



ZADÁNÍ BAKALÁ SKÉ PRÁCE

Název:	Detekce zboží v ruce zákazníka pomocí analýzy snímk z termokamery
Student:	Lukáš Brchl
Vedoucí:	doc. RNDr. Ing. Marcel Ji ina, Ph.D.
Studijní program:	Informatika
Studijní obor:	Teoretická informatika
Katedra:	Katedra teoretické informatiky
Platnost zadání:	Do konce zimního semestru 2018/19

Pokyny pro vypracování

Cílem práce je navrhnut a implementovat algoritmus, který bude detektovat, zda má zákazník v ruce zboží nebo ne. Využita bude termokamera umíst ná nad regálem, kterou je snímán obraz chování zákazníka p ed ním. Termovizní snímek obsahuje informaci o teplot snímaných p edm t . To by m l být rozlišovací faktor mezi chladn jším zbožím a teplejší rukou. Spolu s termokamerou je ze stejného místa snímán i standardní barevný snímek a tzv. hloubková mapa (mapa vzdálenost objekt od kamery).

- 1) Seznamte se s úlohou detekce zboží v ruce zákazníka u regálu. Prove te rešerši stávajících metod pro zpracování snímk z termokamery, které souvisí s ešenou úlohou.
- 2) Navrhn te algoritmus, který umožní detektovat p ítomnost zboží v ruce zákazníka.
- 3) Navržený algoritmus implementujte v Jav a využijte knihovnu OpenCV.
- 4) Navržený a implementovaný algoritmus ov te na reálných datech, která nam íte.
- 5) Dosažené výsledky vyhodno te a diskutujte výhody a nevýhody použitého p ístupu.

Seznam odborné literatury

Dodá vedoucí práce.

doc. Ing. Jan Janoušek, Ph.D.
vedoucí katedry

prof. Ing. Pavel Tvrďk, CSc.
d kan

ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE
FAKULTA INFORMAČNÍCH TECHNOLOGIÍ
KATEDRA TEORETICKÉ INFORMATIKY



Bakalářská práce

**Detekce zboží v ruce zákazníka pomocí
analýzy snímků z termokamery**

Lukáš Brchl

Vedoucí práce: doc. RNDr. Ing. Marcel Jiřina, Ph.D.

16. května 2017

Poděkování

Tímto bych chtěl hlavně poděkovat svému vedoucímu práce doc. Marcelu Jiřinovi za všechny poskytnuté rady a možnost průběžných konzultací během práce. Dále pak Ing. Jakubovi Novákovi za pomoc při řešení nesnází s termokamerou.

Děkuji i své rodině za podporu při studiu a v neposlední řadě také děkuji Michaele Dvořákové, Ing. Iljovi Mašíkovi a Ing. Michaele Čížkové za typografické připomínky.

Prohlášení

Prohlašuji, že jsem předloženou práci vypracoval samostatně a že jsem uvedl(a) veškeré použité informační zdroje v souladu s Metodickým pokynem o etické přípravě vysokoškolských závěrečných prací.

Beru na vědomí, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorského zákona, ve znění pozdějších předpisů. V souladu s ust. § 46 odst. 6 tohoto zákona tímto uděluji nevýhradní oprávnění (licenci) k užití této mojí práce, a to včetně všech počítačových programů, jež jsou její součástí či přílohou a veškeré jejich dokumentace (dále souhrnně jen „Dílo“), a to všem osobám, které si přejí Dílo užít. Tyto osoby jsou oprávněny Dílo užít jakýmkoli způsobem, který nesnižuje hodnotu Díla a za jakýmkoli účelem (včetně užití k výdělečným účelům). Toto oprávnění je časově, teritoriálně i množstevně neomezené.

V Praze dne 16. května 2017

.....

České vysoké učení technické v Praze
Fakulta informačních technologií
© 2017 Lukáš Brchl. Všechna práva vyhrazena.

Tato práce vznikla jako školní dílo na Českém vysokém učení technickém v Praze, Fakultě informačních technologií. Práce je chráněna právními předpisy a mezinárodními úmluvami o právu autorském a právech souvisejících s právem autorským. K jejímu užití, s výjimkou bezúplatných zákonných licencí, je nezbytný souhlas autora.

Odkaz na tuto práci

Brchl, Lukáš. *Detekce zboží v ruce zákazníka pomocí analýzy snímků z termokamery*. Bakalářská práce. Praha: České vysoké učení technické v Praze, Fakulta informačních technologií, 2017. Dostupný také z WWW: <https://github.com/brchlluk/BachelorThesis>.

Abstrakt

Tato práce se zabývá detekcí zboží v ruce zákazníka pomocí analýzy snímků z termokamery. Kamera je umístěna nad regálem, kde snímá oblast bezprostředního dění před ním. Snímaná data z kamery obsahují informace o teplotách jednotlivých předmětů na scéně. Pomocí teplotních dat je následně možné na snímcích segmentovat ruku a zkoumat, zda se na scéně vyskytuje zboží. Cílem práce je průzkum možných řešení této úlohy, návrh vlastního řešení a následně jeho implementace. V práci je k problému přistupováno více způsoby a je navrženo několik možných řešení. Pro snímání dat je použita termokamera FLIR A65 se sadou nástrojů eBUS SDK. Algoritmy jsou implementovány v jazyce Java s využitím knihovny OpenCV. Vyhodnocený algoritmus dynamického odčítání pozadí dosahuje celkové klasifikační přesnosti 88 %.

Klíčová slova segmentace obrazu, segmentace ruky, klasifikace, počítačové vidění, strojové učení, termokamera, FLIR, OpenCV, Java

Abstract

This thesis deals with the detection of goods in the customer's hand by analyzing images from infrared camera. The camera is located above the shelf, monitoring behavior of the customer standing in front of it. Acquired data contains informations about the temperature of the objects in the view. Using these temperature data, it is possible to segment hand and examine the presence of the goods on the scene. The aim of the thesis is to explore possible solutions of this task, to propose a solution and subsequently implement it. There are several approaches to the problem and several possible solutions are proposed. The FLIR A65 thermal camera with the eBUS SDK tool is used to capture data. The algorithms are implemented in Java using the OpenCV library. The evaluated background subtraction algorithm achieves an overall classification accuracy of 88 %.

Keywords image segmentation, hand segmentation, classification, computer vision, machine learning, infrared camera, FLIR, OpenCV, Java

Obsah

Úvod	1
Motivace	2
Cíl práce	2
Struktura dokumentu	3
Návaznost práce	3
1 Termokamera	5
1.1 Co je to termokamera	5
1.2 Princip snímání	5
1.3 Konstrukce	8
1.4 Specifikace a parametry	11
1.5 Získávání dat	11
2 Analýza	17
2.1 Současný stav řešení problematiky	17
2.2 Související práce s termokamerou	17
2.3 Analýza zadaného problému	19
2.4 Nutné předpoklady	19
2.5 Výhody využití termokamery pro tuto úlohu	20
2.6 Překážky při snímkování s termokamerou	21
2.7 Dostupné prostředky	23
3 Průzkum možných řešení	27
3.1 Získávání dat z kamery	27
3.2 Předzpracování dat	31
3.3 Řešení segmentací na základě teplot objektů	34
3.4 Řešení s využitím hranové detekce	35
3.5 Řešení pomocí dynamického odčítání pozadí	38
3.6 Řešení s kombinací RGB obrazu	43

4 Zvolené řešení a jeho vyhodnocení	49
4.1 Shrnutí vyhodnocovaného řešení	49
4.2 Volba atributů	49
4.3 Rozdělení snímků do tříd	51
4.4 Klasifikace	51
4.5 Výsledky	51
4.6 Příklady klasifikovaných snímků	52
4.7 Komentář ke klasifikovaným snímkům	52
5 Implementované nástroje	55
5.1 Upravený eBUS Player	55
5.2 Aplikace s detekčními algoritmy	55
5.3 Aplikace pro anotování snímků	61
6 Zhodnocení	63
6.1 Využitá technika	63
6.2 Navržené řešení	63
6.3 Implementace řešení	64
6.4 Výsledky řešení	64
6.5 Budoucí využití	65
Závěr	67
Literatura	69
A Seznam použitých zkratek	75
B Obsah přiloženého média	77

Seznam obrázků

0.1	Na obrázku (a) je přídavný modul pro chytré telefony a na (b) telefon s již zabudovanou termokamerou	2
1.1	Ukázkový snímek pořízený termokamerou [1]	6
1.2	Elektromagnetické spektrum s vyznačenými pásmi infračerveného záření [2]	7
1.3	Konceptuální černé těleso	8
1.4	Jednoduché schéma hlavních konstrukčních částí termokamery [3] .	9
1.5	Vlevo je snímek bez aplikovaných korekcí signálu a vpravo je snímek po aplikaci vhodných algoritmů zpracování signálu v příslušných obvodech [4]	9
1.6	Schéma bolometru [5]	10
1.7	Koncept GenICam protokolu [6]	13
2.1	Ukázkový snímek optických vad u termokamery FLIR A65 s ohniskovou vzdáleností 25 mm	22
2.2	Termokamera FLIR A65	24
3.1	Neupravený snímek v plné velikost	29
3.2	Výřez oblasti zájmu neupraveného snímku	30
3.3	Schéma umístění kamery	31
3.4	Fotografie umístění kamery	32
3.5	Jednotlivé kroky předzpracování snímku	33
3.6	Vývojový diagram algoritmu s využitím hranové detekce	36
3.7	Ilustrace operací získání popředí snímků a segmentace ruky	39
3.8	Proces zpracování snímku zboží	40
3.9	Vývojový diagram algoritmu dynamického odečítání pozadí	41
3.10	Chybě získané popředí kvůli změně expozice aktuálního snímku .	42
3.11	Ukázkový snímek snímaného obrazového pole webkamery	44
3.12	Snímek z webkamery oříznutý k obrazovému poli termokamery .	45
3.13	Ručně sloučené snímky z webkamery a termokamery	46

3.14	Vývojový diagram algoritmu využívající kombinace běžného a termografického snímku	47
4.1	Příklady správně klasifikovaných snímků kategorie prázdná ruka	53
4.2	Příklady klasifikovaných snímků jako ruka se zbožím, ale ve skutečnosti prázdná ruka	53
4.3	Příklady správně klasifikovaných snímků kategorie ruka se zbožím	54
4.4	Příklady klasifikovaných snímků jako prázdná ruka, ale ve skutečnosti ruka se zbožím	54
5.1	Diagram tříd detekčních algoritmů	57
5.2	Diagram tříd zdrojů snímků	58
5.3	Diagram tříd nastavení parametrů algoritmů	59
5.4	Diagram tříd balíčku image	59
5.5	Aplikace s detekčními algoritmy	60
5.6	Aplikace pro anotování snímků	61

Seznam tabulek

1.1	Obecné rozdělení oblastí infračerveného záření	8
1.2	Klíčové parametry při volbě termokamery.	12
2.1	Klíčové specifikace kamer použitých v této práci	24
3.1	Hodnoty registrů pro výběr oblasti zájmu	30
4.1	Tabulka pozitivní prediktivní hodnoty jednotlivých tříd (class precision) a celkové přesnosti (accuracy) jednotlivých klasifikátorů . .	52
4.2	Matice záměn neuronové sítě vyhodnocena pomocí 10-násobné křížové validace	52

Úvod

Rozpoznávání obrazu je bezesporu jedno z aktuálně nejhojněji diskutovaných a rychle se rozvíjejících témat posledních let. Prakticky denně lze sledovat nové metody a technologie tohoto oboru, které nám usnadňují práci a celkově mění náš život. Příkladem mohou být nově vznikající autonomní vozidla, která ke svému řízení využívají sadu senzorů a kamer, a mají za úkol nás co nejbezpečněji dopravit na potřebné místo. Jiným příkladem mohou být rozsáhlé neuronové sítě, které na základě vstupního obrázku zvládnou detektovat předměty, osoby a dokonce popsat činnosti, které se na snímku odehrávají. Zkrátka, oboru rozpoznávání obrazu se aktuálně věnuje velké množství lidí po celém světě a vzhledem k tempu, jakým se všechny technologie vyvíjejí, je možné brzy očekávat převratné novinky s užitkem pro lidstvo.

Další velmi zajímavou oblastí, která se v posledních letech začíná značně rozšiřovat a přibližovat veřejnosti, jsou termokamery. Pomáhají nám při záchráně osob, v průmyslu (například při monitorování nebezpečných látek), ale dokonce už i při běžných činnostech jako je hledání chyb v tepelné izolaci našich domů.

Oblast termografie má za sebou již dlouhou řadu let výzkumu a vývoje. Zatímco infračervené záření bylo objeveno již kolem roku 1800, první prototypy kamer, které by infračervené záření zvládli zaznamenat, byly vynalezeny až kolem roku 1930. Od této doby technologie termokamer prošla několika generacemi, ale je zřejmé, že možnosti termokamer ještě nejsou zdaleka vyčerpány. Stále se objevují nové inovace a příkladem lze zmínit firmu FLIR Systems (dále jen FLIR), která je jedním z největších výrobců termokamer na celém světě. FLIR postupně rozšiřuje své portfolio termokamer i do neobvyklých oborů, jako jsou již zmíněná autonomní vozidla, ale snaží se zpřístupnit kamery i pro širokou veřejnost. Jejich termografické moduly pro chytré telefony je možné pořídit za velmi lidovou cenu a jsou dostupné skoro pro každého.

O tom, že lidé o obor termografie mají stále větší zájem svědčí i to, že firma FLIR v roce 2014 představila svoji první verzi modulu pro chytré telefony a v roce 2017 je už dostupná jeho třetí generace. V roce 2016 byl dokonce



(a) FLIR One 3rd gen.

(b) CAT S60

Obrázek 0.1: Na obrázku (a) je přídavný modul pro chytré telefony a na (b) telefon s již zabudovanou termokamerou

představen i samostatný chytrý telefon s již zabudovanou termokamerou. Je tedy zřejmé, že termokamery se velmi rychle rozšiřují do podvědomí veřejnosti a postupně se dále budou stávat již běžnou součástí dennodenních zařízení.

Motivace

Využití těchto nově rozvíjejících technologií by pro toto téma mělo přinést nové zajímavé objevy a poznatky k ještě neúplně probádané oblasti.

Tématem detekce zboží v ruce se již zabýval student Oliver Keruľ-Kmec ve své bakalářské práci [7], ale bylo zjištěno, že se jedná o poměrně náročnou úlohu. Jeho práce přináší řešení a poznatky, ale obecně je možné tvrdit, že problém ještě nebyl doposud uspokojivě vyřešen. Má tedy smysl se tématem dále zabývat a to zejména na jiné úrovni obrazu než bylo řešeno, a tím je termografický obraz.

Cíl práce

Cílem práce je seznámení s úlohou detekce zboží v ruce zákazníka, prozkoumání souvisejících řešení úloh v oblasti termokamer. Dále návrh vlastního algoritmu, který bude zboží v ruce zákazníka schopný detektovat. Algoritmus bude implementován v programovacím jazyku Java s využitím volně dostupné knihovny OpenCV. Pro ověření funkčnosti algoritmu je nutné jeho vyhodnocení na vlastnoručně naměřených datech. Na základě dosažených výsledků bude provedena diskuse nad výhodami a nevýhodami využititého řešení.

Struktura dokumentu

V první řadě je nutné obecné seznámení s termokamerami a oblastí termografie. K tomu slouží první kapitola, kde jsou stručně probrány důležité základy pro pochopení práce. Následně je provedena analýza zadaného problému, zavedení nutných předpokladů a představení prostředků práce. Nadcházející kapitola se pak věnuje průzkumu možných řešení a popisu navržených postupů. Ve čtvrté kapitole jsou vyhodnoceny dosažené výsledky a pátá kapitola je věnována implementačním záležitostem. V poslední části je celkové zhodnocení využitých postupů.

Návaznost práce

Tato práce nepřímo navazuje na práci studenta Olivera Keruľ-Kmece *Detekce přítomnosti zboží v ruce zákazníka* [7].

Termokamera

Pro pochopení řešení této práce je nutné se seznámit s technologií termokamer a dalšími důležitými pojmy z termografického oboru. Technologie a způsob použití je proti běžným kamerám rozdílná a má tedy smysl se jejím popisem zabývat.

Oblast termografie je sama o sobě velmi náročná a mohla by být samostatným tématem rozsáhlé vědecké práce. Kapitola se věnuje pouze základním principům a pro detailnější pochopení této oblasti jsou doporučeny práce [8, 9, 10].

1.1 Co je to termokamera

Termokamera je zařízení, které nám umožňuje zobrazit a zaznamenat námi jinak neviditelné infračervené záření. Podobným způsobem jako běžné kamery zaznamenávají viditelné spektrum.

1.2 Princip snímání

Snímání termokamery spočívá v bezdotykovém měření povrchové teploty těles. Tento způsob využívá skutečnosti, že každé těleso, jehož teplota je vyšší než absolutní nula, vyzařuje tepelné záření [11, 2]. Jak se těleso stává teplejší, ve vnitřní struktuře se zrychluje pohyb atomů a molekul, a tím se tvorí více energie, kterou je potřeba vyzářit. Snímač termokamery (detektor) infračervené spektrum zaznamenává formou elektronických signálů, který jsou následně digitalizovány a převedeny na termografický snímek.

Technologicky není možné konstruovat snímače termokamery tak, aby zachytily celý rozsah infračerveného spektra. Proto jsou vždy konstruovány pro konkrétní infračervenou oblast. V dnešní době je to nejčastěji oblast LWIR (více o oblastech v 1.2.3) a to díky nižší ceně technologie a velmi dobrým výsledkům.



Obrázek 1.1: Ukázkový snímek pořízený termokamerou [1]

1.2.1 Absolutní nula

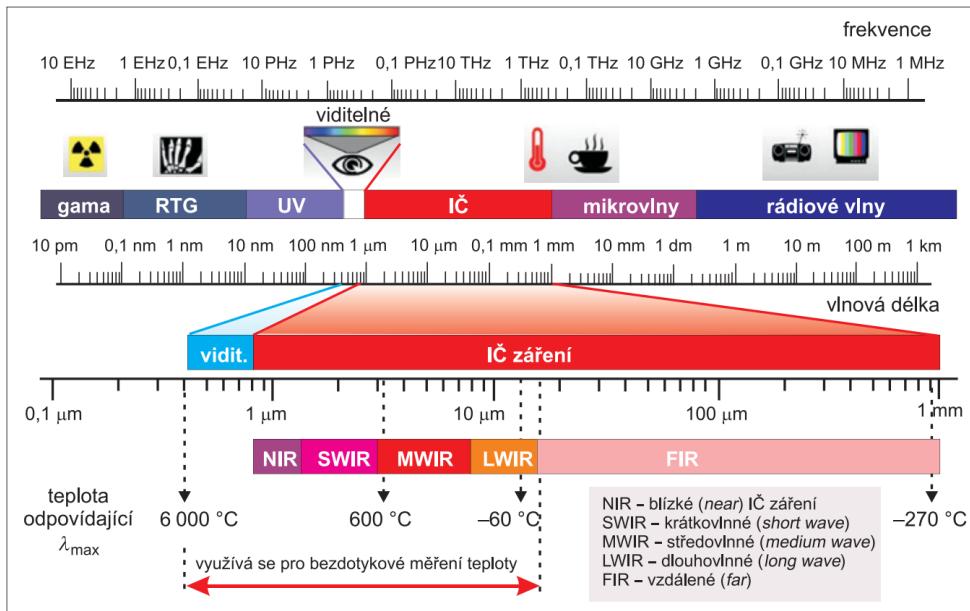
Absolutní nula je teoretický stav tělesa, při kterém se zastaví veškerý pohyb částic a z tělesa již nelze odebrát žádné další teplo. Hodnota absolutní nuly je 0 kelvinů neboli $-273,15^{\circ}\text{C}$. Zatím neexistuje známý způsob, jak 0 K dosáhnout. Jsou pouze známy způsoby a byly provedeny experimenty, které se k dosáhnutí hodnoty absolutní nuly limitně přiblížují.

