

Bakalářská práce



České
vysoké
učení technické
v Praze

F3

Fakulta elektrotechnická
Katedra kybernetiky

Klíčování na barevné pozadí pomocí GPU

Roman Galajda

Vedoucí: Ing. Roman Berka, Ph.D.

Obor: Informatika a počítačové vědy

Studijní program: Otevřená informatika (bakalářský)

Květen 2017

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

Student: Roman Galajda
Studijní program: Otevřená informatika (bakalářský)
Obor: Informatika a počítačové vědy
Název tématu: Klíčování na barevné pozadí pomocí GPU

Pokyny pro vypracování:

Zhodnoťte algoritmy pro klíčování na barevné pozadí s pomocí GPU implementované v rámci předchozího semestrálního projektu a navrhněte možná vylepšení. Zaměřte se především na metodu klasifikace pozadí a na metodu propagace barev pozadí. Zabývejte se rovněž precizováním kritérií pro hodnocení kvality výstupu klíčovacího algoritmu. Navrhněte pak aplikaci umožňující počáteční nastavení a následné dávkové spuštění klíčovacího procesu. Hotovou aplikaci otestujte a zdokumentujte.

Seznam odborné literatury:

- [1] YAMNITSKY, David: Real-Time Chroma Keying on the GPU. White paper. Boris FX, Boston USA, 2009.
- [2] DEBEVEC, Paul, WENGER, Andreas, TCHOU, Chris, GARDNER, Andrew, WAESE, Jamie, HAWKINS, Tim: A Lighting Reproduction Approach to Live-Action Compositing. ACM Trans. Graphics (Proc. ACM SIGGRAPH 2002) 21(3), 547-556 (2002).
- [3] ŠOUREK, Peter: Kontrola barevné korony na zeleném pozadí, Bakalářská práce, ČVUT FEL, 2014.

Vedoucí bakalářské práce: Ing. Roman Berka, Ph.D.

Platnost zadání: do konce letního semestru 2017/2018

L.S.

prof. Dr. Ing. Jan Kybic
vedoucí katedry

prof. Ing. Pavel Ripka, CSc.
děkan

V Praze dne 1. 3. 2017

Poděkování

Chtěl bych poděkovat panu Ing. Romanu Berkovi, Ph.D. za vstřícný přístup k vedení práce, za cenné rady a poskytnuté materiály.

Prohlášení

Prohlašuji, že jsem předloženou práci vypracoval samostatně a že jsem uvedl veškeré použité informační zdroje v souladu s Metodickým pokynem o dodržování etických principů při přípravě vysokoškolských závěrečných prací.

V Praze dne 26. května 2017

.....

Abstrakt

Tato práce se zabývá problematikou klíčování na barevné pozadí. Popisuje současné metody klíčování a představuje nový způsob řešení klíčovací rovnice využívající výpočetních schopností dnešních grafických karet. Dále poukazuje na některé praktické problémy při klíčování na barevné pozadí, jako je barevná korona a nedokonalosti pozadí a navrhuje jejich řešení. Výsledky celého algoritmu jsou nakonec posouzeny a porovnány s referenčním řešením.

Klíčová slova: Klíčování na pozadí, GPU

Vedoucí: Ing. Roman Berka, Ph.D.

Abstract

This thesis focuses on the technique of chroma keying. It describes current methods of keying and brings a new way of solving the keying equation that leverages the compute capabilities of recent graphics cards. It also points out several practical problems of chroma keying, such as color spill and background imperfections and suggests possible solutions. Finally, the results of the algorithm are evaluated and compared to a reference solution.

Keywords: Chroma Keying, Alpha Matting, GPU

Title translation: Chroma Keying on the GPU

Obsah

| | | | |
|----------------------------------------|-----------|-----------------------------------------|-----------|
| 1 Úvod | 1 | 4.2 Odhad barvy pozadí | 14 |
| 1.1 Klíčování na pozadí | 1 | 4.3 Potlačení barevné korony | 15 |
| 1.2 Barevná korona | 2 | 4.4 Barva popředí a kompozice | 16 |
| 1.3 Motivace | 2 | 5 Výsledky a vyhodnocení | 17 |
| 2 Současná řešení | 3 | 5.1 Subjektivní hodnocení | 17 |
| 2.1 Blue screen matting | 3 | 5.2 Objektivní hodnocení | 21 |
| 2.2 Chroma keying | 4 | 6 Závěr | 25 |
| 2.3 Color Spill | 5 | A Literatura | 27 |
| 3 Vlastní řešení | 7 | B Obsah příloženého CD | 29 |
| 3.1 Problematika klíčování na pozadí | 7 | | |
| 3.2 Řešení klíčovací rovnice | 9 | | |
| 3.3 Problém barevné korony | 10 | | |
| 3.4 Odhad barvy pozadí | 11 | | |
| 3.5 Shrnutí | 12 | | |
| 4 Implementace | 13 | | |
| 4.1 Struktura a přehled | 13 | | |

Obrázky

| | |
|------------------------------------------|----|
| 5.1 Subjektivní test - Větrník | 18 |
| 5.2 Subjektivní test - Trol | 19 |
| 5.3 Subjektivní test - Kolo | 20 |
| 5.4 Objektivní test - Kolo | 22 |
| 5.5 Objektivní test - Kočka | 22 |
| 5.6 Objektivní test - Brýle | 23 |
| 5.7 Objektivní test - Kouř | 23 |
| 5.8 Objektivní test - Voda | 24 |

Tabulky

| | |
|-----------------------------------------------------------|----|
| 5.1 Výsledky měření na testovacích obrázcích | 24 |
|-----------------------------------------------------------|----|

Kapitola 1

Úvod

1.1 Klíčování na pozadí

Klíčování na pozadí (anglicky image matting, alpha matting nebo také chroma keying) je ve filmovém a televizním průmyslu často využívanou technikou pro přenos popředí jednoho obrázku, například herce, do jiného obrázku, třeba předem natočené scény. Řešeným problémem je rozdělení popředí a pozadí vstupního obrázku a vytvoření takzvané alpha mapy, která zachycuje úroveň průhlednosti popředí. Problém se rozděluje na dvě úlohy podle vlastností pozadí:

V případě neznámého, obecného pozadí se jedná o takzvaný natural image matting, kdy máme k dispozici často jen částečnou informaci od uživatele o tom, které části obrázku jsou pozadím a které popředím. Uživatel tuto informaci zadává vytvořením takzvané trimapy, která klasifikuje jednotlivé oblasti obrázku jako popředí, pozadí, nebo neznámou část. Natural image matting algoritmus pak využívá barvy ze známých oblastí obrázku a snaží se s jejich pomocí zpřesnit neznámé oblasti. Takovéto algoritmy jsou ale většinou příliš náročné a vytvoření kvalitní trimapy může vyžadovat velké množství času.

Pokud ale víme, že barva pozadí je konstantní nebo téměř konstantní, například při natáčení na zeleném či modrém klíčovacím plátně, pak se stává úloha oproti natural image mattingu jednodušší. Se znalostí barvy pozadí lze snadno klasifikovat oblasti obrázku automaticky, bez pomoci trimapy. Lze také

předpokládat znalost barvy pozadí i v částech obrázku částečně zakrytých popředím. Obecně tedy algoritmy řešící tuto úlohu dosahují srovnatelných nebo i lepších výsledků, a to rychleji a s menším stupněm interakce uživatele. Z těchto důvodů se v praxi využívá spíše tato metoda, kdy je to možné. Nevýhodou je totiž omezení při natáčení, kdy musí být přítomno klíčovací plátno, správné osvětlení a popředí, které je v barevném kontrastu s klíčovací barvou.

V této práci se zabýváme výhradně problematikou klíčování na známé barevné pozadí.

