



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

VÝPOČETNÍ RE-FOTOGRAFIE

COMPUTATIONAL RE-PHOTOGRAPHY

BAKALÁŘSKÁ PRÁCE

BACHELOR'S THESIS

AUTOR PRÁCE

AUTHOR

ADAM ČERVENKA

VEDOUCÍ PRÁCE

SUPERVISOR

Doc. Ing. MARTIN ČADÍK, Ph.D.

BRNO 2017

Vysoké učení technické v Brně - Fakulta informačních technologií

Ústav počítačové grafiky a multimédií

Akademický rok 2016/2017

Zadání bakalářské práce

Řešitel: **Červenka Adam**

Obor: Informační technologie

Téma: **Výpočetní re-fotografie**
Computational Re-Photography

Kategorie: Uživatelská rozhraní

Pokyny:

1. Seznamte se s problematikou re-fotografie, jejímž cílem je zachytit nový snímek odpovídající existující (historické) fotografii.
2. Navrhněte a implementujte systém pro re-fotografii, který s využitím technik počítačového vidění pomůže fotografovi nalézt správnou pozici a orientaci kamery.
3. Se systémem experimentujte, posuďte jeho vlastnosti a diskutujte možnosti budoucího vývoje.
4. Dosažené výsledky prezentujte formou videa, plakátu, článku, apod.

Literatura:

- <http://cadik.posvete.cz/locate/>
- <http://people.csail.mit.edu/soonmin/rephoto/rephoto.pdf>

Pro udělení zápočtu za první semestr je požadováno:

- Body 1 a 2 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: **Čadík Martin, doc. Ing., Ph.D.**, 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 01 Brno, Bcžetěchova 2



doc. Dr. Ing. Jan Černocký
vedoucí ústavu

Abstrakt

Požizování fotografií podle historických snímků je předmětem zájmu jak pro fotografy, tak i badatele. Tímto způsobem lze pozorovat vývoj místa, případně jeho okolí, v časovém rozmezí. Problém může nastat v případě nalezení stejného pohledu jako u referenčního snímku. Nalezení totožného pohledu se obvykle řeší jen manuálně. Fotografové se v takovém případě musí spoléhat na odhad a fotografický um. Tento způsob může být nepřesný a pro fotografa značně nepohodlný. Řešení představuje refotografický algoritmus, který využívá postupy z oboru počítačového vidění.

Abstract

Photographing by historical picture is subject of interest both for photographers and researcher. In this way, it is possible to observe development of site, or its surroundings, within the timeframe. Problem may come to pass when we want to found the same view as the reference picture. Finding the same view is usually solve only by manually. In this case, photographers must rely to estimation and photographic skill. This method may be inaccurate and for photographer considerably inconvenient. The solution is the refutographic algorithm that using computer vision techniques.

Klíčová slova

Refotografie, Android OS, OpenCV, 3D rekonstrukce, Real-time aplikace

Keywords

Rephotography, Android OS, OpenCV, 3D reconstruction, Real-time application

Citace

ČERVENKA, Adam. *Výpočetní re-fotografie*. Brno, 2017. Bakalářská práce. Vysoké učení technické v Brně, Fakulta informačních technologií. Vedoucí práce Doc. Ing. Martin Čadík, Ph.D.

Výpočetní re-fotografie

Prohlášení

Prohlašuji, že jsem tuto bakalářskou práci vypracoval samostatně pod vedením pana Doc. Ing. Martina Čadíka, Ph.D. Uvedl jsem všechny literární prameny a publikace, ze kterých jsem čerpal.

.....

Adam Červenka

17. května 2017

Poděkování

Chtěl bych poděkovat svému vedoucímu bakalářské práce Doc. Ing. Martinovi Čadíkovi, Ph.D. za odborné vedení a cenné rady při vypracování projektu.

Obsah

1	Úvod	2
2	Analýza	3
2.1	Požadavky	4
2.2	Podobné řešení	4
2.3	Seznámení s problémy	5
3	Kalibrace neznámého fotoaparátu	8
3.1	3D rekonstrukce	9
3.2	Optický střed	10
3.3	Pozice historického fotoaparátu	11
4	Navigace uživatele v reálném čase	12
4.1	Robustní odhad	13
4.2	Lehký odhad	13
4.3	Relativní matice	14
5	Implementace	16
5.1	Odhad snímků	17
5.2	Pozice referenčního snímku	19
5.3	Výsledné porovnání	20
5.4	Propojení se systémem android	21
6	Testování	22
6.1	3D rekonstrukce	22
6.2	Optický střed	23
6.3	Pozice fotoaparátů	25
7	Dosažené výsledky	27
7.1	Podoba aplikace	29
7.2	Refotografie	30
7.3	Experimenty	32
8	Závěr	34
	Literatura	35

Kapitola 1

Úvod

Refotografování je moderní trend, který spočívá ve znovu vyfotografování stejného místa zpravidla dle předlohy představující historický snímek. Od prvních fotoaparátů uplynula velká řádka let. První fotografie jsou tedy němými svědky mnohdy více jak stoleté minulosti. Zachycená místa našeho okolí si prošla dlouhou cestou a patřičnými proměnami, aniž bychom si to uvědomovali. Jestliže se porovná sto let stará fotografie se současným místem, lze dospět k pozoruhodným závěrům.

Samotný proces refotografie přináší z pohledu počítačového vidění mnoho problémů a překážek. Už samotná historická fotografie představuje hned několik komplikací. Referenční fotografie je pořízená na neznámém fotoaparátu, a proto se důležité parametry v rámci kalibrace musí dopočítat. Referenční místo může být příliš změněné. Mohou zde být také jiné faktory, které porovnání referenčního a aktuálního snímku komplikují nebo ho naprosto znemožňují.

K řešení zmíněných problémů byly využity znalosti trojrozměrné rekonstrukce scény a postupy z oboru počítačového vidění za využití knihovny OpenCV. Návrh aplikace, zvolené postupy a algoritmy vychází z článku Computational rephotography [2], který se touto problematikou refotografie zabývá. Obsahem článku je především analýza možných problémů a nabízející řešení v aplikování postupů a algoritmů z oblasti počítačového vidění.

Cílem refotografického algoritmu je odhadnout pozici referenčního snímku a poskytnout uživateli navigaci na tuto pozici. Uživateli je tím umožněno pořídit refotografii z totožné pozice jako u historického fotografa. Vstupní hodnotou je historická fotografie reprezentující referenční snímek. Výsledkem je refotografie objektu zaznamenaného na referenčním snímku. Historický snímek s dobře zarovnanou moderní refotografií může sloužit jako pozoruhodná vizualizace ukazující časový vývoj staveb či jejich okolí. Refotografický nástroj tak osloví především seriózní historiky, amatérské badatele a fotografy.

Následující sekce dokumentace popisují problematiku refotografického procesu z pohledu oboru počítačového vidění. Obsah je zaměřen na realizaci refotografického algoritmu od analýzy problému po testování. Před realizací je třeba definovat základní požadavky na refotografický algoritmus a potenciální problémy (viz sekce 2). Návrh byl rozdělen do dvou částí. První se zabývá trojrozměrnou rekonstrukcí a výpočtem kalibračních parametrů (viz sekce 3). Druhá část popisuje výpočet navigace v reálném čase (viz sekce 4). Podrobnější informace o fungování refotografického algoritmu, včetně pseudokódů, jsou popsány v sekci věnující se implementaci (viz sekce 5). Následuje kapitola popisující výsledky testování (viz sekce 6) a výslednou podobu aplikace (viz sekce 7).

Kapitola 2

Analýza

Před realizací projektu proběhla jeho analýza. Byly zformulovány základní požadavky na refotografický algoritmus. Analýzou se také odhalily potenciální problémy algoritmu. Pro tvorbu analýzy posloužili projekty, které mají podobné cíle jako popisovaná práce. Tyto projekty sloužily taktéž jako inspirace během implementace refotografického algoritmu. Pro pochopení celého procesu je uvedena nezbytná teorie zaměřená na odhad pozice fotoaparátu a komplikace spjaté s tímto procesem. Všechny komplikace odhalené analýzou jsou popsány i s možným řešením, které bude podrobně rozebráno v kapitolách zaměřených na návrh aplikace (viz kapitola 2.3). Hlavním cílem kapitoly je uvést potřebnou teoretickou rovinu procesu refotografování a poskytnout základní předpoklady pro pochopení dalších částí dokumentace.

Proces refotografie a tedy i cíl refotografického nástroje lze znázornit na následujícím snímku (viz obr. 2.1). Snímek zachycuje podobu Grandhotelu Brno v době první republiky a v současné době. Ilustrativně byl použit historický fotoaparát pod historickou fotografií, reprezentující tehdejší počátky fotografování a zařízení nVidia SHIELD, na kterém byla pořízena fotografie ze současné doby.

REPHOTOGRAPHY



Obrázek 2.1: Demonstrace refotografie

2.1 Požadavky

Hlavním cílem práce je realizace interaktivního nástroje, který poslouží k vytváření refotografií. Od refotografického algoritmu se očekává, že bude pracovat především s historickými fotografiemi, popřípadě s malbami a kresbami. Jinými slovy vstup algoritmu představuje historický snímek a výstupem je pozice historického fotoaparátu. Uživatel by měl být naveden k pozici historického snímku. Z této pozice by mohl fotograf pořídit výslednou refotografii. Cílem nástroje je tedy nalezení pozice referenčního snímku vzhledem k pozici aktuálně stojícího fotografa a poskytnout navigaci na tuto pozici.

Jakmile fotograf pořídí odpovídající refotografii, naskytne se mu pohled na rozdílnou scénu mnohdy vzdálenou od sebe desítky až stovky let. Fotograf tak pouhým pohledem může učinit hned několik závěrů. Stavební úpravy a vývoj okolí se uživateli promítnou na jedné obrazovce. Smyslem algoritmu je nahradit refotografování pouhým odhadem a umožnit tak badatelům a fotografům jednodušší a přesnější metodu. U fotografů, kteří se snaží refotografii pořídit pouhým odhadem, dochází k nepřesnostem. Refotografii tak nelze brát jako věrohodný zdroj informací, a tudíž z ní nelze vyčíst patřičné závěry. Pro dosažení co největší přesnosti musí fotograf vynaložit mnoho úsilí a času, aniž by měl jistotu, že dosáhl cílového řešení.

Interaktivní nástroj refotografie, který je zde popisován, je určený pro mobilní aplikace s operačním systémem Android. Operační systém Android patří mezi nejrozšířenější systémy pro mobilní aplikace. Tablety nebo mobilní telefony dnes nabízí zabudovaný fotoaparát na vysoké úrovni, který je pro běžného uživatele dostačující. Výsledky jsou prezentovány na obrazovce mobilního zařízení.

2.2 Podobné řešení

Je potřeba uvést také projekty zabývající se alespoň částečně problematikou refotografie. Tyto projekty byly součástí analýzy a na jejich základě se alespoň částečně formovaly specifikace refotografického algoritmu. Vzhledem k charakteru práce se nelze vymezovat jen na obory informatické, ale i práce fotografů, kteří mají stejné cíle jako popisovaný algoritmus.

