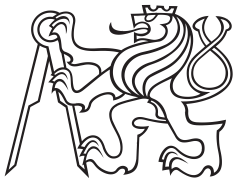


Bakalářská práce



České
vysoké
učení technické
v Praze

F3

Fakulta elektrotechnická
Katedra kybernetiky

Stereo termokamera

Ivan Žižka

Vedoucí: RNDr. Petr Štěpán, Ph.D.

Obor: Základy umělé inteligence a počítačových věd

Studijní program: Otevřená informatika

Květen 2022

I. OSOBNÍ A STUDIJNÍ ÚDAJE

Příjmení: **Žižka** Jméno: **Ivan** Osobní číslo: **492068**
Fakulta/ústav: **Fakulta elektrotechnická**
Zadávací katedra/ústav: **Katedra kybernetiky**
Studijní program: **Otevřená informatika**
Specializace: **Základy umělé inteligence a počítačových věd**

II. ÚDAJE K BAKALÁŘSKÉ PRÁCI

Název bakalářské práce:

Stereo termokamera

Název bakalářské práce anglicky:

Stereo Thermal-Camera

Pokyny pro vypracování:

- 1) Seznamte se s kamerou Flir Lepton 3.5.
- 2) Napište program pro snímání dat z této kamery a kameru zkalibrujte. Zdokumentujte kvalitu provedené kalibrace kamery.
- 3) Vytvořte stereo kameru ze dvou senzorů Flir Lepton 3.5 a kameru zkalibrujte.
- 4) Vytvořte program, který bude detekovat člověka, nebo jiný tepleně výrazný objekt v kameře a zjistěte přesnost detekce na základě vzdálenosti od kamery.

Seznam doporučené literatury:

- [1] FLIR LEPTON® Engineering Datasheet [online]. s. 34-39 [cit. 2021-11-25]. Dostupné z: <https://www.flir.com/globalassets/importedassets/document/flir-lepton-engineering-datasheet.pdf>
- [2] KARLSEN, Robert E., Douglas W. GAGE, Charles M. SHOEMAKER, Grant R. GERHART, Josh HARGUESS a Shawn STRANGE. Infrared stereo calibration for unmanned ground vehicle navigation [online]. In: . 2014-6-3, 90840S- [cit. 2022-01-06]. Dostupné z: doi:10.1117/12.2053537
- [3] Depth Map from Stereo Images. OpenCV [online]. [cit. 2022-01-07]. Dostupné z: https://docs.opencv.org/4.x/d/d53/tutorial_py_depthmap.html

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: **17.01.2022**

Termín odevzdání bakalářské práce: **20.05.2022**

Platnost zadání bakalářské práce: **30.09.2023**

RNDr. Petr Štěpán, Ph.D.
podpis vedoucí(ho) práce

prof. Ing. Tomáš Svoboda, Ph.D.
podpis vedoucí(ho) ústavu/katedry

prof. Mgr. Petr Páta, Ph.D.
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

Poděkování

Děkuji panu doktorovi Petrovi Štěpánovi za vedení této práce, jeho cenné rady a podněty.

Prohlášení

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

V Praze, dne 20.5.2022

Ivan Žížka

Abstrakt

Tato práce se zabývá geometrickou kalibrací stereo termokamery a následného odhadování vzdálenosti teplých objektů od soustavy. Stereo termokamera se skládá ze dvou kamer FLIR LEPTON 3.5 s rozlišením 160x120 pixelů. Stereo vidění u kamer pracujících ve viditelném spektru je velmi dobře prozkoumaný jev, ale využití termokamer pro tento typ vidění nikoliv. V práci je představena metoda, která umožňuje geometrickou kalibraci termokamer a následnou kalibraci celé stereo soustavy. Tato metoda je založena na běžně používaném kalibračním vzoru, který je zahříván pomocí slunečního záření.

Následně je využito kalibrovaných kamer k odhadu vzdálenosti teplých objektů od základny kamery. Tyto objekty jsou detekovány a lokalizovány za pomoci teplotního prahování. Pro odhadování vzdálenosti objektu jsou využity různé metody, které jsou mezi sebou porovnány.

Klíčová slova: termokamera, stereo termokamera, LEPTON 3.5, geometrická kalibrace, detekce

Vedoucí: RNDr. Petr Štěpán, Ph.D.

Abstract

This thesis deals with the geometric calibration of a stereo thermal camera and the subsequent estimation of the distance of warm objects from the system. The stereo thermal camera consists of two FLIR LEPTON 3.5 cameras with 160x120 pixel resolution. Stereo vision u cameras operating in the visible spectrum is a well-studied phenomenon, but the use of thermal cameras for this type of vision is not. In this paper, a method is presented that allows geometric calibration of thermal cameras and subsequent calibration of the entire stereo array. This method is based on a commonly used calibration pattern that is heated by solar radiation.

The calibrated cameras are then used to estimate the distance of the warm objects from the camera base. These objects are detected and located using temperature thresholding. Different methods are used to estimate the object distance and are compared with each other.

Keywords: thermographic camera, thermal camera, stereo thermal camera, LEPTON 3.5, geometric calibration, detection

Title translation:
Stereo Thermal-Camera

Obsah

1 Úvod	1	3.4 Kvalita kalibrace	16
2 Kamera	3	4 Stereo kamera	19
2.1 Dírkový model kamery	3	4.1 Stereo vidění	19
2.1.1 Pozice a rotace kamery	5	4.2 Kalibrace stereokamery	20
2.1.2 Projekční matice	6	4.3 Epipolární geometrie	21
2.2 Distorze	6	4.4 Rektifikace obrazu	22
2.3 Geometrická kalibrace kamer	7	4.5 Stereo matching algoritmy	24
2.4 FLIR LEPTON 3.5	8	4.5.1 Plošné metody	24
2.4.1 Čtení dat z kamery	9	4.5.2 Porovnávání rysů	24
3 Geometrická kalibrace termokamery	13	4.6 Výpočet vzdálenosti	25
3.1 Geometrická kalibrace RGB kamer	13	4.7 Přesnost stereo soustavy	26
3.2 Kalibrace termokamer	13	5 Detekce teplých objektů a odhad vzdálenosti	29
3.2.1 Aktivně zahřívání vzory	14	5.1 Detekce objektů	29
3.2.2 Pasivně zahřívání vzory	14	5.2 Detektor teplých objektů	30
3.3 Úprava a vylepšení kalibračního vzoru	15	5.2.1 Načtení dat z kamer a nastavení parametrů	30
		5.2.2 Prahování teplot	31
		5.2.3 Čištění obrazu	31

5.2.4 Ohraničení oblastí konturami a filtrování	32
5.2.5 Ohraničení kontur bounding boxem	32
5.3 Vlastnosti detektoru	33
5.4 Odhad vzdálenosti teplých objektů	35
5.4.1 Odhad vzdálenosti při použití algoritmu pro párování rysů	35
5.4.2 Odhad vzdálenosti při použití plošného stereo matching algoritmu	36
5.4.3 Odhad vzdálenosti při použití informací o konturách	36
5.5 Zhodnocení jednotlivých algoritmů pro odhad vzdálenosti	38
6 Závěr	41
Literatura	43

Obrázky

2.1 Projekce vzoru do obrazu (2D) [4]	3	4.5 Porovnávání rysů - stereo termokamera 2x FLIR Lepton 3.5	25
2.2 Nákres dírkové kamery s virtuální obrazovou rovinou	4	4.6 Odhad vzdálenosti pomocí triangulace [9]	26
2.3 Radiální distorze, převzato z [12]	7	4.7 Odhad vzdáleností teplého objektu	26
2.4 Tangenciální distorze, převzato z [13]	7	4.8 Ideální průběh disparity v závislosti na vzdálenosti	27
2.5 Termokamera FLIR lepton 3.5 na platformě PureThermal 2	9	4.9 Odhad vzdálenosti při chybě 1 px při určování stereo páru	28
3.1 Šachovnicové kalibrační vzory	14	5.1 Obrázky filtrace masky, převzato z [17]	33
3.2 Odraz teplého objektu na kalibrační desce, ohraničen bílým obdélníkem	15	5.2 Postup detekce dle jednotlivých kroků	34
3.3 Obrázek kalibrační desky po zpracování algoritmem na detekci rohů	16	5.3 Porovnání algoritmů pro určování vzdálenosti pohybujícího se člověka	37
4.1 Stereo termokamera	20	5.4 Vizualizace jednotlivých metod použitých pro odhad vzdáleností (v řádku vždy jedna metoda)	39
4.2 Kalibrační deska ve stereo soustavě - výřez (vlevo levá kamera, vpravo pravá)	20		
4.3 Epipolární geometrie, převzato z [13]	21		
4.4 Proces rektifikace	23		

Tabulky

3.1 Tabulka reprojekčních chyb . . .	17
5.1 Parametry BM algoritmu	36
5.2 Porovnání rychlosti jednotlivých metod	38

Kapitola 1

Úvod

S rozvojem bezpilotních dronů, autonomních aut nebo bezpečnostních systémů rostou nároky na získávání informací o okolním prostředí. Jedním z oborů, který se zabývá primárně získáváním informací o vnějším světě, je počítačové vidění. Tento obor pracuje s velkou množinou senzorů, které mají různé vlastnosti. Metody počítačového vidění jsou velice závislé na vstupních datech, ať už se jedná o samotný obraz, parametry kamery a prostředí nebo vlastnosti snímaných objektů. V této bakalářské práci je hlavním vstupním senzorem dvojice termokamer od společnosti FLIR. Termokamery v obecné rovině mají podobné vlastnosti jako kamery pracující ve viditelném spektru, ale s tím rozdílem, že obraz je generován na základě teploty prostředí. Z této vlastnosti vyplývá množství výhod, ale i nevýhod. Hlavní výhodou je, že se tyto kamery dají použít za špatné viditelnosti nebo v noci. Největší nevýhodou je množství šumu, ztráta textur a kontrastu pokud mají objekty podobnou teplotu.

Stereo soustavy složené ze dvou a více RGB kamer jsou velmi dobře známým a prozkoumaným senzorem v počítačovém vidění. Používají se pro odhad vzdálenosti objektů nebo na detekci objektů a to včetně lidí či překážek. Velkou výhodou je, že zde existuje celá řada algoritmů, které umí tyto úlohy řešit. Použití termokamer do stereo soustav není běžné, hlavním důvodem byla v minulosti jejich vysoká cena, váha a citlivost na vnější podmínky. S příchodem poměrně levných, lehkých a nechlazených termokamer se tento trend začíná obracet.

Typickým použitím termokamer v dnešní době je kombinace jedné či více termokamer se senzorem Light Detection And Ranging (LIDAR) nebo se stereo párem klasických kamer [1]. Použití samotného páru termokamer

není běžné. Jedním z prvních projektů, kde byla nasazena samotná stereo termokamera byla vizuální odometrie [7].

Cílem této práce je sestavit stereo termokameru z levných senzorů FLIR Lepton 3.5, provést geometrickou kalibraci jednotlivých kamer, provést kalibraci stereo soustavy a následně otestovat přesnost odhadu vzdálenosti teplého objektu nebo člověka od základny kamery.

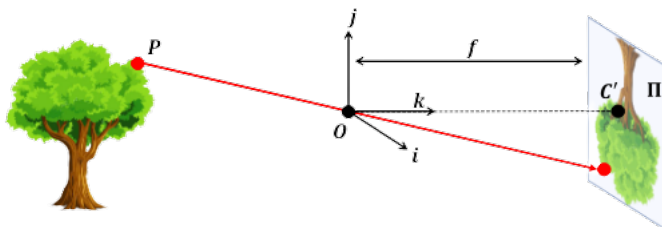
Kapitola 2

Kamera

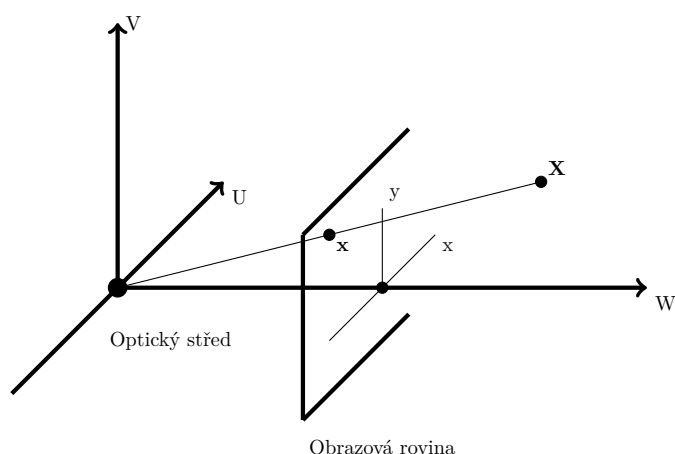
2.1 Dírkový model kamery

Kamery, které jsem používal v této práci, odpovídají dírkovému modelu. Jedná se o velmi jednoduchý, ale přesto hojně používaný koncept, který mapuje body v reálném světě na body ve 2D prostoru. [4]

V reálném světě je dírkový model modelován pomocí uzavřené komory s malým otvorem (pinhole) v přední části. Paprsky od objektů z reálného světa procházejí skrze tuto štěrbinu a formují obraz na tmavé straně komory (obrazová rovina). Z této konstrukce jasně vyplývá, že obraz je převrácen oproti předloze, tento princip je znázorněn na obrázku 2.1 [11]. Místo toho můžeme uvažovat o takzvaném virtuálním obrazu tak, že přesuneme obrazovou plochu před samotnou štěrbinu. Není možné zkonstruovat kameru s virtuálním



Obrázek 2.1: Projekce vzoru do obrazu (2D) [4]



Obrázek 2.2: Nákres dírkové kamery s virtuální obrazovou rovinou

obrazem, ale tento přístup je matematicky ekvivalentní se skutečnou dírkovou kamerou s tím rozdílem, že je přímočařejší na pochopení. Štěrba (pinhole) se taktéž označuje jako optický střed obrazu. Optický střed také definuje počátek kartézského souřadnicového systému, ve kterém jsou body reprezentovány tří-dimenzionálním vektorem $\mathbf{X} = [u, v, w]^T$. Virtuální obrazová plocha je posunuta oproti počátku ve směru osy W (taktéž nazývaná jako optická osa), tento princip je znázorněn na nákresu 2.2. Vzdálenost, o kterou je virtuální obraz posunut od počátku, je rovna ohniskové vzdálenosti $-f$ [11].