1.2.2 Elektromagnetické záření

Elektromagnetické záření je příčné postupné vlnění tvořené elektrickým a magnetickým polem. Lidé nejčastěji vnímají záření nazývané viditelné spektrum neboli světlo, které nám umožňuje vidět. Kromě viditelného spektra bylo postupem času objevováno více druhů záření, které bylo nutné rozdělit do skupin. Tyto skupiny tvoří takzvané elektromagnetické spektrum, kde jsou jednotlivá záření popsána jejich vlnovými délками (respektive frekvencemi). Přechody mezi skupinami jsou plynulé a některé se částečně překrývají.

Záření, které lidskému oku umožňuje vidět se nazývá viditelné spektrum. Viditelné spektrum lze nalézt mezi UV a IR zářením, a to pouze ve velmi úzkém rozsahu vlnových délek od 400 nm do 800 nm. Záření mimo viditelné spektrum je pro lidstvo nejen neviditelné, ale lidé jej nedokáží ani vnímat. Přesto je možné se zářením různých druhů dennodenně setkávat. Příkladem je ultrafialové záření, jehož zdrojem je slunce a lidský organismus ovlivňuje tvorbou kožních skvrn a přispívá na tvorbě vitamínu D. Na rozdíl od lidí však existují živočichové, kteří jsou schopni vnímat nejen viditelné spektrum, ale také UV a IR spektrum. [12, 2]

1.2. Princip snímání



Obrázek 1.2: Elektromagnetické spektrum s vyznačenými pásmi infračerveného záření [2]

1.2.3 Infračervené záření

Infračervené záření je pouze jedna ze skupin patřící do celku elektromagnetických záření. Jeho vlnová délka je přibližně od 760 nm do 1 mm a nachází se mezi viditelným spektrem a mikrovlnami. [13]

Vlnová délka IR záření je závislá na teplotě tělesa, které jej vyzařuje. Obecně platí, že čím je teplota tělesa vyšší, tím se vlnová délka zkracuje [2]. IR záření se dále dělí na jednotlivá pásma, která však nejsou striktně daná a existuje více využívaných rozdělení [13]. Nejčastěji se však využívá obecné rozdělení uvedené v tabulce 1.1. V jiných zdrojích se též vyskytuje:

- rozdělení dle doporučení od Mezinárodní organizace pro osvětlování (CIE),
- rozdělení dle normy ISO 20473,
- astronomické rozdělení.

1.2.4 Absolutně černé těleso

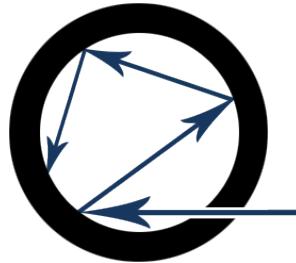
Absolutní černé těleso je ideální teoretické těleso, které pohlcuje veškeré dopadající záření všech vlnových délek. Současně s touto vlastností se chová jako ideální zářič. To znamená, že ze všech možných těles o stejné teplotě vyzařuje nejvíce energie.

V reálném světě se absolutně černému tělesu blíží některé hvězdy a nebo vyrobené přístroje, pomocí kterých jsou termokamery kalibrovány. [2, 14]

Název pásmá	Zkratka	Rozsah od	Rozsah do
blízké infračervené záření	NIR	0.76 μm	1.4 μm
krátké vlnové délky	SWIR	1.4 μm	3 μm
střední vlnové délky	MWIR	3 μm	8 μm
dlouhé vlnové délky	LWIR	8 μm	15 μm
vzdálené infračervené záření	FIR	15 μm	1000 μm

Tabulka 1.1: Obecné rozdělení oblastí infračerveného záření

Černé těleso si lze představit jako dutinu s matným černým povrchem a malým vstupním otvorem stejně jako je ilustrováno na obrázku 1.3. Jakmile záření projde otvorem do trubky, po několika odrazech se úplně pohltí. Tedy veškeré záření, které se do absolutně černého tělesa dostane, se již zpět ven neodrazí. Pohlcené vstupní záření se přemění na energii a je zpět vysíláno pouze ve formě tepelného záření.



Obrázek 1.3: Konceptuální černé těleso

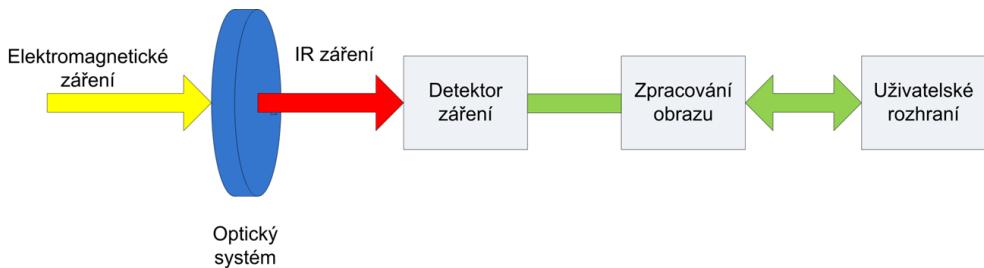
1.3 Konstrukce

Všechny termokamery se skládají z minimálně třech důležitých součástí, kterými jsou: optická soustava, detektor a obvody pro předzpracování obrazu [14, 12, 15]. Jednoduché schéma termokamery je možné vidět na obrázku 1.4.

Základní a nejednoduší dělení termokamer je podle přítomnosti ovládacích prvků [16]. Toto dělení probíhá na:

Stacionární termokamery obvykle využívané v průmyslu, v bezpečnosti, pro monitoring, apod. Stacionární kamery jsou trvale připojeny k počítači, který řídí jejich nastavení a přenos snímků. Obsahují pouze konektor pro připojení a nemají žádné ovládací prvky a většinou ani úložiště a baterii.

Ruční termokamery jsou přenosné a nezávislé na počítači. Využívají se zejména pro snímkování v terénu. Na rozdíl od stacionárních kamer obsahují vestavěnou baterii, ovládací prvky a LCD displej.



Obrázek 1.4: Jednoduché schéma hlavních konstrukčních částí termokamery [3]

1.3.1 Detektor

Detektor plní podobnou činnost jako snímač u běžných RGB kamer. Na detektor je pomocí optiky soustředěno infračervené záření, které je převáděno na elektronický signál. Signál je následně předzpracovaný korekcemi v dalších přídavných obvodech a pak teprve převeden na výsledný snímek. Obvody pro předzpracování signálu jsou již nejčastěji součástí detektoru. [12]

Samotné signály z detektoru bez aplikovaných korekcí lze zobrazit, ale výsledný obraz je velmi nekvalitní a navíc s velkým počtem nepřesností. Pro představu je přiložen obrázek 1.5, kde firma FLIR v roce 2012 prezentovala výsledky svých nových algoritmů pro zpracování signálu z detektoru. Z obrázku je patrné, že obvody pro zpracování signálu jsou velmi důležitou součástí termokamer.



Obrázek 1.5: Vlevo je snímek bez aplikovaných korekcí signálu a vpravo je snímek po aplikaci vhodných algoritmů zpracování signálu v příslušných obvodech [4]

Termokamery lze dělit i podle druhu detektoru. Tyto dvě skupiny jsou výsledkem konstrukční technologie a jedná se o dělení na chlazené a nechlazené detektory.

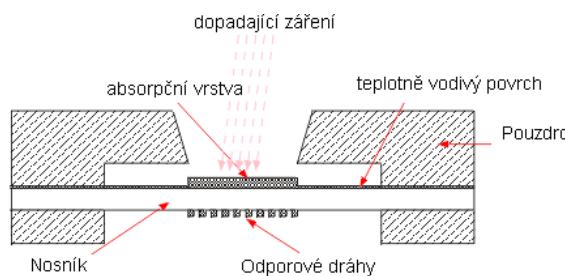
Chlazené detektory, někdy též označovány jako kvantové, přináší velmi přesné měření teplot, vysokorychlostní snímání, rádkové snímání (až

62 000 FPS v případě [17]) a lepší zpracování šumu. V závislosti na materiálu detektoru lze přizpůsobit pro IR oblasti SWIR, MWIR a LWIR. Kvůli nutnému chlazení mají větší rozměry a jejich váha se pohybuje v jednotkách kg. Nevýhodou je několikanásobně vyšší cena oproti nechlazeným detektorům.

Nechlazené detektory jsou konstrukčně jednodušší, spolehlivější, mají menší energetickou spotřebu a jsou cenově dostupné veřejnosti. V dnešní době jsou nejčastější nechlazené detektory na bázi mikrobolometrů. Zajímavostí je fakt, že se dá dosáhnout velmi kompaktních rozměrů. Například detektory FLIR řady Lepton mají rozměry přibližně $1 \times 1 \times 0.7$ cm. Tyto detektory se pouze ojediněle objevují pro jinou IR oblast než je LWIR. Nevýhody nechlazených detektorů jsou méně kvalitní výstup a menší přesnost měření oproti chlazeným detektorům.

Chlazené detektory jsou velmi specifickou záležitostí, týkající se převážně extrémních využití (záchranné a bezpečnostní složky, armáda, vědecké a výzkumné ústavy, apod.), jejich popis a bližší informace nebudou součástí této práce. Podrobné detaily o těchto detektorech lze nalézt v knihách [18, 19].

Naopak s nechlazenými detektory je možné se setkat téměř ve všech běžných případech. Technologie nechlazených detektorů využívá výhradně mikrobolometrického pole [5, 14]. Toto dvourozměrné pole je složeno z miniaturních bolometrů, které jsou citlivé na dopadající infračervené záření. Jednotlivé bolometry mění své elektrické odpory v závislosti na množství tepelného záření snímaného tělesa. Každý odpor znamená jednu hodnotu signálu a jeden výsledný pixel obrazu. Rozlišení nechlazených detektorů je dáno počtem uspořádaných bolometrů v matici.



Obrázek 1.6: Schéma bolometru [5]

Rozlišení termokamer se obecně v porovnání s RGB kamerami pohybuje na skromnějších hodnotách. Od dnešních nejnižších rozlišení 80×60 pixelů až po extrémní třídu s 1024×1024 pixely. Je to z toho důvodu, že pro dosažení dvojnásobného rozlišení termokamery, je nutné čtyřnásobné zvětšení počtu bolometrů v matici.

1.3.2 Optická soustava

Termokamery nemohou využívat běžnou optiku ze skla, tak jako ji využívají klasické kamery. Je to z toho důvodu, že běžná optika infračervené záření převážně odráží než propouští. V dnešní době již existuje více druhů používaných materiálů pro objektivy termokamer, ale nejčastěji se využívá germánium nebo jeho slitina. [20, 13]

Optika z germánia slouží také jako filtr všech ostatních vlnových délek a k samotnému detektoru, v případě germánia, se dostává pouze záření z pásma LWIR nebo MWIR (v závislosti na úpravách optiky). Na optice jsou dále naneseny antireflexní nano vrstvy pro zvýšení propustnosti infračerveného záření. Optika a nano vrstvy termokamer jsou velmi šetrné a neodborným zacházením mohou být způsobena nevratná poškození.

Optika bohužel přichází o své vlastnosti i postupem času. Jedná se zejména o vnější vlivy, které způsobují její oxidaci a tím snižují propustnost infračerveného záření. Více o volbách optických materiálů a jejich porovnání lze nalézt v článku [21].

1.4 Specifikace a parametry

Seznam důležitých parametrů ovlivňující klíčovou funkčnost a možnosti termokamer jsou uvedeny v tabulce 1.2. K jednotlivým parametrům je vždy uveden příklad a vysvětlující komentář.

Dalším velmi důležitým parametrem je aklimatizace kamery. Tento údaj není ve specifikacích často uváděn a přitom je pro měření klíčový. Aklimatizace kamery znamená dobu ustálení, kterou termokamera potřebuje, aby zaznamenávala co nejpřesnější výsledky. V praxi to znamená stabilizování teploty detektoru na předem danou hodnotu, běžně kolem 30 °C u nechlazených detektorů. Dostat se a udržovat stabilní teplotu detektoru má na starosti speciální vnitřní obvod.

1.5 Získávání dat

Práce se stacionární termokamerou probíhá o něco složitějším způsobem než je proti ručním kamerám zvykem. Kamera nemá žádné interní úložiště, display ani ovládací prvky. Všechna práce, jako je ovládání kamery a získávání dat, probíhá pomocí síťového rozhraní a je řízena počítačem. [16]

Uvedené informace v této části, se týkají primárně kamery FLIR řady Ax5 využité v této práci, ale obecně je lze aplikovat na všechny stacionární kamery využívající standardy GenICam a GigE Vision. Tato rozhraní jsou běžně využívány nejen u termokamer od výrobce FLIR, ale i u zařízení ostatních výrobců.

1. TERMOKAMERA

Parametr	Příklady	Komentář
Rozlišení detektoru	240×180 pixelů	Obecně platí, že vyšší je lepší.
Typ detektoru	nechlazený FPA	Chlazený (kvantový) nebo nechlazený, více v 1.3.1.
Snímané spektrum	LWIR, MWIR	Snímaná část IR spektra, více v 1.2.3.
Obnovovací frekvence	30 Hz	Počet snímků, které je kamera schopna za vteřinu pořídit (FPS).
Rozsah měřených teplot	0 - 320 °C	Často více možných voleb v souvislosti s režimem kamery
Úhel záběru	$20^\circ \times 25^\circ$	Často uváděná také ohnisková vzdálenost
Citlivost (NETD)	50 K, 0.05 °C	Vyjadřuje, jaké nejmenší teplotní rozdíly je na povrchu černého tělesa termokamera schopna zaznamenat.
Přesnost	± 5 °C nebo 5 %	Chyba, se kterou je nutné počítat při interpretaci výsledků. Platí horší údaj.
Formát zaznamenávaných snímků	14 bit MONO, 14 bit linear	Výsledný interpretovatelný formát a bitová kvalita.
Provozní teplota	-10 - 50 °C	Důležitý údaj pro využití v nepokojových podmírkách.
Komunikační standardy	GenICam, GigE	V dnešní době již výhradně protokol GenICam s jedním z přenosových standardů.

Tabulka 1.2: Klíčové parametry při volbě termokamery.

1.5.1 GigE Vision

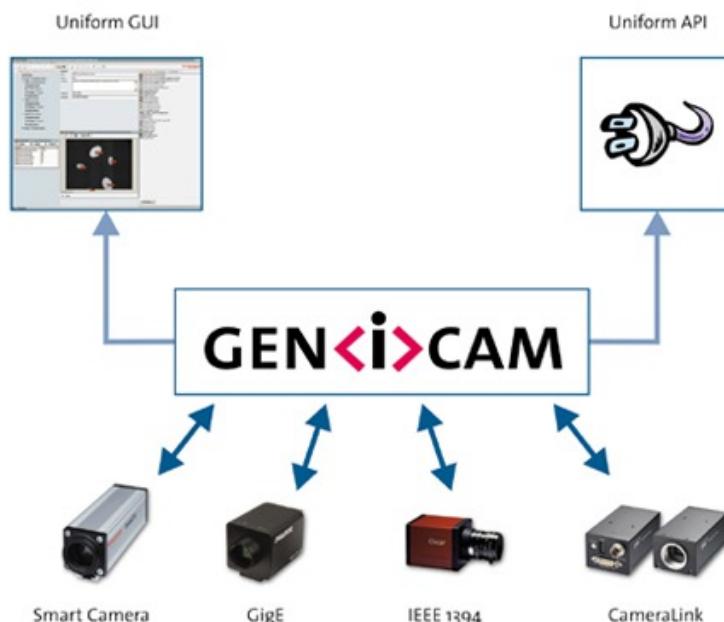
Standard GigE Vision byl v roce 2006 vytvořen pro potřeby vysokých průmyslových kamer. U jeho vzniku stalo celkem 12 firem a nyní je spravován asociací AIA. GigE Vision definuje protokol pro vysokorychlostní přenos obrazu a komunikaci s kamerami. Jeho úkolem bylo unifikovat stávající protokoly pro průmyslové kamery a umožnit lepší kompatibilitu hardwaru a softwaru. Architektura GigE Vision vychází ze standardního internetového protokolu UDP, takže je možné kamery používat v běžné síťové infrastruktuře. [22]

1.5.2 GenICam

Standard GenICam byl v roce 2006 vytvořen sdružením EMVA. Jednalo se primárně o snahu vytvoření univerzálního programového rozhraní pro průmyslové kamery strojového vidění. U zařízení s podporou GenICamu tedy

z hlediska vývoje nezáleží na technologii přenosu, vlastnostech kamery, typu kamery a dokonce ani na výrobci kamery. Kamery mohou být připojeny přes rozhraní GigE Vision, USB3 Vision a přesto jsou způsoby konfigurace, snímání, atd. díky jednotnému rozhraní vždy stejné. [23]

K tomu, aby vše správně fungovalo a bylo možné standard dále rozšiřovat, je GenICam rozdělen do několika nezávislých modulů, které jsou dále popsány přímo na stránkách výrobce. Stejně tak lze na stránkách sdružení EMA nalézt referenční implementaci pod BSD licencí. Standard se dále vyvíjí a rozšiřuje s pomocí požadavků uživatelů.



Obrázek 1.7: Koncept GenICam protokolu [6]

1.5.3 Dostupný software

Existují již hotové programy pro práci se stacionárními termokamerami, ale i v placených verzích těchto softwarů nelze nalézt všechna potřebná nastavení a funkce. Je to z toho důvodu, že tyto softwary jsou většinou pro průmyslové aplikace jako je monitoring, tvorbu reportů a ostatní komerční činnosti.

Příkladem je placený software FLIR Tools+ [24] přímo od výrobců termokamer FLIR. Tento software podporuje přenos obrazu z kamery a jeho následné zobrazení, ale snímky jsou přenášeny pouze v 8-bitovém formátu. Dále nedovoluje ukládání surových dat, nepodporuje kontinuální ukládání a ani zásah do nastavení kamery.

1. TERMOKAMERA

Všechny nalezené alternativní softwary mají též velmi omezené možnosti a funkce. Kvůli tomu je vhodné využít SDK a zabývat se tvorbou vlastní aplikace pro ovládání kamery a získání snímku pro nutné potřeby řešení této práce.

1.5.4 Sada vývojových nástrojů

Lepším řešením, ale bohužel časově náročnějším, je využít vhodnou sadu vývojových nástrojů (SDK) a s kamerou pracovat na úrovni kódu. Kamera FLIR A65 podporuje rozhraní GigE a GenICam a ač byly oba standardy vytvořeny již v roce 2006, dostupných SDK je v řádu jednotek. Nejčastěji využívané jsou:

- Atlas SDK [25],
- JAI SDK [26],
- ActiveGigE [27],
- eBUS SDK [28].

Všechny výše uvedené SDK podporují standardy GigE Vision a GenICam. Liší se použitým programovacím jazykem, kvalitou dokumentace, množstvím ukázkových programů a také dostupností. Stručně shrnuto:

Atlas SDK je dostupné pouze k zakoupeným kamerám značky FLIR. Práce s Atlas SDK probíhá v jazyce C# nebo prostředí MATLABu. Práce s Atlas SDK je velmi snadná, má velmi dobrou dokumentaci a ukázkové kódy. Jedinou nevýhodou je omezený přístup k nastavování jednotlivých parametrů protokolu.

JAI SDK je produkt od výrobců průmyslových kamer JAI. V současné době se firma JAI podílí také na rozvíjení standardů GigE Vision a GenICam. Jejich SDK je možné využívat v jazyce C/C++ nebo v C#. Je zdarma, obsahuje velmi detailní dokumentaci a dostatečné množství ukázkových kódů. Součástí SDK je ovladač Filter driver pro urychlení přenosu dat.

