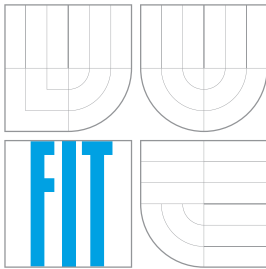


VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY



FAKULTA INFORMAČNÍCH TECHNOLOGIÍ
ÚSTAV POČÍTAČOVÉ GRAFIKY A MULTIMÉDIÍ

FACULTY OF INFORMATION TECHNOLOGY
DEPARTMENT OF COMPUTER GRAPHICS AND MULTIMEDIA

INTERAKTIVNÍ SEGMENTACE POPŘEDÍ/POZADÍ NA MOBILNÍM TELEFONU

INTERACTIVE FOREGROUND/BACKGROUND SEGMENTATION ON MOBILE PHONE

DIPLOMOVÁ PRÁCE

MASTER'S THESIS

AUTOR PRÁCE

AUTHOR

Bc. PETR STUDENÝ

VEDOUCÍ PRÁCE

SUPERVISOR

doc. Ing. ADAM HEROUT, Ph.D.

BRNO 2015

Abstrakt

Práce se zabývá technikami pro interaktivní segmentaci popředí a pozadí na mobilních zařízeních. Cílem je nalézt nebo vytvořit vhodné metody pro oddělení uživatelem označeného objektu od pozadí. V úvahu je brána zejména rychlost segmentace a kvalita výstupu. Jsou popsány principy již existujících metod pro interaktivní segmentaci, je diskutováno jejich využití na mobilních zařízeních. Je navržena a implementována vlastní metoda velmi rychlé segmentace použitá pro náhled z kamery mobilního zařízení v reálném čase.

Abstract

This thesis deals with the problem of foreground extraction on mobile devices. The main goal of this project is to find or design segmentation methods for separating a user-selected object from an image (or video). The main requirement of these methods is the image processing time and segmentation quality. Some existing solutions of this problem are mentioned and their usability on mobile devices is discussed. A mobile application is created within the project, demonstrating the implemented real time foreground extraction algorithm.

Klíčová slova

Interaktivní segmentace obrazu, mobilní zařízení, nalezení řezu grafem, získání vyhlazené alfa masky, editace obrazu, výřez objektu z obrazu, matematická morfologie.

Keywords

Interactive image segmentation, mobile devices, graph cut, alpha matting, image editing, foreground extraction, mathematical morphology.

Citace

Petr Studený: Interaktivní segmentace popředí/pozadí na mobilním telefonu, diplomová práce, Brno, FIT VUT v Brně, 2015

Interaktivní segmentace popředí/pozadí na mobilním telefonu

Prohlášení

Prohlašuji, že jsem tuto diplomovou práci vypracoval samostatně pod vedením pana doc. Adama Herouta. Uvedl jsem všechny literární prameny a publikace, ze kterých jsem čerpal.

.....

Petr Studený
26. května 2015

Poděkování

Děkuji panu docentovi Adamu Heroutovi za poskytnutí veškeré odborné pomoci, nápadů, inspirace a výborné vedení práce na diplomové práci. Děkuji Radku Vrzalovi za odborné rady a pomoc se zhotovením mobilní aplikace.

© Petr Studený, 2015.

Tato práce vznikla jako školní dílo na Vysokém učení technickém v Brně, Fakultě informačních technologií. Práce je chráněna autorským zákonem a její užití bez udělení oprávnění autorem je nezákonné, s výjimkou zákonem definovaných případů.

Obsah

1 Úvod	2
2 Segmentace popředí a pozadí	4
2.1 Metoda rozvodí	5
2.2 Metody založené na hledání řezu v grafu	6
2.2.1 Lazy snapping	7
2.2.2 GrabCut	8
2.3 Image matting	10
2.3.1 Trimapa	11
2.3.2 Rovnoměrné vyhlazení	12
2.3.3 Metoda globálního vzorkování	13
3 Operace používané pro segmentaci	14
3.1 Základní morfologické operace	14
3.2 Označování oblastí	15
3.3 Gaussův filtr	17
3.4 Mediánový filtr	18
4 Vytvořená rychlá segmentační metoda	19
4.1 Klasifikace	21
4.2 Morfologické úpravy	22
4.3 Učení modelu	22
4.3.1 Doučování	23
4.3.2 Přeučování	24
4.4 Vyhledávací okno	24
4.5 Provázání s jinými metodami	26
5 Aplikace	28
5.1 UI pro segmentaci v reálném čase	28
5.2 UI pro kvalitní segmentaci	30
5.3 Akcelerace výpočtů na mobilních zařízeních	30
6 Testování, vyhodnocení	32
6.1 Testování UI rychlé segmentace	32
6.2 Testování kvality výstupu rychlé segmentace	34
7 Závěr	38
A Obsah CD	43

Kapitola 1

Úvod

Výřez objektu z fotografie a jeho následné vložení do jiného snímku. Činnost, nazývaná fotomontáž, nachází uplatnění v oblasti tvorby reklamy (plakáty, letáky, propagace), ve fotoreportážích či jen ve tvorbě obrázků spojených s populární kulturou. Nástroje nabízející funkce pro realizaci této činnosti známe jak z klasických stolních počítačů, tak i z mobilních zařízení. Ovšem proč nevytvořit nástroj, který by nám umožňoval velice snadný způsob vložení objektu do fotografie přímo z náhledu kamery, kterou máme ve svém telefonu zabudovanou? Takový nástroj by dovoloval měnit pozici, velikost a natočení objektu pouhým pohybem telefonu, na kterém běží. Tento prostý, avšak účelný koncept by nabízel vytvoření jednoduché fotomontáže i uživatelům, kteří se kvůli nedostatku dovedností doposud tvorbě fotomontáží vyhýbali, protože jejich tvorba je často značně náročná.

Cílem diplomové je provést průzkum existujících technik, které se zabývají právě výřezem označeného objektu ze snímku, posoudit je z hlediska kvality a výpočetní náročnosti. Pro metody, které jsou účelné v řešení tohoto problému, navrhnout implementaci a integraci do mobilní aplikace, určené pro průměrně výkonný telefon současnosti i pro PC. Mobilní aplikace musí splňovat zejména požadavek jednoduchého a intuitivního ovládání.

Studiem vybraných metod se ukázalo, že je velice obtížné najít metodu, pomocí které by bylo možné vytvářet výřez objektu ze snímků videa v reálném čase na mobilních zařízeních středního výkonu. Byly proto sepsány požadavky na novou metodu, navrženou přímo na míru k řešení tohoto problému. V práci je detailně popsán návrh a způsob implementace nové metody, testování a vyhodnocení rychlosti a kvality výstupu segmentace. Metoda je původně určena k interaktivní segmentaci s minimalistickým uživatelským vstupem na mobilních telefonech, ovšem nachází uplatnění i na PC, například pro záměnu pozadí při videokonferencích.

V rámci semestrálního projektu k této práci bylo splněno studium některých segmentačních technik, zejména pak metod, které jsou založené na hledání řezu v grafu a metody rozvodí. Z těchto dvou druhů technik bylo plánem v semestrálním projektu vytvořit konečné řešení. Použití metody rozvodí pro rychlé rozčlenění obrazu do několika oblastí a vhodné vybrání některých z nich, které představují označený objekt, se zdálo být vhodnou volbou, která splňuje rychlostí požadavky na mobilní zařízení. Získaný výřez objektu, by mohl obsahovat drobné chyby segmentace, ovšem dokázal by zpracovávat v reálném čase náhled z kamery mobilního telefonu. Pro výřez objektu v maximální možné kvalitě měla být použita některá ze segmentačních metod založená na hledání řezu grafem. Taková segmentace by dosahovala vysoké kvality a nebyl by na ni již kladen požadavek běhu v reálném čase nad videem v mobilním zařízení.

Návrh řešení sestavený během semestrálního projektu byl nakonec zaměněn vlastní vytvořenou segmentační metodou a možností jejího provázání s jinými metodami, které se vyznačují vysokou kvalitou výstupu segmentace. Metoda je navržena pro získání masky k výřezu objektu ze snímku, který je oproti svému pozadí dostatečně kontrastní. Pro tyto scény je vytvořená metoda velice rychlá a robustní, přitom vyžaduje od uživatele minimální vstupy. V případě, že metoda dokáže spolehlivě segmentovat objekt ve snímku, pak dokáže i spolehlivě nachystat vstupní data jiným, kvalitnějším metodám bez nutnosti dodatečných uživatelských vstupů.

Text práce je pro přehlednost členěn do kapitol a to následovně. V kapitole 2 je přehled některých segmentačních metod a jejich rozdělení. Zaměřeno je na metody, které jsou podpořené uživatelským vstupem (interakcí), zejména pak takové, které jsou významné pro vytvoření konečného řešení problému nebo jejich studium významně přispělo k dokončení práce. Pro každou představenou segmentační metodu je uveden možný způsob použití v jednotlivých částech projektu. Dále jsou zde k nalezení techniky použitelné ke tvorbě vyhlazeného výřezu objektu z fotografie. V kapitole 3 jsou uvedeny algoritmy, které jsou v návrhu řešení použité jako pomocné pro segmentaci. Pomocné algoritmy se používají pro předzpracování snímku před segmentací, nebo pro zpracování výsledku segmentace. V kapitole 4 je detailní popis vlastní navržené a implementované metody, je provedena její dekompozice, popsány jednotlivé moduly a popsán způsob dosažení kladených požadavků. Kapitola 5 představuje návrh a realizaci mobilní aplikace, která demonstruje funkci vytvořeného algoritmu pro segmentaci náhledu z kamery mobilního zařízení v reálném čase. Zaměřeno je zejména na návrh uživatelského rozhraní, který obaluje vytvořenou metodu. V kapitole 6 je popsán způsob a výsledky testování. Testování je provedeno zejména ve dvou kategoriích. Získání uživatelské zpětné vazby k navrženému uživatelskému rozhraní pro ovládání vytvořené segmentační metody a testy vyhodnocující kvalitu výstupu vytvořené segmentační metody na nachystané datové sadě. Poslední kapitola 7 zhodnocuje dosažené výsledky a představuje možné pokračování v řešení této práce.

Kapitola 2

Segmentace popředí a pozadí

Segmentace obrazu v oboru počítačového vidění je proces rozdělení obrazu na množinu segmentů. Segment je množina pixelů, také nazývaná jako superpixel. Obecným cílem segmentace je změnit nebo zjednodušit reprezentaci obrazu tak, aby byla srozumitelná pro nadcházející analýzu [23]. Segmentačních metod existuje celá řada a můžeme je pro potřeby tohoto projektu rozdělit do dvou základních kategorií.

- Plně automatické segmentační metody
- Interaktivní segmentační metody

Za plně automatické metody považujeme ty, které nevyžadují žádný uživatelský vstup, vyjma zadání parametrů segmentace – uživatel neoznačuje v obraze oblast zájmu. Mezi automatické řadíme metody využívající prahování (náležitost každého pixelu do segmentu je rozhodnut na základě prahu), metody založené na hranách (například Cannyho hranový detektor [4] – nachází uplatnění v případech segmentace obrazu podle hran), regionální metody a některé speciální jako bezmarkerová varianta metody rozvodí či různé varianty metod založené na aktivních konturách (například z [11] – křivky, které se připínají k hranám v obraze. Při dokonalé segmentaci obecné fotografie, by každý segment měl představovat samostatný objekt reálného světa. Ovšem dokonalá automatická segmentace není na obecných fotografiích reálná a segmenty neoznačují samostatné objekty, ale pouze oblasti obrazu, které se vyznačují vysokým stupněm homogenity.

Plně automatické metody nemají díky svým nedokonalým výsledkům v případě této práce tak velké uplatnění, jako metody interaktivní. Ty jsou velice vhodné, protože i malý uživatelský vstup má výrazný vliv na výsledek segmentace obrazu. Cílem práce je studie zejména těch interaktivních metod, jejichž výstupem je obraz segmentovaný do dvou druhů segmentů. Jeden druh reprezentující tvar objektu reálného světa, který byl uživatelským vstupem označený pro oddělení od druhého druhu – jeho pozadí (všechna ostatní oblast, která nepředstavuje segmentovaný objekt). V této práci pro dosažení stanovených cílů se vycházelo zejména z následujících interaktivních metod.

