



**VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ**

BRNO UNIVERSITY OF TECHNOLOGY

**FAKULTA INFORMAČNÍCH TECHNOLOGIÍ**

FACULTY OF INFORMATION TECHNOLOGY

**ÚSTAV POČÍTAČOVÉ GRAFIKY A MULTIMÉDIÍ**

DEPARTMENT OF COMPUTER GRAPHICS AND MULTIMEDIA

**VYHODNOCENÍ METOD STEREOZPRACOVÁNÍ  
OBRAZU**

EVALUATION OF METHODS OF STEREO IMAGE PROCESSING

**BAKALÁŘSKÁ PRÁCE**

BACHELOR'S THESIS

**AUTOR PRÁCE**

AUTHOR

**IVO JURÁČEK**

**VEDOUcí PRÁCE**

SUPERVISOR

**prof. Dr. Ing. PAVEL ZEMČÍK**

BRNO 2017

**Vysoké učení technické v Brně - Fakulta informačních technologií**

Ústav počítačové grafiky a multimédií

Akademický rok 2016/2017

**Zadání bakalářské práce**

Řešitel: **Juráček Ivo**

Obor: Informační technologie

Téma: **Vyhodnocení metod stereozpracování obrazu**  
**Evaluation of Methods of Stereo Image Processing**

Kategorie: Zpracování obrazu

Pokyny:

1. Prostudujte dostupnou literaturu na téma metody stereo zpracování obrazu.
2. Navrhněte metody vyhodnocení přesnosti dosažitelné stereo metodami při měření polohy a/nebo rychlosti.
3. Vyberte vhodnou metodu vyhodnocení a zhodnoťte dosažitelné vlastnosti měření.
4. Implementujte vybranou metodu a demonstруйте funkčnost na vhodném příkladu.
5. Diskutujte výsledky a možnosti pokračování v práci.

Literatura:

- Dle pokynů vedoucího

Pro udělení zápočtu za první semestr je požadováno:

- Body 1 až 3 zadání

Podrobné závazné pokyny pro vypracování bakalářské práce naleznete na adrese

<http://www.fit.vutbr.cz/info/szz/>

Technická zpráva bakalářské práce musí obsahovat formulaci cíle, charakteristiku současného stavu, teoretická a odborná východiska řešených problémů a specifikaci etap (20 až 30% celkového rozsahu technické zprávy).

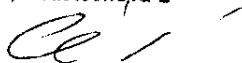
Student odevzdá v jednom výtisku technickou zprávu a v elektronické podobě zdrojový text technické zprávy, úplnou programovou dokumentaci a zdrojové texty programů. Informace v elektronické podobě budou uloženy na standardním nepřepisovatelném paměťovém médiu (CD-R, DVD-R, apod.), které bude vloženo do písemné zprávy tak, aby nemohlo dojít k jeho ztrátě při běžné manipulaci.

Vedoucí: **Zemčík Pavel, prof. Dr. Ing.**, UPGM FIT VUT

Datum zadání: 1. listopadu 2016

Datum odevzdání: 17. května 2017

**VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ**  
Fakulta informačních technologií  
Ústav počítačové grafiky a multimédií  
612 66 Brno, Božetěchova 2



---

doc. Dr. Ing. Jan Černocký  
vedoucí ústavu

## Abstrakt

Tato práce pojednává o určení vzájemné korelace mezi dvěma snímky a její aplikaci na reálný problém s využitím kamery, která je součástí stereo kamery. Naznačuje, jakým způsobem lze využít vzájemnou korelaci pro hledání určitých vzorů v obraze, přičemž lze dopočítat posun mezi dvěma snímky při různém zvětšení obrazu, které je provedeno převzorkováním. Přesnost korelace se poté promítne při kompenzaci snímků pomocí afinní transformace, která slouží jako ověření správnosti vzájemné korelace. Pro další zvýšení přesnosti korelace je využit detektor rohů, který v jednom snímku nalezne výrazné body, které se potom pomocí vzájemné korelace hledají v jiném snímku. Dosažené výsledky touto metodou jsou též zpracovány v této práci.

## Abstract

This thesis addresses the correlation between two pictures and its application to a real example using a camera which is a part of a stereo camera. It suggests how to use correlation to find specific patterns in pictures whereas it is possible to calculate the movement between two scaled pictures which is done by resampling. The accuracy of correlation is then shown when pictures are compensated using affine transformation in order to verify the accuracy of correlation. To further increase the correlation accuracy, it is possible to use a corner detector, it finds main points in a shot and then it searches for a match in another shot using correlation. Final results using this method are also mentioned in this thesis.

## Klíčová slova

Stereo kamera, počítačové vidění, OpenCV, korelace, afinní transformace, Harrisův detektor rohů, převzorkování

## Keywords

Stereo camera, computer vision, OpenCV, correlation, affine transformation, Harris corner detector, resampling

## Citace

JURÁČEK, Ivo. *Vyhodnocení metod stereozpracování obrazu*. Brno, 2017. Bakalářská práce. Vysoké učení technické v Brně, Fakulta informačních technologií. Vedoucí práce Zemčík Pavel.

# Vyhodnocení metod stereozpracování obrazu

## Prohlášení

Prohlašuji, že jsem tuto bakalářskou práci vypracoval samostatně pod vedením pana prof. Dr. Ing. Pavla Zemčíka. Další informace mi poskytl Ing. Adam Široký. Uvedl jsem všechny literární prameny a publikace, ze kterých jsem čerpal.

.....

Ivo Juráček  
16. května 2017

## Poděkování

Rád bych poděkoval svému vedoucímu bakalářské práce, prof. Dr. Ing. Pavlovi Zemčíkovi za vedení a konzultace při vypracování tohoto textu. Dále bych rád poděkoval Ing. Adamovi Širokému za cenné rady pro implementaci řešeného problému.

# Obsah

<b>1</b>	<b>Úvod</b>	<b>3</b>
<b>2</b>	<b>Obraz</b>	<b>4</b>
2.1	Obrazová funkce . . . . .	4
2.2	Digitalizace a histogram . . . . .	5
<b>3</b>	<b>Operace nad obrazem</b>	<b>8</b>
3.1	Konvoluce a korelace . . . . .	8
3.2	Prahování . . . . .	9
3.3	Ekvalizace histogramu . . . . .	9
3.4	Geometrické transformace obrazu . . . . .	10
3.5	Vyhledání významných bodů v obraze . . . . .	14
<b>4</b>	<b>Stereo kamera</b>	<b>16</b>
4.1	Stereo vidění . . . . .	16
4.2	Epipolární geometrie . . . . .	17
<b>5</b>	<b>Zhodnocení stavu a specifikace</b>	<b>19</b>
5.1	Popis problému . . . . .	19
5.2	Návrh řešení . . . . .	20
<b>6</b>	<b>Tvorba testovacích dat</b>	<b>23</b>
6.1	Snímání obrazu . . . . .	23
6.2	Vyznačení oblastí . . . . .	24
6.3	Popis vytvořených sad snímků . . . . .	25
<b>7</b>	<b>Určení vzájemné korelace obrazů a jejich kompenzace</b>	<b>27</b>
7.1	Načtení referenčního snímku a extrakce anotovaných oblastí z referenčního snímku . . . . .	28
7.2	Vytvoření výřezů oblasti ze sady snímků . . . . .	29
7.3	Vyhledání rohů v jednotlivých oblastech . . . . .	30
7.4	Určení vzájemné korelace mezi výřezy oblastí u dvou snímků . . . . .	31
7.5	Kompenzace obrazů mezi referenčním a jiným snímkem ze sady snímků . . . . .	33
7.6	Generování výsledků a kontrola správnosti . . . . .	35
<b>8</b>	<b>Závěr</b>	<b>40</b>
	<b>Literatura</b>	<b>41</b>

<b>Přílohy</b>	<b>43</b>
<b>A Konfigurační soubor</b>	<b>44</b>
<b>B Seznam použitých OpenCV funkcí</b>	<b>46</b>

# Kapitola 1

## Úvod

Pomocí počítačového vidění, což je obor, který se zabývá získáváním informací z obrazu, lze v současné době zkoumat širokou škálu problémů, jako například automatické ovládání vozidel, zaznamenávání trajektorie pohybu lidí či zařazování obrázků do kategorií podle toho, co se na obrázku vyskytuje. Dá se rovněž využít pro sledování dopravní situace na místních komunikacích a v případě problému (autonehoda) může dojít k detekci a vyhodnocení takové situace.

Aby ovšem takové sledování prostoru fungovalo dobře, je nutné mít přesně zkalibrovanou kameru. Vlivem teplotní roztažnosti či vlivem počasí (vítr, déšť) může dojít k mírnému pohybu pevně uchycené kamery, což může způsobit značné odchylky od normálně získávaných hodnot. Výsledné výpočty jsou poté touto odchylkou zkresleny a v extrémních případech se takový systém může stát zcela zbytečným.

Jedna z metod zjišťování takových odchylek je popsána v následujících kapitolách tohoto textu, kdy je využito principu vzájemné korelace dvou, v časově dlouhém horizontu pořízených snímků, kde se určuje vzájemná změna oproti původně zkalibrované hodnotě. Tato změna se poté kompenzuje pomocí afinní transformace.

V první kapitole je definován obraz pomocí matematického modelu a způsob digitalizace obrazu pomocí vzorkování a kvantování pro další zpracování výpočetními systémy. Též je zde popsán i histogram. V druhé kapitole jsou definovány operace nad obrazem. Matematickým modelem je zde popsán princip konvoluce, korelace, prahování a ekvalizace histogramu. Dále jsou definovány geometrické transformace obrazu a princip mapování pixelů při převzorkování obrazu u zvětšení obrazu či afinní transformaci. Další uvedenou operací nad obrazem je vyhledání významných bodů, především rohů, vyskytujících se v obraze. V třetí kapitole je definováno, co to je stereo kamera. Dále zde je uvedeno, co to je stereo vidění a epipolární geometrie. Čtvrtá kapitola obsahuje zhodnocení problému, který řeší tato práce. Jsou zde naznačeny možné postupy pro detekci změny snímané scény z jednoho snímku ze stereo kamery a následný postup pro kompenzaci obrazu. Jsou zde uvedeny možné technické prostředky pro řešení problému a popsán vybraný postup řešení problému. Pátá kapitola popisuje tvorbu testovacích snímků pro řešení tohoto problému. Vysvětluje nároky na jednotlivé snímané scény a popisuje jejich anotaci. Šestá kapitola detailně popisuje jednotlivé kroky a operace, které se pro řešení tohoto problému použily. Na konci jsou uvedeny dosažené výsledky použité metody zpracování obrazu.

# Kapitola 2

## Obraz

Slovo obraz se dá pochopit mnoha způsoby, záleží na konkrétním vědním oboru, který si definici takového slova upravuje pro své potřeby. Dá se chápat jako průmět reálného světa na sítnici oka, jako krajina zachycená na plátno či fotografii nebo displej elektrického zařízení zobrazující nějaké informace. Formálně lze použít definici pomocí matematického modelu.

Obraz se též dá popsat jako grafické vyjádření v nějakém, většinou dvojrozměrném prostoru. Digitální obraz, který je reprezentován v binární soustavě, lze vytvořit buď v grafickém editoru, nebo získat ze sensorů, většinou z fotografických zařízení citlivých na světlo, ale i z radarů, lidarů, tomografických zařízení apod. Ve většině případů jsou výsledná data 2D obraz, ale mohou vznikat i 3D objemy či sekvence obrazů. [17, str. 12]

Sejmutím obrazu se rozumí převedení optické veličiny na veličinu elektrickou. Při snímání je důležité použít snímací zařízení, uspořádání snímané scény, vhodné osvětlení či nastavení příslušného obrazového čidla má na další zpracování vliv. Špatným snímáním dochází ke ztrátě informací, kterou většinou nelze získat, lze se jim pouze přiblížit (například zaostřením pořízeného snímku při špatně zaostřené snímací kameře). Běžně používané zařízení pro snímání obrazu jsou scannery, fotoaparáty či videokamery. [12, str. 36]

### 2.1 Obrazová funkce

Většina snímačů reprezentuje zachycenou scénu pomocí spojitého analogového signálu, jak je uvedeno v [15, str. 39], který se nazývá *obrazová funkce*:

$$z = f(x, y). \quad (2.1)$$

Obrazová funkce je omezená, jelikož obraz má pevně stanovené rozměry, definičním oborem je kartézský součin dvou spojitých intervalů:

$$D = \langle x_{min}, x_{max} \rangle \times \langle y_{min}, y_{max} \rangle, \quad (2.2)$$

kde  $x : x_{min} \leq x \leq x_{max}$  a  $y : y_{min} \leq y \leq y_{max}$ ,  $x, y \in \Re$  jsou souřadnice bodu v kartézském souřadném systému, ve kterých obrazová funkce nabývá hodnot  $z \in H$ . Oborem hodnot  $H$  může být jas, intenzita barvy apod. Hodnota obrazové funkce na určité souřadnici může mít v počítačové grafice více hodnot, nejčastěji se jedná o trojici červené, zelené a modré složky barevného modelu RGB, které obecně tvoří uspořádanou  $n$ -tici  $z = [z_1, z_2, z_3, \dots, z_n]$ . Obor hodnot je zpravidla omezen počtem barev, například ve stupních šedi je 256 hodnot barvy, v modelu RGB je 16 milionů barev apod. Celkově lze tedy obrazovou funkci zapsat takto:

$$f : (\langle x_{min}, x_{max} \rangle \times \langle y_{min}, y_{max} \rangle) \rightarrow (H_1 \times H_2 \times \dots \times H_n). \quad (2.3)$$

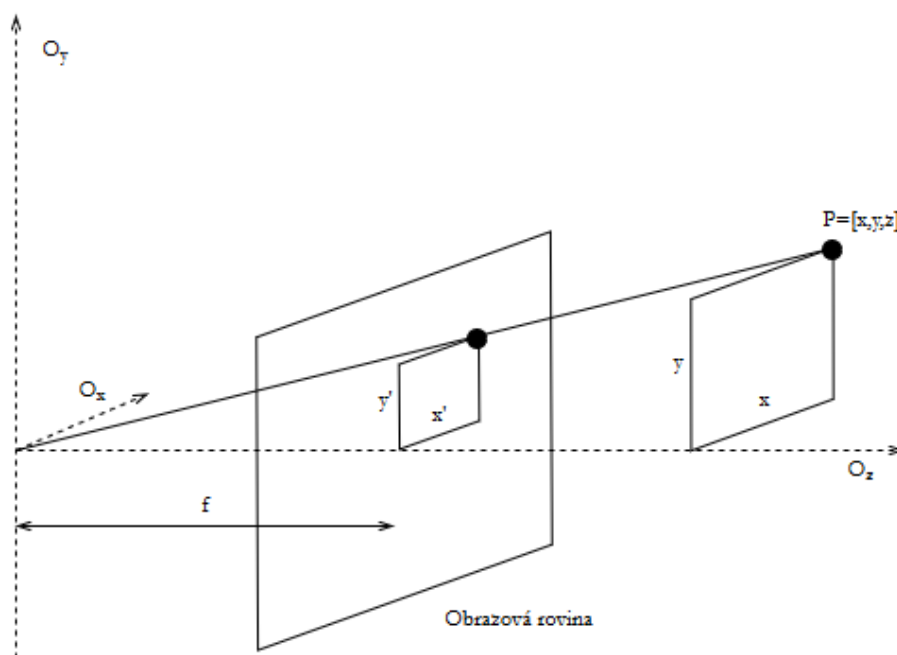


V počítačové grafice se většinou nepracuje se spojitým definičním oborem, ale pracuje se s rastrem. Ten je složen z obrazových elementů, které se nazývají *pixely* (z anglického *picture element*).