1.2 Barevná korona

Při nevhodném osvětlení, nebo pokud jsou objekty příliš blízko klíčovacímu plátnu, může vzniknout efekt barevné korony (v anglické literatuře označován jako *color spill*). Barevná korona je projev odrazu světla od barevného klíčovacího plátna, které dopadá na objekt v popředí. Takto odražené světlo je od klíčovacího plátna zbarvené jeho barvou a tím obarví i popředí. Nejen, že je na výsledném popředí toto zbarvení nežádoucí, ale také komplikuje samotné klíčování. U barvy, která je zbarvena do zelena, v případě zeleného klíčovacího plátna, lze těžko rozlišit, jestli zbarvení pochází z průhlednosti popředí, nebo jestli se jedná o barevnou koronu. V praxi je tedy důležité se vzniku barevné korony vyhnout, ale ne vždy je to možné. Z tohoto důvodu je řešení problému barevné korony nedílnou součástí klíčovacích algoritmů.

1.3 Motivace

Cílem této práce je navrhnout dosud neprozkoumaný postup, který spolehlivě řeší problém klíčování na předem známé barevné pozadí a omezuje výskyt barevné korony, a realizovat ho s využitím akcelerace pomocí GPU. Řešení by mělo být robustní vůči nedokonalostem v odstínu klíčovacího plátna a mělo by správně a spolehlivě odhadnout, zda se v částech obrázku nachází pozadí, popředí, průhledné popředí nebo popředí zasažené barevnou korunou. Výsledkem by měla být mapa průhlednosti, která zachycuje co největší část objektu v popředí, bez toho, aby obsahovala pozadí obrázku. Výsledná barva popředí by měla být zbavena barevného odstínu způsobeného barevnou korunou a ani by neměla být jinak nepřírozně zbarvena.

Kapitola 2

Současná řešení

2.1 Blue screen matting

Klíčování na pozadí má svůj původ v práci Petro Vlahose. Ve svých patentech [5], [6] zavedl problém takzvaného bluescreen mattingu a popsal jeho řešení. Tato metoda je založena na klíčování na modrém klíčovacím plátně, které bylo hojně využíváno ve filmovém průmyslu. Řešení dokázal najít omezením problému několika podmínkami, známé jako Vlahosovy předpoklady. Hlavním z předpokladů je, že barvy v popředí mají více zeleného odstínu, než modrého. Toho využíval jednak k samotnému klíčování, tak i k omezení barevné korony. Přes svoji jednoduchost byl s touto metodou velmi úspěšný a za svojí práci v roce 1995 získal ocenění Oscara.

Na práci Vlahose navázali Smith & Blinn [4], kteří se pokoušeli problém klíčování na modré pozadí vyřešit bez uživatelského vstupu. Použili k tomu další předpoklady, které ovšem možné barvy popředí značně omezují. Prvním z nich je předpoklad, že popředí neobsahuje žádnou modrou barvu. V tom případě lze modrou komponentu barvy použít přímo jako indikátor průhlednosti popředí. Tato metoda ale vyřazuje naprostou většinu barev v popředí, jako jsou odstíny šedi a dvě třetiny všech odstínů. Další Smithova metoda využívá předpokladů podobných těm, které použil Vlahos. Tentokrát ukázal, že lze problém vyřešit pro všechny odstíny šedi a velkou část barevných odstínů, mezi které se řadí i barva kůže.

Poslední řešení, které Smith uvádí, pojmenovává jako triangulaci. Jedná se o velmi obecné řešení a neklade žádné předpoklady ohledně barvy popředí. Nevýhodou je, že vyžaduje, aby byl objekt v popředí snímán před dvěma různými barevnými pozadími. Tím ale dokáže přesně a věrohodně vyřešit problém klíčování na pozadí a zároveň se efektivně zbavit barevné korony.

2.2 Chroma keying

Chroma keying je následovníkem blue screen mattingu a zobecňuje barvu pozadí z čistě modré na jakýkoliv odstín pozadí. Narozdíl od natural image mattingu se ale pořád předpokládá, že barva klíčovacího plátna je stále konstantní. S příchodem digitálního videa modrou barvu pozadí zaměnila barva zelená, na kterou jsou kamery citlivější a dokáží tak vytvořit přesnější výsledky.

S nejnovější technologií přišla společnost Disney [7], která se snaží omezit uživatelský vstup a čas věnovaný hledáním optimálních parametrů pro klíčování na pozadí a ručním domalovávání špatně vyhodnocených detailů. Autoři uvádí, že klíčování v profesionální kvalitě je proto stále nevyřešeným problémem. Také poukazují na důležitost kvality nejen získané průhlednosti popředí, ale i barevného odstínu popředí.

Její metoda se zakládá na takzvaném color unmixingu. Ta tvoří obecnější úlohu, kde se vstupní barva může skládat z lineární kombinace libovolného počtu dílčích barev. Dílčí barvy jsou definovány za pomoci interakce uživatele a aplikace, které spolu vypočítávají pravděpodobnostní model jednotlivých barev. Tuto úlohu pak řeší minimalizací "energie", kde barvy s větší předvěrohodností výskytu jsou preferovány. Tento popis je velmi zjednodušený a nezachycuje komplexitu této metody. I přes klíčování na známém pozadí tuto metodu nelze použít v real-time situacích.

Real-time metodu klíčování využívající GPU popisuje Yamnitsky [8]. Ten rozděluje svůj algoritmus na čtyři dílčí části. Nejdříve vytvoří prvotní klíč popředí podle odstínu barev. S tou pak manipuluje oříznutím pozadí a aplikuje filtr pro odstranění šumu. Následně omezí projev barevné korony. Zajímavý je poslední krok, kdy se pokouší naopak vytvořit barevnou koronu rozmazáním cílového pozadí podél okrajů popředí. Svůj algoritmus ani výsledky ale nezveřejňuje.

Dalším zajímavým řešením je použití klíčovací barvy mimo viditelné spektrum [9], [2]. Pro účely klíčování používají kamerový systém, který pomocí filtrů kromě viditelného světla současně zachycuje i světlo infračervené. Jako klíčovací plátno museli najít materiál, který jednak dobře odráží infračervené světlo, a zároveň málo odráží viditelnou část spektra – tedy je černé. Tím se zbavili jakýchkoli omezení barev v popředí a zároveň eliminovali výskyt nežádoucí barevné korony. Jejich dobré výsledky jsou ale vyváženy obtížností přípravy a praktické využitelnosti.

2.3 Color Spill

Ve filmovém průmyslu je problém barevné korony většinou řešen úpravou osvětlení tak, aby k efektu nedocházelo. Ne vždy je to ale možné nebo praktické, obzvláště pro běžného uživatele ve svém domácím studiu.

Potlačení barevné korony až po natočení snímku je možné několika způsoby. Některé z nich vyžadují podrobné informace scény, jako je rozložení světla a klíčovacího plátna. S jednodušším způsobem přišel už Vlahos [6] a Smith & Blinn [4] pro odstíny bílé a barvu kůže pomocí metody omezení zeleného kanálu. Klíčování na pozadí se ve své bakalářské práci zabýval i pan Peter Šourek [3]. Jedno z jeho řešení využívá pro detekci korony odchylku odstínu od klíčovací barvy v barevném prostoru HSV s velmi dobrými výsledky.

Kapitola 3

Vlastní řešení

3.1 Problematika klíčování na pozadí

Separaci barvy popředí a pozadí závisí na průsvitnosti popředí a popisuje ji tato klíčovací rovnice [4]:

$$C = C_f\alpha + C_b(1 - \alpha) \quad (3.1)$$

Kde C je vstupní barva z klíčovacího obrázku, C_f je barva popředí, C_b je barvou pozadí a α je neprůhlednost popředí. Tedy pokud je $\alpha = 1$, popředí je neprůhledné a C neobsahuje žádnou barvu pozadí. Naopak pokud $\alpha = 0$, popředí je plně průhledné a je rovno barvě pozadí. Platí $0 \leq \alpha \leq 1$.

Pokud jsou C_f a C_b známé, tak pro přenos na popředí druhého obrázku s barvou C_m platí vztah:

$$C_o = C_f\alpha - C_b(1 - \alpha) + C_m(1 - \alpha) \quad (3.2)$$

Řešením je výsledná barva klíčování C_o .