- Markerová varianta metody rozvodí (*marker-controlled watershed transform*) vynikající ve vysoké rychlosti segmentace
- Metody založené na hledání řezu v grafu vynikající ve vysoké kvalitě segmentace

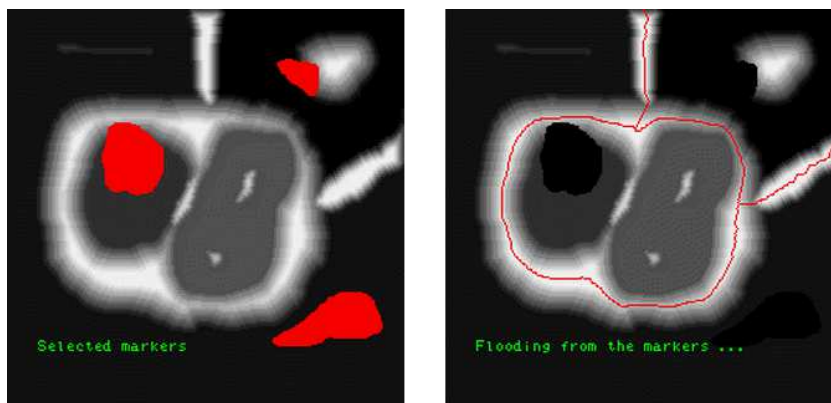
Segmentace metodou rozvodí byla shledána jako účelná, díky své rychlosti a kvalitě, na segmentaci náhledu z kamery mobilního zařízení v reálném čase. Myšlenka aplikace metody

na tento problém je popsána v kapitole 2.1. Pro získání výřezu objektu v plném rozlišení kamery mobilního zařízení při nejvyšší možné kvalitě výsledné segmentace je možné použít některou ze sofistikovanějších metod založených na hledání řezu v grafu. Dvě vybrané metody (*Lazy snapping*, *GrabCut*) jsou zmíněné v následujících podkapitolách. Vybrané jsou z důvodu některých zajímavých prvků, které přináší a mohou být užitečné v průběhu řešení této práce. Mezi tyto prvky patří například: urychlení algoritmu vhodným předzpracováním či iterativní přístup k výpočtu segmentace.

Aby bylo možné výsledek segmentační metody pro výřez objektu z jedné fotografie a jeho následné vložení do jiné je nutné vytvoření vyhlazených okrajů objektu. Důvodem potlačení takzvaných ostrých okrajů kolem vloženého objektu je dosažení maximální vizuální kvality a důvěryhodnosti vytvořené koláže. Způsoby přístupu řešení tohoto problému jsou zavedeny v poslední podkapitole této kapitoly.

2.1 Metoda rozvodí

Metoda rozvodí je postavena na chápání obrazu jako terénu (jas udává výšku terénu) a jeho zaplavování vodou. V nemodifikované metodě jsou výchozí body (nazýváme počátečním povodím), odkud je obraz zaplavován, vybrány jako lokální minima obrazu. Při modifikaci metody (abychom ji mohli považovat za interaktivní) necháme uživatele zadat tyto výchozí body ručně. Několik různých algoritmů pro provedení metody je představeno v [19]. Z počátečních povodí je tedy postupně obraz zaplavován vodou. V místech, kde by mělo dojít ke slítí dvou různých povodí vytvoříme hráz (chápeme jako terén o nekonečné výšce). Postupné zaplavování obrazu je zastaveno, jakmile úroveň vody dosáhne nejvyšší úrovně terénu (úplné maximum obrazu). Výsledný obraz je tedy rozdělen na takové množství regionů, kolik je zadaných souvislých výchozích povodí (obrázek 2.1). Jednotlivé regiony jsou odděleny hrázemi (čára v obraze o tloušťce 1).



Obrázek 2.1: Segmentace obrazu metodou rozvodí. Vstupní šedotónový obrázek s červeně vyznačenými iniciálními markery (vlevo), výstup segmentace – červené čáry značí hranice mezi segmenty (vpravo). Převzato z [1].

Tato metoda vyniká při dobrém zadání počátečních povodí kvalitními výsledky segmentace v dobrém čase – podle [7] snímek o velikosti 321×481 je segmentován na 1GHz ARM procesoru za 47.46 ± 4.99 milisekund, což by mohlo zajistit segmentaci náhledu z kamery mobilního zařízení v reálném čase. Je možné vytvořit iniciální značky pro tuto metodu

jako výstup jednoduché klasifikace pixelů na základě barvy. Obraz, pro který by byla takto vytvořena množina startovních značek, by byl metodou rozvodí značně přesegmentován. Ovšem pomocí metod matematické morfologie je možné odstranit přebytečné hranice mezi segmenty stejného typu (popředí a pozadí) a spojit je ve větší celky.

2.2 Metody založené na hledání řezu v grafu

Metody, které pracují s obrazem jako s grafem, jsou velice populárními pro hledání obrysů objektů v obraze. Podle [3] jsou metody vhodné pro rozdělení obrazu na dvě třídy. Tato vlastnost (často nazývaná omezením) je ideální pro řešení tohoto projektu. Dvě zmíněné třídy jsou v našem případě objekt, který v obraze hledáme, a pozadí, které chceme od objektu oddělit. Představíme dvě metody vycházející z hledání řezu grafem. Protože tyto segmentační metody jsou založené na teorii grafů, zavedeme definice z [6].

Síť (G, s, t, c) je orientovaný graf $G = (V, E)$, který má dva speciální vrcholy: *zdroj* s a *spotřebič* t a každá hrana e má *kapacitu* $c(e)$. Tok je funkce $c : E \rightarrow \mathbb{R}_+$. Spotřebič je někdy označován jako *stok*.

Tok je funkce $f : E \rightarrow \mathbb{R}_+$, která splňuje

$$\text{i) } 0 \leq f(e) \leq c(e) \text{ pro každou hranu } e \in E$$

$$\text{ii) } \sum_{xv \in E} f(xv) = \sum_{vx \in E} f(vx) \text{ pro každý vrchol } v \in V \setminus \{s, t\}$$

Číslu $f(e)$ říkáme tok po hraně e nebo také průtok hranou e . První podmínka říká, že průtok hranou je nezáporný a nemůže překročit kapacitu hrany. Druhá podmínka říká, že co do vrcholu přiteče, to z něj zase musí odtéci.

Velikost toku $|f|$ je množství toku, které protéká ze zdroje do spotřebiče. Protože se tok ve vrcholech ani neztrácí ani nepřibývá, tak ho můžeme spočítat jako $|f| = f(t)$. Tedy jako tok, který přitéká do spotřebiče. Stejně bychom mohli velikost toku měřit jako $-f(s)$, což je velikost toku, který vytéká ze zdroje.

Doplněk množiny $R \subset V$ značíme \bar{R} . Tedy $\bar{R} = V \setminus R$. Množinu orientovaných hran $\delta(R) = \{vw \in E \mid v \in R \& w \in \bar{R}\}$ nazveme řezem určeným množinou $R \subset V$. Řez je (s, t) -řez, pokud $s \in R$ a $t \notin R$. Tedy pokud řez odděluje zdroj od spotřebiče. Řez určený jediným vrcholem v budeme zjednodušeně značit $\delta(v)$ místo $\delta(\{v\})$. *Kapacita řezu* $c(\delta(R)) := \sum_{e \in \delta(R)} c(e)$. Platí věta, že v každém grafu existuje maximální tok.

Pro každý (s, t) -tok f a každý (s, t) -řez $\delta(R)$ platí

$$f(t) \leq c(\delta(R))$$

je důsledkem toho, že pro každý (s, t) -tok f a každý (s, t) -řez $\delta(R)$ platí

$$f(\delta(R)) - f(\delta(\bar{R})) = f(t)$$

Samotná *Fordova-Fulkersonova věta*, která uvádí do vztahu maximální tok a minimální řez v síti oddělující zdroj od stoku a vychází z ní i myšlenka základního způsobu hledání maximálního toku – Fordova-Fulkersonova algoritmu (řešení nalezení maximálního toku v síti) zní: Pokud existuje maximální (s, t) -tok, pak

$$\max\{|f| : f \text{ je } (s, t)\text{-tok}\} = \min\{c(\delta(R)) : \delta(R) \text{ je } (s, t)\text{-řez}\}$$

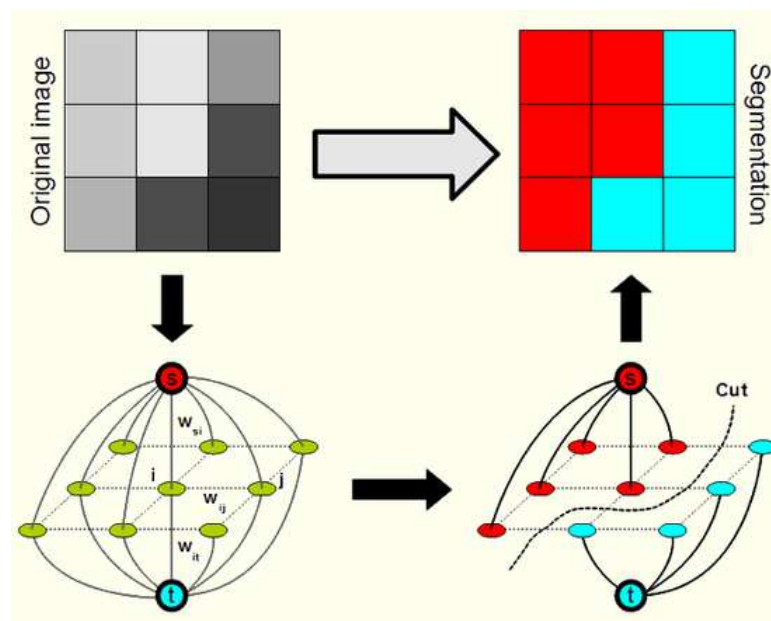
tedy pro každou síť se velikost maximálního toku rovná kapacitě minimálního řezu.

Základní znění *Fordova-Fulkersonova algoritmu* je: začneme s nulovým tokem a postupně budeme hledat vylepšující cesty, podél kterých zvětšíme aktuální tok. Když už nebude existovat vylepšující cesta, tak máme maximální tok. Součet toků f tedy můžeme prohlásit za maximální tok.

Nyní můžeme definovaný algoritmus aplikovat jako metodu zpracovávající obraz. Obrazek reprezentujeme grafem, kde

- každý pixel odpovídá jednomu uzlu grafu
- sousedící pixely (čtyř-okolí, osmi-okolí) propojíme hranou v grafu
- rozdíl intenzit sousedních pixelů považujeme za ohodnocení hran grafu
- pixely označené uživatelem jako popředí propojíme se speciálním uzlem s
- pixely označené uživatelem jako pozadí propojíme se speciálním uzlem t

Takto vytvořený graf na základě vstupního obrazu je znázorněn na obrázku 2.2 pro obraz o rozměrech 3×3 pixely. Cílem této segmentace je rozdělit všechny pixely obrazu do dvou skupin (popředí – hledaný objekt a pozadí). Náležitost pixelu do skupiny je určena propojením se speciálním uzlem po provedení řezu. Z různých modifikací a optimalizací vychází velké množství metod pro interaktivní segmentaci obrazu. Jako zástupce takových metod uvádíme metodu *Lazy snapping* (sekce 2.2.1), která dosahuje dobrých výsledků v průměrném čase, ovšem nese jisté omezení a problémy v některých obecných fotografiích a jako metodu, která dosahuje velice dobrých výsledků segmentace představíme metodu *GrabCut* (sekce 2.2.2).



Obrázek 2.2: Způsob reprezentace obrazu grafem. Vstupní obraz (vlevo nahoře), výstupní segmentovaný obraz (vpravo nahoře), vizualizace grafu reprezentující obraz a nalezení řezu (dole). Převzato z [10].

2.2.1 Lazy snapping

Cílem metody *Lazy snapping*, představené v [13] je minimalizovat množství potřebné uživatelské interakce. Uživatelský vstup sestává ze zadání dvou druhů čar, které definují umístění objektu v obraze a tu část obrazu, kde objekt není (třídy popředí a pozadí). Metoda byla navržena jako rychlý zástupce podobných metod a v původním návrhu je zmíněno, že provedení výřezu objektu může být prováděno po zadání každé čáry.

Vysoká rychlost metody je zajištěna iniciální segmentací, jejíž výstupem je přesegmentovaný obraz. Takto vytvořené malé segmenty (superpixely) jsou vstupem samotné metody nalezení řezu v grafu (popsané v 2.2). Jako způsob iniciální segmentace je tvůrci zvolena metoda rozvodí. Přestože vstupem algoritmu pro hledání řezu v grafu je graf s mnohonásobně menším množstvím uzlů (například dvacetkrát méně superpixelů než pixelů), není celková metoda urychlena stejným násobkem. Nastává zde problém určení hodnoty jednoho superpixelu.

Metoda s sebou nese několik omezení a problémů. Nedokáže segmentovat příliš malé oblasti (superpixely jsou často větší, než hledaný objekt), špatné provedení na obrázcích, kde je podobná kombinace barev v popředí a pozadí. Přestože je metoda uváděna jako velice rychlá, tak segmentace obrázku o velikosti 600×400 je v původní implementaci prováděna v čase 250 ms na stolním počítači. Není tedy vhodná pro segmentaci snímků z kamery mobilního zařízení v reálném čase, ovšem přináší optimalizační myšlenku ve formě superpixelizace.