ActiveGigE je SDK firmy A&B Software. Na rozdíl od jiných SDK ho lze využít ve spoustě různých prostředích a programovacích jazycích. Je dostupné pro: Visual Studio, Visual Basic (VB), Delphi, PowerBuilder, Java, Matlab, Python, QT, Adobe Flash, LabView, GE Fanuc, Indusoft Studio. Zajímavou záležitostí je vkládání ActiveGigE objektů přímo do programů od Microsoftu (Word, PowerPoint) s možností přenosu reálného obrazu. Ze všech porovnávaných SDK obsahuje v základu nejvíce implementovaných funkcí. Samozřejmostí je rozsáhlá dokumentace a ukázkové kódy. Jedinou nevýhodou se zdá být cena, ta je k dubnu 2017 stanovena na 495 dolarů za licenci.

eBUS SDK je od firmy Pleora, která se stejně jako JAI podílí na rozvíjení standardů pro strojové vidění. Je dostupné jak pro jazyk C/C++, tak

1.5. Získávání dat

pro C#. Jeho použití je díky rozsáhlé dokumentaci a ukázkovým zdrojovým kódům velmi snadné. Součástí eBUS SDK je také ovladač Universal Pro Driver pro zvýšení propustnosti síťového rozhraní a snížení zátěže. eBUS SDK je pro kamery FLIR dostupné zdarma, pro kamery ostatních výrobců je licence zpoplatněna.

KAPITOLA **2**

Analýza

V této kapitole jsou popsány existující řešení úlohy detekce zboží v ruce a následně zmíněny práce, které byly shledány jako obsahově zajímavé a užitečné pro řešení této práce. Poté je proveden rozbor zadané úlohy a stanovení nutných předpokladů pro řešitelnost. Dále jsou popsány přínosy, které termokamera v řešení tohoto problému poskytuje a následně i překážky se kterými je nutno počítat. V poslední části jsou představeny dostupné hardwarové a softwarové prostředky.

2.1 Současný stav řešení problematiky

Nebily nalezeny žádné dostupné informace o tom, že by se problematikou zboží v ruce zákazníka ve světě někdo zabýval. Jediné dostupné řešení lze nalézt právě u studenta Olivera Keruľ-Kmece v jeho bakalářské práci [7], který problém řešil s jinou technologií než se zabývá tato práce. Řešený problém je tedy stejný, ale díky rozdílnosti technologií je k němu nutné navrhnout nové postupy a metody.

V jeho řešení je využita kamera Microsoft Kinect, která současně snímá klasický RGB obraz a zároveň k tomu hloubkovou mapu. Jeho řešení využívá dva různé klasifikátory. Vstupem prvního klasifikátoru je segmentovaný záběr ruky pomocí hloubkové mapy. Na tomto vstupu se vypočítá 60 předem daných příznaků a následně je ke klasifikaci využit klasifikátor Gradient boosted trees. Druhý klasifikátor je využit ke zvýšení přesnosti algoritmu a klasifikuje posloupnost tříd, která je získaná prvním klasifikátorem. Posloupnosti jsou rozděleny pomocí algoritmu SIFT na vložení ruky a vyjmutí ruky z regálu.

2.2 Související práce s termokamerou

Existuje samozřejmě velké množství prací, které se zabývají segmentací termografických snímků a detekcí objektů, ale jsou zmíněny jen ty nejzajímavější,

2. ANALÝZA

které byly nalezeny. Jako zajímavé práce jsou shledány zejména ty, které využívají řešení s pomocí více obrazových zdrojů. V termografickém oboru lze nejčastěji nalézt práce, které se zabývají segmentací člověka, detekcí obličeje nebo detekcí gest.

První z nalezených prací je [29], která se zabývá detekcí ruky a sledování její trajektorie s využitím termokamery a kamery Microsoft Kinect. V uvedené práci jsou přijímány dva obrazové zdroje a to hloubková mapa a termografický snímek. Kamery mají rozdílnou optiku, takže je nutné jednotlivé snímky perspektivně upravit. K tomu je využita knihovna OpenCV a transformace pomocí matice homografie, která slouží pro převod mezi jednotlivými snímky. Následně je na snímku získané jeho popředí podle [30]. Na termovizní snímek je aplikováno Otsu prahování [31] pro získání ruky, na které jsou vypočítány konvexní body pro získání prstů. V poslední části jsou vypočteny atributy, které se týkají primárně teplot a snímek je klasifikován pomocí klasifikátoru SVM. V uvedené práci bohužel není ani zmínka o synchronizaci kamer, která se následně v této práci ukázala jako velký problém.

Dvě různé kamery a metoda dynamického odčítání pozadí je využita také v práci [32] pro detekci gest. Zarovnání dvou obrazových zdrojů, je stejně jako v [29], řešeno pomocí transformační matice homografie. Synchronizace bohužel není zmíněna. Segmentace probíhá pouze formou odečtení pozadí. Podle teplot se nesegmentuje, což může být překvapivé. Po tomto kroku už se jen na segmentovaný snímek aplikují morfologické operace [33] a následně je na snímku vyhledána ruka.

Práce [34] se zabývá 3D mapování interierů pomocí kombinací RGB a termografického snímku. V práci se řeší synchronizace jednotlivých zdrojů vypočítáváním posunu časových známek pořízených snímků. Hardwarová synchronizace není přítomna.

V práci [35] jsou kombinovány 3 různé obrazové zdroje pomocí termokamery a kamery Microsoft Kinect. Uvedená práce je zajímavá tím, že popisuje metody korekce perspektivy pomocí kalibrační šachovnice, jak pro Kinect tak i pro termokameru. Zkalibrované snímky jsou na sebe následně zarovnány pomocí transformační matice homografie. Synchronizace není bohužel řešena.

Další nalezenou zajímavou prací je [36], kde se autor zaměřuje na detekci gest a porovnává data získána různými obrazovými zdroji. Porovnáván je klasický RGB snímek, hloubková mapa a termografický snímek. Segmentace ruky na termovizním snímku probíhá pomocí binárního prahování a tento obraz je sloučen s nejbližšími oblastmi hloubkového obrazu. Na tomto obraze jsou dále získané atributy a snímek je klasifikován pomocí Hidden Markov Models (HMM).

V práci [37], jsou shrnutý běžné způsoby segmentace termografických snímků, které pak byly experimentálně vyzkoušeny při řešení této práce.

Ve výše uvedených pracích lze často objevit informaci o tom, že se počítá s tím, že teplota ruky se pohybuje v určitém rozsahu a lze ji segmentovat podle

teplot s pomocí binárního prahování. Tento jednoduchý způsob se následně osvědčil i v této práci.

2.3 Analýza zadaného problému

Zadání problému je vcelku přímočaré. Pro vstupní snímek je nutné rozhodnout, zda se na něm nachází ruka nebo ruka se zbožím a nebo se na něm nenachází nic zajímavého. Je však potřeba počítat s tím, že se nejedná pouze o návrh algoritmu, ale také o získání samotných dat a v neposlední řadě také ověření správného postupu vyhodnocením.

Prvním nezbytným krokem je tedy zaměřit se na to, jakým způsobem budou data získávána. V 1.5 byly popsány možné metody získání dat, ze kterých plyne, že běžně dostupné softwary jsou pro tuto úlohu nedostačující. Je tedy nutné poohlédnout se po vhodném SDK, seznámit se s jeho rozhraním a vytvořit aplikaci pro kontinuální ukládání surových dat. Přijímaná data z kamery není vhodné převádět do komprimovaného formátu z důvodu zachování maximální kvality.

Dalším nutným krokem je průzkum nasnímaných dat a navržení vhodného postupu jejich předzpracování. Tento postup by měl data zejména omezit na určitou oblast zájmu a převést do normalizovaného rozsahu teplot. Podle nutnosti pak záběr patřičně upravit a nakonec redukovat šum.

Na předzpracované snímky aplikovat vhodné algoritmy a tím získat potřebné informace k dalšímu rozhodování o výsledku. S přihlédnutím k faktu, že zboží má nižší teplotu než ruka, by zboží mělo být ve většině případech rozlišitelným. Otázkou zůstává, jak velké předměty je ještě možné detektovat než je dosaženo možností algoritmu a limitu termokamery. S limitem termokamery bude souviset generovaný šum detektoru, se kterým je nutné se vypořádat.

Posledním krokem je klasifikace do předem definovaných tříd. Tyto třídy by mohly být tři a to: pozadí, ruka, ruka se zbožím. Vzhledem k tomu, že určit pozadí lze triviálním způsobem, bude brána v úvahu pouze binární klasifikace třídy ruka a ruka se zbožím.

Je zřejmé, že zboží které na snímku není vidět, například je schované v dlani ruky, nelze tímto způsobem detektovat. V práci se tedy berou v potaz pouze snímky, kde zboží viditelné je. V práci se též neřeší o jaký typ zboží se jedná, protože to se samotnou termokamerou není zjistitelné a existují k tomu lepší metody.

2.4 Nutné předpoklady

Aby byl problém termokamerou řešitelný, musí být bráno v potaz několik předpokladů, které souvisí s teplotou snímaných objektů. Uvedené předpoklady pouze vylučují extrémní případy, které se běžně ve vnitřních prostředí nestávají, ale je nutné je brát na vědomí.

2. ANALÝZA

2.4.1 Předpoklad pro teplotu ruky na snímku

O termoregulaci se v lidském těle stará část mozku s názvem hypothalamus. Povrchová teplota lidského těla není stabilní a v některých situacích se znatelně mění, zato teplota lidského jádra zůstává dlouhodobě neměnná a to dokonce i v případě velkých výkyvů okolních teplot. [38]

Průměrná teplota zdravého člověka na povrchu čela je při běžných pokojových podmínkách 36 - 37 °C [39]. Končetiny, a zvláště prsty, jsou vždy o něco chladnější, protože krev nejdříve putuje přes životně důležité orgány a až poté se dostává do končetin. Rozvádění krve až do končetin je pro tělo náročné, a v rámci vlastních experimentálních měření bylo zjištěno, že při přechodu z chladnějších venkovních podmínek do pokojové teploty je povrchová teplota ruky schopna klesnout i o 10 °C níž, než je její běžný stav. Tento pokles trvá několik minut, dokud se tělo neaklimatizuje.

V této práci se tedy počítá s tím, že snímaný člověk se již nějakou dobu pohyboval v běžné pokojové teplotě a teplota jeho ruky se již ustálila.

2.4.2 Předpoklad pro teplotu zboží a regálu na snímku

Regál je umístěn v místnosti s pokojovou teplotou a jednotlivé zboží má velmi podobnou teplotu jako je teplota vzduchu v místnosti. Teplota zboží v regále musí být rozlišitelná od ruky, jinak by úloha termokamerou nebyla řešitelná.

2.4.3 Předpoklad pro teplotu podlahy na snímku

Pomocí vlastních experimentálních měření bylo zjištěno, že podlaha běžné místnosti je ve většině případů o pár desetin stupňů Celsia studenější než je teplota vzduchu v místnosti. V práci se počítá s tím, že teplota podlahy nepřevyšuje teplotu zboží a tedy ani ruky na snímku.

2.5 Výhody využití termokamery pro tuto úlohu

Tato úloha lze samozřejmě řešit i pomocí jakékoli barevné kamery. Například způsobem, který je uveden v již zmíněné práci Olivera Kerul-Kmece [7]. Všechna řešení pomocí klasické barevné kamery by k segmentaci ruky nejčastěji využívaly její barvu. Barva kůže není přirozeně nikde definovaná a navíc je závislá na světelných podmínkách v dané místnosti. Tento způsob segmentace tedy nemusí vždy fungovat naprostě korektně a obecně není moc dobré se na barvu kůže zcela spoléhat. Dalším možným problémem je situace, kdy má zboží podobnou barvu jako ruka, což se díky různorodosti zboží může běžně stávat.

V 2.4.1 bylo zjištěno, že teplota ruky je silně ovlivněna působením chladného vnějšího vlivu. Tento případ je ale v reálném světě prostředí úlohy velmi málo pravděpodobný, protože člověk se ve vnitřním prostředí často velmi

rychle aklimatizuje. Pokud by přeci jen tento případ opravdu nastal a teplota ruky klesla například z běžných 35 °C na 25 °C, byla by stále rozlišitelná od zboží, které dosahuje pokojové teploty někde mezi 16 až 22 °C. Řešení s termokamerou by tedy mělo být pro segmentaci ruky robustnější, než segmentování ruky podle barvy s běžnou RGB kamerou.

Další výhodou je, že termokamera není závislá na světelných podmínkách. Snímá totiž pouze námi neviditelné infračervené záření. Z toho plyne, že stejné výsledky lze naměřit i v úplné tmě.

Posledním zajímavým poznatkem je, že snímání termokamery neovlivňuje barvy ani vzory, takže výsledný obraz obsahuje méně rušivých elementů než například běžný RGB obraz.

2.6 Překážky při snímkování s termokamerou

Pří snímkování s termokamerou je nutné brát v potaz několik skutečností, které ovlivňují snímaná data a mohou též vést k chybným výsledkům. Tyto skutečnosti mohou být dané konstrukčními vlastnostmi kamery a nebo se jedná o uživatelské chyby nastavení.

2.6.1 Korekce signálů detektoru - NUC

Tak jako běžné kamery i termokamera se při snímkování zahřívá. Změna teploty detektoru pak ovlivňuje jednotlivé mikrobolometry a to tím, že začne vznikat posun hodnot signálů (temperature drift). Tento posun zvládne kamera do určité míry kompenzovat interní logikou, ale když už posun překročí jistý limit, je automaticky spuštěna kalibrační funkce NUC neboli Non-Uniformity Correction. Termokamera vloží mezi objektiv a detektor na krátký okamžik závěrku, která simuluje jednotné homogenní prostředí. S pomocí tohoto uniformního prostředí jsou nastaveny korekce pro normalizaci hodnot signálů.

V době kalibrace termokamera nepřenáší obraz. Tato kalibrace trvá přibližně vteřinu a její četnost je v závislosti na změnách teploty detektoru. Kalibraci NUC tedy může spustit i změna teploty prostředí. Z počátku měření se NUC provádí ve velmi krátkých časových intervalech, ale jak se teplota detektoru stabilizuje, provádí se NUC již méně častěji.

Více informací o NUC a obecně o problému posunu signálů bolometrů se lze dočíst v [40, 41].

2.6.2 Přenos snímků

Protokol GigE pro přenos snímku z kamery vychází z internetového protokolu UDP (User datagram protocol [42]). To znamená, že na rozdíl od TCP (Transmission Control Protocol [43]) protokolu odesílatel neověřuje, zda příjemce data obdržel. Může se tedy stát, že některé snímky nemusí být vůbec

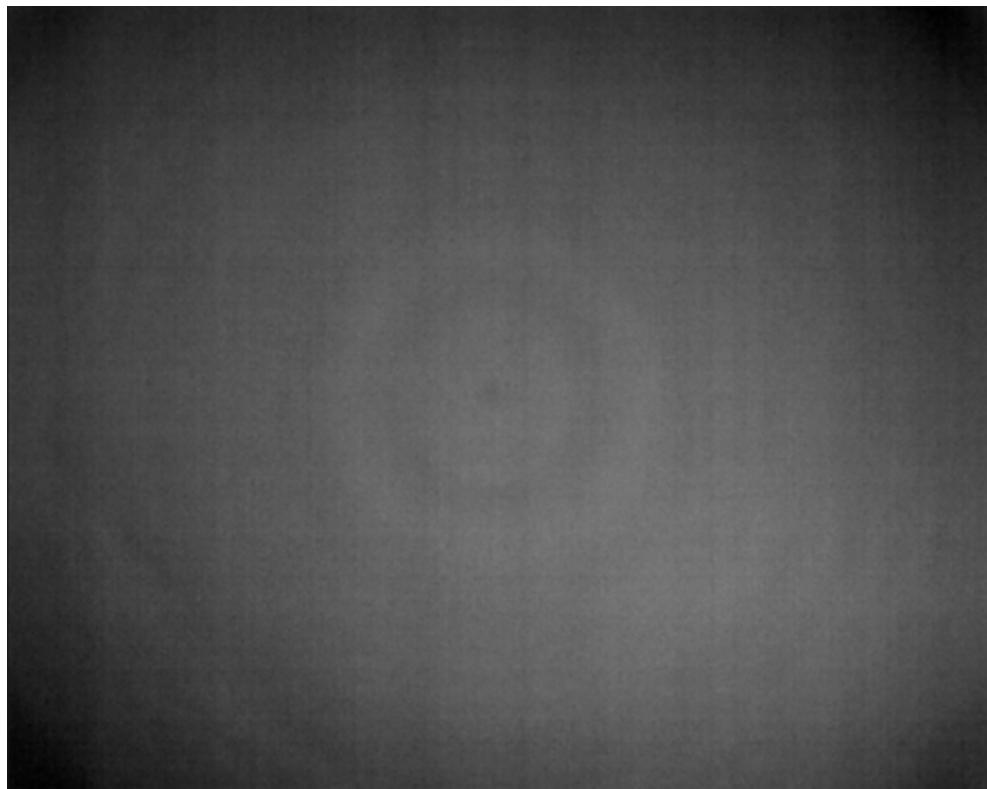
2. ANALÝZA

přijaty. Jedná se spíše o výjimečný případ, ale přesto je potřeba s touto skutečností počítat.

2.6.3 Vady optiky

Optika termokamery není, tak jako žádná optika, dokonalá. Objevují se zde optické vady stejně jako u tradičních objektivů pro běžné kamery. Obecně lze říct, že maximální obrazovou kvalitu lze nalézt v oblasti, která je co nejbližší středu obrazového pole. V těchto místech je obraz nejostřejší, nezkreslený a s největší rozlišovací schopností.

Na obrázku 2.1 je zobrazen neupravený snímek neutrálního prostředí, kde je možné pozorovat optické vady termokamery FLIR A65 s ohniskovou vzdáleností 25 mm. Z obrázku je zřejmé, že největším problémem je vinětace v rozích, která velmi výrazně ovlivňuje výsledné teploty v této oblasti.



Obrázek 2.1: Ukázkový snímek optických vad u termokamery FLIR A65 s ohniskovou vzdáleností 25 mm

2.6.4 Chyby měření a vyhodnocování

Při snímání s termokamerou mohou chyby vznikat už při měření a to zejména chybným nastavením kamery. Kamera na základě těchto chybných nastavení stanovuje neodpovídající teploty. Tyto parametry jsou například: emisivita, zdánlivě odražená teplota a atmosferické vlivy. V rámci práce je nutné teploty od sebe umět rozlišit, ale není nutné znát jejich přesné hodnoty. Proto se nastavení těchto parametrů explicitně nezmiňuje a pro účely této práce nejsou klíčové. Vliv těchto parametrů a postupy, jak je správně nastavit, lze nalézt v související literatuře [44, 11, 12, 16].

2.7 Dostupné prostředky

V této části jsou popsány hardwarové a softwarové prostředky, které jsou v rámci práce dostupné.

2.7.1 Dostupná zařízení

V práci byly v průběhu využity nezávisle dvě termokamery FLIR A65. Kamery se od sebe lišily ohniskovou vzdáleností a snímkovací frekvencí. Jednalo se o kameru FLIR A65 s ohniskovou vzdáleností 25 mm a druhou kameru stejné řady s ohniskovou vzdáleností 13 mm. V částech práce, kde to bude nutné, bude pro zjednodušení o kamerách psáno už jen jako kamera (25 mm) a kamera (13 mm). Klíčové parametry těchto kamer jsou uvedeny v tabulce 2.1, všechny parametry pak lze nalézt přímo ve specifikacích výrobce [44].