Člověk se pohybuje ve trojrozměrném prostoru (3D) a většinou chce obrazovou funkci získat pro dvojrozměrný prostor (2D). Toho lze docílit *perspektivním zobrazením* jisté části 3D prostředí. [12, str. 22] Získaný model je velice realistický a odpovídá získání obrazu z dírkové komory. Je-li bod na souřadnici  $(x, y, z)$  v 3D scéně a  $f$  je ohnisková vzdálenost, tak bod po perspektivním zobrazení bude mít v obrazové rovině souřadnice

$$x' = \frac{xf}{z}, y' = \frac{yf}{z}. \quad (2.4)$$

Geometrie perspektivního zobrazení je ukázána na obr. 2.1



Obrázek 2.1: Perspektivní zobrazení

## 2.2 Digitalizace a histogram

Protože je obraz na obrazovce diskretní, nelze tedy přesně zobrazit spojitou funkci. Je nutné přejít ze spojitě funkce  $f(x, y)$  na funkci diskretní  $I_{i,j}$ , kde je důležité obraz reprezentovat diskretní datovou strukturou, například maticí. [17, str. 14] Tato operace se nazývá *digitalizace* a probíhá ve dvou krocích: *kvantování* a *vzorkování*:

- **kvantování:** obor hodnot  $H$  obrazové funkce  $z = f(x, y)$  je rozdělen na několik intervalů, kde je každý interval reprezentován jedinou zástupnou hodnotou. Může se jednat o rovnoměrné (uniformní) rozložení, kde jsou délky intervalů konstantní, popřípadě se může jednat o nerovnoměrné (neuniformní) rozložení, kde jsou délky

intervalů proměnné. Častěji se používá rovnoměrné kvantování, jelikož je snadnější na realizaci v zařízeních pořizující digitální obraz. [15, str. 41]

Při výběru zástupné hodnoty se nejčastěji využívá průměr celého intervalu, vážený průměr či medián. Tím, že se vybírá jedna reprezentující hodnota z množiny hodnot daného intervalu, dochází ke ztrátě informace, nazývané též jako *kvantizační chyba*. Ta se projevuje jako náhlý skok barev (tzv. Machovy proužky, anglicky *Mach-band effect*) na plochách s malou změnou gradientu. Z tohoto důvodu se vybírá zástupná hodnota tak, aby byla kvantizační chyba co nejmenší.

- **vzorkování:** zaznamenávání hodnot (vzorků) v předem daných intervalech ze spojitě funkce  $f(x, y)$ . Funkci získanou pravidelným vzorkováním  $I_{i,j}$  získáme jako:

$$I_{i,j} = f(x_0 + i \Delta x, y_0 + j \Delta y), i = 0, 1, \dots, j = 0, 1, \dots, \quad (2.5)$$

kde  $\Delta x$  a  $\Delta y$  jsou vzdálenosti dvou vzorků. Původní spojitá funkce  $f(x, y)$  je definována na intervalu  $x \in \langle x_0, x_1 \rangle, y \in \langle y_0, y_1 \rangle$ .

Stanovuje se *vzorkovací frekvence*  $f_s$ , která má hodnotu

$$f_s(x) = \frac{1}{\Delta x}, f_s(y) = \frac{1}{\Delta y}, \quad (2.6)$$

kde  $\Delta x$  a  $\Delta y$  jsou vzdálenosti dvou vzorků. Jednotkou vzorkovací frekvence je počet vzorků za jednotku času (Hz) či počet vzorků na jednotku vzdálenosti (dpi). Čím vyšší bude vzorkovací frekvence, tedy čím větší bude rozlišení obrazu, tím více informací se zaznamená ze spojitě funkce  $f(x, y)$ .

Pro vzorkovací frekvenci existuje vzorkovací teorém (též *Shannonův, Niquistův* či *Kotělníkovův vzorkovací teorém*), který určuje, jak vysoké vzorkovací frekvence je vhodné pro daný typ signálu použít. [14, str. 28] Teorém říká, že při použití vzorkovací frekvence

$$F_s > 2F_{max}, \quad (2.7)$$

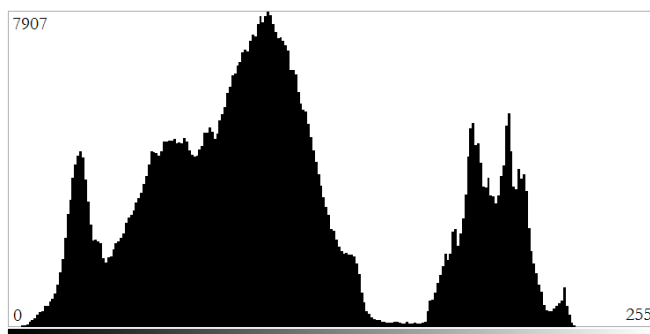
kde  $F_s$  je vzorkovací frekvence, a  $F_{max}$  je mezní frekvence, již nedochází ke ztrátě informací.

Vzorkování se využívá například při natáčení videokamerou, která snímá trojrozměrnou scénu v diskrétních intervalech či zvukový soubor formátu MP3, který je uložen jako posloupnost čísel, reprezentujících úroveň signálu snímanou s vysokou frekvencí. [15, str. 42]

Digitalizace tedy sestává ze vzorkování, kde funkce  $f(x, y)$  je navzorkována do matice s  $M$  řádky a  $N$  sloupci, a z kvantování, kde je každému vzorku přiřazena celočíselná hodnota.

Dále je při digitalizaci důležité vybrat mřížku, do které se uspořádají jednotlivé body. Nejčastěji se používá pravidelná čtvercová mřížka, ale používá se třeba i hexagonální mřížka. Čtvercová mřížka je snadnější k realizaci a odpovídá i reprezentaci pomocí matice, hexagonální mřížka má výhodu v pravidelnosti vzhledem k okolním šesti bodům a jejich stejné vzdálenosti. [12, str. 25]

Pro zobrazení jasové složky v digitálním obraze se využívá histogram jasu, což je vektor s počtem složek odpovídajících jasových úrovní. Hodnoty složek odpovídají četnosti bodů příslušných jasů, zastoupených v obraze. [12, str. 33] Často se zobrazuje jako sloupcový diagram, kde na vodorovné ose jsou jasové úrovně a na svislé četnost vzorků:



Obrázek 2.2: Histogram

Na obrázku 2.2 je ukázka histogramu, který byl získaný z obrázku 7.2. Obsahuje řadu lokálních minim a maxim, přičemž globální maxima reprezentují hodnoty, které se v obraze vyskytují nejčastěji.

## Kapitola 3

# Operace nad obrazem

Operacemi nad obrazem se rozumí provádění takových operací, které při výpočtech používají násobení maticí nebo skalárem či konvoluci nebo korelaci.

### 3.1 Konvoluce a korelace

**Konvoluce** je matematická operace, která se označuje symbolem  $\star$ . [15, str. 46] Konvoluce dvou funkcí  $I(x)$  a  $h(x)$  je definována takto:

$$I(x) \star h(x) = \int_{-\infty}^{\infty} I(x - \alpha)h(\alpha)d\alpha. \quad (3.1)$$

Funkce  $h(x)$  se nazývá *konvoluční jádro* a jednotlivé hodnoty konvolučního jádra určují způsob výpočtu nového pixelu v obraze.

Při práci s digitálním obrazem se používá *diskrétní konvoluce*, která vychází z dvojrozměrné rovnice:

$$I'_{i,j} = I_{i,j} \star h_{i,j} = \sum_{x=-k}^k \sum_{y=-k}^k I_{i-x,j-y}h_{i,j}. \quad (3.2)$$

Při úpravě obrazu pomocí konvoluce, kde konvoluční jádro není skalárem, nastává problém s okrajem obrazu, neboť pro výpočet pixelu upraveného obrazu se zpracovávají i hodnoty jeho blízkého okolí. Jelikož na okraji jsou hodnoty mimo obraz nedefinovány, nastává otázka, jestli je vhodné hodnoty jednotlivých pixelů okrajů duplikovat, využít pixely na protějších stranách okraje či upravit neznámé hodnoty okolí bodu předem nadefinovanou hodnotou. Volba určené metody je individuální podle potřeb uživatele.

**Korelace** je v matematice určení vzájemného vztahu mezi dvěma funkcemi. [13]

$$F \circ I(x) = \int_{-\infty}^{\infty} F(i)I(x + i)di. \quad (3.3)$$

Obdobně jako u konvoluce, i zde je nutno používat diskrétní vzorky:

$$F \circ I(x, y) = \sum_{j=-N}^N \sum_{i=-M}^M F(i, j)I(x + i, y + j) \quad (3.4)$$

Vzájemná korelace (anglicky *cross-correlation*) slouží k určení vzájemné podobnosti dvou funkcí, občas nazývaným jako pohyblivý skalární součin. Využívá se k vyhledávání vzorů v obrazech, kryptoanalýze či neurofyziologii. [9]

Pro spojité funkce  $f$  a  $g$  je definován následovně:

$$(f \star g)(\tau) = \int_{-\infty}^{\infty} f^*(t) g(t + \tau) dt, \quad (3.5)$$

kde  $f^*$  je komplexní sdružení funkce  $f$  a  $\tau$  je posun. Pro diskrétní systémy je použit obdobný zápis [5]:

$$(f \star g)[n] = \sum_{m=-\infty}^{\infty} f^*[m] g[m + n] \quad (3.6)$$

Normalizovaná vzájemná korelace využívá principů vzájemné korelace, ovšem při každém počítání vzorku je provedeno vydělení průměrnou směrodatnou odchylkou: [11]

$$R(x, y) = \frac{\sum_{x', y'} (T(x', y') \cdot I(x + x', y + y'))}{\sqrt{\sum_{x', y'} T(x', y')^2 \cdot \sum_{x', y'} I(x + x', y + y')^2}}, \quad (3.7)$$

Tato normalizovaná vzájemná korelace se používá pro vyhledávání vzorů v obrazech i při řešení tohoto problému.

## 3.2 Prahování

Důležitou operací v počítačové grafice je prahování, což je rozdělení jasové složky do několika množin a nahrazení každé složky jednou hodnotou, která reprezentuje danou množinu. [15, str. 159]

V případě binárního prahování, tedy prahování s jedním prahem a dvěma výslednými množinami, bude předpis vypadat následovně:

$$dst(x, y) = \begin{cases} M & \text{pro } src(x, y) \geq thresh \\ L & \text{pro } src(x, y) < thresh, \end{cases} \quad (3.8)$$

kde  $dst(x, y)$  je výstupní hodnota,  $src(x, y)$  je vstupní hodnota,  $thresh$  je hodnota prahu a koeficienty  $M$  a  $L$  symbolizují dvě výsledné hodnoty, které  $dst(x, y)$  může nabývat.

Prahování se používá například jako mezikrok při detekci hran v obraze či při tvorbě masek, které se používají pro další zpracování v obraze. Tak je tomu i při řešení tohoto problému, kde je vytvořena binární maska obrazu, kde vysoká jasová úroveň této masky reprezentuje body, zahrnující se do výpočtu pro hledání vzájemné korelace.

## 3.3 Ekvalizace histogramu

Ekvalizací (vyrovnáním) histogramu se rozumí nalezení vhodné mapovací funkce, která rovnoměrně rozloží hodnoty histogramu. Po provedení této operace bude výsledkem jasově vyrovnaný obraz. [15, str. 162]

Bude-li mít obraz rozlišení  $x \times y$ , kde bude maximální intenzita označena jako  $max$ , tak ideálně vyrovnaný obraz bude mít hodnotu zastoupené stejné četnosti  $d$  jako průměr všech hodnot

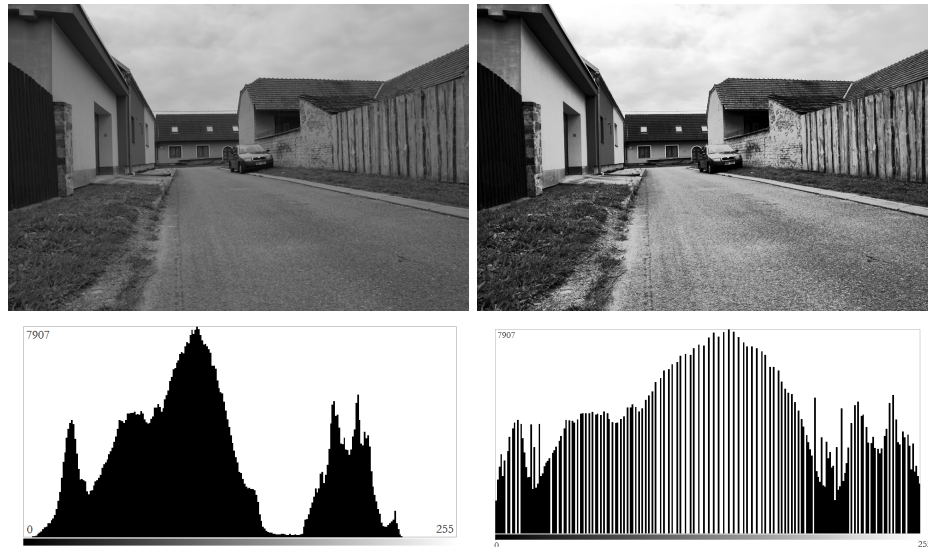
$$d = \frac{x \times y}{max}. \quad (3.9)$$

Některé intenzity v histogramu nelze rozdělit, tedy překračuje-li některá jasová složka ideální hodnotu  $d$ , nelze tuto hodnotu pomocí mapovací funkce zkrátit, ovšem lze dvě a více hodnot sloučit do jedné (mapovací funkce není funkcí prostou).

Pokud obraz, který je například celý černý, obsahuje tmavý šum, tak při ekvalizaci histogramu se tento šum přesune do světlé jasové složky a bude více výrazný. Aby se tomu zamezilo, má smysl provádět ekvalizaci pouze nad částí obrazu, nazývané jako lokální ekvalizace. To se využívá například při úprav fotografií vesmírných objektů.

Při ekvalizaci histogramu v barevném prostoru lze provést ekvalizaci pro každý barevný kanál zvlášť. Potom ovšem může dojít způsobit barevné změny obrázku jako celku (například bylo-li v obraze málo nízkých intenzit modrého kanálu, tak při ekvalizaci se část těchto hodnot přesunula do vysokých intenzit). Proto se provádí ekvalizace histogramu v jiném barevném prostoru, využívá se model YIQ, který používá celkový jas dané barvy. [15, str. 29]

Typickým použitím ekvalizace histogramu je změna jasových poměrů ve fotografiích pořízených v šeru či proti světlu. Přesunutím některých jasových složek do druhé části histogramu totiž zlepši kontrast. Dalším možným použitím je upravení jasových poměrů v matici obsahující korelační koeficienty.



Obrázek 3.1: Obrázek 7.2 před a po ekvalizaci histogramu a odpovídající histogramy

### 3.4 Geometrické transformace obrazu

Transformuje-li se vstupní obraz  $A$  na výstupní obraz  $B$ , tak se při geometrické transformaci přiřazuje pixelu na diskrétní souřadnici  $[i, j]$  nějaká neceločíselná souřadnice  $[x, y]$ . V diskrétním obraze nelze s přesností říct, co se vyskytuje na pozici  $[x, y]$ , proto mohou ve výstupním obraze  $B$  vznikat oblasti s nedefinovanou hodnotou, popřípadě může na pozici pixelu  $B$  příslušet několik bodů z obrazu  $A$ .

Geometrická transformace obrazu je funkce pro dvourozměrný obraz složený z bodů  $[x, y]$ :

$$T(u, v) = [x(u, v), y(u, v)], \quad (3.10)$$

kde vstupy funkcí  $x(u, v)$  a  $y(u, v)$  jsou souřadnice  $u$  a  $v$ , a kde výstupem funkcí je nová poloha pixelu. [15, str. 136]

Způsob, jakým se přiřazují jednotlivé pixely mezi vstupním a výstupním obrazem se nazývá *mapování*. Existují dva přístupy mapování: *dopředné* (u vstupního obrazu  $A$  se prochází jednotlivé pixely a pro každý se hledá se jejich umístění na výstupním obraze  $B$ ) a *zpětné* (u výstupního obrazu  $B$  se prochází jednotlivé pixely a hledá se umístění odpovídajících pixelů na vstupním obraze  $A$ , které se mapují do procházených pixelů). Při dopředném mapování mohou vznikat ve výstupním obraze oblasti s neidentifikovanou hodnotou, protože mapování pixelů se nemusí provádět na celý pixel, ale i na jeho část.

