál užívateľa Cyril Stachniss.
- [44] RUSU, Radu Bogdan. *Semantic 3D Object Maps for Everyday Manipulation in Human Living Environments*. Mníchov, 2009. Dizertácia. Technische Universität München, Fakultät für Informatik.
- [45] Spherical coordinates — pi2 documentation. Documentation of pi2 — pi2 documentation [online]. c2011c [cit. 2024-04-14]. Dostupné z: https://pi2-docs.readthedocs.io/en/stable/spherical_coordinates.html
- [46] Janáková, Ilona. Optické 3D měření [prezentácia]. In: Machine Vision Group [online]. [cit. 2024-04-14]. Dostupné z: <http://vision.uamt.feec.vutbr.cz/?course=POV>
- [47] XU, Yusheng, Xiaohua TONG a Uwe STILLA. Voxel-based representation of 3D point clouds: Methods, applications, and its potential use in the construction industry. *Automation in Construction*. 2021, **126**, 103675. ISSN 0926-5805.
- [48] TAREEN, Shaharyar Ahmed Khan a Zahra SALEEM. A comparative analysis of SIFT, SURF, KAZE, AKAZE, ORB, and BRISK. *2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)*. 2018, 1-10.
- [49] Point cloud outlier removal — Open3D 0.12.0 documentation. *Open3D: A Modern Library for 3D Data Processing — Open3D 0.12.0 documentation* [online]. ©2018 - 2020 [cit. 2024-04-17]. Dostupné z: https://www.open3d.org/docs/0.12.0/tutorial/geometry/pointcloud_outlier_removal.html
- [50] Prepare Camera and Capture Images For Camera Calibration - MATLAB & Simulink. *MathWorks - Makers of MATLAB and Simulink* [online]. © 1994-2024 [cit. 2024-04-24]. Dostupné z: <https://www.mathworks.com/help/vision/ug/prepare-camera-and-capture-images-for-camera-calibration.html>

Zoznam príloh

A Vyhodnotenie testovacích trás	104
A.1 Trasa č. 95 a 99_1	104
A.2 Trasa č. 100 a 101	106
A.3 Trasa č. 102 a 103	108
A.4 Trasa č. 104 a 105	109
A.5 Trasa č. 106 a 107	111
A.6 Trasa č. 111 a 112	113
A.7 Tabuľka parametrov matice zámen	115
B Demonštračná aplikácia	116
C Obsah prílohy na CD	117
D Obsah elektronickej prílohy	119

A Vyhodnotenie testovacích trás

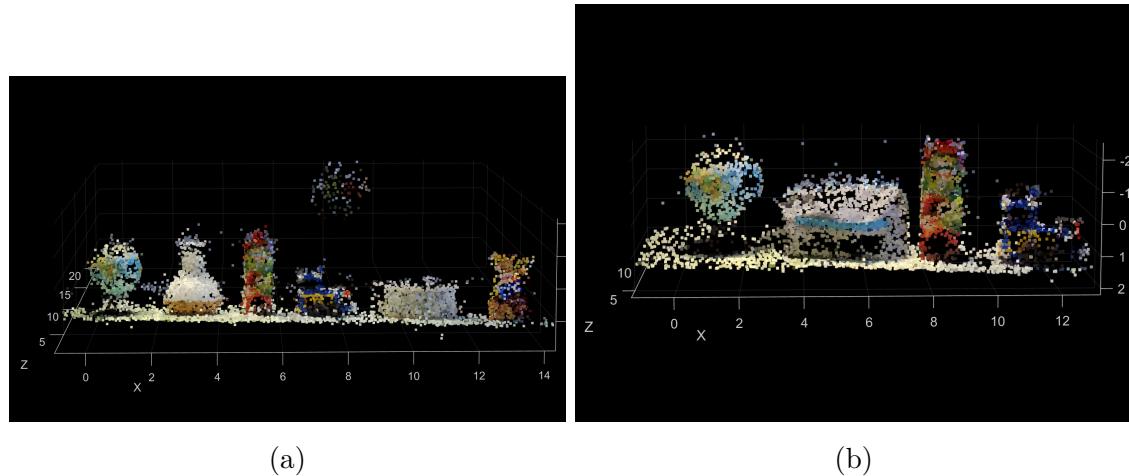
A.1 Trasa č. 95 a 99_1



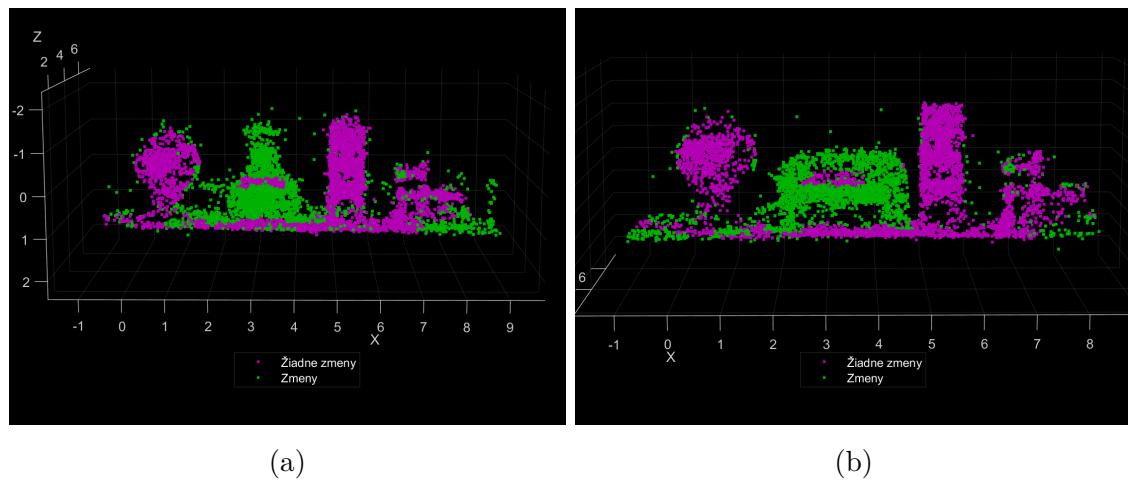
Obr. A.1: Pôvodná testovacia trasa č. 95