Využití druhé kamery bylo nezbytné, protože v průběhu práce nastaly s kamerou (25 mm) technické potíže a bylo nutné ji reklamovat. Aby bylo možné v práci pokračovat, byla zajištěna náhradní kamera stejné řady, bohužel s lehce odlišnými parametry.

Model A65 má vzhledem ke svým malým rozměrům a relativně nízké ceně nadprůměrné rozlišení. Vyšší rozlišení už se objevuje pouze u pár modelů pro extrémní využití za mnohonásobně vysší cenu. Další výhodou je velmi dobrá teplotní citlivost 50 mK, která je v této třídě kamer spíše nadstandard. U kamery (13 mm) je bohužel snímkovací frekvence velmi pomalá. 7.5 FPS je na dnešní poměry málo a pomocí této frekvence nelze běžně zaznamenávat kontinuální pohyb. Ideální by bylo, kdyby kamera (13 mm) měla minimálně 15 FPS, ale ta bohužel nebyla k dispozici. Snímkovací frekvence kamery (25 mm) bohatě dostačuje, v této kategorii lze též považovat za nadstandard.

Pro některé aplikace může být nevýhodou přesnost, která u této řady nepatří mezi nejlepší. Nutno podotknout, že tato řada je primárně určena pro počítačové vidění, kde se očekává kvalitní rozlišovací schopnost (rozlišení NETD) a na přesnosti měřených teplot tolik nezáleží.

Důležitým parametrem je také skutečnost, že kamera zvládá přenášet data v 14-bitovém Temperature linear formátu. To znamená, že jednotlivé pixely

2. ANALÝZA

Specifikace	Kamera (13 mm)	Kamera (25 mm)
Rozlišení detektoru	640 × 512 pixelů	
Teplotní citlivost (NETD)	< 0.05 °C @ +30 °C / 50 mK	
Úhel záběru (FOV)	45° × 37°	25° × 20°
Ohnisková vzdálenost	13 mm	25 mm
Prostorové rozlišení	1.31 mrad	0.68 mrad
Clonové číslo	1.25	
Snímkovací frekvence	7.5 Hz	30 Hz
Zaostřování	Manuální	
Typ detektoru	Nechlazený FPA na bázi mikrobolometrů	
Spektrální rozsah	7.5–13 µm, oblast LWIR	
Časová konstanta detektoru	12 ms	
Rozsah teplot	-25 °C až +135 °C	
Přesnost měření	±5 °C nebo 5 %	
Komunikační standardy	GigE vision, GenICam	
Formát přenášených snímků	8 a 14 bit MONO Signal linear, 14 bit MONO Temperature linear	
Napájení	Power over Ethernet (PoE)	
Provozní teplota	-15 °C až +50 °C	

Tabulka 2.1: Klíčové specifikace kamer použitých v této práci

získaných nezpracovaných dat, po vynásobení konstantou, již odpovídají rovnou teplotám. Kdyby termokamera nepodporovala tento režim, ale pouze Signal linear, tak při obdržení signálu je teplotu nejprve nutno vypočít pomocí vzorce. Více informací o těchto dvou režimech je uvedeno v 3.1.2.



Obrázek 2.2: Termokamera FLIR A65

2.7.2 Softwarové nástroje

Pro získání dat z termokamery je nutné vybrat vhodné SDK. Následně s jeho pomocí vytvořit řídící aplikaci pro příjem dat a ovládání kamery. Ideální pří-

pad by nastal, pokud by bylo možné zvolit SDK s rozhraním pro Javu a integrovat jej do stejné aplikace jako navržený algoritmus.

Pro algoritmus detekce bude vytvořena aplikace v jazyce Java s využitím knihovny OpenCV. Tato aplikace bude mít vytvořené grafické rozhraní pomocí platformy JavaFX [45]. Grafické rozhraní ulehčí manipulaci s daty a usnadní nastavení parametrů jednotlivých algoritmů.

Pro vyhodnocování a klasifikaci dat bude využit nástroj RapidMiner [46], který je v tuto chvíli jeden z nejvyužívanějších nástrojů pro data mining. Tento nástroj lze získat v bezplatné community verzi, která má sice svá omezení, ale je pro potřeby této práce dostačující.

2.7.2.1 Volba SDK

Pro účely této práce by bylo nejhodnější využít SDK ActiveGigE (1.5.4). Poskytuje nejvíce funkcí již v základu a je dostupné pro Javu, takže by jej bylo možné integrovat do aplikace s detekčním algoritmem. Naneštěstí se jedná o placené SDK a tak bylo nutné přejít k alternativní volbě.

Jiné kompatibilní SDK, které by mělo rozhraní pro Javu neexistuje, takže bylo vybíráno na základě jednoduchosti a kvality dokumentace. Rozhodování probíhalo mezi JAI SDK a eBUS SDK, ale nakonec bylo rozhodnuto pro eBUS SDK, které je podstatně jednodušší a nabízí velké množství již implementovaných funkcionalit.

2.7.2.2 OpenCV

OpenCV [47] je volně dostupná knihovna (pod licencí BSD), primárně určená pro zpracování obrazu. Je napsaná v jazyce C/C++ a poskytuje rozhraní pro ostatní programovací jazyky jako je Java a Python. Všechny její operace jsou optimalizovány a navrhnuty pro zpracování dat v reálném čase. Pro maximizaci výkonu využívá hardwarové akcelerace a paralelní výpočty.

KAPITOLA 3

Průzkum možných řešení

Tato kapitola je zaměřena na hledání vhodného řešení zadaného problému. Vzhledem ke skutečnosti, že problém je dost specifický a nikdo jej ještě pomocí termokamery neřešil, jeho referenční řešení není známé. Jsou využity pouze znalosti získané při rešerši a je navrhнуто několik postupů, ze kterých bude vybrán ten nejvhodnější.

3.1 Získávání dat z kamery

Před samotnou implementací algoritmů je nutné se zaměřit na získání dat. V předchozí kapitole bylo popsáno, že se nejedná o triviální úlohu a tak je potřeba vytvořit vlastní aplikaci pro nastavení a příjem dat z kamery.

3.1.1 Aplikace pro přenos dat a komunikaci s kamerou

V rámci tvorby aplikace pro komunikaci s kamerou, bylo nutné se seznámit s poskytovaným rozhraním vybraného eBUS SDK. Dokumentace k tomuto SDK je velmi detailní a tato část probíhala bez větších potíží. Bylo velmi snadné vytvoření jednoduché aplikace v jazyce C++ pro připojení kamery, jejího nastavení a získání jednotlivých snímků. Problém nastal až při implementaci kontinuálního ukládání dat, kdy přes veškerou snahu se nedářilo tuto funkčnost zprovoznit. Aplikace se zasekávala, vracela chybová data a ukládání fungovalo nespolehlivě.

Pro vypořádání se s tímto problémem byla zvolena alternativa, a to využití aplikace eBUS Player z ukázkových C++ kódů. Tato aplikace již obsahuje všechny nutné funkcionality pro správu připojení, nastavení a ukládání dat kamery. Bylo by možné pomocí této aplikace pouze data přenášet z kamery, ukládat na disk a další práci nad daty provádět formou záznamu. Jako elegantnější řešení bylo rozšíření eBUS Playeru pro přenos dat do libovolného programovacího jazyka pomocí síťových soketů. Následné zpracování a třídění

přenesených dat pak může probíhat pohodlněji ve vyšším programovacím jazyce.

V této rozšířené aplikaci je možné nastavit všechna potřebná nastavení kamery pomocí uživatelského rozhraní a navázat komunikaci s klientem. V případě, že je komunikace navázána úspěšně, začnou se přijímaná data z termokamery přenášet skrz soket. Při přenosu se nevyužije naplno propustnost síťového rozhraní, takže přenos není zahlcen a data klientem mohou být přijímány prakticky ihned. Obraz je tímto způsobem možné přenášet téměř v reálném čase.

Nutno podotknout, že součástí odesílaných souborů nejsou žádná metadata ani jiné hlavičky, které by popisovaly příchozí soubory. V rámci práce to nebylo nutné, protože klientská aplikace vždy ví, jaká data má očekávat.

3.1.2 Formát přenášených dat

Aplikace pro přijímání dat z kamery umožňuje ukládat data buď v komprimovaných formátech jako je JPG a BMP, nebo v nekomprimovaných jako je TIFF a nebo v surovém formátu, kde jsou obsahem jednotlivé hodnoty z detektoru. Vzhledem k tomu, že je vždy lepší pracovat s co nejméně komprimovanými daty, je rozhodnuto pro přenos surových dat. Tato nezpracovaná data je z termokamery možné přijímat ve dvou různých formátech. To se řídí dle nastaveného režimu v protokolu GenICam.

Signal linear mode V tomto režimu odpovídají jednotlivé pixely obrazu hodnotám signálu z maticového detektoru. Tento signál toho sám o sobě nic moc neříká, může být pouze převeden na teplotu. Pro převod signálu S na teplotu T v kelvinech slouží vzorec:

$$T_{[K]} = \frac{B}{\ln\left(\frac{R}{S-O} + F\right)}, \quad (3.1)$$

kde hodnoty R , B , F , O jsou uloženy v registrech kamery a lze je zjistit v uživatelském rozhraní přehrávače eBUS a nebo dotazem do GenICam protokolu.

Temperature linear mode Data získaná v tomto režimu při vynásobení konstantou K přímo odpovídají teplotám. Hodnota konstanty K se liší dle typu kamery a také dle rozlišení signálu nastaveného v protokolu GenICam. Pro kameru A65 teplotu T v kelvinech lze ze signálu S získat jako:

$$T_{[K]} = 0.04 \times S \quad (3.2)$$

pro vysoké rozlišení signálu a

$$T_{[K]} = 0.4 \times S \quad (3.3)$$

pro nízké rozlišení signálu.

Z výše uvedených způsobů získání teplot ze surových dat je zřejmé, že využití režimu Temperature linear je výpočetně jednodušší a pohodlnější, protože není potřeba dotazovat protokol a používat výpočetně náročný vzorec s dělením a logaritmem pro každý pixel obrazu. Přenášená data jsou tedy vždy odesílána neupravena, v 14-bitové kvalitě a ve formátu Temperature linear s vysokým rozlišením.

3.1.3 Oblast snímání

Nemělo by smysl snímat celé obrazové pole, protože se v něm vyskytuje regál i osoba. V tomto případě je zajímavé se zaměřit pouze na oblast, kde se může vyskytovat ruka. To znamená, snímat pouze bezprostřední oblast před regálem.

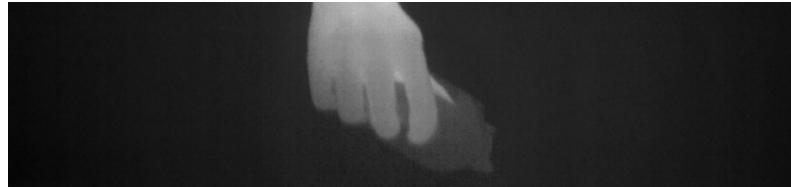
Velikost této oblasti byla experimentálními způsoby stanovena pro obě kamery na hodnotu 640×150 pixelů. Oblast je záměrně o něco širší než by bylo potřeba. Je to z toho důvodu, že během snímání může kdykoliv nastat kalibrace NUC a přerušit tím snímání. Tím, že bude oblast o něco větší, bude získáváno více snímků ruky a tím snížena pravděpodobnost, že NUC překazí snímání v důležitý moment.

V případě, že by kamera (13 mm) měla rychlejší snímkovací frekvenci, byla by pro ni oblast s výškou 70 pixelů dostačující. Příklad výřezu oblasti a jeho neupraveného zdroje je zobrazen na obrázku 3.2 a 3.1.



Obrázek 3.1: Neupravený snímek v plné velikost

3. PRŮZKUM MOŽNÝCH ŘEŠENÍ



Obrázek 3.2: Výřez oblasti zájmu neupraveného snímku

3.1.4 Redukce oblasti

Kamera A65 v sobě mimo jiné nese registry pro možnost snímání pouze výřezu obrazového pole. Díky tomu je možné z kamery rovnou přijímat redukovanou oblast. Tyto registry se nastavují v protokolu GenICam a je možné je změnit v uživatelském rozhraní upraveného eBUS Playeru. Registry pro výběr oblasti jsou označeny jako: Width, Height, OffsetX, OffsetY. Jejich hodnoty pro jednotlivé kamery použité v práci jsou uvedeny v tabulce 3.1. Posun v ose Y je nastaven tak, aby se zaznamenávaná ruka na snímku vyskytovala přibližně uprostřed.

Parametr	Kamera 1 (13 mm)	Kamera 2 (25 mm)
šířka	640 pixelů	
výška	150 pixelů	
posun v ose X	0 pixelů	
posun v ose Y	190 pixelů	170 pixelů

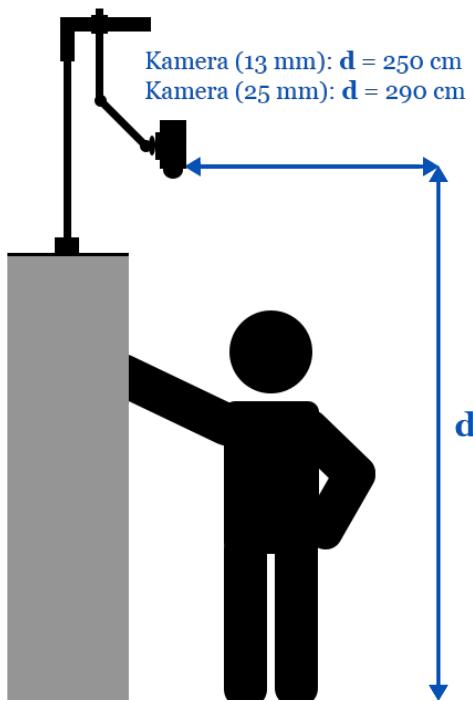
Tabulka 3.1: Hodnoty registrů pro výběr oblasti zájmu

Vyhotovení výřezu obrazového pole rovnou na straně termokamery je výhodné v tom, že je přenos oproštěn od nezajímavých dat a není zbytečně zatěžováno síťové rozhraní. Pokud by se přeci jen přenášel celý snímek, musela by určitá redukce obrazového pole stejně proběhnout na straně klientské aplikace, protože většina algoritmů má velkou výpočetní náročnost a není nutné je provádět nad obrazem v plném rozlišení.

3.1.5 Proces snímání

Snímání probíhalo v Laboratoři zpracování obrazu na FIT ČVUT v Praze. Teplota se v této místnosti pohybovala v rozmezí běžné pokojové teploty a to mezi 19 - 24 °C. Teplota zboží v regále a podlahy v podobných relacích.

Regál je vysoký 200 cm a široký 80 cm. Vzdálenost umístění kamer od země je uvedena na obrázku 3.3. Reálné umístění kamery a podoba regálu je vidět na fotografii 3.4. Kamera je umístěna v souladu s posunem uvedeným v 3.1.4, aby se ruka sahající do regálu vyskytovala přibližně uprostřed, kde obraz dosahuje nejvyšší kvality. Optika kamery je manuálně zaostřena přibližně do 3/4 regálu vzdálenosti od podlahy.



Obrázek 3.3: Schéma umístění kamery

3.2 Předzpracování dat

Ještě před začátkem aplikace algoritmů je data nutné převést z binární podoby na teploty. Tyto teploty normalizovat do nového rozsahu, aby je bylo možné zobrazit. Poté už je možné pracovat se snímkem, který je nutno obrazově předzpracovat.

Zpracování snímku probíhá v následujícím pořadí: převod binárních dat, normalizace teplot, expoziční úpravy, odstranění šumu. Efekt těchto kroků je zobrazen na obrázku 3.5.

3.2.1 Převod binárních dat a normalizace teplot

Binární data z kamery jsou snímána v Temperature linear režimu, takže je možné je snadno převést podle 3.1.2 na teplotu v kelvinech. Teplotu v kelvinech je ještě vhodné převést na běžnější teplotu v °C a to pomocí jednoduchého vzorce:

$$T_{[C]} = T_{[K]} - 273.15, \quad (3.4)$$

kde $T_{[C]}$ je nová teplota v °C a $T_{[K]}$ teplota v K.

3. PRŮZKUM MOŽNÝCH ŘEŠENÍ

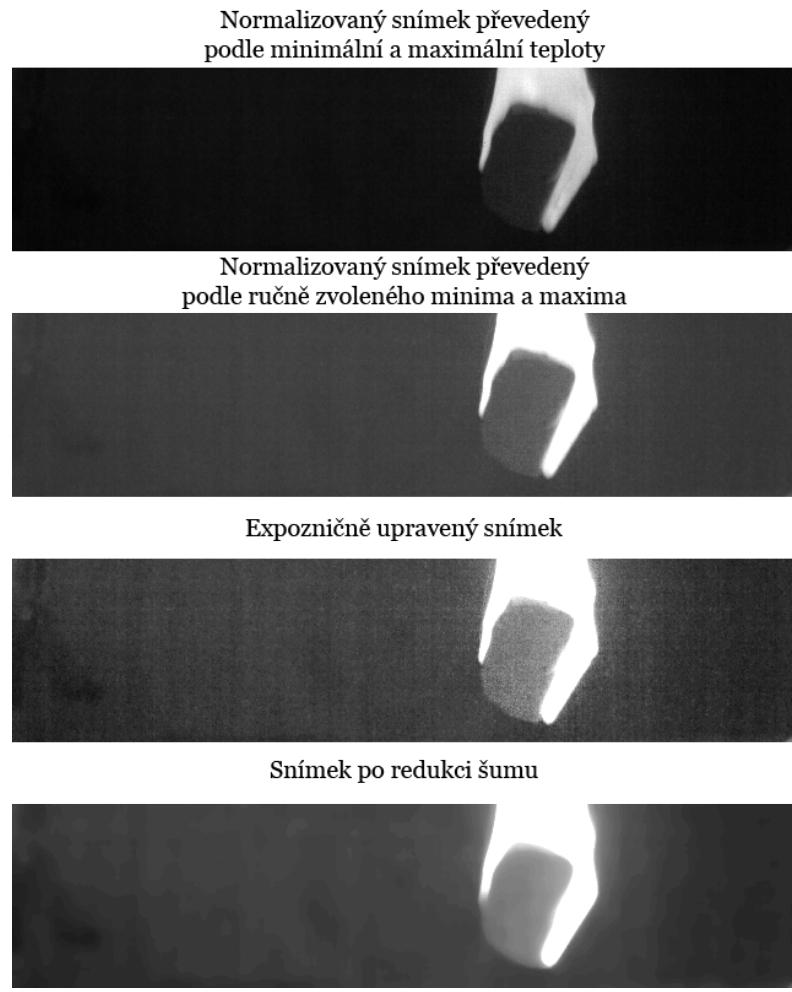


Obrázek 3.4: Fotografie umístění kamery

Tyto teplotní data je nutné normalizovat do nového rozsahu, aby bylo možné je rozumně zobrazit. Nejjednodušším způsobem je ve snímku vyhledat minimální a maximální teplotu a všechny teploty normalizovat pomocí vzorce:

$$x_{norm} = \frac{(newMax - newMin)}{(oldMax - oldMin)} * (x - oldMax) + newMax, \quad (3.5)$$

kde x je hodnota před normalizací, x_{norm} je nová hodnota, $oldMax$ a $oldMin$ jsou nalezené minimální a maximální teploty v snímku a $newMax$ a $newMin$ je nový rozsah hodnot. Většina algoritmů z knihovny OpenCV pracuje pouze s 8-bitovými maticemi, takže je vhodné data převádět rovnou do 8-bitové podoby. Z toho plyne, že hodnota $newMax$ je vždy 255 a $newMin$ vždy 0.



