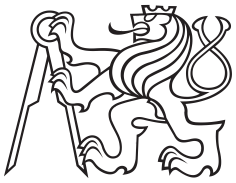
Bakalářská práce



České
vysoké
učení technické
v Praze

F3

Fakulta elektrotechnická
Katedra kybernetiky

SLAM v prostředí stožáru vysokého napětí

Jan Hošťálek

Vedoucí: RNDr. Petr Štěpán, Ph.D.
Obor: Základy umělé inteligence a počítačových věd
Studijní program: Otevřená informatika
Květen 2022

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

I. OSOBNÍ A STUDIJNÍ ÚDAJE

Příjmení: **Hošťálek** Jméno: **Jan** Osobní číslo: **492268**

Fakulta/ústav: **Fakulta elektrotechnická**

Zadávající katedra/ústav: **Katedra kybernetiky**

Studijní program: **Otevřená informatika**

Specializace: **Základy umělé inteligence a počítačových věd**

II. ÚDAJE K BAKALÁŘSKÉ PRÁCI

Název bakalářské práce:

SLAM v prostředí stožáru vysokého napětí

Název bakalářské práce anglicky:

SLAM around High-Voltage Mast

Pokyny pro vypracování:

1) Seznamte se s principy rekonstrukce prostředí ze stereo kamer.
2) Navrhněte algoritmus, který by na základě známé změny polohy dronu byl schopen vytvořit model jednoduchého stožáru vysokého napětí. Uvažujte, že dron se bude pohybovat v konstantní výšce bez náklonů.
3) Navržený algoritmus otestujte v simulačním prostředí gazebo na různě velikých modelech stožáru vysokého napětí. Otestujte algoritmus na modelu zjednodušené termální kamery a na parametrech visuální kamery s rozlišením kolem 2MP.
4) Zjistěte, při jaké nepřesnosti v určení změny polohy dronu bude ještě algoritmus dávat spolehlivé výsledky.

Seznam doporučené literatury:

[1] Hartley, R., & Zisserman, A. (2004). Multiple View Geometry in Computer Vision (2nd ed.). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511811685
[2] Duberg, Daniel, and Patric Jensfelt. "UFOMap: An efficient probabilistic 3D mapping framework that embraces the unknown." IEEE Robotics and Automation Letters 5.4 (2020): 6411-6418.
[3] Nalpantidis Lazaros, Georgios Christou Sirakoulis & Antonios Gasteratos (2008) Review of Stereo Vision Algorithms: From Software to Hardware, International Journal of Optomechatronics, 2:4, 435-462, DOI: 10.1080/15599610802438680

Jméno a pracoviště vedoucí(ho) bakalářské práce:

RNDr. Petr Štěpán, Ph.D. Multirobotické systémy FEL

Jméno a pracoviště druhé(ho) vedoucí(ho) nebo konzultanta(ky) bakalářské práce:

Datum zadání bakalářské práce: **02.02.2022** Termín odevzdání bakalářské práce: **20.05.2022**

Platnost zadání bakalářské práce: **30.09.2023**

RNDr. Petr Štěpán, Ph.D.	prof. Ing. Tomáš Svoboda, Ph.D.	prof. Mgr. Petr Páta, Ph.D.
podpis vedoucí(ho) práce	podpis vedoucí(ho) ústavu/katedry	podpis děkana(ky)

III. PŘEVZETÍ ZADÁNÍ

Student bere na vědomí, že je povinen vypracovat bakalářskou práci samostatně, bez cizí pomoci, s výjimkou poskytnutých konzultací. Seznam použité literatury, jiných pramenů a jmen konzultantů je třeba uvést v bakalářské práci.

. Datum převzetí zadání

Podpis studenta

© ČVUT v Praze, Design: ČVUT v Praze, VIC

Poděkování

Děkuji vedoucímu práce doktoru Petru Štěpánovi za vedení práce, cenné rady a podněty.

Prohlášení

Prohlašuji, že jsem předloženou práci vypracoval samostatně a že jsem uvedl veškeré použité informační zdroje v souladu s Metodickým pokynem o dodržování etických principů při přípravě vysokoškolských závěrečných prací.

V Praze, dne 20. 5. 2022

Jan Hošťálek

Abstrakt

Práce je úvodem do problematiky stereo vidění, věnuje se rekonstrukci stožáru vysokého napětí stereo vizuálním a termálním párem kamer. První částí navrženého postupu je tvorba referenčního modelu na základě známé polohy kamery s využitím knihovny UFOMap, nebo metody Parzen window. Postup respektuje zadané omezení texturové neodlišitelnosti dříku stožáru. Spolehlivost řešení jsme ověřili na vytvořené kolekci snímků stožáru ze vzdálenosti 6 a 8 metrů, pořízených termokamerou s rozlišením 160×120 px a vizuální kamerou s rozlišením 1920×1080 px. V druhé části jsme navrhli postup, kterým jsme schopni najít původní souřadnice kamery při pořízení libovolného snímku s využitím modelu z předchozího kroku. V případě termokamery jsme při snímání stožáru z přibližné vzdálenosti 8 m dosáhli průměrné odchylky do 9 cm, v případě vizuální do 2 cm.

Klíčová slova: stereo vizuální a termální kamera, stožár vysokého napětí, ICP, k-NN, UFOMap, Gazebo, Parzen window, k-means

Vedoucí: RNDr. Petr Štěpán, Ph.D.

Abstract

The thesis is an introduction to stereo vision, it deals with the reconstruction of a high voltage mast by stereo visual and thermal pair of cameras. The first part of the proposed procedure is the creation of a reference model based on a known camera position using the UFOMap library or the Parzen window method. The procedure respects the specified texture indistinguishability constraint of the mast shank. The reliability of the solution was verified on a collection of images of the mast taken from 6 and 8 meters distance with a thermal camera with 160×120 px resolution and a visual camera with 1920×1080 px resolution. In the second part, we proposed a procedure by which we are able to find the original coordinates of the camera from any image using the model from the previous step. In the case of the thermal camera, we obtained an average deviation of up to 9 cm when the mast was imaged from an approximate distance of 8 m, and up to 2 cm in the case of the visual camera.

Keywords: stereo visual and thermal camera, high voltage mast, ICP, k-NN, UFOMap, Gazebo, Parzen window, k-means

Obsah

1 Úvod **1**

2 Matematický aparát **3**

2.1 Základní značení 3

2.2 Rotační matice 3

2.3 Homogenní souřadnice 4

3 Model kamery **7**

3.1 Princip kamery s dírkovou komorou 7

3.1.1 Matice kamery 7

3.1.2 Vnější parametry 8

3.2 Projekce $\boldsymbol{p}_w$ do podprostoru $\mathbf{I}$ . . . 8

4 Rekonstrukce prostředí stereo párem kamer **11**

4.1 Rekonstrukce vzdálenosti s využitím stereo páru kamer 11

4.2 Epipolární geometrie 12

4.3 Standardní metody hledání geometrických protějšků 13

4.3.1 Metody plošné korelace 13

4.3.2 Metody porovnávání charakteristických rysů 13

5 Simulační prostředí **15**

5.1 Stožár vysokého napětí 15

5.2 MRS UAV systém 16

5.3 Stereo vizuální pár tvořený dvěma Mobius kamerami 16

5.4 Stereo termokamera 17

5.5 Dron 18

6 Návrh řešení **19**

6.1 Tvorba referenčního modelu 20

6.1.1 Metodou Parzen window 21

6.1.2 Využitím knihovny UFOMap 21

7 Zpracování snímku s neznámou polohou pořízení **25**

7.1 Navržený přístup 25

7.2 Překrývání stojek stožáru 32

8 Závěr **35**

Literatura **37**

A Ověření navrženého postupu na kolekci snímků zjednodušeného zábradlí **39**

Obrázky

2.1 Značení rotací kolem primárních os. 3

3.1 Projekce bodu **P** do afinního podprostoru obrázku. 10

4.1 Určení vzdálenosti bodu **P** využitím podobnosti trojúhelníků [6]. 12

4.2 Epipolární geometrie - geometrické protějšky $\boldsymbol{x_L}$. 13

5.1 Schéma základních konstrukčních částí a ukázka použitého modelu stožáru vysokého napětí. 15

5.2 Snímky pořízené stereo párem kamer Mobius s rozlišením 1920×1080 px (FOV 90°). 17

5.3 Snímky pořízené stereo párem kamer s parametry termokamery LEPTON 3.5. s rozlišením 160×120 px (FOV 57°). 17

5.4 Dron f550 v simulačním prostředí Gazebo. Stereo pár kamer Mobius jsme umístili nad dron bez držáku. Šířku základny jsme zvolili $b = 35$ cm. 18

6.1 Zjednodušení stožáru dostupného z Aerial-Core project. 19

6.2 Určení nejpravděpodobnějšího umístění základny dříku stožáru metodou Parzen window. Jednotlivá okna jsou zobrazena odstíny černé. Tmavší odstíny představují vyšší hladinu pravděpodobnosti. Pozice kamery při pořizování snímku je znázorněna trojúhelníkem náhodné barvy. 22

6.3 Promítnutí mračna bodů do ufomapy. Stožár rekonstruován z 16 snímků při známé pozici kamery. . 23

7.1 Hledání rotace a posunutí navrženým přístupem. Skutečný snímek pořízen z pohybu po kružnici kolem stožáru pod úhlem 22.5°. . . . 26

7.2 Výsledky navrženého algoritmu. Rekonstruováno z 12 snímků **stereo vizuální kamery** - pohyb po kružnici kolem stožáru. Barvy jsou volené náhodně a slouží pouze k odlišení bodů příslušících rozdílným snímkům. 30

7.3 Výsledky navrženého algoritmu. Rekonstruováno z 12 snímků **stereo termokamery** - pohyb po kružnici kolem stožáru. 31

7.4 Pohled na stožár stereo párem Mobius kamer. 32

7.5 Navržené řešení problému překrývajících se stojek. Zvýšením počtu potenciálních komplementárních dvojic jsme dosáhli výrazně přesnějších výsledků, za cenu vyšší výpočetní náročnosti. 33

A.1 Ukázka několika snímků z kolekce zjednodušeného zábradlí ze vzdálenosti 8 m vizuální kamerou s rozlišením 1 MP. 39

A.2 Výsledky promítnutí odhadovaných původních souřadnic stojek zábradlí. Kolekce obsahuje 5 snímků zábradlí ze vzdálenosti 6 m a 5 snímků ze vzdálenosti 8 m. 40

Tabulky

7.1 Výsledky navrženého přístupu k určení souřadnic základny dříku stožáru. 28

7.2 Zrychlení části algoritmu hledáním vhodné rotace translace pouze dvou bodů, v porovnání s přesnějším ale pomalejším párováním všech bodů modelu. 28

Kapitola 1

Úvod

Autonomní robot k úspěšné navigaci v neznámém prostředí nutně potřebuje mít schopnost zmapovat překážky a lokalizovat svoji polohu vzhledem k těmto překážkám. Této úloze se věnují metody Simultánní Lokalizace a Mapování (SLAM). Relativní polohu vůči startovnímu bodu lze u pozemních robotů určit z odometrických dat. V případě létajících robotů lze na velmi krátkou dobu spočítat polohu z inerciálních jednotek IMU. Kumulací nepřesností v získávaných datech ztrácíme informaci o skutečné poloze dronu. Ke korekci lze využít data ze snímání objektů okolního prostředí. Nalezené souřadnice vybraného objektu, rekonstruovaného z aktuálního snímku, porovnáme se souřadnicemi stejného objektu z předchozích snímků [1]. Cílem této práce je nejdříve vytvořit model základny stožáru vysokého napětí, který následně využijeme k lokalizaci dronu.