Pro mapování lze využít *afinní transformaci*, kde stačí tři páry korespondujících bodů pro nalezení koeficientů:

$$\begin{aligned}x' &= a_0 + a_1x + a_2x, \\y' &= b_0 + b_1y + b_2y.\end{aligned}\tag{3.11}$$

Geometrická transformace aplikovaná na snímek může změnit souřadnicový systém, a proto se zavádí *jakobián*  $J$ , který obsahuje informace o změnách souřadnicového systému

$$J = \left| \frac{\partial(x', y')}{\partial(x, y)} \right| = \begin{vmatrix} \partial x' / \partial x & \partial x' / \partial y \\ \partial y' / \partial x & \partial y' / \partial y \end{vmatrix}\tag{3.12}$$

Pro afinní transformaci se jakobián počítá jako

$$J = a_1b_2 - a_2b_1.\tag{3.13}$$

Afinní transformace se skládá z rotace, translace (posunutí), změny měřítka a zkosení. [17, str. 120]

- **Rotace** (otočení) o úhel  $\alpha$  kolem středu

$$\begin{aligned}x' &= x \cos \alpha + y \sin \alpha, \\y' &= -x \sin \alpha + y \cos \alpha, \\J &= 1.\end{aligned}\tag{3.14}$$

- **Změna měřítka** pro  $a$  na ose  $x$  a  $b$  na ose  $y$

$$\begin{aligned}x' &= a x, \\y' &= b y, \\J &= a b.\end{aligned}\tag{3.15}$$

- **Zkosení** o úhel  $\alpha$

$$\begin{aligned}x' &= x + y \tan \alpha, \\y' &= y, \\J &= 1.\end{aligned}\tag{3.16}$$

Při práci s digitálním obrazem nastává problém při zvětšení obrazu. Při dvojnásobném zvětšení vstupního obrazu bude ve výstupním obrazu každý druhý řádek a sloupec obsahovat neidentifikované hodnoty. Aby se tomu zamezilo, využívá se nalezení spojitě aproximace diskrétního obrazu: *rekonstrukce*. Poté se provede transformace spojitěho obrazu a následuje nové vzorkování. Tento postup se nazývá *převzorkování*. [15, str. 137]

Při rekonstrukci (přechodu od diskrétní funkce  $I_i; i = 1, 2, \dots$  ke spojitě funkci  $f(x); x \in R$ ) je zásadním problémem to, že mezi vzorky se vyskytuje nedefinovaná hodnota. Tuto hodnotu je potřeba aproximovat. Pro rekonstrukci se používá konvoluce, přičemž volba konvolučního jádra  $H(x)$  je klíčová pro výslednou kvalitu či rychlost převzorkování. Omezující podmínkou pro konvoluční jádro je vymezení plochy:

$$\int_{-\infty}^{+\infty} h(t)dt = 1. \quad (3.17)$$

Pro rekonstrukci se využívá interpolace nejbližším sousedem, lineární (bilineární) interpolace, interpolace plochami vyššího stupně či Lanczosovo jádro:

- **Interpolace nejbližším sousedem:** anglicky *nearest neighbour interpolation* nebo *point shift*, je okopírování hodnoty nejbližšího souseda v okolí vzorku.

Výpočet spojitě funkce  $f(x)$  z diskrétní funkce  $I_i; i = 1, \dots, n$  má tvar

$$f(x) = I_i; i - \frac{1}{2} < x \leq i + \frac{1}{2}, \quad (3.18)$$

konvoluční jádro má tvar

$$h(x) = \begin{cases} 1 & \text{pro } 0 \leq |x| < \frac{1}{2} \\ 0 & \text{pro } \frac{1}{2} \leq |x| \end{cases} \quad (3.19)$$

Tato interpolace se implementuje jako zaokrouhlení reálného čísla. Výstupní obraz je velmi nekvalitní, protože při zvětšení obrazu se zvýrazňují skoky u hran s malým skokem, a při zmenšení obrazu mohou zmizet tenké hrany. Využívá se při tvorbě náhledů obrazu, protože výpočet je rychlý a paměťově nenáročný.

- **Lineární (bilineární) interpolace:** anglicky *linear interpolation*, je aproximační metoda, využívající pro výpočet nové hodnoty blízké dva body.

Vezmou-li se dva sousední vzorky  $i$  a  $i + 1$  s hodnotami  $I_i$  a  $I_{i+1}$ , přičemž se zjišťuje hodnota  $f(x)$  v bodě  $x, i < x \leq i + 1$ , a proloží-li se krajní body úsečkou, tak hledaná hodnota v bodě  $x$  se vypočítá jako

$$f(x) = I_i + \frac{x - i}{\Delta x} [I_{i+1} - I_i]. \quad (3.20)$$

Konvolučním jádrem je funkce, která má tvar jehlanového stanu (*tent function*):

$$h(x) = \begin{cases} 1 - |x| & \text{pro } 0 \leq |x| < 1 \\ 0 & \text{pro } 1 \leq |x| \end{cases} \quad (3.21)$$

Při zpracování obrazu se používá *bilineární interpolace*, kde se neberou v potaz pouze dva sousední body, ale celkem čtyři, vždy dva na každém řádku. Výhodou je rychlost výpočtu nových hodnot, ovšem rozmazává původně nerozmazané hrany a přechody. Interpolace se provádí dvakrát, jak pro řádky, tak pro sloupce.

- **interpolace plochami vyššího stupně:** nejčastěji se používá kubická interpolace (*cubic interpolation*), též nazývána jako Parzenovo okno. Pro svůj výpočet vyžaduje



více (často 4x4) vzorků ze svého okolí. Častou metodou použití je aproximace pomocí kubické B-spline křivky. Konvoluční jádro má tvar

$$h(x) = \frac{1}{6} \begin{cases} 3|x|^3 - 6|x|^2 + 4 & \text{pro } 0 \leq |x| < 1 \\ -|x|^3 + 6|x|^2 - 12|x| + 8 & \text{pro } 1 \leq |x| < 2 \\ 0 & \text{pro } 2 \leq |x| \end{cases} \quad (3.22)$$

Při zpracování obrazu se používá konvoluční jádro jak pro řádky, tak pro sloupce. Interpolace se poté nazývá *bikubická*. Tím, že se při výpočtu používá větší okolí, je převzorkování při zvětšení obrazu kvalitnější, ovšem oproti bilineární interpolaci je několikanásobně pomalejší.

- **Lanczosovo jádro:** anglicky *Lanczos kernel*, způsob rekonstrukce, využívající sinc filtr (tzv. Fourierův vzor):

$$\text{sinc}(x) = \frac{\sin(\pi x)}{\pi x}. \quad (3.23)$$

Pro výpočet nové hodnoty se používá rovnice

$$S(x) = \sum_{i=\lfloor x \rfloor - a + 1}^{\lfloor x \rfloor + a} s_i h(x - i), \quad (3.24)$$

kde  $a$  je velikost filtru,  $s_i$  jsou vzorky signálu na pozicích  $i$ ,  $x$  je hodnota na reálné ose a  $\lfloor x \rfloor$  je dolní celá část čísla  $x$ . Konvoluční jádro  $h(x)$  má tvar

$$h(x) = \begin{cases} \text{sinc}(x) \text{sinc}\left(\frac{x}{a}\right) & \text{pro } -a < x < a \\ 0 & \text{jinde} \end{cases} \quad (3.25)$$

respektive se využívá normalizovaná sinc funkce výše uvedeného konvolučního jádra:

$$h(x) = \begin{cases} 1 & \text{pro } x = 0 \\ \frac{a \sin(\pi x) \sin\left(\frac{\pi x}{a}\right)}{\pi^2 x^2} & \text{pro } -a \leq x < a ; x \neq 0 \\ 0 & \text{jinde} \end{cases} \quad (3.26)$$

Hodnota  $a$  je celočíselná hodnota, která značí velikost jádra, většinou je to 2 nebo 3. Při zpracování obrazu, tedy v 2D, je předpis pro konvoluční jádro následující:

$$h(x, y) = h(\sqrt{x^2 + y^2}) \quad (3.27)$$

Převzorkování pomocí Lanczosova jádra při zvětšování obrazu generuje menší počet artefaktů a zachovává ostrost hran. Čím vyšší je velikost jádra  $a$ , tím více se ořeže vnější okraj obrazu. Nevýhodou je, že se mohou tvořit v obraze tzv. *ringing artifacts*, které se projevují u výrazných hran světlými a tmavými světelnými kruhy (anglicky *halos*). Oproti bikubické interpolaci je při zvětšení obrazu výsledek kvalitnější, ovšem samotný výpočet je výpočetně náročný, jelikož se zahrnuje do výpočtů poměrně velké okolí. V praxi se spíše používá zvětšení s bikubickou interpolací, protože představuje kompromis mezi rychlostí výpočtu nových vzorků při zachování vysoké kvality obrazu. [7]

### 3.5 Vyhledání významných bodů v obraze

Při zpracování obrazu se občas hodí vyhledat v obraze výskyt nějakých zajímavých objektů, ať už to jsou hrany, významné body či regiony. Pro provádění geometrických transformací se může s výhodou použít detektor významných bodů a hledat podobné body u dvou podobných snímků. Detektory bodů se též dají použít i při zpracování videa k detekci pohybu objektů. [17, str. 156]

Významnými body mohou být rohy nějakého objektu v obraze. Ty lze najít pomocí lokálních detektorů, kde vstupem je šedotónový obraz a výstupem je obraz, kde jednotlivé body v obraze reprezentují hodnotu pravděpodobnosti, že se jedná o roh. Významné body se z tohoto vygenerovaného snímku získávají prahováním.



Obrázek 3.2: Nejednoznačnost detektoru hran v rohu objektu

Detektory hran jsou nestabilní v rozích objektu, protože gradient v rohu je nejednoznačný, což je znázorněno na obrázku 3.2. Této vlastnosti se využívá právě u detektorů rohů.

Nejjednodušším detektorem rohů je **Moravcův detektor**, který je maximální v pixelech, kde je vysoký kontrast. Tyto body jsou v rozích a ostrých hranách objektu. Moravcův operátor MO se dá zjistiť jako

$$\text{MO}(i, j) = \frac{1}{8} \sum_{k=i-1}^{i+1} \sum_{l=j-1}^{j+1} |f(k, l) - f(i, j)|. \quad (3.28)$$

**Harrisův detektor** je rozšířením Moravcova detektoru rohů vzhledem k rozdílům ohodnocení rohu (což je součet rozdílů čtverců). Mějme šedotónový obraz  $f$ , ve kterém se vyskytuje vzorek obrazu  $W \in f$ , posunutý o  $\Delta x$  a  $\Delta y$ . Suma rozdílů čtverců  $S$  je dána jako

$$S_W(\Delta x, \Delta y) = \sum_{x_i \in W} \sum_{y_i \in W} (f(x_i, y_i) - f(x_i - \Delta x, y_i - \Delta y))^2. \quad (3.29)$$

Roh, který je výrazný, má vysoký výsledek  $S_W(\Delta x, \Delta y)$  pro všechny  $\Delta x$  a  $\Delta y$ . Je-li posunutý vzorek obrazu aproximován pomocí Taylorova rozvoje prvního řádu

$$f(x_i - \Delta x, y_i - \Delta y) \approx f(x_i, y_i) + \left[ \frac{\partial f(x_i, y_i)}{\partial x}, \frac{\partial f(x_i, y_i)}{\partial y} \right] \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix}, \quad (3.30)$$

tak minimální  $S_W$  lze získat analyticky. Dosazením rovnice 3.29 do rovnice 3.30 vznikne

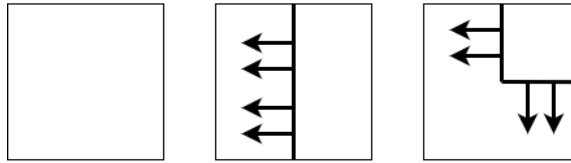
$$\begin{aligned} S(x, y) &= \sum_{x_i \in W} \sum_{y_i \in W} \left( \left[ \frac{\partial f(x_i, y_i)}{\partial x}, \frac{\partial f(x_i, y_i)}{\partial y} \right] \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \right)^2 \\ &= [\Delta x, \Delta y] A_W(x, y) \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix}, \end{aligned} \quad (3.31)$$

kde  $A_W(x, y)$  je Harrisova matice, což je druhá derivace  $S$  okolo bodu  $(x, y) = (0, 0)$ .  $A$  je

$$A(x, y) = \begin{bmatrix} \sum_{x_i \in W} \sum_{y_i \in W} \frac{\partial^2 f(x_i, y_i)}{\partial x^2} & \sum_{x_i \in W} \sum_{y_i \in W} \frac{\partial f(x_i, y_i)}{\partial x} \frac{\partial f(x_i, y_i)}{\partial y} \\ \sum_{x_i \in W} \sum_{y_i \in W} \frac{\partial f(x_i, y_i)}{\partial x} \frac{\partial f(x_i, y_i)}{\partial y} & \sum_{x_i \in W} \sum_{y_i \in W} \frac{\partial^2 f(x_i, y_i)}{\partial y^2} \end{bmatrix}. \quad (3.32)$$

Matice  $A$  je pozitivně definitní a symetrická a reprezentuje okolí. Parciální derivace v ortogonálních směrech jsou vlastní čísla  $\lambda_1, \lambda_2$ , které reprezentují roh. Mohou nastat tři situace:

1. Obě vlastní čísla jsou malá. To značí, že obraz  $f$  je rovný ve vypočteném pixelu a nevyskytuje se zde roh. Na obrázku 3.3 znázorněno úplně vlevo.
2. Jedno vlastní číslo je malé a druhé je velké. Okolí vypočteného pixelu je hřebenovitého tvaru a vyskytuje se zde hrana. Na obrázku 3.3 znázorněno uprostřed.
3. Obě vlastní čísla jsou velká. Znamená to, že byl nalezen roh. Na obrázku 3.3 znázorněno úplně vpravo.



Obrázek 3.3: Ilustrace rozhodování Harrisova detektoru rohů

Výpočet vlastních čísel je přesný, ale výpočetně náročný, a proto stačí vypočítat pouze *response function*

$$R(A) = \det(A) - \kappa \text{trace}^2(A), \quad (3.33)$$

kde  $\det(A)$  je determinant matice  $A$ ,  $\text{trace}(A)$  je součet prvků na hlavní diagonále v matici  $A$ , a  $\kappa$  je parametr, nabývající hodnoty z intervalu  $(0.04; 0.15)$ .

Výhodou použití Harrisova detektoru rohů je necitlivost na 2D posuny a rotace, nízké změny jasu či úhlu pohledu a nízké výpočetní nároky. Naopak při vysoké změně měřítka či úhlu pohledu nebo významných změnách kontrastu se snižuje jeho spolehlivost.

Metod pro detekci rohů existuje široká škála, ať už celkově jiných postupů, než používá Harrisův detektor (například Wang and Brady corner detection) nebo úpravy Harrisova detektoru (například Shi–Tomasi corner detector). Záleží na požadované rychlosti, přesnosti či náročnosti detekce rohů. [4]

## Kapitola 4

# Stereo kamera

Stereo kamera je zařízení, které je složeno ze dvou snímacích zařízení, kde je synchronizována závěrka, clona a ostření. Dvě kamery jsou umístěny na společné základně: bázi, kde jsou jednotlivé kamery horizontálně vedle sebe. [8]

Stereo kamerou se dá získat více informací o snímané scéně, než použitím jedné kamery, jelikož dvě kamery simulují lidské oči (binokulární vidění). Tím lze získat dva různé pohledy na jednu scénu, přičemž vzájemným porovnáním pořízených snímků ze stereo kamery lze získat informaci o relativní hloubce formou mapy disparity, což je rozdíl mezi pozicemi stejných bodů mezi dvěma snímky jedné scény. Hodnoty v mapě disparity odpovídají hloubce scény jednotlivých korespondujících bodů mezi snímky.

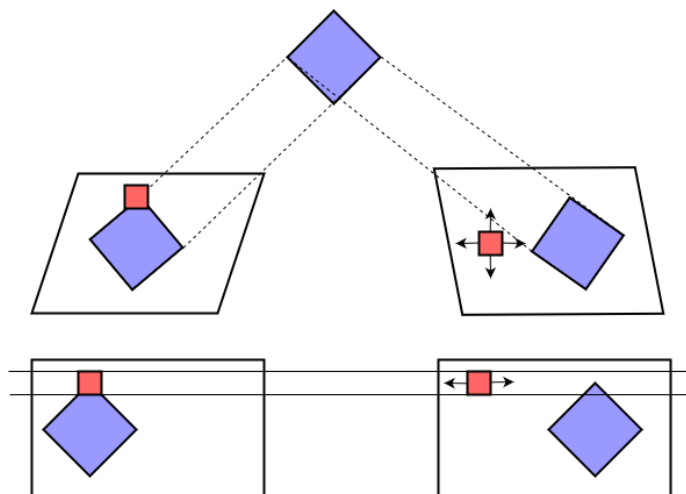
### 4.1 Stereo vidění

Stereo viděním se rozumí získání 3D informace ze snímku, pořízeného snímačem. [17, str. 573] Pokud správně zkalibrované jednotlivé snímače u stereo kamery pozorují jeden bod  $X$  v jedné scéně, tak jeho 3D souřadnice mohou být vypočítané jako průsečík dvou paprsků, což je popsáno *epipolární geometrií*. Základní princip stereo vidění sestává ze tří kroků:

- kalibrace kamery,
- odhad korespondujících bodů mezi páry bodů z levé a pravé kamery (rektifikace) a
- rekonstrukce 3D souřadnic bodů ve scéně.

Kalibrací kamery se rozumí nastavení jednotlivých kamer do výchozí polohy. Kalibrovaná kamera by měla vykazovat stejné hodnoty jako při referenční konfiguraci.