Pro klíčování se známou barvou pozadí C_b lze vyjádřit barva popředí C_f se závislostí na parametru α :

$$C_f = \frac{C - C_b(1 - \alpha)}{\alpha} \quad (3.3)$$

Z praktických důvodů je ale výhodnější vyjádřit $C_f\alpha$ [4]. Barva popředí již předem vynásobená α lze ihned použít pro kompozici a je dobře definovaná i pro $\alpha = 0$. Budeme ji tedy dále používat místo skutečného popředí a zavedeme modifikovanou barvu popředí $C_p = C_f\alpha$. Pro tuto definici popředí vypadá klíčovací rovnice takto:

$$C = C_p + C_b(1 - \alpha) \quad (3.4)$$

Z toho jednoduše odvodíme vztah:

$$C_p = C - C_b(1 - \alpha) \quad (3.5)$$

Předpokládejme, že barva klíčovaného obrázku C a barva pozadí C_b jsou nám známy. Tato rovnice má tedy nekonečně mnoho řešení v závislosti na neznámém parametru α v intervalu:

$$\{\alpha \in R \mid 0 \leq \alpha \leq 1\} \quad (3.6)$$

Při odhadu parametru ovšem může nastat problematický případ, kdy neprůhlednost je nízká a vstupní barva C a barva pozadí C_b si nejsou rovny. V tom případě barva popředí $C_p \approx C - C_b$ vychází nenulově, nebo dokonce záporně pro $C_b > C$. Takové řešení by v striktním případě nemělo být platné, stejně tak pokud neplatí:

$$0 \leq C_f \leq 1; C_f = \frac{C_p}{\alpha} \quad (3.7)$$

I přesto takové řešení může být vyhovující, pokud uznáme, že barevný prostor vstupu nemusí být lineární. Přesněji, že pro zobrazení $f(C)$ do prostoru barev, ve kterém je vstup uchovávan, a pro skutečné barvy a, b platí:

$$f(a + b) \leq f(a) + f(b) \quad (3.8)$$

Klíčovací rovnice za tohoto předpokladu vypadá takto:

$$C_p = f(f^{-1}(C) - f^{-1}(C_b)(1 - \alpha)) \quad (3.9)$$

V ideálním případě by toto zobrazení i jeho inverze, tedy $f^{-1}(C)$, byly známy. Poté by se klíčovací rovnice řešila v původním (lineárním) prostoru barev, kde by podmínka pro barvu popředí mohla být uplatňována. Vstup pro klíčování ovšem většinou pochází z kamer, kde je odhad zobrazení zpátky do tohoto prostoru barev složitý, už jen kvůli přidanému šumu a nepřesnostem komprese.

Ve většině případů zobrazení ani jeho inverzi neznáme, vstup je nepřesný nebo nedokážeme přesně odhadnout barvu pozadí. Poté se musíme s nelinearitou nějak vypořádat. Pro případ, kdy $\alpha = 1$, platí $C_p = f(f^{-1}(C)) = C$. S klesajícím parametrem α se už ale nelinearity a nepřesnosti začínají projevovat a chyba stoupá. Z tohoto důvodu lze z odhadované neprůhlednosti označit nejistotu jako $\sigma = 1 - \alpha$. Naším cílem je tedy minimalizovat nejistotu a tím pádem maximalizovat neprůhlednost α .

3.2 Řešení klíčovací rovnice

Navrhované řešení klíčovací rovnice spočívá ve správném odhadu barvy popředí a ve známosti barvy pozadí ve všech bodech vstupního obrázku. Popis procesu, jak pozadí odhadnout, je uveden v sekci 3.4. Víme, že klíčovací rovnice má nekonečně mnoho řešení, a parametr α chceme maximalizovat v intervalu $[0, 1]$. Pro klíčování na pozadí se ale nespokojíme s jakoukoli výslednou barvou popředí. Definujeme tedy podmínku, kterou musí barva popředí splňovat:

$$H_p(C_p) \in \{0, 1\} \quad (3.10)$$

Kde $H_p(C_p) = 1$ pokud je barva popředí vyhovující, $H_p(C_p) = 0$ pokud ne. Poté lze vyřešit klíčovací rovnici jako:

$$\begin{aligned} \max \alpha \\ H_p(C - C_b(1 - \alpha)) &= 1 \\ 0 \leq \alpha \leq 1 \end{aligned} \quad (3.11)$$

Jinými slovy, hledáme maximální možnou neprůhlednost, pro kterou dostaneme takovou barvu popředí, která splňuje zadanou podmínku.

Abychom toto řešení mohli nalézt efektivně numericky, hodí se, aby byla funkce $H_p(C - C_b(1 - \alpha))$ v závislosti na parametru α monotónní. Poté lze najít řešení například pomocí metody půlení intervalů.

3.3 Problém barevné korony

Barevná korona je projev světla zbarveného klíčovací barvou po odrazu od klíčovacího plátna, které dopadá na objekt v popředí. Výslednou barvu můžeme definovat jako kombinaci přímého světla, světla odraženého od klíčovacího plátna a odrazivost povrchu v popředí. Pokud bereme v úvahu pouze částí světla které se přímo podílí na obrazu, tedy celkové světlo odražené z povrchu v popředí a dopadající na kameru, lze rovnici zjednodušeně popsat takto:

$$C = C_a L + C_a L_k s \quad (3.12)$$

Kde C je vstupní barva, C_a je odrazivost povrchu popředí, L_k je barva světla odraženého od klíčovacího plátna, L je intenzita světla osvětlující scénu a s je parametr určující, kolik světla odraženého z klíčovacího plátna dopadá na povrch. Pro zjednodušení budeme uvažovat, že $L = 1$, tedy že je pro osvětlení použito bílé světlo konstantní intenzity. Dostáváme tedy:

$$C = C_a + C_a L_k s \quad (3.13)$$

Pro potlačení korony chceme vyjádřit C_a :

$$C_a = \frac{C}{(1 + L_k s)} \quad (3.14)$$

Barva světla odraženého od klíčovacího plátna L_k může být přímo zadána uživatelem jako barva klíčovacího plátna. Klíčovací plátno je také osvětleno světlem scény L , ale znovu přepokládáme, že $L = 1$. Tím dostáváme pro parametr s rovnici s nekonečně mnoha řešeními. Pokud chceme úlohu řešit podobným způsobem jako klíčovací rovnici, potřebujeme další podmínku $H_a(C_a) \in \{0, 1\}$, která nám určí, které výsledné barvy C_a jsou přijatelné a které naopak nežádoucí. Naopak od řešení klíčovací rovnice budeme ale parametr s minimalizovat, protože nepřesnost naopak s roustoucím s stoupá.

Další rozdíl oproti řešení klíčovací rovnice je, že parametr s může být v intervalu $[0, \infty]$. Pro $s = 0$ na popředí nedopadá žádné světlo odražené z klíčovacího plátna a efekt barevné korony se zde tedy nevyskytuje. Pro $s = 1$ jsou si světlo scény a odražené světlo sobě rovny. Se zvětšujícím s se poměr světél dále zvětšuje a odražené světlo se stává silnější než světlo scény.

Při současném používání klíčovacího algoritmu a algoritmu na potlačení barevné korony je důležité rozlišit, do jaké míry v barvě C je přítomnost klíčovací barvy způsobena průhledností popředí a do jaké míry jde o efekt barevné korony. Pokud bychom parametr s nijak neomezili, pak by veškerá přítomnost klíčovací barvy byla připisována efektu barevné korony a celý obrázek by byl poté označen jako popředí. Hodí se tedy zavést uživatelský parametr s_{\max} , který možné hodnoty parametru s zhora omezí. Takový parametr lze chápat jako předpoklad, že poměr světla odraženého od klíčovacího plátna a světla scény nepřesáhne určitou hodnotu.

Potlačení korony lze tedy provést vyřešením:

$$\begin{aligned} \min s \\ H_a\left(\frac{C}{(1-L_k s)}\right) &= 1 \\ 0 \leq s \leq s_{\max} \end{aligned} \quad (3.15)$$

Pro monotónní $H_a\left(\frac{C}{(1-L_k s)}\right)$ lze opět nalézt řešení v intervalu $s \in [0, s_{\max}]$ numericky.

3.4 Odhad barvy pozadí

Výše uvedená metoda klíčování na pozadí v první řadě vyžaduje znalost barvy pozadí ve všech bodech vstupního obrázku. V některých případech, kdy je barva pozadí velmi rovnoměrná, lze využít klíčovací barvu jako barvu pozadí, tedy $C_b = C_K$. Často ale v praxi barva klíčovacího plátna není všude stejná, například z důvodu různé úrovně osvětlení a stínů. Proto je vhodné lokální barvu pozadí vhodným způsobem odhadnout.