Pro mapování prostředí v okolí robota se v praxi často využívá senzorů založených na odrazu laserového paprsku. Tyto senzory jsou ale těžké a omezují dobu letu dronů. Proto v naší práci pro rekonstrukci okolního prostředí využíváme stereo vizuálních, nebo termálních kamer. Na rozdíl od paprskových senzorů potřebujeme rekonstruovat hloubkovou mapu, protože na vstupu jsou pouze 2 paralelní RGB snímky dané scény. Využít lze například veřejně dostupné metody knihovny OpenCV [6], podrobněji v podkapitole 6.

I přestože již v dnešní době existuje mnoho spolehlivých metod pro rekonstrukci původní 3D scény [1], tak v naší práci navrhujeme postup přizpůsobený omezením daného prostředí, ve kterém se dron pohybuje. Tím je předpoklad nerozpoznatelnosti odlišností v texturách jednotlivých konstrukčních částí stožáru. Z tohoto důvodu v simulačním prostředí Gazebo používáme barevně

jednolité modely stožáru.

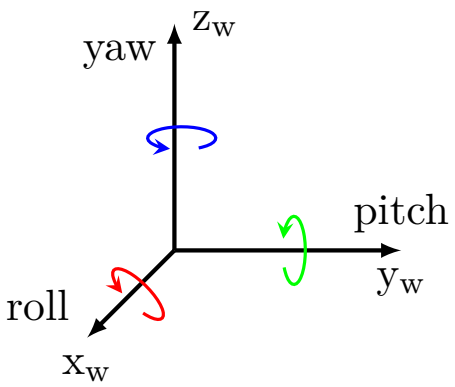
Navržený postup nejdříve, s využitím knihovny UFOMap, nebo metody Parzen window, zpracuje snímky s přibližně známou polohou pořízení a rekonstruuje tak původní model stožáru, viz kapitola 6. Znalosti souřadnic kamery při pořizování snímku můžeme v praxi docílit manuální tvorbou snímků ze známých pozic v okolí stožáru, nebo uchováním dat z gps modulu dronu. V druhé části vytváříme algoritmy, které mapují získaná data ze stereo kamery na model vytvořený v předchozím kroku bez znalosti polohy dronu, viz kapitola 7.

Kapitola 2

Matematický aparát

2.1 Základní značení

Primární osy v $\mathbb{R}^3$ značíme xyz a ve značení úhlů natočení kolem os dodržujeme konvenci yaw (γ), pitch (β) a roll (α) viz obrázek 2.1.



Obrázek 2.1: Značení rotací kolem primárních os.

2.2 Rotační matice

Rotace kolem primárních os lze nejjednodušeji definovat pomocí rotačních matic [3]. Rotaci o úhel α kolem osy x lze zapsat pomocí maticového násobení

maticí

$$\mathbf{R}_x(\alpha) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & cos(\alpha) & -sin(\alpha) \\ 0 & sin(\alpha) & cos(\alpha) \end{bmatrix}. \tag{2.1}$$

Rotaci kolem osy y lze obdobně přepsat do matice

$$\mathbf{R}_y(\beta) = \begin{bmatrix} cos(\beta) & 0 & sin(\beta) \\ 0 & 1 & 0 \\ -sin(\beta) & 0 & cos(\beta) \end{bmatrix}. \tag{2.2}$$

Poslední rotační matice představuje otočení o úhel γ kolem osy z

$$\mathbf{R}_z(\gamma) = \begin{bmatrix} cos(\gamma) & -sin(\gamma) & 0 \\ sin(\gamma) & cos(\gamma) & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{2.3}$$

Zřetězením těchto tří rotací definujeme obecnou rotaci. Příslušnou matici obecné rotace získáme maticovým násobením

$$\mathbf{R}_{xyz}(\alpha, \beta, \gamma) = \mathbf{R}_x(\alpha)\mathbf{R}_y(\beta)\mathbf{R}_z(\gamma) \tag{2.4}$$

$$\begin{bmatrix} cos(\beta)cos(\gamma) & sin(\alpha)sin(\beta)cos(\gamma) - cos(\alpha)sin(\gamma) & cos(\alpha)sin(\beta)cos(\gamma) + sin(\alpha)sin(\gamma) \\ cos(\beta)sin(\gamma) & sin(\alpha)sin(\beta)sin(\gamma) + cos(\alpha)cos(\gamma) & cos(\alpha)sin(\beta)sin(\gamma) - sin(\alpha)cos(\gamma) \\ -sin(\beta) & sin(\alpha)cos(\beta) & cos(\alpha)cos(\beta) \end{bmatrix}.$$

Při řetězení rotací je důležité dodržovat definované pořadí jednotlivých rotací, protože změnou pořadí α, β, γ získáme různé rotace.

2.3 Homogenní souřadnice

Abychom mohli jednoduše maticovým násobením provádět rotace a translace v libovolné dimenzi, používáme homogenní souřadnice. Pro názornost v $\mathbb{R}^2$

$$u = \frac{u'}{w} \qquad\qquad v = \frac{v'}{w}$$

$$\boldsymbol{p} = \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} wu \\ wv \\ w \end{bmatrix} = \begin{bmatrix} u' \\ v' \\ w \end{bmatrix} \tag{2.5}$$

Obdobně pro $\mathbb{R}^3$

$$\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} \rightarrow \begin{bmatrix} \frac{x}{w} \\ \frac{y}{w} \\ \frac{z}{w} \end{bmatrix}. \tag{2.6}$$

Největší výhodou při práci s homogenními souřadnicemi je, že pohyb dronu ve světových souřadnicích, tedy rotaci a translaci, můžeme zapsat pomocí jediné matice

$$\mathbf{A} = \begin{bmatrix} \mathbf{R} & \boldsymbol{t} \\ \mathbf{0} & 1 \end{bmatrix}, \tag{2.7}$$

která představuje orientaci dronu definovanou rotací $\mathbf{R}$ a posunutí dronu o vektor $\boldsymbol{t}$ [4]. Matice $\mathbf{A}$ samozřejmě respektuje všechny maticové operace, proto pokud bychom měli více rotací/translací - více matic $\mathbf{A}$, můžeme je maticovým násobením zřetězit do jedné. Polohu libovolného bodu zapsaného ve tvaru 2.7 transformujeme do globálních souřadnic vzorcem

$$\boldsymbol{v'} = \mathbf{A}\boldsymbol{v}. \tag{2.8}$$

Kapitola 3

Model kamery

3.1 Princip kamery s dírkovou komorou

Princip dírkové komory byl poprvé popsán již několik století před naším letopočtem. Jednoduché optické zařízení je konstrukčně založeno na uzavřené skříňce s malým otvorem na jedné ze stran, kterým prochází paprsky světla odražené od snímaného objektu. Na modelu jednoduché pinhole kamery je v podkapitole 3.2 definováno promítnutí reálného bodu na průmětnu - do souřadnic obrázku.

V práci nejčastěji využíváme tři základní podprostory. Hlavní souřadnicový systém, vůči kterému definujeme pozici stereo kamery a snímaného objektu, značíme **W**. Afinní podprostor kamery značíme **C** a podprostor průmětny **I**. V cizojazyčné literatuře se jedná o world a camera frame a image plane.

3.1.1 Matice kamery

V naší práci předpokládáme znalost základních parametrů používané kamery. Výhodou práce se simulačním softwarem je možnost volby parametrů kamery. Pokud bychom chtěli pracovat s reálnými kamerami, museli bychom parametry hledat kalibrací. Výstupem kalibračního procesu je mimo jiné také matice

kamery

$$\mathbf{K} = \begin{bmatrix} f_x & s & o_x \\ 0 & f_y & o_y \\ 0 & 0 & 1 \end{bmatrix}, \tag{3.1}$$

kde f_x a f_y jsou ohniskové vzdálenosti ve směru osy x a y. Hodnoty se mohou lišit v případě, kdy pixely obrazu nejsou dokonale čtvercové. Parametry (o_x, o_y) určují optický střed (hlavní bod), v cizojazyčné literatuře nazývaný principal point. Optický střed je průsečík průmětny a osy objektivu, obvykle se nachází poblíž středu obrázku [5]. V naší práci neuvádíme, typicky nulové, zkosení obrazu s, které vyjadřuje kolmost os $\hat{x}_i$ a $\hat{y}_i$. Všechny parametry jsou uváděny v pixelech. Konkrétní parametry matice kamery využité v simulaci jsou uvedeny v podkapitole 5.3.

Pokud známe matici kamery, tak matice vnitřních parametrů je pouhé rozšíření matice kamery o sloupcový nulový vektor

$$\mathbf{M}_{\text{int}} = \left[\begin{array}{c|c} \mathbf{K} & \mathbf{0} \end{array}\right] = \begin{bmatrix} f_x & 0 & o_x & 0 \\ 0 & f_y & o_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}. \tag{3.2}$$

3.1.2 Vnější parametry

Vnější (extrinsic) parametry kamery zachycují pozici $\boldsymbol{c}_w$ a orientaci $\mathbf{R}$ kamery v globálním souřadnicovém systému $\mathbf{W}$, viz obrázek 3.1. Takto definovanou pozici kamery v prostoru nejčastěji zapisujeme v homogenních souřadnicích do matice vnějších parametrů

$$\mathbf{M}_{\text{ext}} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{3.3}$$

3.2 Projekce p_w do podprostoru I

$$\boldsymbol{p}_i = \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} \xleftarrow[projekce]{perspektivní} \boldsymbol{p}_c = \begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} \xleftarrow[souřadnic]{transformace} \boldsymbol{p}_w = \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}. \tag{3.4}$$

Předpokladem pro projekci bodu z globálního souřadnicového systému do souřadnicového systému obrázku je znalost vnitřních a vnějších parametrů kamery. Projekci poté můžeme rozložit na dva kroky. Prvním je transformace souřadnic z globálního souřadnicového systému do souřadnicového systému kamery. Ze známé pozice a natočení kamery v prostoru sestavíme matici vnějších parametrů a promítaný bod $\mathbf{P}$, přepsaný do homogenních souřadnic, maticovým násobením zobrazíme do afinního podprostoru kamery[7]. Průmět bodu do afinního podprostoru obrázku, $\mathbb{R}^3 \rightarrow \mathbb{R}^2$, docílíme perspektivní projekcí - maticovým násobením vnitřními parametry kamery

$$\begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} = \begin{bmatrix} x_i' \\ y_i' \\ w' \end{bmatrix} = \begin{bmatrix} z_c x_i \\ z_c y_i \\ z_c \end{bmatrix} = \begin{bmatrix} f_x x_c + z_c o_x \\ f_y y_c + z_c o_y \\ z_c \end{bmatrix} = \begin{bmatrix} f_x & 0 & o_x & 0 \\ 0 & f_y & o_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix}. \quad (3.5)$$