Rektifikací obrazu se rozumí provedení takové transformace obrazu, kde se promítne dva a více snímků do jedné, společné roviny. Nejsou-li pořízené snímky z jedné roviny, tak je nutné vyhledávat korespondence ve dvou směrech: horizontálním a vertikálním. U stereo kamery, která má kamery umístěné na bázi, přičemž kamery snímají stejnou rovinu, lze hledání odpovídajících korespondencí zjednodušit na vyhledání v jednom prostoru, konkrétně po horizontální ose mezi oběma kamerami. Je-li souřadnice bodu z levého snímku známá, lze s jednoduchostí hledat odpovídající bod i ve druhém snímku na stejném řádku. Nejsou-li ovšem kamery perfektně zarovnané, musí se provést rektifikace, aby šlo využít vyhledání korespondencí v horizontální ose. To je znázorněno na obrázku 4.1. [6]

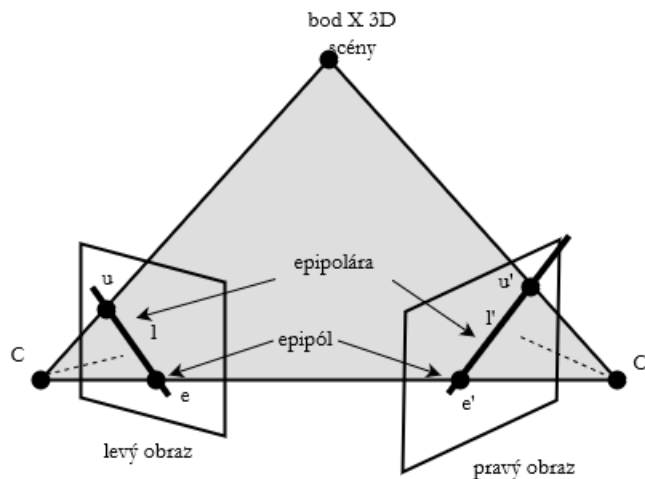


Obrázek 4.1: Prohledávaný prostor před a po rektifikaci

Pro rekonstrukci 3D souřadnic je nutné mít vypočítanou předem mapu disparit. Tato mapa disparit se promítne do 3D point cloud, což je množina modů v 3D prostoru, kde jednotlivé 3D body jde vypočítat, jsou-li známy projekční parametry kamery. [3]

## 4.2 Epipolární geometrie

Geometrie, která je využita u stereo kamery je popsána na obrázku 4.2.



Obrázek 4.2: Geometrie dvou kamer

Přímka spojující středy jednotlivých kamer  $C$  a  $C'$  se nazývá základna. Ta protíná obrazové roviny v epipólech  $e$  a  $e'$ . Epipól je obraz pohledu jedné kamery na druhou kameru,  $e = MC'$  a  $e' = M'C$ . [17, str. 574]

Jakýkoliv bod ve scéně, označený jako  $X$ , který je sledován oběma kamerami, a dva odpovídající paprsky, které vychází z jednotlivých kamer  $C$  a  $C'$ , společně tvoří epipolární rovinu. Tato rovina protíná roviny snímků v epipolárách  $l$  a  $l'$ . Lze též říci, že epipolára je projekce paprsku z jedné kamery do druhé. Všechny epipoláry se protnou v epipólech.

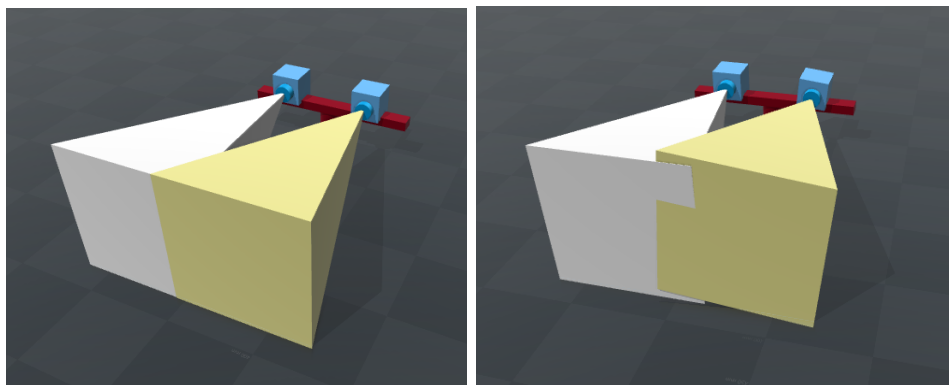
Nechť body  $u$  a  $u'$  jsou projekcí bodu  $X$  pro jednu, respektive druhou kameru. Poté paprsek  $CX$  bude reprezentovat všechny možné pozice bodu  $X$  pro první obraz a bude viděn jako epipolára  $l'$  ve druhém obraze. Bod  $u'$  z druhého obrazu, který odpovídá bodu  $u$  z prvního obrazu, musí ležet na epipoláře  $l'$  ve druhém snímku,  $l'^{\top}u' = 0$ . Tato situace je symetrická i pro druhý obraz, tedy  $l^{\top}u = 0$ . Faktem je, že pozice dvou korespondujících bodů není libovolná, což je známé jako epipolární omezení.

## Kapitola 5

# Zhodnocení stavu a specifikace

### 5.1 Popis problému

Při snímání scény pomocí stereo kamery je důležité, aby jednotlivé snímací aparáty (kamery) byly v pevné poloze. Bude-li například jedna kamera oproti druhé kameře jinak rotovaná, nebude docházet ke snímání jedné scény pouze s předem známou translací, ale přidá se k tomu i rotace. Taktéž může dojít i k jinému perspektivnímu zobrazení dané scény, budou-li jednotlivé kamery špatně seřízeny. Pokud bude například špatně seřízená stereo kamera snímat vozovku, může produkovat špatné výsledky jednotlivých veličin, které se díky ní měří (například rychlost a vzdálenost vozidel či rozpoznání státních poznávacích značek). K tomuto jevu může dojít jak nevhodnou manipulací (při převozu stereo kamery mezi dvěma body může dojít vlivem pohybu a nárazů jednotlivých kamer k vychýlení), tak vlivem přírodních jevů (silný vítr, teplotní roztažnost kovů). Příklad takového natočení je na obrázku 5.1.



Obrázek 5.1: Ilustrace změny natočení části stereokamery

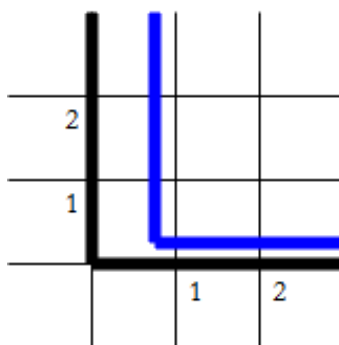
Z tohoto důvodu je dobré zajistit postup, který odhalí, jestli se jednotlivé kamery vůči počáteční konfiguraci neposunuly, a pokud tento posun nastal, tak je dobré zjistit o jakou vzdálenost a kterým směrem vůči svému počátku. Aby se kamera nemusela ručně přenastavovat, měl by se tento posun obrazu kompenzovat.

## 5.2 Návrh řešení

Seřízení jedné kamery ze stereo páru by mohlo probíhat tak, že se vytvoří v době jejího umístění snímek, který se stane výchozím snímkem pro kalibraci. Tento snímek se bude označovat jako *referenční*. Na kopii toho snímku by se vyznačilo několik oblastí, podle kterých se bude kamera dát kalibrovat. I když jde samozřejmě vybrat celý snímek, je to nevhodná strategie, zejména by při operování s celým snímkem docházelo k manipulaci s velkými objemy dat, které by trvalo dlouho zpracovat. Dalším problémem by mohlo být, že s postupem času se mohou objekty na scéně měnit. Jako příklady lze uvést růst rostlin, padání listů, pokrytí plochy sněhem či pohyb způsobený větrem. Proto je vhodné zahrnout do výpočtů pouze části snímků, které se s velkou pravděpodobností příliš nebudou měnit v čase.

Tyto oblasti budou sloužit k hledání podobné oblasti ve snímku, pořízeném někdy v jiné době, kde se dá předpokládat změna proti původnímu stavu. Hledání by mohlo probíhat na základě hledání podobnosti (korelace) této oblasti ve snímku pořízeném v budoucnosti. Protože hledání v celém snímku by bylo paměťově a časově náročné, je lepší vzít pozice vyznačené oblasti v referenčním snímku a na stejných pozicích, zvětšené o blízké okolí, hledat výskyt této oblasti ve snímku pořízeném v budoucnosti.

Jakmile se tyto oblasti detekují, lze zjistit, o kolik se posunuly vůči původnímu stavu, ale jelikož translace není vždy o celé jednotky pixelů, ale může nastat i o desetiny pixelů, je důležité brát tento aspekt v ohled. Jednou z možných variant, jak zjistit posun o desetiny pixelu, je obraz zvětšit (převzorkovat) a v takto zvětšeném obraze lze znovu hledat oblast zájmu, která by byla též zvětšená.



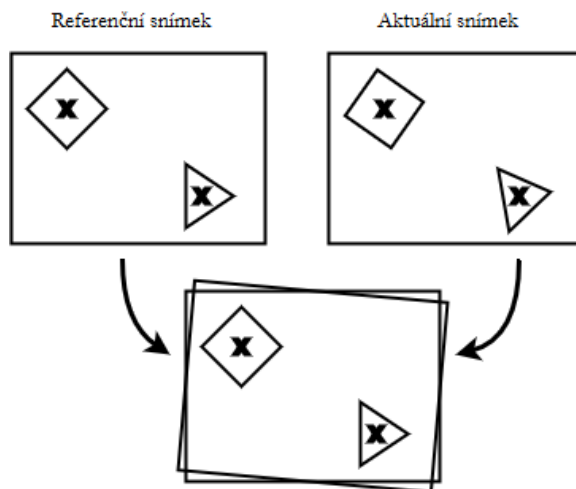
Obrázek 5.2: Ilustrace posunu o méně než pixel

Protože nelze nikdy s jistotou říci, jestli je vypočítaná translace korektní, je vhodné snímek pořízený v budoucnosti promítnout na snímek referenční. Toho lze docílit buď perspektivní transformací, popřípadě afinní transformací. V případě použití perspektivní transformace by došlo k přepočítání snímku do jiné perspektivy a dá se předpokládat, že při minimální změně polohy kamery by byla změna perspektivy zanedbatelná a vypočítání matice pro transformaci by bylo zbytečně paměťově náročné a zdlouhavé. Z tohoto důvodu lze využít afinní transformaci, která perspektivu nemění a dokáže kompenzovat translaci, rotaci, změnu měřítka či zkosení.

Pokud by při se výpočtu transformační matice použilo pouze tolik bodů, kolik by bylo vybráno oblastí, bylo by to sice efektivní co se týče rychlosti, ale bohužel by byla výsledná kompenzace snímku nekvalitní a nepřesná. Z tohoto důvodu je lepší ve vyznačených oblas-



tech vyhledat výrazné hrany či rohy, pro které se budou hledat odpovídající korespondence, čímž se zlepší kvalita afinní transformace.



Obrázek 5.3: Ilustrace aktuálního snímku promítnutého na snímek referenční

Poté, co by se provedla afinní transformace, by se znovu mohly vyhledat vyznačené oblasti z referenčního snímku v částech snímku pořízeného v budoucnosti, přičemž zjištěná translace by ukazovala chybu afinní transformace. Dále by bylo vhodné zjistit, jestli má zvětšení významný vliv na kvalitu kompenzace.

Cílem tedy je vytvořit takovou proceduru, která jednou za určitý časový interval vyzkouší, jestli jednotlivá snímání zařízení ze stereo kamery snímají původní scénu, pro kterou byly kalibrovány, zjistí změny proti původnímu stavu a pokusí se kompenzovat pořízený snímek na původní, čímž se zjistí chyba.

K implementaci řešeného problému lze v podstatě využít dvě možnosti. První možností je použít programovací jazyky založené na práci s maticemi (*Matlab* či *Octave*) a druhou je využít knihovnu *OpenCV* pro běžné programovací jazyky.

- **Matlab:** programovací jazyk a prostředí, které je primárně určeno pro snadné výpočty s maticemi (název vznikl z anglického Matrix Laboratory). Je to komerční produkt, který obsahuje spoustu tzv. Toolboxů, které lze použít pro řešení určitých problémů (například Image Processing Toolbox pro práci s obrazy, Neural Network Toolbox pro práci s neuronovými sítěmi, Partial Differential Equation Toolbox pro práci s parciálními diferenciálními rovnicemi atd.), ovšem většina těchto toolboxů je zpoplatněna. [1]
- **Octave:** programovací jazyk a prostředí, které vychází z Matlabu. Umožňuje práci s maticemi se syntaxí podobnou Matlabu, ovšem ne vše je v Octave implementováno. Při náročnějších výpočtech je aplikace nestabilní a dokáže pracovat jen s omezenou velikostí dočasné paměti. Octave je open source. [16]
- **OpenCV:** knihovna pro práci s počítačovým viděním, která přináší stěžejní datový typ: matici a mnoho operací nad ní. Lze ji použít v mnoha jazycích, jako například v C, C++, Python či Java. Knihovna je open source a spolu s kombinací jazyka C++ lze minimalizovat paměťovou náročnost při výpočtech. [2]

Pro implementaci problému jsem se rozhodl použít jazyk C++ a knihovnu OpenCV verze 3.0. Důvodem je možnost správy paměti a snadná modularita na jiné knihovny. Testovací sady snímků s uměle vytvořenými posuny jsem si pořídil sám.

Vytvořená aplikace (znázorněno na obrázku 7.1) prochází jednotlivé sady snímků a v každé sadě snímků vypočítá translaci zvýrazněných oblastí z jednotlivých snímků z téže sady vůči referenčnímu snímku. Výpočet se opakuje pro různá zvětšení barevně zvýrazněných oblastí, přičemž při převzorkování obrazu lze využít různé způsoby rekonstrukce. Vypočítanou translaci při různém zvětšení oblasti poté zaznamenává do grafů, které jsou uloženy jako obrázky formátu PNG.

Aplikace dále v jednotlivých oblastech u referenčního snímku vypočítá korelační koeficient a detekuje významné body, přičemž nalezne odpovídající korespondence i v jiném snímku ze sady. Tyto informace poté použije k výpočtu afinní transformace, kterou provede na odpovídající snímek ze sady. Následuje vypočítání korelačních koeficientů a nalezení odpovídající korespondence vůči referenčnímu snímku, čímž se zjistí chyba afinní transformace.

Pro zvýrazněné oblasti se při různém zvětšení obrazu vygenerují obrázky, které se uloží do souboru jak u referenčního snímku, tak u dalšího snímku ze sady, a to jak před afinní transformací, tak i po transformaci.

## Kapitola 6

# Tvorba testovacích dat

Pro určení korelace mezi dvěma obrazy bylo nutné nejprve vytvořit sadu scén, ve kterých je několik desítek fotografií stejné scény. Následně se vybral referenční snímek, na kterém se provedlo zvýraznění zájmových oblastí, podle kterých se dále prováděla korelace.

### 6.1 Snímání obrazu

Snímání obrazu probíhalo pomocí digitálního fotoaparátu Sony  $\alpha$  5000, kde výsledné snímky byly uloženy ve formátu JPEG s rozlišením  $5456 \times 3632$  pixelů. Zároveň s tímto snímkem byl vytvořen i další snímek ve formátu RAW. [10]

Při výběru scény ji bylo nutné volit tak, aby se v zachyceném obraze vyskytovaly alespoň tři pevné oblasti, které se vlivem klimatických změn či zásahů člověka příliš neměnily v čase. Vhodnými kandidáty (na obrázku 6.1) na takové oblasti se jevíly:

- členité stěny s výraznými prvky na fasádě bez flóry,
- neomítnuté (nebo částečně omítnuté) stěny postavené z cihel,
- sloupy elektrického vedení bez pohyblivých částí (drátů),
- budovy a objekty na pozadí snímku (rozhledny, kostelní věže, panelová zástavba),
- svodidla podél komunikací, spodní části veřejného osvětlení, stojany nouzového volání podél dálnic a
- výrazné, avšak matné okna a dveře.

Špatnými kandidáty na objekty se stala flóra, která se v čase mění (padá listí ze stromů, pohybuje se větrem, vlivem člověka zásadně mění tvar), dopravní značení a celé pouliční lampy (pohyb při větru, vyšší pravděpodobnost změny tvaru při nárazu vozidla do objektu) či silnice a střechy budov (pokrytí sněhem v zimním období, delší doba vypařování vody po dešti). Při hledání vhodné scény bylo nutné zohlednit fakt, že tyto objekty by ideálně měly být náhodně rozprostřeny po celém obraze.