Refotografický algoritmus je ojedinělým nástrojem k nalezení refotografie. Avšak komplikace s ním spjaté jsou obecně známým tématem v oboru počítačové vidění. Trojrozměrná rekonstrukce je klíčová pro popisovaný algoritmus, ale také se jedná o známé téma, kterým se zabývá mnoho nástrojů. Jedním z nich je Bundler [1]. Bundler představuje nástroj implementovaný v jazyce C++. Zaměřuje se především na trojrozměrnou rekonstrukci, která probíhá na základě pořízených snímků jednoho objektu. Výsledkem je tedy vizualizovaná trojrozměrná rekonstrukce. Zde je podobnost s refotografickým algoritmem, který na základě dvou snímků zrekonstruuje trojrozměrnou rekonstrukci objektu.

Dalším obdobným projektem je projekt s názvem 4D Cities. Cílem 4D Cities je rozpoznat budovy z historických snímků a provést jejich 3D rekonstrukci. Výsledkem je rekonstrukce trojrozměrné scény na základě těchto fotografií. Podobně jako refotografický nástroj, tak i tento projekt pracuje s historickými snímky a rekonstrukcí scény. Tím ovšem podobnost končí. Projekt pracuje na webovém rozhraní. Bližší informace lze najít na webových stránkách 4D Cities (www.cc.gatech.edu/4d-cities).

Pro hledání inspirace a analýzy problému se nelze zaměřit jen na projekty z pohledu informatiky a oboru počítačového vidění. Refotografie je známým tématem mezi komunitou fotografů. Snahou je oslovit tedy i tyto fotografy a zjednodušit jim proces refotografie. Je potřeba dodat, že mnoho z nich dosahuje pozoruhodných výsledků a dokonce uměleckých refotografií. Jejich práce je úctyhodná a pro tento projekt představovala motivaci dosáhnout podobných výsledků.



Obrázek 2.2: Ukázka práce Štefana Brajtera

Za zmínku stojí práce polského fotografa Štefana Brajtera (viz obr. 2.2). Štefan Brajter se zabývá problematikou refotografie. Na základě odhadu vyfotí stejné místo jako na historickém snímku. Výsledek je prezentován jako animace na webovém blogu nebo facebookových stránkách (<http://refotografie.blogspot.cz/>). Inspiruje se především černobílými fotografiemi z prostředí polských měst. Na jeho práci si lze povšimnout, jak velký význam má refotografování a také velkého zájmu ze strany široké veřejnosti. Práce Štefana Brajtera a jemu podobných se staly inspirací pro vytvoření refotografického nástroje.

2.3 Seznámení s problémy

Jak bylo uvedeno výše, refotografický algoritmus si zakládá na odhadu pozice fotoaparátu v prostoru z pořízených snímků a snímání fotoaparátu. Tento proces ovšem přináší mnoho komplikací. Je zapotřebí uvést, že fotografie pořízená fotoaparátem má pouze 2D souřadnice. Třetí souřadnici je potřeba vypočítat vzhledem k pozici fotoaparátu zaměřeného objektu. 3D objekt má vzhledem k fotoaparátu dva pohyby, tj. překlad a rotaci.

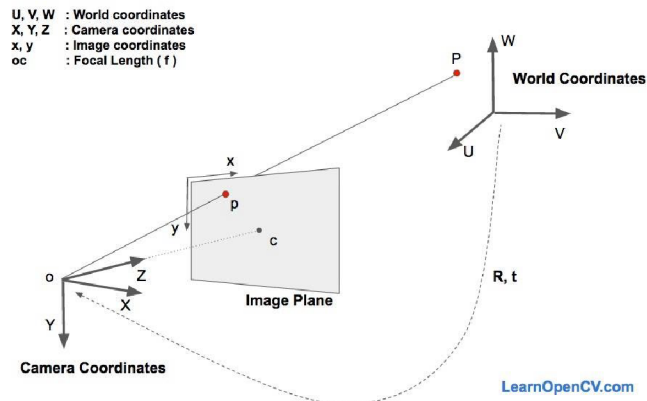
Odhad pozice z 3D objektu tedy znamená vypočítat matici pro překlad a rotaci. Pro výpočet 3D pozice fotoaparátu bude potřeba definovat šest až osm 2D bodů a k nim umístění stejných 3D bodů. Dále pomocí procesu kalibrace fotoaparátu se zjistí zbývající nezbytné parametry, tj. ohnisková vzdálenost a optický střed. Pro úplnost lze dodat, že se k výpočtu připočítá zkreslení, to se ovšem v rámci refotografického nástroje nebere v potaz. Z popisu lze tedy vyvodit velký význam kalibrace pro odhad pozice fotoaparátu (viz rovnice 3.1).

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = R \begin{pmatrix} U \\ V \\ W \end{pmatrix} - t \implies \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = [R|t] \begin{pmatrix} U \\ V \\ W \\ 1 \end{pmatrix}$$

Obrázek 2.3: Rovnice souřadnicového systému

3D souřadnice foceného objektu se promítnou do obrazové roviny s tím, že se 3D objekt promítne jako 2D objekt. Na obrázku níže je vidět zobrazení 3D bodu P do 2D bodu p (viz obr. 2.4).

Předpokládá se, že víme umístění 3D bodu P v globálních souřadnicích s ohledem na parametry kamery. Poté se může vypočítat umístění bodu p ve fotoaparátu souřadnicového systému.



Obrázek 2.4: Odhad pozice fotoaparátu [?]

Tím byla uvedena základní problematika spojená s odhadem pozice fotoaparátu. Ovšem pro realizaci refotografického nástroje lze třeba stanovit dílčí problémy, které přináší komplikace. Po definování základních požadavků a nároků na refotografický algoritmus [2], vyvstanou následující problémy, které lze shrnout do pěti bodů:

1. Výpočet aktuální pozice fotoaparátu v reálném čase může být náročný a aplikace nemusí probíhat plynule.
2. 3D rekonstrukce z fotografie může přinášet značné nejasnosti. Tyto nejasnosti mohou komplikovat výpočet určující pozici aktuální fotoaparátu. Situace se komplikuje s použitím funkce zoom.
3. Relativní odhad pozice se může stát nestabilní v případě dosažení cílové pozice.
4. Historická fotografie může být velmi odlišná od současného stavu v důsledku architektonických úprav, kvality fotografie, počasí atd.
5. Historické fotografie jsou zachyceny pomocí fotoaparátu s neznámými parametry, ohniskovou vzdáleností a optickým středem. Kromě toho byly historické fotografie často zachyceny s nencentralizovaným optickým středem.

První problém sebou přináší vlastnost aplikace běžící v reálném čase (1). Výpočetně náročný odhad pozice aktuální kamery probíhá ve dvou vláknech, kdy první vlákno provádí pomalý robustní odhad a druhé vlákno představuje rychlý lehký odhad pozice uživatele (viz sekce 4). Prokládání obou vláken zajišťuje plynulost výpočtu a uživatel není nikterak omezen pomalou odezvou aplikace.

Aplikace požaduje, aby uživatel pořídil dva snímky objektu s širokou základní linií (*baseline*). To znamená vyfotografovat stejný objekt z odlišného místa nebo úhlu. Pro potřeby této práce se požadují dvě fotografie mírně rozličného úhlu, které zároveň zachycují stejnou scénu jako u referenčního snímku (2). Na základě těchto snímků se provede 3D rekonstrukce scény (viz sekce 3.1). Funkce zoom a ani jiné funkce, které nabízí současné mobilní zařízení se neberou v potaz.

V průběhu navigace uživatele v reálném čase probíhá porovnání aktuálního snímku s prvním pořízeným snímkem na místo referenční fotografie. Tím se předchází nestabilitě a dalším komplikacím (3).

Porovnávat historickou fotografii s fotografií aktuální je příliš komplikované a výsledek nemusí být uspokojivý. Proto je uživatel vyzván k označení šesti až osmi korespondenčních bodů na historické fotografii. Jsou tak rozpoznány společné body u referenčního a aktuálního snímku (4). Výsledek je tak plně závislý na společných bodech, které definoval uživatel.

Pro řešení posledního uvedeného problému se dopočítá optický střed historické fotografie (5). Optický střed se vypočte na základě metody MSAC a práce s úběžníky (viz sekce 3.2). Jako ohnisková vzdálenost se použije hodnota aktuálního fotoaparátu, což přináší částečné omezení refo-
tografického nástroje.

Analýza a řešení těchto problémů velkou měrou přispěly k vytvoření konečného návrhu nefotografického algoritmu. Specifikace byly upraveny tak, aby bylo možné algoritmus realizovat do zcela funkční podoby. Podrobnější řešení těchto problémů je uvedeno v kapitolách níže zabývajících se návrhem aplikace.

Kapitola 3

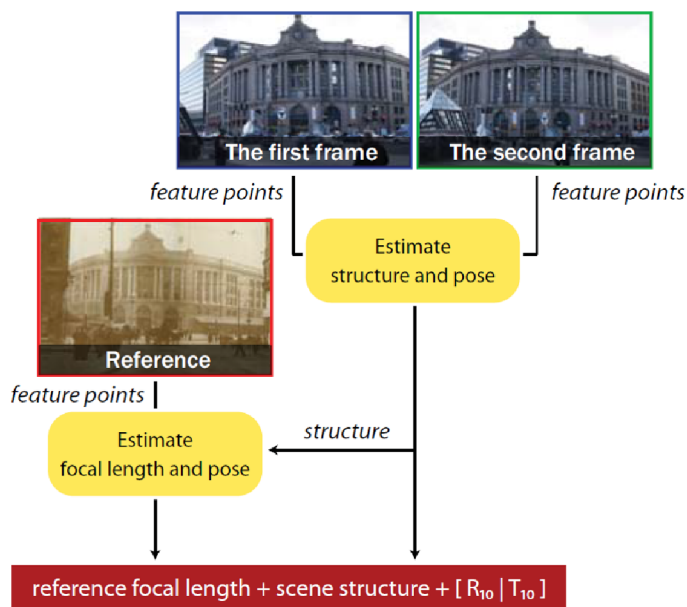
Kalibrace neznámého fotoaparátu

Proces kalibrace je obecně známý pojem v oblasti fotografování a počítačového vidění. Jedná se o postup, jehož výsledkem jsou kalibrační parametry: ohnisková vzdálenost (f_x , f_y) a optický střed (c_x , c_y). Kalibrační matici (K) si lze tak představit následovně:

$$K = \begin{pmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix}$$

Pro potřeby refotografického algoritmu byl proveden výpočet optického středu. Ohnisková vzdálenost byla ponechána na hodnotě ohniskové vzdálenosti dle aktuálního zařízení.

K používání refotografického nástroje by měl mít fotograf k dispozici libovolný referenční snímek neznámého stáří a pořízení na neznámém fotoaparátu. Než algoritmus začne pracovat s referenčním snímkem, musí provést trojrozměrnou rekonstrukci aktuálního objektu, který představuje současný stav na historické fotografii. Návrh tak lze rozdělit na 3D rekonstrukci a výpočet neznámých kalibračních údajů (viz obr. 3.1).



Obrázek 3.1: Kalibrace [2]

Od uživatele se očekává alespoň zdánlivá znalost pozice objektu, který se nachází na referenční fotografii. V rámci správného fungování aplikace nelze porovnávat odlišné místa.

