

Bakalářská práce



České
vysoké
učení technické
v Praze

F3

Fakulta elektrotechnická
Katedra řídicí techniky

Efektivní implementace algoritmů zpracování obrazu z bolometrického pole

Michal Kochman

Vedoucí: doc. Ing. Stanislav Vítek, Ph.D
Srpen 2021

I. OSOBNÍ A STUDIJNÍ ÚDAJE

Příjmení: **Kochman** Jméno: **Michal** Osobní číslo: **465880**
Fakulta/ústav: **Fakulta elektrotechnická**
Zadávací katedra/ústav: **Katedra řídicí techniky**
Studijní program: **Kybernetika a robotika**

II. ÚDAJE K BAKALÁŘSKÉ PRÁCI

Název bakalářské práce:

Efektivní implementace algoritmů zpracování obrazu z bolometrického pole

Název bakalářské práce anglicky:

Efficient implementation of algorithms to process images from bolometer array

Pokyny pro vypracování:

- 1) Prostudujte a popište základní metody zpracování obrazu s ohledem na omezené výpočetní prostředky a paměť, včetně metod využívajících strojové učení
- 2) Navrhněte algoritmy detekce objektů v obraze z bolometrického pole
- 3) Navržené algoritmy implementujte na vhodné embedded platformě (např. SoC ARM)
- 4) Diskutujte stabilitu výsledků s ohledem na vnější podmínky - náhlé změny teploty prostředí nebo objektu, kvalita kalibrace, apod.

Seznam doporučené literatury:

- [1] SUNDARARAJAN, Duraisamy. Digital image processing: a signal processing and algorithmic approach. Springer, 2017.
[2] BISHOP, Christopher M. Pattern recognition and machine learning. Springer, 2006.
[3] KHAN, Muhammad UK, et al. A swift and memory efficient Hough transform for systems with limited fast memory. In: International Conference Image Analysis and Recognition. Springer, Berlin, Heidelberg, 2009. p. 297-306.

Jméno a pracoviště vedoucí(ho) bakalářské práce:

doc. Ing. Stanislav Vítek, Ph.D., katedra radioelektroniky FEL

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

Datum zadání bakalářské práce: **29.01.2021**

Termín odevzdání bakalářské práce: **13.08.2021**

Platnost zadání bakalářské práce: **30.09.2022**

doc. Ing. Stanislav Vítek, Ph.D.
podpis vedoucí(ho) práce

prof. Ing. Michael Šebek, DrSc.
podpis vedoucí(ho) ústavu/katedry

prof. Mgr. Petr Páta, Ph.D.
podpis děkana(ky)

III. PŘEVZETÍ ZADÁNÍ

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

Datum převzetí zadání

Podpis studenta

Poděkování

Kdokoliv má trpělivost s mojí tendencí prozkoumávat slepé vývojové větve a vrhat se do neplodných pokusů bez předchozího uvážení zaslouží být korunován svatozáří. Tu udělit nemohu, nicméně mohu zde jmenovitě zmínit ty, kteří mě podpořili nejvíce - vedoucího své práce, pana doktora Vítka, svého oponenta pana inženýra Kováře a kolektiv firmy Workswell.

Prohlášení

Prohlašuji, že jsem předloženou práci vypracoval samostatně, a že jsem uvedl veškerou použitou literaturu.

V Praze, 13. srpna 2021

Abstrakt

Hlavním cílem této práce je prozkoumat metody digitálního zpracování obrazu za účelem detekce obličeje pro měření teploty člověka termální kamerou (mikrobolometrickým polem), a ověřit jejich implementaci na reálném hardware (termokamera připojená k embedded procesoru architektury ARM). Přináší návrh algoritmu pro detekci obličeje ve snímku založený na vyhledávání elipsy jako vzor tvaru obličeje. Omezení diktovaná slabým výpočetním výkonem použitého hardwaru značně snižují schopnosti navrženého algoritmu, který však zaslouží další zkoumání v případě využití silnější platformy.

Klíčová slova: bolometr, termokamera, měření teploty, zpracování obrazu, detekce obličeje, Houghova transformace

Vedoucí: doc. Ing. Stanislav Vítek,
Ph.D
katedra radioelektroniky FEL

Abstract

The main goal of this paper is to evaluate methods of digital image processing with regards to detection of human face for temperature measurement by thermal camera (microbolometer array), and to test their implementation on real hardware (namely a thermal camera connected to embedded ARM-based microprocessor). The paper suggest an algorithm of face detection based on searching for ellipses as a template for face shape. The limitations brought by the hardware computational capabilities are decremental to the proposed algorithm, which deserves future consideration should more powerful platform be used.

Keywords: bolometer, thermal camera, temperature measurement, image processing, face detection, Hough transform

Title translation: Efficient implementation of algorithms to process images from bolometer array

Obsah

1 Úvod	1	3.3.1 Normalizace, přiřazení barev	12
2 Měření termokamerou	3	3.3.2 Potlačení šumu	12
2.1 Tepelné záření	3	3.3.3 Odstranění šumu ve frekvenční doméně	13
2.1.1 Emisivita	4	3.4 Detekce hran	14
2.2 Bolometr	4	3.4.1 Robertsův hranový detektor	14
2.2.1 Implementace termokamer	5	3.4.2 Prewittové metoda	15
2.2.2 Reálné vlastnosti mikrobolometrických polí	6	3.4.3 Cannyho hranový detektor	15
2.3 Výstup	6	3.5 hledání prvků	17
3 Metody zpracování obrazu	7	4 Metody detekce objektů v obraze	19
3.1 Digitální obraz	7	4.0.1 Detekce tvarů a oblastí	19
3.1.1 Souřadnicová konvence	8	4.1 Metody strojového učení	22
3.1.2 Termogram	8	4.1.1 Konvoluční neuronové sítě	22
3.2 Základní metody práce s obrazem	9	4.1.2 Viola-Jones detektor	23
3.2.1 Prostorová doména	9	5 Návrh řešení a implementace	27
3.2.2 Frekvenční doména	10	5.1 Hardware a software parametry systému	28
3.3 Preprocessing obrazu	11	5.1.1 Software	28

5.1.2 Hardware	28
5.1.3 Snímače	28
5.2 Cíl	29
5.3 Metody strojového učení	29
5.4 Segmentace	30
5.5 Vyhledávání obličeje jako elipsy .	30
6 Výsledky	33
7 Závěr	37
A Literatura	39
B Obsah priloženého adresare	43

2.1 Rozdělení spektra elektromagnetického záření, převzato z [dro20]	5
3.1 Znázornění souřadnicové konvence	8
3.2 Příklad konvoluce matice (obrazu) 3x3 s konvolučním jádrem 3x3 s využitím zarovnání nulami	10
3.3 Srovnání algoritmů pro hledání hran	16
4.1 Příklad segmentace	20
4.2 Příklad Haarových bází, převzato z [DH72]	24
4.3 Výpočet integrálního obrazu, převzato z [Ope]	25
5.1 Srovnání výstupů prahování (segmentace) se stejnými parametry na termogramech tvářích dvou zdravých lidí	30
6.1 Příklady nalezených elips při detekci obličeje (zeleně - maska kalibračního černého tělesa pro výpočet středu, kříž - odhadnutý střed, červeně - nalezené platné elipsy	34



Kapitola 1

Úvod

Žijeme v době, kdy si je lidská civilizace blíže než kdy byla. Cesta z jednoho konce světa na druhý netrvá týdny jako ještě před 100lety, ale je zvládnutelná komerčními linkami za pár dní. Tato propojenost má mnoho výhod, činí nás však zranitelnějšími - například přicházíme o přirozenou karanténní dobu cesty, což může napomáhat šíření nemocí. Na konci roku 2019 se objevily první zprávy o přenosu nového typu viru (SARS-CoV-2) na člověka. Tento virus je přenositelný osobním kontaktem s nakaženou osobou a jedním z nejčastěji vyskytovaných příznaků je horečka a kašel.

Během několika následujících měsíců se virus rozšířil prakticky do celého světa. Jednotlivé státy byly nuceny reagovat mnohými opatřeními s cílem zamezit šíření viru, například omezení pohybu osob a zavedení většího dohledu nad zdravotním stavem jedince (pravidelné testování na přítomnost viru, nebo protilátek v těle).

Pravidelné odběry biologického materiálu a jejich rozbor (testování) jsou pro většinu jedinců nepříjemné a také jsou relativně drahé. Protože jedním z nejčastějších příznaků virových onemocnění je zvýšená tělesná teplota, nabízí se pravidelné měření teploty jednotlivců a následné testování pouze u lidí, u kterých screening tělesné teploty poukáže na možnost nákazy virem. V zájmu minimalizace kontaktu s potenciálním nakaženým člověkem by takové měření bylo ideálně zcela bezkontaktní a automatizované.

Pro bezkontaktní měření teploty je možné využít infračervené teploměry, které je ovšem nutné přesně zaměřit, nebo termokamery, schopné zabrat větší

prostor.

Vyvstala tedy náhlá poptávka po termovizních systémech instalovatelných při vstupech do veřejných budov, firmemních prostor apod. schopných fungovat jako systém výstrahy před vstupem infekčních osob. Pro dosažení maximální možné automatizace je vhodné detekovat obličeje.

