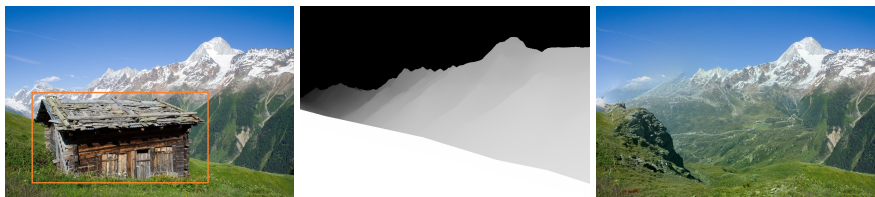


Rekonstrukce digitálních fotografií s využitím hloubkové mapy

Ondřej Valesš*



Abstrakt

Cílem této práce je navrhnout postup, který bude provádět rekonstrukci digitálních fotografií na základě vyhledání podobných fotografií v datasetu, jejichž obsah je poté využit při rekonstrukci. K nalezení podobnosti je navrhnout deskriptor fotografií obohacený o informaci z hloubkové mapy, který umožňuje rychlé vyhledání fotografií v datasetu a nový postup vkládání dat do rekonstruované oblasti, která je na rozdíl od existujících přístupů nejprve rozdělena na podoblasti na základě analýzy hloubkové mapy. To umožňuje řešit rekonstrukci jednotlivých podoblastí zcela odděleně a přináší zlepšení výsledku rekonstrukce, protože podoblasti jsou menší a méně komplexní.

Klíčová slova: digitální fotografie — rekonstrukce — hloubková mapa — deskriptor GIST — Poissonova rovnice — dataset — C++ — OpenCV

Přiložené materiály: [Dotazník](#)

*xvales03@stud.fit.vutbr.cz, Fakulta informačních technologií, Vysoké učení technické v Brně

1. Úvod

Rekonstrukci digitální fotografie lze využít pro opravu poškození, odstranění nežádoucích objektů z fotografie, případně pro doplnění chybějících oblastí při skládání několika fotografií dohromady (příklad rekonstrukce digitální fotografie je na obrázku 1).

Současná řešení mají problémy s rekonstrukcí větších oblastí, protože jediné informace o chybějícím obsahu mohou čerpat pouze z okolí poškozené oblasti a s velikostí poškozené oblasti roste vzdálenost od poškozených pixelů k nejbližšímu zdroji informací.

V dnešní době je časté, že digitální fotografie jsou opatřeny GPS souřadnicemi místa pořízení. To umožňuje generování přesné hloubkové mapy na základě 3D modelu místa pořízení fotografie (například projekt LOCATE [1, 2]). Takováto hloubková mapa přináší informace přímo o poškozené oblasti, které ostatní metody rekonstrukce získávají jen velmi obtížně.

Výsledkem je rekonstrukce, která neodhaduje, co na poškozeném místě mělo být, ale zachovává model původního místa pořízení.

Tato práce je zaměřena zejména na fotografie krajiny, pro které jsou k dispozici kvalitní syntetické hloubkové mapy získané z digitálních modelů terénu (prezentované výsledky využívají 3D model Alp). Digitální výškové modely terénu jsou k dispozici pro celou planetu, ale jejich využití v městských aglomeracích je omezené.

Rekonstrukce je také využívána při vytváření 3D efektu z jediné fotografie (pohled z jiného úhlu vytváří prázdná místa) a při natáčení dobových filmů, kdy je ze záběru potřeba odstranit moderní budovy.

1.1 Průběh rekonstrukce

Na začátku rekonstrukčního procesu je fotografie s vyznačením poškozené nebo nahrazované oblasti. Vý-



Obrázek 1. Odstranění budovy z fotografie je možné považovat za speciální případ opravy poškození.

stupem je fotografie, která zachovává nepoškozenou část a vyznačenou oblast nahrazuje jinými daty tak, aby výsledná fotografie byla věrohodná a nebylo na ní poznat, že byla rekonstruována.

V ideálním případě by rekonstrukce měla vyznačenou oblast nahradit přesně tím, co na poškozeném místě původně bylo. To ale ve většině případů není možné, protože neexistuje zdroj informací přímo o poškozeném místě. Využití hloubkové mapy tento problém částečně odstraňuje.

Snahou je navrhnout autonomní systém, který kromě označení poškozené oblasti nebude vyžadovat záslahu uživatele, bude produkovat co nejvěrohodnější výsledky i pro velké oblasti poškození a nová data nahrazující poškozenou oblast se budou co nejvíce blížit skutečnému původnímu obsahu. Věrohodnost výsledku lze označit jako vizuální správnost, přiblížení původnímu obsahu jako sémantickou správnost [3].

Vizuální správnost je možné otestovat pomocí dotazníku. Pokud nebudou respondenti schopni rozlišit fotografie, které nikdy nebyly poškozeny a ty, které prošly rekonstrukcí, úroveň vizuální správnosti je dostačující. Určení míry sémantické správnosti je možné porovnáním nepoškozeného originálu a rekonstruované fotografie.

1.2 Základní přístupy k rekonstrukci

Přístupy k rekonstrukci digitálních fotografií můžeme rozdělit do dvou základních kategorií: image inpainting a syntéza textur (viz kapitola Existující přístupy 2).

Další možností možností je doplnit poškozenou oblast daty z jiné fotografie. K tomu slouží vyhledání podobných fotografií v datasetu (viz obrázek 2). Tento princip je popsán v práci Scene Completion using Millions of Photographs [4].

Navrhované způsoby využití hloubkové mapy při rekonstrukci vycházejí převážně z práce Scene Completion 2.4. Jedná se o zdokonalení prohledávání

datasetu, které bere v potaz i podobnost hloubkových map. K tomuto účelu je navrženo rozšíření deskriptoru GIST tak, aby obsahoval i informace z hloubkové mapy nebo je používáno dvouúrovňové prohledávání nejprve podle GIST a poté podle hloubkové mapy se sníženým rozlišením.