První krok výpočtu začíná pořízením dvou snímků místa, které odhadem uživatel určí jako referenční místo. Požadavky na pořízené snímky jsou následující: musí být pořízeny ze stejného místa a rozličného úhlu pohledu (odhadem o 20 stupňů) [2]. Výsledkem celého výpočtu je pozice historického fotoaparátu vzhledem k prvnímu pořízenému snímku. Postup je zde uveden jako návrh jednotlivých částí nástroje a může se lišit po stránce implementační.

3.1 3D rekonstrukce

K vytvoření trojrozměrné rekonstrukce je zapotřebí vycházet alespoň ze dvou a více snímků. Jak již bylo zmíněno pro potřeby refotografického algoritmu postačí dva snímky stejného objektu z mírně rozličného úhlu. Po spuštění aplikace jsou tedy pořízeny dva snímky s širokou základní linií (*baseline*). Ta zlepšuje přesnost a stabilitu 3D rekonstrukce.

Předpokládá se, že současná kamera je kalibrována. Kalibrace je důležitým bodem procesu vzhledem k tomu, že kalibrační parametry mají velký vliv na správnou rekonstrukci scény. Na uživatelem pořízených snímcích se detekují a popíší konstantní body pomocí funkce SURF [10]. Funkce SURF nalezne 1000 konstantních neboli neměnných bodů (*Feature Points*). Pro vyhledání nejbližšího souseda je použita metoda Brutal Force matching. Na výsledných bodech je proveden ratio test. Ratio test na základě vzdálenosti vyřadí body, které určil Brutal Force matching. Následuje test symetrie a RANSAC test [7]. RANSAC je iterační metoda pro odhad vhodných bodů (*inlier points*). Výstupem celého procesu je fundamentální matice.

Z fundamentální matice je třeba vyjádřit matice esenciální, jejímž rozkladem se vypočítá matice translace a rotace. Esenciální matice (E) se vypočítá vynásobením fundamentální matice (F) s kalibrační kamerou (K) a s transponovanou kalibrační kamerou (K^{-1}).

$$E = K^{-1} * F * K \quad (3.1)$$

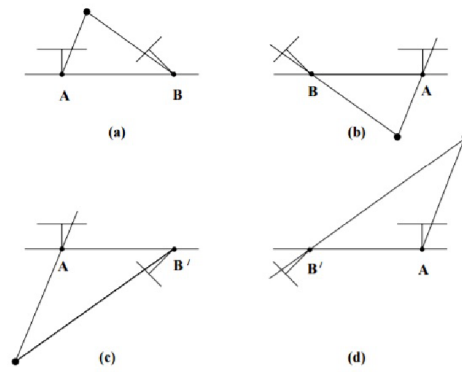
Pro potřeby triangulace je nutné provést rozklad esenciální matice. Výpočtem se získají dvě matice pro rotaci a pro dvě translaci. Prvním krokem výpočtu je rozklad esenciální matice pomocí funkce SVD (Singular Value Decomposition). Funkce SVD zpracuje výsledek do argumentů u , v a w . Výpočet první rotace (R1) a druhé rotace (R2) je uveden níže:

$$R1 = u * \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} * v \quad R2 = u * \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} * v \quad (3.2)$$

Obdobně jako u výpočtu rotační matice probíhá výpočet translace. Řešení reprezentuje matice translace ($t1$) a ($t2$). Translace se vypočítá jako třetí sloupec proměnné u vypočítané funkcí SVD.

$$t1 = u.col(2) \quad t2 = -u.col(2) \quad (3.3)$$

Jelikož výsledkem rozkladu jsou dvě matice rotace a dvě matice translace, mohou nastat čtyři možnosti pohledu obou fotoaparátů. Funkce pro výpočet triangulace se tedy provede čtyřikrát. Směr pohledu fotoaparátů se promítne do výsledných 3D bodů, tudíž správné postavení určí nejvyšší počet kladných čísel v souřadnici z [8] (viz obr. 3.2).

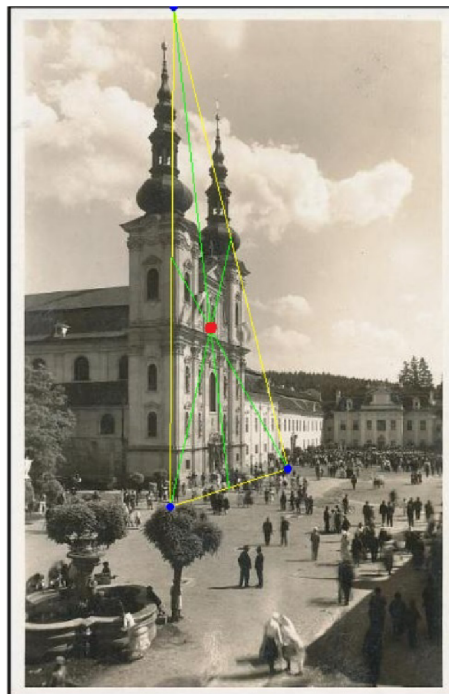


Obrázek 3.2: Různé pohledy kamer [8]

3.2 Optický střed

Následující část popisuje práci s historickou fotografií. Pro úplnost se požaduje referenční fotografie, malba nebo kresba z libovolného období, na nichž je zachycena budova či jiná stavba. Na referenční fotografii nejsou kladeny žádné nároky nebo podmínky.

Zpracované 3D body se promítají do druhého pohledu a zobrazí se uživateli v referenčním snímku. Uživatel je vyzván k označení šesti až osmi bodů. Tímto způsobem je vyřešen největší problém zahrnující proces refotografie. Lokalizovat stejné body na historické a aktuální fotografii je poměrně velký problém v oboru počítačového vidění. Stejně jako správná 3D rekonstrukce je i označení bodů uživatelem prvním předpokladem ke správnému výpočtu refotografického nástroje.



Obrázek 3.3: Výpočet optického středu

Optický střed je jeden z parametrů pro odhad pozice historického fotoaparátu pomocí PnP metody [9]. Jeho výpočet probíhá následovně. Pomocí MSAC metody [7] jsou nalezeny tři úběžníky (*vanish points*), které tvoří pomyslné vrcholy trojúhelníku. Z každého vrcholu je vedena těžnice na protější stranu trojúhelníku a hledaný optický střed leží v průsečíku těchto těžnic.

Pro názornost je celý proces zobrazen na historické fotografii (viz obr. 3.3). Modré body představují nalezené úběžníky (*vanish point*). Při jejich spojení vzniká trojúhelník znázorněný žlutou barvou. Těžnice vedené z vrcholů trojúhelníku jsou vykresleny zelenou barvou. Červený bod na průsečíku všech tří těžnic je výsledný optický střed.

Vnitřní parametry historického fotoaparátu tvoří tedy vypočítaný optický střed a ohnisková vzdálenost aktuálního fotoaparátu.

3.3 Pozice historického fotoaparátu

Smyslem první části je především odhadnout pozici historického fotoaparátu. Na zkruslení se v rámci refotografického algoritmu nebere zřetel. Z předchozího výpočtu jsou známy 3D body a k nim uživatelem definované 2D body. Na základě těchto hodnot se vypočítá projekční matice určující odhad pozice fotoaparátu. Tato matice zahrnuje matici rotace a translace, které jsou pro určení přibližné pozice historického fotoaparátu dostačující. Stejný postup zahrnuje výpočet aktuální pozice fotoaparátu.

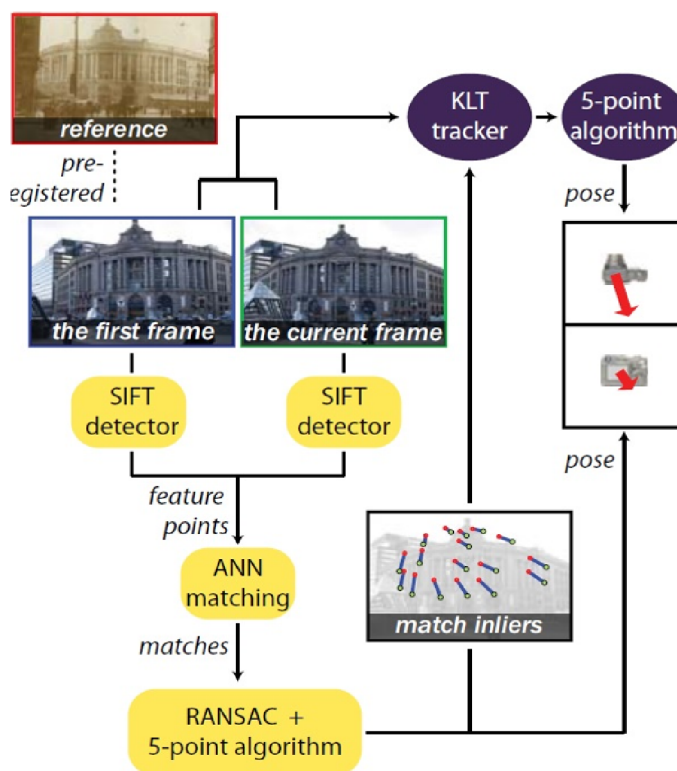
Hledaná pozice historického fotoaparátu je odhadnuta algoritmem PnP [9]. PnP (*Perspective-Point*) metoda řeší problém s odhadem pozice kamery, který byl uveden v kapitole o problémech (viz kapitola 2.3). PnP metoda zahrnuje několik přístupů. Vhodně zvolený přístup závisí na počtu korespondujících bodů a přesnosti jejich určení. Pro korespondenci tří bodů, což je minimální počet, je určena metoda P3P. Pro větší počet bodů je určen přístup EPnP, který počítá s body $n > 3$, kde n je počet korespondenčních bodů [?]. Tyto body se k sobě vážou bez větších odchylek a nepočítá se s odlehlými body. Pro odhad pozice fotoaparátu, u jejichž korespondencí může dojít k odchylkám a extrémům, se používá iterační metoda RANSAC [7]. Odhad pozice kamery určí na základě určených vhodných bodů (*inlier points*), zatímco chybné body označí jako nevhodné body (*outliers points*).

Vzhledem k tomu že PnP metoda je náchylná k chybám, byla použita ve spojení s iterační metodou RANSAC. Jako optimalizační metoda je zvolena Levenberg-Marquardt. Jedná se o metodu nejmenších čtverců určená pro nelineární rovnice. Levenberg-Marquardt minimalizuje reprojekční chybu. Pro následnou navigaci uživatele k nalezené pozici historického fotoaparátu je nutné v reálném čase aktualizovat pozici fotoaparátu (viz sekce 4).

Kapitola 4

Navigace uživatele v reálném čase

Refotografický nástroj poskytuje uživateli v reálném čase navigaci směrem k referenční pozici. K tomu se počítá s relativní představou mezi aktuální a historickou pozicí fotoaparátu. Samotná navigace probíhá pro uživatele v podobě šipek ukazující směr k cílové pozici. Až po nejlepší shodě mezi aktuální a cílovou pozicí se uživateli zobrazí aktuální pohled historické předlohy. K dosažení výkonu v reálném čase se bude prokládat hrubý odhad s odhadem lehkých. Schéma procesu navigace v reálném čase je znázorněno na obrázku níže (viz obr. 4.1).



