# Použité metody a algoritmy

## Registrace uživatelem zadaných bodů

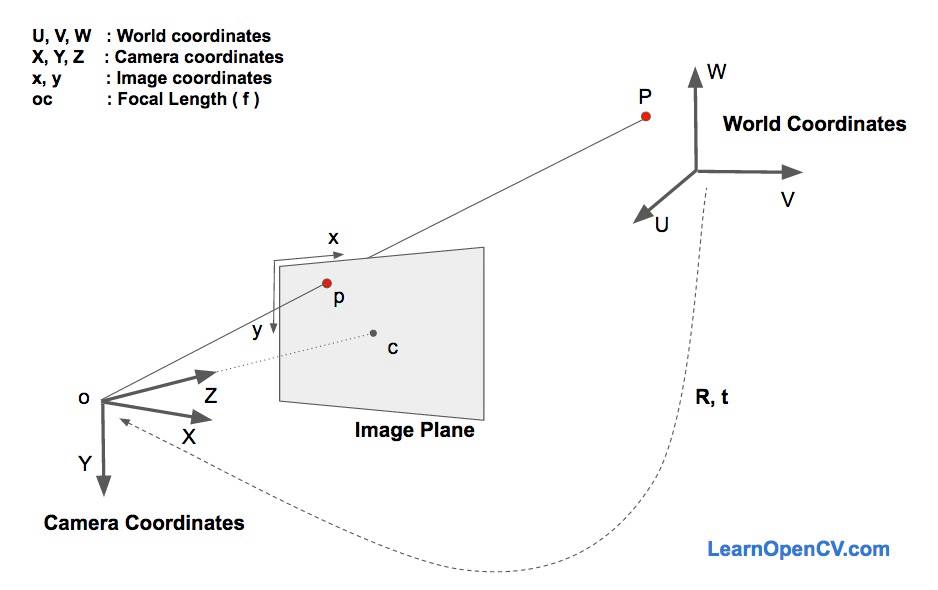
Dle kapitoly 3.1 uživatel bude vyzván, aby zadal 6 – 8 bodů. Algoritmus funguje tak načte obrázek, s kterým se bude pracovat. Inicializuje se RobustMatcher a jako počet registrovaných bodů se stanoví číslo 8. Při každém kliku se zaregistruje jeden bod. Body je třeba zadávat systematicky podle zrovna požadovaného bodu např. 0,0,0 bude nejdolnější bod.

## Odhad pozice

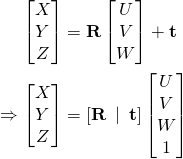
3D objekt má pouze dva pohyby vzhledem ke kameře.

Překlad: Pohyb fotoaparátu z jeho aktuálního umístění do nového umístění. Překlad má tři stupně a to X, Y, Z.

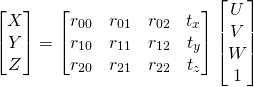
Otočení: Fotoaparát můžeme také otočit a to po ose X, Y, Z.

Takže odhad pózy z 3D objektu znamená najít 6 čísel – tři pro překlad a tři pro rotaci. Pro výpočet 3D póze budeme potřebovat 6-8 2D bodů a k ním umístění stejných 3D bodů a ohniskovou vzdálenost fotoaparátu, optický střed obrazu a radiální parametry zkreslení. Je potřeba kalibrovat fotoaparát. Ve

Klíčové jsou globální souřadnice (world coordinates), souřadnice obrázku (image coordinates) a souřadnice kamery (camera coordinates). 3D souřadnice jsou promítány do obrazové roviny s využitím přirozených parametrů kamery (ohnisková vzdálenost, optický střed). Na obrázku je vidět zobrazení *P* 3D bodu do *p* 2D bodu. Předpokládejme, že víme umístění (U, V, W) 3D bodu P v globálních souřadnicích s ohledem na souřadnice kamery, můžeme vypočítat umístění (X, Y, Z) bodu P ve fotoaparátu souřadnicového systému s použitím následující rovnice.