Zcela novým přístupem k rekonstrukci je rozdělení oblasti na základě hloubkové mapy na části se stejnou hloubkou. Tento přístup staví na předpokladu, že části fotografie stejně vzdálené od kamery obsahují jeden objekt nebo spolu související objekty, což zjednodušuje rekonstrukci. Další novinkou je možnost vygenerovat celou fotografii přímo z hloubkové mapy (StyLit [5]), výřez této syntetické fotografie je možné použít pro rekonstrukci.

1.3 Výhody využití hloubkové mapy

Hlavním přínosem použití hloubkové mapy je získání dalších informací o rekonstruované fotografii. Ty se projeví při prohledávání datasetu, protože vybrané zdrojové fotografie mají podobnou hloubku a tedy i podobnou míru detailu jako poškozená fotografie. To vede na plynulejší navázání dat do rekonstruované fotografie a vizuálně lepší výsledek.

Rozdělení rekonstruované fotografie podle hloubkové mapy zase umožňuje provádět rekonstrukci pro každou část zcela odděleně s tím, že pro jednotlivé části je zachováno pouze relevantní (stejně vzdálené) okolí. K rekonstrukci dílčích částí můžeme využít libovolné, i s hloubkovou mapou nepracující, metody (dokonce pro každou část jinou).

Další výhodou rozdělení podle hloubkové mapy je zjednodušení obsahu dílčích částí, protože hranice objektů na fotografii kopírují hranice v hloubkové mapě, což usnadňuje rekonstrukci dílčích částí.

Posledním přínosem hloubkové mapy z 3D modelu je poskytnutí odrazu skutečného obsahu poškozeného místa [1], rekonstrukce tedy není prováděna pouze s cílem vizuální správnosti, ale je možné poškozené



Obrázek 2. Rekonstrukce s využitím dalšího zdroje dat. Označená část fotografie je doplněna výřezem z jiné fotografie. Jsou prezentovány tři možné výsledky rekonstrukce s lišící se zdrojovou fotografií.

místo doplnit přibližně tím, co na něm původně bylo. Tomu se ostatní přístupy nemohou přiblížit.

1.4 Struktura práce

Tento text se věnuje rozboru existujících přístupů (kapitola Existující přístupy 2), návrhu využití hloubkové mapy při vyhledávání v datasetu (kapitola Hloubková mapa při hledání podobnosti 3) a při přípravě fotografie k rekonstrukci (kapitola Rozdělení fotografie podle hloubky 4) a implementaci a testování rekonstrukčního postupu (kapitola Implementace a výsledky 5).

2. Existující přístupy

2.1 Image inpainting

Image inpainting využívá kopírování pixelů z okolí poškozené oblasti, přičemž se snaží zachovat strukturu zbytku fotografie [3]. Toho se snaží docílit propagací hran vcházejících do poškozené oblasti, podle kterých je potom řízeno vyplňování oblasti.

Modifikace tohoto přístupu se liší hlavně ve způsobu propagace hran, kdy se snaží navazovat dvojice hran přes poškozenou oblast a v heuristikách, které používají pro výběr a modifikaci pixelů pro vyplnění poškozené oblasti. Silnou stránkou tohoto přístupu je rychlost rekonstrukce a autonomnost procesu, protože kromě označení poškozené oblasti není potřeba žádný vstup od uživatele.

Image inpainting je vhodný pouze pro rekonstrukci úzkých oblastí, protože nebere v potaz detaily a texturu rozšiřované do poškozené oblasti [3], což při doplnění velké oblasti vede na rozmazané řešení bez dostatečných detailů.

2.2 Syntéza textur

Nedostatek detailů v řešení se snaží odstranit syntéza textur. K rekonstruované oblasti uživatel vybere (ze stejné nebo i zcela odlišné fotografie) zdrojovou texturu. Rekonstrukce je potom provede vygenerováním syntetické textury, která má podobné vlastnosti (hlavní je vizuální podoba) jako zdroj.

Syntézu textur je možné využít pro libovolně velké oblasti bez ztráty detailu, problémem ale mohou být fotografie s prolínajícími se texturami, kde není k poškozené oblasti možné najít jednotný zdroj textury. Další nevýhodou je potřeba manuálně označit zdrojovou texturu, zvláště u oblasti zasahující do několika objektů s odlišnou texturou. Uživatel potom musí oblast rozdělit na části a ke každé z nich definovat zdrojovou texturu [6]. Modifikace tohoto postupu se zaměřují na automatizaci rozpoznání zdrojové textury a využití fuzzy logiky [7] pro řešení prolínajících se textur.

2.3 Pokročilé metody

Současné metody se snaží zkombinovat oba přístupy a dosáhnout tak současně autonomnosti systému a rekonstrukce s vysokou mírou detailu i pro velké oblasti. Patří sem Exemplar Based Image Inpainting [8], kde se používá interpolace hran do poškozené oblasti, která je tak automaticky rozdělena. Příslušné části jsou potom doplněny texturou z okolí přímo sousedícího s podoblastí. Pro uživatele tak odpadá nutnost řídit rekonstrukci.

Dalším příkladem je Error Concealment [7], kde se využívá sdružování oblastí s podobnou texturou do superpixelů [9]. Celá rekonstrukce je potom provedena na úrovni superpixelů a nakonec je provedena zpětná transformace, kdy je oblast příslušící superpixelu vyplněna jeho texturou. Použití superpixelů umožňuje rekonstrukci oblastí s prolínající se texturou.

Příslušnost pixelu z původní fotografie k superpixelu je vyjádřena pomocí fuzzy logiky, což umožňuje přesah superpixelů – ekvivalent prolínající se textury. Rekonstrukce na úrovni superpixelů se zásadně neliší od neliší od Image Inpainting (krom informací o textuře), základem je propagace hran do rekonstruované oblasti a její vyplnění superpixelu.