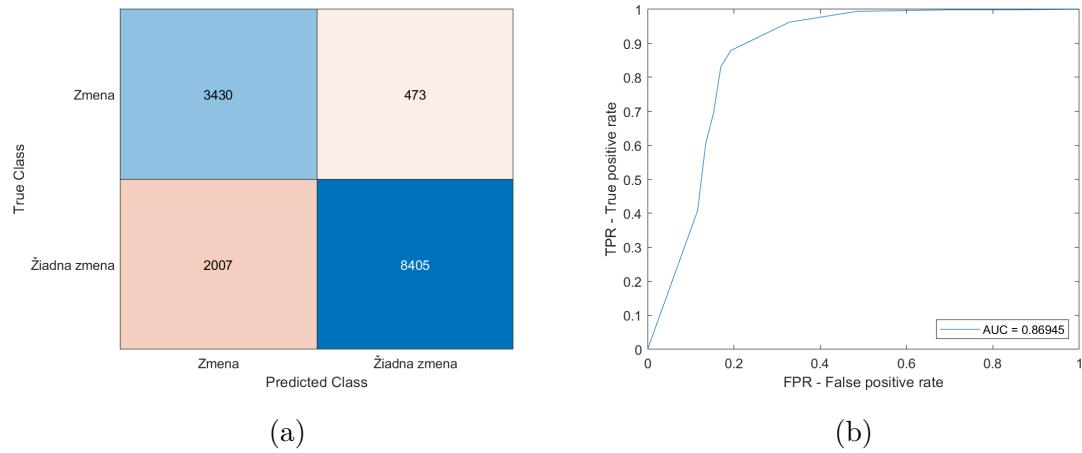
Obr. A.2: Aktuálna testovacia trasa č. 99_1



Obr. A.3: a) Pôvodná testovacia trasa č. 95, b) Aktuálna testovacia trasa č. 99_1



Obr. A.4: a) Zmeny v pôvodnej trase č. 95, b) Zmeny v aktuálnej trase č. 99_1



Obr. A.5: a) Matica zámen pre trasy 95 a 99_1, b) ROC krivka pre trasy 95 a 99_1

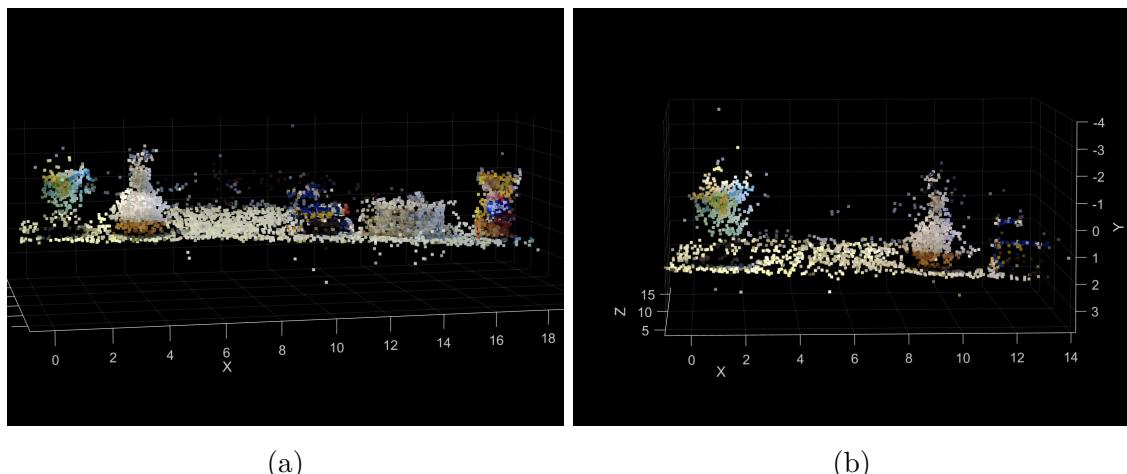
A.2 Trasa č. 100 a 101



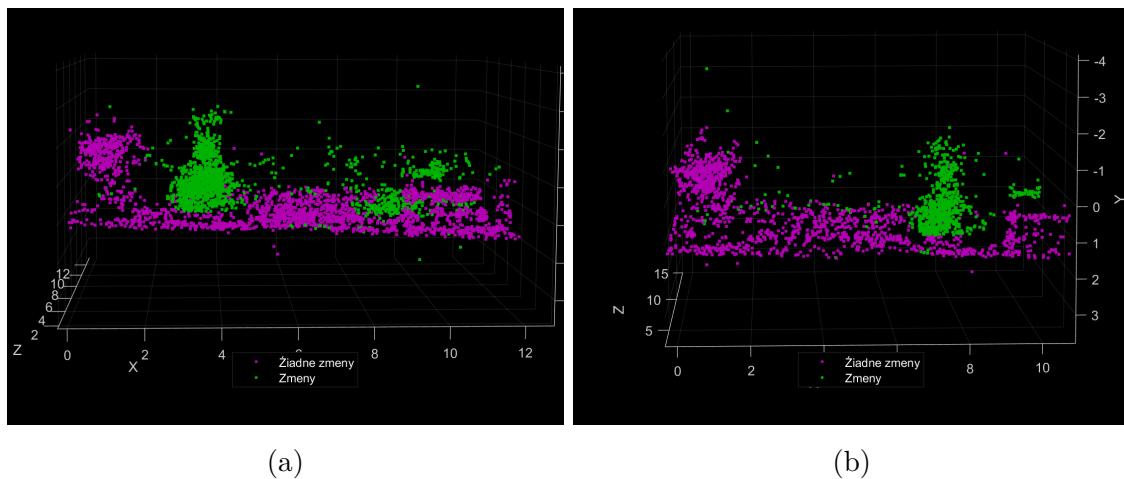
Obr. A.6: Pôvodná testovacia trasa č. 100