Obrázek 4.1: Navigace v reálném čase [2]

Na vyobrazeném schématu navigace (viz obr. 4.1) si lze povšimnout, že porovnání probíhá mezi prvním pořízeným snímkem a aktuálním snímkem. S referenčním snímkem se v druhé části refotografického výpočtu nepracuje. Dalším bodem jsou dva přístupy k výsledné navigaci. První z nich představuje robustní odhad, který nalezne cílové body (*feature points*) a pomocí RANSAC

metody odhadne pozici fotoaparátu [7]. Výsledek (*match inliers*) se předá KL trackeru [11], který reprezentuje lehký odhad a zároveň provede vyhodnocení směru pohybu demonstrováno šipkou. Na základě schématu lze usoudit, že navigace je složená ze dvou částí, tj. z lehkého a robustního odhadu.

Hlavním kritériem algoritmu je především zajištění dostatečné výpočetní rychlosti. Z předchozího výpočtu jsou známy 3D body prvního pořízeného snímku. Ty jsou porovnávány s klíčovými body aktuálního snímku. Pro každý jeden snímek probíhá jedno vyhodnocení refotografického nástroje. Z algoritmů pro vyhledávání nejbližšího souseda byl vyhodnocen jako nejlepší Flann matching pro svou výpočetní rychlost [2]. Po aplikování PnP metody je na výslednou matici rotace a translace aplikován Kalmanův filtr [6]. V rámci lehkého odhadu je použita metoda Lucas-Kanade [4].

4.1 Robustní odhad

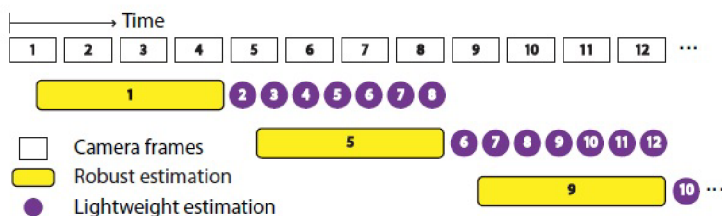
Robustní odhad lze charakterizovat jako výpočet pozice aktuálního fotoaparátu s přesným výsledkem, ovšem s výpočetně pomalým algoritmem. Navigace je plně závislá na robustním odhadu, který určuje konečný výsledek refotografického nástroje.

Výpočet navigace začíná robustním odhadem prvního snímku. Postup algoritmu zahrnuje detekci a popsání bodů metodou SIFT pro aktuální snímek podobně jak bylo uvedeno v kapitole 3D rekonstrukce (viz sekce 3.1). Body aktuálního a prvního pořízeného snímku jsou použity pro nalezení nejbližšího souseda metodou Flann matching. Tato metoda byla použita pro svou výpočetní rychlost [2], čímž se liší například od již použité metody Brute-Force matching. Podobně jako u 3D rekonstrukce jsou i zde provedeny ratio testy a test symetrie.

Po detekci nejbližších sousedů se aplikuje PnP metoda pro odhad pozice aktuálního fotoaparátu. Výpočet aktuální pozice fotoaparátu probíhá stejně, jak bylo uvedeno v předchozí kapitole (viz kapitola 3.3). Výpočet pokračuje výpočtem relativní matice vzhledem k aktuálnímu a historickému pozici fotoaparátu (viz kapitola 4.3). Robustní odhad končí aktualizací Kalmanova filtru.

4.2 Lehký odhad

Robustní odhad je pomalý a pro uživatele by aplikace nepracovala plynule. Z toho důvodu je pomalý robustní odhad prokládán rychlým lehkým odhadem. Robustní odhad předá jako argument shodné body (*match inliers*) lehkému odhadu. Součástí lehkého odhadu není určení přesného směru, ale vypočítání následujícího směru pohybu uživatele. Lehký odhad vychází z prvního výpočtu robustního odhadu, a dále výpočet pokračuje střídáním schodných bodů (*match inliers*) nalezených robustním a lehkým odhadem. Lehký odhad požaduje znalost předchozího snímku, proto se provádí lehký odhad na všech snímcích, krom snímku prvního (viz obr. 4.2).



Obrázek 4.2: Prokládání odhadů [2]

Lehký odhad pracuje s metodou Lucas-Kanade [11], která odhadne směr pohybu. Metoda Lucas-Kanade dokáže odhadnout nové body aktuálního snímku a tím určit další pohyb uživatele. Na základě těchto bodů je proveden odhad pozice fotoaparátu pomocí metody PnP tak, jak bylo uvedeno v kapitole o robustním odhadu (viz sekce 4.1).

Lucas-Kanade je diferenciální metoda pro odhad optického toku. Byla vyvinuta pány Bruce D. Lucas a Takeo Kanade, po nichž nese jméno. Metoda počítá s konstantním tokem a řeší základní optické průtokové rovnice pro všechny pixely v tomto okolí. Ke správnému výpočtu používá metodu nejmenších čtverců [11].

4.3 Relativní matice

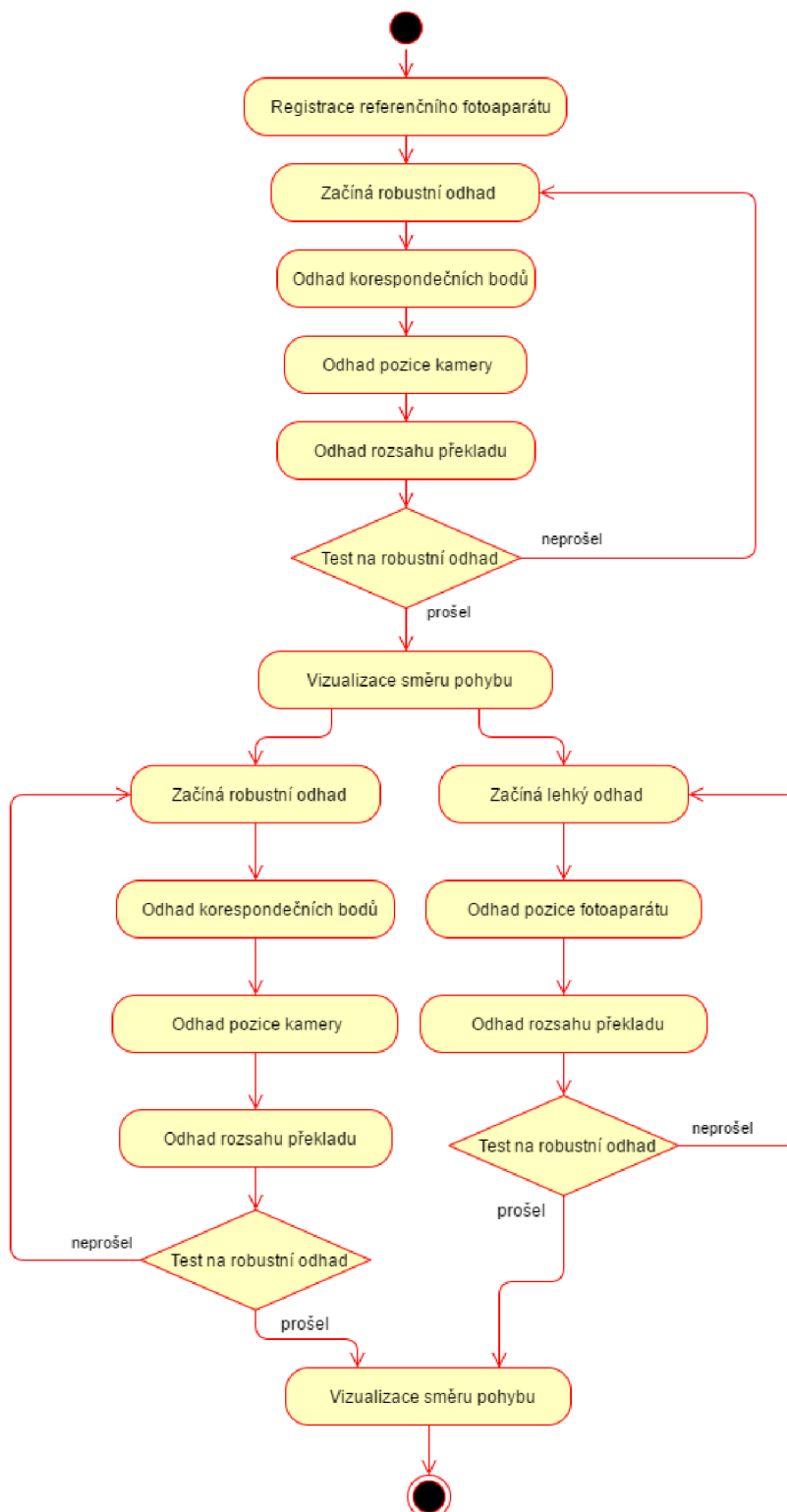
Algoritmus provedl odhad pozice historického a aktuálního fotoaparátu. Následuje odhad relativního rozdílu mezi pozicemi fotoaparátů. Na základě rovnice uvedené níže se provede výpočet relativní matice, která představuje rozdíl mezi pozicemi aktuálního a historického fotoaparátu.

Relativní matici lze demonstrovat jako projekční matici s maticemi rotace a translace. Výpočet relativní matice (rev_T) se provádí rozdílem mezi pozicí historického (ref_T) a aktuálního fotoaparátu ($current_T$):

$$rev_T = current_T^{-1} * ref_T$$

Pro potřeby algoritmu je podstatný hlavně translační vektor, a to především pozice X a Y. Uživateli je ve výsledku zobrazena šipka, která ho navede na cílovou pozici fotoaparátu. Po úspěšném nalezení této pozice, je uživatel schopen pořídit refotografii cílového místa. Práce refotografického algoritmu tím končí.

Pro lepší názornost je níže uvedeno schéma výpočtu navigace uživatele v reálném čase tak, jak probíhá od prvního snímku po vizualizaci (viz obr. 4.3). Výsledky robustního a lehkého odhadu jsou podrobeny testy, kde se kontroluje, zda byl detekován alespoň minimální počet shodných bodů (*matching points*). Při malém počtu bodů by mohl být výsledek nepřesný, zkreslený a uživatel by byl tak naveden na úplně jiný směr, než je jeho cílová pozice fotoaparátu.



Obrázek 4.3: Stavový diagram navigace v reálném čase

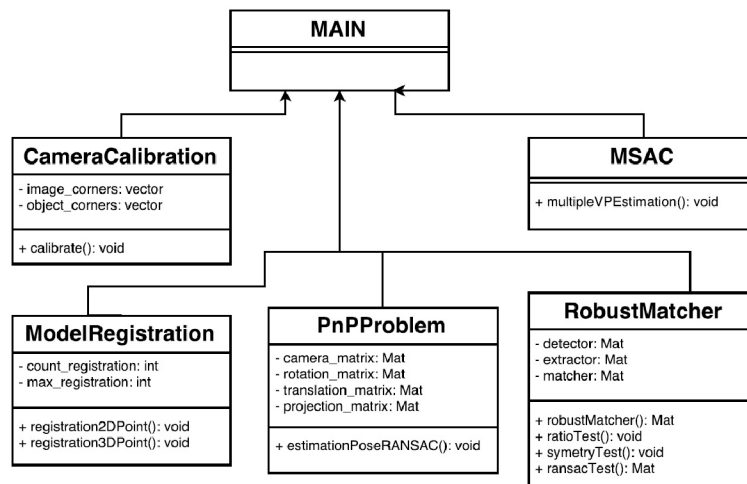
Kapitola 5

Implementace

Následující kapitoly hloupěji nahlédnou do algoritmů zajišťující funkčnost procesu refotografie. Implementace lze rozdělit na tři části. První část představuje 3D rekonstrukci pořízených snímků. Následuje výpočet pozice historického fotoaparátu. Poslední část zahrnuje výpočet aktuální pozice fotoaparátu, která se porovnává s pozicí historického snímku. Implementace je rozvržena dle vypracovaného návrhu uvedeného v sekci Kalibrace neznámého fotoaparátu (viz sekce 3) a Navigace uživatele v reálném čase (viz sekce 4). Logické celky mají jasně vymezené vstupy a výstupy.