Metoda PatchMatch [10] rozdělí rekonstruovanou fotografii na políčka $7px \times 7px$ a vytvoří mapování mezi nejpodobnějšími políčky. Samotná rekonstrukce je potom řízena uživatelem, který ve fotografii označí poškozenou oblast a určí pomocné vodící linky.

Poškozená oblast je doplňována od poškozených políček na vnějšku oblasti, ty jsou doplněny pomocí mapování sousedních nepoškozených políček. Poškozená oblast je tak vyplňována podobnými políčky, ale díky mapování se nestane, že by celá oblast byla vyplněna nakopírováním stejného políčka.

2.4 Nahrazení z jiné fotografie

Řešením, které obchází potřebu odhadovat obsah poškozené oblasti a kopírovat data z okolí nebo generovat texturu, je Scene Completion [4]. To využívá dataset, ve kterém vyhledá fotografii podobnou té rekonstruované, a poškozenou oblast nahradí daty z nalezené fotografie. Fotografie je tedy rekonstruována za pomoci části skutečné scény, lze tak dosáhnout vizuálně i sémanticky správného řešení.

Pravděpodobnost nalezení dostatečně podobné fotografie v datasetu roste s jeho velikostí, pro dosažení uspokojivých výsledků bylo využito datasetu obsahujícího 2 miliony fotografií. To vytváří potřebu vytvoření spolehlivého a rychlého způsobu, jak v takto velkém datasetu nalézt vizuálně podobné fotografie. I při využití deskriptorů předpočítaných pro dataset

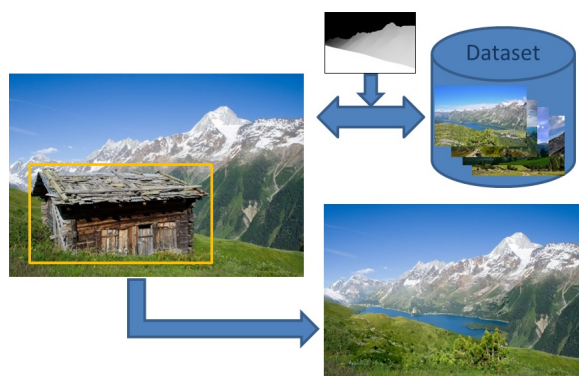
je doba potřebná pro rekonstrukci řádově větší než u ostatních přístupů [4].

Dalším problémem je plynulé navázání dat vložených z jiné fotografie, k tomu je využit Poisson blending a Poisson image editing [11]. Jedná se o nalezení takových dat k doplnění, které mají hranici z původní fotografie a zachovávají barevné poměry zdrojových dat z datasetu.

3. Hloubková mapa při hledání podobnosti

Návrh využití hloubkové mapy při vyhledávání navazuje na Scene Completion [4] a využití deskriptoru GIST. Deskriptor GIST byl vytvořen pro rychlé vyhledávání sémanticky blízkých fotografií a je tvořen vektorem čísel, která reprezentují množství, rychlost přechodu a orientaci hran v jednotlivých částech fotografie, včetně informací o velikosti a poměru stran.

Využití takového vektoru pro vyhledávání bylo předvedeno v práci Evaluation of GIST Descriptors for Web-scale ImageSearch [12, 13]. Vzdálenost vektorů GIST dvou fotografií potom vyjadřuje jejich relativní podobnost. Příklad nalezené podobné fotografie je na obrázku 3.



Obrázek 3. Vyhledání podobné fotografie v datasetu. Nalezená fotografie má podobný vzhled i hloubkovou mapu a je využita jako zdroj dat při rekonstrukci.

3.1 GIST hloubkové mapy

Nejjednodušším způsobem využití hloubkové mapy je její převod do podoby obrázku a výpočet GIST i pro samotnou hloubkovou mapu. Protože hloubková mapa je bez textur a barev, ale reprezentuje základní rozložení scény, vyhledávání pomocí tohoto deskriptoru vrací fotografie, které jsou si podobné rozvržením, ale nikoli barvou. Proto je při vyhledávání potřeba kombinovat GIST hloubkové mapy a samotné fotografie. Efektivně se tím zdvojnásobí délka deskriptoru, což vede na pomalejší vyhledávání. Výsledná podobnost

dvou fotografií se potom spočítá jako:

$$D_{total} = t \times D_{photo} + (1 - t) \times D_{depth}, \quad (1)$$

kde D_{total} je výsledná vzdálenost, D_{photo} je vzdálenost v prostoru GIST fotografie a D_{depth} je vzdálenost v prostoru GIST. Parametr t vyjadřuje váhu kladenou barvě při určování podobnosti.

Nevýhodou takového deskriptoru je jeho zaměření na hrany, nebere tak při vyhledávání v potaz absolutní hloubky v odpovídajících si částech fotografií.

3.2 Mapa průměrných hloubek

Mapa průměrných hloubek je pro účely této práce navrhnutý deskriptor, který je založen na absolutní hloubce fotografie. Jedná se vektor průměrných hodnot hloubky v jednotlivých oblastech fotografie.

Hloubková mapa fotografie je rozdělena na $M \times M$ sektorů a pro každý sektor je vypočtena průměrná hloubka. Průměrné hloubky jsou potom po řádcích umístěny do vektoru deskriptoru. Volba hodnoty M závisí na požadavcích na rychlost a kvalitu prohledání datasetu, pro velké hodnoty je výsledný deskriptor dlouhý a vyhledávání pomalé, pro malé hodnoty je zvýšena rychlost vyhledávání na úkor kvality.

Tento deskriptor dobře popisuje absolutní hodnoty v hloubkové mapě, stejně jako GIST hloubkové mapy však neobsahuje informace o barvě a je nutné jej kombinovat s GISTem fotografie. Protože absolutní hodnoty hloubky jsou řádově vyšší než odezva na filtrem detekující hrany využitý v případě GIST, je potřeba vzdálenost v prostoru map průměrných hloubek normovat:

$$D_{total} = t \times D_{photo} + \frac{(1 - t) \times D_{avg_map}}{N}, \quad (2)$$

kde D_{avg_map} je vzdálenost v prostoru map průměrných hloubek, N je normovací konstanta a význam zbylých členů je stejný jako v případě GISTu hloubkové mapy 3.1.