Cílem této práce je detekce obličeje v obraze termokamery pro následné vyhodnocení teploty v této oblasti.

Kapitola 2

Měření termokamerou

V této kapitole je představen fyzikální princip měření teploty termokamerou, základní omezení a výhody plynoucí z tohoto principu,

2.1 Tepelné záření

Všechny objekty ve fyzikálním světě mající teplotu nad absolutní nulou ($0\text{K} = -273.15^\circ\text{C}$) vyzařují do svého okolí elektromagnetické záření které nazýváme tepelným. Z Wienova posunovacího zákona vyplývá, že pro tělesa s teplotou nižší než řádově tisíce stupňů celsia je maximální energie vyzářena na infračervených vlnových délkách [BB06], pro měření teplot blízkých tělesné teplotě teplokrevných živočichů nás nejvíce zajímají vlnové délky okolo $10\mu\text{m}$, tedy záření v infračerveném spektru (IR záření)

Intenzita tepelného vyzařování (na všech vlnových délkách) je vyjádřena Stefan-Boltzmanovým zákonem [SH02, BB06]

$$I = \epsilon\sigma T^4 \left[\frac{\text{W}}{\text{m}^2} \right] \quad (2.1)$$

Tento zákon udává intenzitu záření I v závislosti na třech parametrech:

- Stefan-Boltzmanově konstantě $\sigma \doteq 5.67 \cdot 10^{-8} \text{Jm}^{-2}\text{s}^{-1}\text{K}^{-4}$ [SH02],

- Termodynamické teplotě tělesa T [K]
- emissivitě ϵ

Pro určení teploty tělesa o známé emissivitě tedy postačuje změřit intenzitu záření.

■ 2.1.1 Emisivita

Emisivita $\epsilon \in (0; 1)$ je parameter vyjadřující poměr mezi intenzitou vyzařování reálného tělesa a vyzařování absolutně černého tělesa (t.j. tělesa pohlcujícího veškeré záření dopadající na jeho povrch) o stejné teplotě. Emisivita obecně záleží na vícero parametrech (např. vlnová délka - závislá na teplotě, materiál povrchu, úhel vyzařování) a bývá určována experimentálně. Těleso které má pro určitý rozsah vlnových délek konstantní emisivitu se nazývá šedé těleso [ISO17], do této kategorie spadá většina objektů běžně měřených termokamerami.

Pro lidskou pokožku je hodnota emissivity velice blízká 1, běžně je uváděna v rozmezí $0.97 - 0.99$ [VMSMCC09, Tog89]. Vysoká hodnota emissivity je výhodná pro přesnost měření, neboť potlačuje rušivý vliv teploty pozadí (tedy odražené teploty).

■ 2.2 Bolometr

V termálních kamerách se k měření intenzity IR záření využívá mikrobolometrického pole, tedy matice malých bolometrických snímačů. Bolometr se skládá z teplotně závislého rezistoru a absorpční vrstvy která se dopadem elektromagnetických paprsků zahřívá (např. zlato). Zahřátí senzoru dopadajícím elektromagnetickým zářením má za výsledek měřitelnou velikost odporu. Pro určení teploty měřeného tělesa (zdroje elektromagnetického záření) je ještě nutné provést kompenzaci vlivu okolní teploty, zpravidla referenčním odporovým teploměrem (NTC termistor).

2.2.1 Implementace termokamer

Infračervená oblast elektromagnetického záření je relativně široká (ve srovnání s viditelným nebo ultrafialovým zářením), zabírá rozsah 700nm – 1mm. Je tedy vhodné si zavést jemnější dělení. Takové dělení poskytuje např [ISO17].

Mluvíme-li o termokamerách může být vhodnější si zavést dělení podle komerčně dostupných senzorů, které se soustředí na průmyslově využitelné vlnové délky. [dro20] Toto dělení nepokrývá ani zdaleka celý IR rozsah:

- LWIR (Long wavelength infrared, 8 – 14 μm). Kamery pracující s LWIR jsou nejrozšířenější, používané hojně pro kontrolu úniků tepla z budov, inspekce solárních elektráren apod a pro měření teploty.
- MWIR (Middle Wavelength Infrared, 3 – 5 μm). Kamery pracující s MWIR najdou využití například pro detekci úniku plynu.
- SWIR (Short Wavelength Infrared, 0.9 – 1.7 μm). Kamery pracující s LWIR jsou z vyjmenovaných nejméně rozšířené, jejich hlavní výhodou je průhlednost záření o této vlnové délce vodními mraky.



Obrázek 2.1: Rozdělení spektra elektromagnetického záření, převzato z [dro20]

Další důležité dělení kamer je podle jejich schopnosti určovat teplotu:

- Radiometrické termokamery jsou schopné určit absolutní teplotu měřeného objektu. Výstupem radiometrické kamery je reprezentace teploty, většinou v Kelvinově stupnici.
- Neradiometrické termokamery nejsou schopny určit absolutní teplotu měřeného objektu. Jejich výstupem je je reprezentace rozdílů teploty v rámci obrazu, který není možné vztáhnout k nějaké stupnici teploty.

■ 2.2.2 Reálné vlastnosti mikrobolometrických polí

Výroba mikrobolometrických polí je relativně náročná. Ačkoliv existují termokamery s rozlišeními FullHD a větší (tedy více jak 1024 x 720), pro svoji snažší cenovou dostupnost jsou nejhojněji využívány kamery s menším rozlišením (např. 640x520, 320x240 ale i tak malá rozlišení jako 80 x 60). I při rozlišení 80 x 60 se na matici bolometrů projevuje její nestejněměrné zahřátí a nestejně převodní charakteristiky jednotlivých pixelů a je nutné pravidelně provádět NUC (z anglického non-uniformity correction). Tento proces je pro většinu komerčních kamer prováděn pomocí shutteru (sejmutí homogenního pole) a úpravy kolísání. Měřené hodnoty se po provedení NUC pomocí shutteru mohou skokově změnit až o desetiny stupně celsia.

■ 2.3 Výstup

Výstupem z termokamery je matice hodnot, ve které každý prvek (i, j) reprezentuje teplotu na souřadnicích (i, j) . Tuto matici lze chápat jako jednonábový obraz (obraz ve stupních šedi). Obrazu se dále věnuje kapitola 3.1

Kapitola 3

Metody zpracování obrazu

3.1 Digitální obraz

Gonzales [GW08] definuje obraz jako dvourozměrnou funkci

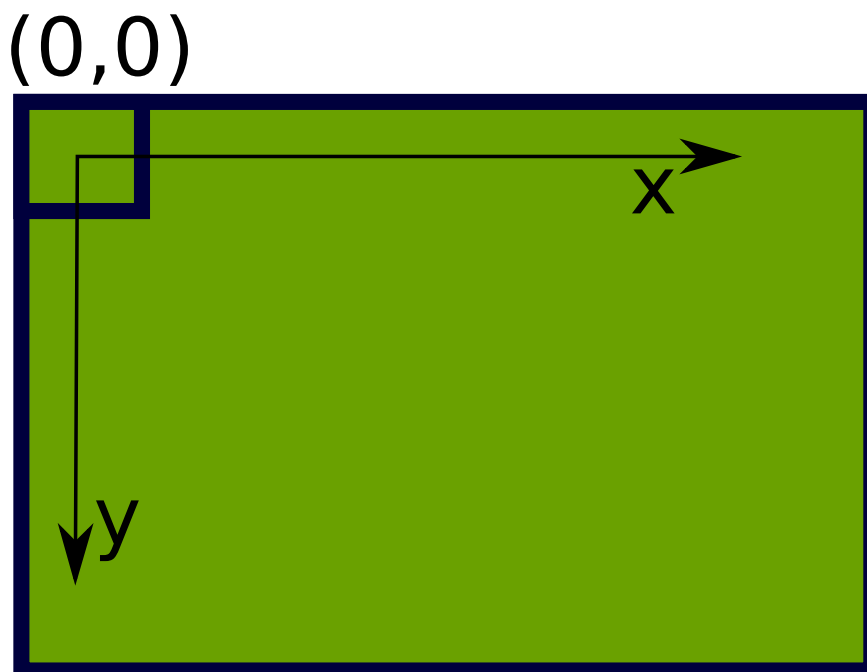
$$f(x, y), \tag{3.1}$$

kde x a y jsou prostorové souřadnice a hodnota f pro dané souřadnice odpovídá barevné intenzitě pro tyto souřadnice. Pokud jsou souřadnice x a y celá čísla a všechny funkční hodnoty konečné a diskrétní mluví o obrazu digitálním. Jednotlivé body na souřadnicích (x, y) pak označujeme jako pixely. Funkci 3.1 budeme dále označovat jako obrazovou funkci.

Pro digitální obraz je tedy funkce $f(x, y)$ zobrazení $\mathbb{N}^2 \rightarrow \mathbb{Z}^n$ kde n je dimenzionalita barevného prostoru, nejčastěji $n = 1$ (jednokanálový obraz - obraz ve stupních šedi označovaný také jako grayscale nebo binární obraz), $n = 3$ (barevný obraz například v prostorech RGB nebo YUV) nebo $n = 4$ (barevný obraz například v prostorech CMYK nebo RGBA). Pole (matici) hodnot funkce $f(x, y)$ se v této práci označuje také symbolem \mathbf{f} . Některé metody zpracování obrazu pracují barevném prostoru (RGB, YUV). Pro snímky získané termokamerou nás zajímají pouze metody pracující s jednokanálovým obrazem (obrazem ve stupních šedi), neboť mikrobolometrické pole má na výstupu vždy jen jeden kanál (reprezentující intenzitu záření v nějaké části IR spektra).