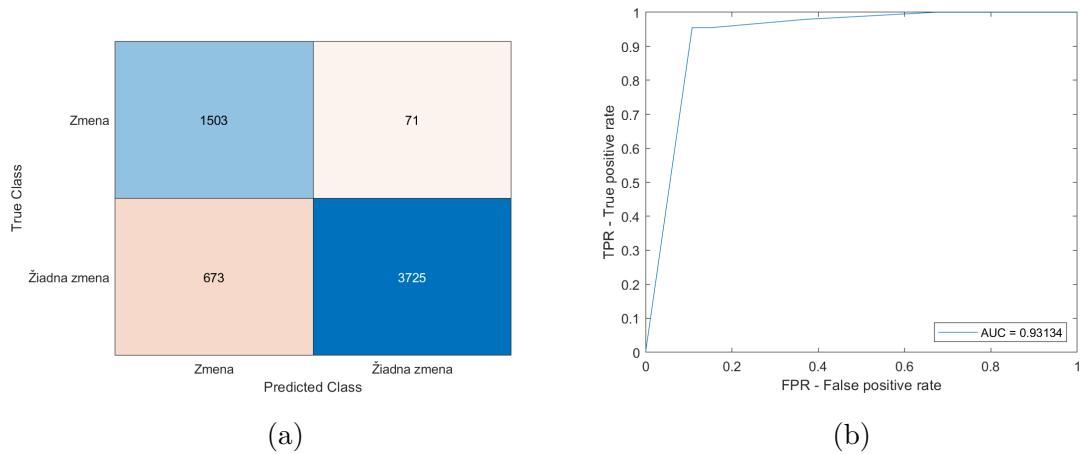
Obr. A.7: Aktuálna testovacia trasa č. 101



Obr. A.8: a) Pôvodná testovacia trasa č. 100, b) Aktuálna testovacia trasa č. 101



Obr. A.9: a) Zmeny v pôvodnej trase č. 100, b) Zmeny v aktuálnej trase č. 101



Obr. A.10: a) Matica zámen pre trasy 100 a 101, b) ROC krivka pre trasy 100 a 101

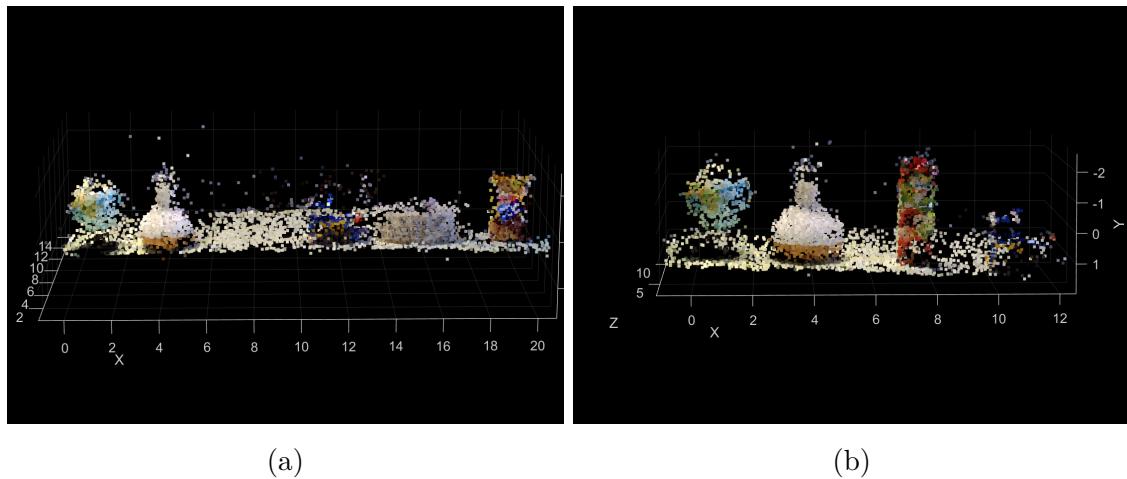
A.3 Trasa č. 102 a 103



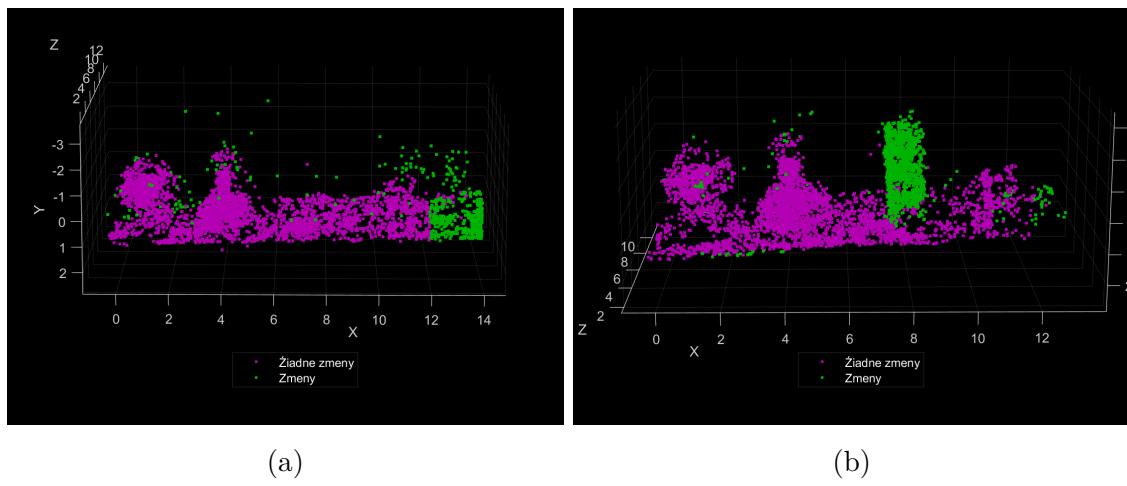
Obr. A.11: Pôvodná testovacia trasa č. 102