Posledním úkolem je určit pozici $\mathbf{x} = [x, y]^T$, kam se zobrazí bod $\mathbf{X} = [u, v, w]^T$. Nejjednodušším způsobem je spojit bod \mathbf{X} s optickým centrem (střed souřadnicového systému) a následně nalézt průsečík s plochou virtuálního obrazu.

Dále uvažujme, že obrazová rovina bude ve vzdálenosti 1 a nikoliv f . Převod bodu \mathbf{X} tedy bude vypadat následovně (světové souřadnice vlevo, obrazové souřadnice vpravo):

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} \longrightarrow \begin{bmatrix} f \cdot u/w \\ f \cdot v/w \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}. \quad (2.1)$$

Souvisejícím problémem je, že při výrobě reálných kamer nemusí být pixely čtvercové, což lze vyřešit různými ohniskovými vzdálenostmi pro různé osy, takže vzniká nový vztah:

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} \longrightarrow \begin{bmatrix} f_x \cdot u/w \\ f_y \cdot v/w \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}. \quad (2.2)$$

Jelikož konvence nepředpokládá začátek souřadnicového systému v optickém středu, ale v levém horním rohu obrazové roviny, je nutné přidat vzdále-

nost bodu od průniku optické osy s rovinou obrazu. Úpravou vztahu 2.2 je získán nový

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} \longrightarrow \begin{bmatrix} f_x \cdot u/w + c_x \\ f_y \cdot v/w + c_y \end{bmatrix}. \quad (2.3)$$

V literatuře se ještě objevuje jeden parametr s (anglicky skew), který dává do vztahu závislost pozici x v závislosti na výšce v ve světových souřadnicích [11]. V této práci budu uvažovat $s = 0$. Rovnice pro převod 2.3 jsou následující:

$$\begin{aligned} x &= \frac{f_x u}{w} + c_x \\ y &= \frac{f_y v}{w} + c_y \end{aligned}. \quad (2.4)$$

Parametry kamery se dají zapsat do matice \mathbf{K} , která se nazývá matice vnitřních parametrů nebo také matice parametrů kamery

$$\mathbf{K} = \begin{bmatrix} f_u & 0 & c_u \\ 0 & f_v & c_v \\ 0 & 0 & 1 \end{bmatrix}. \quad (2.5)$$

■ 2.1.1 Pozice a rotace kamery

Střed souřadnicového systému kamery není vždy ve středu souřadnicového systému okolního světa [11]. Dalším problémem je sdílení souřadnic mezi více kamerami, proto je nutné definovat souřadnicový systém, který budou všechny kamery sdílet. Z tohoto důvodu mohou být předchozí rovnice rozšířeny o transformaci souřadnic kamery do těchto nových (sdílených). Tato transformace souřadnic zahrnuje rotaci bodů a následnou translaci:

$$x' = \mathbf{R}x + \mathbf{t} \quad (2.6)$$

$$x' = [\mathbf{R} \mid \mathbf{t}] \begin{bmatrix} x \\ 1 \end{bmatrix} \quad (2.7)$$

$$\begin{bmatrix} u' \\ v' \\ w' \end{bmatrix} = \begin{bmatrix} \omega_{11} & \omega_{12} & \omega_{13} & t_1 \\ \omega_{21} & \omega_{22} & \omega_{23} & t_2 \\ \omega_{31} & \omega_{32} & \omega_{33} & t_3 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \\ 1 \end{bmatrix}, \quad (2.8)$$

kde x' je transformovaný bod, \mathbf{R} odpovídá matici rotace a \mathbf{t} je translační vektor. Matice $[\mathbf{R} \mid \mathbf{t}]$ se označuje jako vnější parametry kamery.

2.1.2 Projekční matice

Po spojení předchozích rovnic a jejich přepsání do homogenních souřadnic dostáváme nový vztah:

$$\lambda \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f_u & 0 & c_u & 0 \\ 0 & f_v & c_v & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \omega_{11} & \omega_{12} & \omega_{13} & t_1 \\ \omega_{21} & \omega_{22} & \omega_{23} & t_2 \\ \omega_{31} & \omega_{32} & \omega_{33} & t_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \\ 1 \end{bmatrix} \quad (2.9)$$

nebo v maticovém zápisu

$$\lambda \mathbf{x} = \begin{bmatrix} \mathbf{K} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Omega} & \boldsymbol{\tau} \\ \mathbf{0}^T & 1 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \\ 1 \end{bmatrix} \quad (2.10)$$

$$\lambda \mathbf{x} = \mathbf{P} \mathbf{X}. \quad (2.11)$$

Součinem prvních dvou matic vznikne matice \mathbf{P} , která se nazývá projekční matice kamery.

2.2 Distorze

Dírkové kamery se mohou potýkat s různými typy distorze, hlavními jsou radiální a tangenciální. První z nich, radiální distorze způsobuje to, že se rovné křivky (linie) zakulacují. Tímto typem distorze jsou nejvíce ovlivněny body na okrajích, jelikož radiální distorze se zvětšuje směrem od středu do krajů [12]. Tento princip může být popsán pomocí rovnic 2.12. U radiální distorze často rozlišujeme soudkovitost (nafouknutí středu snímku) a poduškovitost (vztažené rohy). Tento rozdíl je zobrazen na obrázku 2.3.

$$x_{dist} = x(1 + k_1 r^2 + k_2 r^4 + k_3 r_6) \quad (2.12)$$

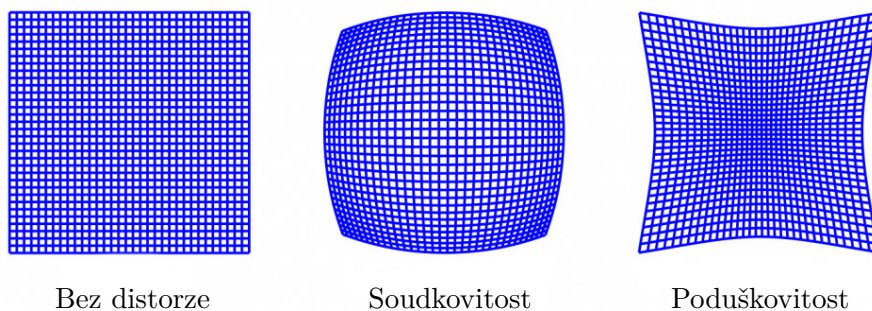
$$y_{dist} = y(1 + k_1 r^2 + k_2 r^4 + k_3 r_6) \quad (2.13)$$

Tangenciální distorze je způsobena tím, že čočka (senzor kamery), není rovnoběžný s projekční rovinou. To způsobuje, že se některé části obrazu zdají být blíže nebo naopak dále než je očekáváno [12]. Opět může být tento jev popsán pomocí rovnic:

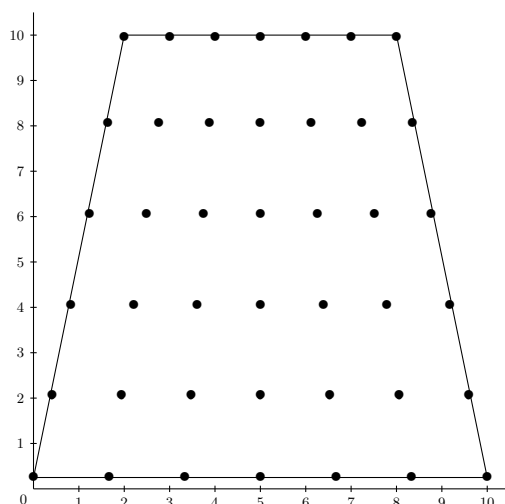
$$x_{dist} = x + [2p_1 xy + p_2(r^2 + 2x^2)] \quad (2.14)$$

$$y_{dist} = y + [p_1(r^2 + 2y^2) + p_2 xy]. \quad (2.15)$$

Ve výsledku je nutné nalézt pět parametrů, které se nazývají distorzní koeficienty: $[k_1 k_2 p_1 p_2 k_3]$ [10].



Obrázek 2.3: Radiální distorze, převzato z [12]

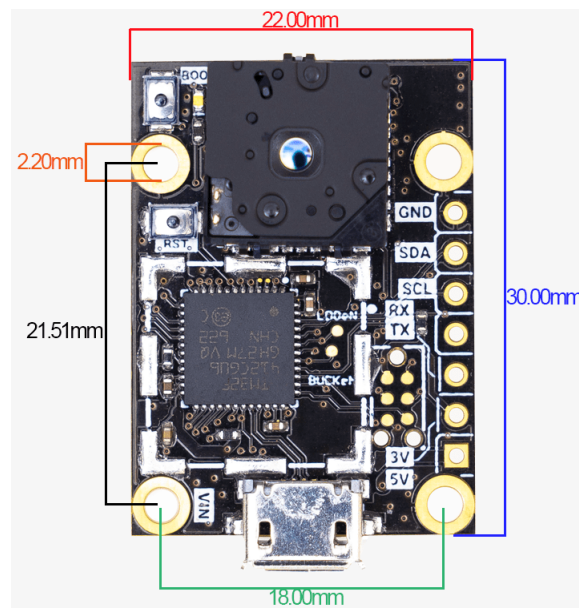


Obrázek 2.4: Tangenciální distorze, převzato z [13]

2.3 Geometrická kalibrace kamer

Kalibrace kamer slouží k nalezení konkrétních parametrů fyzických kamer. Následně lze pomocí nalezených parametrů odstranit distorzi čočky, určit velikost objektů nebo odvodit pozici kamery v prostoru. Tyto informace jsou hojně využívány v počítačovém vidění [5].

Parametry kamery obsahují vnitřní, vnější parametry a distorzní koeficienty. K nalezení těchto parametrů slouží kalibrační vzor, u kterého známe všechny potřebné vlastnosti, jako je například velikost a rozložení. V počítačovém vidění lze nalézt velké množství různých kalibračních vzorů, ale nejčastěji používaným je vzor založený na vzhladu šachovnice. Důležitý je takový vzor, u kterého je možno detekovat určité signifikantní body. Například u šachovnice jsou tím rohy jednotlivých políček. Nasnímáním vzoru z různých úhlů lze parametry kamery odhadnout. Jedním z algoritmů pro odhad parametrů



Obrázek 2.5: Termokamera FLIR lepton 3.5 na platformě PureThermal 2

výsledků kalibrace a minimalizovat reprojekční chyby. Rozlišení kamery nekomplikuje pouze její kalibraci, ale i aplikaci dalších metod počítačového vidění, které byly použity.

Na obrázku 2.5 je kamera FLIR Lepton 3.5, která je připevněna na platformě PureThermal 2 Smart I/O. Tato deska umožňuje čtení dat z kamery pomocí USB nebo pomocí rozhraní UART. V mém případě bylo využito připojení kamery pomocí USB. Deska PureThermal je předkonfigurovaná, tak aby fungovala v režimu plug-and-play.

■ 2.4.1 Čtení dat z kamery

Jak jsem dříve zmínil, kamera se pro operační systém jeví jako každá jiná web kamera, jelikož je takto PureThermal nakonfigurován. Z tohoto důvodu je možné ji bez větších problémů použít za použití například softwaru VLC media player, pro který poskytuje obraz ve falešných barvách. Sama společnost vyrábějící platformu PureThermal, poskytuje vlastní nástroj pro čtení dat z kamery. Tento nástroj je open-source a je možné ho dále využívat.