3.1.1 Souřadnicová konvence

Není-li uvedeno jinak je v tomto dokumentu používána běžná souřadnicová konvence pro obraz, tj. levý horní pixel obrazu má souřadnice $(0, 0)$, osa Y směřuje k levému dolnímu rohu obrazu (reprezentuje výšku obrazu) a osa X k pravému hornímu rohu (reprezentuje šířku obrazu). Grafické znázornění souřadnicové konvence je na obr. 3.1.



Obrázek 3.1: Znázornění souřadnicové konvence

3.1.2 Termogram

Uvedená definice obrazu vyžaduje pro digitální obrazy získané z termokamery pouze malou úpravu - hodnota $f(x, y)$ nese jen informace o intenzitě záření, ale je zároveň reprezentací teploty pro dané souřadnice v obraze. Teplota může být absolutní (pro "radiometrický termogram", tedy snímek z radiometrické kamery který si skutečně s každým pixelem může nést hodnotu teploty) nebo relativní (pro "neradiometrický termogram", tedy snímek z neradiometrické termokamery, kde nemáme informaci o změřené teplotě ale pouze o teplotních rozdílech v rámci snímku).

Digitální výstup ze mikrobolometrického pole je často reprezentován 14bi-

tovými hodnotami teploty. Tyto hodnoty jsou následně často saturovány pro určitý teplotní rozsah a normalizovány na 8 bitů. Tuto reprezentaci lze následně interpretovat jako obraz ve stupních šedi (viditelného spektra).

Obecně lze říct, že termogram je obraz převádějící intenzitu záření z infračervené oblasti do viditelného spektra.

Pro jedodušší orientaci v obraze se často přiřazují jednotlivým intenzitám barvy (namísto využití převodu na jednobarevný obraz) na základě definovaných look-up tabulek (tzv. barevných palet), popřípadně se obarví všechny pixely jejichž teplota splňuje určitou podmínku (takové zbarvení označujeme jako alarm, nebo isothermu).

3.2 Základní metody práce s obrazem

3.2.1 Prostorová doména

Operací na obraze v prostorové doméně rozumíme operace prováděné přímo na hodnotách jednotlivých pixelů, tj. přímo na funkčních hodnotách funkce $f(x, y)$. Takové metody můžeme zapsat jako

$$g(x, y) = T(f(x, y)), \quad (3.2)$$

kde $g(x, y)$ značí novou obrazovou funkci (tj. obrazovou funkci upraveného obrazu) a T je nějaká operace definovaná na pixelech obrazu.

Konvoluce

Velká část operací na digitálním obraze je realizována pomocí dvourozměrné diskrétní konvoluce. Ta se v matematice obvykle značí $*$, tohoto značení se budeme v tomto dokumentu držet. Konvolucí je každému bodu (x, y) je přiřazena hodnota k podle vztahu

$$k(x, y) = f(x, y) * m(i, j) = \sum_{i=0}^{I-1} \sum_{j=0}^{J-1} f(x-i, y-j) \cdot m(i, j) \quad (3.3)$$

ve kterém $m(i, j)$ jsou hodnoty konvoluční masky (jádra) o rozměrech $I \times J$ (většinou $I = J$). Prakticky jsou tedy každému pixelu přiřazeny hodnoty

funkčně (většinou lineárně) závislé na hodnotách funkce $f(x, y)$ v jeho okolí o rozměrech $I \times J$.

Popsaný přístup zjevně narazí na problémy při okrajích obrazu, kde každý pixel nemá něktrým směrem potřebný počet "sousedů". Tyto pixely buď vůbec nemusí být zahrnuty do zpracovaného obrazu (zpracovaný obraz se vůči originálnímu zmenší o daný počet pixelů), mohou být do zpracovaného obrazu překopírovány beze změny nebo se původní obraz zvětší dodáním okraje (většinou o nulových hodnotách) který se použije pro výpočty (tzv. zero-padding, zarovnání nulami).

Příklad výpočtu konvoluce matice o rozměrech 3×3 s jádrem 3×3 s využitím zero-padding je uveden na obrázku 3.2. Konkrétní výpočet hodnoty pixelu o souřadnicích $(x, y) = (0, 0)$

0	0	0	0	0	*	1	0	-1	=	-14	-6	14
0	7	5	7	0		2	0	-2		-17	-14	17
0	2	4	8	0		1	0	-1		-12	-10	12
0	3	4	5	0								
0	0	0	0	0								

Obrázek 3.2: Příklad konvoluce matice (obrazu) 3×3 s konvolučním jádrem 3×3 s využitím zarovnání nulami

3.2.2 Frekvenční doména

Úprava obrazu ve frekvenční doméně se hojně využívá, jeho hlavní výhodou je výpočetní rychlost ve srovnání s výpočtem konvoluce v prostorové doméně.

Pro filtraci ve frekvenční doméně je nejprve nutné převést prostorovou reprezentaci obrazu (funkci 3.1) na reprezentaci frekvenční, tedy na funkci

$$F(u, v). \quad (3.4)$$

Toho se v praxi nejčastěji docílí pomocí diskrétní Fourierovy transformace (DFT).

■ Fourierova transformace

Tu definujeme jako

$$F(u, v) = \mathbb{F}\{f(x, y)\} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-2\pi j \frac{uxN+xyM}{MN}} \quad (3.5)$$

pro všechna $u = 0, 1 \dots M - 1$ a $v = 0, 1 \dots N - 1$ ($M \times N$ značí rozměry snímku). Inverzní operaci, tedy přechod z frekvenční domény zpět do prostorové reprezentace obrazu označujeme jako inverzní DFT (IDFT). Vztah pro IDFT je možné získat přímo z definice DFT:

$$f(x, y) = \mathbb{F}^{-1}\{F(u, v)\} = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) e^{2\pi j \frac{uxN+xyM}{MN}}. \quad (3.6)$$

Namísto výpočetně složité konvoluce v prostorové doméně, můžeme ve frekvenční provádět mnohem jednodušší násobení přenosovou funkcí filtru, označme takovou funkci $H(u, v)$.

Operace ve frekvenční doméně včetně převodu mezi doménami lze tedy zapsat jako

$$g(x, y) = \mathbb{F}^{-1}\{\mathbb{F}\{f(x, y)\} \cdot H(u, v)\}. \quad (3.7)$$

Tato operace odpovídá konvoluci $f(x, y) * \mathbb{F}^{-1}(H(u, v))$ v prostorové doméně.

■ Rychlá fourierova transformace

Pro svoji častou používanost je DFT často spočítána pomocí Rychlé Fourierovy transformace (FFT, Fast Fourier Transform). Jedná se o skupinu algoritmů využívající specifických vlastností spektra pro urychlení výpočtu DFT (převádí problém ze skupiny algoritmů s náročností $O(N^2)$ na problém s náročností $O(n \log n)$).

■ 3.3 Preprocessing obrazu

Preprocessing obrazu je proces běžně užíván pro dva různé účely:

1. Vylepšení kvality obrazu (pro lidského pozorovatele)

2. Předzpracování obrazu pro běh složitějších detekčních či měřicích algoritmů.

3.3.1 Normalizace, přiřazení barev

Dnes hojně používané mikrobolometrické senzory mají rozlišení okolo 14bitů. Pro zobrazení je tento rozsah často normalizován na 8bitů, a to oříznutím krajních ohodnot (změna zobrazeného rozsahu, saturace) a následně přemapováním zbylých hodnot na rozsah 0-255. Takto získané hodnoty lze chápat jako hodnoty obrázku ve stupních šedi, nebo jim přiřadit barvy podle předem definované look-up tabulky.

3.3.2 Potlačení šumu

IR senzory mají typicky mnohem horší SNR (signal-to-noise ratio, poměr signálu a šumu) než digitální senzory pracující ve viditelném spektru [DABL06]

Dnes se i pro filtraci občas využívá metod strojového učení, zejména při práci s obrazem získaným v jiném spektru elektromagnetických vln než viditelném (např. magnetická rezonance, tomografie). [HZH21, KSK⁺20]. Těmto metodám se zde nebudeme věnovat.

Filtrace šumu v obraze má za následek ztrátu ostrosti obrazu (jeho rozmazání).

Odstranění šumu v prostorové doméně

K odstranění šumu v prostorové doméně se využívá konvolučních masek operujících jako průměrování (vyhlazení obrazu). Prkladem filtračního jádra k odstranění šumu tedy může být

$$\mathbf{m} = \frac{1}{9} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}, \quad (3.8)$$

realizující prostý aritmetický průměr na okolí 3x3, nebo

$$\mathbf{m} = \frac{1}{159} \begin{pmatrix} 2 & 4 & 5 & 4 & 3 \\ 4 & 9 & 12 & 9 & 7 \\ 5 & 12 & 15 & 12 & 9 \\ 4 & 9 & 12 & 9 & 7 \\ 2 & 4 & 5 & 4 & 3 \end{pmatrix}, \quad (3.9)$$

realizující tzv. Gaussův filtr.