Obrázek 3.5: Jednotlivé kroky předzpracování snímku

Způsob normalizace snímku pomocí nalezené minimální a maximální teploty se ukázal jako nevhodný, protože zboží bylo stěží rozeznatelné i pro člověka. Hledání hodnot *oldMax* a *oldMin* tedy není automatizováno a jejich volba probíhá ručně.

3.2.2 Úprava expozice a odstranění šumu

Normalizovaný snímek je pro zvýraznění teplotních rozdílů expozičně upraven. Proběhne zvýšení kontrastu a následné snížení jasu. Tímto procesem se na snímku vygeneruje velké množství šumu, které je potřeba zredukovat.

Šum je odstraněn nejprve pomocí mediánového filtru [48] a následně ještě pomocí bilaterálního filtru [49], který zachovává hrany.

3.3 Řešení segmentací na základě teplot objektů

Prvním navrženým řešením je segmentace obrazu pouze na základě teplot. Toto řešení se dá považovat za velmi naivní, ale nemusí být hned jasné, v čem spočívá problém.

Ještě před zpracováním snímku je možné z jednotlivých pixelů získat konkrétní teploty. V oblasti zájmu se může objevit pouze ruka, zboží a podlaha. Pro zamýšlení se nad dalším postupem je nutné zjistit rozmezí teplot těchto tří objektů.

Teplotně se ruka pohybuje přibližně od 26°C a není zde moc co vymýšlet. Problém nastává s určením rozsahu teploty zboží. Zboží se totiž v průběhu manipulace zahřívá od ruky a nebo naopak ochlazuje okolním vzduchem. Pokud se teplota v místnosti změní, některé zboží na tuto změnu reaguje rychleji a některé pomaleji v závislosti na tepelné vodivosti. Zboží má oproti podlaze teplotu průměrně o $0.2 - 0.5^{\circ}\text{C}$ vyšší, ale může se stát, že některé zboží bude mít v některých chvílích i nižší teplotu než má podlaha (díky tepelné vodivosti a změně teploty vzduchu).

Dalším faktorem je nestabilita výstupu z termokamery. Mezi jednotlivými korekcemi NUC (2.6.1) se teplota s postupem času nelineárně mění. Tento posun může být v některých situacích až 1°C za minutu.

Najít přesný teplotní rozsah pro zboží není možné a tato úloha se tímto způsobem řešit nedá.

3.4 Řešení s využitím hranové detekce

Druhým navrženým řešením je využít hranovou detekci pro získání hran objektu a prahování pro získání ruky. Prvním krokem je získání masky ruky z předzpracovaného snímku. Masku je možné získat více způsoby. Byly experimentálně vyzkoušeny tyto postupy: binární prahování, Otsu prahování, adaptivní prahování, segmentace pomocí hranové detekce. Ukázalo se, že čím složitější metoda byla využita, tím byly výsledky horší, protože tyto adaptivní algoritmy se pro povahu úlohy nehodí.

Pokud předzpracovaný snímek obsahuje ruku, tak si lze na obrázku 3.5 všimnout, že pixely ruky jsou velmi jasné. Začínají průměrně od hodnoty 180 (v rozmezí 0 černá a 255 bílá) a je snadné je rozlišit od tmavého pozadí. Pro získání masky ruky je využito jednoduché, ale velmi rychlé, binární prahování s hodnotou prahu 180.

Na snímku je dále patrné, že kolem ruky vzniká záře, která nenese žádnou informaci a lze ji považovat za šum. Tento šum se dá velmi obtížně redukovat, protože vzniká kolem celé oblasti ruky a ovlivňuje i zboží. Alespoň pro částečné odstranění této záře je využito morfologické operace dilatace [33] s obdélníkovým jádrem. Je však nutné počítat s tím, že dilatace zmenšuje oblast pro zboží.

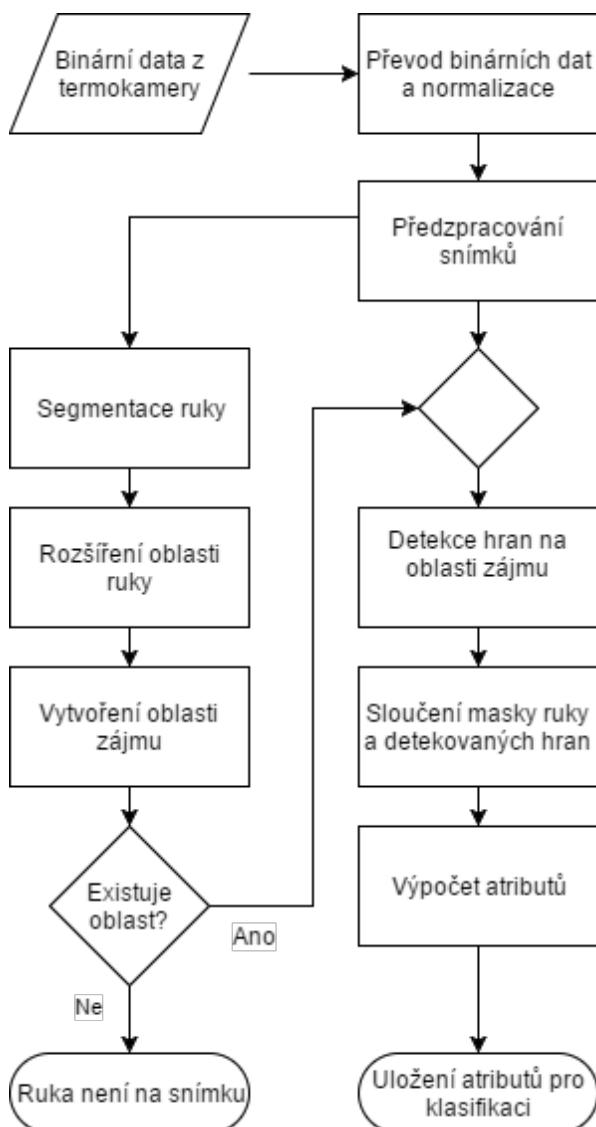
Aby byly hledány pouze relevantní hrany, je na snímku vyhledána oblast zájmu a ohrazena pomocí ohraňujícího obdélníku (bounding boxu). Oblast zájmu je nalezena pomocí vyhledání největší kontury. Pomocí největší kontury se pak řídí šířka oblasti, která je ještě vhodně zvětšena. Výška oblasti je vždy celá velikost snímku, protože zboží dost často přesahuje právě před rukou. Pokud není tato oblast nalezena, ruka se na snímku nevyskytuje a nemá smysl dále vyhodnocovat.

V případě, že je nalezena oblast zájmu, je dalším krokem vyhledání hran. Stejně jako u způsobu segmentace ruky bylo pro detekci hran experimentálně vyzkoušeno několik detektorů (Prewittův, Sobelův, Sacharrův, Laplaceův, Cannyho) s různými způsoby nastavení a předzpracování obrazu. Ve všech případech poskytoval Cannyho hranový detektor [50] nejpřesnější výsledky. V případech velmi malé jasové změny (teplota zboží blízko teplotě podlahy), nacházel též Cannyho detektor nerelevantnější hrany.

Na nalezené hrany je dále v několika iteracích aplikována dilatace a eroze pro spojení blízkých hran a odstranění hran malých, které jsou pravděpodobně šum.

Výsledný snímek hran zboží vznikne jako binární součet snímku hran a inverze binární masky ruky. Na tomto snímku se pak vypočtou předem dané atributy kontur (uvedeny v následující kapitole 4.2), podle kterých je možné ho klasifikovat. Algoritmus je též graficky znázorněn ve vývojovém diagramu 3.6.

3. PRŮZKUM MOŽNÝCH ŘEŠENÍ



Obrázek 3.6: Vývojový diagram algoritmu s využitím hranové detekce

3.4.1 Problémy

Toto řešení se mimo jiné ukázalo jako velmi citlivé na pohyb ruky. V případě, že se ruka pohybuje rychle, na snímku se projeví pohybová neostrost, která díky nižší jasové hodnotě než má ruka, není segmentována. Následně se po aplikaci Cannyho detektoru jeví artefakty pohybové neostrosti ruky jako hrany zboží a tyto snímky jsou chybně klasifikovány. Problém s nadbytečnými hranami způsobuje i záře, která se vyskytuje kolem ruky.

3.4.2 Vyhodnocení

Výsledky tohoto řešení se pohybovali v rozmezí 65 - 75 % celkové přesnosti klasifikace v závislosti na použitém klasifikátoru. Detailnější informace nejsou dále diskutovány, protože toto řešení nebylo shledáno jako konečné a bylo navrženo lepší, které je následně detailněji vyhodnoceno.

3.5 Řešení pomocí dynamického odčítání pozadí

Toto řešení bylo s využitím nových znalostí a postřehů navrženo pro zlepšení předchozích výsledků.

V paměti je uchováváno průběžně aktualizující se pozadí. Je to z důvodu, že obraz z termokamery není plně stabilní a mezi jednotlivými snímky se v čase mění. Největší změna v obraze se vždy projeví po kalibraci NUC a je na ni nutné co nejrychleji reagovat. Z toho důvodu nebylo použité pouze průměrování pozadí, ale je využito akumulovaní pozadí s předem danou váhou. Při akumulovaném pozadí jsou jednotlivé pixely nového pozadí vypočteny jako:

$$acc(x, y) = (1 - \alpha) \times acc(x, y) + \alpha \times src(x, y), \quad (3.6)$$

kde $src(x, y)$ je hodnota vstupního pixelu na souřadnicích (x, y) , $acc(x, y)$ je hodnota pixelu akumulovaného pozadí v paměti a α je váha nového pixelu. Čím větší je α , tím je schopné se pozadí rychleji adaptovat na změnu prostředí. Pro α byly testovány hodnoty od 0,1 až do 0,5 a nakonec se pro tento typ úlohy ukázala jako nejlepší hodnota 0,3.

Od aktuálního snímku je odečteno pomocí [30] průběžně aktualizující se pozadí a z této operace vznikne maska popředí. Tato maska je použita k získání reálného popředí. Na získané popředí je stejně jako v řešení 3.4 aplikováno binární prahování pro zisk masky ruky a tato maska je dilatací rozšířena. V případě, že žádná maska ruky neexistuje, je snímek použit pro výpočet nového pozadí. Efekty těchto kroků je možné vidět na obrázku 3.7.

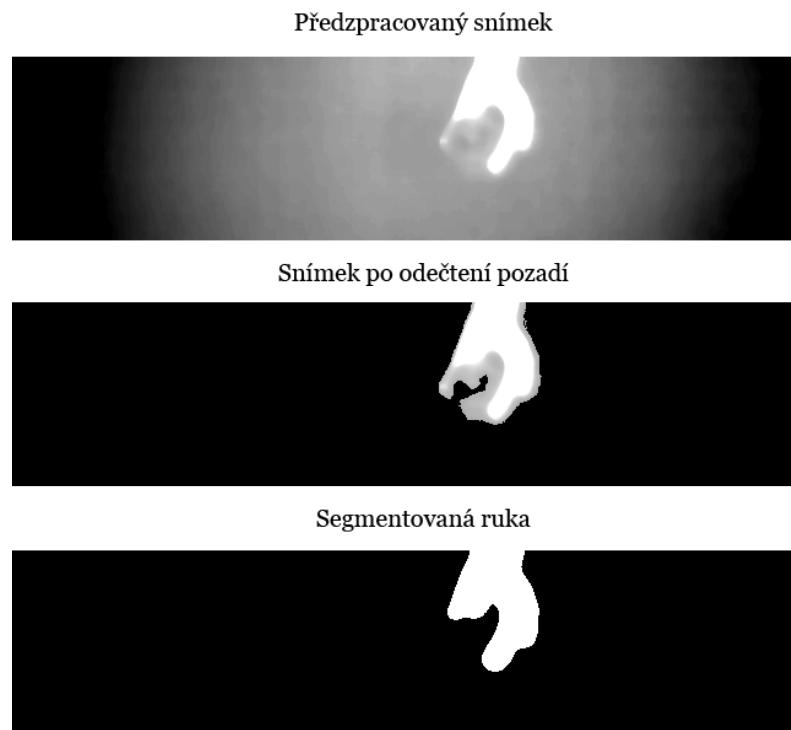
Maska ruky je invertována a použita pro maskování popředí. Tímto způsobem na snímku zůstane šum a možné zboží, které je nutné nějakým způsobem zpracovat. Prvním aplikovaným způsobem je odstranění možného zboží od šumu ve formě záře kolem ruky. Tento postup spočívá v nalezení kontur, jejich approximaci pomocí algoritmu Douglas-Peucker [51], nalezení konvexní obálky a vyhledání konvexních defektů. Mezi konvexními defektami je nalezen ten s největší hloubkou, který se pravděpodobně nachází v oblasti mezi palcem a ostatními prsty. Jednotlivými body defektů jsou protnuty úsečky černé barvy a vhodné velikosti, které v některých případech zajistí oddělení šumu od zboží. Tento postup je ilustrován na obrázku 3.8.

Na oddělené fragmenty obrazu jsou dále v několika krocích vhodně aplikovány morfologické operace a to eroze a poté dilatace. Tento snímek může být zbylé zboží nebo šum a jsou na něm vypočteny předem dané atributy pro klasifikaci. Algoritmus je také popsán ve vývojovém diagramu 3.6.

3.5.1 Úskalí řešení

Největší nevýhodou tohoto řešení je, že je silně závislé na uloženém pozadí v paměti. V případě, že proběhne NUC kalibrace a do paměti se nestáčí načíst dostatečný počet snímků nového pozadí, proběhne operace odečtení pozadí chybně. V tomto případě jsou pak vypočteny nesmyslné hodnoty atributů,

3.5. Řešení pomocí dynamického odčítání pozadí



Obrázek 3.7: Ilustrace operací získání popředí snímků a segmentace ruky

protože se jako zboží tváří skoro celá oblast snímku. Pro ilustraci je tento případ zobrazen v obrázku 3.10.

Dalším problémem je, stejně jako v algoritmu 3.4, rychlý pohyb ruky při kterém se projevuje pohybová neostrost, která rozšiřuje pomyslnou oblast zboží.

3.5.2 Vyhodnocení

Toto řešení bylo vybráno jako finální a jeho vyhodnocení je možné detailněji nalézt v následující kapitole 4.

3. PRŮZKUM MOŽNÝCH ŘEŠENÍ

Snímek s vyznačenými konvexními body



Snímek po maskování popředí maskou ruky



Oddělovací čáry v konvexních bodech (pro ilustraci bílé barvy)

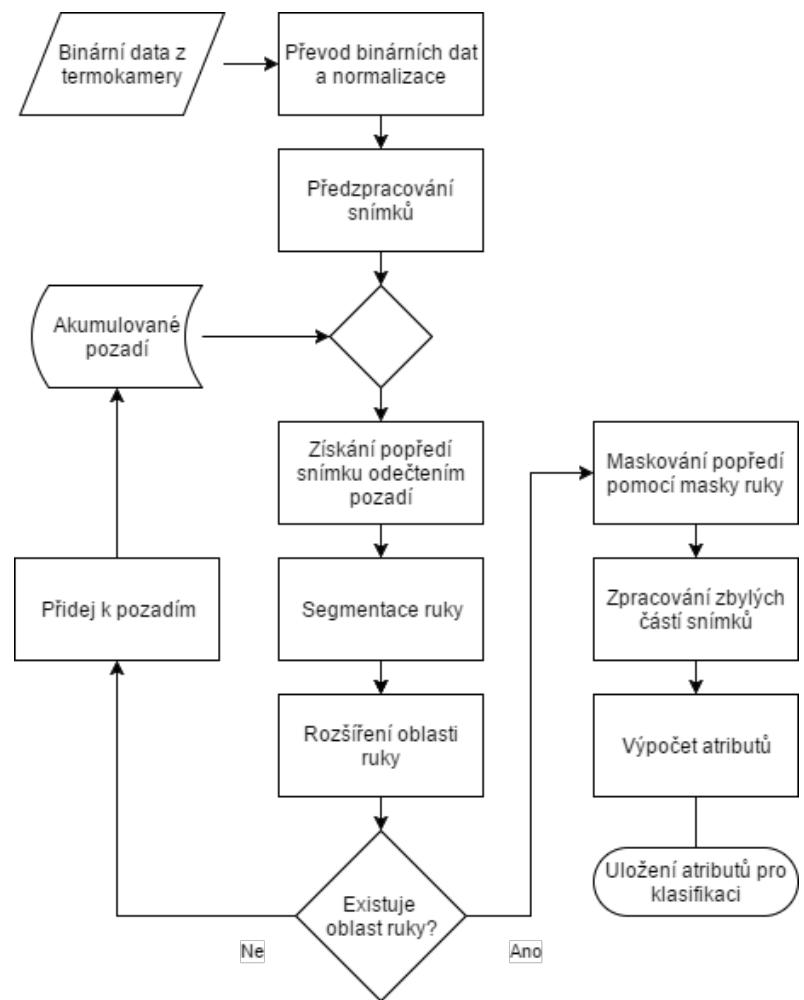


Výsledný snímek po aplikaci morfologických operací



Obrázek 3.8: Proces zpracování snímku zboží

3.5. Řešení pomocí dynamického odčítání pozadí



Obrázek 3.9: Vývojový diagram algoritmu dynamického odečítání pozadí

3. PRŮZKUM MOŽNÝCH ŘEŠENÍ

Pozadí uchovávané v paměti



Aktuální snímek (expozničně světlejší)



Chybné popředí po odečtení pozadí



Obrázek 3.10: Chybně získané popředí kvůli změně expozice aktuálního snímku

3.6 Řešení s kombinací RGB obrazu

U všech řešení pouze s termokamerou je největším problémem rozlišit pozadí (podlahu) od zboží kvůli velmi podobným teplotám a záři, která vzniká kolem ruky. V případě, že by bylo možné striktně oddělit pozadí od popředí (ruka se zbožím), celý problém by se značně zjednodušil, protože segmentovat zboží od ruky lze díky rozdílným teplotám velmi snadno.

V tomto ohledu by bylo lepší využít klasickou kameru s viditelným spektrem pro získání masky popředí. Tuto masku pak aplikovat na termovizní snímek a následně na něm segmentovat ruku a detektovat zboží. Tento snímek pak ještě vhodně zpracovat a vypočítat na něm atributy ke klasifikaci.

Pro otestování tohoto řešení byla využita externí RGB webkamera a ukázalo se, že toto řešení přináší dva velmi podstatné problémy.

3.6.1 Synchronizace kamer

Prvním problémem je sladění snímků z různých kamer. Kamery spolu nejsou nijak propojeny a tak procesorové hodiny tikají na každé kameře jinak. Pořízené snímky v čase jsou rozdílné a je možné nalézt snímky, které při pohybu ruky neodpovídají žádným snímkům z druhé kamery. Jednoduše řečeno, kamery nemají synchronizované snímání a tak je u každé kamery snímek vyhotoven v jiný čas. Navíc u toho u termokamery vstupuje kalibrace NUC, při které kamera žádné snímky nepřenáší a občas se objevují i chyby přenosu, kdy požadovaný snímek nemusí být vůbec obdržen.

Snímky nelze přesně synchronizovat již po přijmutí z kamer, protože by tímto způsobem vznikalo mnohem více chyb a výsledný algoritmus by byl méně přesný než v případě jedné kamery.

Synchronizaci snímků je možné řešit pouze hardwarově. Některé průmyslové kamery bývají vybaveny konektory, které slouží právě k synchronizaci více kamer. Je možné propojit několik kamer, kdy synchronizace snímání probíhá vůči jedné vybrané kameře (master, slave režim). Synchronizaci je možné konfigurovat pomocí protokolu GenICam. Termokamera dostupná v práci má k dispozici synchronizační konektor, ale nebyla k dispozici druhá kamera, která by s ní byla kompatibilní.