Z maticového násobení vyjádříme rovnice pro x_i a y_i

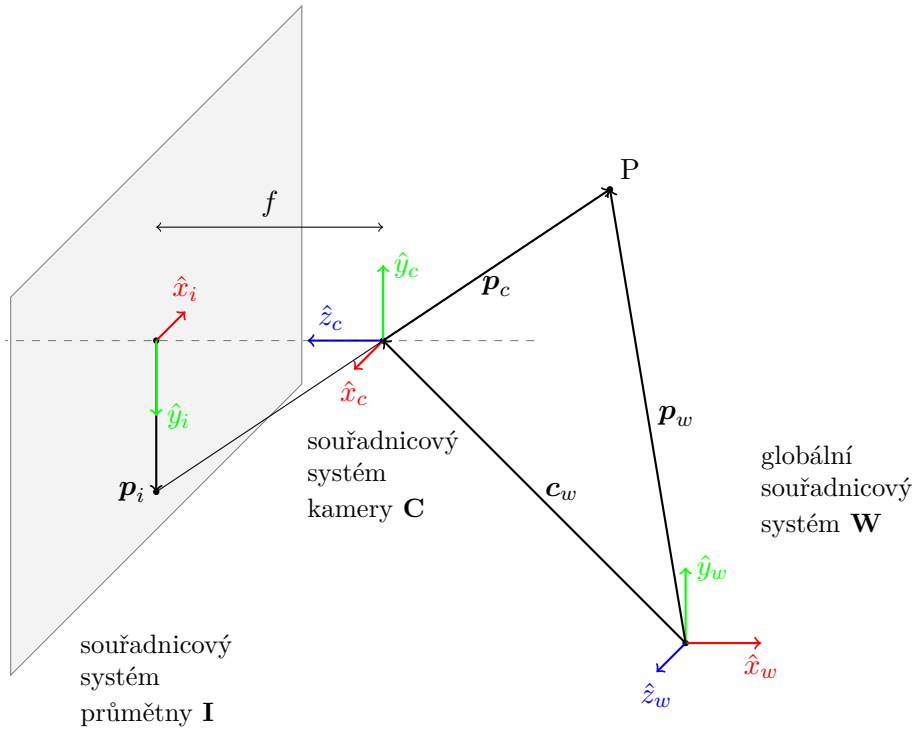
$$x_i = f_x \frac{x_c}{z_c} + o_x \qquad\qquad y_i = f_y \frac{y_c}{z_c} + o_y. \quad (3.6)$$

Výslednou projekci $\boldsymbol{p}_w$ z podprostoru $\mathbf{W}$ do souřadnic obrázku můžeme zjednodušit následovně

$$\boldsymbol{p}_i = \mathbf{M}_{\text{int}} \mathbf{M}_{\text{ext}} \boldsymbol{p}_w = \mathbf{P} \boldsymbol{p}_w, \quad (3.7)$$

kde $\mathbf{M}_{\text{int}}$ je matice vnitřních parametrů kamery, $\mathbf{M}_{\text{ext}}$ reprezentuje pozici kamery v lineárním podprostoru $\mathbf{W}$. Pokud zřetězíme všechna zobrazení do jediné projekční matice $\mathbf{P}$, můžeme jednoduchým přenásobením vektoru $\boldsymbol{p_w}$ z globálního souřadnicového systému dostat souřadnice bodu v rovině obrázku

$$\begin{bmatrix} x_i \\ y_i \\ w \end{bmatrix} = \begin{bmatrix} f_x & 0 & o_x & 0 \\ 0 & f_y & o_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}. \quad (3.8)$$



Obrázek 3.1: Projekce bodu **P** do afinního podprostoru obrázku.

Kapitola 4

Rekonstrukce prostředí stereo párem kamer

V současné době existuje mnoho algoritmů a přístupů k řešení problému rekonstrukce 3D scény. Hlavním cílem je získat z jednoho, nebo více obrázků původní souřadnice objektů v souřadnicovém systému **W**.

4.1 Rekonstrukce vzdálenosti s využitím stereo páru kamer

Informaci o vzdálenosti objektu ve směru osy pohledu kamery nejsme schopni získat pouze z jediného snímku. Možným způsobem řešení je přidání druhé kamery, kterou umístíme do souřadnicového systému první kamery na souřadnice (b,0,0) [7] - v našem případě jsme po otestování různých šířek (35, 70, 90, 110, 130 cm) základny b zvolili hodnotu 35 cm. Především z důvodu, aby při umístění na použitý dron kamera nepřesahovala šířku dronu, viz obrázek 5.4. Vzdálenost objektu, ve směru osy z, od ohniska kamery je možné dopočítat pomocí podobnosti trojúhelníků následovně

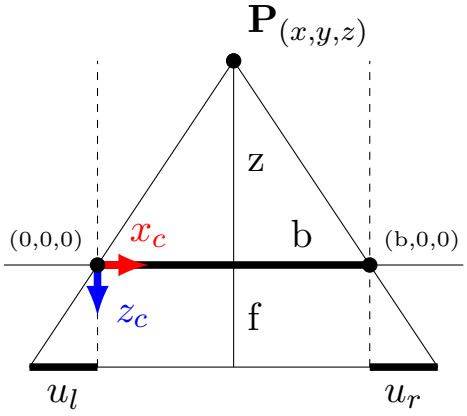
$$\begin{aligned} u_l &= f_x \frac{x}{z} + o_x & v_l &= f_y \frac{y}{z} + o_y \\ u_r &= f_x \frac{x-b}{z} + o_x & v_r &= f_y \frac{y}{z} + o_y, \end{aligned} \tag{4.1}$$

přepsáním do jednoho vektoru

$$(u_l, v_l) = (f_x \frac{x}{z} + o_x, f_y \frac{y}{z} + o_y)$$
$$(u_r, v_r) = (f_x \frac{x-b}{z} + o_x, f_y \frac{y}{z} + o_y). \tag{4.2}$$

Jednoduchými matematickými operacemi vyjádříme původní souřadnice v globálním souřadnicovém systému a dojdeme ke vztahům

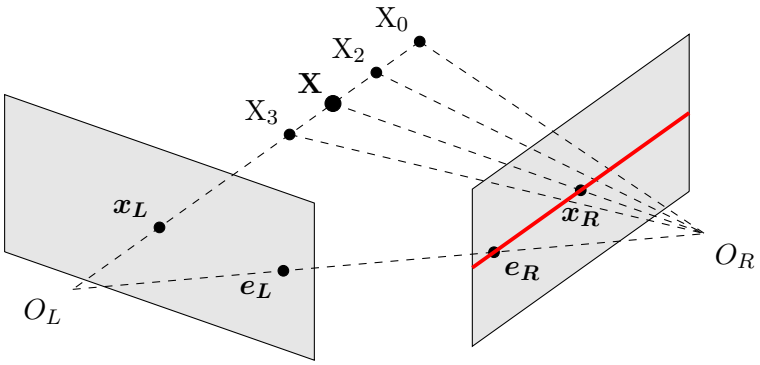
$$x = \frac{b(u_l - o_x)}{(u_l - u_r)} \qquad y = \frac{bf_x(v_l - o_y)}{f_y(u_l - u_r)} \qquad z = \frac{bf_x}{(u_l - u_r)}. \tag{4.3}$$



Obrázek 4.1: Určení vzdálenosti bodu **P** využitím podobnosti trojúhelníků [6].

4.2 Epipolární geometrie

Jedním z nejsložitějších problémů stereo vidění je přiřazení bodu z levého obrázku jeho geometrickému protějšku v obrázku pravém. Je zřejmé, že bychom mohli hledat v celém afinním podprostoru pravého obrázku. Zmenšit prohledávaný podprostor můžeme znalostí vzájemných pozic kamer a následnou úpravou obrázků - rektifikací. Vzhledem k použití simulovaných kamer nemusíme odstraňovat soudkovité zkreslení, které se často vyskytuje u reálných objektivů. Při použití reálných kamer lze využít metod OpenCV, které tyto problémy eliminují. Pokud máme takto připravené snímky, lze hledání geometrického protějšku bodu v levém obrázku zjednodušit do podprostoru $\mathbb{R}^1$, tzn. hledáme příslušný bod na epipolární přímce [5], viz obrázek 4.2.



Obrázek 4.2: Epipolární geometrie - geometrické protějšky $\boldsymbol{x_L}$.

4.3 Standardní metody hledání geometrických protějšků

4.3.1 Metody plošné korelace

Algoritmy pro hledání komplementárních dvojic bodů v levém a pravém snímku založené na hledání korelace mezi oblastmi pixelů nazýváme block matching algoritmy, zkráceně BM algoritmy. Tento typ metod hledá pro oblast v levém obrázku nejpodobnější oblast v obrázku pravém, například na epipolární přímce, s využitím například součtu absolutních rozdílů intenzit [2]. V rané fázi návrhu jsme provedli pokusy rekonstrukce hloubkové mapy nejznámějšími metodami implementovanými v knihovně OpenCV. Jedná se knihovní funkce StereoBM a StereoSGBM [6]. Vzhledem k definovaným omezením detekovaných objektů, viz kapitola 1, jsme ale nedosáhli uspokojivých výsledků.

4.3.2 Metody porovnávání charakteristických rysů

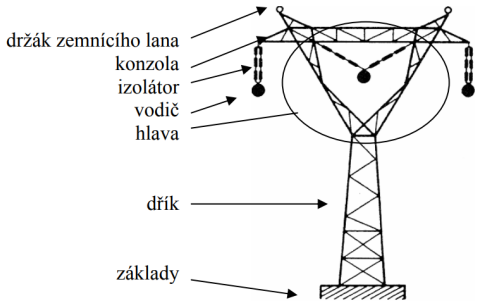
Algoritmy založené na hledání charakteristických rysů nazýváme feature matching algoritmy. V každém z obrázků nalezneme charakteristické body, kandidáty mohou být například rohy objektu, ty pak porovnáme s nalezenými body v opačném obrázku a určíme jejich nejpravděpodobnější komplementární pár. Pro účely testování tohoto přístupu jsme opět použili knihovní funkci OpenCV implementující algoritmus SIFT (Scale Invariant Feature Transform), kromě SIFTu lze ale využít i třeba SURF (Speeded-Up Robust Features) nebo ORB (Oriented FAST and Rotated BRIEF)[8]. Problém tohoto přístupu

byl především v případě využití termokamery, kde byl počet nalezených charakteristických bodů nedostačující.

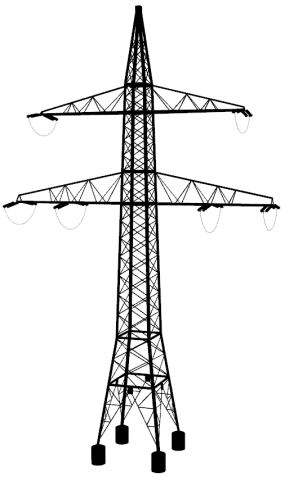
Kapitola 5

Simulační prostředí

5.1 Stožár vysokého napětí



(a) : Schéma stožáru vysokého napětí [10].



(b) : Použitý model Danube.

Obrázek 5.1: Schéma základních konstrukčních částí a ukázka použitého modelu stožáru vysokého napětí.