K odstranění šumu je také možné využít několika snímků stejné scény a provádění prostorových operací majících na vstupu více snímků. Tato metoda je vhodná při práci se stabilní scénou.

■ 3.3.3 Odstranění šumu ve frekvenční doméně

K odstranění šumu ve frekvenční doméně se využívají low-pass filtry (filtry typu dolní propust). Ideální low-pass filter má pouze hodnoty $H(u, v) \in \{0, 1\}$, tedy až do nějakého mezního (u, v) propouští signál nezměněný a dále nepropouští vůbec. Taková přenosová funkce je jen těžko vyjádřitelná analytickým vzorcem, často se tedy používají aproximace, např. Chebyševova nebo Butterworthova.

■ NUC

Téměř vždy používanou formou preprocessingu za účelem odstranění nedokonalostí termogramu je aplikace opravy neuniformity senzoru. NUC (non-uniformity correction) je odstranění rušení vyvolaného vlastní teplotou mikrobolometrického pole - každý mikrobolometr má trochu jinou vlastní teplotu a trochu jinou převodní charakteristiku. Většinou se tedy využívá se tedy periodického odečtu hodnot z homogenního obrazu (většinou vlastní mechanické uzávěrky senzoru, odtud pak pojem "shutter correction") k vypočtení opravy pro každý pixel obrazu. [HCJH13]. Tato filtrace je téměř vždy implementována přímo na snímači. Implementace je často kombinací výše popsaných přístupů.

Protože NUC opravuje teplotní hodnoty jednotlivých pixelů, měřená teplota těsně před a po provedení opravy se může lišit (řádově o desetiny stupně).

3.4 Detekce hran

Hledání hran v obraze je důležitý typ operace, který značně zjednodušuje detekci objektů. Obraz typicky nese velké množství dat, z nichž pro algoritmy vyhledávající v něm nějaké prvky zpravidla většina není důležitá. Detekce hran slouží k odflitrování (pro následné algoritmy) nezajímavých prvků.

Hledání hran je sice do jisté míry inverzní operací k potlačení šumu (při potlačení šumu je obraz vyhlazován a hrany jsou tak potlačeny namísto zvýrazněny) a je většinou implementováno stejným principem.

K hledání hran je možné využít naivní přístup (např. pouhé hledání gradientu) nebo složitější algoritmy (např. Cannyho algoritmus). Existují i metody využívající strojového učení.

Metody pro hledání hran se většinou aplikují na jednobarevný obraz.

3.4.1 Robertsův hranový detektor

Jedná se o detektor pracující v prostorové doméně. Robertsův hranový detektor určí velikost gradientu ze vztahu

$$G(x, y) = \sqrt{(\sqrt{f(x, y)} - \sqrt{f(x + 1, y + 1)})^2 + (\sqrt{f(x + 1, y)} - \sqrt{f(x, y + 1)})^2} \quad (3.10)$$

Jedná se vlastně o geometrický průměr hodnot obrazů získaných dvojí konvolucí s rotovanou maskou (jádrém) 2×2 s hodnotami ± 1 na diagonálách. Získané hodnoty porovná s určitým definovaným prahem T . Je-li výsledek $G[f(x, y)] > T$, pak bodem (x, y) prochází hrana.

Sobelova metoda

Další detektor pracující v prostorové doméně - Sobelova metoda hledání hran je založena na dvou konvolučních maskách (jedna je pouze rotací druhé), detekující hrany v vertikálním a horizontálním směru. Je možné použít i rotace masky pro detekce v diagonálních směrech (použití 4 směrů namísto dvou

je sice výpočetně náročnější, ale vydává lepší výsledky). Uvedené hodnoty masky (Sobelova operátoru) jsou pro okolí 3×3 , nicméně operátor je možné definovat i pro větší okolí.

$$S_x = \begin{pmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{pmatrix}, \quad S_y = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & -0 \\ -1 & -2 & -1 \end{pmatrix}.$$

Konvolucí těchto konvolučních jader s maticí obrazu f získáme matice směrových gradientů G_x a G_y . Velikost gradientu v nějakém bodě je vyjádřena

$$G(x, y) = \sqrt{(G_x(x, y))^2 + (G_y(x, y))^2}. \quad (3.11)$$

Opět platí že tímto bodem prochází hrana právě tehdy když $G(x, y) > T$.

Jedná se o spojení průměru (filtrace) v jednom směru a hledání gradientu ve směru na něj kolmém, jak je zřetelnější z rozepsání konvoluční masky (např. pro svislé hrany) následujícím způsobem:

$$S_x = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & -1 \end{pmatrix}. \quad (3.12)$$

■ 3.4.2 Prewittové metoda

Je téměř shodná se Sobelovou metodou, jen využívá mírně jiné hodnoty konvoluční masky:

$$P_x = \begin{pmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{pmatrix}, \quad P_y = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & -0 \\ -1 & -1 & -1 \end{pmatrix}.$$

Podobných operátorů lišících se pouze v konkrétních hodnotách konvolučního jádra existuje celá řada.

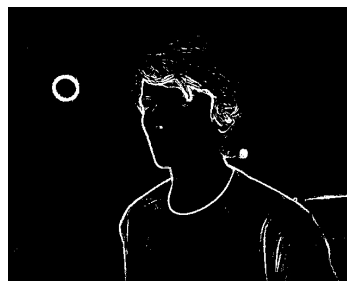
■ 3.4.3 Cannyho hranový detektor

Cannyho detektor, nebo též Cannyho detekční algoritmus je hojně používaný algoritmus pro detekci hran. Byl navržen Johnem Cannyem v roce 1986 [Can86] Pokud není obraz příliš zarušen šumem dosahuje často dobrých výsledků. Průběh Cannyho algoritmu lze popsat v následujících krocích:

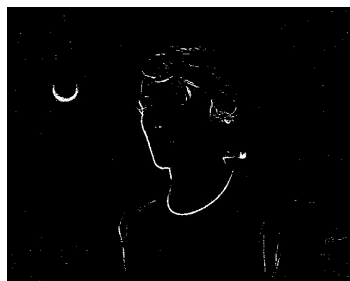
1. Odstranění šumu - většinou se používá jednoduchý Gaussův filter aplikovaný konvoluční maskou.
2. Určení gradientů - např. Sobelevou metodou. Tado metoda nám výhodně určí také směr gradientu (pro následující krok).
3. Nalezení lokálních maxim gradientů, respektive odebrání bodů které maximem nejsou (nonmaximum supression). Tento krok znamená nalezení pixelů jejichž susedé ve směru a proti směru gradientu mají hodnotu gradientu nižší
4. Eliminace nevýznamných hran (thresholding s hysterezí) - předchozími kroky jsme získali všechny (i nevýznamné) hrany. V tomto kroku určíme dvě hranice (označme G_{min} a G_{max}) jako prahy pro hodnotu gradientu. Pokud hodnota gradientu daného pixelu je větší než G_{max} je "uznán" hranou, pokud je menší než G_{min} pak hranou není. Pokud se nachází hodnota gradientu pixelu mezi G_{min} a G_{max} je uznán právě tehdy když přímo susedí s jiným již uznaným pixelem. V potaz je možné brát různě definované okolí (např. 4. pixely přímo susedící s vyhodnocovaným pixelem, častěji se uvažují i pixely susedící diagonálně, tedy 8pixelů).



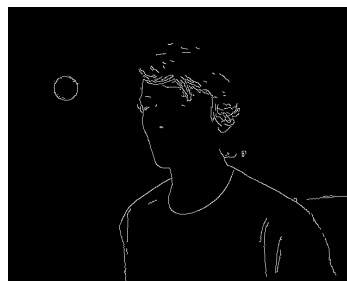
(a) : Původní snímek



(b) : Vyhledání hran Sobelovou metodou



(c) : Vyhledání hranho Robertsovým detektorem



(d) : Vyhledání hran Cannyho algoritmem s použitím Sobelovy metody

Obrázek 3.3: Srovnání algoritmů pro hledání hran

■ 3.5 hledání prvků

Výstup z pre-processingu je vstupem vlastních algoritmů hledajících a rozpoznávajících prvky v obraze. Typicky hledanými jsou dopravní značky, státní poznávací značky vozidel (a text obecně), lidé a obličeje. Hledání prvků a objektů se věnuje následující kapitola.

Kapitola 4

Metody detekce objektů v obraze

Detekce specifických objektů v obraze, ideálně v reálném čase, je obor vzrůstající důležitosti. Mnoho společností dnes investuje do projektů automatizace výroby a logistiky (vzrůstající potřeba detekce výrobků na pásových dopravnících), asistentů řízení a samořiditelných aut (nutnost detekce překážek, ostatních účastníků silničního provozu a dopravního značení), rozpoznávání osob a obličejů v záznamech CCTV kamer atd.

Některé naivní a relativně jednoduché algoritmy uvedené v této kapitole mohou být teoreticky považovány za pouhý preprocesing pro algoritmy složitější, nicméně s omezeným výpočetním výkonem je možné, ne-li nutné, je zkoumat i z pohledu možnosti detekce pouze s jejich využitím.