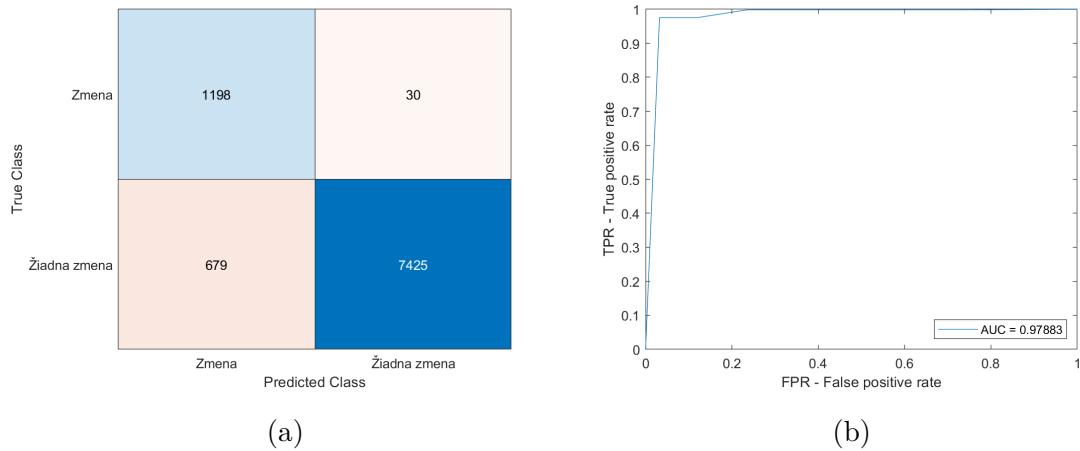
Obr. A.12: Aktuálna testovacia trasa č. 103



Obr. A.13: a) Pôvodná testovacia trasa č. 102, b) Aktuálna testovacia trasa č. 103



Obr. A.14: a) Zmeny v pôvodnej trase č. 102, b) Zmeny v aktuálnej trase č. 103



Obr. A.15: a) Matica zámen pre trasy 102 a 103, b) ROC krivka pre trasy 102 a 103

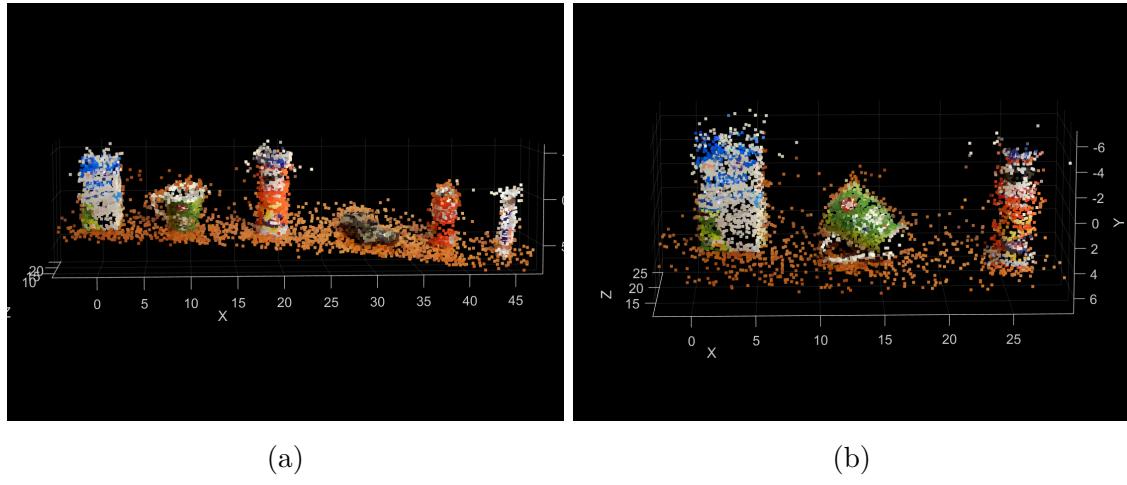
A.4 Trasa č. 104 a 105



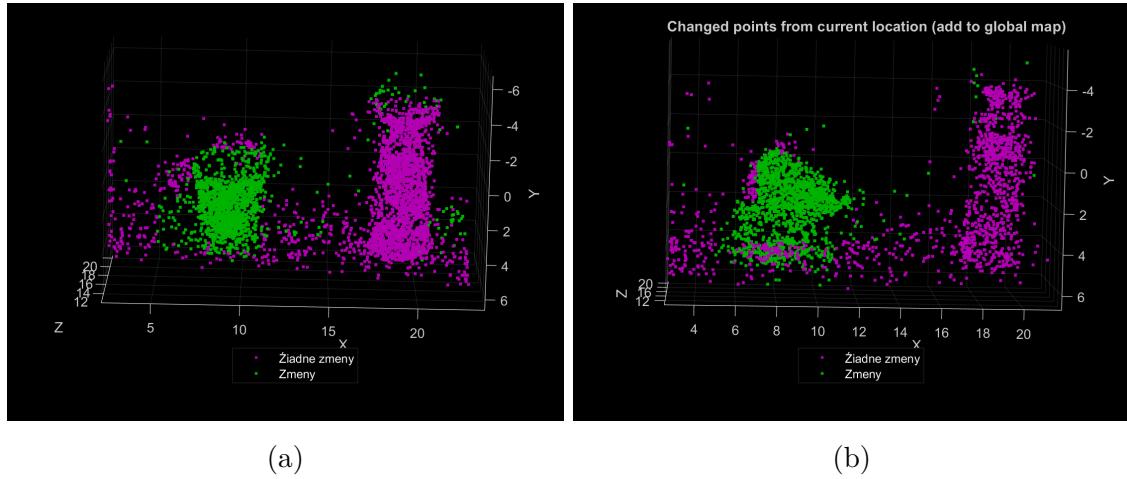
Obr. A.16: Pôvodná testovacia trasa č. 104