Při nalezení vhodné scény ke snímání bylo dalším důležitým kritériem, aby se na snímku vyskytoval zaparkovaný osobní, či dodávkový automobil s jasně čitelnou státní poznávací značkou, které dohromady budou sloužit k výzkumným účelům v budoucnosti. Vhodné bylo snímat scénu při zatažené obloze z důvodu menších odlesků slunečních paprsků na lesklých předmětech.



Obrázek 6.1: Referenční snímek a kandidátní oblasti pro anotace

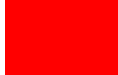





Při snímání scény bylo snímací zařízení umístěno na stativu, aby nedocházelo k rozmazávání výsledného obrazu, a aby nedocházelo ke změně výchozí polohy fotoaparátu při tvorbě sekvenci snímků. Dotažením všech šroubů na stativu bylo zamezeno vertikálnímu posunu, horizontálnímu posunu bylo zamezeno tak, že se stativ nikam v prostoru neposunul. Při fotografování této scény bylo pořízeno několik obrazů s velmi mírným posunem, způsobeným umělým rotačním pohybem fotoaparátu, ukotveným na stativu. Posun byl v rámci jednotek pixelů na výsledném snímku vzhledem k prvnímu vyfotografovanému snímku.

První snímání scény probíhalo v obci Tvrdonice, okres Břeclav. Druhé snímání proběhlo ve městě Brně, převážně v městské části Královo pole. Každá scéna tvoří sadu snímků, ve které je alespoň deset fotografií, kde první snímek byl prohlášen za referenční a ostatní jsou vůči němu mírně posunuty. Veškeré snímky z obce Tvrdonice jsou vytvořeny z výšky od země přibližně 50 cm, snímky z města Brna jsou tvořeny z mostů, nadjezdů a teras tak, aby stojící automobil byl snímán z výšky alespoň 3 metry oproti pozemní komunikaci, na které byl automobil umístěn. Tímto se simuluje jedno z možných využití stereo kamery v dopravním průmyslu jako snímač vozovky.

## 6.2 Vyznačení oblastí

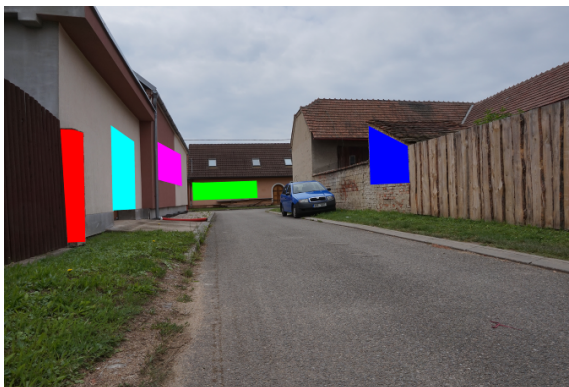
V každé sadě snímků byl první snímek vybrán jako referenční (obrázek 6.1), který byl podroben analýze, kde byly vybrány minimálně tři pevné oblasti a maximálně šest pevných oblastí, které se na snímku jeví jako nejlepší. I když existují automatizované metody pro detekci bodů v obraze, mohlo by se stát, že by algoritmus detekoval špatné body, jako například větve stromu, které se v průběhu roku výrazně mění, dále body blízko okrajům obrazu, které mohou být při významnějším posunu mimo obraz či hrany stínů, které se v průběhu dne pohybují. Dalším důvodem pro manuální anotaci oblastí v obraze je fakt, že se dle lidského úsudku vyberou ideální objekty v obraze, které se z dlouhodobého hlediska s největší pravděpodobností nebudou měnit, a které nebudou do výpočtů zbytečně zahrnovat chybu.

Tyto oblasti jsou jasně označeny barvami:

Barva	Název	RGB hodnota	Barva	Název	RGB hodnota
	Červená	(255, 0, 0)		Žlutá	(255, 255, 0)
	Zelená	(0, 255, 0)		Purpurová	(255, 0, 255)
	Modrá	(0, 0, 255)		Azurová	(0, 255, 255)

Tabulka 6.1: Seznam použitých barev pro zvýraznění oblastí

Oblasti jsou vyznačeny souvisle na bitové kopii referenčního snímku, které slouží jako masky. Vytypovaný anotovaný objekt je označen nepravidelným polygonem. Jednotlivé oblasti, které se mezi sebou liší použitou barevnou výplní, se nedotýkají a mají mezi sebou alespoň deset pixelů mezeru. Velikosti jednotlivých označených oblastí jsou různé. Příklad takto upraveného obrazu je [6.2](#).



Obrázek 6.2: Referenční snímek doplněný o zvýrazněné oblasti

K vytvoření takto anotovaného snímku byl použit běžný rastrový editor GIMP (GNU Image Manipulation Program), který umožňuje tvořit vrstvy obrazu. Podkladovou vrstvou se stala kopie referenčního snímku. Každá oblast byla zaznačena do vlastní vrstvy, především kvůli pozdější editaci tvaru a pozice oblastí. Vytvořený výstupní snímek je formátu PNG (Portable Network Graphics), který využívá bezztrátovou kompresi rastrové grafiky.

### 6.3 Popis vytvořených sad snímků

Celkem touto metodou bylo vybráno a zaznamenáno 27 scén, z toho 16 v obci Tvrdonice, okres Břeclav, a 11 v Brně, především v Králově poli. Ke všem sadám byl vybrán referenční snímek a byl vytvořen i soubor, obsahující anotované oblasti.

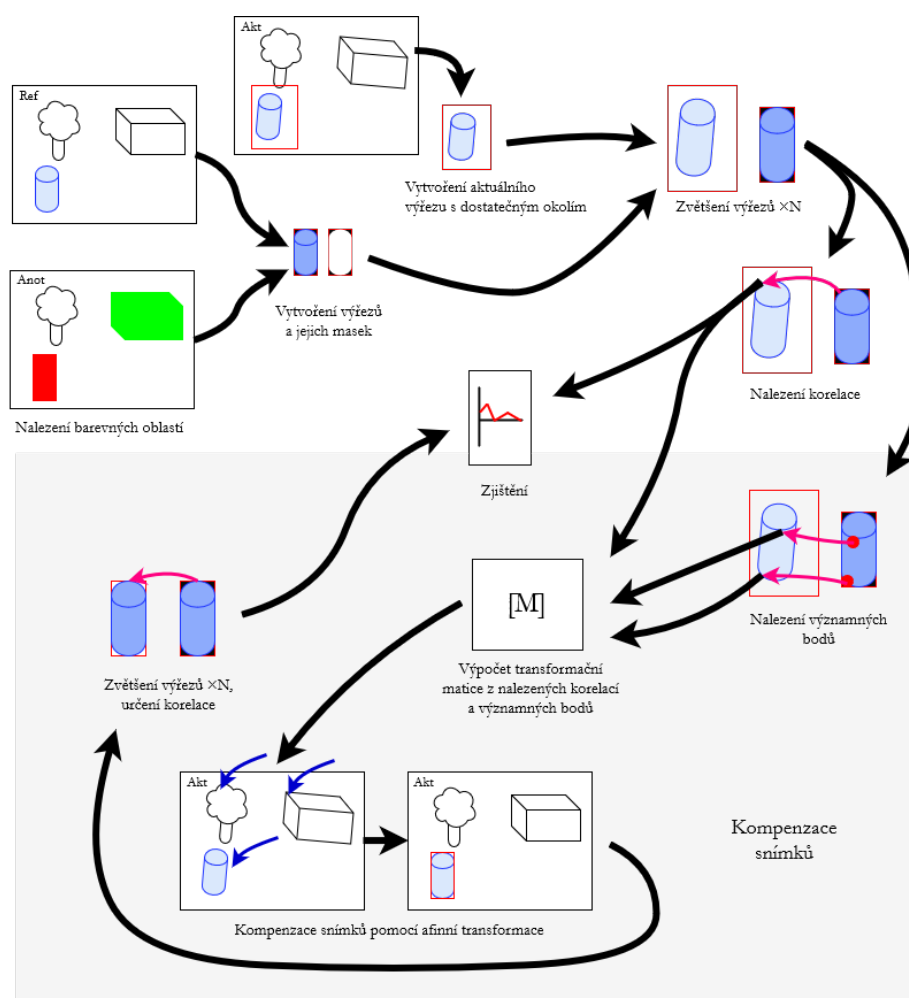
Název	Počet oblastí	Počet fotografií	Poznámka
Tvrdonice (1)	4	48	Snímek je příliš tmavý, označeny neomítnuté stěny
Tvrdonice (2)	5	11	Označeny neomítnuté stěny a dekorace nad okny
Tvrdonice (3)	6	33	Označeny doplňky na domech a elektrické přípojky do domů
Tvrdonice (4)	6	18	Označeny neomítnuté stěny a elektrické přípojky do domů
Tvrdonice (5)	6	25	Označeny dekorace nad okny
Tvrdonice (6)	5	26	Označeny neomítnuté stěny
Tvrdonice (7)	4	35	Označeny neomítnuté stěny a elektrické přípojky do domů
Tvrdonice (8)	5	58	Označeny dekorace u oken
Tvrdonice (9)	6	57	Označeny boky komínů
Tvrdonice (10)	3	43	Označeny členité ploty a části fasády
Tvrdonice (11)	4	43	Označeny neomítnuté stěny
Tvrdonice (12)	4	29	Označeny dekorace na domech
Tvrdonice (13)	4	34	Označen členitý plot a dekorace na domech
Tvrdonice (14)	4	32	Označeny doplňky na domech
Tvrdonice (15)	5	29	Označeny dekorace na fasádě a elektrické přípojky do domů
Tvrdonice (16)	4	32	Označen spodek kmenu stromu, dekorace na fasádě domů
Brno (1)	4	46	Označena členitá stěna
Brno (2)	4	46	Označeny členité plochy na domech a domy v pozadí obrazu
Brno (3)	3	48	Označeny dekorace na domech
Brno (4)	4	48	Označeny dekorace na domech
Brno (5)	4	35	Označeny dekorace na domech
Brno (6)	3	43	Označeny elektrické přípojky do domů a dekorace nad okny
Brno (7)	3	39	Označena členitá stěna a rohy stěn u oken
Brno (8)	4	42	Označeny dekorace na domech
Brno (9)	4	50	Označeny okna a dekorace nad okny
Brno (10)	4	49	Označeny dekorace nad okny
Brno (11)	3	45	Označeny okna a dekorace nad okny

Tabulka 6.2: Seznam vytvořených sad snímků a jejich informace

Tento dataset se nepoužívá pouze pro řešení tohoto problému. Bude se využívat i pro zpracování obrazu pomocí HDR, detekci vozidel či rozpoznávání státních poznávacích značek na vozidlech.

## Kapitola 7

# Určení vzájemné korelace obrazů a jejich kompenzace



Obrázek 7.1: Diagram znázorňující algoritmus pro určení korelace a kompenzace dvou snímků

Pro určení vzájemné korelace, detekce posunu a kompenzaci dvou snímků je vždy jedním vstupem referenční obraz a druhým je snímek z adresáře posunutých snímků. Postup pro určení vzájemné korelace a následné kompenzace, znázorněn na obrázku 7.1, je následující:

Nejprve je načten snímek referenční, spolu se snímkem, obsahující barevně zvýrazněné oblasti. Z referenčního snímku se získá výřez, který je dán barevnou anotací. Poté se získá prahováním maska výřezu.

V adresáři, kde jsou další, natočené snímky (na obrázku 7.1 znázorněn jako *act*, aktuální snímek), se postupně jednotlivé snímky projdou a vytvoří se výřezy s okolím vyznačených objektů (neví se přesná poloha, proto se přidává okolí). Získané výřezy a masky se poté zvětší na  $N$  násobek a hledá se korelace mezi výřezy.

Po určení korelace se nalezne nejvyšší korelační koeficient a vypočítá se translace výřezů. Vypočtené hodnoty se uloží jak v textové podobě, tak ve formě grafu, vždy pro aktuální snímek pro jednotlivé oblasti při různém zvětšení.

Dále se pro provedení kompenzace naleznou ve výřezích významné body, které spolu s nejvyššími korelačními koeficienty slouží k výpočtu transformační matice. Proveďte se afinní transformace aktuálního snímku. Následuje další vytvoření výřezů, tentokrát na transformovaném obraze, přičemž se provede zvětšení výřezů, ve kterých se určí vzájemná korelace. Určí se pozice nejvyššího korelačního koeficientu a vypočítá se translace, což značí chybu afinní transformace. Ta se zanesou do grafu a spolu s kompenzovanými výřezy se uloží.

## 7.1 Načtení referenčního snímku a extrakce anotovaných oblastí z referenčního snímku

Aby se mohlo začít provádět hledání vzájemné korelace mezi snímky, je nutné nejprve snímky načíst do paměti. K této akci je k dispozici v knihovně OpenCV, která je použita k řešení problému, mechanismu pro otevření obrazu jako matice, kde jednotlivé buňky matice reprezentují barevnou informaci jednoho bodu v obraze.



Obrázek 7.2: Referenční snímek převedený na šedotónový obraz

Důležité je správně nastavit způsob načtení a následné reprezentace obrazu. Při načítání barevného snímku dochází k převodu obrazu na tříkanálový obraz reprezentovaný BGR (blue green red, česky modrým zeleným červeným) modelem. Naopak je-li snímek načten jako šedotónový, tak je původní obraz reprezentován pouze jedním kanálem, obsahující pouze odstíny šedi. Soubory, které lze zpracovat, mohou mít mnoho formátů, jako například BMP, JPG, PNG, HDR či TIF.



Extrakcí anotovaných oblastí se rozumí nalezení spojitých regionů ve snímku s označenými oblastmi, které slouží jako masky pro získání výřezů okolí označených oblastí z referenčního snímku.



Obrázek 7.3: Vytvořený výřez a jeho maska

Přes snímek obsahující barevně anotované oblasti se postupně iteruje po jednotlivých buňkách matice, reprezentující obraz, kde se jednotlivé pixely porovnávají na výskyt barvy uvedené v tabulce 6.1. Je-li taková barva v obraze zastoupena, zaznamenají se čtyři informace: první a poslední výskyt této barvy zleva a shora. Tím jsou získány čtyři body, které tvoří obdélník a těsně obepínají hledanou oblast, která může být extrahována z referenčního snímku. Toto vyhledání je provedeno pro každou barvu anotovaných oblastí.

V dalším kroku následuje vytvoření maskovaného výřezu z referenčního snímku. Tím se rozumí vytvoření takového obrazu, kde jsou vyobrazeny jednotlivé ručně anotované oblasti z referenčního snímku bez jejich okolí.

Pozice čtyř bodů, které ohraničují barevně zvýrazněnou oblast výřezu, byly zjištěny předchozím krokem. V něm se změnil ty pixely, které nebyly zahrnuty v barevné oblasti, na černé pixely. Vznikne nepravidelný šedotónový polygon, který je umístěn na černém pozadí (obrázek 7.3). Tento výřez se využívá při hledání vzájemné korelace mezi referenčním a dalším snímek.

Dále je metodou prahování vytvořena dvoubarevná maska z výřezu, kde všechny odstíny šedi kromě černé jsou změněny na bílou barvu (obrázek 7.3).

## 7.2 Vytvoření výřezů oblasti ze sady snímků

Původní testované postupy prováděly počítání korelačních koeficientů mezi výřezem a celým snímek ze sady. I když se v drtivé většině případů našla přibližná poloha vzoru (výřezu) na obraze, doba výpočtu byla řádově vyšší, stejně tak i množství alokované dočasné paměti.

Proto bylo efektivnější neprovádět výpočet korelačních koeficientů nad celým snímkem, ale pouze nad jeho částmi.



Obrázek 7.4: Aktuální výřez spolu s okolím

Kromě vytvoření výřezů z referenčního snímku následuje vytvoření dalších výřezů, tentokrát ze souboru ze sady snímků, pro snazší orientaci v textu označených jako **aktuální výřezy**.

Aktuální výřez je tvořen tak, že se vezme levá horní souřadnice referenčního výřezu, přičte se k ní šířka a výška výřezu a takto vypočtený obdélník je dále rovnoměrně zvětšen o okolí s velikostí  $N$ . Rozšíření obdélníku je z důvodu hledání odpovídající korespondence mezi výřezem referenčním a výřezem aktuálním, platí, že čím je zahrnuté okolí větší, tím větší je paměťová, ale i časová náročnost pro další výpočty. Celočíselný parametr  $N$  se doporučuje použít v jednotkách pixelů. Tvoření těchto aktuálních výřezů je provedeno tolikrát, kolik je anotovaných oblastí. Znázorněno na obrázku 7.4.

Každému referenčnímu výřezu tedy náleží černobílá maska a výřez aktuální, který má jednotlivé rozměry rozšířené o  $2N$  než odpovídající referenční výřez.