Princip odhadu barvy pozadí je, že odstíny, které jsou součástí pozadí, jsou přítomny ve velké části obrázku, narozdíl od průhledných detailů objektu v popředí. Způsob, který zde předkládám, z tohoto důvodu pro odhad pozadí využívá 2D mediánový filtr, který zachovává hodnotou většiny a není ovlivněn extrémními hodnotami.

Nejdříve si definujme další klasifikační funkci, tentokrát pro barvu pozadí: $H_b(C) \in \{0, 1\}$. Ta určuje, jestli daná barva je pozadím či nikoliv. Barvy, pro které platí $H_b(C) = 1$, můžeme označit jako barvu pozadí, tedy $C_b = C$. Poté aplikujeme mediánový filtr na obrázek tak, že výsledkem pro každý pixel je medián barev pixelů v jeho blízkosti, pro které platí $H_b(C) = 1$.

Tím dostaneme dobrý odhad barvy pozadí i v oblastech obrázku částečně zakryté popředím. Další výhodou tohoto postupu je, že i na barvy, které jsou klasifikovány jako pozadí, je aplikován filtr. Pokud je tedy pixel ve skutečnosti drobným detailem popředí s velkou průhledností, například konečky vlasů, není jeho barva použita jako barva pozadí, i když tak může být klasifikována, protože neprojde mediánovým filtrem.

Tento filtr je aplikován několikrát za sebou, dokud není pro všechny částečně průhledné oblasti obrázku vytvořen odhad barvy pozadí. Samotný mediánový filtr je aplikován zvlášť pro každou složku barvy, tedy pro červenou, modrou a zelenou. Nejde tedy striktně o mediánový filtr, ale tento způsob stále zachovává jeho kvality.

3.5 Shrnutí

Celý navrhovaný algoritmus probíhá tak, že ze vstupního obrázku je vytvořena takzvaná mapa pozadí procesem ze sekce 3.4. Tato mapa obsahuje barvu pozadí minimálně v oblastech, kde je očekáván výskyt částečně průhledného popředí. Poté je ze vstupního obrázku i z mapy pozadí odstraněna barevná korona způsobem popsáním v sekci 3.3. Klíčovací proces předpokládá, že vstup neobsahuje barevnou koronu, a tedy je nutné ji odstranit ještě před klíčováním. Zároveň je třeba stejnou operaci provést i na mapě pozadí, aby odpovídala vstupu. Pro obrázek je poté vyřešena klíčovací rovnice 3.11 pro parametr α , kde C je barva obrázku po odstranění korony a C_b je barva z mapy pozadí.

Z parametru α odvodíme barvu popředí C_p pomocí vztahu 3.5. Výsledná barva je pak vytvořena kombinací barvy popředí C_p , parametru neprůhlednosti α a barvy cílového obrázku C_m vztahem:

$$C = C_p + C_m(1 - \alpha) \quad (3.16)$$

Kapitola 4

Implementace

Záměrem konzolové aplikace je vytvoření prostředí pro rychlé dávkové spuštění předloženého algoritmu a jednoduché testování jeho možností a dovedností. Aplikace je napsána v jazyce c++14 a využívá knihovnu ffmpeg pro načítání a zapisování video souborů a obrázků. Samotný algoritmus je implementován pomocí compute shaderů v jazyce GLSL, které jsou přeloženy do formátu SPIR-V a spuštěny na grafické kartě pomocí grafického rozhraní Vulkan. Pro spuštění této aplikace je tedy třeba, aby byla dostupná grafická karta s řadiči podporující Vulkan. Žádné z těchto technologií ale nejsou nezbytně nutné pro vlastní implementaci uvedeného řešení. Aplikace podporuje pouze rozhraní Windows s architekturou x64.

4.1 Struktura a přehled

Program je rozdělen do několika dílčích celků, které tvoří abstrakci kolem použitých knihoven. První z nich je třída `VulkanDevice`, která provádí potřebnou inicializaci řadiče a grafické karty, jako je hledání příkazových front a paměťových bloků. Třída `FrameProcessor` pak inicializuje vše potřebné pro zpracování jednoho snímku videa podle zadaných parametrů. Třída alokuje paměť, načítá shadery a tvoří příkazové buffery popisující příkazy pro zpracování snímku. Díky architektuře rozhraní Vulkan stačí příkazové buffery vytvořit pouze jednou a pak je opakovaně posílat k provedení na grafické kartě.

O to se také stará třída `VideoProcessor`, která má na starosti zpracování celého videa. Pro efektivní paralelní výpočet se jednotlivé snímky videa vkládají do fronty a jeden po druhém se na grafické kartě zpracovávají. Během toho může zbytek aplikace dekodovat další snímek ke zpracování nebo zapisovat snímek už zpracovaný.

Třídy `InputVideoStream` a `OutputVideoStream` využívají knihovnu `ffmpeg` pro načtení, respektive zapsání, jednoho snímku videa z či do souboru na disku. Třída `AlphaMatting` všechny tyto dílčí části kombinuje do jednoho celku, který efektivně zpracovává video soubory z disku. Obsahuje hlavní cyklus aplikace manipulující s frontou snímků. Od `VideoProcessor` zažádá o již zpracovaný snímek, zapíše jej na disk, na jeho místo načte další snímek ze vstupu a pošle ho ke zpracování na grafickou kartu.

Vstupem programu jsou dvě videa nebo obrázky a konfigurační soubor. První video obsahuje snímky ke klíčování, ze kterého chceme oddělit popředí. Druhé video nebo obrázek je pak pozadím, na které se má popředí přenést, takzvané *matte*. Konfigurační soubor obsahuje kromě volitelných parametrů i typ výstupu. Tím může být jednak klasický výstup algoritmu, tak i jednotlivých mezikroků, jako je odhadovaná mapa pozadí, mapa neprůhlednosti zachycující parametr α nebo vstupní obrázek po potlačení barevné korony. Tyto výstupy jsou obzvlášť užitečné pro zkoumání a testování uvedeného algoritmu.

Na straně grafického procesoru v průběhu programu běží tři compute shadery. První z nich klasifikuje pozadí a inicializuje tím mapu barev pozadí. Druhý shader pak slouží k propagaci barev pozadí. Poslední shader pak zajišťuje zbytek. Vstupní barva a barva *matte* jsou načteny a linearizovány převodem z barevného prostoru *sRGB*. Poté je aplikováno potlačení barevné korony na vstupní barvu a barvu pozadí z předchozího shaderu a řešením klíčovací rovnice je nalezena barva popředí a její průhlednost. Vše je nakonec zkombinováno a převedeno zpět do prostoru *sRGB*.

4.2 Odhad barvy pozadí

Prvním krokem odhadu barvy pozadí je inicializace mapy pozadí. Ta je zredukována na čtvrtinovou velikost vstupního obrázku pro rychlejší zpracování. Pro určení, zda je barva součástí pozadí je třeba definovat klasifikační funkci pozadí $H_b(C)$.

Všechny klasifikační funkce v této implementaci používají funkci založenou na kosinu úhlu mezi odstíny dvou barev:

$$\text{hue}_{\Delta}(C_a, C_b) = \frac{1}{2} - \frac{1}{2} \frac{\left(C_a - \frac{1}{3} \sum C_a\right) \cdot \left(C_b - \frac{1}{3} \sum C_b\right)}{\left\|C_a - \frac{1}{3} \sum C_a\right\| \left\|C_b - \frac{1}{3} \sum C_b\right\|} \quad (4.1)$$

Klasifikační funkce pro barvu pozadí je navíc ještě omezena saturací. Barva, která je velmi tmavá nebo velmi světlá, pravděpodobně není součástí klíčovacího plátna i když má podobný odstín. Saturací funkce je převzata z definice barevného prostoru HSV:

$$\text{sat}(C) = \max(C_r, C_g, C_b) - \min(C_r, C_g, C_b) \quad (4.2)$$

Kde C_r , C_g a C_b jsou příslušné barevné složky barvy C . Poté jako klasifikační funkci pozadí můžeme použít:

$$H_b(C) = \begin{cases} 1, & \text{pokud } \text{hue}_{\Delta}(C, C_K) \leq p_h \wedge \text{sat}(C) \geq p_s \\ 0, & \text{jinak} \end{cases} \quad (4.3)$$

Kde C_K je klíčovací barva, p_h je parametr určující maximální povolenou vzdálenost odstínu a p_s je parametr určující minimální povolenou saturaci barvy.