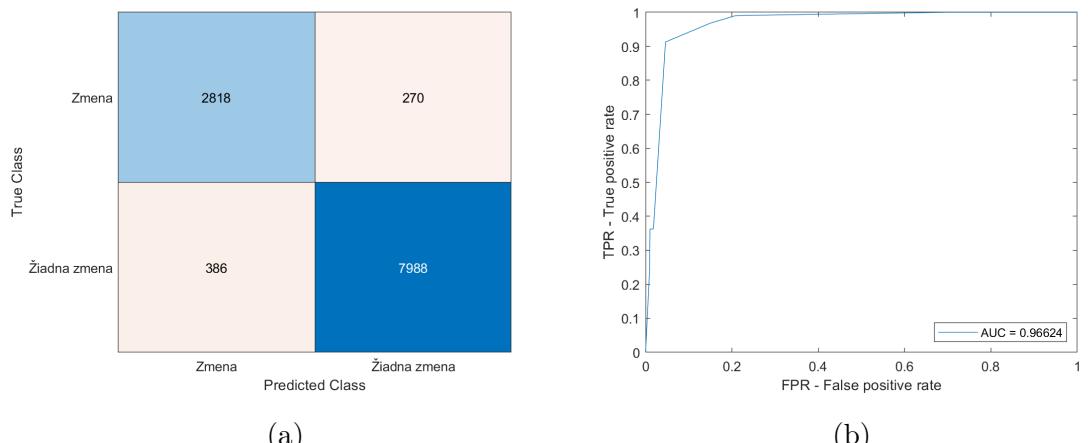
Obr. A.17: Aktuálna testovacia trasa č. 105



Obr. A.18: a) Pôvodná testovacia trasa č. 104, b) Aktuálna testovacia trasa č. 105



Obr. A.19: a) Zmeny v pôvodnej trase č. 104, b) Zmeny v aktuálnej trase č. 105



Obr. A.20: a) Matica zámen pre trasy 104 a 105, b) ROC krivka pre trasy 104 a 105

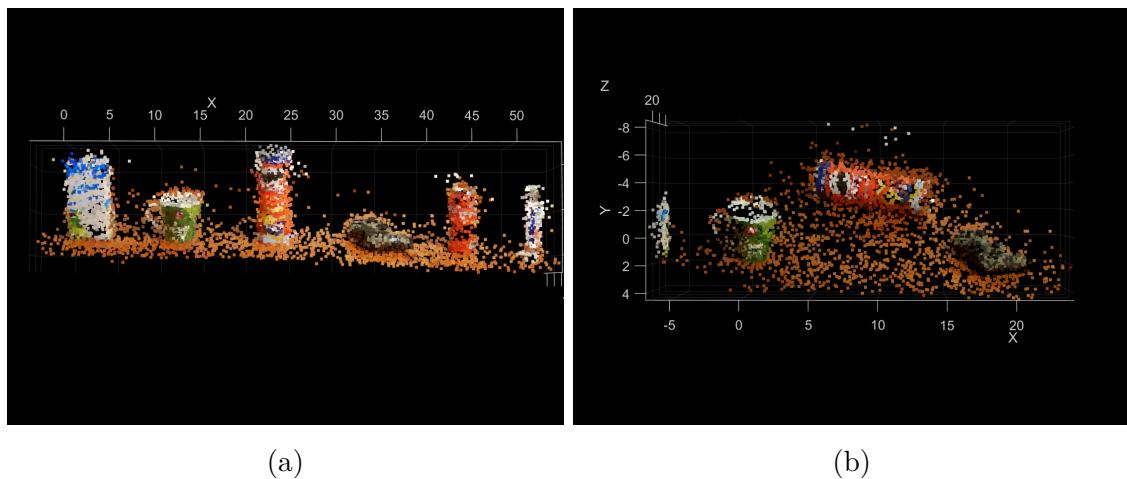
A.5 Trasa č. 106 a 107



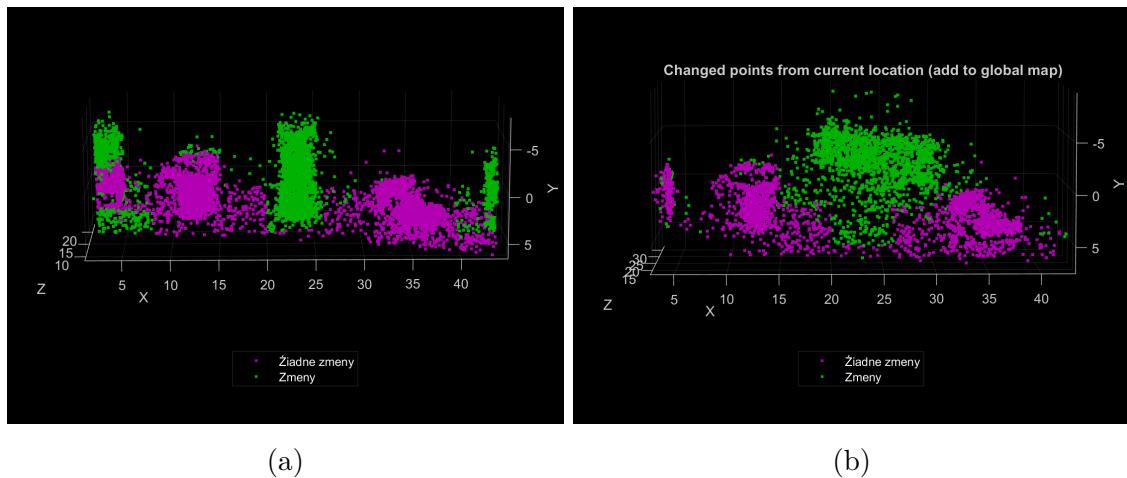
Obr. A.21: Pôvodná testovacia trasa č. 106