V mém případě jsem využíval knihovny openCV pro python, která nabízí velké množství funkcí na práci s obrazem obecně. Pro čtení dat z kamery jsem používal komponentu *VideoCapture* z openCV, které je potřeba nastavit cestu

ke kameře. Ta se může lišit v závislosti na operačním systému, v mém případě (ubuntu 20.04 LTS) se jedná například o `/dev/video2` nebo `/dev/video4` v závislosti na tom, zda se jedná o levou či pravou kameru. Typické chování mého operačního systému bylo takové, že záleželo na pořadí připojení kamer. První zapojená kamera se mapovala na `/dev/video2` a druhá na `/dev/video4`. Dalším důležitým krokem je nastavení parametrů souvisejícími s formátem čtení. Jelikož `VideoCapture` musí mít nastaveno kódování, ve kterém bude daná data číst. Jedná se konkrétně o parametry `CAP_PROP_FOURCC` a `CAP_PROP_CONVERT_RGB`. Pro čtení teplotních dat musí být způsob kódování (kodek) nastaven na Y16 a pro čtení ve falešných barvách YUVY.

```
# reading image from camera false colors / raw data
cam = cv2.VideoCapture(cam_path)
if raw: #temperature data
    cam.set(cv2.CAP_PROP_FOURCC,
            cv2.VideoWriter.fourcc('Y', '1', '6', ' '))
    cam.left_cam.set(cv2.CAP_PROP_CONVERT_RGB, 0)
else: #false colors
    cam.set(cv2.CAP_PROP_FOURCC,
            cv2.VideoWriter.fourcc('Y', 'U', 'V', 'Y'))
    cam.set(cv2.CAP_PROP_CONVERT_RGB, 1)
```

```
#read img from camera
succ, img = cam.read()
```

Když je zvolen teplotní formát, tak data nelze přímo zobrazit v obrazové formě, jelikož se jedná o 16 bitové hodnoty. Z toho důvodu je důležité je nějakým způsobem převést na barevnou škálu. Těchto barevných škál můžeme mít velkou řadu a každá odpovídá jinému použití. Jedním z nejpřímochařejších zobrazení je to, kde jsou teplotní data převedeny do odstínů šedé. Existuje množství takovýchto transformací, ale ideálním příkladem, který jsem využíval i já, je lineární mapování. Zde je vybrán maximální a minimální prvek z teplotních dat a následně jsou ostatní převedeny do rozmezí 0 - 255 (hodnota šedi).

```
def linear_mapping(data, rgb_output=True):
    cv2.normalize(data, data, 0, 255, cv2.NORM_MINMAX)
    if rgb_output:
        out = cv2.cvtColor(np.uint8(data), cv2.COLOR_GRAY2RGB)
        return out
    return np.uint8(data)
```

I když má kamera rozlišení 160x120 pixelů, teplotní data mají velikost 160x122. To je způsobeno tím, že poslední dva řádky obsahují telemetrická data, jako je čas běhu kamer, status a další [8].

Kapitola 3

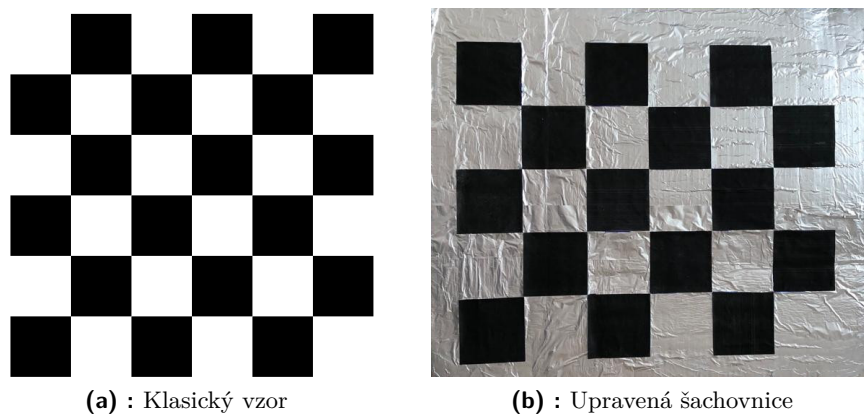
Geometrická kalibrace termokamery

3.1 Geometrická kalibrace RGB kamer

Geometrická kalibrace kamer pracujících ve viditelném spektru je velice dobře zdokumentovaná a známá věc. Existuje řada metod, kterými lze nalézt projekční matici kamery. Jednou z nejrozšířenějších knihoven používaných v počítačovém vidění je openCV, jejíž součástí je i implementace algoritmu pro geometrickou kalibraci kamery. Součástí této knihovny je taktéž detektor signifikantních bodů kalibračního vzoru, pro šachovnicové vzory to jsou rohy jednotlivých políček. Implementace detekce rohů v openCV vyžaduje, aby šachovnicový vzor byl černo-bílý s tím, že vzor je na bílém podkladu (okolo šachovnice je bílý okraj).

3.2 Kalibrace termokamer

Na rozdíl od RGB kamer je kalibrace termokamer o něco složitější, jelikož tyto senzory generují výstupní obraz na základě teploty. Proto zde nelze použít běžně používané kalibrační vzory jako u běžných kamer. Obecně se nabízejí dvě metody, jak dosáhnout toho, aby byl kalibrační vzor viditelný. První metodou je pasivní zahřívání a druhou aktivní zahřívání.



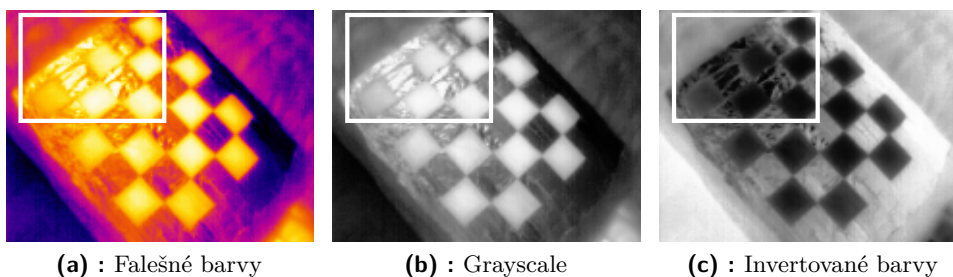
Obrázek 3.1: Šachovnicové kalibrační vzory

3.2.1 Aktivně zahříváné vzory

Aktivně zahříváné zdroje jsou většinou konstrukčně náročnější a existuje zde velká variabilita. Příkladem takového vzoru může být mřížka složená ze žárovek, které generují teplo [3]. Souvisejícím problémem je nalezení signifikantních bodů kalibračního vzoru. V uvedené publikaci [3] Zoetgnande s týmem tento problém vyřešil pomocí prahování a následným dopočítáním těžišť jednotlivých segmentovaných bodů. Obecně jsou aktivně zahříváné vzory robustnější, jelikož zde není přímá závislost na prostředí.

3.2.2 Pasivně zahříváné vzory

Hlavní výhodou pasivního zahřívání je jeho jednoduchost, jelikož je možné používat klasické kalibrační vzory a následně využívat všech výhod knihoven, které se používají v počítačovém vidění. Obecně tento přístup využívá různých vlastností materiálů a barev, ze kterých jsou kalibrační vzory vyrobeny. Úplně nejjednodušším a nejprůchoďejším přístupem je využití klasické černo-bílé šachovnice, která je zahřívána externím zdrojem, tím může být lampa nebo sluneční svit [2]. Vlastnosti barev zapříčiní to, že jednotlivé barevné části, černá a bílá, nebudou dosahovat stejné teploty [3]. Díky tomuto jevu je možné nasnímat kalibrační vzor tak, aby byl viditelný na termokameře.

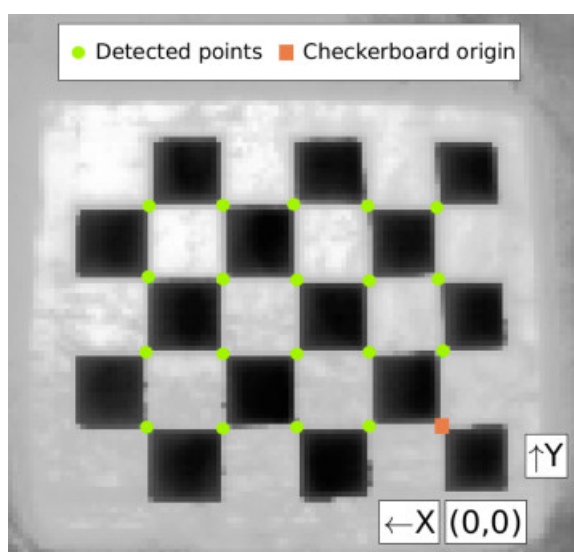


Obrázek 3.2: Odraz teplého objektu na kalibrační desce, ohraničen bílým obdélníkem

3.3 Úprava a vylepšení kalibračního vzoru

Pro kalibraci kamery, kterou jsem používal, jsem zvolil pasivně zahříváný vzor šachovnice. V první fázi jsem použil neupravenou šachovnici o velikosti pole 2 cm a s rozložením 5x6 polí. Použití neupraveného vzoru vedlo k velmi špatné detekci rohů políček a následně ke špatným výsledkům celkové kalibrace. Proto musel být vzor vylepšen, aby rozdíly teplot mezi sousedními políčky byl významnější. Úprava kalibrační desky byla provedena za pomoci hliníkové fólie, kterou byly všechny bílé části polepeny. Důvodem použití této fólie je, že má velmi vysokou odrazivost. Využití této vlastnosti a efektu studené oblohy [7] pomohlo zvýšení kontrastu pořizovaných snímků. Takto upravenou desku je možné použít pro kalibraci, přičemž nejlepších výsledků je dosahováno, když je kalibrační deska umístěna na místo, kde odráží chladnou oblohu (venku) a zároveň na ni mohou dopadat sluneční paprsky. Na obrázku 3.3 je zobrazena zahřátá kalibrační deska s následným aplikováním algoritmu na detekci rohů jednotlivých polí. Tento postup je náchylný na umístění vzoru, jelikož při snímání kalibrační desky kamerou může hliníková fólie odrážet i teplé objekty, jako jsou například budovy, na obrázku 3.2 je takový odraz vidět. Tento jev může být jednoduše přirovnán k odrazu objektů pomocí zrcadla. Nejlepší podmínky pro kalibraci kamer jsou dny s malou oblačností, silným slunečním světlem a bez větru, kdy je velký teplotní rozdíl mezi zahřátými černými políčky od slunce a políčky, které jsou pokryté fólií a odrážejí oblohu.

Tato kombinace jevů umožňuje nasnímat kalibrační vzor v dostatečně velkém kontrastu, aby byl zpracovatelný běžně používanými algoritmy pro kalibraci kamer. V mém případě byla použita knihovna openCV. Jelikož použité kamery mají poměrně malé rozlišení, musely být výstupní obrázky zvětšeny tak, aby je mohl zpracovat algoritmus na detekování rohů políček šachovnice. Konkrétně byly obrázky zvětšovány koeficientem 2, 3, 4 a 5, následně byly souřadnice detekovaných rohů opět převedeny do původního měřítka. Výsledky kalibrace pro různá zvětšení jsou uvedena v tabulce 3.1.



Obrázek 3.3: Obrázek kalibrační desky po zpracování algoritmem na detekci rohů

Pro zvětšování obrázků jsem používal následující funkci z openCV:

```
img = cv2.resize(img, None, fx=MULT_FACTOR, fy=MULT_FACTOR,
interpolation=cv2.INTER_CUBIC)
```

Knihovna na detekci rohů na vstupu požaduje černé políčka na bílém podkladu, jelikož jsem používal výstup kamery ve falešných barvách, musel jsem barvy invertovat. Z tohoto důvodu jsou černá políčka černá, i když by měla být teplá (a proto světlá). Druhým problémem je šum, způsobený nedokonalostí samotné kamery [6]. Ve svém přístupu jsem se tento jev snažil zmírnit použitím velkého data-setu fotografií zachycující kalibrační vzor.

Výhodou použití takto upraveného kalibračního vzoru je možnost použít stejný vzor jak pro termokamery, tak pro kamery pracující ve viditelném spektru. Těto vlastnosti se dá využít například při kombinaci stereo termokamer s RGB kamerou nebo kamerami.

3.4 Kvalita kalibrace

Když jsou parametry kamery nalezeny, je důležité určit zda byla kalibrace provedena dobře nebo ne. Hlavním ukazatelem je průměrná reprojekční chyba (MRE), která je definovaná jako průměrná euklidovská vzdálenost mezi souřadnicemi projektovaných bodů a těmi, které byly detekovány na

Zvětšení	1	2	3	4	5
průmerné MRE v datasetu	0.14	0.14	0.08	0.05	0.05
max MRE v datasetu	0.24	0.20	0.17	0.10	0.10
poměr detekcí rohů	50/70	57/70	60/70	60/70	55/70

Tabulka 3.1: Tabulka reprojekčních chyb

referenčním snímkem. Tato průměrná chyba je měřena pro jeden konkrétní snímek. Zde je využito toho, že kalibrační vzor má svůj levý horní roh ve středu globálních souřadnic. Jelikož je známá velikost tohoto vzoru a rozložení, například šachovnice s velikostí pole 7 cm a rozložením 4x5 polí, můžeme tyto body z reálného světa projektovat do souřadnic v obrázku. MRE je měřena v pixelech, obecně se za dobře kalibrovanou kameru považuje odchylka, která je menší než 1 px.

Hodnoty průměrných reprojekčních chyb (MRE) pro jednotlivé zvětšení na stejné sadě snímků jsou uloženy v tabulce 3.1, taktéž je zde zobrazena maximální spočítaná hodnota MRE v datové sadě.