Barvám klasifikované jako pozadí je přiřazena váha 1, ostatním váha 0. Barvy spolu s váhami jsou pak pro každý 4x4 blok pixelů zprůměrovány a zapsány do mapy pozadí. Poté následuje propagace barev pozadí z oblastí s vyšší vahou do oblastí s vahou nižší. Jak bylo popsáno v sekci 3.4, pro tento účel je použit vážený mediánový filtr [1]. Takový filtr je ale příliš náročné počítat s velkým rozsahem, a tak ho aproximujeme několika aplikacemi menšího filtru. V našem případě je použit filtr o 9-ti vstupech ve tvaru kříže. Vážený medián o 9-ti hodnotách lze efektivně spočítat seřazením hodnot pomocí řadících sítí komparátorů (<http://pages.ripco.net/~jgambler/nw.html>) a následnou iterací přes seřazené váhy. Hodnota, pro kterou suma předchozích vah přesáhne polovinu sumy všech vah, je váženým mediánem pro tyto hodnoty.

4.3 Potlačení barevné korony

Dalším krokem je odstranění korony jak ze vstupní barvy, tak i ze získané barvy pozadí, aby nebyl porušen klíčovací vztah. Tento krok by bylo možné provést najednou ještě před propagací barvy pozadí, ale klasifikace pozadí by se poté stala těžší, jelikož se rozdíl mezi barvami popředí a pozadí po potlačení korony snižuje.

Funkce pro klasifikaci barvy korony využívá funkci $hue_{\Delta}(C_a, C_b)$ 4.1 vracející vzdálenost mezi odstíny dvou barev. Tu porovnává s pevně zadaným úhlem 90° , rozdělující prostor barev na dvě poloviny. Tento úhel dává subjektivně nejpřirozenější výsledky. Místo toho je volitelný parametr s_{\max} v řešeném vztahu 3.15.

$$H_a(C) = \begin{cases} 1, & \text{pokud } hue_{\Delta}(C, C_K) \leq \frac{1}{2} \\ 0, & \text{jinak} \end{cases} \quad (4.4)$$

Řešení hledáme numericky pomocí metody půlení intervalu $[0, s_{\max}]$. Tato metoda je vhodná pro použití na grafických kartách, protože vede k pevně danému počtu kroků a předvídatelné přesnosti a je jednoduchá na implementaci. Pro dostatečně přesný výsledek po linearizaci stačí 16 kroků metody půlení intervalů.

4.4 Barva popředí a kompozice

V tomto kroku algoritmus pomocí klíčovací rovnice odhaduje barvu popředí za pomoci celkové barvy C a odhadnuté barvy pozadí C_b . Předpokládáme, že obě tyto barvy už prošly procesem odstranění korony. I pro tento krok zavedeme klasifikační funkci $H_p(C)$ pro maximalizační úlohu 3.11 s parametrem p_p určující minimální povolenou vzdálenost od odstínu barvy pozadí.

$$H_p(C) = \begin{cases} 1, & \text{pokud } hue_{\Delta}(C, C_b) \geq p_p \\ 0, & \text{jinak} \end{cases} \quad (4.5)$$

I tato úloha je řešena numericky pomocí metody půlení intervalu $[0, 1]$ o 16 krocích.

Posledním krokem je samotná kompozice barvy popředí C_p , barva cílového pozadí C_m a průhlednost popředí α . Ten je proveden jednoduše výrazem 3.16:

$$C = C_p + C_m(1 - \alpha) \quad (4.6)$$

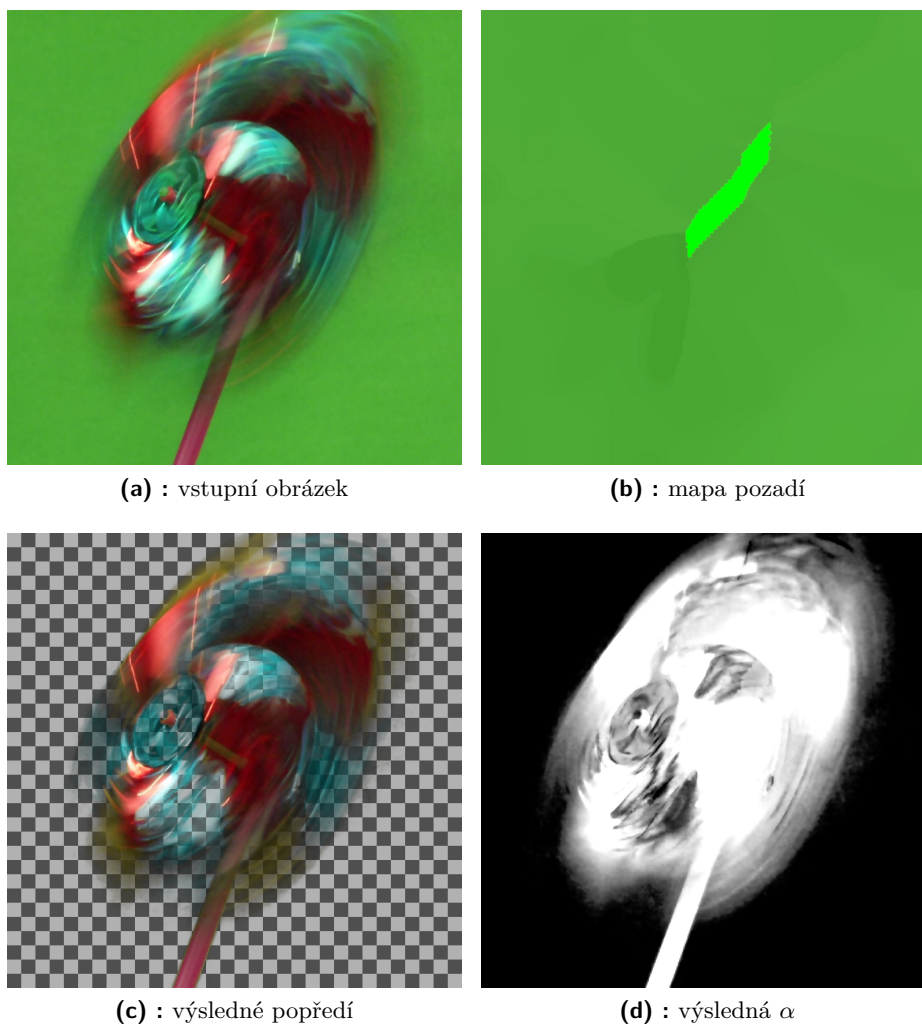
Kapitola 5

Výsledky a vyhodnocení

5.1 Subjektivní hodnocení

Vyvinutý program byl otestován na několika zákeřných příkladech vstupu, které dobře ilustrují jeho dovednosti a slabiny. Následující příklady jsou z důvodu chybějícího referenčního řešení hodnoceny pouze subjektivně. Pozornost je věnována hlavně nesprávné klasifikaci popředí a pozadí, což lze dobře pozorovat na mapě průhlednosti, respektive na mapě pozadí. Barevné odstíny, které jsou přeneseny na mapu pozadí se nebudou objevovat jako součástí popředí. Dále je sledována také barva výsledného popředí, která by se neměla příliš odlišovat od zbarvení objektu ve vstupním obrázku.

U každého příkladu je uveden krátký popis a slovní ohodnocení. Testovací snímky pocházejí z videí dodané panem Ing. Romanem Berkou, Ph. D. Příložen je vždy vstupní obrázek (vlevo nahoře), odhadnutá barva pozadí (vpravo nahoře), výsledné popředí zobrazené na čtverečkovaném pozadí (vlevo dole) a nakonec mapa parametru α (vpravo dole).