3.6.2 Perspektiva snímků

Druhým problémem je rozdílná perspektiva snímků. Různé kamery nikdy nebudou mít stejnou optickou soustavu. Tento fakt je způsoben různými optickými vadami objektivů a často se objektivy liší i v ohniskové vzdálenosti. Tyto skutečnosti pak ovlivňují výsledné zarovnání snímků. Na obrázku 3.11 lze nejdříve vidět rozměrově neupravený snímek z webkamery, který je následně oříznutý na oblast termokamery 3.12. V posledním obrázku 3.13 jsou

3. PRŮZKUM MOŽNÝCH ŘEŠENÍ

oba snímky prolnuty a lze snadno pozorovat, že i přes stejnou scénu na sebe snímky, právě díky rozdílné perspektivě, nelze přesně zarovnat.



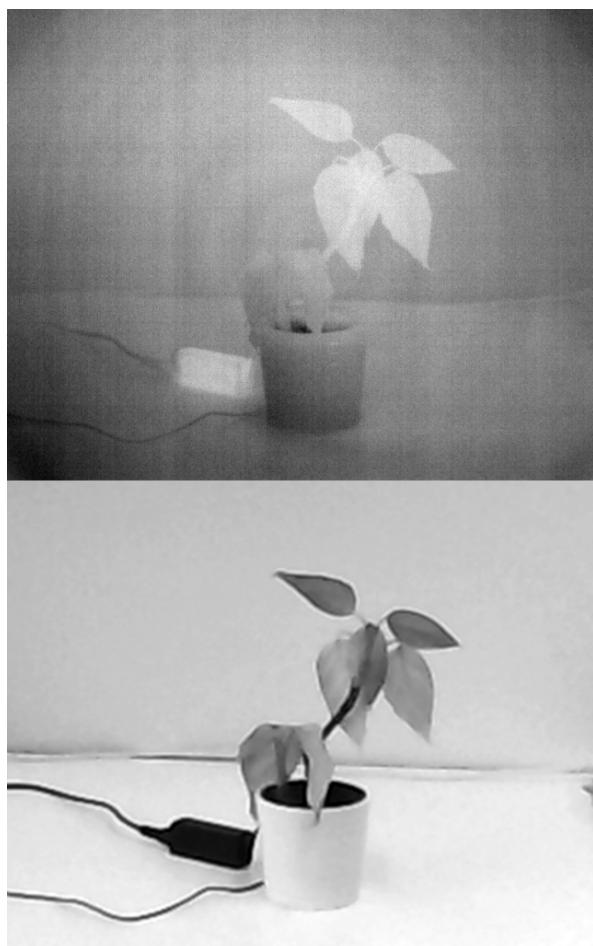
Obrázek 3.11: Ukázkový snímek snímaného obrazového pole webkamery

Řešením je nalezení transformační matic homografie pro mapování obrazu webkamery na plochu obrazu termokamery. V problému detekce zboží v ruce to znamená hledání několika transformačních matic, protože perspektiva se mění v závislosti na vzdálenosti snímaného objektu a ruka na snímku se může objevit v několika rovinách. Čím více je nalezeno transformačních matic pro různé vzdálenosti, tím jsou výsledky mapování přesnější. V knihovně OpenCV je hledání transformační matice homografie již implementováno metodou findHomography z balíku calib3d.

3.6.3 Shrnutí

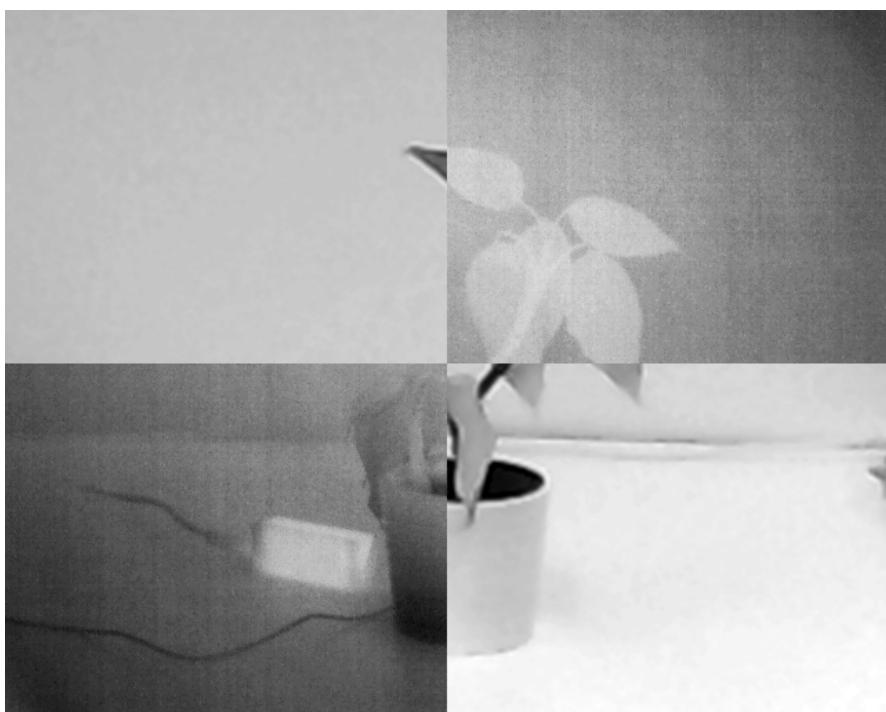
Toto řešení nakonec nebylo implementováno hlavně kvůli problémům se synchronizací kamer, které v rámci práce nebylo možné vyřešit. Pro případné navázání na tuto práci je alespoň uveden vývojový diagram 3.14 algoritmu.

3.6. Řešení s kombinací RGB obrazu

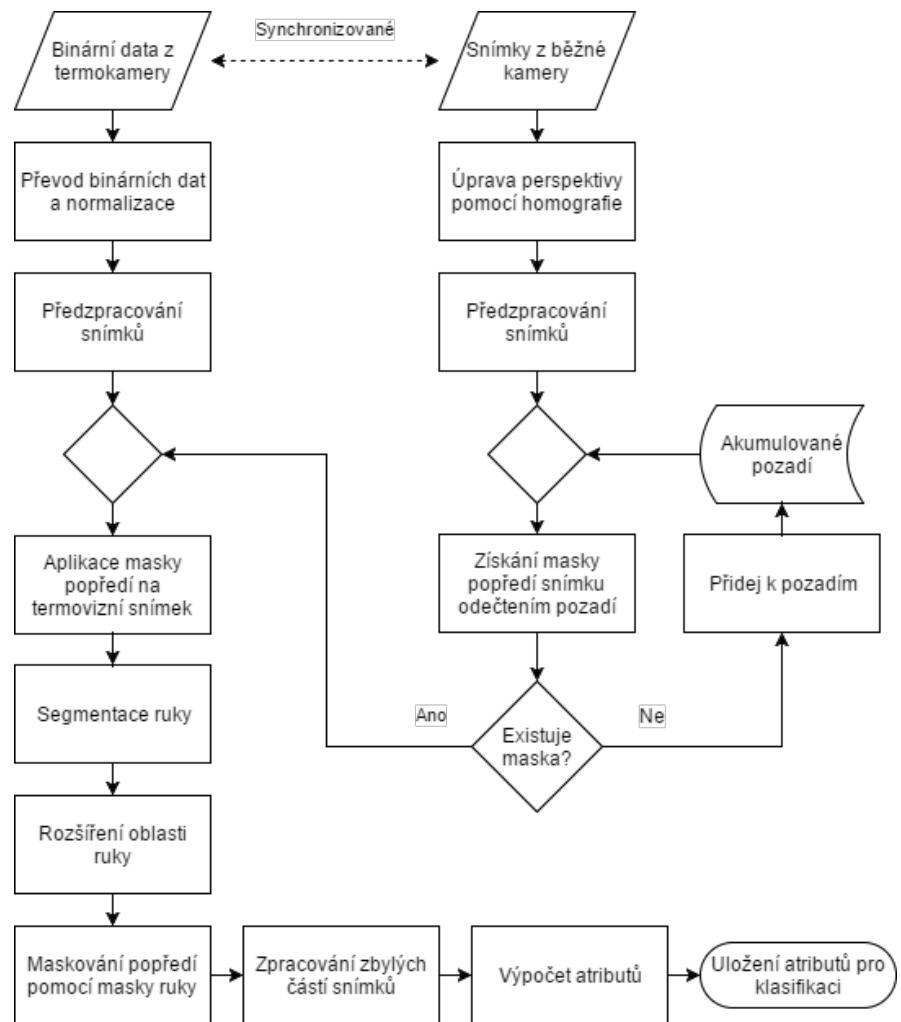


Obrázek 3.12: Snímek z webkamery oříznutý k obrazovému poli termokamery

3. PRŮZKUM MOŽNÝCH ŘEŠENÍ



Obrázek 3.13: Ručně sloučené snímky z webkamery a termokamery



Obrázek 3.14: Vývojový diagram algoritmu využívající kombinaci běžného a termografického snímku

Zvolené řešení a jeho vyhodnocení

Finálním vybraným řešením pro vyhodnocení je řešení s termokamerou a algoritmem dynamického odčítání pozadí. Detaily algoritmu jsou uvedeny v 3.5 a zde je uvedeno pouze jeho stručné shrnutí.

4.1 Shrnutí vyhodnocovaného řešení

V paměti je průběžně aktualizováno pozadí, vůči kterému se vyhodnocuje právě načtený snímek. V případě, že je snímek detekován jako pozadí, aktualizuje se s jeho pomocí to v paměti. V opačném případě je získané popředí snímku, na kterém se dále segmentuje ruka. Pomocí invertované masky ruky je získáno popředí bez ruky, na kterém proběhne redukce artefaktů ve formě šumu a jako poslední krok je výpočet atributů pro klasifikaci.

4.2 Volba atributů

Pro řešení bylo zvoleno celkem 20 atributů zaměřující se na vlastnosti kontur. Atributy související s jasem, histogramem a podobnými vlastnostmi obrazu nebyly použity, protože výstup z kamery se v čase mění a není možné zajistit jeho stabilitu.

Zvolené atributy jsou rozděleny do dvou skupin. Na atributy týkající se vlastností největší nalezené kontury a souhrnné atributy všech nalezených kontur. Atributy vypočítávány na největší nalezené kontuře jsou:

Délka kontury (Perimeter)

Plocha kontury (Area)

Minimální průměr vepsané kružnice (Minimum diameter)

4. ZVOLENÉ ŘEŠENÍ A JEHO VYHODNOCENÍ

Maximální průměr vepsané kružnice (Maximum diameter)

Délka konvexní obálky (Convex perimeter)

Plocha konvexní obálky (Convex area)

Formfaktor (Formfactor) získán jako

$$\frac{4 \times \pi \times \text{Area}}{\text{Perimeter}^2} \quad (4.1)$$

Kulatost (Roundness) získána jako

$$\frac{4 \times \text{Area}}{\pi \times \text{Maximum diameter}^2} \quad (4.2)$$

Poměr stran (Aspect ratio) získán jako

$$\frac{\text{Maximum diameter}}{\text{Minimum diameter}} \quad (4.3)$$

Konvexita (Convexity) získána jako

$$\frac{\text{Convex perimeter}}{\text{Perimeter}} \quad (4.4)$$

Pevnost (Solidity) získána jako

$$\frac{\text{Area}}{\text{Convex area}} \quad (4.5)$$

Kompaktnost (Compactness) získána jako

$$\frac{\sqrt{\frac{4}{\pi} \times \text{Area}}}{\text{Maximum diameter}} \quad (4.6)$$

Rozsah (Extent) získán jako

$$\frac{\text{Area}}{\text{Maximum diameter} \times \text{Minimum diameter}}. \quad (4.7)$$

Atributy vypočítávané na všech nalezených konturách jsou:

Počet kontur

Součet ploch kontur

Minimální plocha kontury

Průměrná plocha kontury

Součet délek kontur

Minimální délka kontury

Průměrná délka kontury.

4.3 Rozdělení snímků do tříd

Před klasifikací je nejdříve nutné snímky rozdělit do několika skupin. Klasifikačními skupinami jsou prázdná ruka a ruka se zbožím, ale pro správnou funkčnost uvedeného algoritmu je ještě nutné v paměti uchovávat pozadí. Čím více snímků pozadí bude při vyhodnocování k dispozici, tím budou výsledky přesnější.

Snímky jsou uchovávány jako binární soubory obsahující data v podobě signálů, takže je nutno vytvořit speciální aplikaci pro jejich zobrazení a následné rozřazení do tříd. Pomocí této aplikace bylo ručně anotováno celkem 910 snímků prázdné ruky a 973 snímků ruky se zbožím. Všechna anotovaná pozadí byla získána pomocí aplikace s algoritmy a ručně zkонтrolovány pomocí aplikace pro anotování. Těchto snímků pozadí je celkem 2173.

Navzdory tomu, že byly při práci využity dvě kamery, budou klasifikovány snímky pouze z kamery (13 mm). Je to z toho důvodu, že v nasnímaných datasetech kamery (25 mm) se vyskytuje množství chyb, které vznikaly nevypočitatelně, kvůli její technické závadě.

Celkem se tedy bude vyhodnocovat 1244 anotovaných snímků z kamery (13 mm), z toho 741 snímků ruky a 503 snímků ruky se zbožím, s pomocí 2031 snímků pozadí.

4.4 Klasifikace

Před samotnou klasifikací jsou ještě jednotlivé hodnoty atributů normalizovány. Následně je spouštěna detekce anomalií a z datasetu jsou některé záznamy odstraněny. Odstranění probíhá, pokud se součet délek kontur přibližuje celkové možné ploše obrazového pole. Tyto záznamy jsou očividně chybné a vznikají nejčastěji ihned po NUC, kdy ještě není dostatečně aklimatizováno uložené pozadí a při operaci odčítání pozadí se celý snímek bere jako pořadí (ilustrováno v předchozí kapitole 3.5.1). Tyto případy jsou považovány za chyby měření a jsou odstraněny.

Ke klasifikaci je využit program RapidMiner, který poskytuje velké množství různých klasifikátorů a ostatních užitečných nástrojů pro data mining.

4.5 Výsledky

K vyhodnocení výsledků je použita 10-násobná křížová validace (10-fold cross validation) a bylo vyzkoušeno celkem 10 různých klasifikátorů. Výsledky jednotlivých klasifikátorů jsou uvedeny v tabulce 4.1. Jako nejpřesnější klasifikátor se jeví neuronová síť s celkovou přesností 88.66%. Podrobnější výsledky k tomuto klasifikátoru je možné vidět v matici záměn 4.2.

4. ZVOLENÉ ŘEŠENÍ A JEHO VYHODNOCENÍ

Klasifikátor	Prázdná ruka	Ruka se zbožím	Celková přesnost
Naivní bayes	91.29%	78.31%	85.37%
3-NN	88.19%	88.39%	88.27%
5-NN	87.33%	88.57%	87.78%
10-NN	87.40%	89.52%	88.18%
Lineární regrese	91.29%	78.31%	85.37%
SVM	89.36%	85.98%	88.02%
Rozhodovací strom	87.12%	87.39%	87.22%
Rozhodovací les	89.31%	86.83%	88.35%
GBT	86.83%	89.04%	87.62%
Neuronová síť	87.91%	90.39%	88.83%

Tabulka 4.1: Tabulka pozitivní prediktivní hodnoty jednotlivých tříd (class precision) a celkové přesnosti (accuracy) jednotlivých klasifikátorů

Klasifikace	Skutečnost		<i>class precision</i>
	prázdná ruka	ruka se zbožím	
prázdná ruka	691	95	87.91%
ruka se zbožím	44	414	90.39%
<i>class recall</i>	94.01%	81.34%	

Tabulka 4.2: Matice záměn neuronové sítě vyhodnocena pomocí 10-násobné křížové validace

4.6 Příklady klasifikovaných snímků

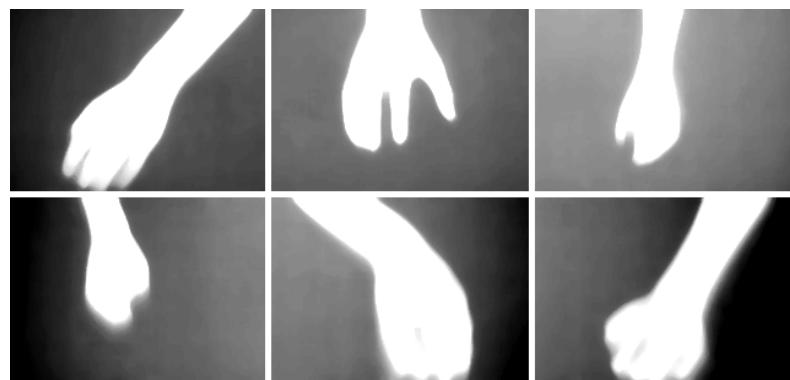
V této části je uvedeno několik názorných příkladů klasifikace, tak jak je určila neuronová síť. Pro větší názornost jsou tyto snímky již předzpracovány. Na jednotlivých snímcích si lze též všimnout, jak se snímané prostředí poměrně zásadně mění. Jsou uvedeny příklady správně klasifikovaných (4.1, 4.3), ale i chybně klasifikovaných (4.2, 4.4).

4.7 Komentář ke klasifikovaným snímkům

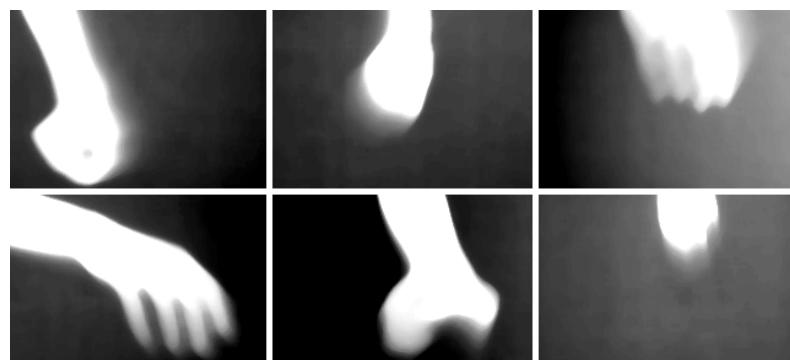
Obecně by mělo být snadné detektovat ruku, protože na její teplotu se dá vždy spolehnout. Problém nastává při jejím rychlém pohybu, kdy pohybová neostrost zabrání jejímu správnému segmentování. Většina chybně klasifikovaných rukou jsou pohybově neostré snímky, případně snímky, před kterými došlo ke změně pozadí, kdy se uložené pozadí ještě nestihlo adaptovat této změně.