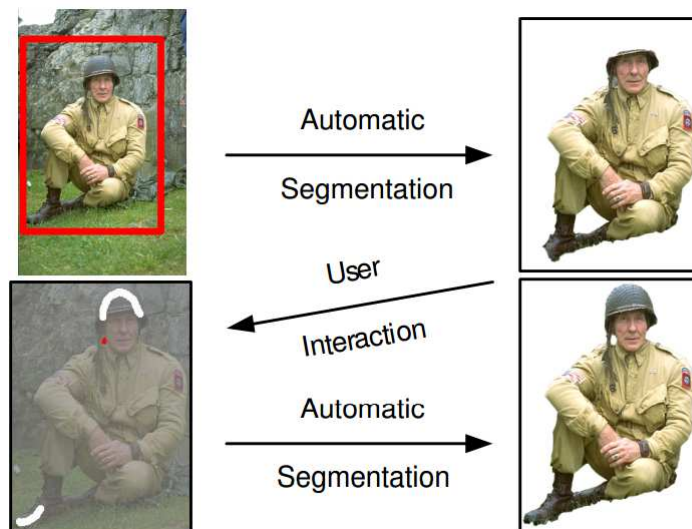


Obrázek 2.3: Ukázka použití metody *Lazy snapping*. Vstupní obrázek, kde žluté čáry jsou uživatelské vstupy značící popředí a modré čáry jsou vstupy značící pozadí (vlevo), výstupní obrázek po výřezu (vpravo). Převzato z [13].

2.2.2 GrabCut

Metoda GrabCut, představená v [20] řeší úlohu oddělení objektu od pozadí v barevném obraze. Uživatelská interakce je v této metodě rozdělena na dva fáze. První uživatelský vstup (povinný) je takový, že ve vstupním obraze uživatel označí ohraničující obdélník, který vymezuje polohu objektu v rámci obrazu. Vnější část obdélníku značí oblast, která je zcela jistě pozadím. Vnitřní část obdélníku značí oblast, kde dochází ke kombinaci popředí a pozadí. Po zadání tohoto jediného obdélníku je metoda spuštěna nad obrazem a pokud je

výsledek segmentace neuspokojivý, je možné zadat druhou (nepovinnou část) uživatelských vstupů. V druhé fázi uživatel značí pomocí dvou druhů čar (jedna pro popředí, druhá pro pozadí) segmentační chyby, vzniklé po první fázi. Princip uživatelských vstupů je znázorněn na obrázku 2.4.



Obrázek 2.4: Princip metody GrabCut. Vstupní obraz s vyznačeným ohraničujícím oknem (vlevo nahoře), výstup první fáze automatické segmentace pouze na základě ohraničujícího okna (vpravo nahoře), uživatelská interakce – bílé čáry doznačují popředí a červené pozadí (vlevo dole), výstupní obrázek z celého procesu segmentace. Převzato z [20].

Metoda GrabCut na rozdíl od ostatních metod, založených na hledání řezu v grafu, používá iterativní způsob hledání řezu, který má větší výpočetní nároky, ale přináší mnohem uspokojivější výsledky segmentace. Už v původním konceptu je navržena metoda pro práci s barevnými obrazy, některé starší metody jsou vytvořené pro práci s monochromatickým obrazem, což snižuje jejich možnosti využití. Před první iterací je sestaven graf reprezentující obraz a v něm jsou všechny pixely označené na základě uživatelského vstupu. Hranám v grafu (sousedící pixely) jsou přiřazeny váhy na základě barevných komponent, popisující třídy popředí a pozadí. Pomocí metody hledající minimální řez grafem (sekce 2.2) je určena segmentace obrazu a všem pixelům je přiřazeno nové označení, které je vstupem pro další iteraci. Tento postup je možné opakovat, protože použitý algoritmus konverguje. V poslední části metody dochází k vyhlazování okrajů nalezeného objektu. V každé iteraci algoritmu jsou provedeny následující tři kroky:

- Každý pixel vstupního obrazu je přiřazen jedné ze tříd popředí, nebo pozadí.
- Parametry barevných modelů pro třídy jsou přepočítány na základě příslušnosti pixelů. Tím jsou určeny váhy pro grafovou reprezentaci obrazu.
- Je použit algoritmus pro hledání řezu v grafu, jehož výsledkem je segmentace obrazu na dva druhy segmentů (binární maska pro výřez), která je vstupem pro další iteraci metody nebo je úplným výsledkem segmentace.

Podle [7] snímek o velikosti 321×481 je segmentován na 1GHz ARM procesoru za 7709 ± 1569 milisekund v množství iterací potřebných ke správnému výsledku segmentace,

proto není možné tuto metodu použít pro segmentaci náhledu z kamery mobilního zařízení v reálném čase, ovšem je možné ji použít pro vysoce kvalitní segmentaci konečného snímku v plném rozlišení z důvodů dobré kvality výsledné segmentace a z důvodu iterativního přístupu metody. Iterativní přístup, kde výstupem každé iterace je zároveň vstup pro následující iteraci, může být s výhodou použita pro průběžné zobrazování výsledků segmentace, nebo pozastavení běhu algoritmu a pozdější navázání na průběžný výsledek.

2.3 Image matting

Výraz *image matting* z důvodu těžkého nalezení českého ekvivalentu nebudeme v této práci překládat a použijeme jej v původním znění. Obecně *matting* (podle [24]) je technika přesného určení popředí v rámci obrazu nebo videa. Techniku interaktivní *image matting* zdefiniujeme z [12]. Jedná se o proces získání objektu v popředí z obrazu, založeném na omezeném množství uživatelské interakce. Je velice důležitým prvkem pro editaci obrázku (popřípadě videa). Z pohledu počítačového vidění se jedná o velice náročný úkol z důvodu své nejednoznačnosti – pro každý pixel vstupního obrazu musí být určena nejen náležitost pixelu do popředí či pozadí, ale zároveň neprůhlednost popředí pouze z jediné informace a to je barva daného pixelu. Proces vytváří takzvanou alfa masku (toto je vyobrazeno na obrázku 2.5), jedná se o obrázek ve stupních šedi, ve kterém bílý pixel značí zcela neprůhledné popředí, černý zcela průhledné a úrovně šedé barvy částečnou průhlednost. Podle hodnot alfa masky může být objekt popředí z jednoho snímku smíchán se snímkem jiným, čímž dojde k záměně pozadí extrahovaného objektu. Kvalita vytvořené alfa masky velice ovlivňuje vizuální kvalitu a důvěryhodnost takto vytvořené kompozice.

Metody zmíněné v předchozích sekcích vytvářely masku pro výřez, která sestávala pouze z dvou druhů pixelů bílé pro popředí, černé pro pozadí. Při použití takové masky pro výřez a vložení objektu popředí do jiného obrazu vznikne kompozice, ve které je znatelná ostrá, nedůvěryhodná hrana kolem vloženého objektu. Techniku získání takové masky budeme nazývat *cutting*.



Obrázek 2.5: Rozdíl mezi maskou pro výřez objektu (uprostřed) a alfa maskou pro míchání objektu s jiným pozadím (vpravo). Vstupní obrázek, na který jsou zobrazené masky aplikovatelné (vlevo)

Obecným přístupem tvorby alfa masky je následující postup:

- Použití některé segmentační techniky pro vygenerování takzvané *trimapy* (sekce 2.3.1). Ta je vstupním prvkem dalšího kroku postupu. Pro vygenerování může být použita libovolná metoda, která rozděluje vstupní obraz do na segmenty dvou druhů – popředí a pozadí.

- Použití nějakého algoritmu pro *image matting*, jehož vstupem je vstupní obraz a z něj vygenerovaná trimapa.
- V případě, že je zpracováváno video, je nutné aby byly potlačeny nekonzistentní prvky z temporálního pohledu.

Pro vygenerování alfa masky ze vstupního obrazu a vstupní trimapy existuje celá řada technik. Podle použitého přístupu se rozdělují na dvě kategorie: techniky založené na parametrickém vzorkování a techniky nezaložené na vzorkování.

Mezi metody založeném na vzorkování řadíme například metodu *Ruzan and Tomasi* [21], která je považována za jednu ze základních technik pro tvorbu alfa masky. Přístup předpokládá, že neznámý region je úzký pás mezi oblastmi popředí a pozadí a kostra neznámého regionu může být reprezentována jako řetězec pixelů. Řetězec je rozdělen na několik lokálních oken, ve kterých separovaně probíhá výpočet alfa hodnot na základě sestrojených barevných modelů (ve formě Gaussiánů popisujících barvy).

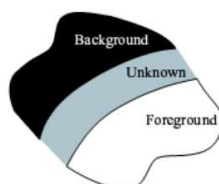
Druhou kategorií jsou metody, které nejsou založeny na vzorkování. Mezi takové metody řadíme například *Poisson Matting* [26], kde autoři využívají řešení Poissonovy rovnice k nalezení alfa hodnot v neznámé oblasti trimapy. Tím se snižuje chyba způsobená nesprávnou klasifikací barevných vzorků v komplexních scénách. Metoda má předpoklad, že změna intenzity mezi oblastí popředí a pozadí je vždy hladká.

V praktické části tohoto projektu je v kapitole o vlastní navržené metodě 4 ukázáno možné provázání s metodou pro *image matting*. V implementaci vlastního algoritmu je využita metoda *Global sampling method* [8] (popsaná v 2.3.3) jako metoda navazující pro získání alfa masky ve vysoké kvalitě. Vytvořená metoda pro rychlou (ovšem nižší kvality) segmentaci dokáže produkovat dostačující trimapu této metodě.

2.3.1 Trimapa

V předchozím sekci bylo zmíněna nutnost generování takzvaných *trimap*. Trimapa je obraz, velikostí odpovídající velikosti vstupního obrazu, sestávající ze třech druhů pixelů:

- pixely, označené za zcela jisté popředí
- pixely, označené za zcela jisté pozadí
- neznámé pixely

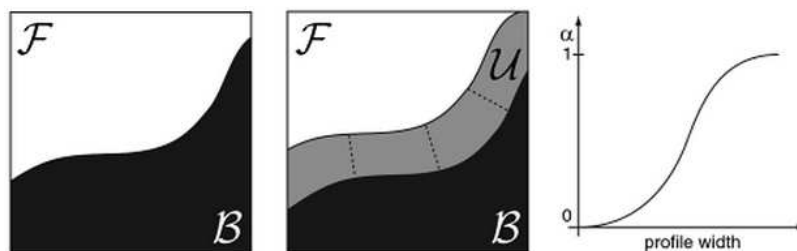


Obrázek 2.6: Trimapa sestává z pixelů tří druhů – jisté pozadí, jisté popředí a neznámé pixely. Převzato z [24].

Neznámé pixely jsou ty, které se nachází v oblasti, která odděluje regiony popředí od regionů pozadí. Trimapa může být vytvořena manuálně uživatelem, nebo automaticky vygenerována některou segmentační metodou. Manuální vytváření v této práci nebude bráno v úvahu z důvodů příliš velkého množství uživatelské interakce – tu potřebujeme maximálně minimalizovat. Metody pro vytvoření alfa masky nacházejí správnou hodnotu alfa, určující průhlednost vysegmentovaného objektu, na těch pozicích vstupního obrazu, které jsou trimapou označeny jako neznámé.

2.3.2 Rovnoměrné vyhlazení

V mnoha případech se využívá nejjednodušší možný způsob získání alfa masky pro smíšení segmentovaného objektu s jiným pozadím (takzvaný *matting*) z binární masky pro výřez objektu. Přístup je popsán v [18], způsob získání je ilustrován na obrázku 2.7. Spočívá v úpravě binární masky pomocí morfologických operací, která vede k vytvoření trimapy s rovnoměrně širokými neznámými regiony, které oddělují jisté popředí od jistého pozadí. Hodnoty alfa jsou poté určeny na nejkratších úsečkách v neznámém regionu, které spojují oblast popředí a pozadí. Na těchto přímkách je zvolena hodnota alfa pomocí funkce připomínající funkci sigmoida. Je nutné upozornit, že tento přístup, vytváří rovnoměrně široký neznámý region a tedy i rovnoměrně široké vyhlazení okrajů masky. Není vhodný v situacích, kde je region vyžadující vyhlazení různě široký (příkladem obrázek 2.8).



Obrázek 2.7: Nejjednodušší způsob vytvoření alfa masky z nevyhlazeného výstupu kterékoli segmentační metody. Výstup segmentační metody (vlevo), morfologickými operacemi vytvořená trimapa (uprostřed), funkce pro vyhlazení profilu neznámého regionu v trimapě (vpravo). Hodnota alfa je určena funkcí z intervalu $[0, 1]$, kde 0 znamená neprůhledné popředí a 1 znamená zcela průhledné popředí. Obrázek převzat z [18]

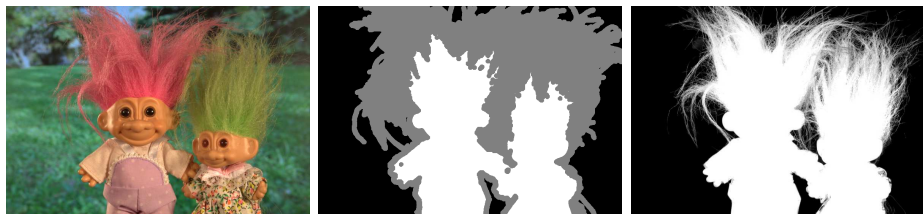
Protože je tento přístup nesrovnatelně méně náročný než zmíněné algoritmy pro získání kvalitních alfa masek, je využit v podobné verzi pro vyhlazování masky, která je výstupem vytvořené rychlé segmentační metody.