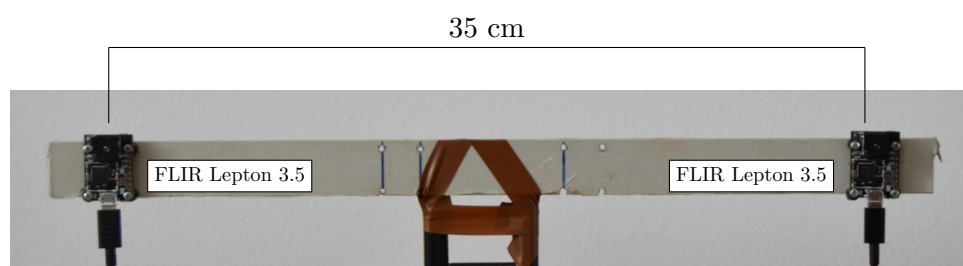
Kapitola 4

Stereo kamera

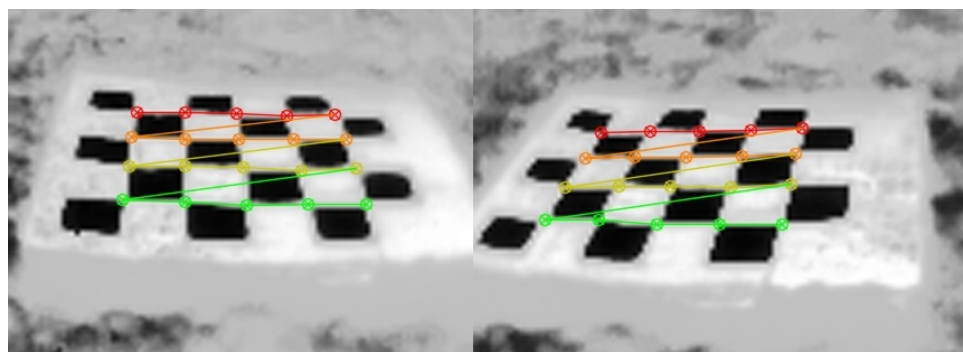
4.1 Stereo vidění

Koncept stereo vidění je inspirován biologií člověka, jelikož při projekci bodu z reálného světa do obrazových souřadnic dochází ke ztrátě informace o vzdálenosti bodu od kamery, to stejné se děje u člověka, když se dívá pouze jedním okem. Systém dvou a více kamer, které jsou od sebe vzdáleny o určitou vzdálenost značenou jako b , se v počítačovém vidění nazývá stereo soustava (stereo kamera) a snaží se informaci o vzdálenosti bodu od kamery odhadnout.

Stereo kamery mohou mít velké množství konfigurací (vzdálenost od sebe, počet kamer nebo úhel mezi obrazovými plochami kamer), jedna z konfigurací je vidět na obrázku 4.1. V obecném smyslu se často jedná o dvě a více kamer, které mají stejné vlastnosti, těmi jsou například rozlišení, ohnisková vzdálenost nebo barevná hloubka. Samozřejmě existují soustavy, kde je využito rozdílných kamer, ale jedná se převážně o specifické použití. V této práci jsem použil dvojici kamer se stejnými parametry (FLIR Lepton 3.5), které byly umístěny v jedné rovině. Často se tato konfigurace nazývá jako jednoduchý stereo systém. Výstupem stereo kamery je rozdílová mapa (disparity map), která poskytuje informaci o tom, o kolik pixelů se liší pozice jednotlivých objektů v levém a pravém obrázku. Obecně platí, že objekty které jsou blíže k základně kamery se liší o větší vzdálenost než ty, které jsou dále. Objekty, které leží ve velké vzdálenosti (jsou nekonečně daleko), by se neměly vůbec vzdálenostně lišit. Ke generování hloubkových map z dvojice obrázků se používají stereo matching algoritmy.



Obrázek 4.1: Stereo termokamera



Obrázek 4.2: Kalibrační deska ve stereo soustavě - výřez (vlevo levá kamera, vpravo pravá)

4.2 Kalibrace stereokamery

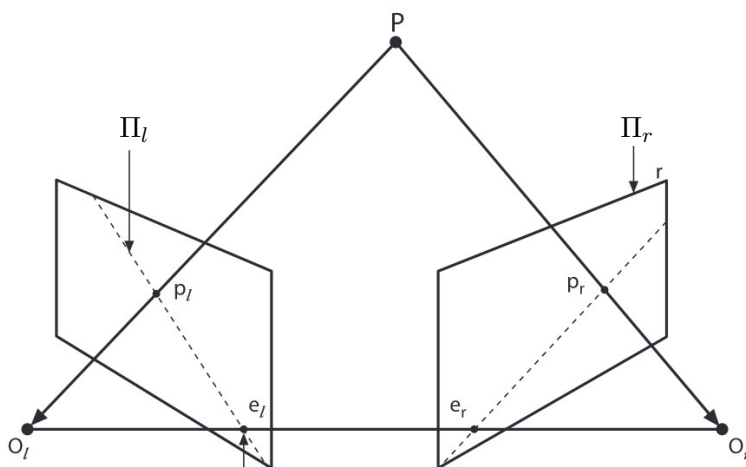
K tomu aby bylo možné správně rekonstruovat 3D prostor je nutné provést kalibraci stereo kamery. Ta je opět založena na stejném kalibračním vzoru jako geometrická kalibrace jedné kamery s tím rozdílem, že soustava musí snímat stejnou kalibrační desku ve stejný čas, toto zobrazuje obrázek 4.2. V mém přístupu jsem už při snímání kalibrační desky pro jednu kameru ukládal snímky z obou kamer, tak aby bylo možno se stejným datasetem kalibrovat jednu kameru nebo obě zároveň. Zároveň není nutné opět detekovat rohy šachovnice a je možné použít výstupy kalibrací jednotlivých kamer.

Parametry, které chceme o stereo soustavě zjistit jsou rotace a posunutí jednotlivých kamer. Tyto informace je možné zjistit, jelikož známe velikost kalibračního vzoru a jeho umístění v prostoru. I když používám jednoduchou stereo soustavu není rotace nulová, jelikož upevnění samotných kamer ke konstrukci není přesné, taktéž samotné uložení FLIR kamery v platformě PureThermal není rovněž zcela přesné. Dalším parametrem, který nalezneme stereo kalibrace je fundamentální matice. Ta je důležitá pro další vlastnost, kterou stereokamera má, tou jsou epipolární roviny. S nalezenou rotací, translací, maticemi obou kamer a distorzními parametry můžeme provést rektifikaci neboli opravu. Pro kalibraci soustavy byla využita funkce z knihovny openCV

stereoCalibrate, jejíž vstupem jsou souřadnice bodů v prostoru (levý horní roh šachovnice je v počátku), detekované rohy šachovnice na levém a pravém obraze, parametry kamer (matice kamery a distorzní parametry). Jelikož jsem prováděl stereo kalibraci s již známými parametry kamery, používal jsem přepínač pro použití těchto parametrů¹.

4.3 Epipolární geometrie

Základní geometrií ve stereo systému se nazývá epipolární geometrie. Tato geometrie kombinuje dva dírkové modely (jeden pro každou kameru). Pro každou kameru máme vlastní centrum projekce O_l a O_r , a také příslušné projekční roviny Π_l a Π_r (viz obrázek 4.3). Bod P z reálného světa se projektuje na každou z těchto ploch (značené jako p_l a p_r). Jako epipóly jsou označeny body, které je obrazem spojnice středů projekce obou kamer (značení e_l a e_r). Plocha, která vznikne v prostoru mezi epipóly a bodem P se nazývá epipolární rovina (plocha). Přímkou na kterých leží p_l respektive p_r a e_l respektive e_r se nazývají epipoláry [13].



Obrázek 4.3: Epipolární geometrie, převzato z [13]

¹funkce *stereoCalibrate* umí tyto parametry spočítat, pokud nejsou známy

Nejdůležitějším důsledkem epipolární geometrie je mapování bodu v levém obrázku na přímku v pravém (a opačně). Toto je určeno vnitřními geometrickými vlastnostmi, jež jsou interpretovány tak zvanou fundamentální maticí F [15]. Pro každé dva body, x_r a x_l , které jsou projekcí stejného bodu X (z reálného světa) platí:

$$x_r^T F x_l = 0 \quad (4.1)$$

$$l_r = F x_l \quad (4.2)$$

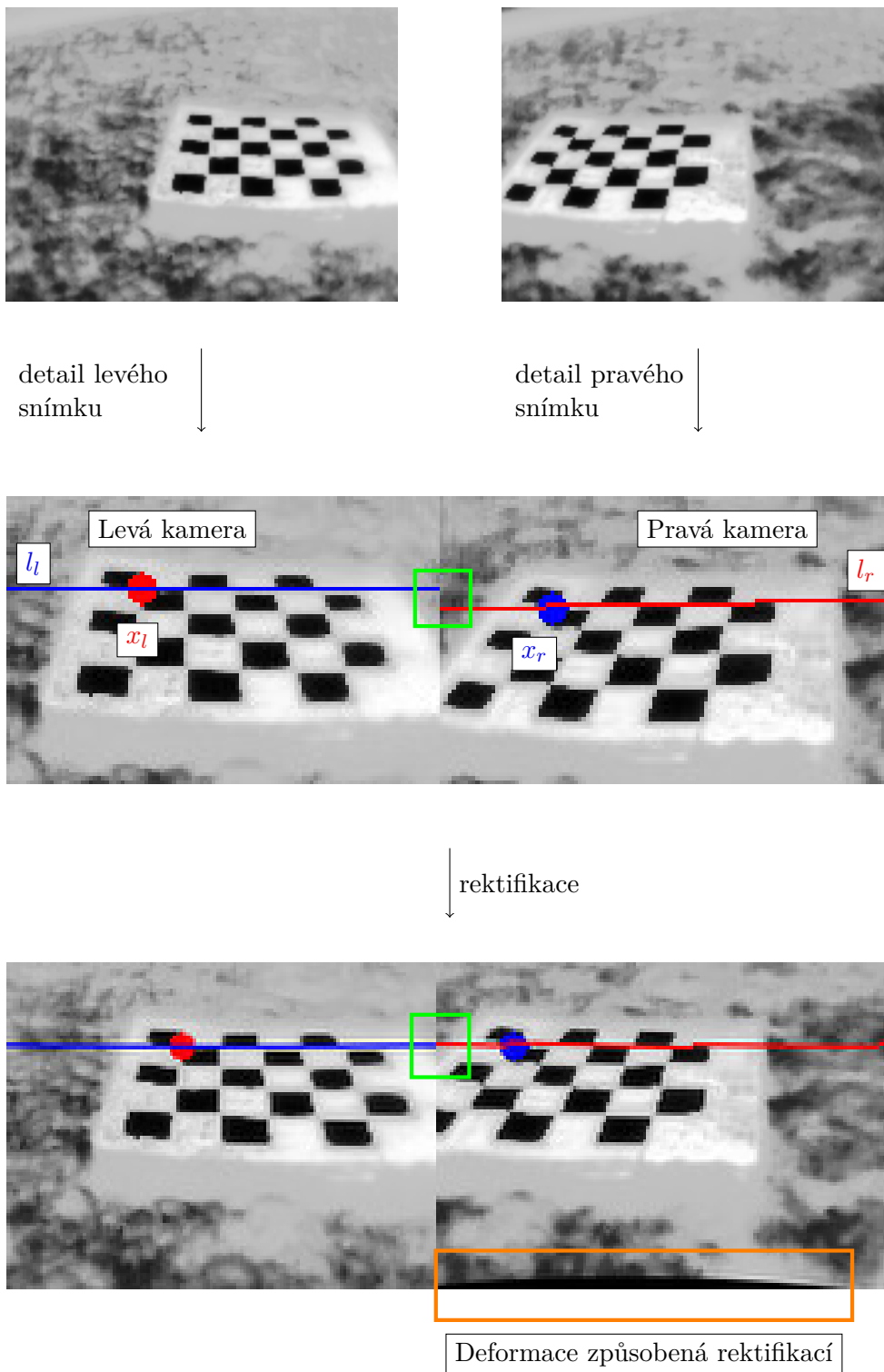
$$l_l = F x_r, \quad (4.3)$$

kde l_r respektive l_l jsou epipoláry odpovídající bodu x_l respektive x_r . Pokud jsou obě optické osy rovnoběžné, pak odpovídající body z levého obrázku leží ve stejném řádku pravého obrázku [14]. Tím pádem jsou epipoláry rovnoběžné se spodním a horním okrajem obrázku.

4.4 Rektifikace obrazu

Ze znalosti rotace a translace pomocí stereo kalibrace soustavy a znalosti parametrů kamer je možné provést stereo rektifikaci. Rektifikace slouží k tomu, aby epipoláry byly rovnoběžné se spodním okrajem obrazu. V důsledku této transformace je možné hledat korespondující body jen v řádku a výpočet se tím velice zjednoduší. Opět byla využita funkce z knihovny openCV, konkrétně se jedná o *stereoRectify*. Parametry, které jsou tímto získány je možné z pořízeného obrázku odstranit distorzi a následně vyrovnat epipoláry, tento princip je ukázán na obrázku 4.4.

Pro účely znázornění principu rektifikace v obrázku 4.4, musely být původní obrázky zvětšeny a oříznuty tak, aby bylo možné ukázat detailně rozdíl v epipolárách.



Obrázek 4.4: Proces rektifikace

4.5 Stereo matching algoritmy

Existuje množství metod pro nalezení stereo párů bodů, obecně je dělíme do dvou kategorií a to plošné a ty založené na rysech. Z epipolární geometrie je známo, že dva korespondující body v levém a pravém obraze musí ležet na epipoláře. Dále se používají i jiná omezení jako je částečná spojitost rozdílůvých map, jedinečnost - každému pixelu z levého obrazu odpovídá právě jeden z pravého a pořadí pixelů (pořadí z levého odpovídá tomu v pravém). Vstupem pro tyto metody, často bývají rektifikované snímky tak, aby epipolární přímky na sebe navazovaly a hledání stereo páru bodů se zjednodušilo na hledání odpovídajícího bodu v řádku.