Projekt byl vypracován v jazyce C++ s využitím objektově orientovanými principy programování. Výpočetní rychlost jazyka C++ je dostatečně rychlá pro navigaci uživatele v reálném čase. Pro projekt byly klíčové krom standardních knihoven především knihovny OpenCV 3. 1 [?] nabízející velké množství funkcí pro práci v oboru počítačového vidění. Uživatelské rozhraní je implementováno v jazyce Java, který je určený pro operační systém Android. Implementace se řídí metodikou zvanou SOLID [3], jejichž dodržení konvence zajišťuje efektivnější a přehlednější zdrojový kód.

Model tříd lze popsat jako hlavní třídu *Main* a ostatní podtřídy, které se zabývají konkrétními problémy. Diagram tříd je schématicky znázorněn na obrázku níže (viz obr. 5.1).



Obrázek 5.1: Diagram tříd

5.1 Odhad snímků

Na vstupu této sekce jsou požadovány aktuálně pořízené dva snímky objektu ze stejného místa a rozličného úhlu. Výstupem je množina 3D bodů, které jsou společné pro oba snímky. Z hlediska implementace zde byly uvedeny postupy pro standardní výpočet 3D bodů neboli 3D rekonstrukce na základě triangulace. Schématický postup je uveden následovně:

1. Vstup: Dva uživatelem pořízené snímky
2. Kalibrace
3. feature points + matching
4. Fundamentální a esenciální matice
5. Triangulace
6. Výstup: Společné 3D body snímků

Předpokládá se, že uživatel zná údaje o optickém středu a ohniskové vzdálenosti fotoaparátu. Pro implementační potřeby projektu a zjištění těchto parametrů na testovacím zařízení bylo nutné vytvořit kalibrační funkci. Ta standardně vychází ze snímků šachovnicové plochy, na které jsou detekovány vnitřní rohy. Pro testovací zařízení byla použita šachovnicová hrací plocha o velikosti 7x9 na celkem 28 snímcích. Správné hodnoty získané kalibrací se projeví na 3D rekonstrukci.

Pro lepší názornost procesu je uveden alespoň pseudokód pro výpočet triangulace. Jedná se o kombinaci několika postupů, které byly upraveny pro potřeby projektu. Následující pseudokód je uveden schématicky, nejedná se o kompletní kód a jsou zde uvedeny jen podstatné části. Jelikož standardní postupy jsou obecně známé a používané, jsou naznačeny jen v podobě volání funkce.

Algorithm 1: Rozklad esenciální matice

```
1 function robustMatcher (first frame, second frame);  
   Output: matches points - shodné body  
2 for iterace_matches_points do  
3   | x = keypoints first frame[queryIdx].x;  
4   | y = keypoints first frame[queryIdx].y;  
5   | detectionpoints first frame.push.Points(x, y);  
6   | x = keypoints second frame[trainIdx].x;  
7   | y = keypoints second frame[trainIdx].y;  
8   | detectionpoints second frame.push.Points(x, y)  
9 end  
10 function findFundamentalMat (detectionpoints, detectionpoints);  
   Input : points of first frame, points of second frame  
   Output: fundamental matrix  
11 Essencialmatrix = K.t() * F * K;  
   Input : K: camera matrix, F: fundamental matrix  
12 W = 0, -1, 0, 1, 0, 0, 0, 0, 1;  
13 z = 0, -1, 0, 1, 0, 0, 0, 0, 0;  
14 Rotationvector = W.t() * v.t() * u;  
15 Transformationvector = v * z * v.t();
```

Výše uvedený algoritmus představuje výpočet rotačního a transformačního vektoru na základě nalezené fundamentální matice a shodných bodů. Funkce *robustMatcher()* zahrnuje postup od vyhledávání a popisu klíčových bodů po nalezení shody mezi prvním a druhým snímkem. Shodné body jsou v iteracích rozděleny mezi dvě matice, které reprezentují shodnou detekci pro oba snímky. Nalezené body slouží jako vstup pro funkci z knihovny OpenCV *findFundamentalMat*, a tudíž k vyhledání fundamentální matice. Rozkladem fundamentální matice se vypočte esenciální matice dle vzorců k tomu určených. Výstupem je matice o velikosti 3x3 pro rotaci a 3x1 pro transformaci. Pro dodržení velikosti matic je určena funkce *Rodrigues()* z knihovny OpenCV.

Algorithm 2: Výpočet 3D bodů

Input : K: camera matrix, T: transformation matrix, R: rotation matrix

```

1 for  $i = 0 \rightarrow 4$  do
2   if  $i = 0$  then
3     |  $rotationVector = R1;$ 
4     |  $translationVector = t;$ 
5   end
6   if  $i = 1$  then
7     |  $rotationVector = R2;$ 
8     |  $translationVector = t;$ 
9   end
10  if  $i = 2$  then
11   |  $rotationVector = R1;$ 
12   |  $translationVector = -t;$ 
13  end
14  if  $i = 3$  then
15   |  $rotationVector = R2;$ 
16   |  $translationVector = -t;$ 
17  end
18  Rotation-transforamtion matrix =  $hconcat(rotationvector, translationvector);$ 
19  New camera matrix =  $K * [R|T];$ 
20  function triangulatePoints (Cameramatrix, detectionpoints);
    Output: 3D points of images
21 end

```

Triangulace požaduje nalezené body z předchozího algoritmu a projekční matici. Projekční matice o velikosti 3x4 se skládá z rotační a transformační matice vypočítaných z rozkladu esenciální matice. Celý proces proběhne čtyřikrát, tudíž výsledkem jsou čtyři množiny 3D bodů. Je to způsobeno tím, že dva fotoaparáty mohou být otočeny po z-souřadnici směrem k pohledu nebo naopak směrem od pohledu. To se projeví tak, že z-souřadnice je v záporných nebo kladných číslech. Po nalezení množiny 3D bodů se proto zkontroluje počet záporných hodnot u z-souřadnice. U výsledku, kde počet takových hodnot je nejmenší, se jedná o požadovanou množinu bodů.

Množina bodů odpovídá trojrozměrné rekonstrukci scény pořízené na fotografiích. To lze v rámci testování ověřit v nástroji MeshLab. Seznam všech 3D bodů je nainportován do programu v textovém souboru. Výsledkem je vizualizace bodů v trojrozměrném prostoru.

5.2 Pozice referenčního snímku

Zatímco cílem první části algoritmu je rekonstrukce trojrozměrné scény, neboli výpočet množiny 3D bodů, cílem druhé části je nalezení projekční matice vzhledem k prvnímu snímku. Uživatel má k dispozici pořízený snímek s vyznačenými body (3D body převedené do 2D podoby) a referenční snímek, na kterém stejné body uživatel vyznačí. Algoritmus zaznamenává uživatelem definované body a k nim vázané 3D body, popřípadě nabídne uživateli vyznačit další vhodný bod. Správnost výpočtu je závislý na uživatelově určené poloze bodů, proto má uživatel možnost svou volbu změnit nebo jinak upravit.

1. Vstup: historický snímek, 3D body pořízených snímků
2. Registrace korespondenčních bodů
3. Odhad optického středu
4. Odhad pozice historického fotoaparátu pomocí PnP metody
5. Výstup: pozice historického fotoaparátu

Schématicky je znázorněn jen algoritmus pro registraci bodu na referenčním snímku. Ostatní body jsou otázkou volání již použité funkce v knihovně OpenCV nebo obecně známé metodiky, tudíž je není potřeba blíže rozebírat.

Algorithm 3: Registrace korespondenčních bodů

```
1 while true do
2   3DPoint = listOf3dPoints[index];
3   if !endregistration then
4     draw2DPointOnFirstFrame(registration2DPoints);
5     if lastCountRegistrationPoint != currentlyCountRegistrationPoint then
6       register3DPoint(point3f);
7       lastCountRegistrationPoint = currentlyCountRegistrationPoint;
8     end
9   else
10    break;
11  end
12 end
```

Algoritmus proběhne jen jednou za celý výpočet refotografie. Pozice historického fotoaparátu se porovnává s pozicí aktuálního snímku vypočítanou v nadcházejícím algoritmu. Na závěr se pomocí PnP metody vypočítá pozice historického fotoaparátu. Na vstupu jsou požadovány 3D body vypočítané v předchozím algoritmu a uživatelem definované 2D body na referenčním snímku.

Rotační a transformační matice vypočítaná z PnP metody je dále upravená podle postupu uvedeného níže. Pro rotační matici je provedena transpozice, zatímco transformační matice je vypočtena podle vzorce $T = -R * T$.

Podoba matice určující pozici fotoaparátu je uvedena níže. Jedná se o matici 4x4 zahrnující vypočítanou upravenou matici rotace (uvedenou jako R) a transformace (uvedenou jako T).

Algorithm 4: Výpočet pozice historického fotoaparátu

```
1 function estimatePoseRANSAC (list3Dpoints, list2Dpoints);  
   Input : 3D bodů získaných pomocí triangulace, 2D body určené uživatelem  
   Output: rotační a transformační matice  
2  $T = -rotanMatice * transformanMatice$ ;  $Pozice = [R|T|0001]$ 
```

5.3 Výsledné porovnání

Nadcházející algoritmus na rozdíl od předchozích algoritmů pracuje v reálném čase, tudíž jeho hlavní požadavek je, aby byl co nejméně náročný na výpočet. Na vstupu jsou požadovány aktuální snímky, které mohou představovat pohled fotoaparátu. Podobně jako u prvního algoritmu následuje vyhledávání cílových bodů (*feature points*) a vhodných bodů (*inlier points*) s prvním pořízeným snímkem a aktuálním snímkem. Pozice aktuálního fotoaparátu je zjištěna pomocí PnP metody. Opět je uveden schématický postup algoritmu.

1. Vstup: Aktuální snímek
2. Feature points + matching
3. Odhad pozice aktuálního fotoaparátu pomocí PnP metody
4. Výstup: rozdíl pozice historického a aktuálního fotoaparátu

Schématicky je uveden pseudokód pro metodu *robustMatcher()*. Jejím cílem je především zjistit společné body u u prvního a aktuálního snímku.