Z nepřeberného množství konstrukčních typů stožárů jsme využili veřejně dostupný model z projektu Aerial-Core, na kterém se podílí i skupina MRS[1], viz obrázek 5.1b. V naší práci se věnujeme především nalezení správných souřadnic opěrných stojek dříku stožáru. Dříkem stožáru je myšlena spodní

[1]Zdrojové soubory jsou veřejně dostupné z repozitáře `https://github.com/ctu-mrs/aerialcore_simulation`

rovnoběžná, nebo rozšiřující se část stožáru, viz schéma na obrázku 5.1.

5.2 MRS UAV systém

Pro tvorbu kolekcí snímků a následnou částečnou vizualizaci výsledků v naší práci využíváme veřejně dostupný balíček systémů vyvíjený skupinou MRS [13] pro systém ROS [14].

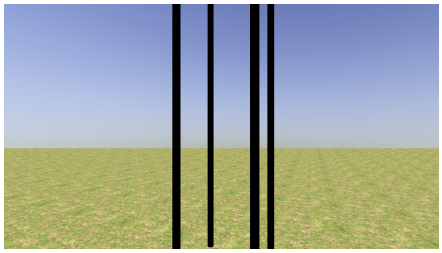
Pro vizualizaci a manipulaci s modely v globálních souřadnicích propojeného simulačního prostředí využíváme software Gazebo. Pro účely naší práce jsme vytvořili dva odlišné světy. První obsahuje model reálného stožáru, druhý jeho zjednodušenou verzi, viz kapitola 6.

Dalším 3D vizualizačním softwarem, určeným primárně pro vizualizaci dat z použitých senzorů v systému ROS, je RViz. Kromě zobrazení snímků z pohledu kamer, RViz v naší práci využíváme k zobrazení mračen bodů (pointcloud) a voxelů UFOMapy. Výhodou zobrazení výstupu knihovny UFOMap v RVizu je především možnost v reálném čase měnit její nastavení, konfigurovat promítání mračna bodů a modifikovat práh obsazenosti pro zobrazení voxelů. Konkrétnímu využití knihovny UFOMap se věnujeme v podkapitole 6.1.2.

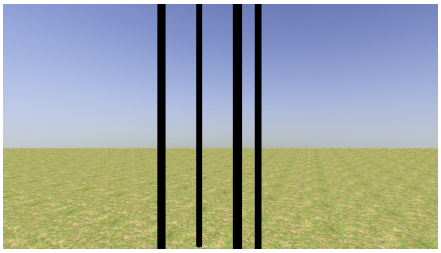
5.3 Stereo vizuální pár tvořený dvěma Mobius kamerami

V nabízených modulech umístitelných na dron bohužel nebyla pro naší práci vhodná stereo kamera. Proto jsme vytvořili vlastní stereo pár ze dvou samostatných Mobius modulů, viz obrázek 5.4. Výhodou je možnost úpravy šířky základny mezi kamerami, volba vnitřních a distorzních parametrů obou senzorů. Pro tvorbu vizuálních snímků jsme využili, kromě výchozího FOV 131°, především 90° FOV a nulové distorze. Dle zadání je výstupní rozlišení 1920×1080 px, což odpovídá přibližně 2 MP. Použitá matice vnitřních parametrů je při 90° FOV je následující

$$\mathbf{K}_{\text{viz}} = \begin{bmatrix} 960 & 0 & 960 \\ 0 & 540 & 540 \\ 0 & 0 & 1 \end{bmatrix}. \tag{5.1}$$



(a) : Pohled levé kamery.



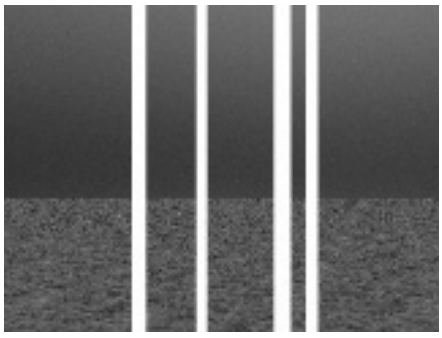
(b) : Pohled pravé kamery.

Obrázek 5.2: Snímky pořízené stereo párem kamer Mobius s rozlišením 1920×1080 px (FOV 90°).

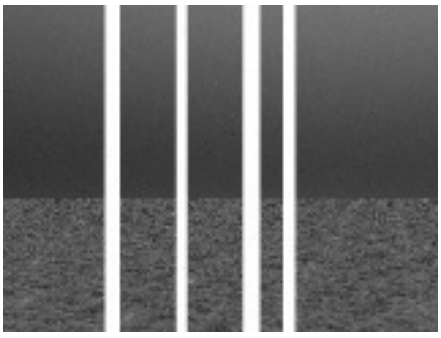
5.4 Stereo termokamera

Simulační prostředí kromě vhodné stereo vizuální kamery neposkytuje ani vhodně implementovanou stereo termokameru. Proto jsme se po konzultaci s vedoucím rozhodli nahradit teplotní mapu upraveným vizuálním výstupem kamery, viz obrázek 5.3b. Na navrhovaný algoritmus tento fakt nemá vliv, protože se nezabýváme rozpoznáváním stožáru v pořízeném snímku. Rozpoznání pixelů příslušících stojce dříku provádíme manuálně, nalezením oblastí stejné barvy. Hlavním rozdílem od vizuálního páru je ale konfigurace kamery. Využili jsme parametrů reálného typu stereo termokamery Lepton 3.5. Rozlišení na výstupu Lepton kamery je 160×120 px a horizontální FOV činí 57°. Příslušná matice vnitřních parametrů

$$\mathbf{K}_{\text{ter}} = \begin{bmatrix} 146.44 & 0 & 960 \\ 0 & 109.83 & 540 \\ 0 & 0 & 1 \end{bmatrix}. \tag{5.2}$$



(a) : Pohled levé kamery.

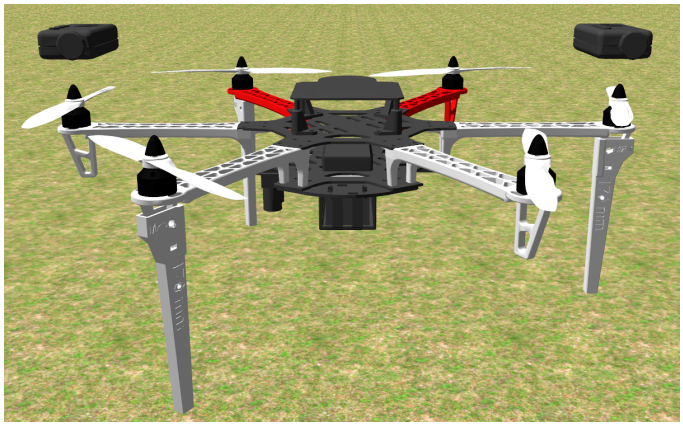


(b) : Pohled pravé kamery.

Obrázek 5.3: Snímky pořízené stereo párem kamer s parametry termokamery LEPTON 3.5. s rozlišením 160×120 px (FOV 57°).

5.5 Dron

Abychom mohli vytvořeným stereo párem kamer pohybovat v okolí stožáru, umístili jsme jej do prostoru nad dron, konkrétně jsme vybrali dostupný model f550. V praxi bychom museli kamery připevnit na konstrukční prvek, tak aby vzdálenost mezi kamerami byla požadovaných 35 cm. Pohybovat dronem v prostoru je možné manuálně, zadáním cílového bodu, nebo především zadáním požadované trajektorie, zadáním klíčových bodů. Připravili jsme trajektorii ve tvaru kružnice a dle zadání se dron pohybuje v konstantní výšce. Modifikace dronu, simulačního prostředí a konfiguraci kamer jsme prováděli na vzorovém příkladě waypoint-flier[2]



Obrázek 5.4: Dron f550 v simulačním prostředí Gazebo. Stereo pár kamer Mobius jsme umístili nad dron bez držáku. Šířku základny jsme zvolili $b = 35$ cm.

[2]Dostupný z repozitáře `https://github.com/ctu-mrs/example_ros_packages`

Kapitola 6

Návrh řešení

V rané fázi návrhu postupu jsme se věnovali promítnutí bodů ze souřadnicového systému obrázku (průmětny) do prostoru globálního souřadnicového systému. K úspěšnému nalezení příslušného bodu v prostoru, kterému odpovídá vybraný pixel, je zapotřebí nalézt hloubku ve které se původní bod nacházel.

(a) : Detail Danube stožáru. **(b) :** Zjednodušený stožár.

Obrázek 6.1: Zjednodušení stožáru dostupného z Aerial-Core project.

Vzhledem k ověření principů rekonstrukce opakujícího se vzoru, realistický model stožáru zjednodušíme odstraněním příčných konstrukčních prvků a zaměříme se na rekonstrukci stojek stožáru, viz obrázek 6.1. Namísto odhadu jednoho správného geometrického protějšku výše uvedenými metodami promítneme do původního podprostoru všechna možná přiřazení význačných bodů v levém obraze podobným bodům v obraze pravém. Jinými slovy, pro pixely příslušící stojce viditelné levou kamerou existuje v pravém snímku na epipolární přímce mnoho potenciálních protějšků, to má za důsledek průmět stojky do různé hloubky. Je zřejmé, že tímto vzniknou i kombinace promítnuté do záporné hloubky (za kamerou), o kterých můžeme s jistotou říci, že jsou falešné a můžeme je proto odstranit.

Hlavní myšlenkou navrženého algoritmu je, že ze všech kombinací se ty správné potvrdí průmětem opět všech kombinací rekonstruovaných ze snímku, který zabírá stožár z jiného úhlu. Využité metody, včetně podrobnějšího postupu tvorby referenčního modelu, popisujeme v podkapitole 6.1.

Druhá část navrženého algoritmu využívá vytvořený referenční model k určení polohy kamery a tím i ke správnému umístění nových průmětů bodů bez apriorní znalosti pozice a natočení kamery při pořízení rekonstruovaného snímku. Konkrétní postup a výsledky jsou uvedeny v kapitole 7.

Před procesem tvorby referenčního modelu snímky z kolekce nejdříve upravíme. Nejprve sjednotíme barevné spektrum obrázků z obou typů kamer. Snímky z vizuální kamery převedeme do odstínů šedé a již černobílé snímky z termokamery invertujeme, protože teplejší barvy jsou vyobrazeny světlejšími odstíny. V dalším kroku snížíme počet odstínů šedé aplikováním metody k-means, s například 3-4 centroidy. Tento krok nám zjednoduší hledání stojek v případě, že by z jiného úhlu pohledu měl stožár odlišný odstín.

Abychom snížili počet promítaných bodů na minimum, nehledáme geometrické protějšky pro všechny pixely příslušící jedné stojce, ale pouze pro střed stojky. Samotná detekce stojek - rozpoznání stožáru ve snímku, není předmětem naší práce, proto středy hledáme jednoduše jako středy nejtmavších oblastí. Snížení počtu potenciálních komplementárních dvojic je podstatné především pro snímky z vizuální kamery, protože pokud bychom hledali všechny kombinace geometrických protějšků pro všechny pixely dané stojky, časová náročnost by zamezila nalezení jakéhokoliv řešení.