Obrázek 5.1: Subjektivní test - Větrník

Tento obrázek větrníku obsahuje velký rozsah hodnot průhledností v průběhu pohybu. Klíčování ztěžují nazelenalé odlesky větrníku, které algoritmus částečně klasifikoval jako průhledné. Pozadí je jednoduché a algoritmus nemá problém ho replikovat i za větrníkem, až na malý úsek, který je příliš daleko od vhodného pozadí, a tak je vyplněn klíčovací barvou.



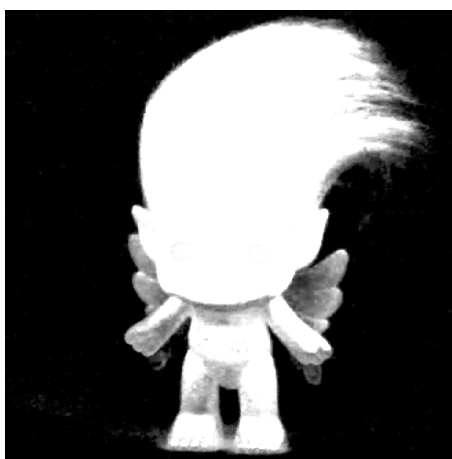
(a) : vstupní obrázek



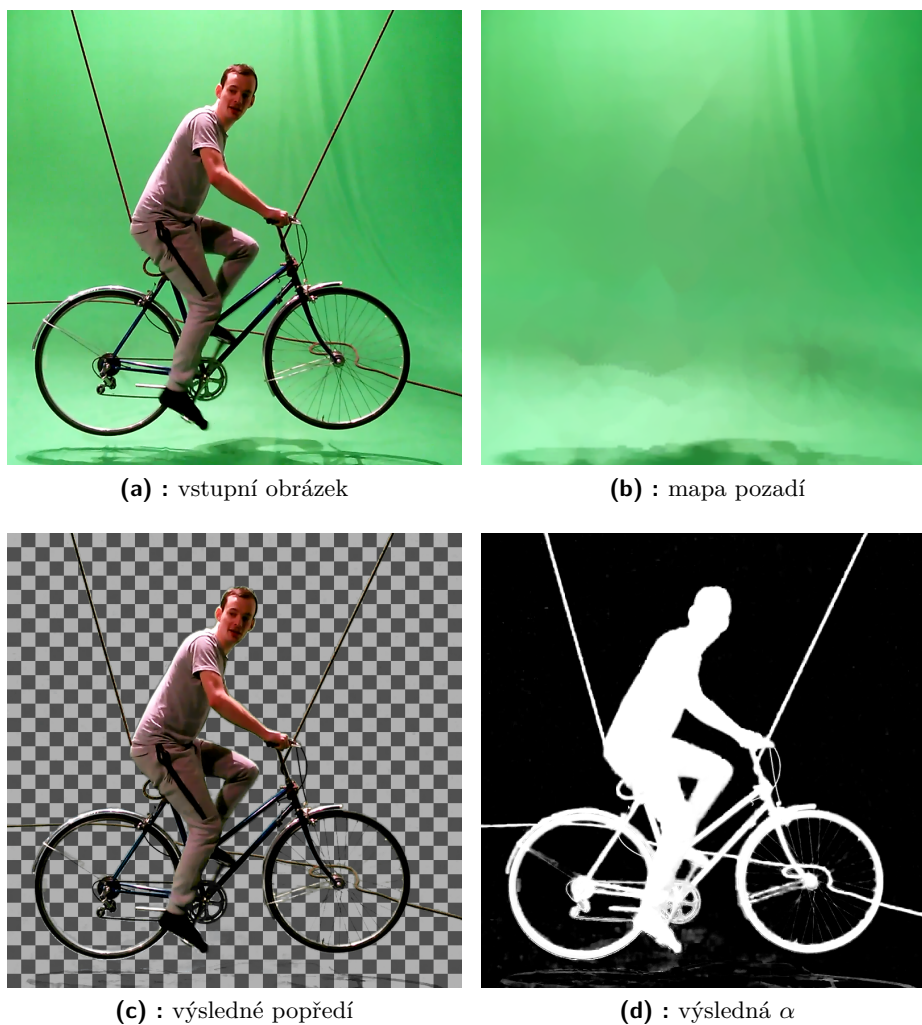
(b) : mapa pozadí



(c) : výsledné popředí

(d) : výsledná α **Obrázek 5.2:** Subjektivní test - Trol

Tentokrát je přítomno velké množství barevné korony, se kterou si algoritmus i přes ruční volbu parametrů nedokáže poradit. Průhledná křídla a vlasy figurky obsahují velmi podobné barvy jako krční oblasti zasažené barevnou korunou. Korona je sice úspěšně potlačena a barvy výsledného popředí nejsou zbarvené do zelena, ale místo toho jsou klasifikovány jako průhledné. Problém také dělá šum, který je obtížné barevně rozlišit od konečků vlasů. Pozadí je odhadnuto smysluplně. Silné stíny u nohou figurky nejsou součástí mapy pozadí, ale větší stíny v pozadí ano a jsou tedy z popředí odstraněny.

**Obrázek 5.3:** Subjektivní test - Kolo

Poslední příklad testuje hlavně schopnost algoritmu klíčovat na pozadí, které nemá konstantní barvu, ale obsahuje spoustu stínů a přesvětlených oblastí. Barva pozadí je ale i přesto odhadnuta správně a zakryté oblasti kopírují nedaleké odstíny pozadí. Algoritmus se snaží v pozadí zachytit i stín kola, ale část jeho detailů byla buď ztracena ve čtyřnásobném zmenšení, nebo neprošla skrze mediánový filtr. U nohou jezdce se také vyskytuje efekt barevné korony, tentokrát ale v menší míře, kterou algoritmus dokáže bez problému odstranit.

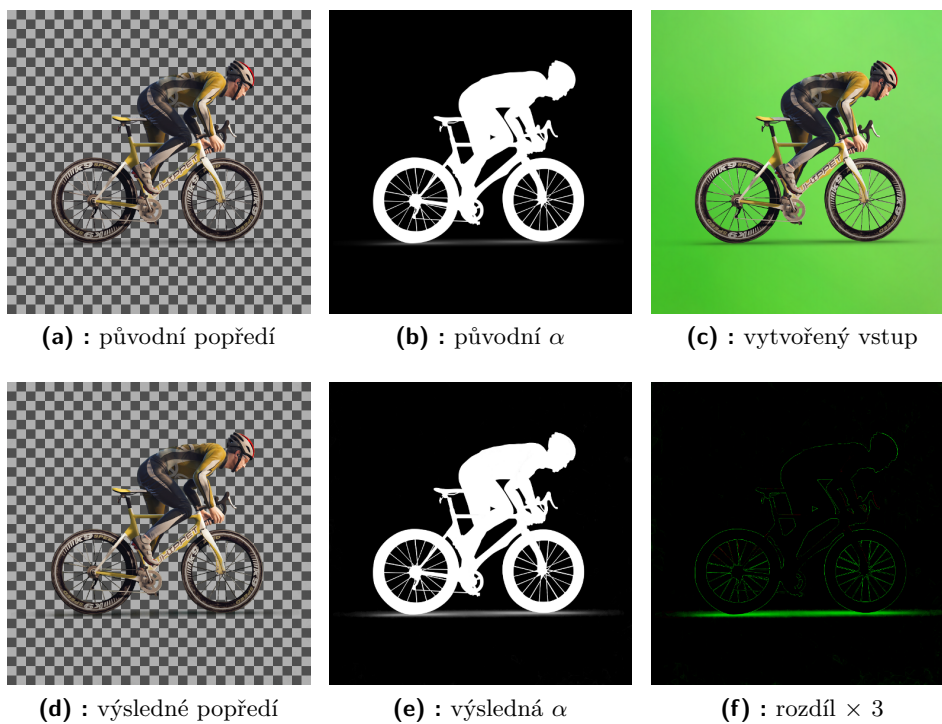
5.2 Objektivní hodnocení

Objektivní testování klíčovacích algoritmů je nelehký úkol. Protože jsou z naprosté většiny aplikovány na natáčená videa nebo fotografie, neexistuje pro ně žádné objektivně správné referenční řešení. Je tedy třeba taková referenční řešení nějakým způsobem vytvořit. Jedním z nich je snímání objektu před monitorem, se kterým je možné měnit pozadí bez změny v popředí. Poté lze získat dostatečně kvalitní referenční řešení pomocí zmíněné metody triangulace. Další metodou, která byla zvolena v této práci, je klíčování počítačem vytvořených scén se známou průhledností objektů v popředí.