4.5.1 Plošné metody

Při použití plošných metod pro hledání stereo páru bodů, popisuje tento pixel jeho malé okolí. Když je vstupem rektifikovaný obrázek, je možné následně prohledat odpovídající řádek v druhém obrázku a najít takový, který má co nejpodobnější popis. Pro nalezení odpovídajícího korespondenčního bodu slouží různé funkce, příkladem takovéto funkce může být součet absolutních rozdílů intenzit (SAD - sum of absolute intensity differences) [18].

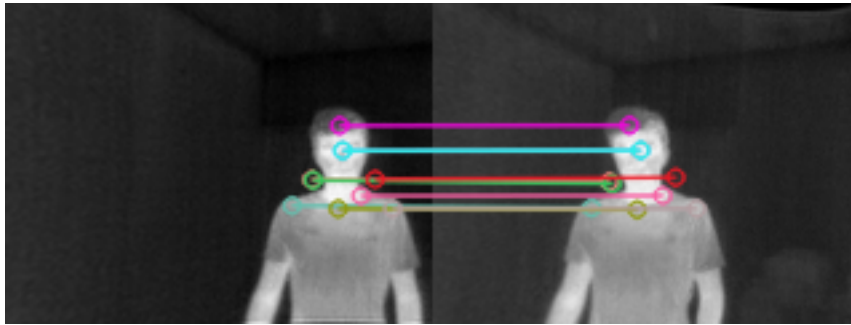
$$\sum_{(x,y) \in W} |I_R(x,y) - I_L(x+d,y)| \quad (4.4)$$

Kde $I(x,y)$ je intenzita daného pixelu a d je posun ve druhém obrázku. Jak už jsem zde uvedl, existuje řada modifikací jako součet kvadrátu rozdílů intenzit a další. Obecně tento typ algoritmů funguje velmi dobře, ale potýká se s problémy na obrázcích, které obsahují opakující se vzor a je tak velmi těžké určit odpovídající páry.

4.5.2 Porovnávání rysů

Při použití této metody, je důležité nalezení signifikantních bodů, které jsou na obou snímcích stejné. Můžou to být hrany nebo rohy následně můžeme tyto větší celky rozložit na menší a u nich určit o jakou vzdálenost se liší. Pro následný popis těchto klíčových bodů můžeme využít například SIFT, SURF nebo BRIEF [19]. Dalším krokem je jejich párování, jelikož nemáme jistotu, že klíčových bodů bude stejné množství, nebo že každý klíčový bod bude mít odpovídající korespondující bod, ve druhém obrázku musíme tyto

omezení nějak vyřešit. Opět zde hrají velkou roli epipoláry, na kterých daný bod musí ležet. Následně můžeme použít pomocí popisovacích algoritmu najít ten, který leží na epipolární rovině a nejvíce odpovídá danému bodu (pomocí K-NN)[14]. Největším problémem, při hledání signifikantních bodů může být to, že jich není nalezeno dost nebo jich je jen malé množství. Na obrázku 4.5 lze vidět, jak funguje porovnávání rysů na snímcích ze stereo termokamery.



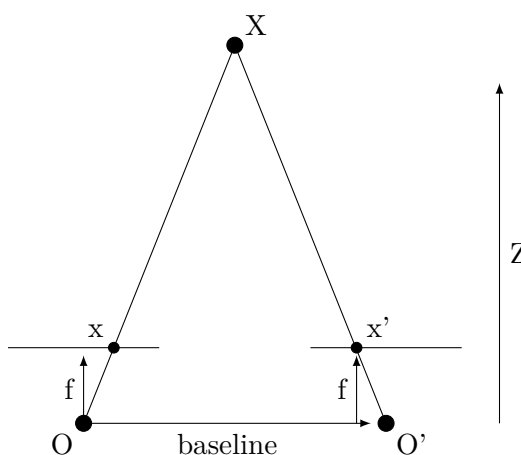
Obrázek 4.5: Porovnávání rysů - stereo termokamera 2x FLIR Lepton 3.5

4.6 Výpočet vzdálenosti

Konfigurace používaná v této práci je složena ze dvou kamer v jednoduchém stereo systému, které jsou od sebe vzdálené o 35 cm značená jako b (to se nazývá báze nebo baseline), jedná se o soustavu z obrázku 4.1. Pomocí rozdílových map, které generují stereo matching algoritmy je možné zjistit o kolik pixelů se liší jednotlivé objekty v levém a pravém obraze. Ze znalostí parametrů stereo soustavy jako je ohnisková vzdálenost, velikosti základny a podobnosti trojúhelníků (které vyplývají z epipolární geometrie) je možné odvodit následující vztah:

$$\begin{aligned} x - x' &= \frac{bf}{Z} \\ Z &= \frac{f}{x - x'} = \frac{bf}{d}, \end{aligned} \quad (4.5)$$

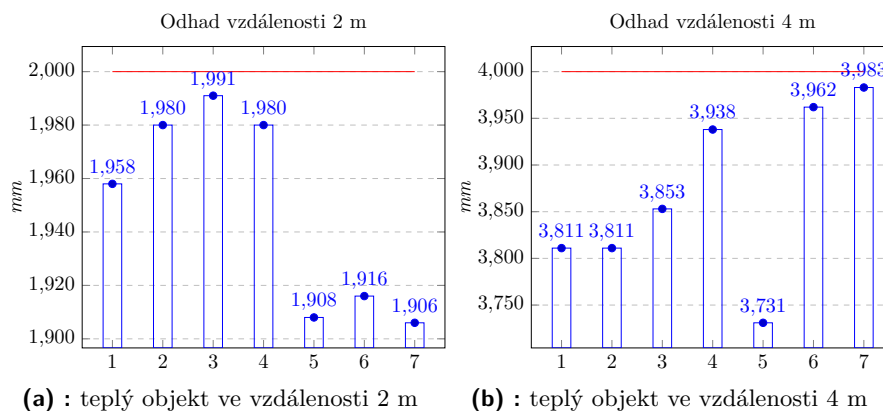
kde Z odpovídá vzdálenosti bodu od kamery, x je x-ová pozice v levém obrázku, x' je x-ová pozice v pravém obrázku a d může být získán z rozdílové mapy (viz. obrázek 4.6).



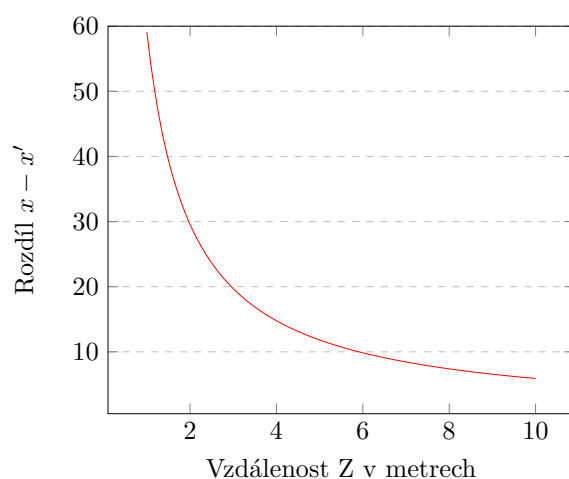
Obrázek 4.6: Odhad vzdálenosti pomocí triangulace [9]

4.7 Přesnost stereo soustavy

Důležitým měřítkem stereo kamery je to, jak dobře dokáže soustava odhadovat vzdálenost bodu od základny kamery. K takovému měření jsem využil teplý objekt, v mém případě se jednalo o opakovatelně použitelné tepelné polštářky na ohřátí rukou (který se používá v outdoorových sportech). Ten dosahuje teploty až 54°C a jeho rozměry jsou $90\text{ mm} \times 90\text{ mm}$. Následně jsem tento objekt nasnímal z obou kamer a spočítal těžiště tohoto teplého objektu v každém obrázku. Jelikož se jednalo o jediný teplý bod na scéně, bylo možné pomocí prahování separovat tento objekt. Z parametrů získané stereo kalibrací a rozdílů středů objektů jsem získal vzdálenost. Tato vzdálenost je vynesena v grafu 4.7, konkrétně se jedná o objekt ve vzdálenosti 2 a 4 metry.



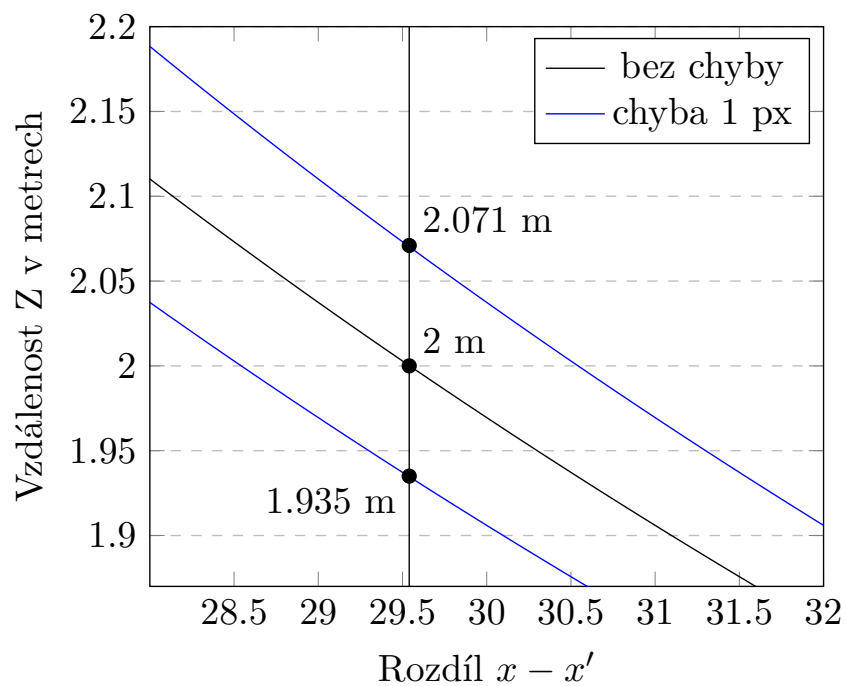
Obrázek 4.7: Odhad vzdáleností teplého objektu



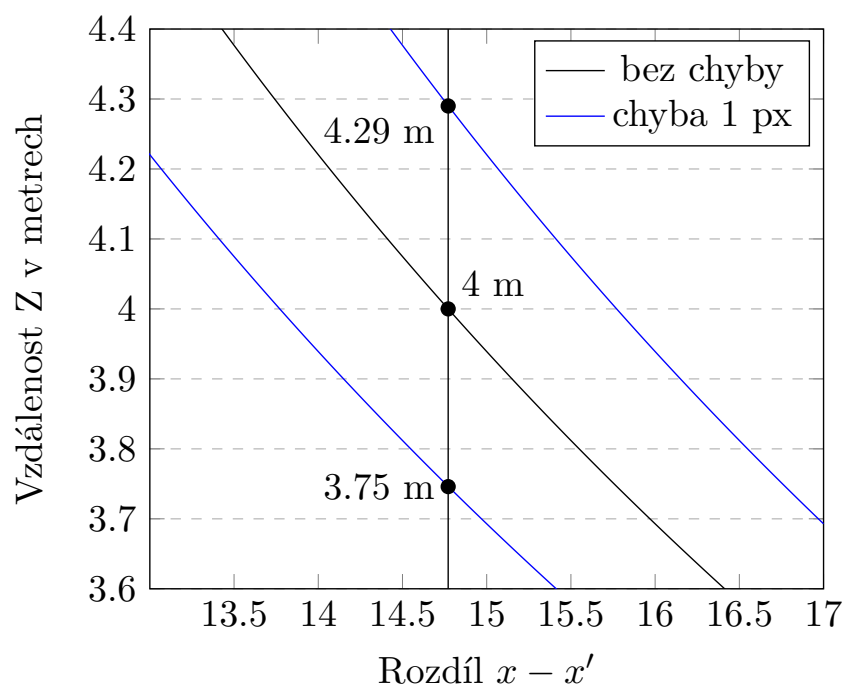
Obrázek 4.8: Ideální průběh disparity v závislosti na vzdálenosti

Ideální průběh rozdílu $x - x'$ v závislosti na vzdálenosti je pro naši stereo soustavu zobrazen na obrázku 4.8. Z grafu lze vyčíst, že rozdíl korespondujících bodů se ve větší vzdálenosti bude lišit jen mírně a tím pádem se i přesnost odhadu vzdálenosti bude zhoršovat. To klade velké nároky na stereo matching algoritmy, aby správně a přesně určovali odpovídající stereo páry bodů. Další důležitou věcí je, že rozdíl $x - x'$ by měl nabývat jen celočíselných hodnot. V mém případě předchozí věta není zcela pravdivá, jelikož jsem používal rozdíl těžišť, která ve většině případů nevycházejí celočíselně.

Jak už jsem již dříve zmínil, hledání korespondujících bodů z levého a pravého snímku by mělo být maximálně přesné. Na grafech 4.9 je zobrazeno, jak se změní odhad vzdálenosti objektu, pokud je chyba detekce rozdílu x-ových souřadnic rovna ± 1 px. Zde je jasně vidět, že s klesající disparitou roste vliv chyby na přesnost odhadu vzdálenosti.