4.0.1 Detekce tvarů a oblastí

Tyto metody nevyžadují žádné trénování, jejich funkčnost je založena čistě na navržení postupu využívajícího znalost typické vlastnosti hledaného objektu, typicky např. tvar a barva. Oproti metodám využívajících strojové učení jsou výpočetně snazší, nicméně mnohem náchylnější k nespolehlivé detekci, nicméně zpřesňování detekce úpravou vstupních parametrů je mnohem intuitivnější.

■ Segmentace

Segmentace znamená hledání segmentu (oblasti) obrazu, která splňuje určitou podmínku. Základními typy segmentace jsou:

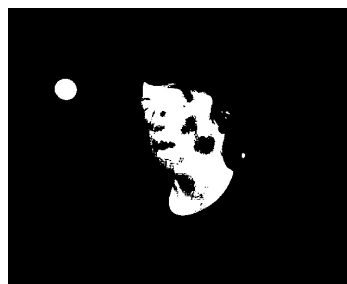
- **Prahování (thresholding):** Prahováním označujeme detekci určitých oblastí zájmu v obraze, pokud se dá předpokládat že pro dané oblasti platí nějaká nerovnost $\alpha < f(x, y) < \beta$ (nebo nerovnost opačná, záleží na konkrétních vlastnostech hledané oblasti). α a β jsou volené konstanty (obecně n-tice konstant při práci s vícekanálovým obrazem). Lze používat také složitější hysteretní prahování, kde jsou definovány prahy 2 (T_2, T_1). V takovém případě se pixel mající hodnotu mezi uvedenými prahy zahrne do oblasti (segmentu) právě tehdy když má nějakého souseda který již v segmentu je.

Pomocí prahování je možné v barevných obrazech například zvýraznit oblasti které pravděpodobně obsahují rostliny (zelená barva), nebo v případě termogramů oblasti splňující určitou teplotu. Prahování se využívá při práci s termogramem pro určení oblastí zájmu (alarmy).

- **Hledání uzavřených oblastí:** Další metodou jak rozdělit obraz na skupinu oblastí je nalézt oblasti uzavřené nějakou hranou, tedy použít algoritmy hledání hran, hrany topologicky prozkoumat (nalézt uzavřené křivky) a označit oblasti uvnitř uzavřených křivek.



(a) : Původní snímek



(b) : Binární segmentace prahováním

Obrázek 4.1: Příklad segmentace

■ Houghova transformace

Po nalezení hran je možné mezi nalezenými hranami hledat konkrétní tvary. Pokud jsou hledané tvary snadno matematicky parametrizovatelné (tedy zpravidla jedná-li se o základní geometrické tvary jako kuželoščky) je možné

využít varianty na Houghovu transformaci. Houghova transformace byla navržena v roce 1962 pro detekci úseček (resp. přímek) v obraze, nicméně dá se zobecnit pro práci s libovolným analyticky definovatelným útvarem v rovině. [DH72]

Houghova transformace pracuje s hlasováním v tzv. prostoru parametrů. Hledané útvary v obraze je nutné analyticky parametrizovat, např. přímka se parametrizuje dvojicí parametrů (r, θ) , kde r je vzdálenost přímky od počátku souřadnicového systému a ρ je úhel který přímka svírá s osou x . Pokud bereme v potaz pouze úhly v základním intervalu (např. $\langle 0, 2\pi \rangle$) je touto parametrizací každá přímka určena jednoznačně. Prostor parametrů je plocha určená právě souřadnicemi (r, θ) .

Vstupem Houghovy transformace je binární obraz hran (např. výstup z Cannyho detekčního algoritmu) a prázdný prostor parametrů (tj. prostor parametrů kde každému bodu (r, θ) je přiřazena váha $v(r, \theta) = 0$). Pro každý bod (x, y) v obraze který je hranou lze určit množinu všech přímek které jím mohou procházet, tedy množinu možných (r, θ) . Všem takovým dvojicím přiřadíme novou hodnotu $v_{new}(r, \theta) = v(r, \theta) + 1$. Následně vybereme takové body (r, θ) které splňují $v(r, \theta) > \alpha$, kde α je zvolený parametr (práh) nebo nějaký počet nejvyšších hodnot $v(r, \theta)$ - tyto body nám definují nalezené prvky (v tomto případě přímky).

S rostoucí složitostí hledaných objektů (resp. jejich parametrického vyjádření) roste rozměr prostoru parametrů a tedy i výpočetní náročnost algoritmu. Jako jedna z možností urychlení výpočtu se používá tzv. náhodná Houghova transformace [XOK90]. Namísto iterací přes všechny pixely a pro každý pixel určování možných parametrů jsou z obrazu náhodně nebo sekvenčně vybírány n -tice hranových bodů (kde n musí být rovno minimálnímu počtu bodů definujícímu daný geometrický útvar, tedy např. při hledání přímek $n = 2$, pro kružnice $n = 3$ atd.). Pro ty jsou určeny hodnoty parametrů definující stejný útvar jako vybrané body. Namísto ukládání vah v prostoru parametrů se využívá akumulátor. Pokud je nějaká podobný útvar v akumulátoru (podobný ve významu lišící se jen o definované Δ) je v akumulátoru nahrazen útvarem tvořící (vážený) průměr původního útvaru v akumulátoru a nalezeného útvaru s váhou zvýšenou o jedna, pokud v akumulátoru takový útvar není je do něj vložen s váhou 1. Algoritmus končí buď při nalezení -nté elipsy s váhou více než threshold, nebo při vyčerpání počtu iterací.

4.1 Metody strojového učení

Tyto metody vyžadují natrénování klasifikátoru na předem připravených datech. Využívají tedy míru podobnosti hledaného typu objektu k jiným výskytům tohoto objektu.

Trénování je výpočetně náročné, navíc většinou vyžaduje předzpracovanou trénovací množinu snímků (tedy množinu snímků u kterých je známý požadovaný výsledek). Pro urychlení trénování bylo navrženo mnoho algoritmů (např. adaboost), nicméně většina detektorů stále není trénována od začátku, ale využívá předtrénované klasifikátory, natrénované na velkých obecných datasetech za využití specializovaného hardware.

4.1.1 Konvoluční neuronové síť

Laické veřejnosti nejznámější a jedna z nejpokročilejších metod vyhledávání prvků v obraze je průchod snímku natrénovanou neuronovou sítí. Konvoluční neuronová síť (KNN) je typ umělé neuronové sítě navržený specificky pro práci s digitálním obrazem. [ON15]

Velkou výhodou metod založených na neuronových sítích je možnost hledat téměř libovolné prvky (tyto metody nejsou závislé na matematické parametrizaci hledaných tvarů či barev). Naopak nevýhodou je velká výpočetní náročnost, využití neuronových sítí se dostalo do popředí až v poslední době s rozšířením dostupnosti grafických čipů (GPU). V rámci této práce je zmiňována jen pro svoji značnou rozšířenost v povědomí laiků, pro výpočty na hardwaru se značně omezenou výpočetní kapacitou je zcela nevhodná.

Další nevýhodou je nutnost natrénování neuronové sítě, které kromě velkého výpočetního výkonu požaduje také poměrně velké množství lidské práce (ruční označení a výběr trénovací množiny).

Konvoluční neuronové síť se skládají z mnoha vrstev, které lze rozdělit následovným způsobem [ON15]

1. Vstupní vrstva (input layer) - vlastní hodnoty pixelů zkoumaného obrazu

2. Konvoluční vrstva (convolutional layer) - základní stavební prvek KNN. Konvoluční jádro je natrénované - své hodnoty získá z učení, tedy z předkládání snímků pro které je známý požadovaný výstup. Výstupem konvoluční vrstvy je tzv. příznaková mapa. Konvoluční jádro použité v této vrstvě se někdy označuje též jako detektor příznaků.
3. Sdružovací vrstva (pooling layer) - provede podvzorkování (downsampling), typicky sníží rozměry (dimenzionalitu) příznakové mapy při zachování důležitých informací filtrační operací (např. maximum či průměr okolí pixelu). Tím se docílí jednak rychlejšího výpočtu v další vrstvě, jednak se model stává odolnějším vůči zkreslením na vstupu (která by měla za následek menší změny v příznakové mapě)
4. Plně propojená vrstva (fully-connected layer) - vyprodukuje závěrečné "skóre", na které lze nahlížet jako na pravděpodobnost že hledaný prvek je v obraze přítomen. Plně propojené vrstvě na konci sítě se říká klasifikátor.

Konvolučních a sdružovacích vrstev zpravidla bývá v neuronové síti několik (označují se dohromady jako "skryté vrstvy"), běžně se střídá konvoluční-sdružovací až po výstupní plně propojenou vrstvu.

Ačkoliv pro detekci lidí, obličejů (a mnoha dalších předmětů) v obraze získaném ve viditelném spektru existuje dnes již nemalé množství předtrénovaných sítí (např VGG-19, GoogLeNet [RDS⁺15]), pro data získané ve spektru IR je výběr mnohem omezenější. Podle některých prací [JAaDSV17], že pro detekci obličejů v IR spektru nejsou neuronové sítě vůbec vhodným nástrojem.