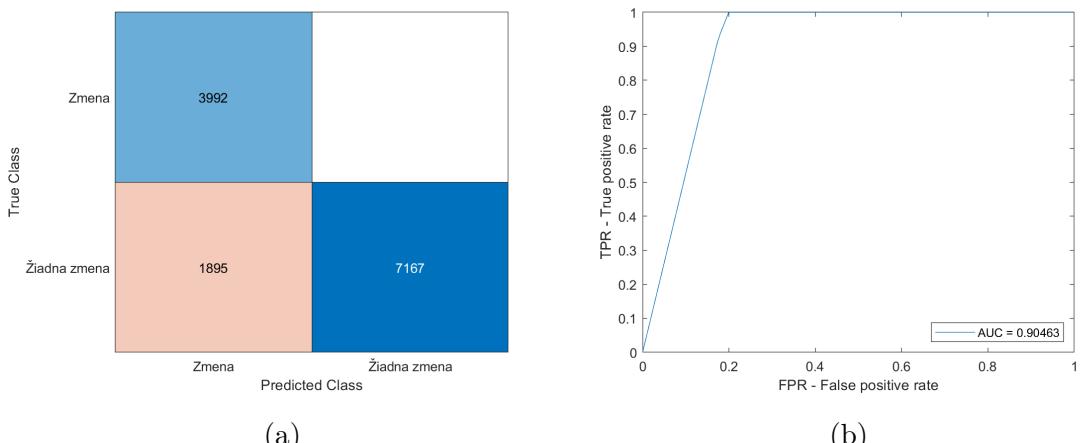
Obr. A.22: Aktuálna testovacia trasa č. 107



Obr. A.23: a) Pôvodná testovacia trasa č. 106, b) Aktuálna testovacia trasa č. 107



Obr. A.24: a) Zmeny v pôvodnej trase č. 106, b) Zmeny v aktuálnej trase č. 107



Obr. A.25: a) Matica zámen pre trasy 106 a 107, b) ROC krivka pre trasy 106 a 107

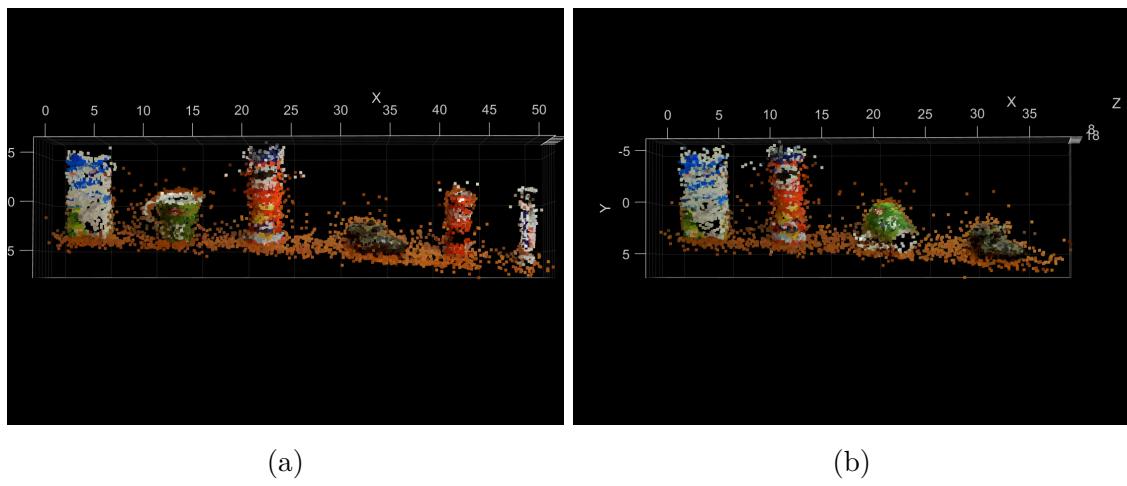
A.6 Trasa č. 111 a 112



Obr. A.26: Pôvodná testovacia trasa č. 111



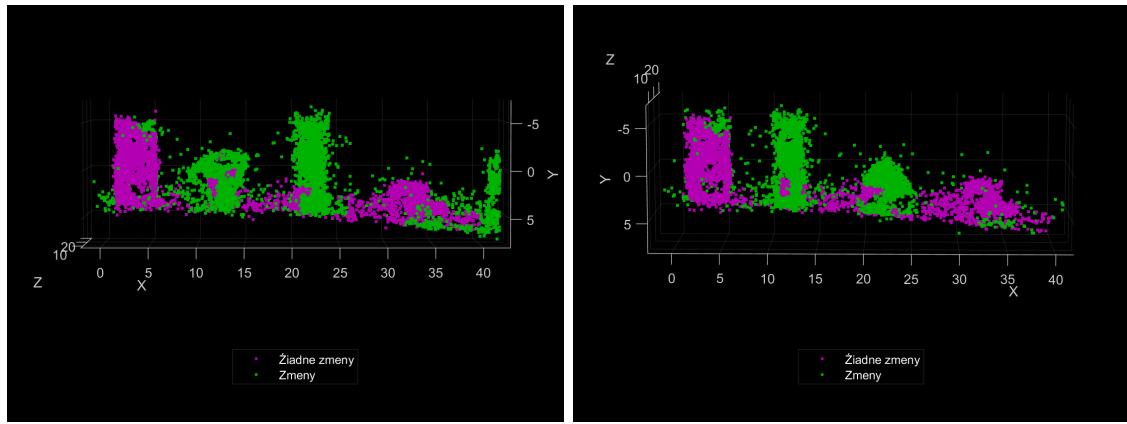
Obr. A.27: Aktuálna testovacia trasa č. 112