6.1 Tvorba referenčního modelu

K odhadu matice vnějších parametrů kamery, která snímá stožár z neznámého úhlu, viz kapitola 7, využíváme v naší práci referenční model stožáru. K tvorbě takového modelu, ze kterého následně zjistíme nejpravděpodobnější umístění stojek dříku, jsme použili dva odlišné přístupy. Prvním z nich je Parzen window, kdy pracujeme pouze v jednom vodorovném řezu dříkem, tzn. v $\mathbb{R}^2$. Alternativní přístup je odhad pomocí UFOMapy. Základním předpokladem pro tvorbu modelu je znalost externích parametrů kamery při tvorbě kolekce snímků. To znamená, že potřebujeme několik stereo snímků stožáru se známou pozicí v globálním souřadnicovém systému. Obě využité metody pracují i s odchylkou v nepřesnosti určení souřadnic, například použitého gps zařízení.

6.1.1 Metodou Parzen window

Jednodušším, ale výpočetně náročnějším přístupem je použití Parzen window. Parzen winow je neparametrická metoda určená pro odhad hustoty pravděpodobnosti $\hat{f}(x)$

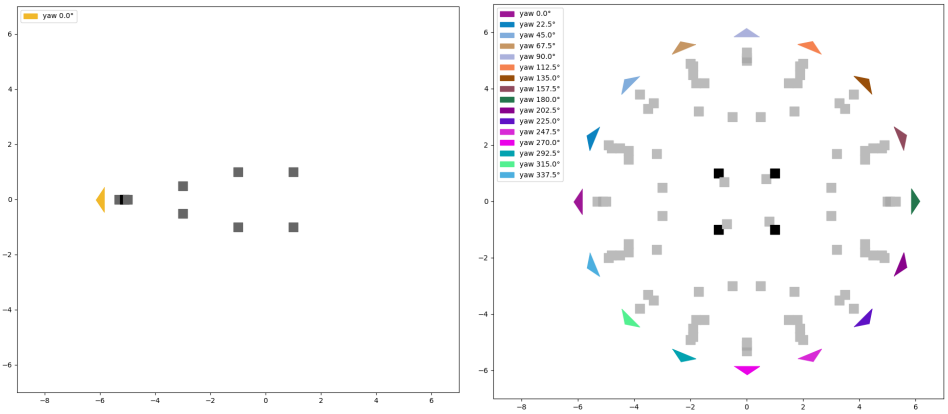
$$\hat{f}_h(x) = \frac{1}{n}\sum_{i=1}^{n} K_h(x - x_i), \tag{6.1}$$

kde K_h je symetrická funkce, jejíž integrál je roven jedné. Parametr h je velikost jádra, $x_i \in \{x_1, ..., x_n\}$ jsou všechny přípustné kombinace geometrických protějšků [11].

V naší práci je jádrovou funkcí jednoduchá hyperkrychle. Konkrétně pro jeden vodorovný řez dříkem se jedná o čtverec. Velikost jádra (délku hrany čtverce) jsme zvolili stejnou jako velikost voxelu ufomapy 5 cm. Pro názornost výsledků jsme vyhodnotili pravděpodobnosti pro všechny body v oblasti všech promítnutých bodů, viz obrázek 6.2. Jednotlivá okna jsou na obrázku reprezentována čtverci s odlišnými odstíny, dle hladiny pravděpodobnosti. Pro přehlednost jsme odstranili okna s nejnižší hladinou pravděpodobnosti. Jinými slovy, tmavší čtverce obsahují více promítnutých bodů, vyšší hladinu pravděpodobnosti, než ostatní. Z obrázku můžeme jednoduše získat správné souřadnice základny dříku vybráním bodů s nejvyšší pravděpodobností. Dle očekávání se jedná o čtyři nejtmavší čtverce blízko středu se souřadnicemi X_{ref} ={(-1,-1),(-1,1),(1,-1),(1,1)}, které přesně odpovídají souřadnicím, na které jsme stojky v Gazebu umístili.

6.1.2 Využitím knihovny UFOMap

Názornější a výpočetně méně náročnou metodou, kterou jsme pro určení správných souřadnic základny dříku a eliminování falešných průmětů stožáru použili, je promítnutí získaných bodů do UFOMapy [9]. Stejně jako v případě metody Parzen window jsme nejprve postup otestovali pouze v jednom vodorovném řezu dříkem. Později jsme, vzhledem k velmi dobrým výsledkům, implementovali promítání bodů do původní 3D scény, viz obrázek 6.3a. I přestože cílem naší práce není představení detailů modulu UFOMap, je důležité zmínit alespoň základní princip. Spojité prostředí rozdělíme do konečného počtu krychlí (voxelů), s předem definovanou délkou hrany. Každý voxel si udržuje pravděpodobnost, že se v něm nachází objekt, kterou zvyšujeme počtem bodů do něj promítnutých. Jinými slovy, do takto definovaného prostoru můžeme promítnout rekonstruované mračno bodů a zvýšit tak příslušným voxelům pravděpodobnost, že se v nich nachází objekt.



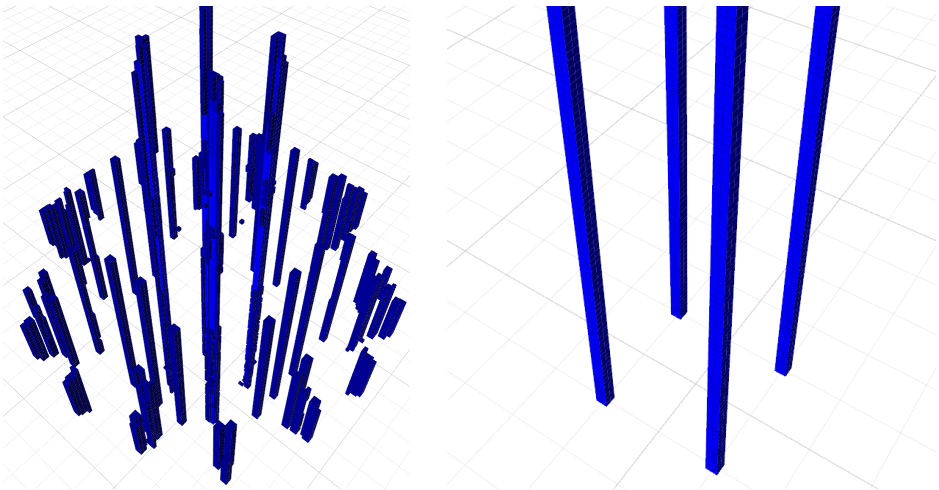
(a) : Výsledek po zpracování 1 snímku stožáru - všechny stojky mají stejnou hladinu pravděpodobnosti.

(b) : Výsledek po zpracování 16 odlišných snímků stožáru.

Obrázek 6.2: Určení nejpravděpodobnějšího umístění základny dříku stožáru metodou Parzen window. Jednotlivá okna jsou zobrazena odstíny černé. Tmavší odstíny představují vyšší hladinu pravděpodobnosti. Pozice kamery při pořizování snímku je znázorněna trojúhelníkem náhodné barvy.

V naší práci jsme konkrétně využili veřejně přístupný modul pro komunikaci s ROSem a mračno bodů posíláme jako zprávu typu /python/pointcloud2. Výslednou 3D mapu lze zobrazit v RVizu, protože UFOMap modul zveřejňuje mapu jako zprávu /ufomap_mapping_server_node/map_depth_1. Po promítnutí všech kombinací potenciálních stojek dříku stožáru z několika odlišných úhlů do UFOMapy, je vhodným prahováním možné zachovat pouze voxely s nejvyšší pravděpodobností. Jinými slovy, voxely příslušící falešným průmětům stojek stožáru mají pravděpodobnost obsazení objektem výrazně nižší, než ty správné.

Velikost voxelu je vhodné volit v závislosti na nepřesnosti určení souřadnic, ze kterých byl snímek pořízen a v závislosti na rozměrech detekovaného objektu. Vzhledem k tomu, že pracujeme se simulací, kde je nepřesnost minimální, zvolili jsme voxel size 5 cm. Spolehlivé výsledky jsme, při adekvátním prodloužení hrany voxelu na 10–20 cm, dostali i pro nepřesně určené souřadnice kamery - náhodně jsme modifikovali souřadnice s odchylkou v rozmezí 5–10 cm. V případě výraznějšího zašumění bychom měli volit ještě větší velikost voxelu, což ale vzhledem k rozteči a velikosti stojek již postrádá smysl.



(a) : Potenciální průměty stojek ze všech úhlů na kružnici kolem stožáru.

(b) : Odfiltrované falešné kombinace vhodným prahováním - zobrazení voxelů s nejvyšší pravděpodobností obsazenosti - výsledný model stožáru.

Obrázek 6.3: Promítnutí mračna bodů do ufomapy. Stožár rekonstruován z 16 snímků při známé pozici kamery.

Kapitola 7

Zpracování snímku s neznámou polohou pořízení

Existuje mnoho robustních metod hledajících správnou rotaci-translaci vytvořeného pointcloudu (nebo polygonové sítě), s cílem minimalizovat chybu překrývajících se ploch s dosavadním modelem. Jednou z nejjednodušších metod je ICP (Iterative Closest Point).

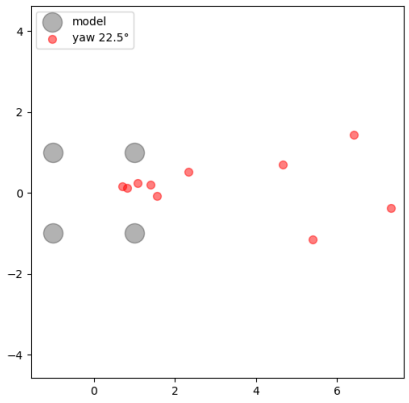
Většina implementací základního ICP nejdříve přesune těžiště získaných bodů do těžiště modelu a poté hledá nejvhodnější rotaci. Iterative napovídá, že tento proces opakujeme a snažíme se snižovat kritérium chyby. Pro naší práci je ale nevhodný především inicializační krok translace bodů. Model má v našem případě v půdorysu 4 body, promítnuté kombinace ale mohou mít několikanásobně více bodů. Především je problém, že kombinace blíže ohnisku kamery jsou blíže u sebe, a tudíž přesun těžiště na těžiště modelu nevede ke správným výsledkům.

7.1 Navržený přístup

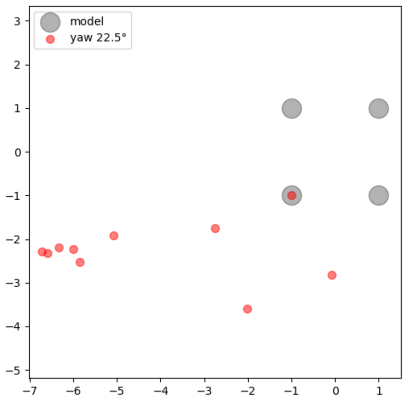
Vytvořený referenční model v druhé části navrženého postupu využíváme k nalezení původních souřadnic kamery. Postup lze zjednodušit do několika kroků následovně:

1. Snímání stožáru stereokamerou v simulačním prostředí Gazebo
2. Předzpracování a úprava vytvořené kolekce snímků
3. Nalezení středů každé ze stojek stožáru
4. Tvorba všech potenciálních průmětů do podprostoru **W**
5. Hledání správné rotace/translace minimalizující e_{nn}

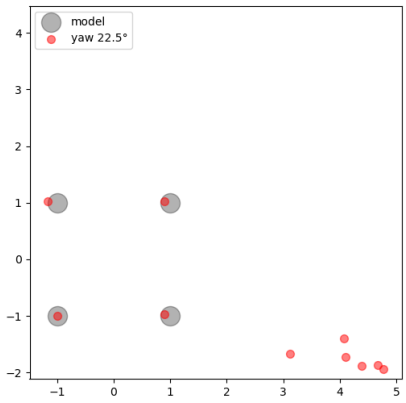
Při návrhu našeho řešení jsme se ICP algoritmem inspirovali a navrhli jsme následující modifikace. Nejprve definujme chybu e_{nn}, kterou budeme minimalizovat. Jak je zmíněno v minulé kapitole, součet vzdáleností všech bodů vůči modelu je nevhodný.