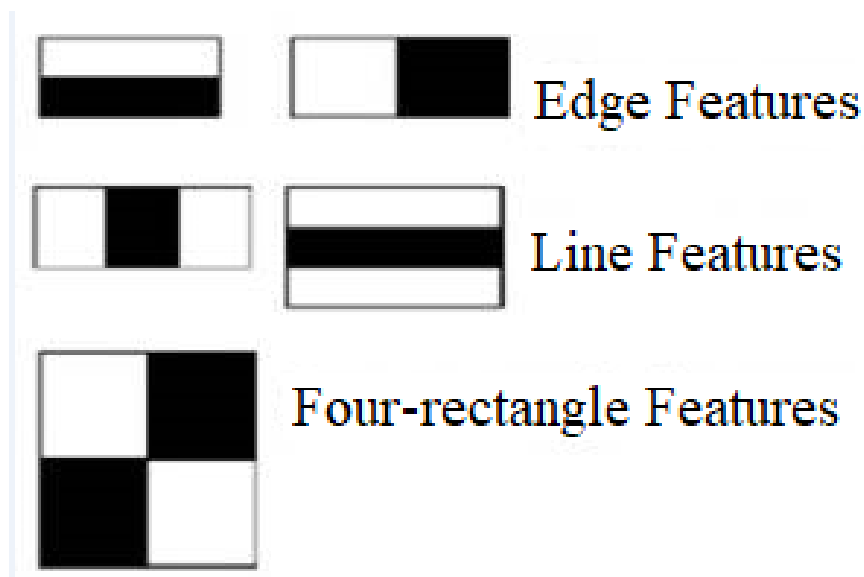
4.1.2 Viola-Jones detektor

V roce 2001 představili Paul Viola a Michael Jones [VJ01] detektor objektů založený na kaskádě předučených klasifikátorů využívajících tzv. Haarovy příznaky. Dnes se jedná o velice využívaný algoritmus, jehož implementaci hledající obličej v obraze (ve viditelném spektru) nalezneme v mnoha mobilních telefonech a fotoaparátech.

Haarovy příznaky

Haarovy příznaky (Haar-like features) jsou pojmenovány podle matematika Alfreda Haara, neboť jsou založeny na jeho návrhu Haarovy vlnky (Haar

wavelet). Tyto funkce Haar zkonstruoval již v roce 1909. Jedná se o příznaky digitálního obrazu, klasifikátor tedy nezpracovává obraz jako takový ale namísto toho pracuje s hodnotami příznaků. To obvykle vede k nižší výpočetní náročnosti.



Obrázek 4.2: Příklad Haarových bází, převzato z [DH72]

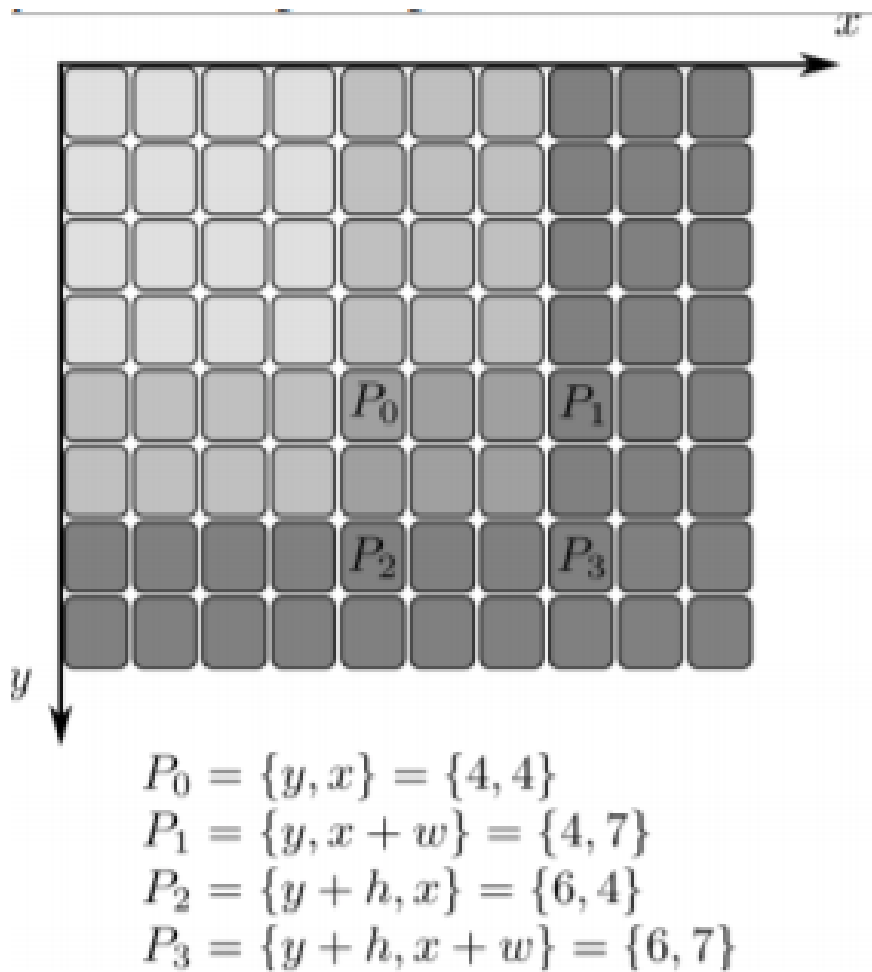
Na obrázku 4.2 je zobrazeno několik Haarových bází. Výpočet hodnoty Haarova příznaku v obraze je roven rozdílu sum hodnot pixelů maskovných částí příznaku zobrazenou bíle a hodnot pixelů maskovaných částí zobrazenou černě (Haarovy příznaky si lze představit jako konvoluci s částí konvoluční maksy o hodnotách -1 a částí o hodnotách 1)

■ Integrovaný obraz

Obyčejné sčítání pixelů pomocí konvoluce je časově náročné. Namísto toho lze jako vstup pro Haarovy příznaky použít integrovaný obraz - tento obraz je maticí o stejném rozměru jako původní obraz, která v každém bodě má sumu všech pixelů nalevo a nahoru od svého indexu.

■ Klasifikátory

Klasifikátor je funkce, která co nejlépe rozděluje klasifikační třídy a obvykle rozhoduje podle nějaké hraniční hodnoty příznaků. Viola-Jones detektor



Obrázek 4.3: Výpočet integrálního obrazu, převzato z [Ope]

používá kaskádu "slabých"klasifikátorů na hodnotách Haarových příznaků. Slabým klasifikátorem rozumíme klasifikátor který má velké množství falešně pozitivních výsledků - jeho výstup se tedy chápe spíš jako "obraz objekt neobsahuje/ obraz objekt možná obsahuje"spíše než "obsahuje/neobsahuje". Pokud je výstup jednoho klasifikátoru pozitivní (tj "možná"), je obraz podstoupen dalšímu klasifikátoru v řadě. Projde-li obraz celou kaskádou s jednotlivými výstupy "možná"považuje se objekt za detekovaný.

Tento přístup výrazně urychluje zpracování zejména obrazů které hledaný objekt neobsahují (obraz může být vyhodnocen jako "negativní"jakoukoliv vrstvou kaskády a nadále se nezpracovává). To je velkou výhodou máme-li zpracovávat datasey kde se hledaný objekt vyskytuje s malou frekvencí.

Kapitola 5

Návrh řešení a implementace

Návrh řešení detekce obličejů pro měření lidské teploty termokamerou byl vypracován pro komerčně dostupný all-in-one systém české společnosti Workswell, kameru Workswell MEDICAS.

Jedná se o systém kamery pracující v IR spektru (bolometrické jádro FLIR TAU2), kameru pracující ve viditelném spektru a výpočetní jednotku zpracovávající a zobrazující výsledky v reálném čase.

Při používání systému Workswell MEDICAS se v zorném poli snímače trvale nachází kalibrační černé těleso o teplotě 40°C pro kontinální kalibraci snímače a zajištění co nejpřesnějšího měření. Pozice tohoto tělesa ve snímaném obraze je nutné ručně nastavit (pro odběry kalibračních vzorků teploty) a je tedy snadné jej z obrazu odfiltrovat (zadefinováním ignorované oblasti okolo daných souřadnic).

Další známou informací je předpokládaná vzdálenost člověka od snímače, jsme tedy schopni odhadnout velikost hledaného obličeje.

■ 5.1 Hardware a software parametry systému

■ 5.1.1 Software

Workswell MEDICAS [Wor] je all-in-one systém využívající operační systém založený na Linuxovém jádře s běžící aplikací psanou v jazycích C/C++. Na systému je nainstalována open-source knihovna pro počítačové vidění OpenCV.

■ 5.1.2 Hardware

■ Výpočetní jednotka

Výpočetní jednotka je realizována na vícejádrovém procesoru architektury ARM s časováním 1GHz.

Za běžného použití systému bez detekce objektů v obraze je vytížení výpočetní jednotky okolo 50% až 60%, při ukládání videa ze snímačů vytížení stoupne na 70% (určeno z výstupu linuxového příkazu `top`)

■ 5.1.3 Snímače

K výpočetní jednotce je připojena termální kamera s rozlišením 640 x 512 pixelů, snímající záření o vlnových délkách 7.5 – 13.5 μm (SWIR) s objektivem 45°x 37°.

Držým snímačem je kamera pro viditelné spektrum s rozlišením 1920x1080 (fullHD).