Z chybně klasifikovaných snímků ruky se zbožím je zřejmé, že některé zboží mělo teplotu podobnou podlaze a na snímku je velmi obtížně rozlišitelné. Druhým případem je zboží, které je sice viditelné, ale má menší rozměry, takže

4.7. Komentář ke klasifikovaným snímkům



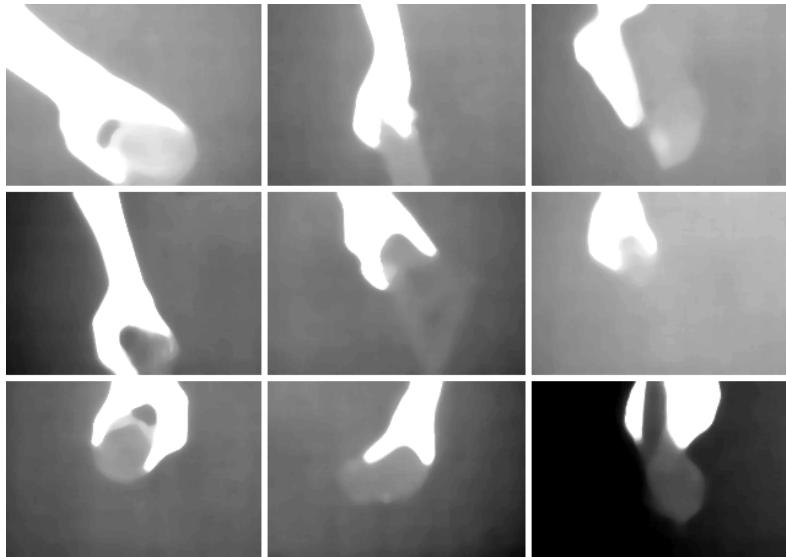
Obrázek 4.1: Příklady správně klasifikovaných snímků kategorie prázdná ruka



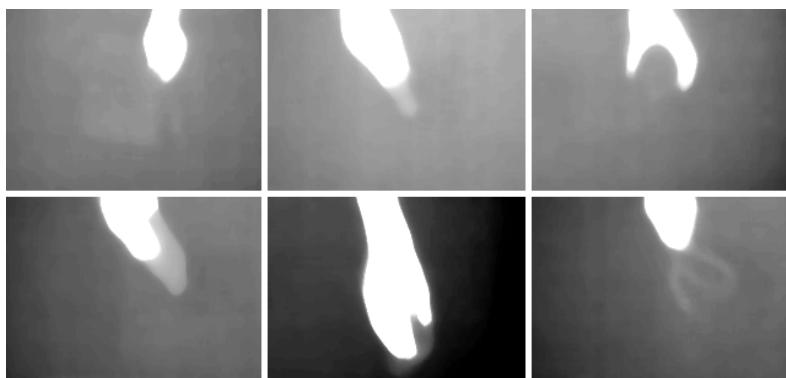
Obrázek 4.2: Příklady klasifikovaných snímků jako ruka se zbožím, ale ve skutečnosti prázdná ruka

po aplikaci morfologických operací již není rozpoznatelné.

4. ZVOLENÉ ŘEŠENÍ A JEHO VYHODNOCENÍ



Obrázek 4.3: Příklady správně klasifikovaných snímků kategorie ruka se zbožím



Obrázek 4.4: Příklady klasifikovaných snímků jako prázdná ruka, ale ve skutečnosti ruka se zbožím

Implementované nástroje

V rámci práce bylo nutno vytvořit několik nástrojů, bez kterých by práce nemohla být splněna. Tato kapitola se tedy věnuje obecnému popisu nástrojů a implementačním detailům detekční aplikace. Ke každému nástroji jsou dostupné zdrojové kódy, které je možné nalézt na přiloženém paměťovém médiu a také v online repozitářích [52, 53, 54].

5.1 Upravený eBUS Player

Tento nástroj slouží ke konfiguraci a získávání dat z termokamery. Proti jeho původní verzi byl modifikován přidáním síťové komunikace pro přenos dat. Původní aplikaci lze nalézt v ukázkových kódech nástroje eBUS SDK. Zdrojový kód je psán v jazyce C++ a je poměrně složitý k pochopení (obsahuje přibližně 50 tříd), proto zde nebude více popisován.

5.2 Aplikace s detekčními algoritmy

Tuto aplikaci lze považovat za hlavní a obsahuje implementaci algoritmů pro zpracování vstupního obrazu z termokamery. Aplikace využívá algoritmy z knihovny OpenCV. Obraz může být buď přijímán skrze soket rovnou z kamery (pomocí modifikovaného eBUS Playeru), nebo načten z předem definované složky.

Aplikace je koncipovaná tak, aby v ní bylo možné algoritmy co nejlépe odladit. Obsahuje tlačítko pro pozastavení na vybraném snímku a pomocí uživatelského rozhraní je možné měnit parametry algoritmů a získat ihned náhled výsledku na konkrétním snímku. Důležitou vlastností je ukládání příchozích snímků ze soketu a poté možné jejich přehrávání s předdefinovanou rychlostí. K náhledu snímků slouží čtyři panely pro různé fáze zpracování. V panelech se tedy objevuje: náhledový (předzpracovaný) snímek, snímek po odečtení pozadí, segmentovaná ruka, potencionální zboží.

5. IMPLEMENTOVANÉ NÁSTROJE

K nastavování parametrů algoritmů slouží čtyři záložky, které jsou rozděleny do logických celků. Tyto nastavení je možné exportovat a importovat, využitý je formát XML.

5.2.1 Implementace

Cílem této práce není detailní popis návrhu a architektury, takže je uvedeno pouze pár základních informací, pro případné snadnější navázání na stávající řešení. Implementace je poměrně rozsáhlá a celá aplikace obsahuje přibližně 30 tříd, které jsou logicky rozděleny do několika balíků.

Celé uživatelské rozhraní zastřešuje ovladač MainController, který deleguje požadavky uživatele dál k příslušným třídám. Tyto požadavky jsou nejčastěji na aktuálně využívaný obrazový detektor (AbstractDetector), aktuálně čtený zdroj (DataReciever) nebo na změnu nastavení parametrů algoritmu uložených v paměti (Settings).

5.2.2 AbstractDetector

Třída AbstractDetector je předek všech obrazových detektorů s abstraktní metodou detect(). Potomci této třídy jsou celkem 3 a to:

BackgroundDetector slouží pro detekování snímků pozadí a jejich následné ukládání do definované složky.

MogDetector vybraný a vyhodnocený algoritmus detekce pomocí odčítání pozadí uvedený v 3.5.

EdgeDetector algoritmus detekce využívající Cannyho hranový detektor uvedený v 3.4.

Pro detailnější informace o struktuře těchto objektů je uveden diagram tříd 5.1.

5.2.3 DataReciever

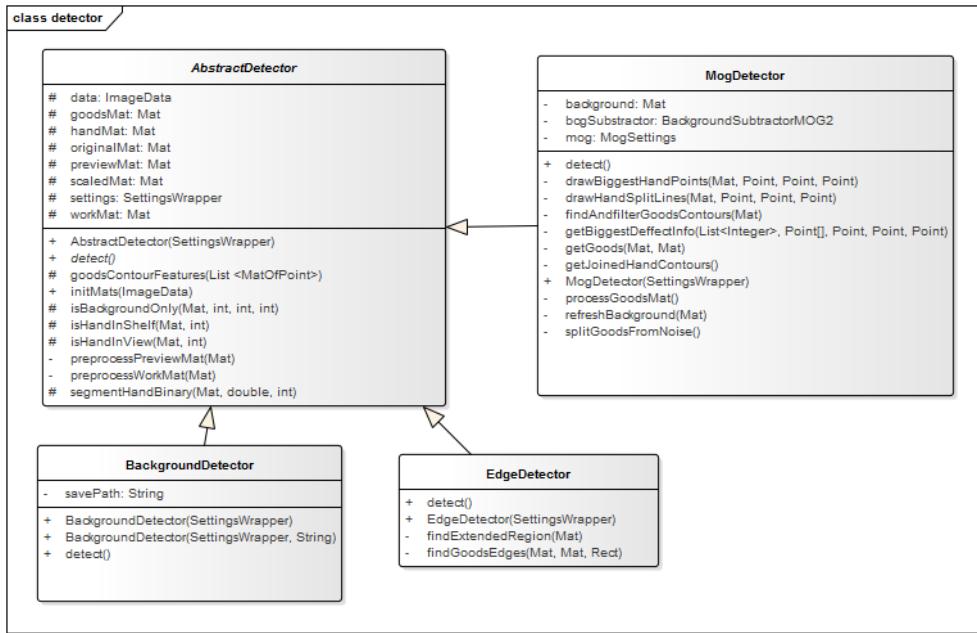
Pro příjem dat z více zdrojů slouží abstraktní generická třída DataReciever. Jedním z potomků této třídy je FlirDataReciever, který je využíván pro příjem dat z termokamery a druhým potomkem je třída pro příjem z webkamery nebo jiné USB kamery. Obě třídy implementují funkčnost ukládání a načtení snímků z úložiště.

5.2.4 Settings

V paměti aplikace je uchováváno aktuální nastavení parametrů využívaných algoritmů. Tyto nastavení jsou rozděleny do čtyř skupin:

PreviewSettings slouží k aktivaci efektů algoritmů na náhledový snímek.

5.2. Aplikace s detekčními algoritmy



Obrázek 5.1: Diagram tříd detekčních algoritmů

PreprocessingSettings obsahuje jednotlivé parametry pro předzpracování obrazu před aplikací algoritmů.

MogSettings jsou nastavení pro algoritmus odčítání pozadí.

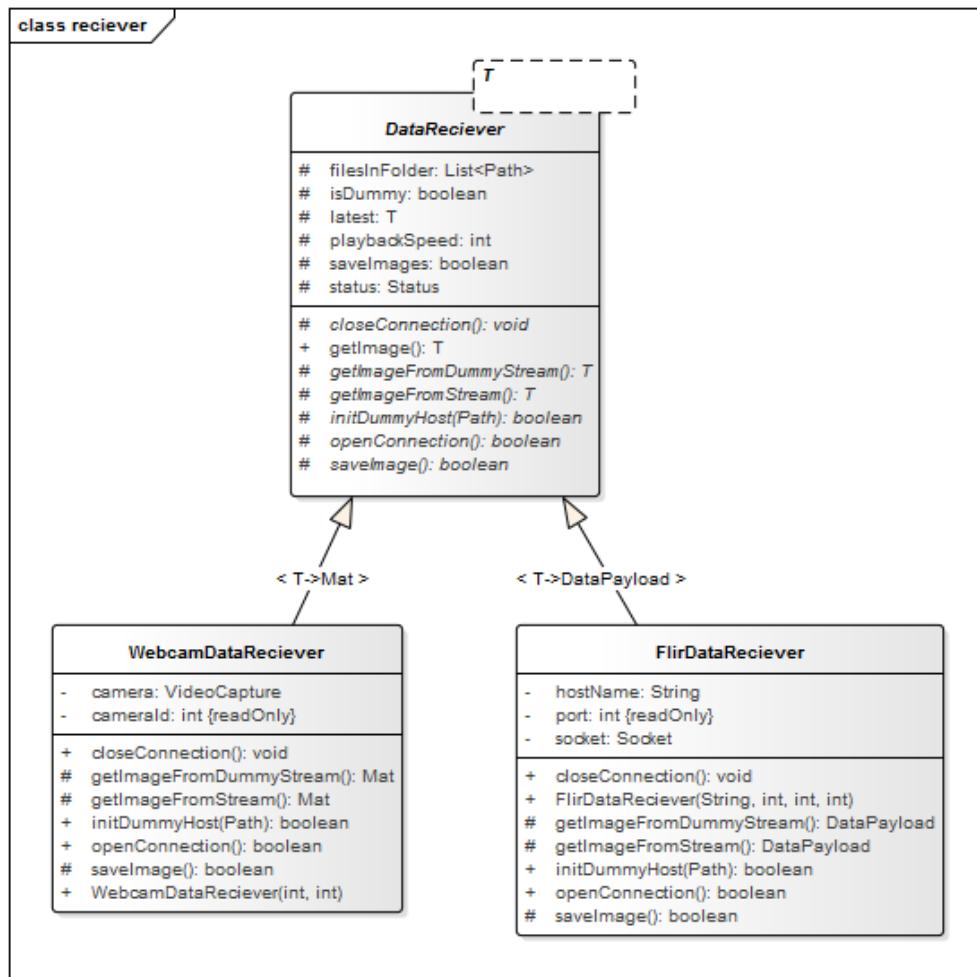
EdgeDetectSettings jsou nastavení parametrů algoritmu hranové detekce. Implementace těchto nastavení není dokončena.

Jednotlivé objekty nastavení jsou k dispozici v objektu **SettingsWrapper**, jehož instance je předávána všem detektorům. Všechna nastavení je možná exportovat do xml pomocí statické třídy **SettingsManager**.

5.2.5 Práce s obrazem

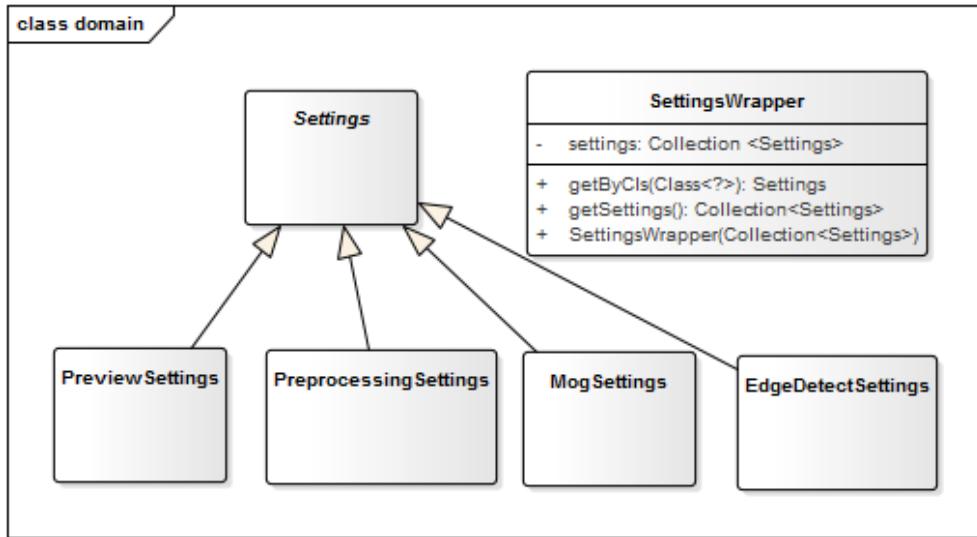
Poslední částí, která stojí za zmínku je balíček **image**, jehož obsahem jsou dvě statické třídy **MatOperations** a **ImageConvertor**. Třída **MatOperations** je pomocná třída pro práci s OpenCV maticemi, která obsahuje volání knihovny, ale oproti běžnému volání nikdy nemodifikuje vstupní data. **ImageConvertor** obsahuje optimalizované metody pro převod mezi binárními daty, OpenCV maticemi a JavaFX obrázkem.

5. IMPLEMENTOVANÉ NÁSTROJE

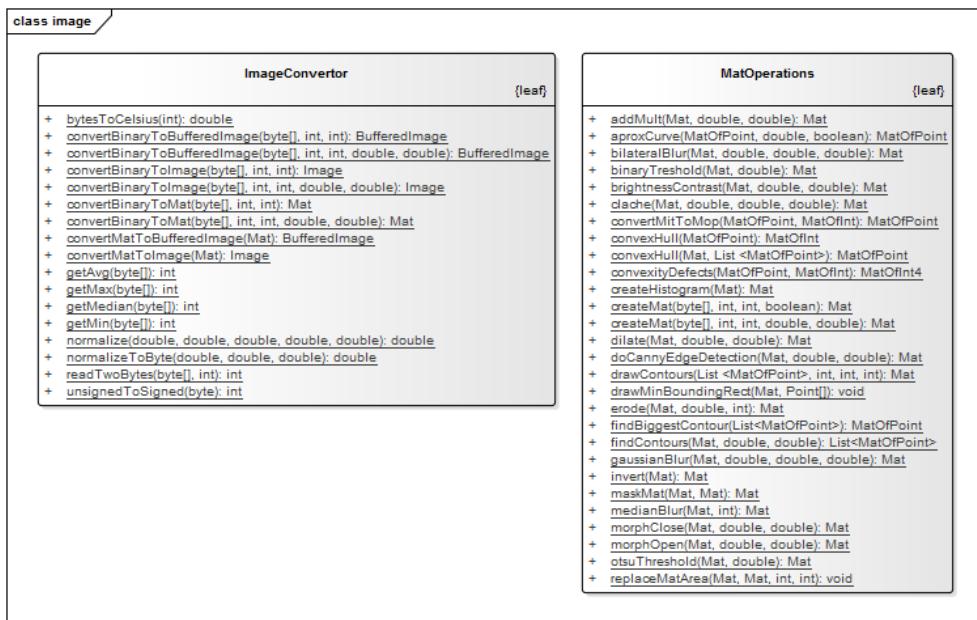


Obrázek 5.2: Diagram tříd zdrojů snímků

5.2. Aplikace s detekčními algoritmy

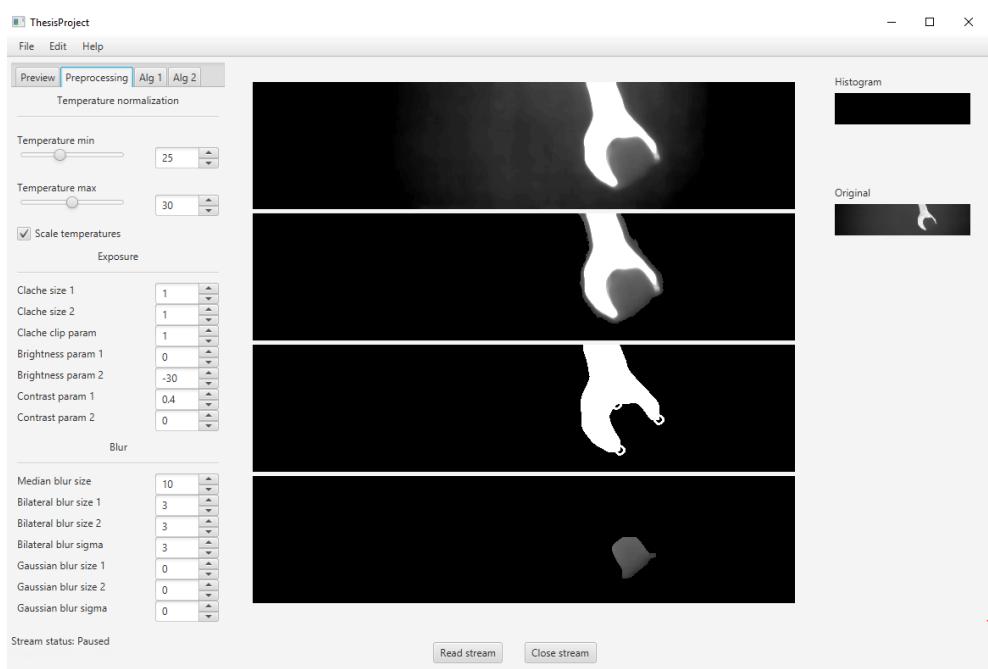


Obrázek 5.3: Diagram tříd nastavení parametrů algoritmů



Obrázek 5.4: Diagram tříd balíčku image

5. IMPLEMENTOVANÉ NÁSTROJE



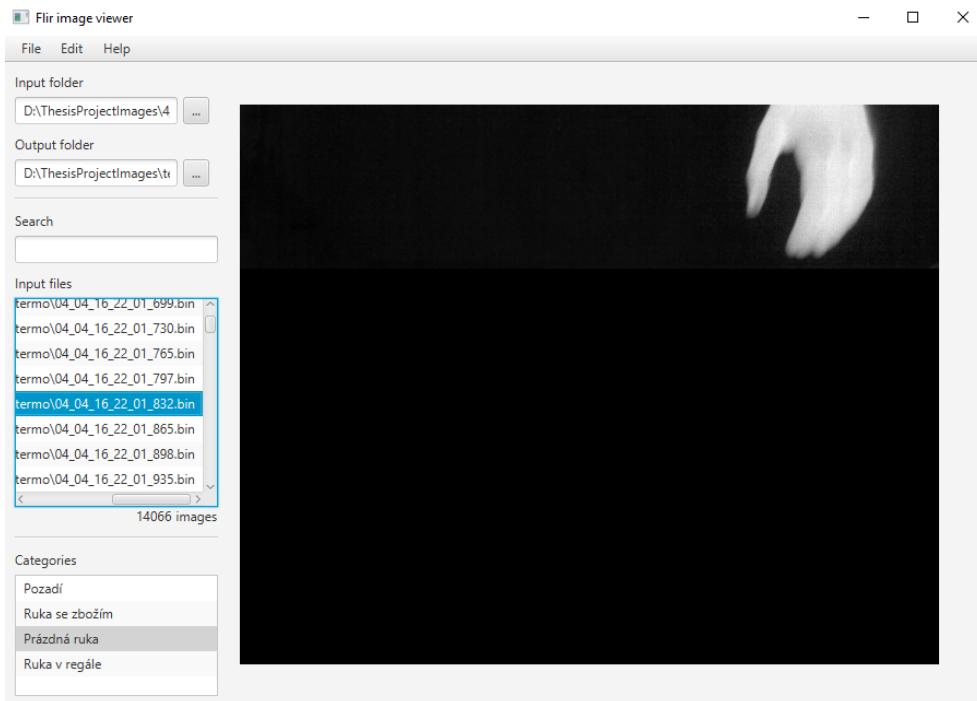
Obrázek 5.5: Aplikace s detekčními algoritmy

5.3 Aplikace pro anotování snímků

Pro potřeby vyhodnocování bylo nutné jednotlivé snímky rozdělit do skupin. K tomu byl vytvořen tento nástroj, který lze snadno ovládat pomocí klávesnice a tím rychle rozřazovat jednotlivé snímky.