(a) : Body promítnuté pohledem kamery z (0,0) ve směru osy x.



(b) : Posunutí bodů na jeden z rohů modelu.



(c) : Nalezení rotace využitím SVD rozkladu.

Obrázek 7.1: Hledání rotace a posunutí navrženým přístupem. Skutečný snímek pořízen z pohybu po kružnici kolem stožáru pod úhlem 22.5°.

Proto v práci využíváme algoritmu nejbližšího souseda, cizojazyčně k-NN (k-Nearest Neighbors), a měříme odchylku od bodu modelu pouze k nejbližšímu z promítnutých bodů.

Podobně jako metoda Parzen window je k-NN neparametrickou metodou odhadu hustoty pravděpodobnosti. Na rozdíl od Parzen window neklade omezení na velikost jádra ale na požadovaný počet bodů obsažených v každém okně. V naší práci využíváme 1-NN a pro každý bod referenčního modelu hledáme nejbližší potenciální geometrický protějšek. V práci jsme využili knihovny sklearn implementující algoritmus nejbližšího souseda pomocí kd stromů [12]. Kritérium chyby pro vybraný bod modelu $\boldsymbol{m_i}$ lze poté definovat následovně

$$e_{nn}(\boldsymbol{m_i}) = \min_{\boldsymbol{x_i} \in \mathbf{X}} ||\boldsymbol{m_i} - \boldsymbol{x_i}|| \tag{7.1}$$

Ve výsledcích uvádíme i průměrnou odchylku od modelu, kterou lze definovat v závislosti na počtu stojek stožáru následovně

$$e_{avg} = \frac{1}{|\mathbf{M}|} \sum_{\boldsymbol{m_i} \in \mathbf{M}} e_{nn}(\boldsymbol{m_i}) \tag{7.2}$$

Navržený postup nalezení polohy kamery ve světových souřadnicích lze popsat následovně, v každé iteraci vyberme jeden bod $\boldsymbol{x_i}$ z množiny všech kombinací $\mathbf{X} = \{\boldsymbol{x_1}, ..., \boldsymbol{x_n}\}$, vyberme bod $\boldsymbol{m}_i$ z množiny bodů modelu $\mathbf{M} = \{\boldsymbol{m_1}, ..., \boldsymbol{m}_k\}$ (v případě stožáru je k=4) a posuňme všechny body množiny $\mathbf{X}$ o translační vektor $\boldsymbol{t} = \boldsymbol{m}_i - \boldsymbol{x}_i$, viz obrázek 7.1b).

V druhém kroku hledáme vhodnou rotaci kolem bodu $\boldsymbol{m_i}$. Abychom nemuseli vhodnou rotaci hledat iterativně, využijeme SVD rozkladu. Nejdříve ale provedeme posun bodu $\boldsymbol{m_i}$ do počátku souřadnic, translační vektor nazvěme například $\boldsymbol{t}_0 = -\boldsymbol{m}_i$. SVD rozklad není možný nad maticemi $\mathbf{A} \in \mathbb{R}^{2\times 4}$ a $\mathbf{B} \in \mathbb{R}^{2\times k}$, v naší práci uvádíme ve výsledcích výhody a nevýhody hledání rotace na každé podmnožině o velikosti 4 a 2, viz tabulka 7.2. Výstupní matice SVD rozkladu značíme $\mathbf{USV^T}$.

Rotační matice je $\mathbf{R} = \mathbf{SIV^T}$, kde $\mathbf{I}$ je jednotková matice. Všechny body $\mathbf{X}$ rotujeme maticí $\mathbf{R}$ následovně $\mathbf{X'} = \mathbf{X} \cdot \mathbf{R}$. Poté už stačí jen všechny body posunout zpět o $-\boldsymbol{t}_0$ a dostaneme výsledek na obrázku 7.1c). Tento postup provedeme pro všechna $\boldsymbol{x_i}$ a $\boldsymbol{m_i}$ z příslušných množin.

Výsledky doplňkové části, kdy se snažíme najít vhodné vnější parametry kamery bez znalosti souřadnic dronu při pořízení snímku, uvádíme pro vizuální kameru při rozlišení 1920×1080 (2 MP) a FOV 90° a termokameru při rozlišení 160×120 a FOV 57°. Nad rámec jsme ještě otestovali navržený

	zašuměná matice vnějších parametrů		neznámá poloha stereo kamery	
	e_{avg}	e_{max}	e_{avg}	e_{max}
1920×1080 vizual	1,3 cm	12 cm	1,24 cm	10 cm
1280×720 vizual	3,6 cm	12 cm	3,7 cm	10,8 cm
160×120 termo	6,7 cm	43,9 cm	8,2 cm	33,45 cm

Tabulka 7.1: Výsledky navrženého přístupu k určení souřadnic základny dříku stožáru.

	e_{avg4}	t_{avg4}	e_{avg2}	t_{avg4}
1920×1080 vizual	0,92 cm	18,43 s	0,94 cm	1,01 s
1280×720 vizual	3,8 cm	17,82 s	4,06 cm	1 s
160×120 termo	7,1 cm	17,99 s	8,67 cm	0,98 s

Tabulka 7.2: Zrychlení části algoritmu hledáním vhodné rotace translace pouze dvou bodů, v porovnání s přesnějším ale pomalejším párováním všech bodů modelu.

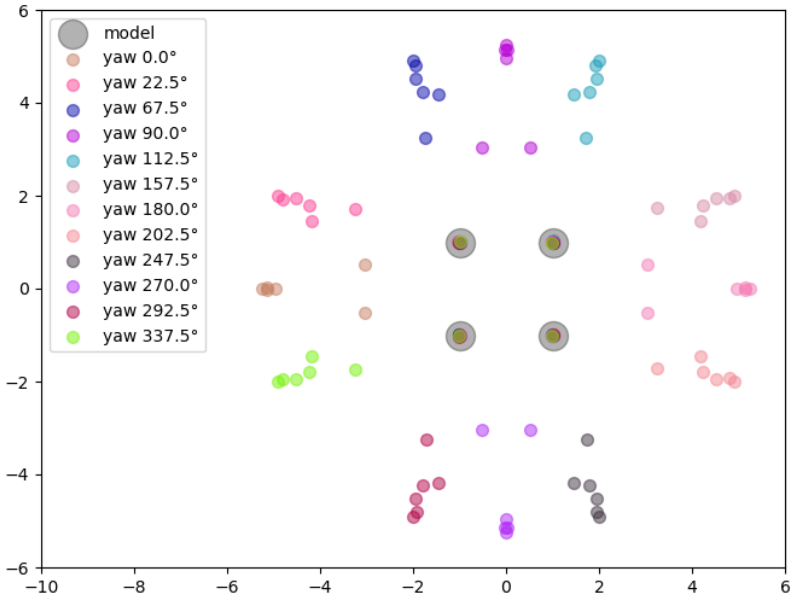
přístup na vizuální kameře s menším rozlišením 1280×720 (1 MP). Kromě průměrné chyby, kterou jsme definovali v 7.2, pro úplnost ještě uvádíme největší vzniklou odchylku od bodu modelu k nejbližšímu sousedovi e_{max}. V tabulce 7.1 jsou uvedeny výsledky pro všechny úhly pohledu, včetně snímků, kdy se stojky překrývají. Do obrázků 7.2 a 7.3 jsme ale body příslušící pohledům s překryvy nezaznamenali, především z důvodu přehlednosti, protože pro kritické oblasti promítáme mnoho kombinací a obrázek by potom byl nepřehledný. Pro představu by se jednalo o čtyřnásobek promítnutých bodů, viz obrázek 7.5c, protože překrývající se části vidíme vždy z úhlu pohledu $k \cdot 45°$, kde $k = 1, 2, 3, 4$.

Z hodnot lze vyčíst, že čím větší má použitá kamera rozlišení, tím menší průměrné odchylky jsme schopni dosáhnout. I přestože má termokamera rozlišení v řádku pouze 160 pixelů, průměrná chyba činí pouze 7 cm. Nejlepší průměrné odchylky kolem 1 cm jsme dosáhli na kolekci snímků z vizuální kamery při 2 MP rozlišení.

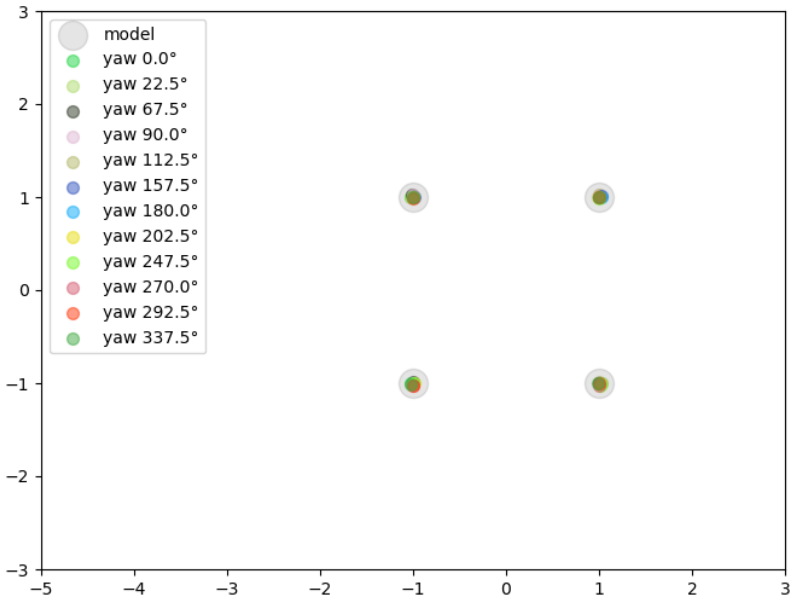
Největší odchylka od modelu vznikla, dle očekávání, na kolekci snímků z termokamery. Je především způsobena tím, že pomocí 160 px musíme do obrázku promítnout vzdálené objekty, tzn. odchylka jednoho pixelu může v globálních souřadnicích znamenat i skok několika desítek centimetrů.

Experimenty ukázaly, že informace o umístění kamery při focení snímku nemá na výsledek téměř žádný vliv, tudíž můžeme na vstup umístit libovolné snímky s neznámou polohou pořízení.

Jak je již zmíněno v podkapitole 7.1 hledání rotace je možné pomocí všech bodů modelu, tzn. v případě stožáru se 4 stojkami musíme využít všechny podmnožiny velikosti 4 z naměřených bodů. Tento postup je pomalý, ale lze ho zrychlit za cenu mírného zhoršení přesnosti využitím podmnožin menší velikosti, v našem případě jsme zvolili velikost 2. Pro zjištění rotace a posunutí se tedy netestují všechny kombinace 4 bodů modelu a 4 bodů z aktuálního měření, ale jen kombinace 2 bodů modelu a 2 bodů z aktuálního měření. V tabulce 7.2 uvádíme zrychlení této části algoritmu a zhoršení průměrné odchylky.