5.2 Cíl

Podle normy ISO je měření lidské teploty termokamerou pro medicínské účely možná za splnění určitých podmínek. Při využití našeho HW by měření mělo probíhat ze vzdálenosti asi 80cm u kořene nosu (ve vnitřních koutech oka) [ISO16], kde teplota povrchu těla nejlépe odpovídá teplotě tělesného jádra.

Aby byla tato podmínka zajištěna a při překročení bezpečné teploty (teploty neidnikující zdravotní potíže) došlo k automatické akci (varování obsluhy, neotevření turniketu pro vstup do areálu atp.) je vhodné detekovat obličej člověka stojícího v definované vzdálenosti od kamery v obraze a teplotu vyhodnocovat pouze v nalezeném obličej.

5.3 Metody strojového učení

Jak bylo zmíněno v teoretické části, nemusí být pro detekci v IR spektru použití rozšířených metod strojového učení nejvhodnějším přístupem. S tímto názorem se pojí další překážka takového přístupu, tedy fakt že existuje jen velmi malé množství klasifikátorů předtrénovaných na obrazy získané v jednotlivých částech IR spektra (tedy MWIR, SWIR a LWIR termokamer). Např. [JAaDSV17] sice prezentuje srovnávání detekcí v těchto částech elektromagnetického spektra, nicméně použité datasety a klasifikátory nejsou veřejně dostupné.

Konfigurace použitého systému umožňuje využít pro detekci obraz z kamery ve viditelném spektru, použít transformaci potlačující paralax. Vedoucí této bakalářské práce, pan docent Vítek, provedl několik testů na nafačeném datasetu s využitím open source knihovny strojového vidění OpenCV a předtrénovaných kaskád klasifikátoru Viola-Jones v jazyce Python se slibnými výsledky. Navržený algoritmus byl následně přepsán do jazyka C++ a proveden test detekce obličej přímo na cílovém hardware s využitím dat z RGB snímače v reálném čase. Při vypnutí funkcí systému nutné pro celkový běh systému Workswell Medicas ale nikoliv pro samotnou detekci za účelem zvýšení rychlosti detekce bylo dosaženo rychlosti zpracování snímků v řádech jednoho snímku za řádově desítky sekund.

Po této zkušenosti nebyly metody strojového učení dále testovány.

5.4 Segmentace

Využití segmentace soustředící se na vyhledávání oblastí odpovídajících rozmezí teplot přibližně 35°C - 39°C a rozměrově odpovídajících obličejů vyniká svojí snadnou implementací. Bohužel naráží na několik závažných problémů.

Ukázalo se, že míra rovnoměrnosti povrchové teploty lidské tváře je mnohem individuálnější než by se zdálo. Zatímco platí že maximální teplota na povrchu tváře je celkem vypovídající o teplotě organismu, prohřátí zejména tváří je ovlivněno nejen vnějšími vlivy, ale také individuálním prokrvením a dalšími fyziologickými parametry. Měřená povrchová teplota některých jedinců v oblasti tváří klesá i pod uvedených 35°C , a obličej tedy nemusí být v po jedoduché prahové segmentaci součástí jediného segmentu.

Zároveň má velký dopad na velikost segmentu délka a styl účesu vlasů a (u mužů) vousů, přítomnost roušky a mnohdy i tvar a velikost výstříhu, přítomnost brýlí... Množství parametrů ovlivňujících tvar a velikost segmentu obsahujícího obličej, jako i množství segmentů obličeje je příliš velké pro rozumně fungující detekci na tak jednoduchém základě jako je prahování.



Obrázek 5.1: Srovnání výstupů prahování (segmentace) se stejnými parametry na termogramech tváří dvou zdravých lidí

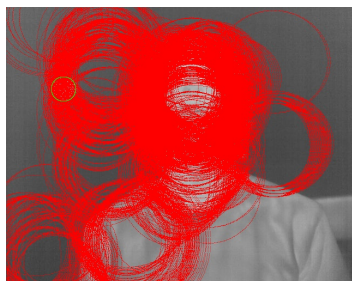
5.5 Vyhledávání obličeje jako elipsy

V této práci je uvedena Houghova metoda jako algoritmus vyhledávající parametrizovatelné tvary v obraze. Tvar obličeje jde v jistém zjedodušení považovat za elipsu, nabízí se tedy vyhledávat elipsu v obraze. Pokud neuvažujeme natočení elipsy vůči souřadným osám má elipsa parametrické

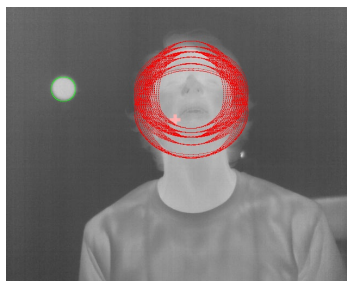
vyjádření

$$x = s_x + a \cos \alpha \quad (5.1)$$

$$y = s_y + b \sin \alpha, \quad \alpha \in \langle 0, 2\pi \rangle \quad (5.2)$$



(a) : Všechny elipsy s více jak 500 hlasy



(b) : Všechny elipsy s více jak 1000 hlasy

Využití náhodné Houghovou transformace by pravděpodobně vedlo k vytváření velmi rozsáhlého akumulátoru elips, resp. že nacházení velké množství elips s malým počtem hlasů, které nemají s tvarem obličeje nic společného a jsou vytvářeny z hran detekovaných na oblečení či pozadí. V závislosti na složitosti obrazu by pak docházelo k velice kolísavému vytížení dostupných HW prostředků, které může způsobit (a při testování jiných algoritmů na daném HW často způsobilo) nečekané zpomalení ostatních úkolů systému, například generování grafického výstupu nebo kontinuální teplotní kalibraci na černém tělese. Tento přístup byl tedy nahrazen původním Houghovým návrhem iterace přes prostor parametrů.

Prostor parametrů pro houghovu transformaci tedy má 5 dimenzí (s_x, s_y, a, b, α) . Uvažujeme-li všechny možné elipsy v obraze pak pro s_x a a připadají v úvahu všechny kladné integery (celá čísla) do šířky obrazu, pro s_y a b pak všechny kladné integery v rozahu do výšky obrazu, možností pro α je teoreticky nekonečno (všechna reálná čísla v intervalu $\langle 0, 2\pi \rangle$), nicméně při výpočtech se samozřejmě použije nějaká diskrétní podmnožina tohoto intervalu. Přes tato omezení je zřejmé že tradiční iterační hlasování v takto rozsáhlém prostoru parametrů v reálném čase je mimo možnosti používaného HW. Je tedy nutné prostor poněkud omezit.

Prvním krokem je omezit se na elipsy o rozměrech které nás skutečně zajímají. Při měření na vzdálenosti $0.8m$ obraz obličeje zabírá průměrně oblast 130×200 pixelů - omezíme tedy intervaly pro poloosy a a b aby odpovídaly tomuto rozměru. Úhly omezíme na diskrétní krok 2 deg . Protože hlasování je implementováno jako vnořené cykly iterující přes prostor parametrů, ukazuje se toto optimalizační omezení stále málo. Jediné zatím neomezené parametry jsou pozice středu elipsy. Určitě nemusíme vyhledávat elipsy se středem mimo obraz, nebo velice blízko jeho okraji. Taková implementace se svoji rychlostí

Kapitola 6

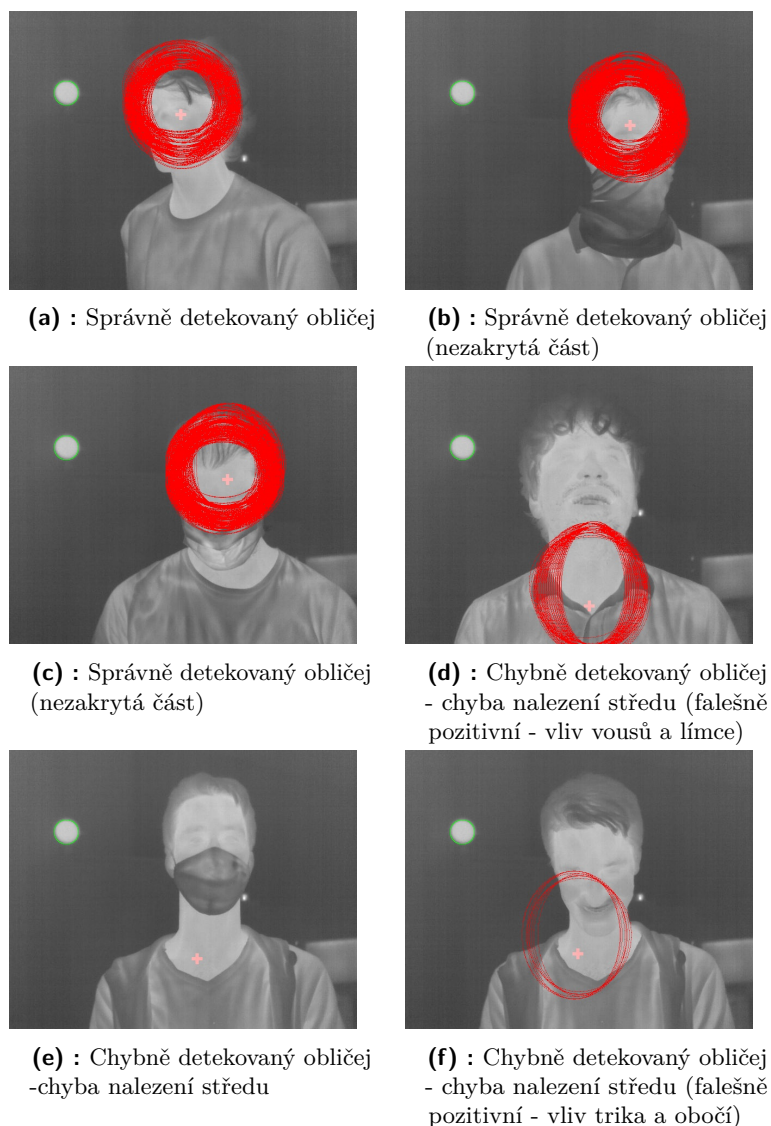
Výsledky

Na nafoceném datasetu 120 snímků obličejů má navržený algoritmus míru úspěšné detekce (nalezené elipsy pokrývají většinu obličejů) okolo 50% v závislosti na zvolených prazích segmentace (pixelů uvažovaných pro výpočet středu elipsy). Tento výpočet je velmi ovlivněn případnou přítomností roušky či jiného způsobu zakrytí dýchacích cest, právě snímky s rouškami mají největší podíl mezi nerozpoznanými obličejů.