V rozšířené formě vypadá následovně



## Levenberg-Marquardt optimalizace

Je jedním z nejčastěji používaných algoritmů využívaných k optimalizaci. Dosahuje totiž největší rychlosti nalezení výsledku. Nevýhodou je ovšem jeho složitá implementace.

## SURF

Jedná se o zrychlenou verzi SIFT algoritmu. Slouží k detekci a k popisu klíčových bodů.

## RANSAC

Jedná se o iterační metodu pro odhad parametrů matematického modelu ze sady pozorovaných dat, který obsahuje nekorespondující body, které nemají žádný vliv na hodnoty odhadů. Z toho to důvodu může být interpretován jako způsob detekce odlehlých hodnot. Jedná se o nedeterministický algoritmus ve smyslu, že produkuje přiměřený výsledek jen s určitou pravděpodobnosti. Se zvýšením počtu iterací se zvyšuje pravděpodobnost.

Data se skládají z „inliers“, tedy z dat, jejichž rozdělení lze vysvětil nějakým souborem parametrů modelu. „Outliers“ jsou data, které jsou mimo model a nemají žádný vliv na výsledek.

## Triangulace

Zajišťuje metoda triangulatePoints.

## Kalibrace kamery

Kalibrace je proces, při kterém se získají různé parametry kamery.

# Postup

## Kalibrace

Prvním krokem našeho výpočtu re-fotografického nástroje je kalibraci kamery a zjištění parametrů neznámého historického fotoaparátu. Provedeme 3D rekonstrukci scény na nichž byla historická fotka pořízena. Uživatel pořídí dva snímky, následně je provedena optimalizace parametrů neznámé fotoaparátu, aby se minimalizovala projekční chyba.

### 3.1 3D rekonstrukce

Uživatel začíná tím, že zachytí dva obrazy s širokou základní linie. Široký základ zlepšuje přesnost a stabilita 3D rekonstrukce. Předpokládáme, že současná kamera je kalibrována, a poté se provede SFM, což registruje dvě fotky a rekonstruuje 3D strukturu. Konkrétně, využíváme robustní algoritmus pro odhad pózy. Shodné body mezi těmito dvěma pohledy ve vzorkovací smyčce se vypočte pomocí RANSAC. Pak, vzhledem k projekci dvou kamer, můžeme rekonstruovat 3D souřadnic každého bodu pomocí triangulace. Tyto 3D body se pak promítá do druhého pohledu a zobrazí je uživateli podél referenčního snímku. Uživatel je vyzván ke kliknutí 6-8 bodů. Tyto body se používají k registraci referenční kamery v dalším kroku registrace

## 3.2 Referenční snímek

Dále se nás týká referenční snímek na rekonstruované scéně z prvních dvou fotografií pořízených uživatelem. K zjištění důležitých parametrů referenčního fotoaparátu je použita Levenberg-Marquardt optimalizace, čímž se minimalizuje počet projekčních chyb porovnaných bodů. Předpokládáme nulové zkreslení a devět stupňů volnosti: jeden pro ohniskovou vzdáleností, dva pro hlavní bod, tři pro otáčení a tři pro překlad. Inicializujeme rotaci matice k matici identity, překlad matice na nulu, a ohnisková vzdálenost na ohniskové vzdálenosti aktuální kamery. Inicializace hlavní bodu je popsáno v kapitole 3.2.1.

3.2.1 Princip odhadu bod.

Hlavním bodem je střed optické osy v obrazové rovině. Hlavní bod není ve středu snímku, a musí být vypočítaný. Analýza mizející bodů poskytuje silné podněty pro odvozování umístění hlavního bodu. Vzhledem k mizejícím bodům tří kolmic je hlavní bod umístěn na trojúhelníku, jehož vrcholy jsou mizející body.