(a) : chyba 1px při 2 m



(b) : chyba 1 px při 4 m

Obrázek 4.9: Odhad vzdálenosti při chybě 1 px při určování stereo páru

Kapitola 5

Detekce teplých objektů a odhad vzdálenosti

5.1 Detekce objektů

Detekce objektů je v počítačovém vidění velmi žhavé téma, které je předmětem řady zkoumání. Obecně se problém dělí do dvou částí a to lokalizace a klasifikace. Existuje řada metod, jak detekce objektů dosáhnout. V posledních letech se problém detekce často řeší pomocí neuronových sítí, které dokáží při správné architektuře a trénovacích datech velice dobře generalizovat a díky tomu detekovat i neznáme objekty. Příkladem takovéto hojně využívané neuronové sítě může být YOLO network (You Look Only Ones), která se podle mnohých řadí k state-of-the-art metodám v detekci objektů. Problémem je jejich poměrně vysoký požadavek na výpočetní sílu hardwaru. Samozřejmě existují i metody, které nevyžadují takové množství výpočetního výkonu. Mezi tyto se řadí například Viola–Jones object detector, který využívá haarovské filtry a kaskádovou architekturu.

Jelikož se výše zmíněné metody používají na snímky pořízené senzory pracující ve viditelném spektru, je velmi dobrý důvod se domnívat, že budou fungovat i u termálních snímků. Jak neuronová síť YOLO [20, 22], tak Viola–Jones detektor [21] byly již pro tuto úlohu využity. Zde je nejdříve důležité rozhodnout s jakými hodnotami se bude pracovat, zda přímo s teplotními daty (kde každý pixel má přiřazenou teplotu reprezentovanou 2 bajtovým celým číslem) nebo zda bude teplotní obrázek mapován na nějaké barevné spektrum.

Obecně všechny zmíněné metody pracující s RGB vstupem a rozlišují objekty na základě tříd, které jsou hledány a na kterých jsou učeny. Pokud je vstupem barevný obrázek, je velice těžké rozhodnout zda je nějaký objekt teplý či nikoliv, jelikož se často barevné spektrum obrázku odvíjí od minimální a maximální teploty na scéně. Z tohoto důvodu je logické sáhnout po teplotním formátu dat a využívat je.

5.2 Detektor teplých objektů

Pro konstrukci detektoru teplých objektů jsem zvolil velice přímočarý přístup, který je založen na prahování. Jelikož mám přístup k teplotním datům, je možné jednoduše filtrovat pixely, které přesahují předem definovanou teplotní hodnotu. Tento přístup je velice přímočarý a z toho důvodu existuje už řada takových implementací. Příkladem takového detektoru je [16], kde bylo využito teplotního prahování k detekci obličeje. Druhým příkladem je [22], kde byla stejná konstrukce použita pro detekci osob.

Můj detektor se skládá z několika částí, které jsou za sebou napojeny, kde výsledkem jsou obdélníky ohraničená teplá místa, včetně jejich kontury. Rozdíl je, že tuto detekci dělám na obou kamerách, tak abych měl kompletní informaci o scéně. Průběh detekce (taktéž znázorněn v jednotlivých krocích na obrázku 5.2):

1. načtení dat z kamer a nastavení parametrů
2. prahování teplot
3. čištění obrazu
4. ohraničení oblastí konturami
5. filtrování kontur na základě plochy
6. ohraničení kontur nejmenším možným obdélníkem (bounding boxes)

5.2.1 Načtení dat z kamer a nastavení parametrů

Načtení dat z kamery probíhá v teplotním formátu, zpracování načítání je popsáno v 2.4.1. Následně jsou oba obrázky zbaveny distorze a rektifikované.

Parametry, které je nutné nastavit pro to, aby algoritmus dobře fungoval jsou minimální teplota objektu, maximální teplota objektu, minimální plocha objektu a případně maximální plocha. Ve výchozím nastavení jsou hodnoty nastaveny na 20°C , 40°C , minimální plocha je nastavena na 200 a maximální hodnota není omezena.

■ 5.2.2 Prahování teplot

Na základě nastavených parametrů teploty se provede prahování, které je možné popsat následujícím matematickým výrazem.

$$\text{Mask}(x, y) = \begin{cases} 1, & \text{když } T_{min} \leq T(x, y) \leq T_{max} \\ 0, & \text{jinak} \end{cases} \quad (5.1)$$

Kde T_{min} je minimální teplota, T_{max} je maximální teplota, $T(x, y)$ je teplota na pozici (x, y) v obrázku. Tímto vznikne maska, která je následně použita v dalším procesu.

■ 5.2.3 Čištění obrazu

Zde je využito binární matematické morfologie, konkrétně se jedná o otevření a zavření. Implementace těchto funkcí je součástí knihovny *openCV*. Jednotlivé funkce jsem aplikoval v následujícím pořadí:

1. otevření s jádrem o velikosti 3x3
 2. zavření s jádrem o velikosti 5x5
- Pythonovský kód této funkce vypadá následovně:

```
def clear_image(self, mask):
    kernel = np.ones((3, 3), np.uint8)
    mask = cv2.morphologyEx(mask, cv2.MORPH_OPEN, kernel)
    kernel = np.ones((5,5), np.uint8)
    mask = cv2.morphologyEx(mask, cv2.MORPH_CLOSE, kernel)
    return mask
```

■ Eroze

I když erozi při čistění přímo nepoužívám je důležité ji zmínit. Základní myšlenka eroze je stejná jako eroze půdy, pouze eroduje hranice objektu (vždy se snaží zachovat bílé oblasti). Obdobně jako při 2D konvoluci, jádro eroze se přiloží na každý pixel obrázku a tento pixel bude považován za 1 pouze v případě, že všechny pixely pod jádrem mají hodnotu 1, v opačném případě se eroduje (vynuluje)[17]. Princip je znázorněn na obrázku 5.1.

■ Dilatace

Je to pravý opak eroze. Často se používá na odstraňování šumu, ale díky této vlastnosti rozšiřuje bílé oblasti. Taktéž se hodí pro spojování bílých oblastí, tak aby vznikl větší celek [17]. Princip fungování je vidět na obrázku 5.1.

■ Otevření a zavření

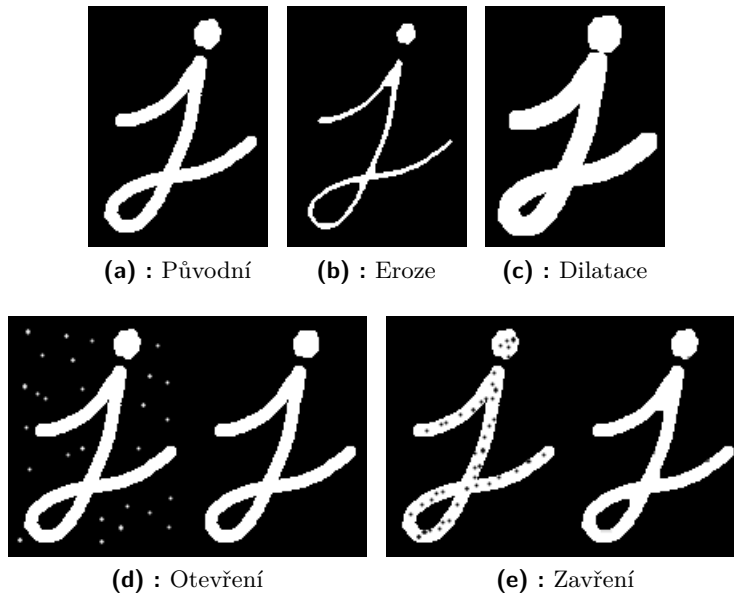
Obecně se dá říci, že otevření je jiný název pro erozi následovanou dilatací a naopak zavření je dilatace následovaná erozí. Obecně se používají na odstranění šumu z binárního obrazu, jak je vidět na obrázku 5.1.

■ 5.2.4 Ohraničení oblastí konturami a filtrování

V této fázi detekce je vstupem již vyčištěná maska, na kterou jsou pomocí *openCV* nalezeny jednotlivé kontury (bílé oblasti). Následně jsou ohraničeny křivkou a spočítána plocha kterou zabírají. Pokud-li je plocha větší než minimální plocha a zároveň menší než maximální dovolená plocha je tato kontura uložena.

■ 5.2.5 Ohraničení kontur bounding boxem

V poslední fázi jsou kontury ohraničeny bounding boxy, které mohou být následně použity na původním obrázku k ohraničení teplých objektů. Další věci



Obrázek 5.1: Obrázky filtrace masky, převzato z [17]

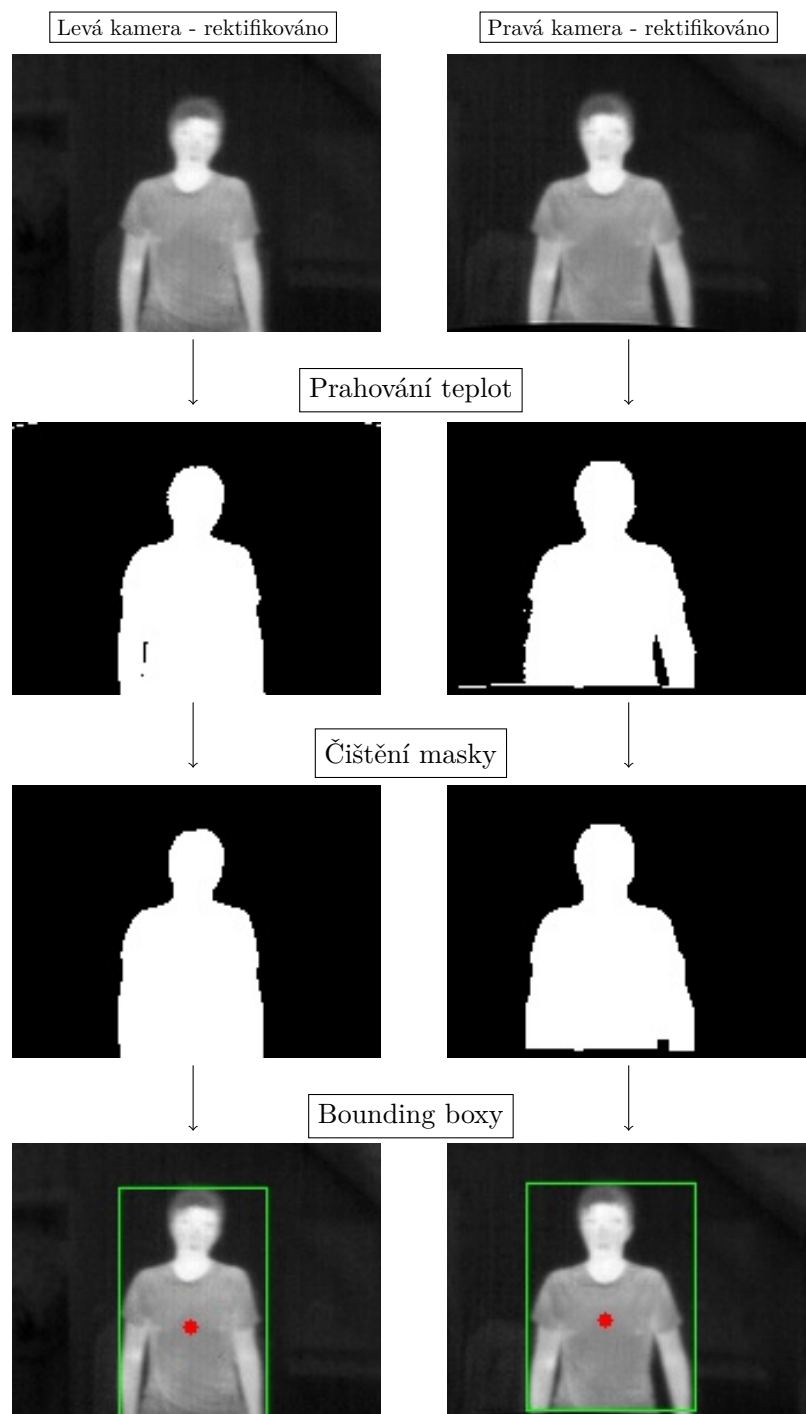
je, že ze znalosti kontur můžeme jednotlivé teplé objekty v levém a pravém obrázku propojit a určit tak korespondenční body, ze kterých můžeme odhadnout jejich vzdálenost. Těmito jsou například těžiště jednotlivých kontur.

5.3 Vlastnosti detektoru

Navržený algoritmus detekce je nezávislý na tvaru objektů, ale jen na jejich teplotě a velikosti. Výhodou tohoto přístupu je malá výpočetní náročnost, která dovoluje i při použití standardního procesoru detekovat objekty v reálné čase na obou snímcích. Obecně je zde velké míra přesnosti, jelikož jsou ohraničeny jen ty oblasti, které přesahují určitou teplotu.