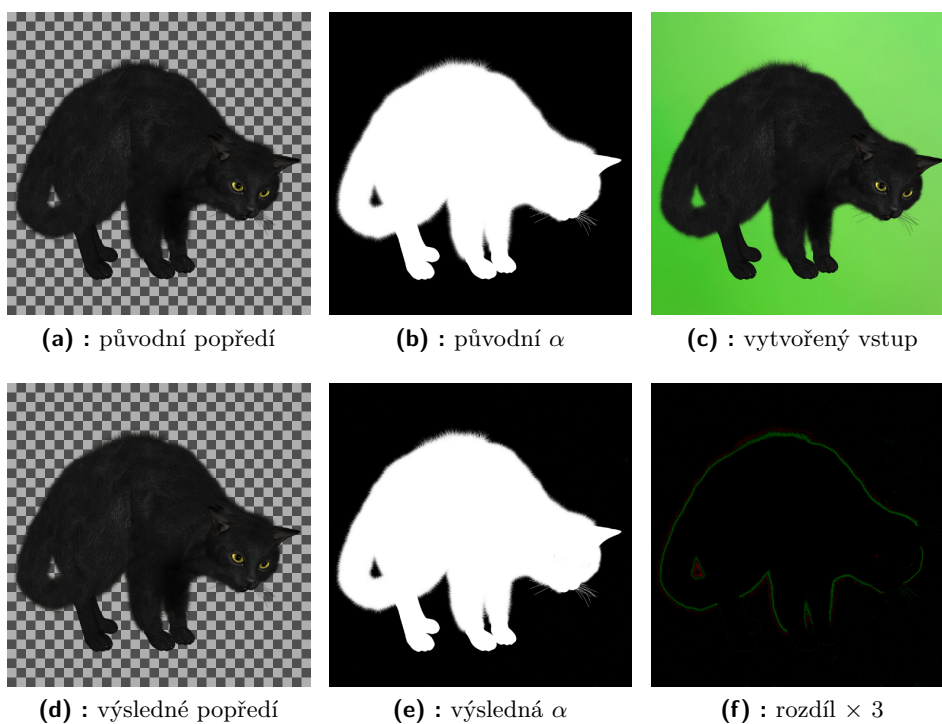
Problémem je také nedostupnost profesionálních klíčovacích programů se kterými lze nový algoritmus porovnat. Většina z nich nabízí zkušební verzi, ale ta je často limitována vodoznakem nebo nemožností exportovat snímek bez komprese. Tyto vady brání objektivnímu porovnání a test by tak nemohl být věrohodný. Výsledky jsou tedy porovnány pouze s referenčním řešením. I tak lze ale posoudit kvalitu výsledků.

Každý test obsahuje obrázek s mapou průhlednosti. Ty jsou získané z webové stránky <http://nobacks.com> a jsou použitelné pro nekomerční účely. Tyto obrázky jsou poté zkombinovány s vytvořeným pozadím s měnícími se odstíny zelené, simulující nedokonalosti osvětlení klíčovacího plátna. Výsledkem jsou mapy popředí a průhlednosti získané uvedeným algoritmem ze vstupního obrázku. Zobrazen je také rozdíl v parametru α oproti referenčnímu řešení. Zelená znamená, že výsledný parametr má vyšší hodnotu než referenční řešení, červená poukazuje naopak na hodnotu nižší. Čím tmavší je barva rozdílu, tím je řešení kvalitnější.

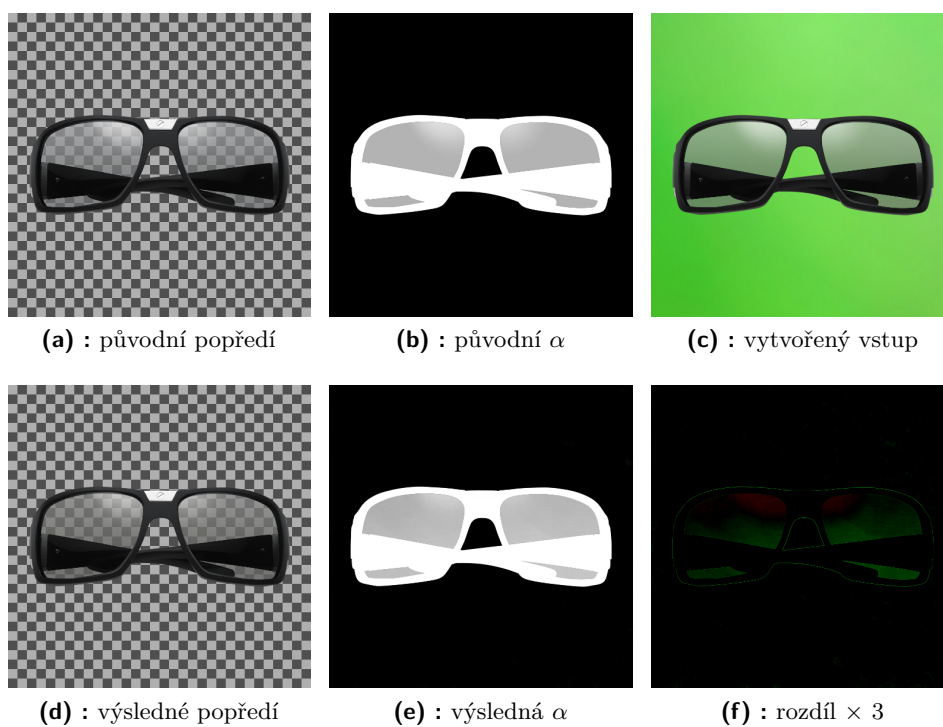
Testován je průměr druhých mocnin chyb parametru α (MSE) oproti referenčnímu řešení a čas, který algoritmus potřeboval pro vyřešení jednoho snímku. Čas byl změřen na grafické kartě NVIDIA GeForce GT 650M jako průměr pěti spuštění. Naměřené hodnoty jsou uvedeny v tabulce na konci spolu s rozměry obrázku.