Žádáme uživatele ke kliknutí na třech paralelních linkách ve stejném směru, i když dvě paralelní linky jsou dostatečné pro výpočet, žádáme o tři ke zlepšení robustnosti. Vypočítá se průsečík z rovnoběžek. Lokalizují se jednotlivé úběžníky na základě váženého průměr tří průsečíků. Hmotnost je úměrná úhlu mezi dvěma řádky, poněvadž umístění z úběžníku stává méně spolehlivý při menších úhlech. Vyřadí se úběžný bod, pokud je součet těchto tří úhlů nižší než 5 stupňů.

Během Levenberg-Marquardt nelineární optimalizace, se inicializuje hlavní bod jako optický střed. Máme-li jeden konečný a dva nekonečné mizející body.

V souhrnu lze konstatovat, že výsledek z výše uvedených metod je 3D rekonstrukce scény z první a druhé fotky, stejně jako kalibrace referenčního zobrazení a jeho relativní představovat od prvního pohled. Tato informace je pak použita v dalším stupni.

4. VEDENÍ UŽIVATELE V REÁLNÉM ČASE

Náš re-fotografický nástroj poskytuje uživateli v reálném čase navádění směrem k referenčnímu hledisku. K tomu počítáme s relativní představou mezi aktuálním pohledem a referenčním pohledem a s vizualizací potřebnou pro 3D překlad a pro uživatele v podobě dvou šipek, jak je znázorněno na obrázku 9. Ukážeme aktuální pohled na uživatele až po nejlepším vyrovnání rotace mezi aktuálním pohledu a odkazem pohled, který byl použit. Tato rotační stabilizace umožňuje uživateli zaměřit se na 3D překladu a vyhnout se starostí s přesnou rotací fotoaparátu. K dosažení výkonu v reálném čase, budeme prokládat hrubým odhadem, ale relativní představují výpočty s rychlejší, lehké aktualizace schéma. Schéma našeho poradenského přístupu v reálném čase jak je znázorněno na obrázku 6.

## 4.1 Silná pozice fotoaparátu

V našem procesu hrubého odhadu, odhadujeme fotografickou relativní představu na první snímek namísto odkazu, aby se zabránilo degenerace v odhadu, kdy uživatel přiblíží požadovaný hledisko. Protože víme, referenční umístění kamery vzhledem k první snímku [R10 | T10] můžeme odvodit relativní pozici mezi aktuální a referenční fotografií. [Rij | Tij] je j-tý kamera umístěná vzhledem k umístění i-tý kamery. R je jeho rotační složka a T je jeho translační složku. U každého rámce n, my výpočet aktuální pozice kamery vzhledem k prvnímu umístění kamery [R 1N | T1n]. Translační složka T0n proudu kamery umístěné vzhledem k referenčnímu fotoaparátu.

T0n = T1n - R 1N \* R? 10 \* T10.

V naší případě, budeme prokládat tento hrubý odhad s lehkým odhadem. Představujeme podrobnosti v kapitole 4.2.

*4.1.1 Korespondence odhadu*. Chcete-li zjistit shodu mezi první a aktuální fotografií, používáme funkci určenou pro detekci bodů. SIFT je navržen tak, aby byl neměnný v měřítku změny a lineární změny jasu. To je také částečně neměnný na pohledu změny. Pro rychlost, používáme implementaci GPU [Sinha et al. 2006]. Vstupní obrazy mají kolem jednoho megapixelu a náš downsample obrazy od dvou na rychlost. Pro pře vzorkování obrazu, SIFT detekuje kolem jednoho tisíce hraných bodů. Používáme přibližný vyhledávání metoda ANN [Arya et al. 1998] k nalezení shody. Stanovili jsme práh testu druhé poměr [Lowe 2004] poměrně striktně na 0,6 udržovat pouze důvěryhodné korespondence.