2.3.3 Metoda globálního vzorkování

Metoda globálního vzorkování (*Global sampling matting method*) [8] je metoda pro vytvoření kvalitní alfa masky z vstupního snímku a trimapy, patří do kategorie metod založených na vzorkování. Zatímco ostatní metody využívají vzorky z malých lokálních oblastí a získávají hodnotu alfa separovaně v daných lokálních oblastech, tato metoda využívá vzorky získané z celého vstupního obrazu. Mohou nastat případy, kdy pro výpočet alfa hodnoty na pozici neznámého pixelu v trimapě nemusí být nalezen vhodný vzorek v žádné ze známých oblastí při použití okna. Autoři využívají jednoduché a efektivní cenové funkce

k řešení nejednoznačnosti během procesu výběru vzorku. Funkce uvažuje nejen barvu pixelu, ale také jeho pozici. Pro nalezení korespondence mezi neznámými pixely a pixely ze známých oblastí popředí a pozadí je využita zobecněná funkce pro prohledávání velkých stavových prostorů.

Pokud jsou pro neznámý pixel určeny vzorky popředí a pozadí je možné provést výpočet alfa hodnoty. Výsledná hodnota alfa je získána stejným způsobem jako u ostatních metod založených na vzorkování.



Obrázek 2.8: Metoda Global sampling matting. Vstupní obraz metody (vlevo), vstupní trimapa (uprostřed), výstupí alfa maska(vpravo).

Pro tuto práci je metoda globálního vzorkování zmíněna zejména proto, že v metoda navržená pro segmentaci náhledu z kamery (kapitola 4) je určena k propojení s některou metodou, která produkuje vysoce kvalitní alfa masku. Jelikož navržená metoda dokáže produkovat dostatečně dobré trimapy pro tuto metodu, je využita jako metoda navazující.

Kapitola 3

Operace používané pro segmentaci

Aplikace segmentačních metod v praxi vyžaduje řadu podpůrných operací a algoritmů. Protože vstupní obraz může obsahovat různé druhy šumů, či jiné prvky, které mohou mít negativní vliv na výsledek použité segmentační metody, využívají se operace pro předzpracování obrazu ještě před samotnou segmentací. Mezi takové operace řadíme aplikace filtrů pro odstranění šumu. Další nutné operace jsou pro práci s výstupem segmentačních metod. Mezi důvody použití operací nad výstupem se řadí:

- drobná úprava tvaru výstupních segmentů
- odstranění zanedbatelně malých segmentů
- výběr jednoho konkrétního souvislého segmentu

V této kapitole budou zavedeny ty pomocné operace, které jsou využívány ve vlastní navržené segmentační metodě (kapitola 4). Mezi ně patří základní morfologické operace pro práci s tvarem objektu, algoritmy pro značení souvislých oblastí, Gaussův filtr a mediánový filtr.

3.1 Základní morfologické operace

Matematická morfologie vychází z vlastností bodových množin. V počítačovém vidění jejich nejčastější použití spočívá v aplikaci na binární obraz pro odstranění šumu, zjednodušení tvaru objektu, zdůraznění struktury objektu nebo získání obrysu objektu. V tomto projektu, kde výstupem segmentace je právě binární obraz (1 – značí popředí, 0 – značí pozadí) je možné některé základní morfologické operace s výhodou použít na práci s tvarem masky, získané ze segmentace. Mezi základní čtyři operace matematické morfologie podle [22] řadíme: erozi, dilataci, morfologické otevření a morfologické uzavření.

První základní operací je dilatace, která skládá body dvou množin pomocí vektorového součinu. Objekty v binárním obraze jsou tedy po aplikaci dilatace změněny právě o velikost strukturního elementu. Používá k zaplnění děr v obraze. Dilatace obrazu X strukturním elementem B je definována jako

$$X \oplus B = \bigcup_{b \in B} X_b \quad X \oplus B = \{d \in E^2 : d = x + b, x \in X, b \in B\}$$

Druhá základní morfologická operace je eroze, který skládá body dvou množin pomocí vektorového rozdílů. Objekty v binárním obraze jsou po aplikaci eroze menší o velikost

strukturního elementu. Může se tedy stát, že objekty spojené tenkou čarou jsou rozděleny na více objektů, nebo příliš tenké objekty se ztratí úplně. Eroze je duální operace k dilataci a je dána následujícím vztahem.

$$X \ominus B = \bigcap_{b \in B} X_b \quad X \ominus B = \{d \in E^2 : d + b \in X, \text{ pro } \forall b \in B\}$$

Další dvě základní morfologické operace jsou vytvořené složením předchozích dvou. Morfologické uzavření se aplikuje na obraz jako dilatace po erozi. Uzavření způsobí spojení dvou objektů, které jsou oddělené dírou menší než strukturní element a vyhladí obrys objektu. Uzavření obrazu X strukturním elementem B je definováno jako

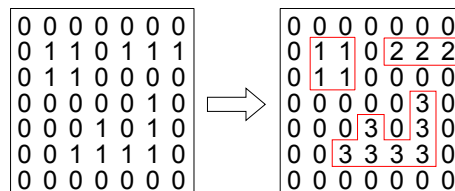
$$X \bullet B = (X \oplus B) \ominus B$$

Morfologické otevření obrazu je vytvořené jako eroze následovaná dilatací. Používá se k zjednodušení struktury objektu tak, že oddělí objekty spojené šijí tenčí, než strukturní element. V segmentaci může mít praktické využití například k odstranění hranic mezi oblastmi značící pozadí po segmentaci metodou rozvodí (spojí sousední oblasti pozadí do jediného). Otevření obrazu X strukturním elementem B je definováno jako

$$X \circ B = (X \oplus B) \ominus B$$

3.2 Označování oblastí

Označování spojených oblastí (*Connected-component labeling*) je algoritmus aplikovaný z teorie grafů, kde je množina spojených oblastí unikátně označena na základě dané heuristiky. Nejedná se o segmentační metodu, pouze při segmentaci nachází dobré uplatnění. V počítačovém vidění se zpravidla používá pro unikátní označení jednotlivých segmentů. Vycházíme z toho, že segment je tvořen souvislou oblastí o stejné hodnotě, ale ta ovšem v rámci celého obrazu není jedinečná (obrázek 3.1). Pro přiřazení jedinečného identifikátoru každé oblasti je možné použít dvou-průchodový algoritmus z [25] nebo jedno-průchodový algoritmus z [5]. Označení oblastí v dvou průchodech bývá v mnoha případech efektivnější, protože nevyžaduje žádné speciální datové struktury, ovšem v některých ojedinělých případech může být označení jedním průchodem výhodnější. Vhodná volba algoritmu bude diskutována v kapitole 4.

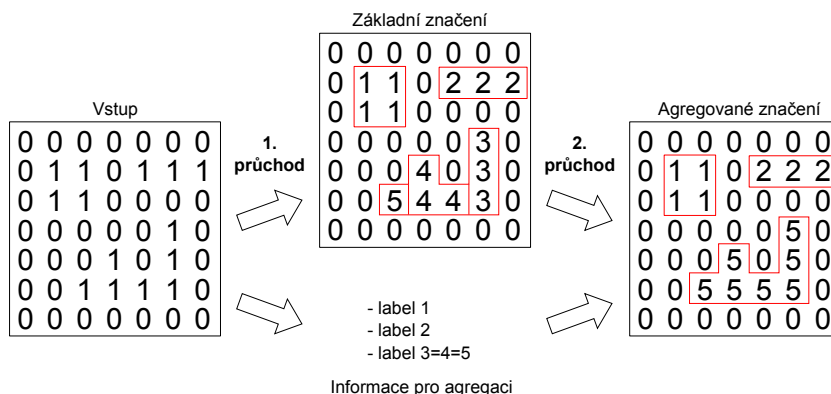


Obrázek 3.1: Označování souvislých oblastí. Vstupní binární obraz – 1 pro popředí, 0 - pro pozadí (vlevo), výstupní obraz s unikátně označenými spojitými oblastmi (vpravo)

Jako první bude představen princip dvou-průchodového algoritmu. V prvním průchodu jsou všechny pixely, nepatřící do pozadí označeny na základě sousedních pixelů (používá se čtyř-okolí, nebo osmi-okolí). Pokud žádný ze sousedních pixelů není označený, dostává

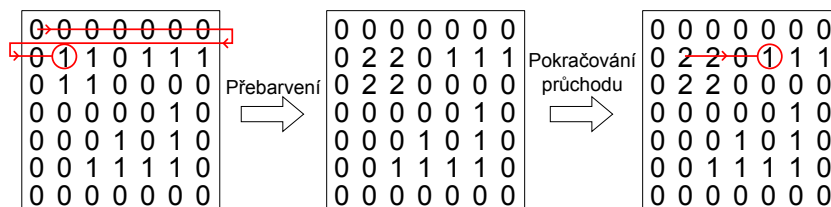
hodnotu zcela novou. Pokud jsou sousední pixely označené různými hodnotami, je pixel označen nejnižší hodnotou a pro vyšší hodnoty je nejnižší zaznamenána jako rodičovská.

V druhém průchodu dochází k agregaci. Tedy všechna označení, kterým bylo přiřazeno jiné (nižší) rodičovské označení jsou shlukována právě do rodičovského. V druhém průchodu je možné zjistit přesný počet pixelů pro každou souvislou oblast. Na základě toho mohou být příliš malé oblasti (způsobené přesegmentováním) odstraněny, nebo může dojít k výběru největší spojitě oblasti pro případ, že výsledkem segmentace je nalezení jediného objektu. Funkce dvou-průchodového algoritmu je znázorněna na obrázku 3.2.



Obrázek 3.2: Algoritmus pro unikátní označení oblastí ve dvou průchodech. První průchod vytváří základní značení a informace pro agregaci. Druhý průchod přeznačuje oblasti podle informací pro agregaci.

Jedno-průchodový algoritmus vychází z použití semínkového vyplňování. To vyžaduje pro implementaci použití zásobníku, nebo rekurzivního přístupu. Algoritmus při značení oblastí prochází postupně všechny obrazové pixely a pokud narazí na pixel splňující kritéria pro značení (například *hodnota* = 255 pro přiřazení unikátních identifikátorů bílým oblastem v osmi bitovém obraze) je pixel vložen na zásobník a spuštěno takzvané přebarvování. To představuje vytažení pixelu z vrcholu zásobníku, změnu jeho hodnoty na unikátní a vložení jeho čtyř-okolí (případně osmi-okolí) na zásobník, pokud opět splňují kritéria pro značení. Přebarvování je prováděno, dokud zásobník není prázdný. Po přebarvení následuje pokračování průchodu obrazem od poslední pozice. Problémem při jednom průchodu je nemožnost odstranění příliš malých oblastí, ovšem velikosti oblastí je možné počítat v průběhu přebarvování. Funkce jedno-průchodového algoritmu je znázorněna na obrázku 3.3.



Obrázek 3.3: Algoritmus pro unikátní označení oblastí v jediném průchodu. Pro přebarvování využívá techniku známou jako semínkové vyplňování.

Algoritmus je možné velice snadno rozšířit o test přítomnosti hledané oblasti, který předchází průchodu přes všechny pixely. Test spočívá ve vložení jednoho konkrétního pixelu na zásobník, pokud splňuje kritéria přebarvení. Tímto testem můžeme zjistit údaje o oblasti, jejíž součástí je námi vybraný pixel, před samotným průchodem obrazu. Mezi údaje, které lze pomocí testu prvním přebarvením zjistit patří:

- Velikost oblasti jako počet pixelů, ze kterých sestává.
- Pozice a velikost ohraničujícího obdélníku.

Na základě těchto údajů je možné rozhodnout, zda oblast je námi hledaná oblast. Pokud ne, je algoritmus spuštěn standardním postupem a oblast obarvená testem již není znovu přebarvována. Přítomnost testu tedy nezvyšuje časovou složitost algoritmu.