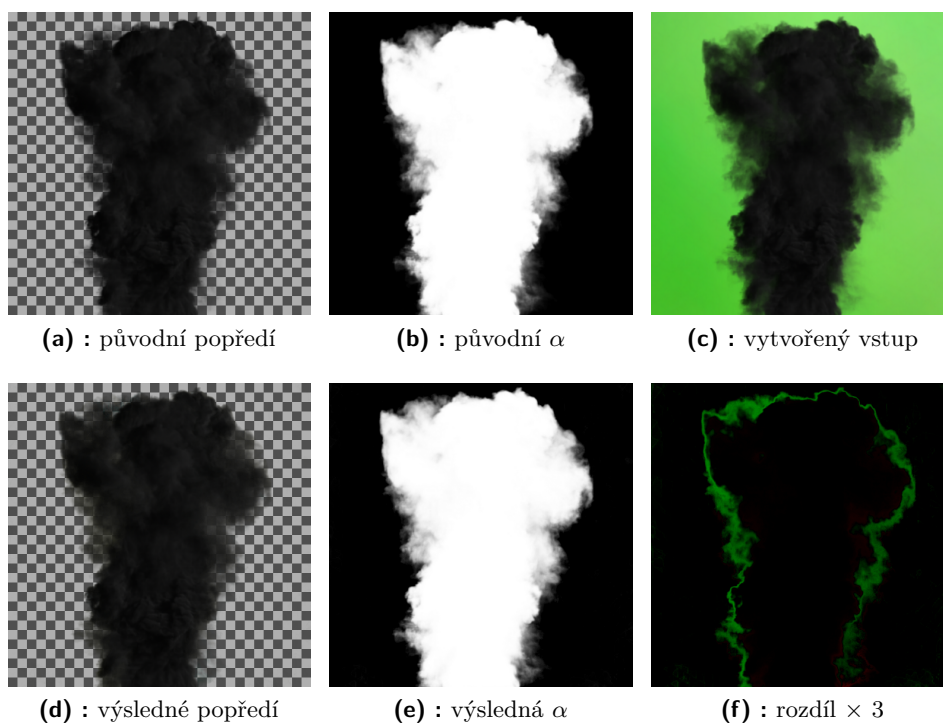
Obrázek 5.4: Objektivní test - Kolo



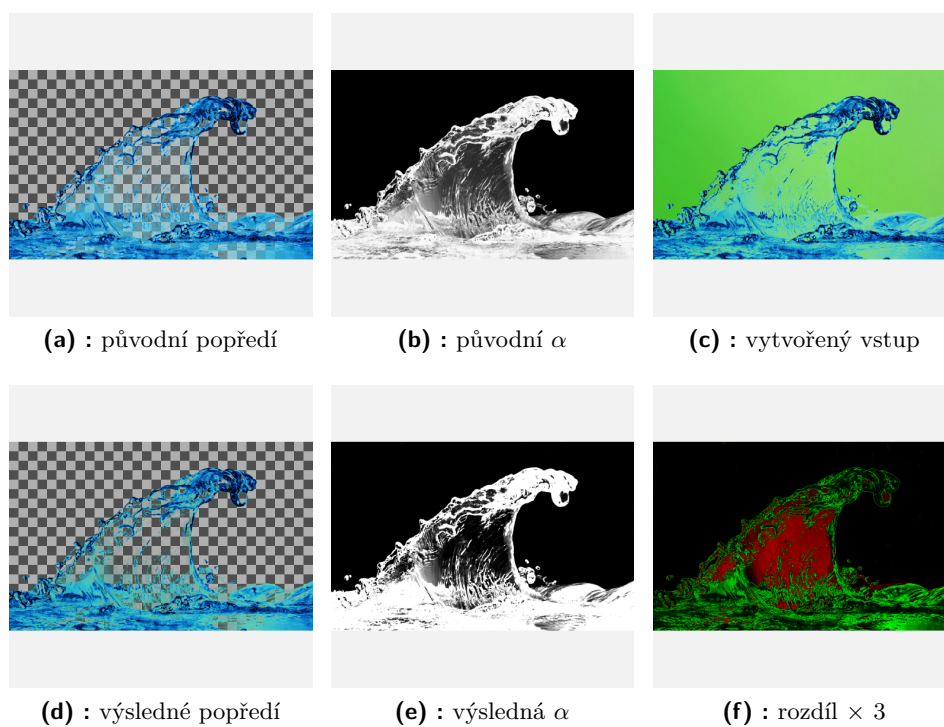
Obrázek 5.5: Objektivní test - Kočka



Obrázek 5.6: Objektivní test - Brýle



Obrázek 5.7: Objektivní test - Kouř



Obrázek 5.8: Objektivní test - Voda

| Příklad | Rozměry | MSE | Čas (s) |
|-----------|--------------------|-------|---------|
| 5.4 Kolo | 800 \times 800 | 0.196 | 0.11 |
| 5.5 Kočka | 1480 \times 1480 | 0.034 | 0.35 |
| 5.6 Brýle | 1000 \times 1000 | 0.06 | 0.17 |
| 5.7 Kouř | 660 \times 660 | 0.217 | 0.08 |
| 5.8 Voda | 1280 \times 798 | 2.462 | 0.17 |

Tabulka 5.1: Výsledky měření na testovacích obrázcích



Kapitola 6

Závěr

Byl předveden další ze způsobů řešení klíčovací rovnice a omezení projevu barevné korony. Výsledek se zdá být přijatelný a je vhodný pro praktické použití. Klasifikace barev pomocí jednoho uživatelem zadaného odstínu a vzdálenosti od něj je ale omezující jak pro algoritmus, tak i pro uživatele. Jednou z možností by bylo rozlišení barev pozadí pomocí “tahů štětcem” uživatele. Program by tak dostal k dispozici více barev pozadí i popředí a mohl by podle nich lépe vyhodnotit další, neznámé barvy.

V případě synteticky vytvořených testovacích obrázků bez nepřesností způsobených kamerou algoritmus pro většinu vstupů dosahuje velmi dobrých výsledků, ale i tak jsou oproti referenčnímu řešení zatíženy chybou. U reálných snímků je chyba větší z důvodu nelinearity, šumu a dalších nepřesností. Bylo by vhodné prozkoumat metodiky, jak se proti artefaktům ve vstupu bránit a učinit tak algoritmus robustnější. K tomu by mimo jiné mohla pomoci aplikace filtru na výslednou průhlednost podobným způsobem jako v [8], nebo zavedení dalších dodatečných úprav vstupů a výstupů.

Program ve své podobě není schopen pracovat v reálném čase kvůli náročné propagaci pozadí. Pro účely editace videa je ale dostatečně výkonný a zpracování jednoho snímku zabírá méně než vteřinu i při vysokém rozlišení a slabší grafické kartě. Přesto by bylo možné vytvořit alternativní algoritmus, který dokáže vstup klíčovat v reálném čase. Zároveň by se dalo zpříjemnit používání aplikace grafickým uživatelským prostředím.

Celkový dojem výsledků je ale velice pozitivní. Metoda odhadu barvy pozadí předčila očekávání a je pro algoritmus velkým přínosem. Potlačení barevné korony i samotný klíčovací algoritmus využívají podobný způsob řešení, který ukázal, že přináší velmi dobré výsledky.

Příloha A

Literatura

- [1] BROWNRIGG, D. R. K. The weighted median filter. *Commun. ACM* 27, 8 (Aug. 1984), 807–818.
- [2] DEBEVEC, P., WENGER, A., TCHOU, C., GARDNER, A., WAESE, J., AND HAWKINS, T. A lighting reproduction approach to live-action compositing. *ACM Trans. Graph. (Proc. ACM SIGGRAPH 2002)* 3, 21 (2002), 547–556. Dostupné z: <http://gl.ict.usc.edu/Research/LS3/debevec-s2002-preprint-compressed.pdf>.
- [3] ŠOUREK, P. Kontrola barevné korony na zeleném pozadí. Bakalářská práce, ČVUT FEL, 2014.
- [4] SMITH, A., AND BLINN, J. Blue screen matting. *SIGGRAPH* (1996). Dostupné z: <http://alvyray.com/papers/cg/blusig96.pdf>.
- [5] VLAHOS, P. Electronic composite photography with color control, 02 1977. Dostupné z: <https://www.google.com/patents/US4007487>.
- [6] VLAHOS, P. Comprehensive electronic compositing system, 08 1982. Dostupné z: <https://www.google.com/patents/US4344085>.
- [7] YAGIZ, A., TUNÇ, A., POLLEYFEYS, M., AND SMOLIC, A. Interactive high-quality green-screen keying via color unmixing. *ACM TOG Disney* (2016). Dostupné z: http://zurich.disneyresearch.com/~taydin/pdf/yagiz_tog16.pdf.
- [8] YAMNITSKY, D. Real-time chroma keying on the gpu. *Boris FX* (2009). Dostupné z: <https://pdfs.semanticscholar.org/15fa/3d9759f0bb61850a506666765aaf0cd9432d.pdf>.

- [9] YASUDA, K., NAEMURA, T., AND HARASHIMA, H. Thermo-key: human region segmentation from video. *IEEE Computer Graphics and Applications* (2004). Dostupné z: <http://ieeexplore.ieee.org/document/1255805/?arnumber=1255805>.



Příloha B

Obsah přiloženého CD

1. **bin/** - Obsahuje spustitelné soubory přeložené pro platformu Windows x64.
2. **docs/html/** - Obsahuje dokumentaci doxygen v html.
3. **examples/run_examples.bat** - Batch soubor, který spustí přednastavený příklad a uloží výsledné obrázky do **examples/output/**
4. **images/** - Obsahuje použité obrázky.
5. **src/** - Obsahuje zdrojový kód aplikace a projektový soubor pro Visual Studio 2015.
6. **tex/** - Obsahuje zdrojový kód této bakalářské práce.
7. **readme.txt** - Uživatelský manuál aplikace (anglicky).
8. **alpha_matting.pdf** - Tato bakalářská práce v PDF.