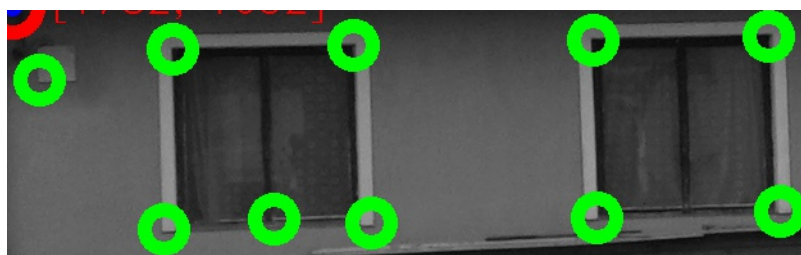
### 7.3 Vyhledání rohů v jednotlivých oblastech

Pro přesnější vypočtení transformační matice pro geometrickou transformaci digitálního obrazu nestačí pouze tolik bodů popisujících obrázky, kolik je celkem anotovaných oblastí. Z většího množství vzorků jde lépe takovou transformační matici odhadnout.

Je-li ručně označená barevná oblast dobře zvýrazněna (tedy neobsahuje problematické pohyblivé či nevýrazné objekty), lze uvnitř těchto oblastí využít automatizovaných metod pro detekci hran, vrcholů či významných bodů.

Pro detekci hran se využívá část Harrisova detektoru rohů, kdy se vypočítají vlastní čísla matice gradientů, popsané rovnicí 3.32. Tím, že se vypočítají přesné hodnoty, je dosaženo větší pravděpodobnosti správného rozhodnutí, že testovaný pixel obsahuje roh, ovšem výpočet je náročnější kvůli parciálním derivacím.

V každé anotované oblasti se vyhledá maximálně deset významných bodů, které se uloží do vektoru obsahující souřadnice rohů všech oblastí. Důvodem pro zvolení této hodnoty je kompromis, že při hledání rohů na budově, kde se pravidelně opakuje nějaký prvek, může být při vyhledávání více rohů nalezeno několik velmi podobných bodů, které mohou špatně ovlivnit kompenzaci posunutého snímku vůči referenčnímu snímku. Naopak bude-li počet hledaných bodů velmi nízký, bude snímek popsán méně body a kompenzace bude méně přesná. Vezme-li se v potaz, že je na obrázku zvýrazněno šest oblastí, tak bude obrázek před kompenzací obrazu popisovat až šedesát významných bodů.



Obrázek 7.5: Výřez, ve kterém se detekovalo deset bodů

Jsou-li detekovány rohy, které nejsou příliš výrazné, tak se neberou jako plnohodnotné rohy a neberou se v potaz při dalších výpočtech.

Jakmile jsou získány souřadnice všech významných bodů ve vyznačených oblastech, separuje se pro každý nalezený bod jeho okolí o velikosti 60x60 pixelů, přičemž uprostřed se vyskytuje nalezený významný bod. Tyto obrazy jsou poté uloženy do vektoru ve stejném pořadí, v jakém jsou uloženy souřadnice rohů ve vektoru obsahující významné body.

## 7.4 Určení vzájemné korelace mezi výřezy oblastí u dvou snímků

Aby se našla přibližná pozice vzoru (referenčního výřezu) v obraze (aktuálního výřezu), je využito metody určení vzájemné korelace dvou obrazů. Pro každý pixel v obraze se provede výpočet korelačního koeficientu, který znázorňuje podobnost mezi oběma snímky. Čím vyšší hodnota korelačního koeficientu je, tím vyšší podobnost se ve zpracovávaném bodě nachází. Následuje vyhledání maximální hodnoty korelačního koeficientu, kde jeho pozice znázorňuje souřadnice přibližného výskytu v obraze. Z této nalezené pozice lze vypočítat posunutí mezi těmito výřezy. Aby se zvýšila přesnost určení posunu mezi snímky na desetiny pixelu, jsou jednotlivé výřezy několikanásobně zvětšeny a mezi zvětšenými snímky se hledá vzájemná korelace.

Prvním krokem při hledání vzájemné korelace je několikanásobné zvětšení všech výřezů a masek pomocí převzorkování obrazu. Zvětšení je z důvodu zvýšení přesnosti hledání vzájemné korespondence, protože aktuální snímek ze sady snímků nemusí být vždy posunut o celé jednotky pixelu, ale mimo jiné i o desetiny pixelu. Zvětšením obrazů se sice zvýší přesnost určení vzájemné korelace, ovšem též dochází k několikanásobnému zvětšení objemu dat, které je nutno zpracovat a porovnat.

Pro zvětšování se dopředu vypočítají nové rozměry obrazu, na které se původní obraz převzorkuje. Při převzorkování lze použít různé metody výpočtu chybějících pixelů, přičemž jednotlivé metody se liší svou přesností, kvalitou či rychlostí.

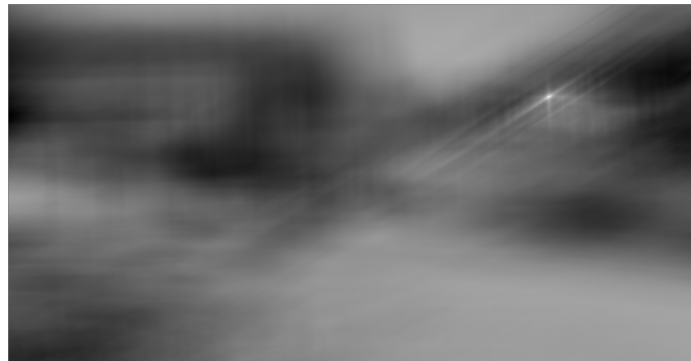
Pro určení přibližné pozice referenčního výřezu v aktuálním výřezu se používá vzájemná korelace. Obrazem se stává aktuální výřez, vytvořený z aktuálního snímku ze sady snímků, vzorem se stává výřez, vytvořený z referenčního snímku, který má svou masku, protože tvar výřezu je nepravidelný. Masky se využívá z toho důvodu, že zvýrazněné oblasti ohraničené nepravidelným polygonem mohou ve své blízkosti mít pohyblivé objekty, které mohou negativně narušovat korelaci.

Použitou metodou určení korelačního koeficientu je hledání vzájemné korelace, normalizované, která vychází z předpisu:

$$R(x, y) = \frac{\sum_{x', y'} (T(x', y') \cdot I(x + x', y + y'))}{\sqrt{\sum_{x', y'} T(x', y')^2 \cdot \sum_{x', y'} I(x + x', y + y')^2}}, \quad (7.1)$$

kde  $T$  je vzor,  $I$  je obraz, ve kterém se  $T$  hledá a  $R$  je výsledek.

Princip hledání vzoru v obrazu je takový, že se postupně prochází obraz  $I$ , ve kterém se určují korelační koeficienty vzoru  $T$ . Vypočtené hodnoty jsou poté uloženy do výsledné matice  $R$ . Koeficienty jsou normalizovány. Jakmile jsou vypočteny všechny korelační koeficienty, vyhledáním maximální hodnoty v matici  $R$  se zjistí pozice vzoru na obrazu.



Obrázek 7.6: Korelační koeficienty

Na obrázku 7.6 je zobrazena matice korelačních koeficientů pro každý pixel mezi referenčním snímkem a výřezem jako obrázek, po ekvalizaci histogramu. Světlejší části obrázku symbolizují vyšší vzájemnou korelaci, tmavší části zase nižší korelaci. V pravé části obrázku se nachází viditelný světlý bod, kde je vypočítaná korelace nejvyšší.

Hledání vzájemné korelace není provedeno pouze pro celé výřezy vytvořené z barevně anotovaných oblastí, ale mimo jiné i pro automaticky nalezené rohy a jejich blízké okolí, reprezentované čtvercovým výřezem, uloženým do vektoru, jak bylo podrobněji popsáno v 7.3. Tyto snímky se také hledají v aktuálním výřezu, kde se nalezené pozice maximálního korelačního koeficientu ukládají do dalšího vektoru ve stejném pořadí, v jakém se vyhledávali.

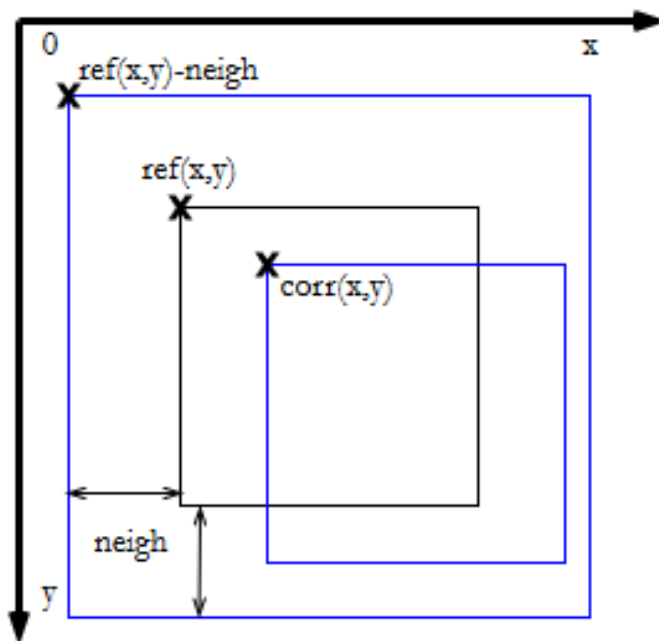
## 7.5 Kompenzace obrazů mezi referenčním a jiným snímkem ze sady snímků

Pro kompenzaci obrazů se využívají levé horní souřadnice výskytu jednotlivých barevně anotovaných oblastí s nalezenými významnými body ve vyznačených oblastech a jejich nalezené souřadnice ze vzájemné korelace se snímkem aktuálním. Ze všech získaných bodů se poté vypočítá matice afinní transformace, která se aplikuje na celý aktuální snímek a v obrazu po transformaci se poté znovu hledají odpovídající korespondence, aby se určila kvalita afinní transformace.

Po určení polohy souřadnic maximální hodnoty korelačního koeficientu mezi obrazem a vzorem dochází k přepočítání souřadnic na aktuální snímek. Musí se zohlednit fakta, že výřezy se několikanásobně zvětšily a jsou doplněny o blízké okolí:

$$\text{shift}(x, y) = \left( \text{ref}(x, y) - \left( \text{ref}(x, y) - \text{neigh} + \frac{\text{corr}(x, y)}{\text{mul}} \right) \right), \quad (7.2)$$

kde  $\text{ref}(x, y)$  je horní levá souřadnice výskytu výřezu v referenčním snímku,  $\text{neigh}$  je velikost přidaná k referenčnímu výřezu (blízké okolí),  $\text{corr}(x, y)$  je nalezená pozice vzoru v aktuálním výřezu, a  $\text{mul}$  je parametr udávající násobek zvětšení vzoru. Na obrázku 7.7 je znázorněn referenční výřez (černý čtverec) a aktuální výřez (malý modrý čtverec), doplněný o okolí (velký modrý čtverec).

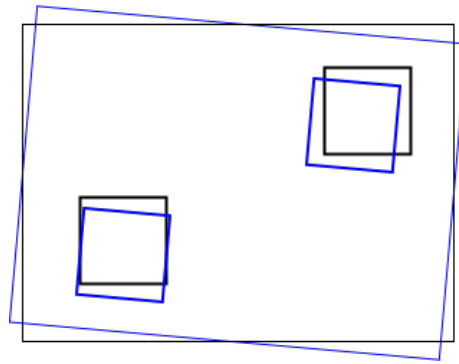


Obrázek 7.7: Ilustrace znázorňující výpočet translace při zanedbání zvětšení výřezů

I když se k referenčnímu výřezu rovnoměrně přičítá hodnota okolí jak vlevo a nahoru, tak i vpravo a dolů, postačí hodnota velikosti okolí přidaná doleva a nahoru, protože se posun počítá od počáteční hodnoty referenčního výřezu, tedy levého horního rohu. Vypočtenou hodnotu posunu je poté nutno vydělit koeficientem zvětšení, aby se zachovalo stejné měřítko mezi kompenzovanými snímky.

Jakmile se provede výpočet posunu, provede se test, jestli je posun jednotlivých barevně anotovaných oblastí vůči nalezeným korespondencím na jiném snímku nad velikost maximálně deset pixelů. V případě kladné odpovědi nemá smysl provádět kompenzaci obrazu, protože při tak vysokém posunu již může docházet k zachycení jiné roviny než původní a afinní transformace by již byla nedostatečnou pro kompenzaci, ale muselo by se využít principů perspektivní transformace.

Jelikož při snímání obrazu nelze zajistit, že při mírné změně úhlu fotoaparátu bude střed rotace uprostřed snímku, vzniká zde problém s rotací obrazu. Jednotlivé výřezy, u kterých se vypočítá posun podle rovnice 7.2, tak mají mezi odpovídajícími výřezy z referenčního snímku různé hodnoty, tedy například levá část obrazu je natočena více dolů, než pravá nahoru pro vodorovnou osu a podobně i pro svislou osu. Podobná situace je znázorněna na obrázku 7.8. Z tohoto důvodu nelze kompenzaci provádět pouhou rotací snímku, protože při snímání může dojít k mírnému posunu. I když je snímek rotovaný, lokálně se to projevuje jako posun.



Obrázek 7.8: Ilustrace znázorňující různý posun a rotaci jednotlivých oblastí v rámci jednoho snímku

Aktuální snímek je proto třeba transformovat. Pro transformaci se používá maticové násobení mezi obrazem a transformační maticí.

I když se pro odhad afinní transformace dá použít dvou různých snímků, ve kterých se automaticky naleznou významné body, ze kterých se poté provede odhad, nelze této možnosti prakticky využít. Důvodem je fakt, že se na snímku mohou vyskytovat objekty, které mohou měnit polohu či svoji velikost, a které zároveň mohou na svém povrchu obsahovat významné body. Z tohoto důvodu jsou do výpočtu transformační matice zahrnuty veškeré nalezené významné body z výřezů, uložené ve vektoru (viz 7.3) jejich nalezené korespondence, které jsou uloženy v jiném vektoru (viz 7.4).

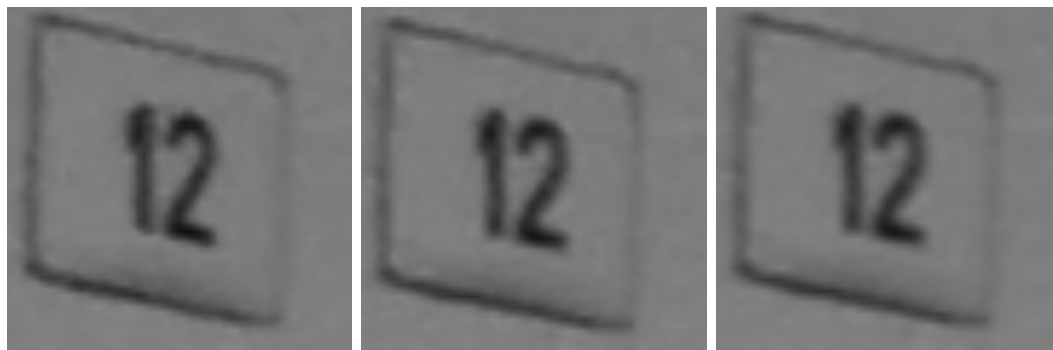
Jednotlivé elementy z obou zmíněných vektorů jsou postupně vkládány do dalších dvou vektorů, určených k výpočtu transformační matice, přičemž vkládání předchází výpočet rozdílu jednotlivých dvojic souřadnic. Je-li rozdíl příliš velký, pravděpodobně byla nalezena špatná korespondence a oba testované body se poté nehodí pro odhad transformace. Odstranění bodů z původních vektorů by mohlo způsobovat ztráty informací pro jiné snímky ze sady snímků (například nalezený roh bude překryt nějakým objektem, tím se odpovídající korespondence nenalezne, popřípadě bude nepřesná). Další body, které se do těchto vektorů přidávají, jsou souřadnice bodů, reprezentující jednotlivé barevně zvýrazněné oblasti a jejich odpovídající korespondence.



## 7.6 Generování výsledků a kontrola správnosti

Poté, co je nalezena hodnota vzájemných posunů u výřezů získaných z referenčního snímku a aktuálního snímku, a provede se kompenzace obrazu pomocí afinní transformace, jsou generovány výstupní soubory všech kompenzovaných výřezů. Dále se generují grafy, obsahující údaje při různém převzorkování jednotlivých snímků. Nakonec se provede výpočet chyby afinní transformace, který se rovněž zaznamená do grafu.

Aby se vygenerovaný obraz mohl dále analyzovat, je dobré si získané výřezy uložit jako soubory pro pozdější využití. Nové snímky se ukládají převzorkované, a proto s větším rozlišením snímku roste i velikost výřezů.