3.3 Gaussův filtr

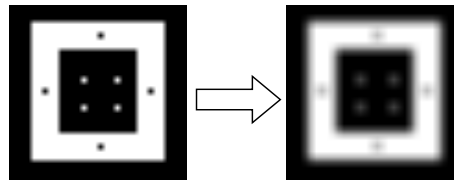
Gaussův filtr (Gaussián) je lineární filtr, který modifikuje vstupní signál (například obraz) konvolucí s jádrem vytvořeným Gaussovou funkcí. Podle definice z [23] je konvoluční jádro pro signál o dvou dimenzích definováno vztahem

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}.$$

Frekvenční analýzou je možné prokázat, že funkcí filtru je útlum vyšších frekvencí vstupního signálu. Díky tomuto efektu je filtr vhodný zejména pro potlačení šumu, což ovšem způsobí vyhlazení vstupní informace – říkáme, že je obraz po konvoluci rozmazaný. Nejčastější použití v oboru zpracování obrazu nachází filtr s velikostí konvolučního jádra 3×3 , 5×5 a 7×7 . Podle zvolené velikosti konvolučního jádra je určen parametr σ pro výpočet jednotlivých hodnot jádra. Výpočet σ vychází z empirického *pravidla tří sigma* [16], příkladem vztah použitý pro výpočet parametru v knihovně OpenCV s velikostí jádra *ksize*, který je aproximací pravidla tří sigma.

$$\sigma = 0,3(0,5(ksize - 1) - 1) + 0,8$$

V rámci této práce není použitý pro odstranění šumu ze vstupního obrazu, ale je využita jeho vlastnost vyhlazení. Při použití filtru na osmibitový obraz, ve kterém se nachází dva druhy pixelů (bílé a černé), vytvoří rovnoměrně vyhlazené přechody mezi dvěma zmíněnými barvami (ilustruje obrázek 3.4). Velikost přechodů je ovlivněna rozměry konvolučního jádra. Může být vhodný pro jednoduché získání alfa masky z binární masky pro výřez. Tímto lze provádět jednoduše jistou formu naivní verze techniky *image matting*.



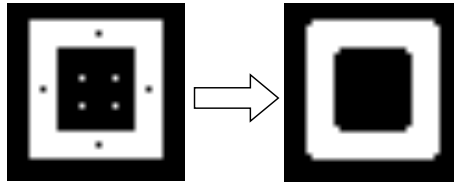
Obrázek 3.4: Použití filtrace Gaussiánem (konvoluční jádro 3×3) na černobílý obraz o rozměrech 32×32 . Výstupní šedotónový obraz má vyhlazené přechody mezi oblastmi.

3.4 Mediánový filtr

Mediánový filtr je technika nelineární filtrace obrazu, často používaná pro odstranění šumu. Bude zaveden popis filtru uvedený v [9]. Hodnota pixelu výstupního obrazu je určena jako medián pixelu vstupního obrazu a jeho určitého okolí. Mediánem rozumíme takovou hodnotu, pro kterou je stejný počet větších a stejný počet menších hodnot v dané množině. Mezi významné vlastnosti mediánového filtru patří:

- odstranění výstřelového šumu ze vstupního obrazu
- zachování hran ve filtrovaném obraze

Díky těmto vlastnostem, může být mediánový filtr použit pro úpravu masky pro výřez objektu popředí (zmíněné v předchozí kapitole) a sloužit tak jako částečná náhrada morfologického otevření a odstranění zanedbatelně malých oblastí, při kterém dochází k zaoblení ostrých rohů oblastí.



Obrázek 3.5: Použití mediánové filtrace (jádro 3×3) na černobílý obraz o rozměrech 32×32 . Výstupní šedotónový obraz je zbaven zanedbatelně malých oblastí a jsou zachovány jeho hrany.

Kapitola 4

Vytvořená rychlá segmentační metoda

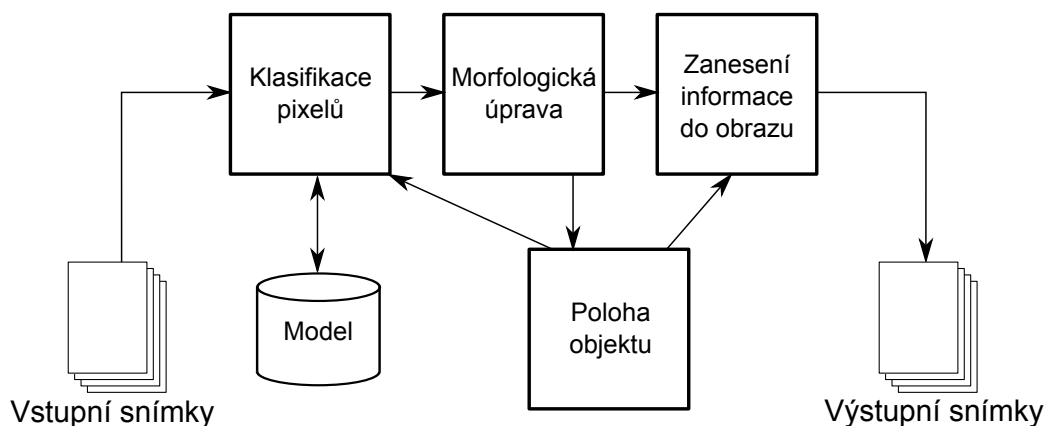
V předchozí kapitole byly představeny vybrané metody pro interaktivní segmentaci obrazu na dva druhy segmentů (popředí a pozadí). Ovšem žádná ze zmíněných metod zcela nesplňuje požadavky na segmentaci náhledu z kamery mobilního zařízení v reálném čase, jako jsou: minimální uživatelský vstup, vysoké rozlišení, vysoká rychlost, dobrá kvalita výsledné segmentace. Proto byla průběhu studia segmentačních metod vytvořena metoda nová pro jedno konkrétní využití. Návrh koncipován tak, aby metoda mohla předcházet některé ze sofistikovanějších metod, zmíněných v kapitole 2. Požadavky, které metoda musí nutně splňovat jsou shrnuty v následujících bodech

- Minimalizuje uživatelský vstup – ideálním případem je hrubé označení objektu, který chceme oddělit od pozadí, bez jakéhokoli jiného uživatelského vstupu. Tento požadavek dovoluje v implementované aplikaci minimalizovat prvky uživatelského rozhraní a zjednodušuje ovládání.
- Je výpočetně nenáročná. Účelem metody je segmentovat v reálném čase náhled z kamery současných mobilních zařízení.
- Spolehlivě segmentuje objekt, který je barevně kontrastní oproti pozadí.

Oproti tomu jsou požadavky, které metoda zcela nutně splňovat nemusí:

- Nepožadujeme zcela dokonalou segmentaci označeného objektu, nalezený segment v obraze nemusí být zcela shodný s reálným objektem na snímku. Spokojíme se s přiměřenou kvalitou, kdy segment bude plochou odpovídat minimálně z osmdesáti procent segmentovanému objektu.
- Nepožadujeme robustní metodu segmentace snímků, kde objekt je oproti pozadí v nízkém kontrastu.

V obrázku 4.1 je uvedeno blokové schéma navržené a vytvořené metody. Princip spočívá v klasifikaci jednotlivých pixelů pouze na základě barvy, žádné hrany a textury nejsou uvažovány. Náležitost pixelu do třídy *popředí* nebo *pozadí* je určena podle hodnot ve vytvořeném barevném modelu. Protože v každém dalším snímku se mohou velice rychle měnit světelné podmínky v obraze je model klasifikátoru průběžně upravován. Výstupem bloku pro klasifikaci je maska, která může obsahovat řadu chyb a nedokonalostí, které jsou v bloku pro



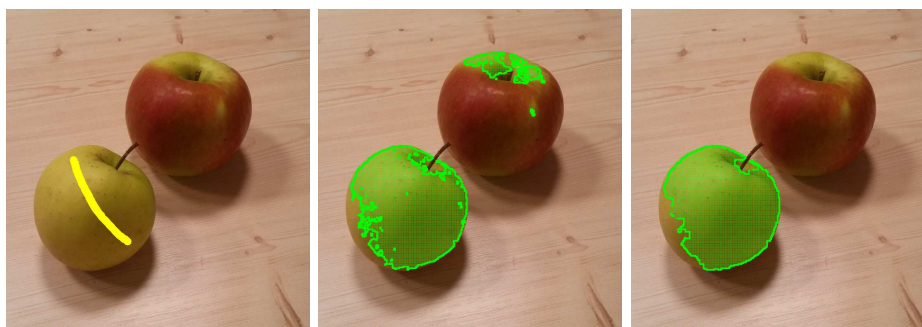
Obrázek 4.1: Schéma vytvořené segmentační metody, založené na klasifikaci pixelů podle barvy

morfológickou úpravu odstraněny. Výstupní masky jsou promítnuty do vstupního snímku na obrázku 4.2.

Barevný model uchovává informaci potřebnou ke klasifikaci ve formě čtyř histogramů.

- Histogram popředí
- Histogram pozadí
- Normalizovaný histogram popředí
- Normalizovaný histogram pozadí

Použité histogramy jsou třírozměrné, jedna dimenze pro každý barevný kanál. Aby nebyly vytvořené histogramy příliš řídké, jsou ve všech dimenzích histogramy kvantizovány po dvou i více hodnotách.



Obrázek 4.2: Výstupy modulů segmentační metody. Vstupní obrázek s žlutým označením objektu (vlevo), výstup modulu pro klasifikaci pixelů na základě barvy (uprostřed), výstup modulu pro morfológickou úpravu (vpravo)

4.1 Klasifikace

Klíčovou částí celé vytvořené metody pro rychlou segmentaci obrazu na mobilním zařízení je jednoduchý a velice rychlý způsob klasifikace – rozřazení pixelů do dvou kategorií, jedna kategorie pro pixely náležící do popředí (segmentovaný objekt), druhá pro pixely náležící do pozadí (vše ostatní). Jedním z požadavků použité klasifikace je nutnost, aby učení modelu (například za účelem zpřesnění) v průběhu segmentace nemělo významný vliv na rychlost běhu celé metody. Snažíme se dodržet tento požadavek, tedy metoda musí fungovat na mobilním zařízení v reálném čase i tehdy, když je průběžně upravován model.

Z hlediska požadavků byl pro klasifikaci zvolen způsob vycházející z Bayesova teorému. Zavedeme definici Bayesova teorému a z něj vycházející klasifikátor podle [28]. Klasifikátor je funkce, která přiřazuje třídu pro vzorek. Z pravděpodobnostního pohledu, podle Bayesova teorému, je pravděpodobnost vzorku $E = (x_1, x_2, \dots, x_n)$, že náleží do třídy c definována jako

$$p(c|E) = \frac{p(E|c)p(c)}{p(E)}.$$

Vzorkem v případě této metody je jeden obrazový pixel vstupního snímku a jeho parametry jsou barevné složky, ze kterých se skládá. Klasifikátor je tedy funkcí, jejímž parametrem je jeden obrazový pixel. Výstupem funkce je třída, do které pixel náleží. Třída $c \in C$, kde $C = \{foreground, background\}$ je množina tříd klasifikátoru. Pro zjednodušení následujícího popisu vytvořeného klasifikátoru provedeme přepis výrazu do prostého textu, tak dostáváme

$$posterior_c = \frac{prior_c \times likelihood_c}{evidence},$$

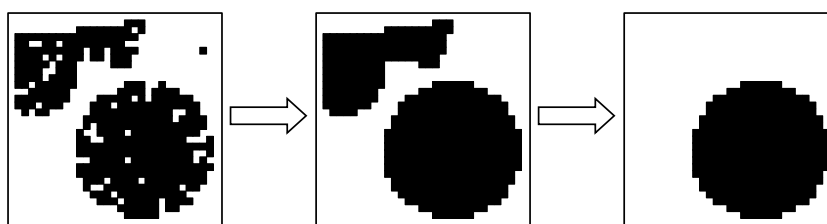
kde *posterior* značí aposteriorní pravděpodobnost náležitosti pixelu do třídy d , *prior* značí apriorní pravděpodobnost, tedy předpokládané množství pixelů náležící do třídy c zapsané jako poměr. Parametr *prior* je volen jako poměr množství pixelů klasifikovaných v minulém snímku jako třídy c proti celkovému množství pixelů vstupního snímku. Předpokládáme, že segmentovaný objekt bude přibližně stejně velký, jako byl v minulém snímku. Pro první snímek je *prior* volen jako konstanta. Ze sestrojených normalizovaných třírozměrných histogramů získáme hodnotu *likelihood*, jako hodnotu konkrétního zásobníku histogramu na souřadnicích udaných hodnotami jednotlivých barevných kanálů pixelu. Obdobným způsobem může být získán *evidence* z normalizovaného histogramu celého obrazu. Ovšem protože se obraz klasifikuje pouze do dvou tříd, vyčíslíme *posterior_{foreground}* i *posterior_{background}* pro každý pixel s hodnotou *evidence* = 1, čímž se vyhneme časově náročnému dělení a výslednou třídu volíme na základě porovnání dvou spočítaných aposteriorních pravděpodobností.