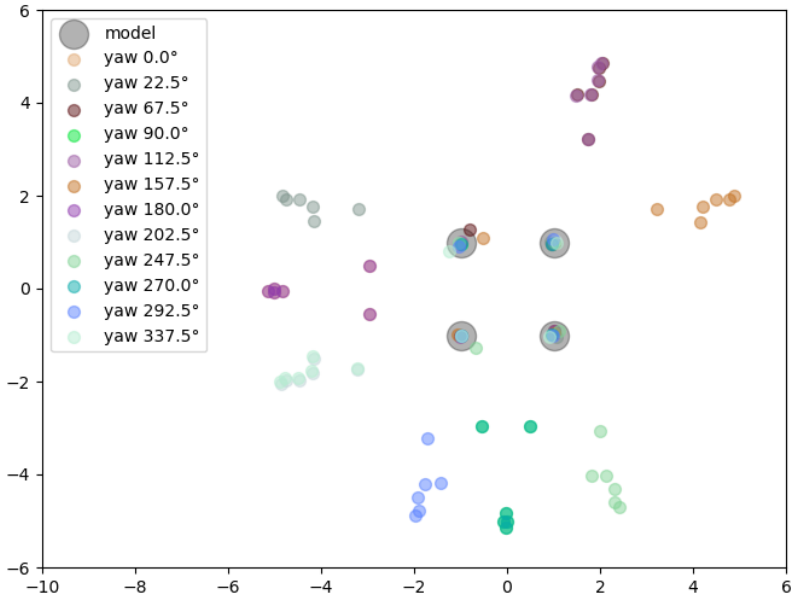


(a) : Vizuální kamera - Výsledek navrženého algoritmu.

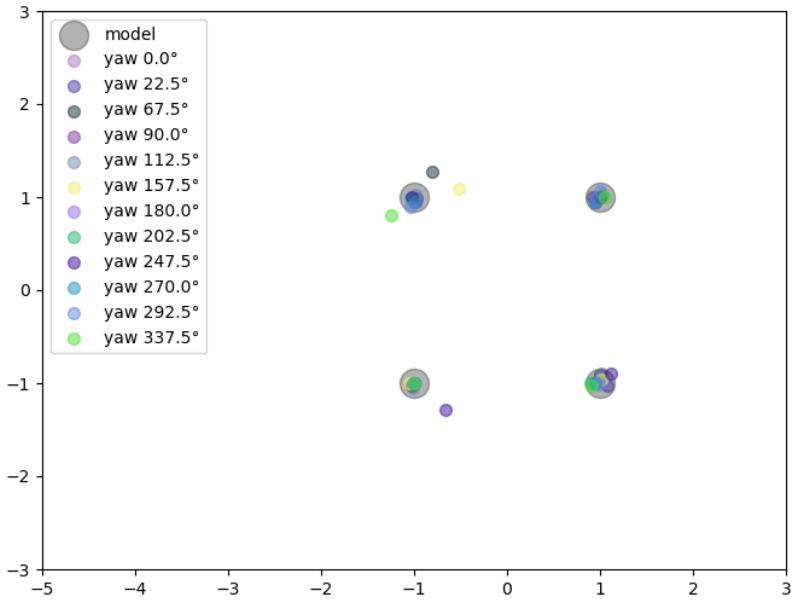


(b) : Vizuální kamera - Metodou 1NN odebrány falešné průměty.

Obrázek 7.2: Výsledky navrženého algoritmu. Rekonstruováno z 12 snímků **stereo vizuální kamery** - pohyb po kružnici kolem stožáru. Barvy jsou volené náhodně a slouží pouze k odlišení bodů příslušících rozdílným snímkům.



(a) : Termokamera - Výsledek navrženého algoritmu.



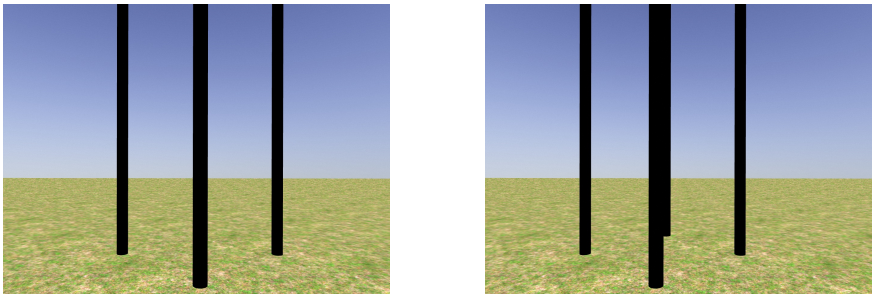
(b) : Termokamera - Metodou 1NN odebrány falešné průměty.

Obrázek 7.3: Výsledky navrženého algoritmu. Rekonstruováno z 12 snímků **stereo termokamery** - pohyb po kružnici kolem stožáru.

7.2 Překrývání stojek stožáru

Počet potenciálních průmětů stojek stožáru jsme v kapitole 6 výrazně snížili tím, že hledáme geometrické protějšky pouze středů stojek. Hledání stojek pouze jako středu nejtmavší oblasti přináší komplikaci v případě snímání stožáru z pohledu, kdy se stojky překrývají viz obrázek 7.4. Odchylka od reálného středu stojky má výrazný vliv na následný průmět bodu do původního souřadnicového systému. Jak je znázorněno na obrázku 7.5b, ze dvou překrývajících se stojek nalezneme pouze jednu, navíc s velmi významnou chybou. Příklad je uvedený pro kolekci tvořenou snímky stožáru ze vzdálenosti 6 m, odchylka od modelu je v tomto případě v desítkách centimetrů, konkrétně $e_{max} = 35,5$ cm. Klíčovou informací pro možnost eliminace této chyby je znalost očekávaného počtu stojek z referenčního modelu. Je nutné podotknout, že tento problém nemá vliv na počet nalezených referenčních stojek při tvorbě modelu.

Implementaci jsme ještě zjednodušili předpokladem, že se v jednom snímku překrývají vždy maximálně dvě stojky. Pokud bychom tento přístup chtěli aplikovat na jiný model stožáru, není problém hledat n překrývajících se oblastí - porovnáme počet nalezených hledáním středů s počtem referenčních. Na obrázku 7.5c jsou znázorněny všechny potenciální kombinace geometrických protějšků, oblast překrytí dvou stojek (levá horní, pravá dolní) jsme nereprezentovali pouhým středem, ale interpolovali jsme ji předem definovaným počtem bodů. Počet bodů jsme pro účely testování přizpůsobily původní šířce kritické oblasti v pixelech. V případě zpracování kolekce snímků z vizuální kamery jsme zvolili 50 bodů, v případě zpracování kolekce snímků z termokamery jsme zvolili 10 bodů.

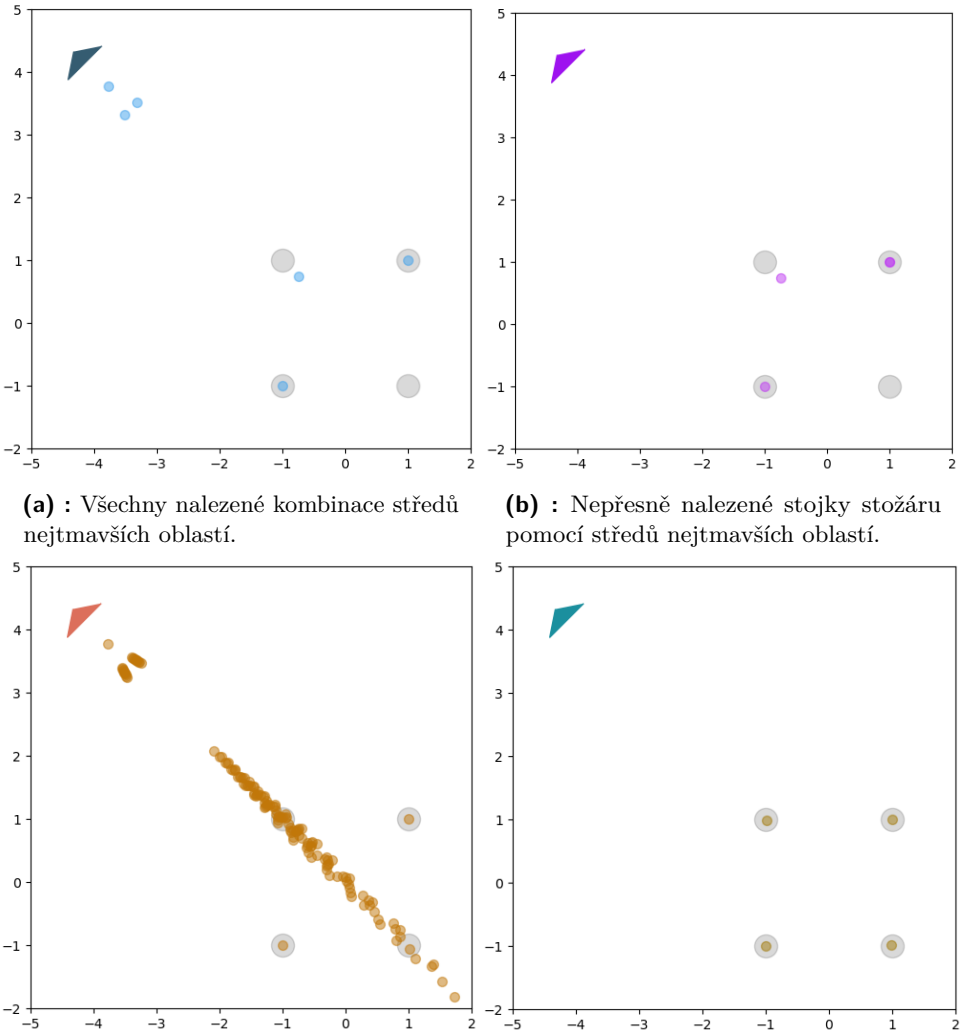


(a) : Levý snímek překrývajících se stojek. **(b) :** Pravý snímek překrývajících se stojek.

Obrázek 7.4: Pohled na stožár stereo párem Mobius kamer.

Již na první pohled je i z obrázků 7.5b) a 7.5d) zřejmé přesnější umístění levého horního bodu. Konkrétně jsme tímto postupem byly schopni maximální vzniklou odchylku od referenčního modelu snížit až o jeden řád. Pro uvedený příklad byla původní $e_{max} = 35{,}5$ cm, navrženou modifikací algoritmu jsme

byly schopni snížit e_{max} na 9,3 cm. Je nutné podotknout, že za cenu výrazného snížení kritéria maximální odchylky, platíme především zvýšením výpočetní náročnosti. Je zřejmé, že na průměrnou odchylku e_{avg} tato modifikace nemá příliš výrazný vliv, protože v ostatních snímcích se stojky nepřekrývají.



(a) : Všechny nalezené kombinace středů nejtmavších oblastí.

(b) : Nepřesně nalezené stojky stožáru pomocí středů nejtmavších oblastí.