Algorithm 5: Výpočet 3D bodů

```
1 function robustMatcher (first frame, currently frame);  
   Output: detection points of first image, detection points of currently image  
2 goodMeasurement = false;  
3 if goodMatches.size() > 0 then  
4   | estimatePoseRANSAC(listPoints3DModelMatch, listPoints2DSceneMatch);  
5   | result = PoseRefCamera – PoseCurrentlyCamera;  
6   | if inliersIdx.rows >= minInliersKalman then  
7   | | goodMeasurement = true;  
8   | | fillMeasurements(measurements, translationMeasured, rotationMeasured)  
9   | end  
10 end  
11 updateKalmanFilter(measurements, translationEstimated, rotationEstimated);  
12 return = goodMeasurement;
```

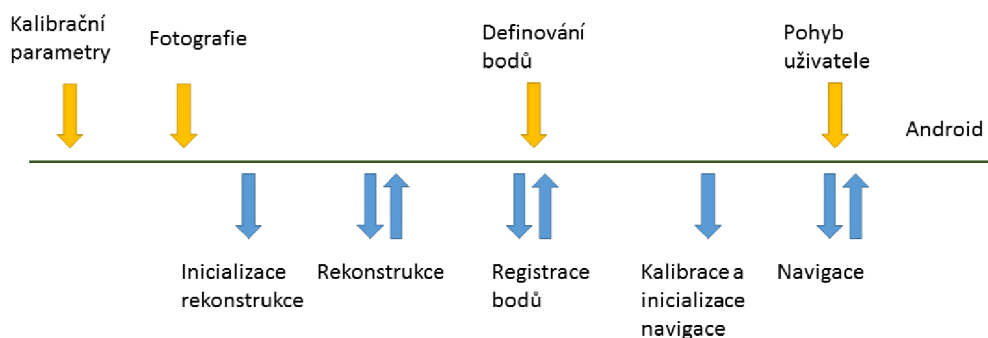
Algoritmus pokračuje výpočtem pozice aktuálního fotoaparátu dle postupu uvedeného v sekci, která se věnuje výpočtu pozice historického fotoaparátu (viz sekce 3.3). Jestliže je k dispozici matice určující pozici historického a aktuálního fotoaparátu, vypočítá se mezi nimi relativní matice určující jejich rozdíl. Níže je uveden vzorec pro výpočet relativní matice.

$$REV_T = CUR_T^{-1} * REF_T.$$

V uvedeném vzorci představuje CUR_T invertovanou matici určující pozici aktuálního fotoaparátu. REF_T představuje matici určující pozici referenčního fotoaparátu. Výsledkem je matice relativního rozdílu (REV_T) mezi oběma maticemi. Výsledná matice tedy poskytuje informace o pozici fotoaparátu v podobě dvou matic, kdy matice 3×3 představuje rotační matici a matice 3×1 představuje matici translace. Pro účely refotografie se s maticí rotace neuvažuje a počítá se jen s maticí translace. Matice translace zahrnuje pozice X, Y, Z. Pozice Z se nepřipouští, jelikož se počítá s rozkladem matic a uživatel by měl stát na stejném místě. Pozice X znázorňuje pohyb doleva a nebo naopak doprava. Stejně tak pozice Y znázorňuje pohyb nahoru nebo dolů. Na základě těchto údajů program vykreslí šipku jakým směrem je třeba posunout fotoaparát.

5.4 Propojení se systémem android

Refotografický algoritmus je určen pro mobilní aplikaci se systémem android. Pro výpočetní výkon jazyka C++ byl zvolen postup, který oddělil grafické rozhraní od samotného výpočtu refotografie. Oba systémy musely být nějakým způsobem propojeny a bylo zapotřebí zrealizovat komunikaci mezi nimi.



Obrázek 5.2: Schéma komunikace Android aplikace (modrá - komunikace přes nativní metody, žlutá - komunikace prostřednictvím uživatele)

Komunikace probíhá v rámci nativních metod, tj. metody, které mají svého reprezentanta jak v systému Android, tak v refotografickém algoritmu. Snahou bylo, aby komunikace mezi oběma přístupmi byla co nejmenší a objem přenášených dat byl co nejjednodušší (viz obr. 5.2). Pro potřeby projektu se přenáší jen číselné hodnoty, které jsou dále zpracovány v systému Android. Obě části jsou tudíž na sobě nezávislé a lze je implementovat odděleně. Mostem, který je spojuje je třída *OpenCVNative*, zahrnující zmíněné nativní metody.

Pro konfiguraci a propojení obou přístupů byl zvolen nástroj *Gradle 3.3* pro překlad celého projektu a *Android NDK r14* pro překlad nativní části projektu. Simulace probíhala na zařízení nVidia SHIELD s operačním systémem Android.

Kapitola 6

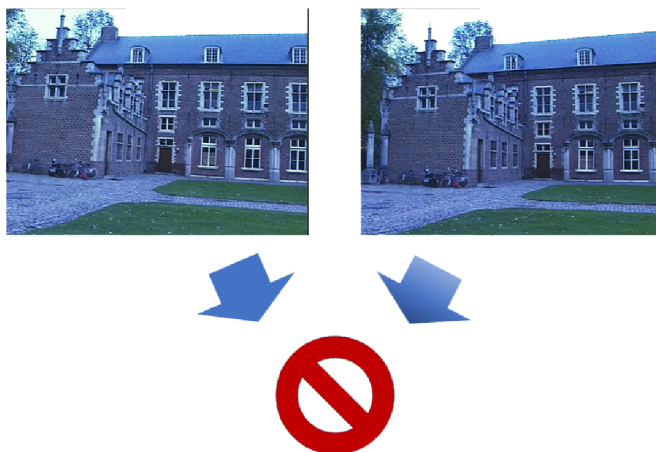
Testování

Testování lze rozdělit na tři fáze: ověření 3D rekonstrukce, test výpočtu optického středu, ověření pozice fotoaparátu a případně výpočet odchylky. Následující kapitoly se zaměří na způsob testování a vyhodnocení výsledků. Konečné výsledky dosažené refotografickým algoritmem jsou shrnuty v kapitole Dosažené výsledky (viz sekce 7).

Počáteční fáze testování probíhala na pořízených fotografiích mobilním zařízením nVidia SHIELD. Snímání fotoaparátu bylo simulováno videozáznamem pořízeným stejným zařízením. Pro testování bylo použito celkem 150 fotografií a 23 videozáznamů historických budov z okolí Brna. Historické fotografie zde byly zastoupeny 19 snímky.

6.1 3D rekonstrukce

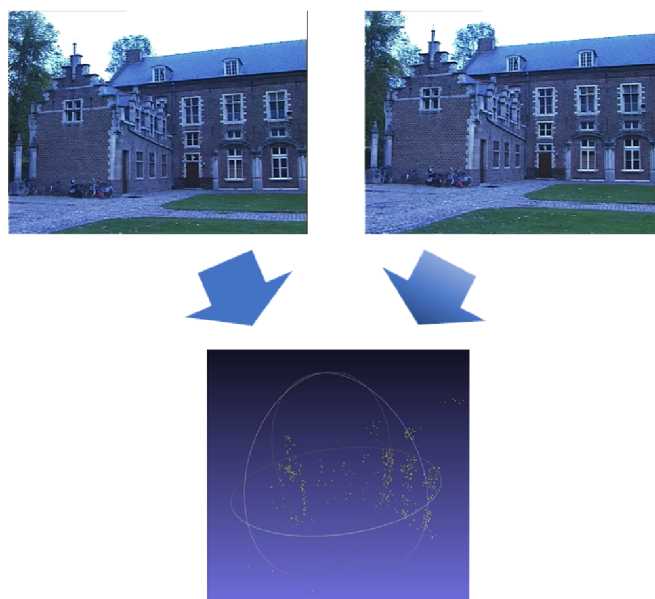
Výsledek 3D rekonstrukce lze snadno prezentovat na základě nalezených 3D bodů. Pro testovací účely byl použit program Meshlab, který dokáže zrekonstruovat body v prostorové rovině. Počet nalezených 3D bodů závisí na rozlišení pořízených snímcích a na způsobu zachycení daného objektu. Toto tvrzení si lze snadno ověřit na datové sadě Marca Pollefeys (<https://www.inf.ethz.ch/personal/marc.pollefeys/>).



Obrázek 6.1: Špatně provedená triangulace

Na první ukázce jsou dva snímky stejného objektu (viz obr. 6.1). První snímek je mírně otočen a snímáný objekt je na úplném okraji fotografie. Druhý snímek se o moc neliší, přesto

výsledek triangulace je neúspěšný. Shluk nalezených 3D bodů nezobrazuje strukturu objektu. Zde platí pravidlo snažit se zachytit objekt ve středu snímku bez otočení. Ke snímanému objektu je lepší přistupovat spíše blíže pro zachycení většího počtu klíčových bodů (*Feature Points*).



Obrázek 6.2: Dobře provedená triangulace

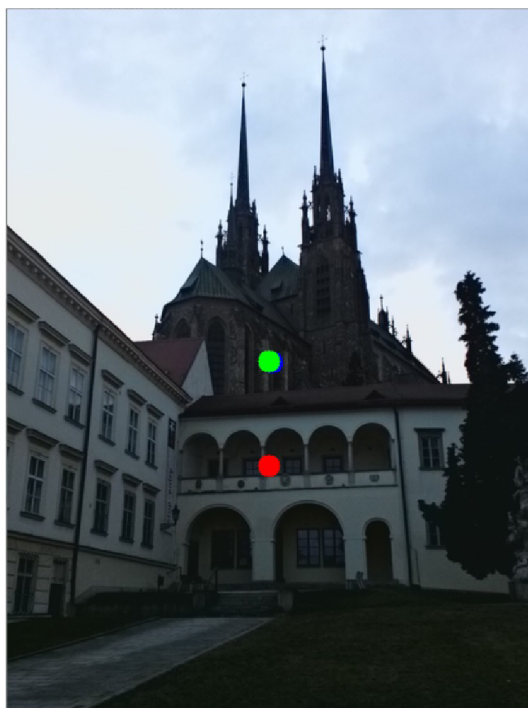
Druhá ukázka ukazuje správnou interpretaci trojrozměrné rekonstrukce. Opět se jedná o dva snímky stejného objektu. Snímky se od sebe liší drobným posunutím. Výsledkem triangulace je ovšem struktura objektu, především je zachycen vnější roh a viditelné stěny domu. Z toho lze usuzovat, že tyto snímky jsou vhodné pro vytvoření trojrozměrné rekonstrukce obrazu, což je první předpoklad ke správnému výpočtu refotografie.

Z vypořizovaných výsledků lze usoudit hned několik postřehů. Byla prokázána závislost rekonstrukce na kalibraci. Rekonstrukce u snímku pořízeného s použitím zoom nebo na jiném zařízení neodpovídala očekávané podobě. Důležitým faktorem bylo i například rozlišení zařízení a pohyb uživatele v prostoru při pořizování snímků. Pořízení takových snímků, jejichž rekonstrukce bude úspěšná může pro běžné uživatele představovat komplikace a snímaný objekt bude třeba vyfotografovat několikrát.

6.2 Optický střed

Vypočítaný optický střed lze demonstrovat na zařízení, u kterého jsou známy kalibrační parametry. Referenční snímek tedy nepředstavovala historická fotografie, ale snímky pořízené zařízením nVidia SHIELD a snímky z datové sady od Marca Pollefeyse (<http://www.cs.unc.edu/~marc/>). Vypočítaný optický střed lze porovnat jak s geometrickým středem, tak se středem vypočítaným pomocí kalibrační metody, tj. detekce kalibračních parametrů na základě šachovnice (viz obr. 6.3).