(a)

(b)

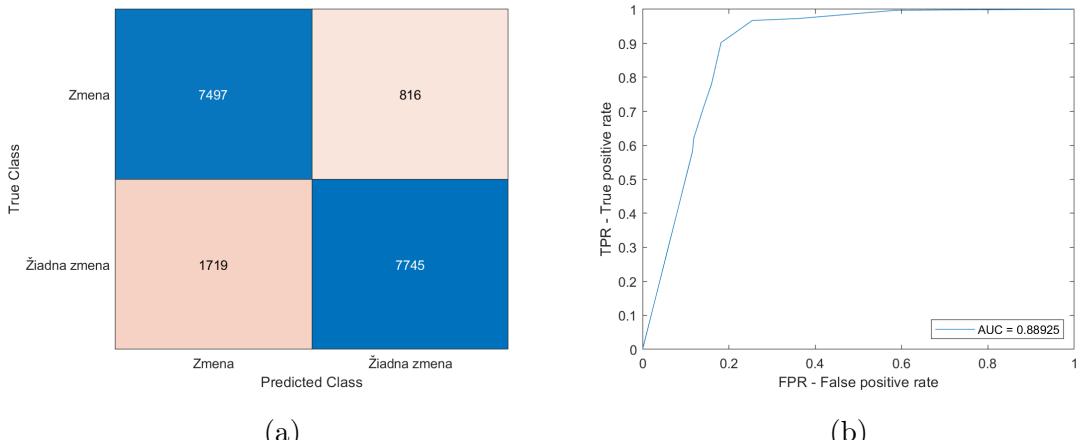
Obr. A.28: a) Pôvodná testovacia trasa č. 111, b) Aktuálna testovacia trasa č. 112



(a)

(b)

Obr. A.29: a) Zmeny v pôvodnej trase č. 111, b) Zmeny v aktuálnej trase č. 112



Obr. A.30: a) Matica zámen pre trasy 111 a 112, b) ROC krivka pre trasy 111 a 112

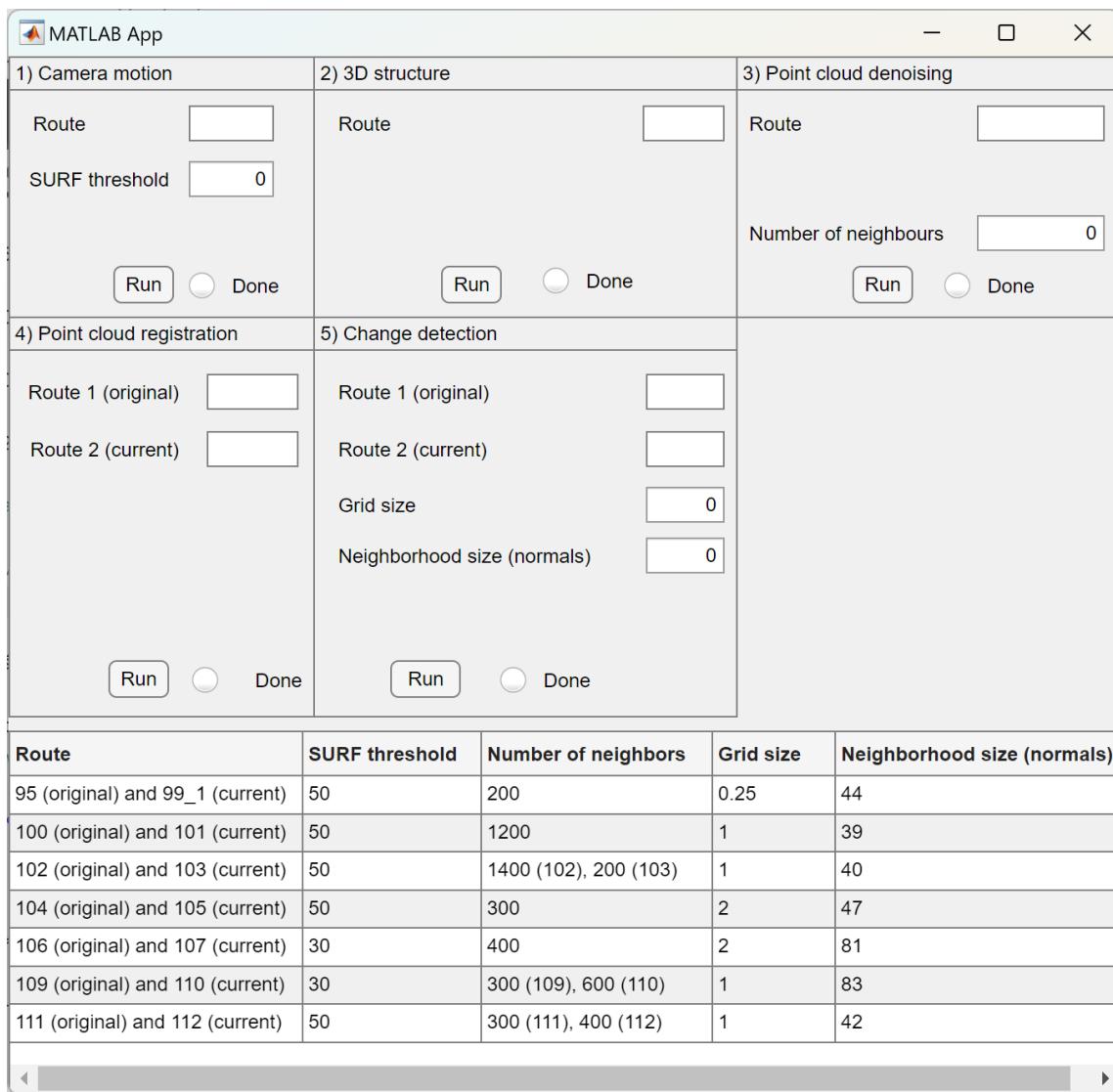
A.7 Tabuľka parametrov matice zámen

Tab. A.1: Tabuľka parametrov matice zámen testovacích trás

Pôvodná trasa	Aktuálna trasa	Senzitivita (TPR)	Špecificita (TNR)	Presnosť	AUC
95	99_1	0,879	0,807	0,827	0,869
100	101	0,955	0,847	0,875	0,931
102	103	0,976	0,916	0,924	0,979
104	105	0,913	0,954	0,943	0,966
106	107	1	0,791	0,855	0,905
109	110	0,973	0,798	0,896	0,892
111	112	0,902	0,818	0,857	0,889