Zvolení vhodného středu je pro současnou podobu algoritmu kritické. Omezení diktovaná hardwarem neumožňují přílišné rozšíření prostoru parametrů, a omezení možných středů elips je pro udržení zpracování v reálném nezbytné, chybně určený střed však zaručuje selhání algoritmu. Při přítomnosti chirurgické roušky může vést k chybné detekci oblasti mezi lemem oděvu (trička) a lemem roušky, tedy prakticky krku, jako brady (asi 25% případů v datasetu).

Důležitým parametrem pro výpočet středu je teplota okolí. Jednak teplota objektů v pozadí -pokud se v obraze vyskytuje velké množství ploch s teplotou blízkou teplotě obličejů, dochází k ovlivnění odhadovaného středu elips, tak teplota vzduchu. Lidská tvář v prostředí o 20°C vykazuje jiné charakteristiky prohřátí než lidská tvář v -5°C, v těsném okolí očí teplota odpovídá teplotě jádra ale povrchová teplota čela, uší a lící se může velmi lišit.

Pokud není na obličejů přítomna rouška a není-li obličej natočen vůči kameře vykazuje pokusy v reálném provozu vysoké úspěšnosti. Při praktických testech na hardware s navrženým algoritmem nedochází k zpomalení ostatních funkcí systému, vyžití procesoru se drží okolo 85%.



Obrázek 6.1: Příklady nalezených elips při detekci obličeje (zeleně - maska kalibračního černého tělesa pro výpočet středu, kříž - odhadnutý střed, červeně - nalezené platné elipsy)

Dochází k určitým nesrovnalostem mezi výsledky naloženými v datasetu a zpracováním přímo na kameře. Při zpracování přímo na kameře je dosaženo lepších výsledků díky většímu teplotnímu rozlišení (hloubce) dostupného obrazu. Při provozu na kameře nejsou pozorovány časté falešně pozitivní detekce jako v případě přístupu založeného čistě na segmentaci prahováním.

Z obrázku ?? je zřejmé, že navržený algoritmus je schopný obličej detekovat určením oblasti zakryté daným množstvím elips, nicméně jeho výpočetní náročnost je srovnatelná s metodami klasifikátorů založených na strojovém učení. Na rozdíl od nich lze navržený algoritmus zjednodušit za cenu predikovatelné

ztráty přesnosti.



Kapitola 7

Závěr

V rámci této práce byly prozkoumány algoritmy pro zpracování obrazu, se zvláštním soustředěním na obraz získaný z mikrobolometrického pole. Pozornost byla věnována algoritmům věnujícím se vlastnostem obrazu, jako je prahování a hledání hran, hledání geometrických tvarů a jejich možného využití pro detekci lidských obličejí v termogramu.

Navržený algoritmus, který je kombinací prahování a vyhledávání elips v obraze přístup je možné za dodržení "dostatečně nízké" (do 30°C) teploty pozadí použít k detekci nezakrytých obličejů s nízkým číslem falešně detekovaných obličejů, při vyšších teplotách pozadí však naráží na svá omezení a selhává. Na hardware s vyšším výpočetním výkonem by bylo možné využít dalších možností pro jeho zlepšení, nebo jej použít jako součást rozsáhlejší kaskády detekčních algoritmů.



Příloha A

Literatura

- [BB06] Stephen J. Blundell and Katherine M. Blundell, *Concepts in thermal physics*, Oxford niversity Press, 2006.
- [Can86] John Canny, *A computational approach to edge detection*, IEEE Transactions on Pattern Analysis and Machine Intelligence **PAMI-8** (1986), no. 6, 679–698.
- [DABL06] A. Dawoud, M.S. Alam, A. Bal, and C. Loo, *Target tracking in infrared imagery using weighted composite reference function-based decision fusion*, IEEE Transactions on Image Processing **15** (2006), no. 2, 404–410.
- [DH72] Richard O. Duda and Peter E. Hart, *Use of the hough transformation to detect lines and curves in pictures*, Commun. ACM **15** (1972), no. 1, 11–15.
- [dro20] *What are swir, mwir, and lwir - and what do they mean?*, Dec 2020, [Online][cit 11.7.2021].
- [GW08] Rafael C. Gonzalez and Richard E. Woods, *Digital image processing (3rd ed.)*, Pearson Prentice Hall, 2008.
- [HCJH13] Wang Huawei, Ma Caiwen, Cao Jianzhong, and Zhang Haifeng, *An adaptive two-point non-uniformity correction algorithm based on shutter and its implementation*, 2013 Fifth International Conference on Measuring Technology and Mechatronics Automation, 2013, pp. 174–177.
- [HZH21] Yong Huang, Nan Zhang, and Qun Hao, *Real-time noise reduction based on ground truth free deep learning for optical*

- coherence tomography*, Biomedical optics express (2021), no. 12, 4.
- [ISO16] ISO, *Iso/iec 80601*, Standard, ISO, Březen 2016.
- [ISO17] ———, *Nedestruktivní zkoušení – infračervená termografie – slovník*, Standard, ISO, Březen 2017.
- [JAaDSV17] Aishwarya Pant Janhavi Agrawal and, Tejas I. Dhamecha, Richa Singh, and Mayank Vatsa, *Understanding thermal face detection: Challenges and evaluation*, SPRINGER, 2017.
- [KSK⁺20] Masafumi Kidoh, Kensuke Shinoda, Mika Kitajima, Kenzo Isogawa, Masahito Nambu, Hiroyuki Uetani, Kosuke Morita, Takeshi Nakaura, Machiko Tateishi, Yuichi Yamashita, and Yasuyuki Yamashita, *Deep learning based noise reduction for brain mr imaging: Tests on phantoms and healthy volunteers*, Magn Reson Med Sci. (2020), 195–206.
- [ON15] Keiron O’Shea and Ryan Nash, *An introduction to convolutional neural networks*, ArXiv e-prints (2015).
- [Ope] OpenCV, *Opencv2 reference manual*, <http://www.opencv.org.cn/opencvdoc/2.3.1/opencv2refman.pdf>.
- [RDS⁺15] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg, and Li Fei-Fei, *Imagenet large scale visual recognition challenge*, International Journal of Computer Vision (IJCV) (2015), 211–252.
- [SH02] Robert Siegel and John Howell, *Thermal radiation heat transfer*, Taylor & Francis, 2002.
- [Tog89] T Togawa, *Non-contact skin emissivity: measurement from reflectance using step change in ambient radiation temperature*, Clinical Physics and Physiological Measurement **10** (1989), no. 1, 39–48, [Online][cit 14.6.2021].
- [VJ01] Paul Viola and Michael Jones, *Robust real-time object detection*, vol. 57, 01 2001.
- [VMSMCC09] C. Villasenor-Mora, F.J. Sanchez-Marin, and S. Calixto-Carrera, *An indirect skin emissivity measurement in the infrared thermal range through reflection of a co2 laser beam*, Rev. Mex. Fís. (2009), 387–392, [Online][cit 14.6.2021].
- [Wor] Workswell, *Workswell medicas datasheet*, <https://workswell-thermal-camera.com/thermal-imaging-camera-airport-flue-virus-screening/>.

- [XOK90] Lei Xu, Erkki Oja, and Pekka Kultanen, *A new curve detection method: Randomized hough transform (rht)*, Pattern Recognition Letters **11** (1990), no. 5, 331–338.



Příloha B

Obsah prilozeneho adresare

- data - složka obsahující možný testovací dataset snímků z termokamery (126 snímků ve formátu JPG)
- pythonCodes - testované algoritmy v jazyce python
 - edgeDetect.py - algoritmy detekce hran
 - hough.py - algoritmy detekce elpis
- cCodes - algoritmy vybrané algoritmy v jazyce C++ pro testování přímo na HW
 - eyeDetect (.h, .cpp) - algoritmy Viola-Jones pro detekci ve viditelném obraze
 - edgeDetection (.h, .cpp) - algoritmy Houghovy transformace pro detekci v termoobraze