Výsledky testování lze demonstrovat následující tabulkou (viz tabulka 6.1). Pro porovnání jsou uvedeny tři sady testovacích vzorků. První reprezentuje snímek s malým rozlišením, druhý naopak s velkým rozlišením. Na závěr jsou uvedeny výsledky na snímku z datového balíčku Marca Pollefeyse (<http://www.cs.unc.edu/~marc/>). Pro každý reprezentativní vzorek jsou uvedeny dva výsledky, z nichž každý reprezentuje jinou orientaci.



Obrázek 6.3: Zobrazení optických středů (zelená - geometrický střed, modrá - střed vypočítaný kalibrací, červená - střed vypočítaný MSAC metodou)

Hodnoty	Geometrický střed	Pomocí kalibrace	Pomocí MSAC	Rozdíl
cx	240	268	274	2,24 %
cy	320	352	339	4,31 %
cx	320	323	319	1,24 %
cy	240	240	239	0,42 %
cx	972	1006	971	3,48 %
cy	1296	1329	1295	3,66 %
cx	1296	1286	1295	0,70 %
cy	972	1001	971	3,00 %
cx	288	288	287	0,35 %
cy	384	384	383	0,27 %
cx	384	384	383	0,27 %
cy	288	288	287	0,35 %

Tabulka 6.1: Výsledky testování optického středu

Rozdíl mezi hodnotami vypočítanými metodou MSAC a pomocí kalibrace z 20 testovacích vzorků byl nižší jak 5 % ze základu hodnoty vypočtené z kalibrace. Z toho lze konstatovat, že výsledek se příliš neliší od hodnot vypočítaných jinými metodami. Nejmenší rozdíl a tudíž nejlepší výsledek byl u vzorků z datového balíčku Marca Pollefeyse (<http://www.cs.unc.edu/~marc/>), kde rozdíl nebyl větší jak 1 %. Jelikož snímky s různě velkým rozlišením udávaly různé rozdíly, lze konstatovat, že rozlišení nemá vliv na výpočet optického středu. Výpočet metodou MSAC ovšem

ovlivňuje především scéna zaznamenaná na snímku a na tom, nakolik MSAC úspěšně detekuje úběžníky.

6.3 Pozice fotoaparátů

Cílem práce je odhadnout pozici historického fotoaparátu a poskytnout uživateli navigaci na tuto pozici. Z toho plyne, že odhad pozice je dalším důležitým aspektem výpočtu. Pozici fotoaparátu si lze představit jako projekční matici skládající se z matice rotace a translace, jak již bylo mnohokrát uvedeno v předchozích kapitolách. Pro testování jsou důležité body pozice x a y .

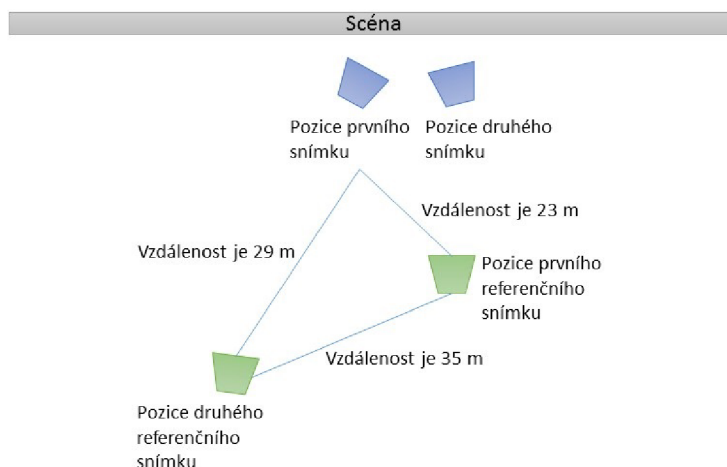
Na úvod je představen jednoduchý test, kdy první pořízený snímek a referenční snímek je totožný. Výsledek lze zobrazit v následující tabulce (viz tabulka 6.2).

Pozice	První snímek	Druhý snímek	Odhadovaná pozice referenčního snímku	Rozdíl v procentech
x	0,390	0,355	0,365	6,41 %
y	-0,324	-0,336	-0,322	0,62 %

Tabulka 6.2: Pozice, kdy první a referenční snímek je totožný

Souřadnice pozice prvního snímku a odhadu pozice referenčního snímku se liší zanedbatelně. Jedná se o ideální situaci, kdy lze označit přesně body na referenčním snímku, jelikož se jedná o totožný snímek jako ten první. Ovšem taková situace nikdy nenastane.

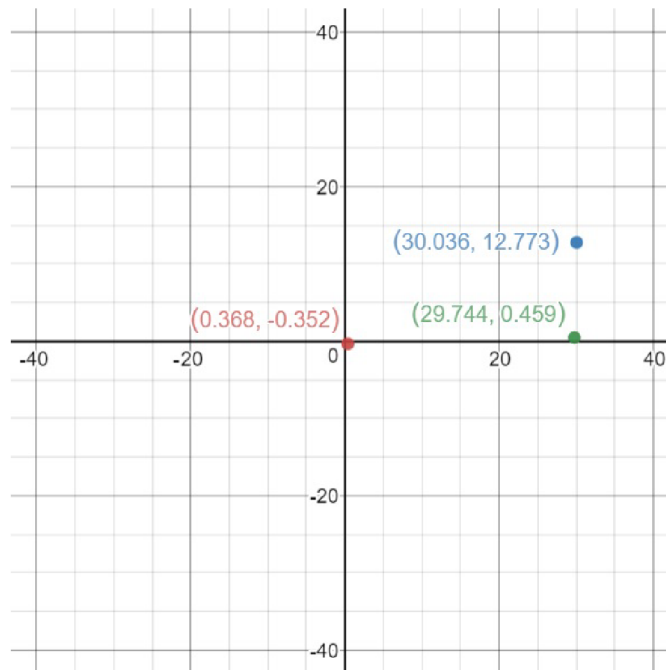
Odhadnutá poloha fotoaparátem pořízených snímků lze porovnat s GPS souřadnicemi. Snímky byly pořízeny se zapnutím ukládání polohy, tudíž lze snadno dopočítat skutečnou polohu v jednotkách metrů. Pro bližší vysvětlení je představena následující situace (viz obr 6.4).



Obrázek 6.4: Test pozice fotoaparátu

Jedná se o modelovou situaci, kdy modré lichoběžníky představují pozici prvního a druhého snímku pořízeného uživatelem. Zelené lichoběžníky představují pozici referenčních snímků. Jelikož je známa GPS souřadnice všech pořízených snímků, snadno se získají hodnoty reprezentující vzdálenost v jednotkách metrů.

Po spuštění refotografického algoritmu je situace namodelovaná následovně (viz obr. 6.5).



Obrázek 6.5: Test pozice fotoaparátu

V souřadnicovém systému jsou zakresleny důležité body. Červený bod znázorňuje uživatelem pořízený první snímek. Zelený bod reprezentuje odhad pozice prvního referenčního snímku a modrý bod představuje odhad druhého referenčního snímku. Po vypočítané vzdálenosti mezi některým z bodů se zjistí, kolik jednotek představuje jeden metr a naopak. Porovnáním výsledků a dopočítáním rozdílu se zjistí odchylka výpočtu.

Snímek A	Snímek B	Vzdálenost vypočítaná odhadem	Skutečná vzdálenost
První snímek	Odhad pozice prvního referenčního snímku	29,387 m	23 m
První snímek	Odhad pozice druhého referenčního snímku	32,454 m	29 m

Tabulka 6.3: Porovnání vzdáleností pozic fotoaparátu

Z výpočtu si lze snadno odvodit za pomoci trojčlenky, že chyba se pohybuje kolem tří až čtyř metrů (viz tabulka 6.3). Chyba u vzorových snímků, tj. u snímků, kde byla triangulace nejúspěšnější, se pohybuje do pěti metrů. U horších snímků, kde detekovaných bodů bylo málo nebo nebyly korektně definované, se chyba pohybuje řádově několik desítek metrů a více.

Kapitola 7

Dosažené výsledky

Celý vývoj aplikace sebou nesl spoustu úskalí a komplikací, s kterými se bylo třeba vypořádat. Výsledek lze definovat jako aplikaci určenou pro mobilní zařízení s operačním systémem Android, která umožní uživateli zjistit pozici uživatelem určeného refotografického snímku. Tato práce představuje možné řešení refotografického problému v oboru počítačového vidění a možný způsob, jak lze zpracovávat historické fotografie výpočetní technikou. Následující kapitoly popisují finální podobu projektu tak, jak byla odevzdána v rámci bakalářské práce. Je třeba konstatovat nakolik byly naplněny kladené cíle, dosažené úspěchy a výsledky experimentů a pozorování.

Pro názornost je třeba si připomenout postup s uživatelského pohledu a celý proces refotografie od pořízení fotografie po vyhodnocení závěrečné fotografie. Samotný proces je stručně komentovaný. Práce je zaměřená na historické stavby a budovy, z toho důvodu byly zvoleny významné památky nacházející se v centru města Brna a jeho okolí. Pro ukázkou byla zvolena černobílá fotografie Biskupského dvora (viz obr. 7.1).



Obrázek 7.1: Historický snímek Biskupského dvora

Historická fotografie byla použita ze serveru FOTOHISTORIE (<http://fotohistorie.cz/>). Fotograf zde zachytil Biskupský palác a v pozadí Katedrálu svatého Petra a Pavla. V takovém případě fotograf přijde na přibližné místo pořízené fotografie a pořídí dva snímky (viz obrázek 7.2). Pro zachycení velkého počtu klíčových bodů (*feature points*) se doporučuje přistoupit blíže k objektu.



Obrázek 7.2: Pořízené snímky Biskupského dvora

Na obrázku je vidět, že druhý snímek je mírně z odlišného úhlu. Tyto dva snímky poslouží k vytvoření trojrozměrné rekonstrukce objektu, v tomto případě Biskupského paláce. Hlavní koncentrace klíčových bodů (*feature points*) se nachází na věžích Baziliky svatého Petra a Pavla a u vnějšího rohu Biskupského dvora. Pro uživatele to znamená, že pro správný výpočet je třeba označit body s co největší plochou a určit body na místě, kde byla rekonstrukce nejméně úspěšná.



Obrázek 7.3: Ukázka navigace

Následně je fotograf navigován k cíli pomocí popsaného algoritmu (viz obr. 7.3). Navigace probíhá prostřednictvím šipek, které se zobrazí uživateli na displeji. Vizualizace může znázorňovat až osm různých směrů a jeden chybný stav v podobě žlutého křížku v černém poli. Kombinace černého čtverce s bílým rámem a žlutým znakem umožní dobrou čitelnost bez ohledu na snímaný obraz.

Z cílové pozice je pořízena fotografie Biskupského dvora, která může být považována za refotografii (viz obr. 7.4). Refotografický snímek může být následovně porovnán či dále zpracován na výpočetním zařízení.