Přesouvání mezi jednotlivými snímky probíhá pomocí klávesnicových šipek a přesunutí snímku do předem definované složky pomocí čísla 1-4 v závislosti na kategorii. Snímky je též možné vyhledávat pomocí textového pole a pomocí klávesy D odstraňovat.

V implementaci nejsou žádné složitosti. Aplikace se skládá pouze z jednoho ovladače (MainController), který deleguje požadavky uživatelského rozhraní. Soubory jsou udržovány v objektu InMemoryFiles, který snímky v paměti spravuje a poskytuje je zpět do GUI podle nutnosti. Všechny metody pro převod binárního souboru na snímek poskytuje statická třída ImageConversions.



Obrázek 5.6: Aplikace pro anotování snímků

KAPITOLA **6**

Zhodnocení

Práce se nakonec ukázala o něco složitější, než se na začátku zdálo. Přes všechny možné nástrahy od poruchy původní kamery až nestabilitu obrazu se nakonec povedlo najít rozumné řešení, které vzhledem k dostupným prostředkům poskytuje uspokojivé výsledky.

6.1 Využitá technika

Během práce se bohužel vyskytla závada na původní kameře (25 mm) a bylo nutné využít záložní kameru (13 mm) s jinou ohniskovou vzdáleností a horší snímkovací frekvencí. Tato kamera (13 mm) měla širší úhel záběru, takže ji bylo nutné více přiblížit k regálu. Původní kamera (25 mm) byla pro tuto úlohu vhodnější, protože delší ohnisková vzdálenost přispěla k tomu, že kolem regálu nebyl snímán zbytečně velký prostor. Jedinou nevýhodou delší ohniskové vzdálenosti je, že v snímané oblasti regálu byla menší hloubka ostrosti, než při snímání s kamerou (13 mm).

7.5 snímků za vteřinu u náhradní kamery (13 mm) je obecně velmi málo, ale i přesto se dalo při rozšíření oblasti zájmu dosáhnout poměrně dobrých výsledků. V případě že by bylo vyhodnocování provedeno na funkční kameře (25 mm), byly by výsledky lepší, protože díky vyšší snímkovací frekvenci se algoritmus zvládne rychleji adaptovat na změnu prostředí. Dostačující by byla i kamera se snímkovací frekvencí 15 Hz.

6.2 Navržené řešení

Finální navržený algoritmus se i přes nestabilitu obrazu z kamery jeví, proti ostatním navrženým řešením, jako robustní. Je schopný se poměrně rychle adaptovat na změny prostředí, která vznikají v malé míře při postupném zahřívání detektoru a ve velké míře při kalibraci NUC.

6. ZHODNOCENÍ

Stávající řešení by bylo možné ještě různými způsoby vylepšovat, ale je nutné se zamyslet nad tím, zda se jedná opravdu o nejlepší možný způsob řešení. Adekvátním vylepšením by bylo zaměřit se na stabilizování snímané scény. Jedním z faktorů, které by toto vylepšení ovlivnilo je algoritmus, pro co nejlepší stanovení hodnot *min* a *max* použitých při normalizaci snímku. V tuto chvíli jsou tyto hodnoty nastavovány ručně. Ruční nastavení poskytuje mnohem lepší výsledky, než normalizace pomocí nalezení minimální a maximální teploty v snímku.

Vylepšení, které by přineslo největší přínos je zaměřit se na více obrazových zdrojů. Termokamera se hodí perfektně pro segmentaci ruky, ale segmentovat pomocí termokamery předměty, které mají velmi podobné teploty jako pozadí, se ukázalo jako velmi obtížné. V 3.6 bylo navrženo řešení využívající viditelný a termovizní snímek, ale nakonec nebylo implementováno kvůli vzniklým problémům se synchronizací.

6.3 Implementace řešení

Vytvořená aplikace s algoritmy detekce slouží též jako klient pro příjem snímků z aplikace eBUS Player. Mezi přenosem dat těchto dvou aplikací vzniká jen minimální zpoždění a je možné nad daty pracovat téměř v reálném čase. Toto bylo bráno v potaz při implementaci jednotlivých metod, které jsou maximálně optimalizovány. Případným rozšířením o klasifikátor lze tedy ihned vyhodnocovat. Aplikace je též připravena na příjem více datových zdrojů, jako je například obraz z RGB kamery.

Aplikace má vytvořené uživatelské rozhraní s velmi dobrou správou nastavení parametrů algoritmů. Návrh aplikace je smysluplný, podstatné části jsou komentovány a je tedy snadné ji dále rozšiřovat.

6.4 Výsledky řešení

Výsledky byly lehce ovlivněny tím, že mezi manipulací se zbožím nebyly nechávány dostatečné přestávky. Tyto nedostatečné pauzy způsobily, že při změně obrazového prostředí (NUC, teplotní drift) nebyl získán dostatečný počet snímků pro adaptaci uloženého pozadí. To způsobuje zhoršení výsledků operace získání popředí pozadí, která na scéně zahrne větší oblast popředí než reálně existuje a prázdná ruka pak může být klasifikována jako ruka se zbožím. Tento problém byl bohužel zjištěn až v poslední fázi při vyhodnocování, kdy už byla nasnímána všechna data.

Klasifikace probíhala pomocí nástroje RapidMiner a bylo vyzkoušeno mnoho klasifikátorů a různých nastavení. Ke klasifikaci bylo využito 20 atributů zaměřených na vlastnosti kontur. V případě že by byl k dispozici stabilní obrazový výstup, dávalo by smysl přidat atributy zaměřující se na vlastnosti obrazu. Přidání dalších užitečných atributů by výsledky klasifikace zlepšilo.

K dispozici byl poměrně skromný data set s přibližně 1200 anotovanými snímky. Při využití většího datasetu by byl vytvořen robustnější model. Klasifikovány byly snímky pouze z kamery (13 mm) a celková přesnost klasifikace se pohybovala v rozmezí 85 - 88% v závislosti na využitém klasifikátoru.

6.5 Budoucí využití

V budoucnu by bylo možné řešení ještě dále rozvíjet, protože místa pro zlepšení stále existují. Jedno z budoucích témat by mohlo být například detekce pomocí termokamery a kamery Microsoft Kinect s využitím více obrazových zdrojů. Kombinací obrazových zdrojů by bylo možné stávající řešení zdokonalit, protože každé spektrum přináší o snímaném obrazu nové informace. Bohužel kombinace obrazů z různých zdrojů se ukázalo jako velmi obtížné téma a přináší s sebou spoustu dalších komplikací. Toto téma by se tedy převážně hodilo pro budoucí diplomovou práci.

Závěr

Cílem práce bylo vytvoření algoritmu pro detekci zboží v ruce zákazníka pomocí snímků z termokamery. K tomu bylo nutné se seznámit s technologií termokamer a oborem termografie. V rešeršní části byly nabity znalosti pro návrh vlastního řešení, které následně bylo s jejich pomocí implementováno. Na řešení úlohy bylo nahlíženo různými způsoby a byly navrhnuty celkem čtyři různé algoritmy. Implementovány jsou dva z nich a jsou k nalezení v přiložené aplikaci. Vytvořená aplikace poskytuje uživatelské rozhraní pro snadnou práci s daty a konfiguraci podstatných parametrů aplikovaných algoritmů. Pro vyhodnocení výsledků práce je použito řešení s využitím algoritmu dynamického odcítání pozadí. K vyhodnocování se využívají vlastnoručně naměřená data z Laboratoře pro zpracování obrazu na FIT ČVUT s pomocí termokamery FLIR A65 umístěnou nad regálem. Na základě výsledků algoritmu je provedena diskuze, kde jsou uvedeny souhrnné informace o finálním řešení a jeho možné budoucí vylepšení.

Literatura

- [1] FLIR Systems: Flir Water heater sample image. [online], [cit. 2017-04-03]. Dostupné z: http://www.flir.com.hk/uploadedImages/Thermography_USA/Products/Ex-Series/E8-Water-Heater.jpg
- [2] Kadlec, K.: Teoretické základy bezdotykového měření teploty (část 1). [online], [cit. 2017-04-02]. Dostupné z: http://automa.cz/Aton/FileRepository/pdf_articles/52890.pdf
- [3] Sova, J.: Konstrukce termokamery. [online], [cit. 2017-04-25]. Dostupné z: <http://www.termokamera.cz/konstrukce-termokamery/>
- [4] Mathworks: FLIR Accelerates Development of Thermal Imaging FPGA. [online], [cit. 2017-04-25]. Dostupné z: https://www.mathworks.com/company/user_stories/flir-accelerates-development-of-thermal-imaging-fpga.html
- [5] Vojáček, A.: Co jsou BOLOMETRY a MIKROBOLOMETRY. [online], [cit. 2017-04-25]. Dostupné z: <http://automatizace.hw.cz/clanek/2005111601>
- [6] STEMMER IMAGING Ltd: GigE Vision and GenICam. [online], [cit. 2017-04-28]. Dostupné z: <https://www.stemmer-imaging.co.uk/en/technical-tips/gige-vision-and-genicam/>
- [7] Kerul'-Kmec, O.: *Detekce přítomnosti zboží v ruce zákazníka*. B.S. thesis, České vysoké učení technické v Praze. Vypočetní a informační centrum., 2016.
- [8] Kaplan, H.: *Practical applications of infrared thermal sensing and imaging equipment*, ročník 75. SPIE press, 2007.
- [9] Gaussorgues, G.; Chomet, S.: *Infrared thermography*, ročník 5. Springer Science & Business Media, 2012.

LITERATURA

- [10] Vollmer, M.; Möllmann, K.-P.: *Infrared thermal imaging: fundamentals, research and applications*. John Wiley & Sons, 2010.
- [11] Puhl, D.: *Bezkontaktní měření teploty*. B.S. thesis, České vysoké učení technické v Praze. Vypočetní a informační centrum., 2015.
- [12] Smrž, M.: *Bezkontaktní měření teploty*. B.S. thesis, Vysoké učení technické v Brně. Fakulta elektrotechniky a komunikačních technologií, 2013.
- [13] Stupňánková, M.: *IR záření v současnosti*. Dizertační práce, Masarykova univerzita, Lékařská fakulta, 2009.
- [14] Jakl, O.: *Experimentální měření s termovizní kamerou*. B.S. thesis, Vysoké učení technické v Brně. Fakulta elektrotechniky a komunikačních technologií, 2011.
- [15] Malík, D.: *Zpracování dat z termokamery*. Diplomová práce, Vysoké učení technické v Brně. Fakulta elektrotechniky a komunikačních technologií, 2012.
- [16] Kužel, K.: *Bezkontaktní měření teplot*. B.S. thesis, Vysoké učení technické v Brně, Fakulta strojního inženýrství, 2010.
- [17] FLIR Systems: FLIR SC 7000 datasheet. [online], [cit. 2017-04-15]. Dostupné z: http://www.flirmedia.com/MMC/THG/Brochures/RND_017/RND_017_US.pdf
- [18] Daniels, A.: Field Guide to Infrared Systems, Detectors, and FPAs. SPIE, 2010.
- [19] Rogalski, A.: *Infrared detectors*. CRC press, 2010.
- [20] Kadlec, K.; Sova, J.: Termokamery a pyrometry – princip měření, vlastnosti a využití. [online], [cit. 2017-04-02]. Dostupné z: http://www.allforpower.cz/UserFiles/file/termokamery_1.pdf
- [21] Vandervlugt, C.: Choosing Optical Materials for Long Wave Infrared (8–12 μ m) Applications. 2009.
- [22] Automated Imaging Association: The World's Largest Machine Vision Trade Association. [online], [cit. 2017-04-25]. Dostupné z: <http://www.visiononline.org/vision-standards-details.cfm?type=5>
- [23] European Machine Vision Association: GenICam standard. [online], [cit. 2017-04-25]. Dostupné z: <http://www.emva.org/standards-technology/genicam/genicam-downloads/>
- [24] FLIR Systems: Flir Tools+ software. [online], [cit. 2017-04-15]. Dostupné z: <http://www.flir.com/instruments/display/?id=60416>

- [25] FLIR Systems: Atlas SDK. [online], [cit. 2017-04-25]. Dostupné z: http://flir.custhelp.com/app/answers/detail/a_id/1043/~atlas-sdk
- [26] JAI: JAI SDK and Control Tool. [online], [cit. 2017-04-25]. Dostupné z: http://www.jai.com/en/support/jai_sdk_and_control_tool
- [27] A&B Software LLC: Active GigE. [online], [cit. 2017-04-25]. Dostupné z: <http://www.ab-soft.com/activegige.php>
- [28] Pleora Technologies Inc.: eBUS SDK. [online], [cit. 2017-04-25]. Dostupné z: <http://www.pleora.com/our-products/ebus-sdk>
- [29] Saba, E. N.; Larson, E. C.; Patel, S. N.: Dante vision: In-air and touch gesture sensing for natural surface interaction with combined depth and thermal cameras. In *Emerging Signal Processing Applications (ESPA), 2012 IEEE International Conference on*, IEEE, 2012, s. 167–170.
- [30] Itseez: *OpenCV - Background Subtraction Methods*. 2015, [cit. 2017-05-02]. Dostupné z: http://docs.opencv.org/3.1.0/d1/dc5/tutorial_background_subtraction.html
- [31] Otsu, N.: A threshold selection method from gray-level histograms. *Automatica*, ročník 11, č. 285-296, 1975: s. 23–27.
- [32] Zeng, B.; Wang, G.; Lin, X.: A hand gesture based interactive presentation system utilizing heterogeneous cameras. *Tsinghua Science and Technology*, ročník 17, č. 3, 2012: s. 329–336.
- [33] Serra, J.: *Image analysis and mathematical morphology, v. 1*. Academic press, 1982.
- [34] Vidas, S.; Moghadam, P.; Bosse, M.: 3D thermal mapping of building interiors using an RGB-D and thermal camera. In *Robotics and Automation (ICRA), 2013 IEEE International Conference on*, IEEE, 2013, s. 2311–2318.
- [35] Palmero, C.; Clapés, A.; Bahnsen, C.; aj.: Multi-modal RGB–Depth–Thermal Human Body Segmentation. *International Journal of Computer Vision*, ročník 118, č. 2, 2016: s. 217–239.
- [36] Appenrodt, J.; Al-Hamadi, A.; Michaelis, B.: Data gathering for gesture recognition systems based on single color-, stereo color-and thermal cameras. *International Journal of Signal Processing, Image Processing and Pattern Recognition*, ročník 3, č. 1, 2010: s. 37–50.
- [37] Duarte, A.; Carrão, L.; Espanha, M.; aj.: Segmentation algorithms for thermal images. *Procedia Technology*, ročník 16, 2014: s. 1560–1569.

LITERATURA

- [38] Príkopa, Z.: Navrhněte teplotní model prochládání lidského organismu. 2011.
- [39] Mlčák, R.; Pavlek, M.: Mapování teplotních polí na povrchu lidského těla pomocí termovizní kamery. In *Revelling in Reference: XXVI. Conference Hydro-Termo*, 2007, s. 62–64.
- [40] Olbrycht, R.; Wiecek, B.: New Approach to thermal drift correction and gain determination in Microbolometer thermal cameras. In *12th Quantitative InfraRed Thermography Conference, QIRT*, 2014, s. 7–11.
- [41] Riou, O.; Berrebi, S.; Bremond, P.: Nonuniformity correction and thermal drift compensation of thermal infrared camera. In *Defense and Security*, International Society for Optics and Photonics, 2004, s. 294–302.
- [42] Postel, J.: User datagram protocol. Technická zpráva, 1980.
- [43] Postel, J.: Transmission control protocol. 1981.
- [44] FLIR Systems: FLIR A65 datasheet. [online], [cit. 2017-05-01]. Dostupné z: <http://www.mds-flir.com/data/bbsData/Ax5.pdf>
- [45] Corporation, O.: [online], [cit. 2017-05-01]. Dostupné z: <http://docs.oracle.com/javase/8/javase-clienttechnologies.htm>
- [46] RapidMiner: [online], [cit. 2017-05-02]. Dostupné z: <https://rapidminer.com/>
- [47] Bradski, G.: OpenCV library. [online], [cit. 2017-05-01]. Dostupné z: <http://opencv.org/>
- [48] Huang, T.; Yang, G.; Tang, G.: A fast two-dimensional median filtering algorithm. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, ročník 27, č. 1, 1979: s. 13–18.
- [49] Tomasi, C.; Manduchi, R.: Bilateral filtering for gray and color images. In *Computer Vision, 1998. Sixth International Conference on*, IEEE, 1998, s. 839–846.
- [50] Canny, J.: A computational approach to edge detection. *IEEE Transactions on pattern analysis and machine intelligence*, , č. 6, 1986: s. 679–698.
- [51] Itseez: *The OpenCV Reference Manual - AproxPolyDP*. 2014, [cit. 2017-05-02]. Dostupné z: http://docs.opencv.org/2.4/modules/imgproc/doc/structural_analysis_and_shape_descriptors.html#approxpolydp
- [52] Brchl, L.: Modified eBUS Player. [online], [cit. 2017-05-15]. Dostupné z: <https://github.com/brchlluk/Modified-eBUS-Player>

- [53] Brchl, L.: Thesis project application. [online], [cit. 2017-05-15]. Dostupné z: <https://github.com/brchlluk/ThesisProject>
- [54] Brchl, L.: FLIR Image viewer. [online], [cit. 2017-05-15]. Dostupné z: <https://github.com/brchlluk/FLIR-image-viewer>

Seznam použitých zkrátek

- AIA** Automated Imaging Association
BSD Berkeley Software Distribution
CIE International Commission on Illumination
EMVA European Machine Vision Association
FPS Frames per second
FPA Focal Plane Array
GBT Gradient boosted trees
GUI Graphical user interface
HMM Hidden Markov Model
IR Infrared
MOG Mixture of gaussians
NUC Non-Uniformity Correction
PoE Power over Ethernet
UDP User Datagram Protocol
SVM Support vector machines
RGB Red, Green, Blue
UV Ultraviolet
XML Extensible markup language

Obsah přiloženého média

readme.txt	stručný popis obsahu přiloženého média
src	
src	
impl	adresář s implementovanými nástroji
eBUSPlayer	zdrojové kódy upravené aplikace eBUS Player
ThesisProject	zdrojové kódy hlavní aplikace s algoritmy
FlirBinaryViewer	zdrojové kódy nástroje pro prohlížení a anotování snímků
thesis	zdrojová forma práce ve formátu LATEX
text	text práce
thesis.pdf	text práce ve formátu PDF
image_examples	několik ukázkových snímků z termokamery
dataset	složka s vyhodnocovaným datasetem