(c) : Všechny nalezené kombinace po proložení bodů v oblasti překrytí.

(d) : Nalezené správné pozice stojek stožáru vysokého napětí.

Obrázek 7.5: Navržené řešení problému překrývajících se stojek. Zvýšením počtu potenciálních komplementárních dvojic jsme dosáhli výrazně přesnějších výsledků, za cenu vyšší výpočetní náročnosti.

Kapitola 8

Závěr

Kromě úvodní teorie rekonstrukce scény ze snímku stereo kamery jsme se v naší práci věnovali především návrhu algoritmu, nebo spíše postupu řešení problému rekonstrukce stožáru vysokého napětí. Nezanedbatelný díl praktické části jsme věnovali tvorbě kolekcí snímků. Snímky jsme pořizovali vlastní konfigurací stereo páru Mobius kamer v simulačním prostředí Gazebo.

V první části navrženého postupu připravíme každý pořízený snímek stožáru k následnému zpracování. Konkrétně měníme barevné spektrum a využíváme metod k-means pro snížení počtu obsažených odstínů šedé. V takto před připraveném snímku následně nalezneme stojky stožáru a vytvoříme všechny potenciální průměty stojek ve světovém souřadnicovém systému. Falešné i správné umístění stojek v původním prostoru v podobě mračna bodů promítneme do pravděpodobnostní mapy (UFOMapy). Při zpracování několika snímků zachycujících stožár z vícero předem známých pozic jsme schopni vybráním voxelů s nejvyšší hladinou pravděpodobnosti určit referenční souřadnice stojek stožáru. V našem případě jsme díky simulačnímu prostředí měli velmi přesné pozice dronu, ze kterých byl snímek pořizován, a proto jsme byli schopni dosáhnout velmi dobrých výsledků i pro menší 5cm voxely, viz podkapitola 6.1.2. Uvádíme také alternativní, výpočetně náročnější, ale implementačně přímočarý přístup k tvorbě referenčního modelu metodou Parzen window v jednom vodorovném řezu stožárem, viz podkapitola 6.1.1. Ověřili jsme, že správnou volbou velikosti voxelu, případně velikosti okna u Parzen window, jsme schopni dostat spolehlivé výsledky i pro nepřesné souřadnice pořízeného snímku.

Dalším cílem naší práce byla rekonstrukce na základě neznámé/nepřesné

pozice dronu. Navrhli jsme funkční přístup pro nalezení původní pozice dronu bez apriorní znalosti pozice pořízení snímku. Inspirovali jsme klasickým ICP algoritmem a nejpravděpodobnější vnější parametry kamery při pořizování snímku hledáme postupem popsaným v podkapitole 7.1. Posuneme jeden z promítnutých bodů na jeden z bodů modelu a následně pomocí SVD rozkladu nalezneme nejvhodnější rotaci. Nalezený výsledek ohodnotíme součtem vzdáleností od každého bodu referenčního modelu k nejbližšímu sousedovi. Nejnižší průměrné odchylky od referenčního modelu $e_{nn} \doteq 1$ cm jsme dosáhli na kolekci snímků ze stereo vizuální kamery s rozlišením 1920×1080, výsledky v kapitole 7.1. Vzhledem k časové náročnosti se omezujeme na minimální počet potenciálních průmětů. Tím dosáhneme výrazného zrychlení, ale vznikne nezanedbatelný problém v případě překrývajících se stojek, řešení nalezením a následnou interpolací určitého počtu bodů do kritické oblasti jsme představili v podkapitole 7.2.

V budoucích vylepšeních bychom chtěli navrhnout algoritmy, které by vhodněji využily nepřesné informace o poloze dronu. S cílem zpřesnit lokalizaci dronu v případech, kdy jsou snímky z vícero úhlů pohledu nerozlišitelné, například protože je snímaný objekt symetrický. Dále bychom chtěli navrhnout zjednodušení procesu tvorby referenčního modelu stožáru.

Literatura

[1] KHAIRUDDIN, A. R., M. S. TALIB a H. HARON. Review on simultaneous localization and mapping (SLAM). In: *2015 IEEE International Conference on Control System, Computing and Engineering (ICCSCE)*. 2015, s. 85-90. Dostupné z: doi:10.1109/ICCSCE.2015.7482163

[2] LAZAROS, Nalpantidis, Georgios Christou SIRAKOULIS a Antonios GASTERATOS. Review of Stereo Vision Algorithms: From Software to Hardware: From Software to Hardware. *International Journal of Optomechatronics*. Taylor & Francis, 2008/11/26, 2(4), 435-462. ISSN 1559-9612. Dostupné z: doi:10.1080/15599610802438680

[3] Wikipedia Contributors. Rotation matrix. *Wikipedia - the free encyclopedia*. 2022. Dostupné z: https://en.wikipedia.org/wiki/Rotation_matrix

[4] Wikipedia Contributors. Homogeneous coordinates. *Wikipedia - the free encyclopedia*. 2022. Dostupné z: https://en.wikipedia.org/wiki/Homogeneous_coordinates

[5] HARTLEY, Richard a Andrew ZISSERMAN. Multiple view geometry in computer vision. 2nd ed. Cambridge: Cambridge University Press, 2003. ISBN 978-0-521-54051-3.

[6] Depth Map from Stereo Images. *OpenCV: Open Source Computer Vision* [online]. 2022. Dostupné z: https://docs.opencv.org/4.x/dd/d53/tutorial_py_depthmap.html

[7] Simple Stereo: video lecture. In: *First Principles of Computer Vision* [online]. 2021. Dostupné z: https://fpcv.cs.columbia.edu/

[8] Introduction to SIFT (Scale-Invariant Feature Transform). *OpenCV: Open Source Computer Vision* [online]. 2022. Dostupné z:`https://docs.opencv.org/4.x/da/df5/tutorial_py_sift_intro.html`

[9] DUBERG, Daniel a Patric JENSFELT. UFOMap: An Efficient Probabilistic 3D Mapping Framework That Embraces the Unknown. *IEEE Robotics and Automation Letters*. 2020, 5(4), 6411-6418. ISSN 2377-3766. Dostupné z: doi:10.1109/LRA.2020.3013861

[10] KOSTKA, Tomáš. *Mechanika venkovních vedení*. 2020, 3. Dostupné z: `https://www.ssph.cz/vyuka/wp-content/uploads/2020/03/ee-mech_v_v.pdf`

[11] KWAK, N. a CHONG-HO CHOI. Input feature selection by mutual information based on Parzen window. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2002, 24(12), 1667-1671. ISSN 0162-8828. Dostupné z: `doi:10.1109/TPAMI.2002.1114861`

[12] Scikit-learn developers. Nearest Neighbors. *Scikit-learn* [online]. 2022. Dostupné z: `https://scikit-learn.org/stable/modules/neighbors.html`

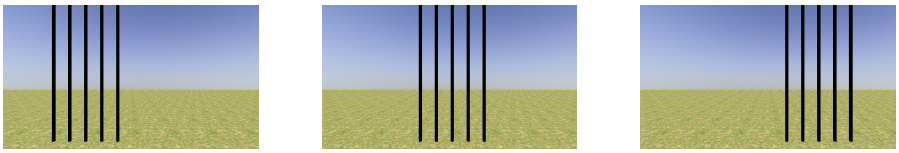
[13] BACA, Tomas, Matej PETRLIK, Matous VRBA, Vojtech SPURNY, Robert PENICKA, Daniel HERT a Martin SASKA. The MRS UAV System: Pushing the Frontiers of Reproducible Research, Real-world Deployment, and Education with Autonomous Unmanned Aerial Vehicles: Pushing the Frontiers of Reproducible Research, Real-world Deployment, and Education with Autonomous Unmanned Aerial Vehicles. *Journal of Intelligent & Robotic Systems*. Springer Science and Business Media, 2021, 102(1). Dostupné z: `doi:10.1007/s10846-021-01383-5`

[14] QUIGLEY, Morgan, Ken CONLEY, Brian GERKEY, Josh FAUST, Tully FOOTE, Jeremy LEIBS, Rob WHEELER a Andrew NG. *ROS: an open-source Robot Operating System: an open-source Robot Operating System*. 3. 2009/01/01.

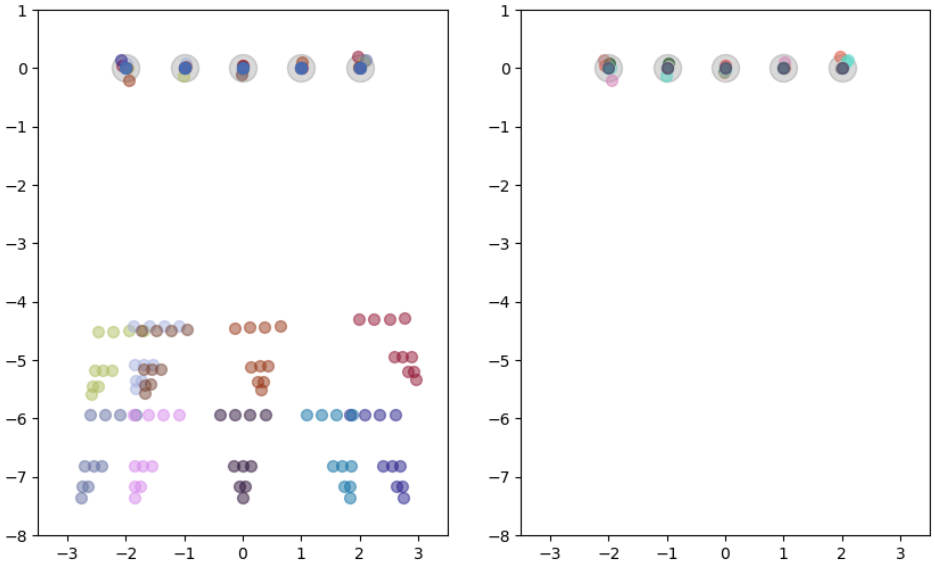
Příloha A

Ověření navrženého postupu na kolekci snímků zjednodušeného zábradlí

Vzhledem k symetričnosti čtvercové základny dříku stožáru není vyloučeno, že při hledání vnějších parametrů kamery za nejpravděpodobnější původní souřadnice kamery vybereme nesprávné, protože pohledy pod úhly $k \cdot 90°$, $k = 0,1,2,3$ jsou neodlišitelné. Nejen z tohoto důvodu jsme proto vytvořili i zjednodušený model zábradlí. Ukázka z kolekce snímků je na obrázku A.1. Snímky byly pořízeny ze dvou odlišných vzdáleností 6 a 8 m. Dosáhli jsme průměrné odchylky $e_{avg} = 3,3$ cm, maximální odchylka stojky od jejích reálných souřadnic v simulačním prostředí činila $e_{max} = 17$ cm.



Obrázek A.1: Ukázka několika snímků z kolekce zjednodušeného zábradlí ze vzdálenosti 8 m vizuální kamerou s rozlišením 1 MP.



(a) : Všechny potenciální původní souřadnice stojek zábradlí.

(b) : Metodou 1-NN odstraněné falešné průměty.

Obrázek A.2: Výsledky promítnutí odhadovaných původních souřadnic stojek zábradlí. Kolekce obsahuje 5 snímků zábradlí ze vzdálenosti 6 m a 5 snímků ze vzdálenosti 8 m.