B Demonštračná aplikácia

Za účelom testovania algoritmov bola vytvorená aplikácia, na obrázku B.1, ktorá postupne prevedie užívateľa cez všetky moduly systému detekcie zmien. V dolnej tabuľke sa nachádzajú hodnoty nastaviteľných parametrov. Je možné spustiť ktorýkoľvek modul, nemusí sa teda nasledovať stanovené poradie. Aplikácia sa nachádza na priloženom pamäťovom médiu.



Obr. B.1: Aplikácia demonštrujúca vytvorené algoritmy

C Obsah prílohy na CD

```
/..... koreňový adresár priloženého archív  
└─ change_detection.adresár, ktorý obsahuje implementáciu systému detekcie  
    zmien  
      └─ change_detection_demonstration.mlapp..... demonštračná aplikácia  
      └─ modules..... jednotlivé moduly systému  
        └─ sfm_odometry..... modul na výpočet pohybu kamery  
          └─ sfmOdometry.m  
          └─ functions  
            └─ calcPointsDistance.m  
            └─ detectAndExtractFeatures.m  
            └─ estimateRelativePose.m  
            └─ find3Dto2DCorrespondences.m  
            └─ checkParallax.m  
            └─ saveOdometry.m  
        └─ sfm_structure..... modul na určenie 3D štruktúry  
          └─ sfmStructure.m  
          └─ functions  
            └─ detectAndExtractFeatures.m  
            └─ estimateEipolarInliers.m  
            └─ checkParallax.m  
            └─ pointCloudColor.m  
            └─ saveStructure.m  
        └─ pc_denoise ..... modul na odstránenie šumu z mračna bodov  
          └─ pcDenoise.m  
          └─ functions  
            └─ savePCDenoise.m  
        └─ pc_registration..... modul registrácie mračien bodov  
          └─ pcRegistration.m  
          └─ functions  
            └─ savePCRegistration.m  
            └─ select2DFeaturesInView.m  
            └─ tformAbsoluteOrient.m  
    └─ change_detection ..... modul detekcie zmien  
      └─ changeDetection.m  
      └─ functions  
        └─ addPoints.m  
        └─ createOH.m  
        └─ detectSurfaceChange.m  
        └─ drawVoxels.m  
        └─ evaluateChangeDetection.m  
        └─ matchSurface.m  
        └─ padding.m  
        └─ pointCloudShow.m  
        └─ pointCloudShowPair.m
```

```
    ptCloudNormals.m
    removePoints.m
    removePointsInVoxels.m
    selectFromROI.m
    selectGTPoints.m
    selectPoints.m
    transformPointsInROI.m
    voxelize.m
data.. Vstupno-výstupné dátá, ktoré si medzi sebou predávajú jednotlivé
moduly
evaluation_ground_true
images
key_frames_indices
point_clouds
denoised
raw
transformations
vSet_odometry
vSet_structure
wpSet_structure
calibrationSession.mat
support_scripts.....Podporné skripty, ktoré nie sú súčasťou modulov
groundTrueMarking.m
img_preprocessing.m
```

D Obsah elektronickej prílohy

```
/..... koreňový adresár priloženého archív  
└─ change_detection.adresár, ktorý obsahuje implementáciu systému detekcie  
    zmien  
      └─ modules..... jednotlivé moduly systému  
        └─ sfm_odometry..... modul na výpočet pohybu kamery  
          └─ sfmOdometry.m  
          └─ functions  
            └─ calcPointsDistance.m  
            └─ detectAndExtractFeatures.m  
            └─ estimateRelativePose.m  
            └─ find3Dto2DCorrespondences.m  
            └─ checkParallax.m  
            └─ saveOdometry.m  
        └─ sfm_structure..... modul na určenie 3D štruktúry  
          └─ sfmStructure.m  
          └─ functions  
            └─ detectAndExtractFeatures.m  
            └─ estimateEipolarInliers.m  
            └─ checkParallax.m  
            └─ pointCloudColor.m  
            └─ saveStructure.m  
        └─ pc_denoise ..... modul na odstránenie šumu z mračna bodov  
          └─ pcDenoise.m  
          └─ functions  
            └─ savePCDenoise.m  
        └─ pc_registration..... modul registrácie mračien bodov  
          └─ pcRegistration.m  
          └─ functions  
            └─ savePCRegistration.m  
            └─ select2DFeaturesInView.m  
            └─ tformAbsoluteOrient.m  
      └─ change_detection ..... modul detekcie zmien  
        └─ changeDetection.m  
        └─ functions  
          └─ addPoints.m  
          └─ createOH.m  
          └─ detectSurfaceChange.m  
          └─ drawVoxels.m  
          └─ evaluateChangeDetection.m  
          └─ matchSurface.m  
          └─ padding.m  
          └─ pointCloudShow.m  
          └─ pointCloudShowPair.m  
          └─ ptCloudNormals.m
```

```
    └── removePoints.m
    └── removePointsInVoxels.m
    └── selectFromROI.m
    └── selectGTPoints.m
    └── selectPoints.m
    └── transformPointsInROI.m
    └── voxelize.m
└── support_scripts..... Podporné skripty, ktoré nie sú súčasťou modulov
    ├── groundTrueMarking.m
    └── img_preprocessing.m
```