Podle výsledků klasifikace je vytvářena obrazová matice o poloviční velikosti jako vstupní snímek, do které jsou průběžně zapisovány černé (značící pozadí) a bílé (značící popředí) pixely, dále bude tento obraz nazývat jako *maska*. Poloviční velikost masky oproti vstupnímu obrazu je volena z důvodů dalšího urychlení celé metody. Podvzorkování vstupního obrazu (neklasifikujeme každý pixel snímku, ale jeden ze čtyř) výrazně urychlí nejen blok klasifikace, ale i blok morfologických úprav. Při zpětném zapisování výsledku metody do snímku je ovšem nutný průchod přes všechny obrazové pixely. Vytvořená maska je vhodná pro následující zpracování, kde dochází k nalézání a úpravám regionů, které v průběhu klasifikace vznikly.

4.2 Morfologické úpravy

Klasifikované pixely jsou pomocí nelineárních filtrů a morfologických operací upraveny, vyhlazeny a jsou vyplněné drobné díry v objektech. Protože metoda nalezne i všechny ostatní objekty ve scéně, které sestávají ze stejné nebo podobné kombinace barev, je z nalezených objektů vybrán ten, který velikostí a polohou přibližně odpovídá objektu z minulého snímku. V případě úplné ztráty objektu se metoda zaměří na největší objekt ve scéně.

Vstupem bloku je černobílá maska, ve které bílé pixely značí místa, která byla ze vstupního obrazu klasifikována jako popředí a černé pixely místa, která byla klasifikována jako pozadí. Na masku jsou aplikovány morfologické operace pro uzavření děr jak v regionech popředí, tak v regionech pozadí. Před samotným aplikováním je vstupní maska vyhlazena mediánovým filtrem pro úplné odstranění osamocených pixelů. Použití mediánového filtru nahrazuje částečně nutnost použití morfologického otevření, odstraní značné množství zanedbatelně malých oblastí a vyhladí ostré rohy v masce, která je výsledkem klasifikace.



Obrázek 4.3: Ilustrace funkce bloku pro morfologickou úpravu. Vstupní obraz je nejprve vyhlazen, jsou uzavřeny díry a následovně je ze vzniklých regionů vybrán jediný. Z důvodů přehlednosti ilustrace jsou zde barvy pixelů značící popředí a pozadí uvedeny inverze, než je uvedeno v textu.

Z předchozích operací vzniknou v masce spojené regiony o stejné barvě (takzvané *connected components*), ze kterých je pomocí algoritmu pro značení oblastí (sekce 3.2) vybraná jediná oblast. V metodě je využita jednodprůchodová varianta algoritmu. Volba jednodprůchodové varianty urychlí nalezení objektu, pokud kamera výrazně nemění svoji pozici oproti segmentovanému objektu. Metoda přenáší některé informace o poloze a velikosti segmentovaného objektu a proto volí počáteční pixel masky, který je vložený na zásobník pro přebarvení, na pozici objektu v minulém snímku. Průchod přes masku je proveden až tehdy, pokud se objekt na pozici již nenachází.

4.3 Učení modelu

Učení modelu je sestavení histogramů popředí a pozadí a výpočet jejich normalizovaných ekvivalentů. Podle hodnot v normalizovaných histogramech probíhá klasifikace pixelů v prvním bloku metody. Učení modelu v metodě rozdělíme do třech kategorií:

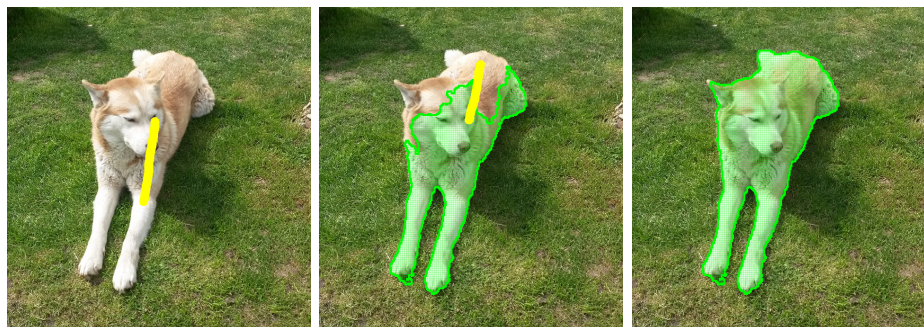
- Vytvoření modelu z prvního uživatelského vstupu
- Doplnění modelu popředí z dalších uživatelských vstupů – *doučování*
- Průběžná úprava modelů popředí i pozadí z důvodů zpřesnění klasifikace – *přeučování*

Pro první učení jsou do histogramu popředí zaneseny všechny obrazové pixely vstupního snímku, které mají průnik s uživatelským vstupem (uživatelé vytvořená trajektorie značící reálný objekt v obraze). Ovšem nemůžeme jednoznačně říct, které pixely snímku náležejí do třídy pozadí. Histogram pozadí je tak rovnoměrně naplněn všemi barvami z celého prostoru barev. Toto je nejkritičtější část metody, protože reálné pozadí může být často klasifikováno jako popředí z důvodu výskytu několika pixelů určité barvy na reálném popředí, ze kterých reálné pozadí z větší části sestává. Robustnější funkcionalitu metoda vykazuje v případě, kdy je objekt označen větším množstvím kratších čar. Všechny uživatelské vstupy kromě prvního jsou pak zpracovávány jako *doučování*, které může tyto případy lépe ošetřovat, protože již existuje sestavený histogram z pixelů pozadí. Po fázi prvního naučení je obraz klasifikován a na základě výsledků klasifikace je histogram pozadí předělán do správných hodnot – při prvním učení je tedy histogram sestavován nadvakrát.

4.3.1 Doučování

Doučování modelu je úprava histogramů popředí i pozadí taková, že do histogramů popředí je informace přidána a z histogramu pozadí je informace odebrána. Doučování je účelné, pokud objekt ve vstupním snímku není segmentován zcela, ale pouze částečně. Situaci ilustruje obrázek 4.4, kde pro správnou segmentaci jsou provedeny dva uživatelské vstupy (druhý vstup je *doučovací*).

Ve fázi doučování jsou již oba histogramy správně sestaveny. Pixely, které jsou v průniku s doučovacím uživatelským vstupem jsou přičítány do histogramu popředí a jsou odečítány z histogramu pozadí. Kdyby záznam o pixelech určité barvy byl z histogramu pozadí úplně odstraněn, docházelo by ke stejnému efektu, který byl zmíněn u učení na prvním uživatelském vstupu. Takto je model přesvědčen, že určitá barva patří do popředí až tehdy, pokud je doučením označeno dostatečné množství pixelů dané barvy. Může se ovšem v této fázi stát, že některou část objektu nebude možné vůbec klasifikovat jako popředí, pokud se skládá z barev, které jsou v pozadí v významném množství.

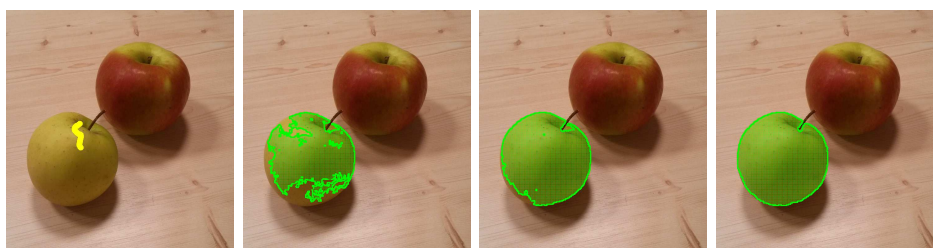


Obrázek 4.4: Doučování barevného modelu. Vstupní obrázek s výchozím uživatelským vstupem (vlevo), výstup metody – je viditelná chybějící část označeného objekt, která je uživatelem doznačena (uprostřed), výstup metody po doučení (vpravo)

4.3.2 Přeučování

V průběhu segmentování jednotlivých snímků náhledu z kamery se mění světelné podmínky nebo jas je upravován vyvážením bílé a to pouhou drobnou změnou polohy kamery. Tyto

změny musí být v metodě maximálně ošetřeny, aby nebylo nutné příliš časté uživatelem zadávané manuální doučování. Proces přeučování je prováděn automaticky v druhém bloku metody. V průběhu morfologických úprav a nalézání regionu je nalezeno malé množství pixelů (nejvýše desítky), které jsou použity pro úpravu modelu. Tyto pixely jsou získány nejmenším možným morfologickým uzavřením (dilatace následovaná erozí s jádrem o velikosti 3×3 tvaru kříže) masky, která se v průběhu druhého bloku utváří a následným rozdílením s maskou původní. Tyto pixely představují nejdrobnější díry, které se nacházely v regionech pozadí, nebo ve vybraném regionu popředí. Hodnoty vybraných pixelů jsou přičteny do histogramů popředí a odečteny z histogramů pozadí, pokud se jednalo o díry v regionu popředí. Pro pixely, které představovaly díry v pozadí je provedena obdobná operace.



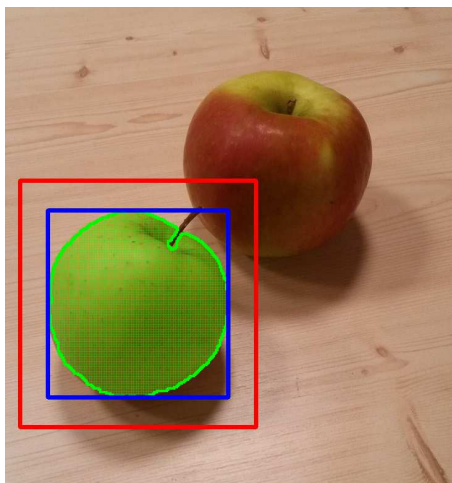
Obrázek 4.5: Automatické přeučování modelu. Vstupní snímek s uživatelským vstupem (vlevo nahoře), ostatní obrázky (směrem doprava) znázorňují výstup metody po různém počtu iterací přeučování (žádná, po dvou a po čtyřech). Další iterace již nezpůsobí žádnou změnu na výstupu.

Celé přeučování je navrženo, aby bylo použito iterativně. Experimentálně bylo prokázáno, že přeučování konverguje ve třech až deseti iteracích pro statický snímek, následující iterace přeučování nemají žádný další vliv na výsledek metody (znázorněno na obrázku 4.5). Přeučování je v metodě automaticky prováděno po každé manuální úpravě modelu a to ve třech iteracích. Další automatické provádění je nutné zavést při zpracování náhledu z kamery alespoň jednou za sekundu, aby byl částečně potlačen vliv měnících se jasových podmínek.

4.4 Vyhledávací okno

Jeden z nejdůležitějších prvků přispívajících k rychlosti celého algoritmu je pohyblivé vyhledávací okno. Jeho použití je inspirováno z [27], kde autoři uvažují, že algoritmus nemusí prohledávat celý snímek, ale pouze oblast posledního výskytu sledovaného objektu rozšířenou o určité okolí. Oblast je získána jako objekt ohraničující obdélník, rozšířená pro příští snímek o osminu šířky snímku a osminu výšky snímku. Nutná velikost rozšíření je zvolena experimentálně s ohledem na možnost efektivní implementace dělení mocninami dvou.

Pohyblivé vyhledávací okno je znázorněno na obrázku 4.6. Modrým orámováním je vyobrazena velikost ohraničujícího obdélníku, která je využita v posledním bloku metody a červeným orámováním je vyznačena velikost vyhledávacího okna pro následující snímek sekvence (videa nebo náhledu z kamery). Velikost a poloha ohraničujícího obdélníku je zjištěna během značení oblastí. Použití vyhledávacího okna má různý vliv na urychlení různých bloků metody. Klasifikace pixelů probíhá zpravidla na menším množství pixelů, než na celém snímku, morfologické a regionální úpravy jsou prováděny na ještě menším množství



Obrázek 4.6: Zobrazení plovoucích oken segmentační metody. Zeleně zvýrazněný je segmentovaný objekt, modré rámování značí okno pro zanesení výsledků do obrazu, červené rámování značí okno pro klasifikaci pixelů v příštím snímku.

než klasifikace. Celé urychlení metody je znázorněno v tabulce 4.1, kde jsou porovnávány časy metody, která využívá vyhledávacího okna na statických snímcích z různou velikostí segmentovaného objektu s variantou bez použití vyhledávacího okna.

Snímek	O	$t_{no-window}$	t_{window}	Zrychlení
image15.jpg	68%	8.89 ms	6.44 ms	1.38-krát
image09.jpg	54%	8.16 ms	6.17 ms	1.31-krát
image12.jpg	41%	8.31 ms	6.01 ms	1.38-krát
image07.jpg	33%	9.27 ms	6.14 ms	1.32-krát
image01.jpg	18%	8.48 ms	4.38 ms	1.93-krát
image18.jpg	8%	8.30 ms	4.26 ms	1.94-krát

Tabulka 4.1: V tabulce jsou uvedeny časy klasifikace několika testovacích snímků s různě velkými vyhledávacími okny a jejich porovnání s variantou bez použití vyhledávacího okna. O značí poměr velikosti segmentovaného objektu proti velikosti snímku. $t_{no-window}$ čas zpracování jednoho snímku bez použití vyhledávacího okna. t_{window} čas zpracování jednoho snímku s použitím vyhledávacího okna. Použité snímky jsou obsahem příloženého CD.