Tento způsob detekce má i jisté nevýhody, které mají různé příčiny. První velkou nevýhodou je více teplých objektů na jednom místě, kde se jejich kontury spojí a celá kontura je označena jako jeden objekt. Další nevýhoda je spojena s vlastností samotné termokamery, jelikož se jedná o nechlazenou kameru, tak samotný snímač se při používání zahřívá a tím mírně ovlivňuje získávané teplotní hodnoty. Samotná kamera má zabudovaný mechanismus na redukci tohoto jevu, ten spočívá v zavření závěrky kamery a upravení vnitřních parametrů podle teploty snímače. Následně je závěrka otevřena, ale po otevření trvá několik sekund než je kamera opět zcela funkční a ukazuje teploty správně. Tento jev se projeví tak, že krátký okamžik posune všechny

teploty o nějakou hodnotu (výše nebo níže). Pokud používáme mapování teplot na barevné spektrum pomocí maximální a minimální teploty, tento jev se v obraze neprojeví.



Obrázek 5.2: Postup detekce dle jednotlivých kroků

5.4 Odhad vzdálenosti teplých objektů

V této sekci porovnám výše uvedený detektor, který jsem použil pro odhad vzdálenosti teplých objektů od kamery. Pro porovnání přesnosti jednotlivých metod bylo použito teplého objektu ve známé vzdálenosti od kamery, který se pohyboval. V tomto případě se jednalo o pohybující ho se člověka ve vzdálenosti 2m a 4m. Osoba se pohybovala po kruhové výseči tak, aby byla celou dobu na snímcích obou kamer. Porovnání jednotlivých metod je ukázáno v grafu 5.3, kde jsou vyneseny naměřené vzdálenosti všemi metodami. Zde je důležité podotknout, že velké výkyvy jsou způsobené časovou prodlevou mezi sejmutím levého a pravého obrázku. V tomto datasetu je snímek pořízen nejdříve levou kamerou a následně pravou. V některých případech se stane, že čtení druhého snímku trvá delší dobu a to vede k časové prodlevě (kdy se osoba pohne), která je nejspíše způsobená samotnou kamerou nebo deskou pureThermal. Tento problém jsem se snažil vyřešit skrze více vláken, ale to často vedlo k větší desynchronizaci obrazu z kamer.

5.4.1 Odhad vzdálenosti při použití algoritmu pro párování rysů

Metoda založena na hledání významných bodů na obou snímcích v obecné rovině funguje velmi dobře. Problémem tohoto přístupu je malé množství nalezených bodů, pro snímky pořízené kamerou pracující ve viditelném spektru, jsou schopny algoritmy pro hledání významných bodů nalézt velké množství těchto bodů. Situace pro teplotní snímky převedené do odstínů šedé je ale horší a množství bodů je razantně menší a v některých případech nejsou detekovány žádné body. Dalším problémem je, že není nikterak zaručeno, že se budou tyto významné body nacházet na člověku. Z tohoto důvodu je nutné kombinovat tuto metodu s detektorem osob, tím může být teplotní (založený na prahování) nebo neuronová síť YOLO. V mém případě bylo využito teplotního detektoru. Následně jsou vybrány jen ty body, které leží uvnitř obdélníkového ohraničení člověka a na základě této informace vypočítána vzdálenost osoby.

Pro nalezení a popis jednotlivých bodů jsem používal algoritmus SIFT, který je součástí knihovny openCV. Pro párování jednotlivých bodů jsem použil hledání nejbližšího souseda (k-NN) s omezením. Toto omezení spočívalo v tom, že body mají stejnou pozici na ose Y na rektifikovaných obrázcích.

5.4.2 Odhad vzdálenosti při použití plošného stereo matching algoritmu

Stejně jako u předchozí metody je nutno nejdříve lokalizovat člověka (nebo libovolný objekt) a následně nalézt v rozdílové mapě tyto hodnoty a vypočítat vzdálenost. Jelikož je rozdílová mapa nespojitá, byly vybrány všechny nenulové hodnoty, které se nacházejí uvnitř ohraničujících obdélníku okolo teplého objektu a následně z nich byl vypočítán průměr. Pro tvorbu rozdílové mapy jsem používal `cv2.StereoBM_create()`, která se řadí do plošných stereo matching algoritmů. Pro co nejlepší výsledek bylo nutno nastavit parametry algoritmu tak, aby bylo možné vůbec detekovat člověka, jelikož výstup byl velmi zašuměný. Nastavení těchto parametrů je k nalezení v tabulce 5.1. Hodnoty pro pohybující objekt v grafu 5.3 odpovídají červené barvě.

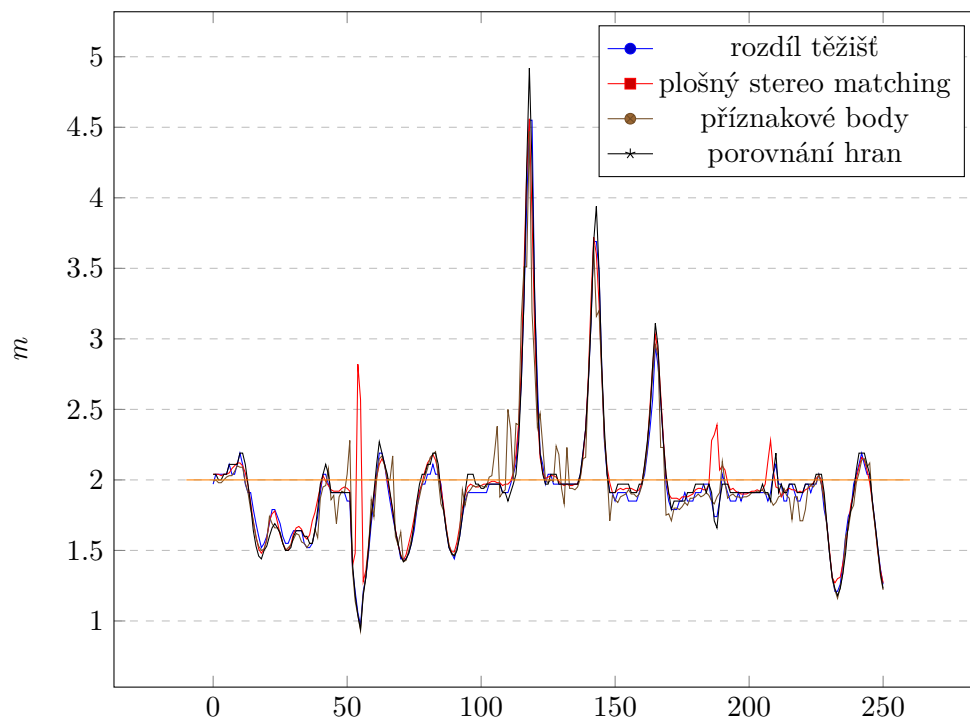
Název	Hodnota
NumDisparities	48
BlockSize	15
UniquenessRatio	10
SpeckleRange	1
SpeckleWindowSize	2
Disp12MaxDiff	10
MinDisparity	1

Tabulka 5.1: Parametry BM algoritmu

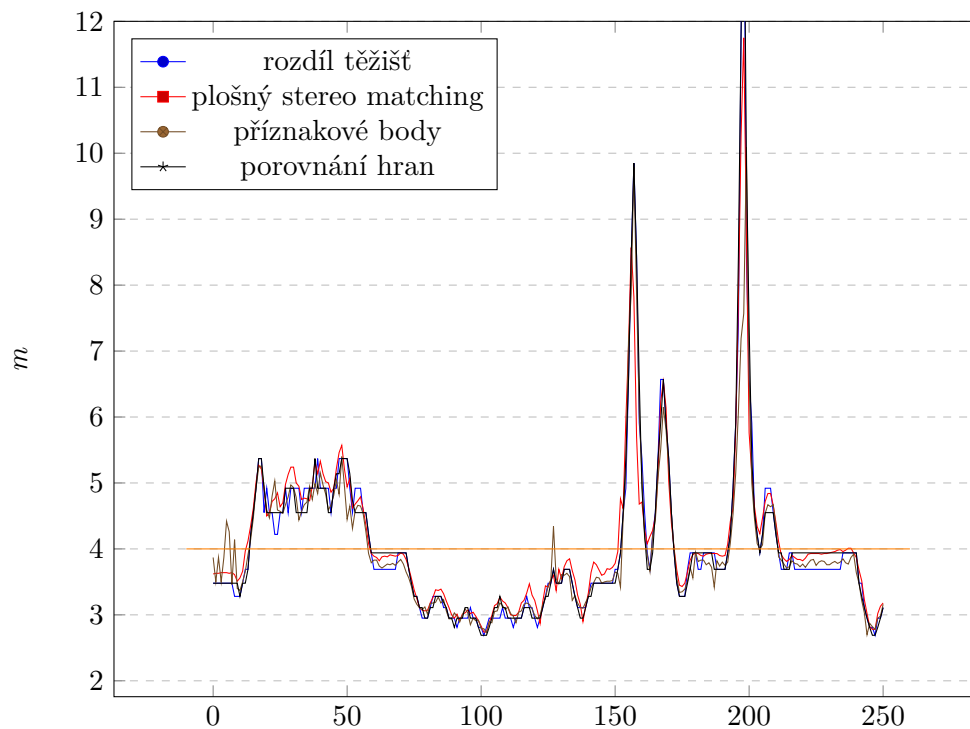
5.4.3 Odhad vzdálenosti při použití informací o konturách

Prvním krokem tohoto detektoru je nalezení odpovídajících kontur v pravém a levém snímku, tento postup byl popsán v sekci 5.2. Následně bylo vypočítáno těžiště jednotlivých kontur a objekty z levého snímku byly spojeny s odpovídajícími tvary z pravého snímku. Byl vypočítán rozdíl těžišť a vypočítána vzdálenost objektů. Hodnoty pro pohybující objekt v grafu 5.3 odpovídají modré barvě.

Dále jsem použil ještě druhý způsob odhadu vzdálenosti založený na informacích o konturách. Zde jsem vzal odpovídající konturu z levého a pravého snímku a spočítal vzdálenost hrany kontury na levé a pravé straně v každém řádku. Následně jsem z každého řádku na kterém se nacházela kontura vypočítal dvě vzdálenosti (levý kraj, pravý kraj) a vypočítal medián za všech řádků. Tuto hodnotu jsem taktéž vynesl do grafu 5.3, kde odpovídá černé barvě.



(a) : 2 metry



(b) : 4 metry

Obrázek 5.3: Porovnání algoritmů pro určování vzdálenosti pohybujícího se člověka

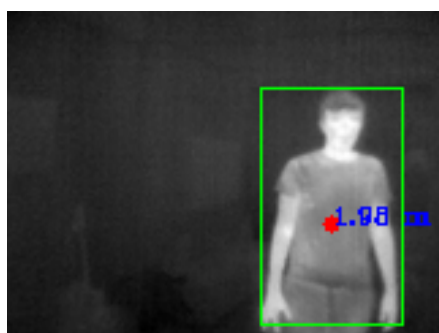
5.5 Zhodnocení jednotlivých algoritmů pro odhad vzdálenosti

Jak už jsem zmínil v sekci 5.4, největší chybu generuje desynchronizace snímků mezi kamerami. Chyba při pohybu objektu ve vzdálenosti dva metry může dosahovat více jak 2 metrů. Ve výše zmíněném grafu 5.3, konkrétně mezi snímky 100 až 170 se osoba pohybovala zprava doleva, a tím pádem vzniká tato chyba. Pro objekty ve vzdálenosti 4 metry způsobuje tento jev daleko větší chybu, konkrétně je tato chyba vidět mezi snímky 150 až 175 a okolo snímku 200.

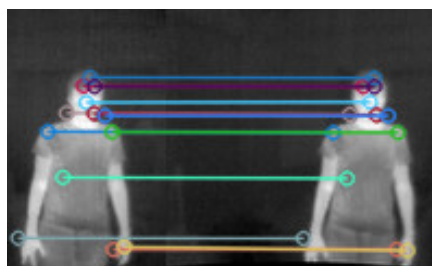
Naopak se dá konstatovat, že všechny výše navržené algoritmy dosahují podobných výsledků a není mezi nimi velký rozdíl. Toto tvrzení podkládají grafy v 5.3, kde je vidět vysoká korelace mezi metodami. Obecně se dá také konstatovat, že na statických scénách tato stereo soustava funguje dobře a dosahuje dobrých výsledků. Vizualizace jednotlivých metod je zobrazena na obrázku 5.4. Porovnání náročnosti výpočtu a jejich doba trvání je k nalezení v tabulce 5.2, měření bylo provedeno na procesoru Intel® Core™ i5-8265U. Vzhledem k tomu, že maximální obnovovací frekvence kamer je menší než 9 Hz je možné použít všechny metody bez omezení.