*4.1.2 Essential Matrix Odhad*. Nyní vypočítáme relativní fotoaparát představovat mezi prvním pohledu a aktuálního rámce, protože základní je široká, což zabraňuje pohybu degeneraci, když uživatel dosáhne cíle. Vzhledem k tomu, kamera uživatele je již kalibrován, my jediné, co potřebujete odhadnout základní matrici, která se týká kalibrovanou obrazů. Používáme algoritmus Stew'enius je pětibodový [2007], který odhaduje základní matrice mezi dvěma kalibrovaných kamer v reálném čase. Jezdíme MSAC (m-odhadce vzorek konsenzuální) [Torr a Zisserman 2000] najít inliers a nejlepší padnoucí základní matrici. MSAC je podobný RANSAC, ale modifikuje nákladové funkce tak, že extrémy jsou uvedeny pevnou trest, zatímco inliers jsou zaznamenány na, jak dobře se hodí data. Přesnost MSAC se nachází v blízkosti MLESAC (Maximum likelihood konsensus) bez ztráty rychlosti [Torr a Zisserman 2000] WE opravit počet iterací při 1000. Určíme inliers a nejlépe padnoucí zásadní matice s použitím

symetrické epipolární vzdálenost [Hartley a Zisserman 2000]. Náš prahová hodnota je 0,01 za normalizovaných souřadnic bodu.

## 4.2 Real-time kamera Pose Odhad

Chceme poskytovat spolehlivé výsledky, ale také komunikovat s uživateli reálný čas. Naše robustní odhad poskytuje spolehlivé výsledky, ale jeho výpočet je drahý a trvá jen pár sekund. Poskytovat v reálném čase zpětná vazba, budeme prokládat náš hrubý odhad s lehkým odhadem, což není tak silný, ale levné. V našem lehkém odhadu, nebudeme aktualizovat korespondence, ale sledovat generovaný nejnovější sady inliers pomocí funkce sledování a přepočítat relativní kamery představují v jednom průchodu.

Používáme Birchfield v KLT implementační [2007] sledovat funkci body. Provede se kontrola konzistence afinní [Shi a Tomasi 1994] a provádí sledování více škálových, který je vynikající vlastnost bod místa z hrubé až jemné rozlišení.

Naše robustní odhad a lehký odhad jsou prokládány, jak je znázorněno na obrázku 7. robustní odhad rozpozná funkci body, najde zápalky, a odhaduje novou sadu inliers a epipolární geometrii použitím robustní statistiky. To trvá kolem dvou sekund, zatímco Naše lehký odhad běží na více než 10 snímků za sekundu. Tento proces umožňuje prokládané přesnost inliers být zachovány a poskytuje uživatelům s aktualizací v reálném čase.

## 4.3 Interleaved Scheme

Naše proložený potrubí je realizován ve třech závitech: jeden komunikuje s fotoaparátem, druhá vede robustní odhad, a další provádí lehké odhad. Na konci každého robustní odhad, sada inliers je předán do lehkého odhadu vlákno. Uložíme následných rámců klíčového rám, na němž robustní odhad počítá inliers. Když světlo odhad je aktualizována s inlier sady z robustního odhadu, začíná sledování z dalšího snímku klíčového snímku místo proudu frame fotoaparátu. Vzhledem k tomu, lehký odhad používá optický tok pro dynamické body, tam by neměl být velký rozdíl mezi klíčové rámem, kde jsou inliers počítán a první snímek při sledování se spustí. Je-li sada inlier obnoví s novým hrubým odhadem uživatelského výsledku, lze pozorovat jednu sekundu zpoždění. Nicméně, to je zanedbatelné ve srovnání se celý proces re-fotografie, a to dělá, nemá vliv na výkon uživatele nebo výsledné kvality re-fotografie. Naše prokládané verze funguje jako na obrázku 8.