V případě, že dojde ke ztrátě sledovaného objektu, například rychlou změnou polohy kamery oproti objektu, je ztráta detekována v bloku morfologických úprav při značení oblastí. Případ ztráty je ošetřen nezměněním polohy a velikosti vyhledávacího okna, okno je pouze o osminu velikosti rozšířeno. Je tak zajištěno, že krajně v osmi snímcích bude probíhat klasifikace (hledání objektů z barev naučených v modelu popředí) nad úplně celým snímkem. Velikost a poloha pohyblivého vyhledávacího okna je jedna ze dvou znalostí o objektu, které metoda propaguje do následujících snímků. Druhá znalost je místo, kde začne algoritmus značení oblastí s vyplňováním masky. Tyto informace jsou v blokovém schématu (obrázek 4.1) značené jako blok *poloha objektu*

Při zadání uživatelských vstupů je nutné, aby byla velikost pohyblivého vyhledávacího okna přizpůsobena zadaným vstupům. Při prvním vytváření modelu, kdy ještě žádná klasifikace neprobíhá, je vyhledávací okno nastaveno jako obdélník ohraničující uživatelský vstup. Pro každý následující uživatelský vstup je kontrolováno, zda se uživatelské vstupy nachází uvnitř, nebo vně používaného vyhledávacího okna. V případě, že jsou uživatelské vstupy vně, je okno rozšířeno tak, aby svou plochou zcela pokrývalo okno z minulého snímku a zároveň ohraničovalo uživatelský vstup. Současně je upravena pozice, na které algoritmus pro značení oblastí začíná – tento prvek umožňuje díky uživatelskému vstupu vybrat jiný objekt ve vstupním snímku bez nutnosti čištění celého barevného modelu.

4.5 Provázání s jinými metodami

V požadavcích na tuto metodu bylo zmíněno, že hlavním požadavkem je rychlost celého algoritmu. Rychlosti je ovšem dosaženo na úkor kvality výsledné segmentace. Aby byla možné metodu zakomponovat do aplikace pro tvorbu fotografické koláže je nutné, aby bylo umožněno propojení s některou z metod zmíněných v kapitole 2. Metody založené na řezech v grafu dosahují významně lepších výsledků segmentace. Pro tyto metody umožňuje vytvořená segmentační metoda přichystat dostatek vstupních informací, aby mohla být na snímek ve vysokém rozlišení aplikována jedna z metod vynikající ve vysoké kvalitě výstupu (metody zmíněné v kapitole 2).



Obrázek 4.7: Provázání rychlé segmentační metody s metodou Global sampling matting pro získání kvalitní alfa masky. První dva obrázky ukazují vstupní obrázek se zaznačeným uživatelským vstupem a výstup navržené metody pro rychlou segmentaci. Na třetím obrázku je zobrazena automaticky vygenerovaná trimapa, která je vstupem další použité metody. Poslední dva obrázky ukazují výstupní alfa masku vytvořenou metodou Global sampling matting a výsledné smíšení s jiným pozadím na základě získané alfa masky.

Na obrázku 4.7 je zobrazena metodou automaticky vytvořená trimapa pro následující zpracování metodou *Global sampling matting*. Tímto je demonstrováno zamýšlené využití

vytvořené metody v praktické aplikaci. Pomocí jej můžeme zpracovávat náhled z kamery mobilního zařízení v reálném čase (a to i pro vysoká rozlišení), metoda v přijatelné kvalitě segmentuje objekt, můžeme s okamžitým výsledkem měnit polohu kamery, ale pro vytvoření kýženého výsledku se použije sofistikovanější metoda, aniž bychom byli nuceni jí dávat dodatečné uživatelské vstupy.

V implementované aplikaci pro mobilní telefon i pro PC jsou vytvořena provázání na metody *GrabCut* i *Global sampling matting*. Vstupními daty pro metody *Global sampling matting* je trimapa, která je snadno získána morfologickými operacemi nad výstupní maskou pro výřez objektu vytvořené metody. V případě metody *GrabCut* jsou vstupními daty ohraničující okno (získané jako velikost a pozice plovoucího vyhledávacího okna) a dodatečné vstupy získané z výstupní masky pro výřez objektu. Experimentálně bylo nastaveno, že metoda *GrabCut* je spuštěna ve dvou iteracích po zadání ohraničujícího okna a v dalších třech po dodání dodatečných vstupů.

Kapitola 5

Aplikace

Metoda popsaná v předchozí kapitole je implementována ve dvou verzích. Verze pro PC, umožňující zejména rychlé testování při změnách parametrů a verze pro mobilní zařízení integrovaná do demonstrační aplikace. Jako cílová platforma demonstrační aplikace je zvolena platforma Android. Jediným prvkem pro přizpůsobení metody cílové platformě je změna dimenzí třírozměrných histogramů barevného modelu. Namísto formátu RGB je ve většině zařízeních platformy Android použit formát YCrCb s kódováním NV21 pro uložení náhledu z kamery zařízení. Záměnou použitých modelů se nezmění funkce vytvořené metody.

Původní návrh pro segmentaci videa v reálném čase využíval metodu rozvodí. Tento návrh byl realizován pouze na PC. Ve verzi pro mobilní zařízení je použit výpočetně méně náročný algoritmus, který dosahuje stejné kvality výsledné segmentace.

5.1 UI pro segmentaci v reálném čase

Jedním z cílů této práce je navrhnout uživatelské rozhraní, které je jednoduché a intuitivní pro vytvořenou metodu. U metod zmíněných v kapitole 2 je dobře vidět, že různé metody vyžadují různé druhy uživatelských vstupů. Účelem vytvořené metody je minimalizace uživatelského vstupu. Shrňme několik druhů uživatelských vstupů. Různé druhy použité interakce poskytnou metodě různé množství informací o popředí a pozadí, ale také jsou různě náročné pro samotné zadání uživatelem. Tyto druhy jsou vyobrazené na obrázku 5.1 a jsou následující:

- **Zadání hranic trimapy:** Uživatel pomocí dvou druhů čar vyznačí hranice neznámé oblasti v trimapě (2.3.1) Oblasti, které se nenacházejí mezi dvěma druhy čar se označují jako jisté popředí a jisté pozadí. Tento způsob značení poskytne použité cílové metodě největší množství informací o objektu i pozadí. Ovšem zadávání je pro uživatele nejnáročnější ze zde zmíněných.
- **Zadání ohraničujícího obdélníku:** Uživatel zadá ohraničující obdélník, který obsahuje celý segmentovaný objekt. Způsob poskytne cílové metodě dobrou znalost o pozadí (vše vně ohraničujícího obdélníku je zcela jisté pozadí) a malou znalost o popředí (vše uvnitř ohraničujícího obdélníku směřuje pozadí a popředí). Pro zpřesnění výsledku může být tento způsob rozšířen možností zadávání polygonu pro zmenšení množství pixelů pozadí v neznámé oblasti. Tímto rozšířením ovšem roste uživatelská náročnost zadání vstupů.

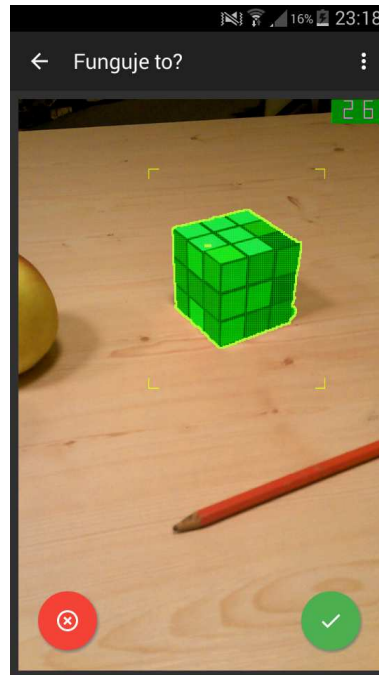
- **Hrubé tahy přes popředí a pozadí:** Uživatel pomocí dvou druhů čar vyznačí hrubě pozadí a popředí. Z pixelů, které jsou průnikem snímku a zadaných čar jsou sestaveny barevné modely pro popředí i pozadí. Problémem tohoto zadávání je případ, kdy uživatel některé barvy (část objektu o významně jiné barvě než zbytek objektu) čarou neoznačí.
- **Hrubé tahy přes popředí:** Uživatel pomocí tahu označí pouze segmentovaný objekt. Tento způsob není v metodách pro výřez používán, protože zcela chybí jakákoli informace o pozadí. Ovšem vytvořená metoda (kapitola 4) je navržena právě pro tento druh uživatelského vstupu.



Obrázek 5.1: Různé druhy uživatelské interakce. Zadání hranic trimapy (levá horní), zadání ohraničujícího obdélníku (pravá horní), hrubé tahy přes popředí a pozadí (levá dolní), hrubé tahy přes popředí (pravá dolní).

Vytvořená metoda je od začátku navrhována pro použití s nejmenším možným uživatelským vstupem a je do určité míry ošetřena chybějící informace o pozadí. Proto je zvolen do uživatelského rozhraní poslední zmíněný přístup k zadávání značení objektu, díky čemuž postačují k řízení metody pouze dva prvky uživatelského rozhraní: značení objektu tahem a tlačítko umožňující vynulování modelu. Na obrázku 5.2 je znázorněna ukázka navrženého minimalistického uživatelského rozhraní. Prvky, které jsou do znázorněného uživatelského rozhraní zavedeny jsou následující:

- **Plocha s náhledem kamery:** tahem na ploše je vytvořena čára – uživatelský vstup pro učení modelu metody.
- **Potvrzovací tlačítko:** pro pořízení snímku, vygenerování trimapy a spuštění metody globálního vzorkování pro pořízení kvalitní alfa masky.
- **Stornovací tlačítko:** pro vynulování barevného modelu metody.
- **Navigační lišta:** pro vedení uživatele v provádění akcí.



Obrázek 5.2: Sejmutý snímek obrazovky z mobilní aplikace pro provedení alfa testů. Důležité prvky UI: potvrzovací tlačítko, stornovací tlačítko a plocha s náhledem kamery.

5.2 UI pro kvalitní segmentaci

Původní návrh aplikace, představený v rámci semestrálního projektu uvažoval využití metody rozvodí pro segmentaci náhledu z kamery v reálném čase a metody GrabCut pro segmentaci snímku s vysokou kvalitou výstupu segmentace. Implementované řešení vyžadovalo při navázání metod dodatečné vstupy pro metodu GrabCut. S představenou nově vytvořenou metodou pro segmentaci videa na mobilním zařízení v reálném čase nejsou žádné dodatečné vstupy potřeba, pokud je dosaženo dostatečné kvality segmentace. Je vytvořeno navázání na metodu GrabCut a metodu globálního vzorkování. Případy, kdy vstupy metod nevytváří požadované výstupy, jsou zapříčiněny nízkou kvalitou segmentace předřazené vlastní metody (nevhodná scéna, špatné uživatelské vstupy) a vstupy je nutné upravit ručně.

5.3 Akcelerace výpočtů na mobilních zařízeních

Navržená metoda je výpočetně nenáročná, ovšem používá několik standardních operací pro zpracování obrazu, nabízí se tak možnost využít knihovny, která obsahuje potřebné operace v pokud možno optimální implementaci či hardwarově akcelerované. Pro implementaci metody na systému Android se nabízí využití knihovny FastCV[17] přímo od společnosti Qualcomm Technologies, výrobce hardwaru pro velký podíl zařízení využívající systém Android. Knihovna FastCV nabízí množství operací počítačového vidění optimalizované pro mobilní zařízení. Je navržená pro maximální efektivitu na všech procesorech založených na technologii ARM. Funkce jsou zejména vyladěny pro výhodné využití na procesorech Qualcomm Snapdragon.

Operace pro zpracování obrazu, které využívá navržená metoda z knihovny FastCV ve své implementaci pro mobilní telefony jsou následující:

- Gaussův filtr s konvolučním jádrem o velikosti 5×5 .
- Mediánový filtr s jádrem o velikosti 3×3 .
- Morfologické operace eroze a dilatace.
- Logické operace mezi obrazy pracující po jednotlivých bitech.

Tabulka 5.1 uvádí průměrnou dobu zpracování jednoho snímku operacemi z knihovny FastCV. Pro měření je použit testovací obraz (obrazy) v rozlišení 1920×1080 , tato velikost je největší možná, kterou obě použitá zařízení dovolují použít pro náhled z kamery. Použité testovací zařízení:

- Samsung Galaxy Note 3 N9005, procesor Qualcomm Krait 400 (Snapdragon 800), grafický procesor Qualcomm Adreno 330
- Samsung S4 mini 9195i, procesor Qualcomm Krait 300 (Snapdragon 400), grafický procesor Qualcomm Adreno 305

Naměřené časy ukazují, že i přes velké použité rozlišení testovacího snímku provádí knihovna operace velice rychle. Nejdůležitějším údajem jsou velice nízké hodnoty u morfologických operací, protože jsou v metodě použity opakovaně pro každý snímek.