3.3 Zachování absolutních hodnot i citlivosti na hrany

Ideální způsob vyhledávání by umožňoval podle potřeby přiřazovat prioritu jak barvě, tak hloubce a u hloubky navíc rozlišovat, jestli má být vyhledávání zaměřeno na rozložení scény (citlivost na hrany a jejich orientaci) nebo na absolutní vzdálenost částí scény. To vyžaduje zkombinování všech tří deskriptorů (GIST, GIST hloubkové mapy, mapa průměrných hloubek) do jediné výsledné vzdálenosti použité při vyhledávání:

$$D_{total} = (1 - t_{depth}) \times D_{photo} + t_{depth} \times D_{total_depth}, \quad (3)$$

kde $1 - t_{depth}$ určuje váhu přiřazenou barvám a t_{depth} váhu hloubkové mapy. Člen D_{total_depth} určující podobnost hloubkových map se vypočítá:

$$D_{total_depth} = t_{abs} \times \frac{D_{avg_map}}{N} + (1 - t_{abs}) \times D_{depth}, \quad (4)$$

kde t_{abs} určuje váhu absolutních hodnot a $1 - t_{abs}$ váhu rozložení scény.

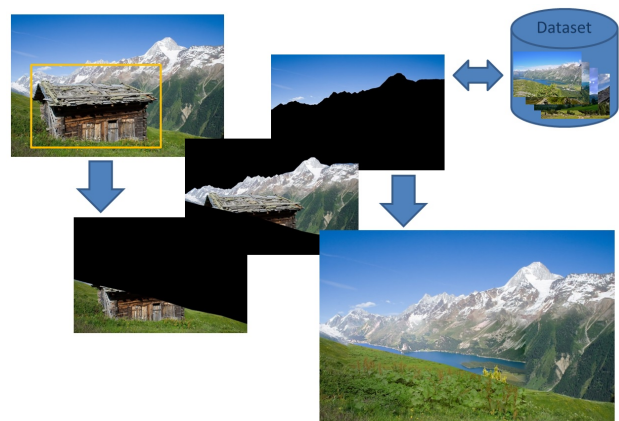
Vzniká tak univerzální způsob vyhledávání kombinující jak barvu, tak hloubku, pro každé vyhledání lze navíc samostatně určit, na jaké aspekty podobnosti má být kladen důraz. Protože vyhledání kombinuje všechny tři deskriptory, je doba potřebná k vyhledání podobných fotografií delší (přibližně suma dob potřebných k vyhledání pro každý dílčí deskriptor).

Porovnání pomocí GIST je provedeno s každou fotografií v datasetu zvlášť, proto je výsledná třída časové složitosti pro jednoduchý deskriptor $O(mn)$, kde m je délka GIST a n počet fotografií v datasetu. Kombinovaný deskriptor navržený v této práci využívá tři nezávislé deskriptory (stejně délky jako GIST), třída složitosti se proto nemění a je pouze násobena konstantou. Výsledkem je složitost $O(3mn)$.

Prohledání datasetu o velikosti 3000 fotografií trvá s využitím kombinovaného deskriptoru jednotky sekund, pro dataset 100 000 fotografií je to 1 až 2 minuty.

4. Rozdělení fotografie podle hloubky

Protože obtížnost rekonstrukce roste se složitostí poškozené oblasti, je nutné navrhnout robustní rekonstrukční algoritmus, který se s tímto problémem vypořádá. Existující přístupy 2 řeší složitost oblasti jako celek, který se snaží zjednodušit abstrakcí na superpixely, mapováním textur, případně interpolací hran. Tedy



Obrázek 4. Fotografie je rozdělena na části, které jsou rekonstruovány samostatně. Výsledná složená rekonstrukce zachovává přesně původní hloubkovou mapu.

v průběhu rekonstrukce je problém postupně rozkládán na dílčí úlohy.

Hloubková mapa umožňuje zvolit přístup opačný. Její využití vychází z předpokladu, že související objekty se ve fotografii nacházejí přibližně ve stejné vzdálenosti. Jinak řečeno, pokud se v hloubkové mapě vyskytuje skoková změna hodnoty, pixely na jedné a druhé straně spolu nesouvisí a náleží k jinému objektu.

Podél této hranice v hloubkové mapě vede ve fotografii linie přechodu mezi objekty, které na sebe navzájem vizuálně nenavazují a tato nespojitost není rušivá, protože se vyskytuje v místě kde je očekávána.

Podél těchto linií je možné celou fotografii (spolu s poškozenou oblastí) rozdělit na části, se kterými je možné pracovat odděleně. Odlišnosti vzniklé odděleným zpracováním se potom vyskytují v místě nespojitosti původní fotografie a nepůsobí rušivě (viz obrázek 4).

Celá rekonstrukce tak probíhá směrem od dílčích částí, které jsou rekonstruovány zvlášť (i s využitím rozdílných metod rekonstrukce), jejichž složením vznikne výsledné řešení.