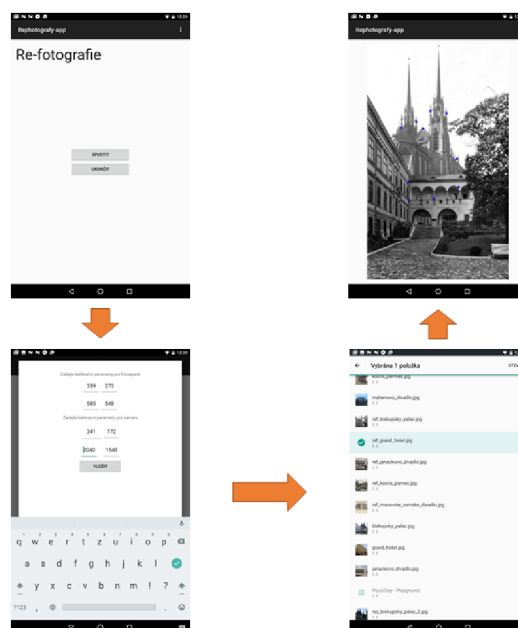
Obrázek 7.4: Refotografie Biskupského dvora

Výsledek závisí na fotografově určené počáteční místo. Nutností je správné zachycení místa na dvou mírně odlišných fotografiích. Výsledek se může lišit podle toho, nakolik přesně fotograf určí korespondenční body. Tyto body ovlivní výslednou fotografii a její určení. Faktorů ovlivňujících výsledek je ovšem mnohem více (viz kapitola 7.3).

7.1 Podoba aplikace

Aplikaci uživatel obsluhuje prostřednictvím grafického rozhraní, které mu umožňuje jednoduchou manipulaci s refotografickým algoritmem. Po spuštění aplikace je zapotřebí zadat kalibrační údaje. Zde jsou již předvyplněné hodnoty pro zařízení nVidia SHIELD, které slouží pro implementační a testovací část projektu. Uživatel má možnost pořídit dva snímky prostřednictvím fotoaparátu zabudovaného v zařízení nebo použít již vyfotografované snímky. Povinností je určit také referenční snímek. Uživatel pak definuje korespondenční body, na jejichž základě proběhne navigace.

Podoba aplikace využívá základní možnosti, které poskytují knihovny určené pro systém Android. Pro základní zpracování obrazu byla použita knihovna *OpenCV 3.1 SDK* a pro dialogové okno určené k výběru souboru, byla použita knihovna *aFileChooser*. Výsledná podoba aplikace je znázorněna na obrázku níže (viz obr. 7.5).



Obrázek 7.5: Uživatelské grafické rozhraní

7.2 Refotografie

Smyslem refotografie je porovnání stejného místa v různých časových obdobích, stejně jako smyslem refotografického nástroje je ulehčit uživateli pořizování těchto refotografií.

Ukázkové refotografie jsou pořizeny za pomoci mobilního zařízení SHIELD. Jedná se o známé historické budovy z okolí Brna. Jako referenční fotografie byly použity snímky z první poloviny 20. století, které jsou volně dostupné na serveru FOTOHISTORIE (<http://fotohistorie.cz/>). Výsledné refotografie byly vybrány jako nejvhodnější snímky z videí simulující pohyb uživatele s fotoaparátem. Výsledky cílové pozice za pomoci navigace v mobilním zařízení nebyly v rámci práce dostatečně uspokojivé pro demonstraci.



Obrázek 7.6: Grandhotel Brno (předloha)



Obrázek 7.7: Grandhotel Brno (refotografie)



Obrázek 7.8: Biskupský dvůr (předloha)



Obrázek 7.9: Biskupský dvůr (refotografie)



Obrázek 7.10: Janáčkovovo divadlo (předloha)



Obrázek 7.11: Janáčkovovo divadlo (refotografie)



Obrázek 7.12: Kašna Parnas a Divadlo Reduta (předloha)



Obrázek 7.13: Kašna Parnas a Divadlo Reduta (refotografie)



Obrázek 7.14: Mahenovo divadlo (předloha)



Obrázek 7.15: Mahenovo divadlo (refotografie)

7.3 Experimenty

Realizace projektu zahrnovala experimentování s různými metodami a přístupy a vyhodnocování jejich výsledků. Z pohledu realizace bylo důležité různé chování algoritmů poskytující knihovna OpenCV 3.1. Pozorování taktéž zahrnovalo sledování, jak se chová algoritmus na různých typech fotografiích a kresbách. Experimentování se částečně podílelo jak na implementaci, tak na testování refotografického nástroje.

Cílem experimentování je odhalit ideální situaci, kdy refotografický nástroj poskytuje nejlepší výsledky. Pro uživatele to představuje především návod, jak ovládat refotografický nástroj. Z experimentování lze vyvodit hned několik poznatků:

- Pozice uživatele by měla být bližší k objektu, než je cílová pozice
- Pro bezpečnější určení cílové pozice je lepší určit více bodů
- Definované body musí být určeny přesně
- Definované body musí být označeny na co největší ploše
- Lepších výsledků se dosáhne se snímky s vyšším rozlišením
- První dva snímky musí být pořízeny z přibližné pozice

Výsledky experimentů mnohdy potvrdily výsledek, který se očekával na základě chování refotografického algoritmu. Experimenty tak posloužily k ověření chování nebo zjištění nejvhodnějšího způsobu použití algoritmu.

Důležitou úlohou experimentů bylo také odhalit vhodné referenční snímky. Jednalo se především o historické snímky a kresby zachycující městské části nebo budovy. Výsledky lze shrnout následovně. Vhodné referenční snímky splňující následující vlastnosti.

- Snímek by měl obsahovat dobře čitelnou strukturu stavby
- Může být použita i kresba nebo malba, zachycuje-li skutečné proporce stavby
- Zachycený objekt nesmí být příliš pozměněn nebo zastavěn
- Okolí zachyceného objektu musí být přístupné pro pohyb uživatele

Výsledky experimentování lze shrnout do následujících vět. Určení refotografie závisí především na uživateli samotném. Jeho definování bodů a pořízení počátečních snímků na samém počátku procesu refotografie výrazně ovlivní výsledek. Pro budoucí vývoj bude důležité především

omezit tento vliv nebo počítat s mírnými odchylkami v určení bodů. Historické stavby zahrnující dekor jsou lépe identifikované a vhodnější pro definování bodů. U moderních staveb mohou nastat problémy v podobě identifikace bodů.

Kapitola 8

Závěr

Tato práce si klade za cíl usnadnit proces refotografie a nabídnout tak možnost porovnání míst v různém časovém období. Cílem bylo pomocí existujících metod počítačového vidění vytvořit plně funkční refotografický algoritmus, který umožní uživateli pro danou historickou fotografii pořídit co možná nejvíce podobný aktuální snímek.

Fotograf je v mnoha ohledech omezen při pořizování potřebných snímků. Původní místo pořízení historické fotografie nemusí být vlivem stavebních prací dostupné. Při pohledu na fotografický objekt se může objevit překážka. Častým jevem jsou vzrostlé stromy nebo nově vybudované stavby. S kresbami a malbami problémy narůstají. Aplikace vyžaduje interaktivitu ze strany fotografa pro nalezení odpovídající refotografie (uživatel musí manuálně zadat body korespondence). Implementovaná aplikace dospěje k požadovanému závěru jen za určitých podmínek: nástroj nepripouští zoom a na počátku procesu je nutné správně pořídit dva snímky dle instrukcí.

Samotný nástroj najde své využití jak u amatérských badatelů, přes historiky, až po výzkumníky mapující vývoj důležitých míst. Pátrání po minulosti je předmětem seriózních vědeckých článků, tak i obecních kronikářů. Refotografický nástroj je jedna z možností, jak lépe představit stavební vývoj a proměny životního prostředí v průběhu času.

Budoucí vývoj se může ubírat hned několika směry. První směr by mohl vést k výpočtu pozice fotoaparátu i v mnohem komplikovanějších případech. Opěrným bodem by byla jen malá část scény zachycená na historické fotografii a zároveň zachovalá do současnosti. Druhý směr by mohl představovat zpracování výsledné fotografie vzhledem k referenčnímu snímku. Aktuální snímek by mohl být převeden do historického snímku a naopak. Výsledné transformace by mohly sloužit jak k porovnávání, tak k vyhodnocování možných závěrů. Tímto tématem se zabývá článek Generalized As-Similar-As-Possible Warping with Applications in Digital Photography [5].

Literatura

- [1] Bundler: Structure from Motion (SfM) for Unordered Image Collections.
URL <http://www.cs.cornell.edu/~snavely/bundler/>
- [2] Bae, S.; Agarwala, A.; Durand, F.: Computational Rephotography. *ACM Trans. Graph.*, ročník 29, č. 3, Červenec 2010: s. 24:1–24:15, ISSN 0730-0301, doi:10.1145/1805964.1805968.
URL <http://doi.acm.org/10.1145/1805964.1805968>
- [3] BOB"), R. C. M. U.: Principles Of OOD. 1999.
URL <http://butunclebob.com>
- [4] Bruhn, A.; Weickert, J.; Schnörr, C.: Lucas/Kanade Meets Horn/Schunck: Combining Local and Global Optic Flow Methods. *International Journal of Computer Vision*, ročník 61, č. 3, 2005: s. 211–231, ISSN 1573-1405, doi:10.1023/B:VISI.0000045324.43199.43.
URL <http://dx.doi.org/10.1023/B:VISI.0000045324.43199.43>
- [5] Chen, R.; Gotsman, C.: Generalized As-Similar-As-Possible Warping with Applications in Digital Photography. *Computer Graphics Forum*, 2016, ISSN 1467-8659, doi:10.1111/cgf.12813.
- [6] Faragher, R.: Understanding the Basis of the Kalman Filter Via a Simple and Intuitive Derivation [Lecture Notes]. *IEEE Signal Processing Magazine*, ročník 29, č. 5, Sept 2012: s. 128–132, ISSN 1053-5888, doi:10.1109/MSP.2012.2203621.
- [7] Fischler, M. A.; Bolles, R. C.: Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Commun. ACM*, ročník 24, č. 6, Červen 1981: s. 381–395, ISSN 0001-0782, doi:10.1145/358669.358692.
URL <http://doi.acm.org/10.1145/358669.358692>
- [8] Hartley, R.; Zisserman, A.: *Multiple View Geometry in Computer Vision*. New York, NY, USA: Cambridge University Press, druhé vydání, 2003, ISBN 0521540518.
- [9] Lepetit, V.; F.Moreno-Noguer; P.Fua: EPnP: An Accurate O(n) Solution to the PnP Problem. *International Journal Computer Vision*, ročník 81, č. 2, 2009.
- [10] Lowe, D. G.: Distinctive Image Features from Scale-Invariant Keypoints. *Int. J. Comput. Vision*, ročník 60, č. 2, Listopad 2004: s. 91–110, ISSN 0920-5691, doi:10.1023/B:VISI.0000029664.99615.94.
URL <http://dx.doi.org/10.1023/B:VISI.0000029664.99615.94>
- [11] Lucas, B. D.; Kanade, T.: An Iterative Image Registration Technique with an Application to Stereo Vision. In *Proceedings of the 7th International Joint Conference on Artificial*

Intelligence - Volume 2, IJCAI'81, San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1981, s. 674–679.

URL <http://dl.acm.org/citation.cfm?id=1623264.1623280>