Obrázek 7.9: Zvětšený výřez z referenčního snímku a nalezené korespondence ze sady snímků před a po provedení afinní transformace

Ke každé sadě snímků je vytvořen adresář `out`, ve kterém se vygeneruje jedenáct adresářů, kde je jeden adresář určen pro generování grafů (pojmenován `graphs`), a zbylých deset adresářů je určených pro uložení výstupních souborů, zejména výřezů. Adresáře jsou pojmenovány `x_resize`, kde `x` je celočíselná hodnota, značící míru zvětšení. V jednotlivých adresářích je tolik adresářů, kolik oblastí bylo na referenčním snímku zvýrazněno, pojmenovaných anglickým překladem českých názvů, uvedených v 6.1.

V každém z těchto adresářů je uloženo několik souborů, které obsahují výřezy, získané předchozími kroky. Vždy je vytvořen soubor, který obsahuje výřez z referenčního snímku, dále se mohou vytvářet snímky, obsahující nalezené odpovídající korespondence (jsou-li nalezeny) a jejich kompenzace vůči snímku referenčnímu. Způsob pojmenování souborů je podle vzorce: `id_mul_(x_dis,y_dis)_type.jpg`, kde `id` je kladné celé číslo, určující pořadí snímku ze sady snímků, přičemž číslování začíná číslicí jedna. Nula značí referenční snímek. Následuje kladné celé číslo `mul`, které značí míru zvětšení snímku, vypočtený horizontální a vertikální posun oblasti, a řetězec `type`, který obsahuje buď slovo `REF`, značící, že se jedná o výřez z referenčního snímku, slovo `NONE`, značící, že se jedná o výřez nalezené odpovídající korespondence, nebo slovo `COMP`, značící výřez nalezené korespondence po kompenzaci snímku.

Na obrázku 7.9 se vlevo nachází zvětšený uložený výřez z referenčního snímku, uprostřed se nachází nalezená korespondence výřezu před afinní transformací a vpravo se nachází nalezená korespondence výřezu po afinní transformaci. Prostřední snímek je vůči ostatním více rotován, což lze pozorovat u obvodových hran nalezeného objektu. Obrázek po afinní transformaci je vlivem interpolace mírně rozostřen vůči stavu před transformací.

Mimo jiné je pro každé zvětšení vytvořen i soubor (obrázek 7.10), který obsahuje referenční snímek, na kterém jsou barevně zvýrazněny vyhledané body. Červenou kružnicí

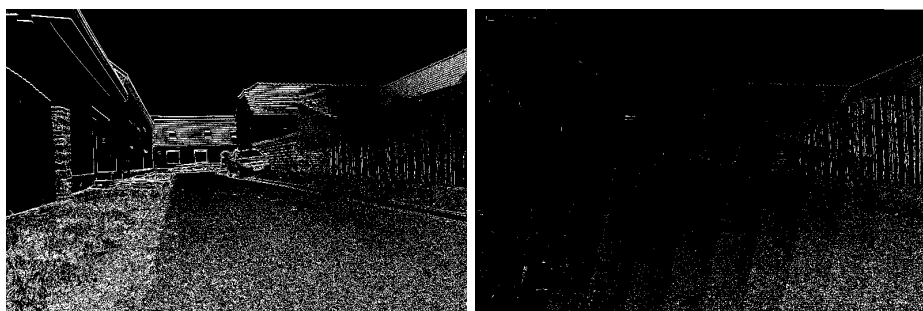
jsou zvýrazněny body, reprezentující barevně zvýrazněné oblasti, zelenou kružnicí jsou zvýrazněny body, které reprezentují významné body, nalezené v barevně zvýrazněné oblasti a modrou kružnicí je zvýrazněn bod, reprezentující barevně zvýrazněné oblasti (podobně jako červený, ovšem po afinní transformaci).

Pro každou zmíněnou barvu kružnic existují její dva různé odstíny, přičemž světlejší odstín značí nalezené body na referenčním snímku a tmavší odstín značí pozici odpovídajících korespondencí na aktuálním snímku. U červených bodů se v jejich blízkosti nachází i textový popisek, obsahující pozice zvýrazněného bodu. Z těchto popisků lze v rámci celého snímku vyčíst, že každá zvýrazněná oblast nemá stejný posun.



Obrázek 7.10: Zvětšený výřez z referenčního snímku, obsahující zvýrazněné body

Dalším souborem, který se generuje v adresářích obsahující výřezy, je aktuální snímek po afinní transformaci. Tento snímek se tvoří k manuální kontrole snímku, jestli se afinní transformace snímku, provedla v pořádku.



Obrázek 7.11: Rozdíl referenčního snímku a aktuálního snímku před a po provedení afinní transformace

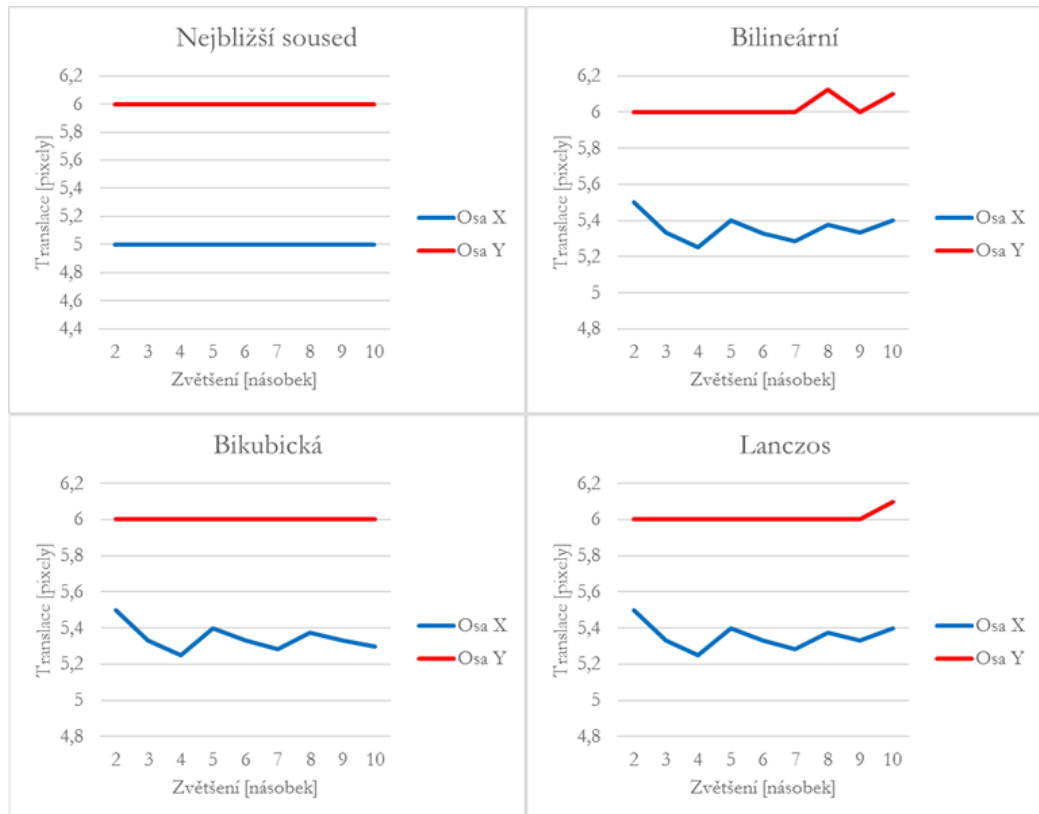
Jedním ze způsobů, jak lze správnost transformace ověřit, je provedení rozdílu mezi referenčním snímekem a transformovaným aktuálním snímekem. V případě, že jsou jednotlivé pixely obou snímků totožné, je jejich vzájemný rozdíl nulový, což se zobrazí jako černý snímek. V opačném případě jsou rozdíly jednotlivých pixelů nenulové, což se projeví jako odstín šedi, protože snímky jsou šedotónové.

Protože pořízené snímky nejsou totožné a jsou vzájemně rotované a posunuté, rozdílem všech pixelů se zvýrazní především hrany a objekty, které se pohybují. Je-li vzájemný posun větší, je i výraznější hrana zobrazena v rozdílu.



Na obrázku 7.11 jsou zobrazeny dva rozdíly výše zmíněných snímků, kde v levé části je zobrazen snímek před provedením afinní transformace a v pravé části je po provedení afinní transformace. V ideálním případě by byl rozdíl snímků po afinní transformaci nulový, tedy snímek rozdílu by byl celý černý. Pro účely tisku bylo provedeno u obou obrázků prahování stejným prahem.

Dalším možným způsobem, jak lze zjistit přesnost určení vzájemných posunů, je porovnat vypočítané translace při různém převzorkování a sledovat trend jednotlivých posunů. Toho lze docílit tím, že se vytvoří pro každý snímek několik grafů (vždy pro každou barevně anotovanou oblast), které tyto informace obsahují.

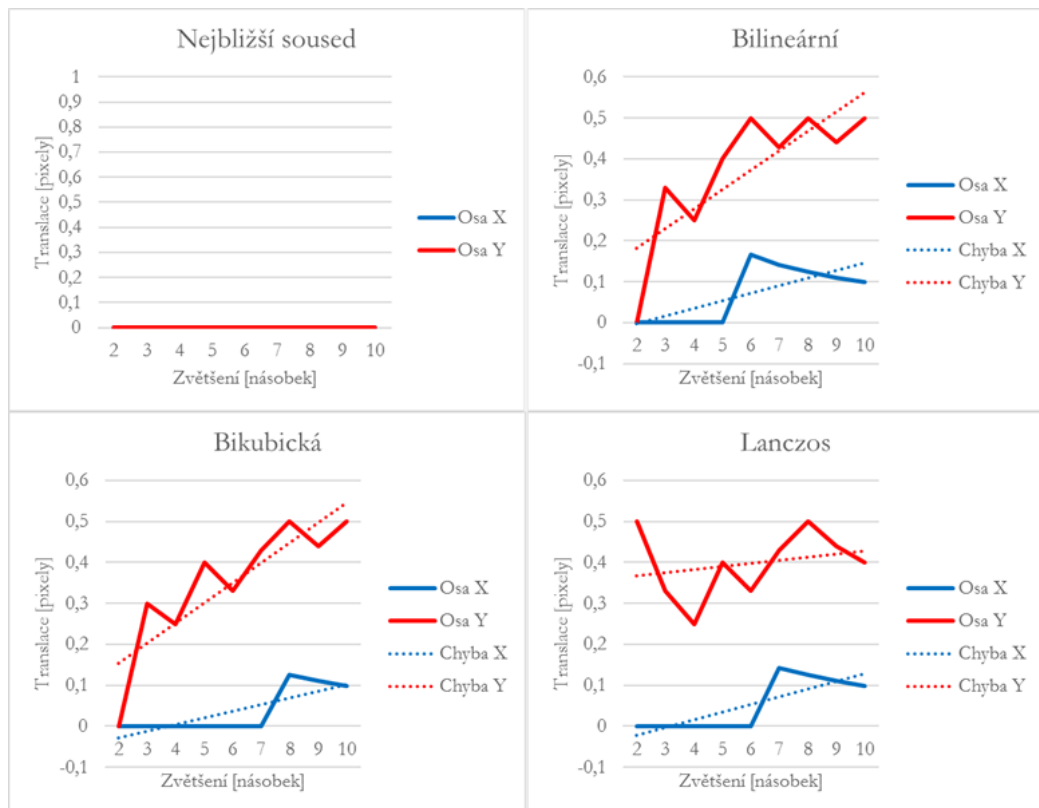


Obrázek 7.12: Zjištěný posun při různém zvětšení pomocí (zleva nahoře) metody nejbližšího souseda, bilineárního, bikubického a Lanczosova převzorkování

Na obrázku 7.12 jsou zobrazeny vygenerované grafy pro oblast zvýrazněný purpurovou barvou ve stejné sadě snímků, jaká je na obrázku 7.2. Jednotlivé grafy znázorňují velikost translace v pixelech (vertikální osa grafu), v závislosti na zvětšení obrazu (horizontální osa grafu). Vypočítaný posun snímku v horizontálním směru snímku je označen tmavě modrou křivkou. Posun ve vertikálním směru snímku je znázorněn červenou křivkou. Při zvětšení obrazu o  $N$  násobek je dosažitelná přesnost  $1/N$ .

V levém horním grafu je zjištěný posun s použitím převzorkování metodou nejbližšího souseda. Tato velmi rychlá metoda detekovala posun obrazu, ovšem ne s přesností menší než celý pixel. Na dalším grafu je použito převzorkování bilineární. Tato metoda je oproti předchozí přesnější, ale produkuje horší výsledky než bikubické převzorkování, která je na dalším grafu. I když je výpočetně náročnější proti předchozím dvěma zmíněným převzorkováním, není tak výpočetně náročná jako Lanczosovo převzorkování, uvedené v posledním

grafu. Lze říci, že bikubické převzorkování má větší smysl použít z hlediska velmi podobné přesnosti s Lanczosovým převzorkováním.

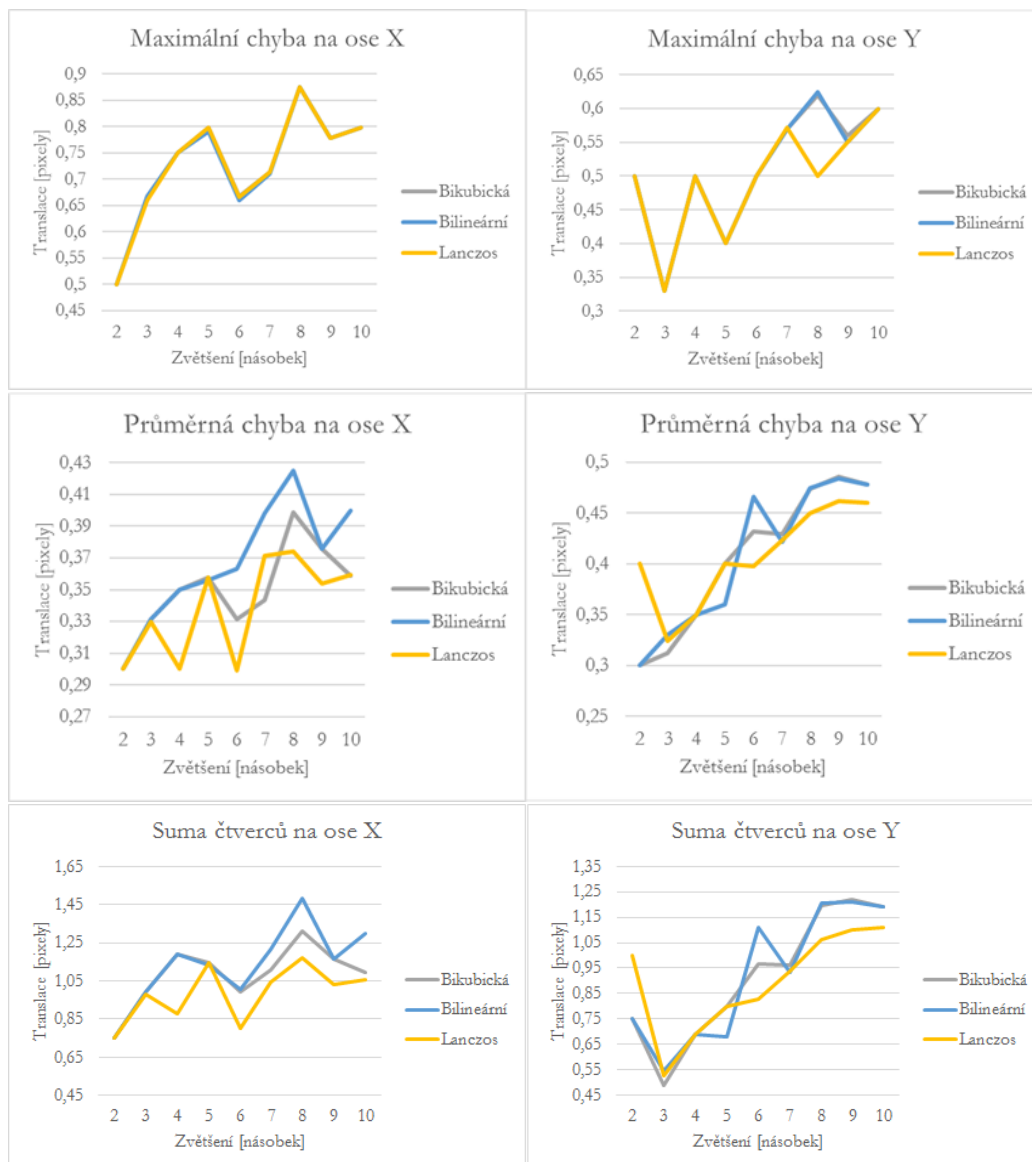


Obrázek 7.13: Zjištěný posun při různém zvětšení pomocí (zleva nahoře) metody nejbližšího souseda, bilineárního, bikubického a Lanczosova převzorkování v obraze po provedení afinní transformace

Na obrázku 7.13 jsou zobrazeny vygenerované grafy pro oblast zvýrazněný purpurovou barvou ve stejné sadě snímků, jaká je na obrázku 7.2, po provedení afinní transformace, což reprezentuje chybu. Osy jsou stejné jako u 7.12. Vypočítaný posun snímku v horizontálním směru snímku je označen modrou křivkou. Posun ve vertikálním směru snímku je znázorněn červenou křivkou. Při zvětšení obrazu o  $N$  násobek je dosažitelná přesnost  $1/N$ .