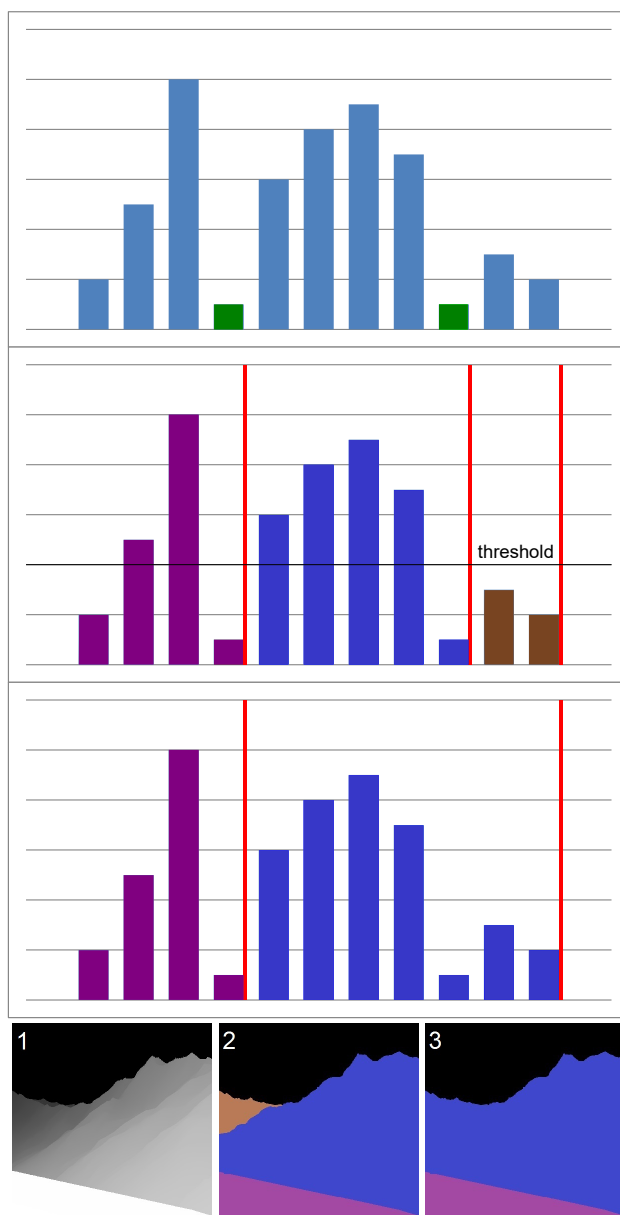
Rekonstrukce dílčích částí probíhá stejným způsobem jako rekonstrukce nerozdělené oblasti s výjimkou hranic, kde spolu sousedí dvě rekonstruované oblasti. Takováto hranice je brána jako volná a neklade žádné požadavky na návaznost s okolím (dílčí rekonstrukce tedy musí navazovat pouze na nerekonstruovanou část fotografie). Výsledná fotografie je bez dalších úprav složena z těchto částí.

4.1 Proces dělení

Od rozdělení očekáváme, že bude kopírovat hloubkovou mapu, bude generovat souvislé podoblasti (s jistou tolerancí) a počet podoblastí bude odpovídat složitosti scény. Také bychom chtěli, aby podoblasti měly definovanou minimální velikost a zabránily tak vygenerování příliš jemného dělení (aby například hory v pozadí jenž se postupně vzdalují tvořily po rozdělení jedinou část).

K rozdělení je používána analýza histogramu hloubkové mapy. Ta nevyžaduje zadání výsledného počtu podoblastí, který tak může odrážet složitost fotografie. Umožňuje slučování sousedních intervalů, pokud podoblasti nemají minimální velikost, a na rozdíl od alternativy K-means clustering [14] nepotřebuje počáteční body v hloubkové mapě.

Hloubkové mapy mají v podstatě neomezený rozsah hodnot – od několika málo centimetrů pro objekty přímo před fotoaparátem až po nekonečno v případě oblohy. Toto je v praxi omezeno na rozsah 4 bytového floating point čísla, přesto nastává problém s volbou intervalů histogramu, který musí mít vypovídající hodnotu jak pro blízké, tak vzdálené objekty a rozumný

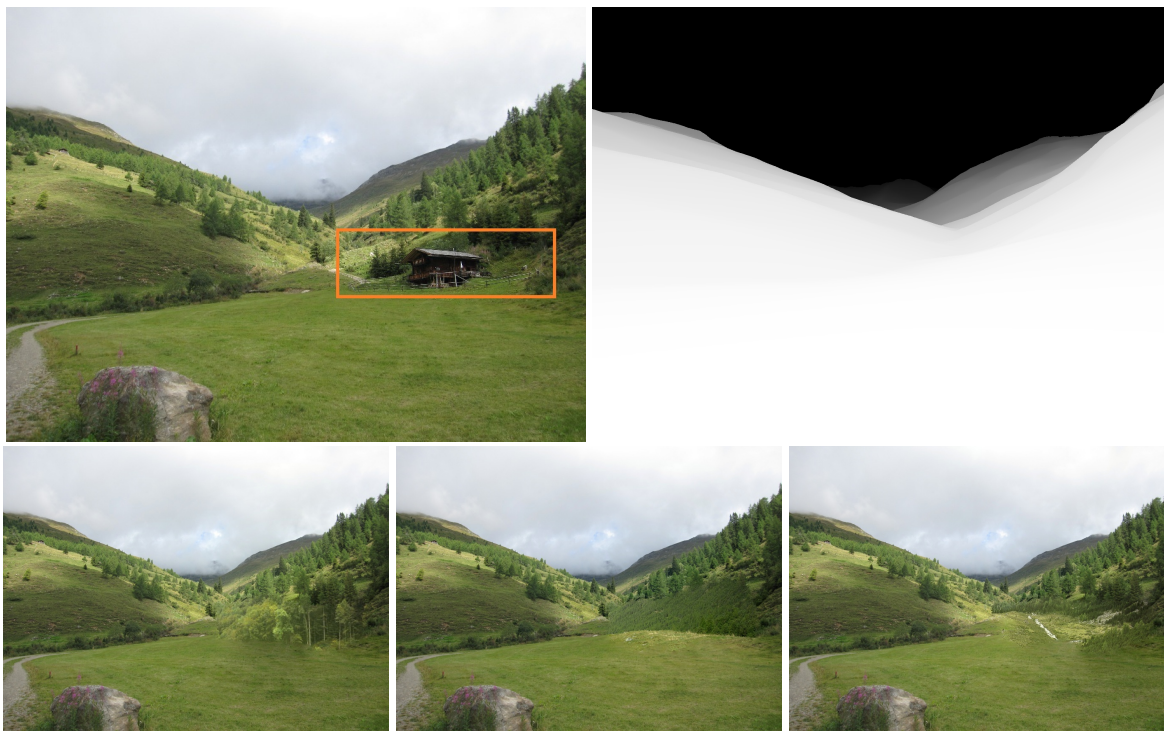


Obrázek 5. Ilustrace rozdělení histogramu na části, v prvním kroku jsou nalezena minima (vyznačeno zeleně), ve druhém kroku je vytvořeno základní dělení na intervaly, ve třetím kroku je sloučen interval (označen hnědě) s nedostatečnou velikostí se svým sousedem.