	Rozdíl těžišť	Plošná metoda	Příznakové body	Porovnání hran
Doba výpočtu jednoho páru snímků	0.005 s	0.006 s	0.067 s	0.01 s
FPS	200	166.67	14.93	100

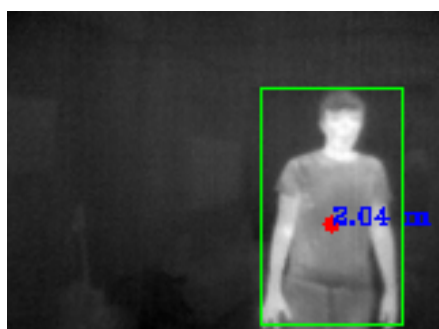
Tabulka 5.2: Porovnání rychlosti jednotlivých metod



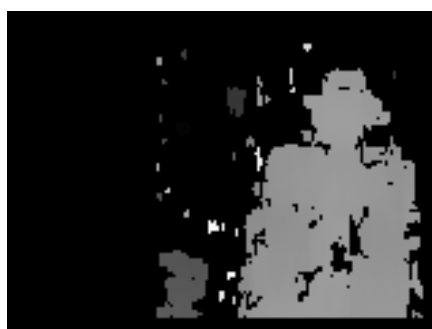
(a) : detekce teplého objektu (levá kamera)



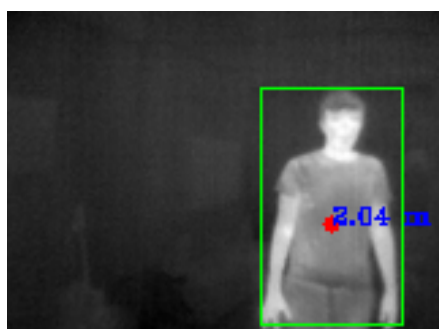
(b) : stereo páry bodů



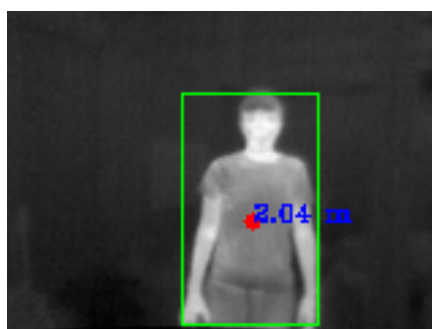
(c) : detekce teplého objektu (levá kamera)



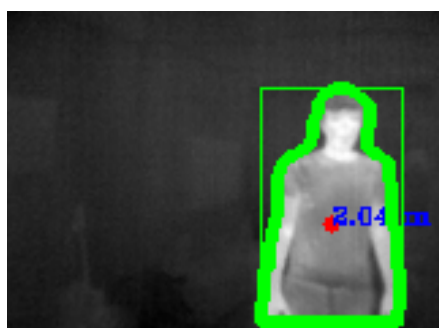
(d) : disperzní mapa



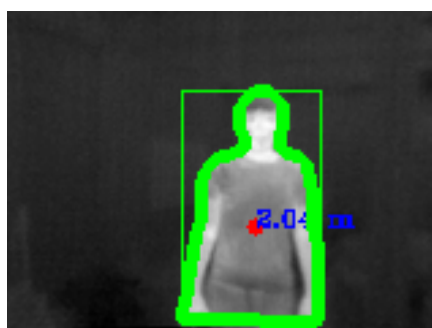
(e) : detekce teplého objektu (levá kamera)



(f) : detekce teplého objektu (pravá kamera)



(g) : detekce kontur (levá kamera)



(h) : detekce kontur (pravá kamera)

Obrázek 5.4: Vizualizace jednotlivých metod použitých pro odhad vzdáleností (v řádce vždy jedna metoda).

Kapitola 6

Závěr

V této práci jsem se seznámil s termokamerou FLIR Lepton 3.5 a následně využil znalosti o této kameře ke tvorbě stereo termokamery, která je složena ze dvou těchto kamer. Výstupem této kamery je možnost odhadu vzdálenosti teplých objektů od základny stereo soustavy.

Pro práci s obrazem jsem využíval hlavně nástroje z knihovny OpenCV. Pro přesný odhad vzdálenosti objektů je důležité provést kalibraci jednotlivých kamer a taktéž kalibraci celé stereo soustavy. Jelikož termokamera generuje obraz na základě teploty, musel být vzor používaný pro kalibraci upraven tak, aby byl viditelný na termokameře. Bylo zvoleno, že vzor bude zahříván pasivně pomocí slunečního světla. Jako kalibrační vzor byl zvolen vzor na základě šachovnice o velikosti 5x6 polí. Každé pole mělo velikost 7 cm a všechny původně bílé části byly potaženy hliníkovou fólií, která má velmi vysokou odrazivost a z tohoto důvodu, zde vzniká dostatečně velký teplotní rozdíl, aby mohl být vzor použit pro kalibraci. Zároveň lze stejný vzor použít i pro kalibrace kamer pracujících ve viditelném spektru. Mnou používané termokamery mají poměrně malé rozlišení 160x120 pixelů, které působí problémy při detekci šachovnice a následné kalibraci. Tento problém jsem vyřešil zvětšením výstupu z kamery na čtyřnásobek a následně jsem přepočítal pozice detekovaných rohů do původních souřadnic. Jak zvětšování tak samotná kamera vnáší do obrazu šum, jeho vliv na kalibraci jsem se snažil minimalizovat použitím velkého množství snímků. Průměrná reprojekční chyba kalibrací obou kamer byla menší než 1 px.

Pro detekci teplých objektů nebo oblastí bylo zvoleno teplotní prahování, kde byly vybrány jen ty oblasti které přesahovali určitou teplotu. Tento přístup

umožňuje provádět detekci v reálném čase s malou výpočetní náročností. Výhodou oproti jiným detektorům, kde je vstupem obraz, je absence trénovacích dat. Jelikož obecně nedokážeme definovat, který objekt může být teplý či nikoliv, běžně se detektory učí detekovat odlišné předměty/třídy. Naopak detektor zaostává v situacích, kde se teplé objekty jen dotýkají nebo překrývají. V takovém případě mnou navržený detektor označí oba objekty jako jeden celek.

Pro detekci specifických objektů, jako jsou třeba lidé, je lepší použít jiný detektor například YOLO nebo Viola-Jones detektor. Výhodou tohoto přístupu je, že existuje obsáhlý dataset termálních dat, vytvořený přímo společností FLIR, který je možné použít k učení odlišných detektorů objektů.

Co se týče odhadu vzdálenosti objektů od základny kamery, porovnal jsem odlišné metody pro tento odhad. První metodou bylo použití plošného matchování, druhá metoda využívala body nalezené algoritmem SIFT a následného párování těchto bodů a poslední dvě metody využívaly informace o jednotlivých konturách. Lze konstatovat, že v jednotlivých metodách nejsou velké rozdíly a odhad vzdálenosti funguje poměrně přesně. Naopak největší problémy působí čtení dat z kamer, které způsobuje velké skokové chyby při odhadu vzdáleností. Jádrem problému je v tom, že jsou data načtena nejprve z levé kamery a následně z pravé. Skoková chyba se objevuje v době, kdy čtení z jedné kamery trvá lehce delší dobu než z první, a zároveň se snímáný objekt pohybuje. Tento nedostatek jsem se snažil vyřešit vícevláknovým přístupem, ale to vedlo k větším a častějším skokovým chybám. Vše nasvědčuje tomu, že desynchronizace je způsobena senzorem a způsobem jeho činnosti.

Obecně je možné tuto stereo soustavu použít pro odhad vzdálenosti do 4 m, jelikož ve větší vzdálenosti velice roste vliv chyby na odhad vzdálenosti. Toto je způsobeno malým rozlišením kamer, kde již chyba jeden pixel, při určování korespondenčních bodů, působí velké nepřesnosti.



Literatura

- [1] KARLSEN, Robert E., Douglas W. GAGE, Charles M. SHOEMAKER, Grant R. GERHART, Josh HARGUESS a Shawn STRANGE. *Infrared stereo calibration for unmanned ground vehicle navigation* [online]. In: . 2014-6-3, 90840S- [cit. 2022-01-06]. Dostupné z: doi:10.1117/12.2053537
- [2] LIU, Ruixuan, Hengrui ZHANG a Sebastian SCHERER. *Multiple Methods of Geometric Calibration of Thermal Camera and A Method of Extracting Thermal Calibration Feature Points* [online]. [cit. 2021-11-28]. Dostupné z: <https://henryzh47.github.io/assets/documents/multiple-methods-geometric.pdf>
- [3] ZOETGNANDÉ, Yannick, Alain-Jérôme FOUGÈRES, Geoffroy CORMIER, Jean-Louis DILLENSEGER, Dmitry P. NIKOLAEV, Petia RADEVA, Antanas VERIKAS a Jianhong ZHOU. *Robust low resolution thermal stereo camera calibration*. In: Eleventh International Conference on Machine Vision (ICMV 2018) [online]. SPIE, 2018, 2019-3-15, s. 115- [cit. 2021-10-12]. ISBN 9781510627482. Dostupné z: doi:10.1117/12.2523440
- [4] HATA, Kenji a Silvio SAVARESE. *CS231A Course Notes 1: Camera Models* [online]. s. 1-5 [cit. 2021-11-25]. Dostupné z: https://web.stanford.edu/class/cs231a/course_notes/01-camera-models.pdf
- [5] What Is Camera Calibration? [online]. [cit. 2021-11-25]. Dostupné z: <https://www.mathworks.com/help/vision/ug/camera-calibration.html>
- [6] YU, Zhang, Shen LINCHENG, Zhou DIANLE, Zhang DAIBING a Yan CHENGPING. *Camera Calibration of Thermal-Infrared Stereo Vision System*. In: 2013 Fourth International Conference on Intelligent

- Systems Design and Engineering Applications [online]. IEEE, 2013, 2013, s. 197-201 [cit. 2021-10-07]. ISBN 978-1-4799-2792-0. Dostupné z: doi:10.1109/ISDEA.2013.449
- [7] MOUATS, T., N. AOUF, L. CHERMAK a M. A. RICHARDSON. *Thermal Stereo Odometry for UAVs*. IEEE Sensors Journal. 2015, 15(11), 6335-6347. ISSN 1558-1748. Dostupné z: doi:10.1109/JSEN.2015.2456337
- [8] FLIR LEPTON® Engineering Datasheet [online]. s. 34-39 [cit. 2021-11-25]. Dostupné z: <https://www.flir.com/globalassets/imported-assets/document/flir-lepton-engineering-datasheet.pdf>
- [9] Depth Map from Stereo Images. *OpenCV* [online]. [cit. 2022-01-07]. Dostupné z: https://docs.opencv.org/4.x/dd/d53/tutorial_py_depthmap.html
- [10] Camera Calibration. *OpenCV* [online]. [cit. 2022-01-07]. Dostupné z: https://docs.opencv.org/3.4/dc/dbb/tutorial_py_calibration.html
- [11] PRINCE, Simon J. D. *Computer vision: models, learning, and inference*. New York: Cambridge University Press, 2012. ISBN 978-1-107-01179-3.
- [12] Understanding Lens Distortion. *LearnOpenCV* [online]. [cit. 2022-04-25]. Dostupné z: <https://learnopencv.com/understanding-lens-distortion/>
- [13] Bradski, G.; Kaehler, A.: *Learning OpenCV: Computer Vision with the OpenCV Library*. O'Reilly Media, 2008, iISBN 978-0-596-51613-0.
- [14] Epipolar Geometry. *OpenCV* [online]. [cit. 2022-04-25]. Dostupné z: https://docs.opencv.org/3.4/da/de9/tutorial_py_epipolar_geometry.html
- [15] Homografie A Epipolární Geometrie. *TRILOBIT: odborný vědecký časopis* [online]. [cit. 2022-04-25]. Dostupné z: <http://trilobit.fai.utb.cz/homografie-a-epipolarni-geometrie>
- [16] BERGHMANS, Francis, Sarun SUMRIDDETCHKAJORN, Kosom CHAITAVON, Anna G. MIGNANI, Antonello CUTOLO, Patrick P. MEYRUEIS a Thomas P. PEARSALL. *A noninvasive human temperature screening system with multiple detection points* [online]. 2008-4-25 [cit. 2022-04-05]. Dostupné z: doi:10.1117/12.779147
- [17] Morphological Transformations. *OpenCV* [online]. [cit. 2022-04-25]. Dostupné z: https://docs.opencv.org/3.4/d9/d61/tutorial_py_morphological_ops.html
- [18] HAMZAH, Rostam Affendi, Afifah Maheran Abdul HAMID a Sani Irwan Md SALIM. The Solution of Stereo Correspondence Problem Using Block Matching Algorithm in Stereo Vision Mobile Robot. In: *2010 Second International Conference on Computer Research and Development* [online]. IEEE, 2010, 2010, s. 733-737 [cit. 2022-04-26]. ISBN 978-0-7695-4043-6. Dostupné z: doi:10.1109/ICCRD.2010.167

- [19] Feature Matching. *OpenCV* [online]. [cit. 2022-04-26]. Dostupné z: https://docs.opencv.org/4.x/dc/dc3/tutorial_py_matcher.html
- [20] IVAŠIĆ-KOS, Marina, Mate KRIŠTO a Miran POBAR. Human Detection in Thermal Imaging Using YOLO. In: *Proceedings of the 2019 5th International Conference on Computer and Technology Applications* [online]. New York, NY, USA: ACM, 2019, 2019-04-16, s. 20-24 [cit. 2022-04-26]. ISBN 9781450371810. Dostupné z: doi:10.1145/3323933.3324076
- [21] SETJO, Christian Herdianto, Balza ACHMAD a FARIDAH. Thermal image human detection using Haar-cascade classifier. In: *2017 7th International Annual Engineering Seminar (InAES)* [online]. IEEE, 2017, 2017, s. 1-6 [cit. 2022-04-26]. ISBN 978-1-5386-3111-9. Dostupné z: doi:10.1109/INAES.2017.8068554
- [22] CHARVÁT, Michal. *System for People Detection and Localization Using Thermal Imaging Cameras* [online]. Brno, 2020 [cit. 2022-05-15]. Dostupné z: <http://hdl.handle.net/11012/194978>. Diplomová práce. Vysoké učení technické v Brně. Fakulta informačních technologií. Ústav inteligentních systémů.