V levém horním grafu je zjištěný posun s použitím převzorkování metodou nejbližšího souseda. Bylo zjištěno, že údajně k žádnému posunu nedochází, což bylo experimentálně otestováno rozdílem referenčního výřezu a aktuálního výřezu po transformaci. Bylo zjištěno, že zjištěné posuny nejsou přesné. I když další metody výpočtu produkují velmi podobné a relativně přesné výpočty, z posledního grafu je patrné, že Lanczosovo převzorkování dokázalo přesněji odhalit chybu již při nízkém zvětšení.

Dále jsou u grafů, kde není průběh po celou dobu nulový, vyznačeny chyby, vypočítané pomocí metody nejmenších čtverců, proložené přímkou. Tyto přímky jsou označeny tečkovaně, přičemž barevně odpovídají křivkám translace. Podle strmosti těchto křivek lze zjistit, že Lanczosovo převzorkování, které je má červenou tečkovanou přímkou nejméně strmou, je zatíženo chybou nejméně.



Obrázek 7.14: Maximální a průměrná chyba a suma čtverců bilineárního, bikubického a Lanczosova převzorkování v obraze po provedení afinní transformace pro všechny oblasti

Vezmou-li se v potaz veškeré oblasti, které jsou v obraze zastoupeny, tak vůči sobě má každá oblast jiné hodnoty chyby transformace. Proto je žádoucí udělat statistickou analýzu pro celý snímek dohromady.

Na obrázku 7.14 je několik grafů, kde je znázorněna chyba transformace (vertikální osa grafu) v závislosti na zvětšení (horizontální osa grafu). Na řádku jsou vždy dva grafy, jeden pro translace na ose X, druhý pro osu Y. Na prvním řádku je zobrazena maximální chyba při různém zvětšení. Až na pár hodnot jsou průběhy funkcí velmi podobné.

Na druhém řádku je průměrná chyba při různém zvětšení. V grafu jde vidět, že bilineární převzorkování má vyšší zatížení chybou, než ostatní převzorkování. Nejlépe dopadlo Lanczosovo převzorkování.

Na posledním řádku je znázorněn průběh funkcí pro sumu čtverců. I zde má Lanczosovo převzorkování nejlepší výsledky.

# Kapitola 8

## Závěr

Cílem práce bylo implementovat aplikaci, která detekuje, jestli se posunula uchycená kamera ze stereo kamery pomocí korelace. Cíl byl splněn.

V případě, že nastala změna oproti původní konfiguraci, tak se pokusit kompenzovat tuto změnu pomocí afinní transformace z bodů v obraze, nalezených pomocí korelace a automatické detekce významných bodů. Dále provést ověření, jestli vypočítaný pohyb kamery je správný a zjistit, jaký vliv má zvětšení snímku na přesnost výpočtu posunů, přičemž všechny výsledky uložit do souborů.

V teoretické části bylo popsáno, co je to obraz, jak je reprezentován, princip digitalizace pomocí vzorkování a kvantování, operace a transformace nad obrazem, především byl popsán princip detekce rohů, konvoluce, korelace a převzorkování. Dále bylo uvedeno, co to je stereo kamera a epipolární geometrie.

V praktické části bylo uvedeno, jakým způsobem probíhala tvorba sad snímků, uvedeny byly požadavky na jednotlivé snímání scény a dále důvody ruční anotace jednotlivých sad, včetně popisu vhodných kandidátních objektů pro anotování. Dále byl detailně popsán postup zjištění posunu mezi dvěma snímky pomocí vzájemné korelace při různém zvětšení obrazů a charakterizován postup nalezení výrazných bodů v obraze. Též zde byla uvedena afinní transformace jako nástroj pro ověření správnosti vypočtených vzájemných posunů snímků. V závěru praktické části byly uvedeny dosažené výsledky.

Osobně si myslím, že všechny cíle kladené na řešení tohoto problému byly splněny. Myslím si, že dosažené výsledky jsou dobré, a lze tedy tuto metodu využít i v praxi u stereo kamer, které jsou součástí složitějších systému pro zpracování obrazu.

V budoucnu by bylo dobré upravit tuto aplikaci jako knihovnu, která by byla snadno přenositelná mezi operačními systémy a dala by se použít pro kompenzaci obrazu přímo na stereo kameře.

# Literatura

- [1] MATLAB - MathWorks. Natick, Massachusetts: The MathWorks, 1994-2017, [Online; navštíveno 13.05.2017].  
URL <https://www.mathworks.com/products/matlab.html>
- [2] OpenCV library. Bradsky, G., 2000-2017, [Online; navštíveno 13.05.2017].  
URL <http://opencv.org/>
- [3] Computer stereo vision. San Francisco (CA): Wikimedia Foundation, 2001-2017, [Online; navštíveno 13.05.2017].  
URL [https://en.wikipedia.org/wiki/Computer\\_stereo\\_vision](https://en.wikipedia.org/wiki/Computer_stereo_vision)
- [4] Corner detection. San Francisco (CA): Wikimedia Foundation, 2001-2017, [Online; navštíveno 13.05.2017].  
URL [https://en.wikipedia.org/wiki/Corner\\_detection](https://en.wikipedia.org/wiki/Corner_detection)
- [5] Cross-correlation. San Francisco (CA): Wikimedia Foundation, 2001-2017, [Online; navštíveno 28.01.2017].  
URL <https://en.wikipedia.org/wiki/Cross-correlation>
- [6] Image rectification. San Francisco (CA): Wikimedia Foundation, 2001-2017, [Online; navštíveno 13.05.2017].  
URL [https://en.wikipedia.org/wiki/Image\\_rectification](https://en.wikipedia.org/wiki/Image_rectification)
- [7] Lanczos resampling. San Francisco (CA): Wikimedia Foundation, 2001-2017, [Online; navštíveno 09.05.2017].  
URL [https://en.wikipedia.org/wiki/Lanczos\\_resampling](https://en.wikipedia.org/wiki/Lanczos_resampling)
- [8] Stereo camera. San Francisco (CA): Wikimedia Foundation, 2001-2017, [Online; navštíveno 13.05.2017].  
URL [https://en.wikipedia.org/wiki/Stereo\\_camera](https://en.wikipedia.org/wiki/Stereo_camera)
- [9] The Discrete Fourier Transform, Part 6. *The Journal of Object Technology*, ročník 9, č. 2, 2010: s. 17–18, ISSN 1660-1769, doi:10.5381/jot.2010.9.2.c2, [Navštíveno 28.01.2017].  
URL [http://www.jot.fm/contents/issue\\_2010\\_03/column2.html](http://www.jot.fm/contents/issue_2010_03/column2.html)
- [10] Sony a5000. Praha: Sony Europe Limited, 2014, [Online; navštíveno 28.01.2017].  
URL <https://www.sony.cz/electronics/fotoaparaty-s-vymennymi-objektivy-2/ilce-5000-kit>
- [11] OpenCV: Object Detection. Bradsky, G., 2015, [Online; navštíveno 28.01.2017].  
URL [http://docs.opencv.org/3.0.0/df/dfb/group\\_\\_imgproc\\_\\_object.html#ga586ebfb0a7fb604b35a23d85391329be](http://docs.opencv.org/3.0.0/df/dfb/group__imgproc__object.html#ga586ebfb0a7fb604b35a23d85391329be)

- [12] Hlaváč, V.; Šonka, M.: *Počítačové vidění*. Praha: Grada, první vydání, 1992, ISBN 80-854-2467-3.
- [13] Jacobs, D.: *Correlation and Convolution*. University of Maryland, 2005, [Online; navštíveno 28.01.2017].  
URL <http://www.cs.umd.edu/~djacobs/CMSC426/Convolution.pdf>
- [14] Jan, J.: *Číslíková filtrace, analýza a restaurace signálů*. Brno: VUTIUM, druhé vydání, 2002, ISBN 80-214-1558-4.
- [15] Žára Jiří; Beneš, B.; Sochor, J.; aj.: *Moderní počítačová grafika*. Brno: Computer Press, vyd 1. vydání, 2004, ISBN 80-251-0454-0.
- [16] John W. Eaton, S. H., David Bateman; Wehbring, R.: *GNU Octave version 3.8.1 manual: a high-level interactive language for numerical computations*. CreateSpace Independent Publishing Platform, 2014, ISBN 1441413006.  
URL <http://www.gnu.org/software/octave/doc/interpreter>
- [17] Sonka, M.; Hlavac, V.; Boyle, R.: *Image processing, analysis, and machine vision*. Toronto: Thompson Learning, třetí vydání, c2008, ISBN 978-0-495-08252-1.

# Přílohy

# Příloha A

## Konfigurační soubor

V některých případech není nutné, aby se využilo všech přednastavených konstant a nastavení pro správný chod aplikace. Z tohoto důvodu se dá využít textový soubor `configuration.txt`, který je umístěn ve stejném adresáři jako spustitelný soubor.

Tento konfigurační soubor umožňuje uživateli upravit vlastnosti chodu programu. Lze jím upravit následující přednastavené konstanty pomocí atributů:

- *corners*: hodnota reprezentující maximální počet automaticky hledaných významných bodů v obraze. Zadaná hodnota musí být celočíselná a z intervalu  $\langle 1; 20 \rangle$ . Při zadání jiné hodnoty se použije přednastavená hodnota *corners* = 10 (viz 7.3).
- *mul*: hodnota, která reprezentuje maximální možné zvětšení jednotlivých výřezů. Zadaný koeficient zvětšení musí být celočíselný a z intervalu  $\langle 1; 10 \rangle$ . V případě zadání jiné hodnoty se použije přednastavený koeficient zvětšení *mul* = 10 (viz 7.4).
- *affine*: přepínač mezi povolením a zakázáním generování snímku z adresáře `img` po afinní transformaci (viz 7.6). Může obsahovat pravdivostní hodnoty *true* (povolení generování snímku) a *false* (zakázání generování snímku). Výchozí hodnotou je *affine* = *true*.
- *show\_labels*: přepínač umožňující nastavení možnosti přidání do obrazu obsahující zvýrazněné body (viz 7.6) textové popisy: souřadnice bodů, reprezentující barevně anotované oblasti. Lze použít logické hodnoty *true* (zobrazení) a *false* (nezobrazení), přičemž výchozí hodnotou je *show\_labels* = *true*.
- *interp*: hodnota reprezentující způsob převzorkování snímku při zvětšení. Lze použít převzorkování bilineární (*linear*), nejbližším sousedem (*nearest*), Lanczos (*lanczos*) či bikubické (*cubic*), které je i výchozí hodnotou.
- *from* a *to*: hodnoty značící adresáře, které se mají kompenzovat (viz 6.2). První zadaná hodnota *from* musí být celočíselná a reprezentuje první adresář ze sady snímků. Druhá zadaná hodnota *to* musí být taktéž celočíselná a reprezentuje poslední adresář ze sady snímků. Sady mezi těmito krajními budou při spuštění programu kompenzovány, přičemž *from*  $\leq$  *to*. Výchozí hodnoty jsou *from* = 1 a *to* = 9.

V konfiguračním souboru lze psát i komentáře, které se při zpracování požadavků ignorují. Komentář je uvozen symbolem `#` a je ukončen znakem konce řádku. Jako komentář se může brát i text, který neobsahuje od začátku řádku název atributu, uvedený ve výčtu v této kapitole.



Při zadání nepodporovaných hodnot pro jednotlivé atributy se použijí hodnoty, které jsou přednastavené v programu. Dále se přednastavené hodnoty použijí i v případě absence některého atributu. Jednotlivé atributy a jejich hodnoty jsou závislé na velikosti písmen (case sensitive). V případě absence celého konfiguračního souboru se vygeneruje nový soubor při prvním spuštění aplikace.

## Příloha B

# Seznam použitých OpenCV funkcí

- `void cv::warpAffine(InputArray src, OutputArray dst, InputArray M, Size dsize, int flags = INTER_LINEAR, int borderMode = BORDER_CONSTANT, const Scalar & borderValue = Scalar())`
- `void cv::resize(InputArray src, OutputArray dst, Size dsize, double fx = 0, double fy = 0, int interpolation = INTER_LINEAR)`
- `Mat cv::estimateRigidTransform(InputArray src, InputArray dst, bool fullAffine)`
- `Mat cv::imread(const String & filename, int flags = IMREAD_COLOR)`
- `bool cv::imwrite(const String & filename, InputArray img, const std::vector<int> & params = std::vector<int>())`
- `static Vec<Tp, cn> cv::normalize(const Vec<Tp, cn> & v)`
- `void cv::matchTemplate(InputArray image, InputArray templ, OutputArray result, int method, InputArray mask = noArray())`
- `void cv::goodFeaturesToTrack(InputArray image, OutputArray corners, int maxCorners, double qualityLevel, double minDistance, InputArray mask = noArray(), int blockSize = 3, bool useHarrisDetector = false, double k = 0.04)`
- `double cv::threshold(InputArray src, OutputArray dst, double thresh, double maxval, int type)`
- `void cv::cvtColor(InputArray src, OutputArray dst, int code, int dstCn = 0)`
- `Mat cv::cvarrToMat(const CvArr *arr, bool copyData = false, bool allowND = true, int coiMode = 0, AutoBuffer<double> *buf = 0)`

# Seznam obrázků

2.1	Perspektivní zobrazení . . . . .	5
2.2	Histogram . . . . .	7
3.1	Obrázek 7.2 před a po ekvalizaci histogramu a odpovídající histogramy . . .	10
3.2	Nejednoznačnost detektoru hran v rohu objektu . . . . .	14
3.3	Ilustrace rozhodování Harrisova detektoru rohů . . . . .	15
4.1	Prohledávaný prostor před a po rektifikaci . . . . .	17
4.2	Geometrie dvou kamer . . . . .	17
5.1	Ilustrace změny natočení části stereokamery . . . . .	19
5.2	Ilustrace posunu o méně než pixel . . . . .	20
5.3	Ilustrace aktuálního snímku promítnutého na snímek referenční . . . . .	21
6.1	Referenční snímek a kandidátní oblasti pro anotace . . . . .	24
6.2	Referenční snímek doplněný o zvýrazněné oblasti . . . . .	25
7.1	Diagram znázorňující algoritmus pro určení korelace a kompenzace dvou snímků . . . . .	27
7.2	Referenční snímek převedený na šedotónový obraz . . . . .	28
7.3	Vytvořený výřez a jeho maska . . . . .	29
7.4	Aktuální výřez spolu s okolím . . . . .	30
7.5	Výřez, ve kterém se detekovalo deset bodů . . . . .	31
7.6	Korelační koeficienty . . . . .	32
7.7	Ilustrace znázorňující výpočet translace při zanedbání zvětšení výřezů . . .	33
7.8	Ilustrace znázorňující různý posun a rotaci jednotlivých oblastí v rámci jednoho snímku . . . . .	34
7.9	Zvětšený výřez z referenčního snímku a nalezené korespondence ze sady snímků před a po provedení afinní transformace . . . . .	35
7.10	Zvětšený výřez z referenčního snímku, obsahující zvýrazněné body . . . . .	36
7.11	Rozdíl referenčního snímku a aktuálního snímku před a po provedení afinní transformace . . . . .	36
7.12	Zjištěný posun při různém zvětšení pomocí (zleva nahoře) metody nejbližšího souseda, bilineárního, bikubického a Lanczosova převzorkování . . . . .	37
7.13	Zjištěný posun při různém zvětšení pomocí (zleva nahoře) metody nejbližšího souseda, bilineárního, bikubického a Lanczosova převzorkování v obraze po provedení afinní transformace . . . . .	38

7.14 Maximální a průměrná chyba a suma čtverců bilineárního, bikubického a Lanczosova převzorkování v obraze po provedení afinní transformace pro všechny oblasti . . . . .	39
---	----

# Seznam tabulek

6.1	Seznam použitých barev pro zvýraznění oblastí . . . . .	25
6.2	Seznam vytvořených sad snímků a jejich informace . . . . .	26