celkový počet intervalů. Proto je využíván histogram s nesterjně velkými intervaly, kde hodnoty hranic intervalů rostou exponenciálně.

To vychází z předpokladu, že s rostoucí celkovou vzdáleností nám přestává záležet na přesnosti dělení (hory na pozadí můžou být vzdálené 5000 metrů nebo 5500 metrů a pořád budou patřit ke stejnému logickému celku, ale objekty vzdálené 10 metrů a 510 metrů už budou tvořit celky různé). Exponenciálně se zvětšující intervaly tak řeší velký rozsah hodnot a současně zachovávají potřebnou přesnost.

Pro rozsah hodnot hloubkové mapy je nejprve



Obrázek 6. Příklad rekonstrukce velké oblasti s komplexní strukturou poškozené oblasti. Při prohledávání datasetu byl použit GIST hloubkové mapy a mapa průměrných hloubek. Zobrazeno několik možných výsledků rekonstrukce.

připraven histogram s příslušnými intervaly, který je v prvním kroku naplněn hodnotami. Poté je celý histogram rozdělen podle lokálních minim s tím, že interval, ve kterém se nachází lokální minimum, je přiřazen k sousednímu intervalu s menší hodnotou.

Ve třetím kroku jsou potom nalezeny intervaly, které nesplňují podmínku minimální velikosti. Ty jsou sloučeny se sousedním intervalem stejně jako minima. Třetí krok je opakován, dokud existují intervaly s nedostatečnou velikostí. Příklad rozdělení histogramu je na obrázku 5. Fotografie je rozdělena podle příslušnosti hloubek do intervalů pro jednotlivé pixely.

5. Implementace a výsledky

V rámci projektu byly použity datasety Alps100k [2] a jeho podmnožina GeoPose3K [15] s hloubkovými mapami. Pro účely vyhledávání byly pro celý dataset předpočítány deskriptory GIST, GIST hloubkové mapy a průměrné hloubky. Dataset není součástí řešení, jsou zahrnuty pouze skripty pro výpočty deskriptorů.

Byl implementován program `repair` pro rekonstrukci jedné fotografie s přiloženou maskou poškození a hloubkové mapy. Pro rekonstrukci je využito vyhledání podobných fotografií v datasetu, pro které uživatel volí váhy přiřazované fotografií a poměrné a absolutní hloubky, jak je popsáno v kapitole Hloubková mapa při hledání podobnosti 3.

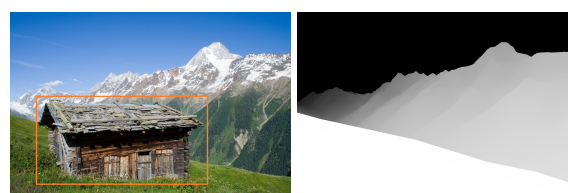
Výsledkem vyhledání je 100 nejpodobnějších fo-

tografií, ze kterých je z každé vytvořena rekonstruovaná fotografie vyřešením Poissonovy rovnice [11] s možnou úpravou okolních pixelů, jak je popsáno ve Scene Completion [4]. Jako výsledek je uživateli prezentováno 10 různých verzí rekonstrukce zadané fotografie (viz obrázek 6).

Druhý program `slice` provádí dělení podle hloubkové mapy, využívá techniku popsanou v kapitole Rozdělení fotografie podle hloubky 4. Jako základ pro exponenciální dělení intervalů používá číslo 2, které bylo vybráno na základě provedení experimentů. Po rozdělení používá pro rekonstrukci dílčích částí program `repair`, jiné metody dílčí rekonstrukce nebyly implementovány. Takto opravené části jsou poté bez dalších úprav složeny.

5.1 Výsledky

Testování `repair` bylo provedeno se sadou 50 fotografií k rekonstrukci. Použité hodnoty t_{depth} (rovnice 3) byly 0,3; 0,5; a 0,7 a $t_{abs} = 0,5$ (rovnice 4). Toto



Obrázek 7. Rekonstruovaná fotografie a její hloubková mapa.



Obrázek 8. Rekonstrukce bez hloubkové mapy.



Obrázek 9. Rekonstrukce s hloubkovou mapou.

testování pracovalo s datasetem obsahujícím hloubkové mapy (3000 fotografií).

Pro účely porovnání byl proveden také test s $t_{depth} = 0,0$ (bez vlivu hloubkových map) a to pro oba datasety.

Provedené testy ukazují, že vizuálně nejlepších výsledků je dosaženo s $t_{depth} = 0,3$. Výsledné fotografie z této sady dobře navazují na neopravené okolí a na rozdíl od rekonstrukce nevyužívající hloubkové mapy dobře zachovávají původní rozložení rekonstruované oblasti.

Příklad porovnávací výsledky je na obrázcích 8 a 9. Fotografie na obrázku 8 byla rekonstruována bez využití hloubkové mapy. Po vizuální stránce je v pořádku, ale nerespektuje horizont v poškozené oblasti. Rekonstrukce fotografie na obrázku 9 využila hloubkovou mapu, což vedlo na lepší zachování skutečného obsahu oblasti.