*4.3.1 Sanity testování*. U každé výsledné pozice provádíme tři testy, abychom se ujistili, že naše vizualizace je spolehlivá. Porovnáváme 3D struktura rekonstruován z každého snímku s naší počáteční 3D rekonstrukce z prvních dvou snímků. Měříme chybu 3D všech bodů a ignorovat odhad představovat, pokud je medián 3D chyba je více než 10%. Typicky je střední chyba menší než 5%.

Kromě toho se kontroluje, zda aktuální kamera představuje konzistentní výsledek s těmi předchozími. Zjistili jsme, že jednoduchý filtr funguje, ačkoli Kalmanův filtr [Kalman 1960] by pravděpodobně generovat dobrý výsledek stejně. Měříme průměr a směrodatná odchylka z míst fotoaparát na předchozích deseti rámců a Potvrzení že současná odhadovaná umístění fotoaparátu do 4 standardních odchylek od střední hodnoty. Předpokládáme, že pohyb kamery je hladký a představované variace jsou malé. Výše uvedené dva testy obvykle detekují špatnou odpověď zhruba jednou za 100 snímků.

A konečně, testujeme strukturu degenerace způsobenou, když všechny inliers pocházejí z jedné jediné rovině ve scéně. Najdeme bestfitting homography pomocí RANSAC s 1,5 pixelu průměrnou mapování chyby do 500 iterací. V případě, že počet homography inliers je více než 70% epipolární inliers geometrie, ignorujeme pózy výsledku odhadu. Vzhledem k tomu, používáme velký, dostatečně výchozí stav, Tato chyba se nevyskytuje často.

Když náš odhad výsledku neprojde výše uvedených testů, my prostě neaktualizujeme vizualizaci. Vzhledem k tomu, že špatné odpovědi se nevyskytují často, nemá to žádný výrazný vliv na uživatelský komfort.

4.4 Měřítko odhadu

Po výpočtu relativní pozice, problém zůstává: Měřítko překladu mezi současným rámem a prvním snímkem je nejednoznačný. Proto měřítko je zachovat soulad mezi iterací. V počátečním kroku kalibrace jsme zrekonstruovaná 3D struktura mezi prvním a druhým snímkem pomocí triangulace. V následné iteraci n jsme zrekonstruovali 3D strukturu mezi prvním a druhým rámem. Měřítko mezi těmito dvěma rekonstrukcemi by měla být odlišný konstantní faktor. My můžeme váhy konzistence odhadu měřítka, který způsobuje, že vzdálenost mezi první kamerou a 3D scénou jsou rovnocenné mezi těmito dvěma rekonstrukcemi. K tomu, klademe první fotoaparát na původu pro obě rekonstrukce. Pak jsme vypočítat střední poměr vzdálenosti původu pro každý bod v 3D rekonstrukci první a n-té rekonstrukce. Nakonec jsme násobili délku překladu vektoru podle poměru, což umožní, že délka naší šipky z vizualizace je smysluplná a konzistentní napříč rámy.

4.5 Rotace a stabilizace

Používáme také výsledek relativního odhadu představující rotaci před zobrazením pro stabilizaci aktuálního snímku. Vzhledem k tomu, uživatelé ji najít náročné současně řídit se pokyny, což naznačuje, jak translační a rotační pohyby, my místo toho jen komunikovat překlad pro uživatele. My automaticky vypočítat nejlepší kameru rotace mezi aktuální a referenční názory, a platí to rotace jako osnovu před zobrazením aktuální snímek. Tato rotace uspořádání umožňuje uživateli soustředit se na překládání fotoaparát správný směr, aniž by se snažil držet fotoaparát v právu orientace.

Účinek 3D natočení kamery a zoom mohou být popsány s nekonečným homografem [Hartley a Zisserman 2000]. Nekonečný homograf je podtřídou obecného homografu, as ním je omezena na tuhé rotace kamery a zoom. Používáme algoritmus Brown a kol. [2007] počítat nekonečné homografy, které se hodí všechny epipolární geometrie inliers se metoda nejmenších čtverců chyby.