V případech, kdy je na hloubkovou mapu kladen příliš velký důraz ($t_{depth} > 0,5$), jsou rekonstrukce po vizuální stránce špatné, protože podobnost hloubkových map převáží podobnost samotných fotografií. To vede na výsledky, kdy je oblast doplněna daty s jinou barvou nebo texturou. Příklad rekonstrukce s příliš velkým důrazem na hloubkovou mapu je na obrázku 10.

Využití hloubkových map tedy přináší zlepšení



Obrázek 10. Rekonstrukce s přílišným důrazem na hloubku.

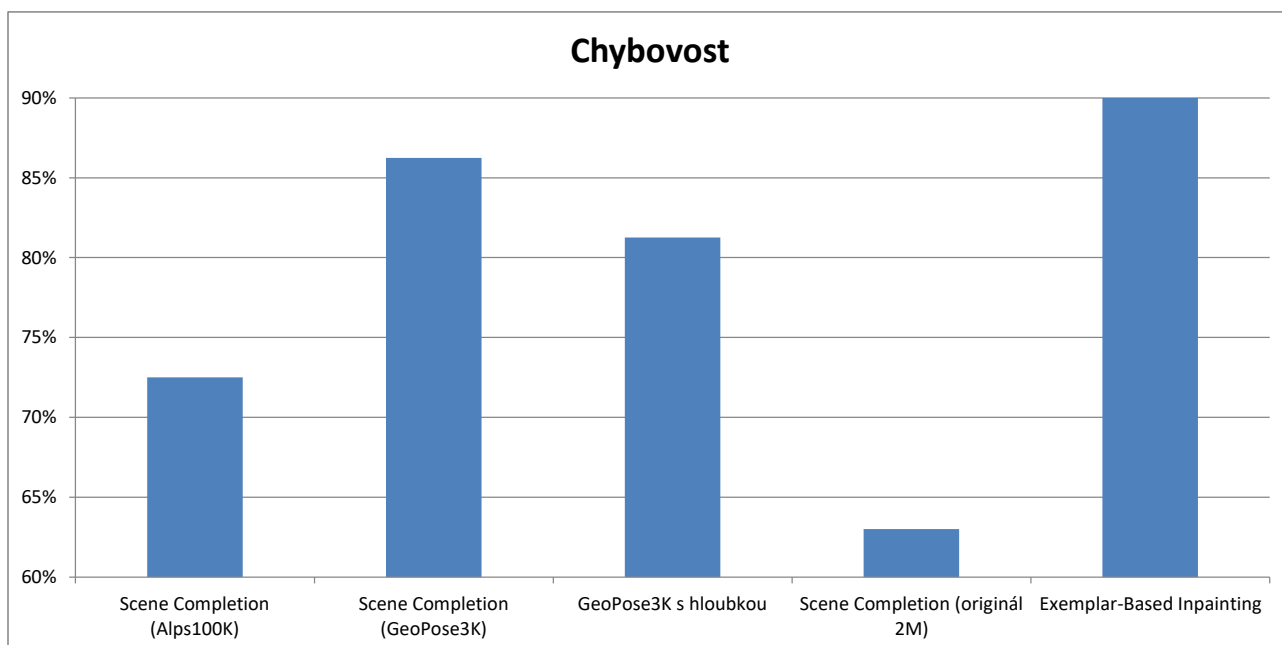
výsledků ve smyslu lepšího přiblížení se původnímu obsahu. Je však závislé na barevné informaci z fotografie a pokud je informace o barvě nevyužita, rekonstrukce na základě hloubkové mapy selže.

Testování *slice* bylo provedeno se stejnou testovací sadou a datasetem GeoPose3K. Výsledky ukazují, že hloubkovou mapou řízená rekonstrukce výborně zachovává rozložení fotografie. Jediným nedostatkem je absence další úpravy dat v průběhu skládání.

Rekonstrukce pomocí *slice* dosahuje vysoké sé-



Obrázek 11. Dva možné výsledky rekonstrukce pomocí *slice*. Oba výsledky přesně kopírují hloubkovou mapu. Na obrázku dole jsou jasně patrné stopy skládání dílčích oblastí.



Obrázek 12. Výsledky dotazníku s přidanou chybovostí Scene Completion (originál 2M) a Exemplar-Based Inpainting [4], převzatých z původního článku. Tyto hodnoty jsou výsledkem testování s podstatně větší sadou výsledků a testovací skupinou.

mantické správnosti, která se velmi blíží skutečnému původnímu obsahu poškozeného místa. Po vizuální stránce zaostává za aplikací `repair`, protože přestože se hranice oblastí nachází v očekávaných místech, stále jsou příliš výrazné. Příklad rekonstrukce pomocí `slice` je na obrázku 11.

5.2 Vyhodnocení dotazníku

Pro účely testování byl vytvořen dotazník obsahující 20 fotografií. U každé fotografie respondent zadával, jestli si myslí, že fotografie byla upravena nebo ne. Dotazník obsahoval originální fotografie, fotografie rekonstruované bez hloubkové mapy s využitím datasetu Alps100K, fotografie rekonstruované bez hloubkové mapy s využitím datasetu GeoPose3K a fotografie rekonstruované s hloubkovou mapou s využitím datasetu GeoPose3K ($t_{depth} = 0,3$). Výsledky¹ dotazníku jsou na obrázku 12.

Dotazník testoval pouze vizuální správnost, sémanticky chybné rekonstrukce (asi 10% všech výsledků) se v dotazníku nevyskytovaly. Dataset Alps100K sloužil jako kontrolní a jeho výsledky (chybovost přibližně 73%) by měly být obdobné jako u konkurenčních metod (konkrétně SceneCompletion s chybovostí přibližně 63% [4]).

Tato implementace tedy nedosahuje stejné kvality jako Scene Completion, což může být způsobeno využitím menšího datasetu a skutečností, že v této

implementaci je pro řešení Poissonových rovnic a optimalizaci hranic využit software třetích stran, který není primárně určen pro tento účel.

Zvýšení chybovosti (86%) u datasetu GeoPose3K bez hloubek je očekávané, protože tento dataset je podstatně menší. Při využití stejně velkého datasetu v kombinaci s hloubkovými mapami však chybovost klesne o přibližně 5% na 81%. Přidání informací z hloubkové mapy tedy (při stejné velikosti použitého datasetu) vede na zlepšení výsledků. Na základě grafu 12 lze předpokládat, že pro větší datovou sadu budou výsledky lepší, než Scene Completion (originál 2M).

6. Závěr

Tato práce se věnovala rekonstrukci fotografií s využitím hloubkové mapy, v jejím průběhu byly navrženy a otestovány způsoby využití hloubkové mapy při hledání v datasetu na základě absolutních hodnot hloubky, relativních poměrů v hloubkové mapě a jejich kombinace. Dále byla hloubková mapa použita k rozdělení komplexní rekonstruované oblasti na dílčí části, což zjednodušilo rekonstrukci rozsáhlých poškození s komplexním okolím.

Rekonstrukce na základě hloubkové mapy přináší možnost poškozené místo nahradit daty podobnými původnímu obsahu, na rozdíl od konkurenčních metod, které mohou původní obsah poškozené oblasti pouze odhadovat na základě okolí. Pomocí dotazníku bylo ukázáno, že při stejné velikosti datasetu vede využití hloubkových map na snížení chybovosti výsledků.

¹ Jedná se o předběžné výsledky. V době tvorby tohoto textu dotazník ještě nevyplnil dostatečný počet respondentů.

Mezi další cíle, které by měly přinést zlepšení výsledků, patří rozšíření datasetu s hloubkovými mapami, přidání podpory dalších metod rekonstrukce pro dílčí oblasti po rozdělení na základě hloubkové mapy a zpracování dat v průběhu skládání dílčích částí do celku.

Poděkování

Chtěl bych poděkovat vedoucímu mé bakalářské práce docentu Martinu Čadíkovi za jeho pomoc, kterou mi poskytl během práce na tomto tématu a při psaní tohoto článku.

Literatura

- [1] Lionel Baboud, Martin Čadík, Elmar Eisemann, and Hans-Peter Seidel. Automatic photo-to-terrain alignment for the annotation of mountain pictures. In *Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition*, CVPR '11, pages 41–48, Washington, DC, USA, 2011. IEEE Computer Society.
- [2] Martin Čadík, Jan Vašíček, Michal Hradiš, Filip Radenović, and Ondřej Chum. Camera elevation estimation from a single mountain landscape photograph. In *British Machine Vision Conference 2015*, pages 1–12. The British Machine Vision Association and Society for Pattern Recognition, 2015.
- [3] Marcelo Bertalmio, Guillermo Sapiro, Vincent Caselles, and Coloma Ballester. Image inpainting. In *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '00, pages 417–424, New York, NY, USA, 2000. ACM Press/Addison-Wesley Publishing Co.
- [4] James Hays and Alexei A. Efros. Scene completion using millions of photographs. In *ACM SIGGRAPH 2007 Papers*, SIGGRAPH '07, New York, NY, USA, 2007. ACM.
- [5] Jakub Fišer, Ondřej Jamriška, Michal Lukáč, Eli Shechtman, Paul Asente, Jingwan Lu, and Daniel Sýkora. StyLit: Illumination-guided example-based stylization of 3d renderings. *ACM Transactions on Graphics*, 35(4), 2016.
- [6] Jian Sun, Lu Yuan, Jiaya Jia, and Heung-Yeung Shum. Image completion with structure propagation. In *ACM SIGGRAPH 2005 Papers*, SIGGRAPH '05, pages 861–868, New York, NY, USA, 2005. ACM.
- [7] Ivana Ilčíková. Texture Aware Video and Image Error Concealment [online]. Master's thesis, Univerzita Komenského v Bratislave, 2016. [Online; navštíveno 29.9.2016].
- [8] A. Criminisi, P. Perez, and K. Toyama. Region filling and object removal by exemplar-based image inpainting. *Trans. Img. Proc.*, 13(9):1200–1212, September 2004.
- [9] Radhakrishna Achanta, Appu Shaji, Kevin Smith, Aurelien Lucchi, Pascal Fua, and Sabine Susstrunk. Slic superpixels compared to state-of-the-art superpixel methods. *IEEE Trans. Pattern Anal. Mach. Intell.*, 34(11):2274–2282, November 2012.
- [10] Connelly Barnes, Eli Shechtman, Adam Finkelstein, and Dan B Goldman. Patchmatch: A randomized correspondence algorithm for structural image editing. *ACM Trans. Graph.*, 28(3):24:1–24:11, July 2009.
- [11] Patrick Pérez, Michel Gangnet, and Andrew Blake. Poisson image editing. In *ACM SIGGRAPH 2003 Papers*, SIGGRAPH '03, pages 313–318, New York, NY, USA, 2003. ACM.
- [12] Aude Oliva and Antonio Torralba. Modeling the shape of the scene: A holistic representation of the spatial envelope. *Int. J. Comput. Vision*, 42(3):145–175, May 2001.
- [13] Matthijs Douze, Hervé Jégou, Harsimrat Sandhawalia, Laurent Amsaleg, and Cordelia Schmid. Evaluation of gist descriptors for web-scale image search. In *Proceedings of the ACM International Conference on Image and Video Retrieval*, CIVR '09, pages 19:1–19:8, New York, NY, USA, 2009. ACM.
- [14] Siddheswar Ray and Rose H. Turi. Determination of number of clusters in k-means clustering and application in colour segmentation. In *The 4th International Conference on Advances in Pattern Recognition and Digital Techniques*, pages 137–143, 1999.
- [15] J. Brejcha and M. Čadík. Geopose3k: Mountain landscape dataset for camera pose estimation in outdoor environments, 2017. *Submitted to: Image Vision and Computing, Elsevier*.