	Note 3 [ms]	Note 3 FPS	S4 mini [ms]	S4 mini FPS
Gaussův filtr 5×5	17.49	57.2	20.07	49.8
Mediánový filtr 3×3	16.92	59.1	19.47	51.3
Logický AND po bitech	3.88	257.7	4.26	234.7
Eroze	3.26	306.7	4.09	244.5
Dilatace	3.18	314.5	3.94	253.8

Tabulka 5.1: Naměřené časy zpracování snímku o rozměrech 1920×1080 operacemi z knihovny FastCV na dvou různých mobilních zařízeních.

Další významným krokem, který by vedl k urychlení celé metody popsané v kapitole 4, je přesun funkce celého posledního funkčního bloku (obrázek 4.1), který zaznamenává výsledek segmentace, do grafické karty společně s renderováním snímku. V současné verzi jsou výsledky zaneseny do snímku v procesoru a poté je snímek renderován grafickou kartou.

Kapitola 6

Testování, vyhodnocení

V předchozích kapitolách byla popsána vytvořená segmentační metoda s vysokou rychlostí segmentace určená k provozu na mobilním zařízení ve vysoké snímkové frekvenci (kapitola 4) a její následné zakomponování do aplikace pro mobilní telefon, která disponuje minimalistickým uživatelským rozhraním pro zanesení uživatelských vstupů vytvořené segmentační metodě (kapitola 5). Předpoklady vytvořeného řešení jsou tedy následující:

- metoda zpracovává video na mobilním zařízení v reálném čase
- metoda dosahuje dostatečné kvality, produkuje minimální počet chyb segmentace
- uživatelské rozhraní je jednoduché a intuitivní

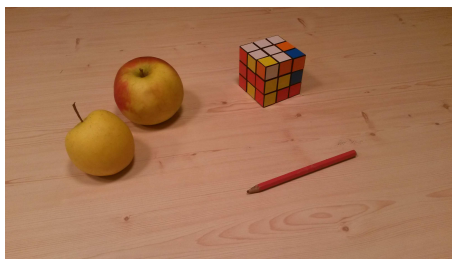
Pro potvrzení (nebo vyvrácení) stanovených předpokladů bylo nutné provedení testů vytvořené metody na datové sadě několika snímků a získání zpětné vazby přímo od potenciálních uživatelů pro posouzení kvality návrhu uživatelského rozhraní. V následujících podkapitolách jsou popsány techniky testování těchto prvků, jejich výsledky a vyhodnocení.

6.1 Testování UI rychlé segmentace

Vyhodnocení kvality návrhu a použitelnosti uživatelského rozhraní bylo hodnoceno technikou zvanou alfa testy. Alfa test je definován v [2] jako simulované, nebo reálné testování provozu na potenciálních uživateli (zákaznících) nebo na nezávislém testovacím týmu ve vývojové společnosti, kteří ovšem do vývoje nezasahují. Alfa testy bývají společně prováděny jako forma akceptačních testů. Díky alfa testům může vývojář zjistit, jak k testovanému softwaru přistupuje uživatel, který se neúčastnil vývoje.

Pro provedení alfa testů byla připravena umělá scéna obsahující několik objektů (z toho dva o totožných barvách) na kontrastním pozadí a vytvořena sada úkolů pro uživatele provádějící test. Testovány byly tyto případy užití:

- Označení jediného objektu, který je z výchozího pohledu na scénu rozměrný.
- Označení jediného objektu, který je z výchozího pohledu na scénu drobný.
- Označení objektu, jeho ztráta z náhledu kamery způsobená změnou polohy kamery, návrat kamery do původní pozice a znovu označení ztraceného objektu.
- Označení jednoho ze dvou stejnobarevných objektů, označení druhého, znovu označení prvního.



Obrázek 6.1: Nachystaná scéna pro provedení alfa testů obsahuje čtyři objekty pro plnění vytvořených úkolů. Pro testování jsou objekty zajímavé z hlediska vlastností: jeden rozměrný složený z více barev, jeden drobný, dva stejnobarevné.

Označením objektu rozumíme vytvoření takového výsledku segmentace o kterém je uživatel přesvědčen, že je dostatečný pro výřez objektu ze snímku a jeho vložení do jiného – provedení fotomontáže. Při plnění jednotlivých úkolů ze sady byly pozorovány následující parametry, na základě kterých je provedeno vyhodnocení kvality návrhu uživatelského rozhraní:

- doba nutná se splnění
- počet fatálních chyb segmentace – případ, kdy uživatel do modelu popředí zanesl příliš mnoho informací z pozadí
- počet použití restartovacího tlačítka – návratů k prázdným modelům popředí a pozadí
- kvalita segmentace jako počet chybějících pixelů segmentovaného objektu (false negative), při které uživatel označil úkol za splněný
- kvalita segmentace jako počet přebývajících pixelů segmentovaného objektu (false positive), při které uživatel označil úkol za splněný

Poslední dva body slouží k získání znalosti o přístupu uživatele k chybám segmentace. Z hlediska navázání metody rychlé segmentace na jiné, kvalitnější metody (viz 4.5) způsobují větší problém pixely, které jsou nesprávně označené jako popředí (false positive), než pixely, které jsou nesprávně označené jako pozadí (false negative).

	čas	zničený model	restart	false negative	false positive
úkol 1	4.3 s	0.4-krát	1.2-krát	7.69%	2.44%
úkol 2	8.8 s	1.2-krát	1.6-krát	11.86%	2.72%
úkol 3	18.3 s	-	-	8.63%	2.26%
úkol 4	23.5 s	-	-	9.03%	3.30%

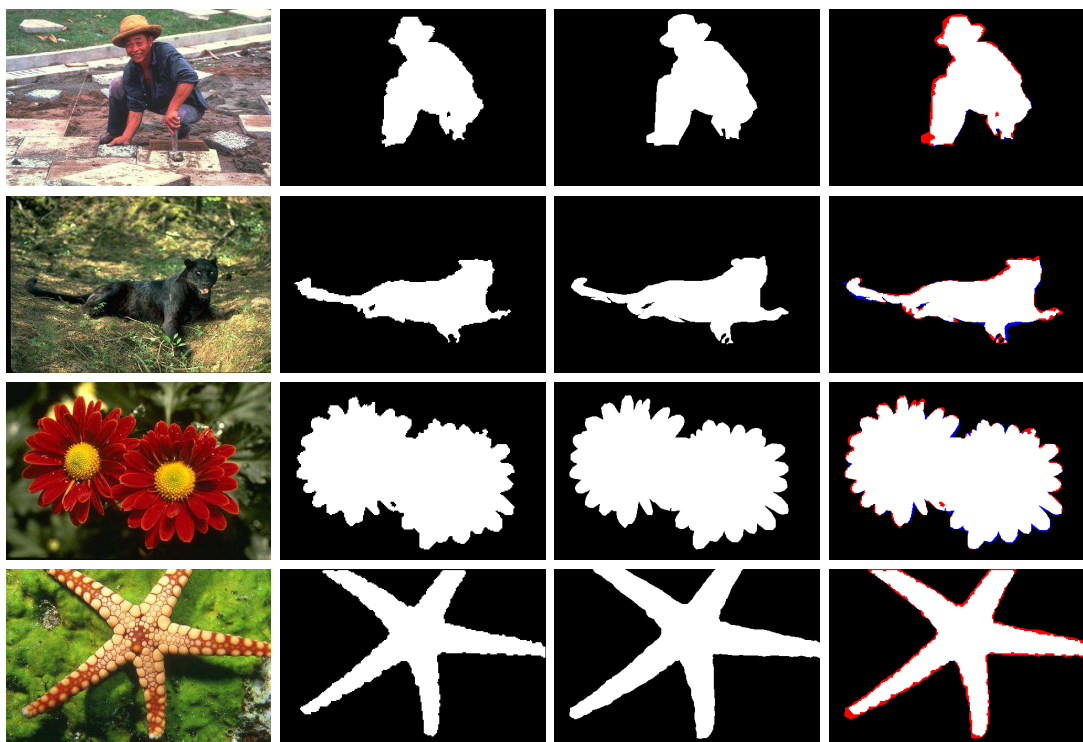
Tabulka 6.1: Výsledky alfa testů. Některé hodnoty nebyly pro některé úkoly zaznamenávány. Hodnoty uvedené v tabulce jsou průměrné získané z celé skupiny testerů.

V tabulce 6.1 jsou uvedené průměry naměřených hodnot. Skupina potenciačních uživatelů podstupující alfa testy sestávala z deseti uživatelů různých věkových kategorií, kteří nepracují a nestudují v oboru zpracování obrazu. Pozitivním výsledkem uživatelské zpětné

vazby je dostatečně malé množství chyb typu *false positive*. Dříve bylo řečeno, že vysokém množství těchto chyb způsobuje nemožnost navázat vytvořenou metodu na jiné metody. Vysoké průměrné časy pro splnění stanovených úkolů jsou způsobeny některými extrémními hodnotami, kdy si uživatel s úkolem nevěděl rady. Problém spočíval hlavně v příliš široké značící čáře, oproti drobnému objektu nebo jeho tenké části. Největším poznatkem z alfa testů je nutnost při použití metody a navrženého uživatelského rozhraní v aplikaci pro širokou veřejnost vytvoření videoprůvodce, nebo animace pro seznámení se způsobem, jak objekt označit, doznačit či změnit označení na jiný objekt.

6.2 Testování kvality výstupu rychlé segmentace

Pro vyhodnocení kvality výstupu vytvořené segmentační metody byl jako kritérium zvolen rozdíl výstupní masky metody oproti referenční, ručně anotované masce. Pro testy byly využity obrázky a částečně nachystané ruční anotace z testovací sady *Berkeley dataset*[14]. Berkeley dataset je rozsáhlá sada fotografií, které byly manuálně segmentovány. Tato testovací sada je volně dostupná a slouží jako trénovací množina algoritmům pro strojové učení nebo jako měřítko pro porovnávání různých segmentačních algoritmů a algoritmů pro detekci hranic objektů.



Obrázek 6.2: Ukázka několika testů na kvalitu výstupu vytvořené segmentační metody. V prvním sloupci jsou vstupní obrázky, v druhém sloupci je nevyhlazená výstupní maska pro výřez, ve třetím sloupci je vzorová maska – ruční anotace, v posledním sloupci je ukázán rozdíl mezi výstupní a referenční maskou – bílé pixely značí správně určené popředí, černé pixely správně určené pozadí, červené pixely nesprávně určené popředí jako pozadí (*false negative*), modré pixely nesprávně určené pozadí jako popředí (*false positive*).

Po účely testování tohoto projektu byla datová sada omezena pouze na podmnožinu použité sady. Účelem vytvořené metody je provést segmentaci obrazu pro nalezení jediného objektu (popředí) a možnost jeho oddělení od zbytku snímku (pozadí). Vybraná podmnožina datové sady, obsahující viditelně jeden oddělitelný objekt, byla rozdělena na dvě kategorie:

- fotografie s objektem, který je barevně kontrastní oproti pozadí
- fotografie s objektem, který sestává (byť částečně) ze stejných barev, jako pozadí

Snímek	O	$ErrC_{FP}$	$ErrC_{FN}$	$ErrC_{Total}$	Err_{image}	Err_{object}
image01.jpg	17.8%	158	2565	2723	1.76%	9.92%
image02.jpg	29.3%	1088	14801	15889	10.29%	35.12%
image03.jpg	35.3%	1793	4055	5848	3.79%	10.72%
image04.jpg	24.7%	30	5673	5703	3.69%	14.90%
image05.jpg	25.4%	530	2402	2932	1.90%	7.46%
image06.jpg	17.6%	16219	3202	19421	12.58%	71.36%
image07.jpg	33.9%	1425	3569	4994	3.23%	9.54%
image08.jpg	14.9%	882	1670	2552	1.65%	11.05%
image09.jpg	54.9%	748	6621	7369	4.77%	8.68%
image10.jpg	26.7%	2767	7244	10011	6.48%	24.21%
image11.jpg	16.4%	128	2908	3036	1.97%	11.94%
image12.jpg	41.6%	109	4550	4659	3.02%	7.24%
image13.jpg	12.1%	388	5652	6040	3.91%	32.16%
image14.jpg	9.1%	119	2316	2435	1.58%	17.26%
image15.jpg	68.4%	257	4475	4732	3.06%	4.48%
image16.jpg	42.1%	3540	2962	6502	4.21%	9.98%
image17.jpg	44.1%	2168	5430	7598	4.92%	11.14%
image18.jpg	8.1%	7514	1916	9430	6.11%	75.14%
image19.jpg	44.9%	1188	1917	3159	2.05%	4.55%
image20.jpg	26.5%	17	5320	5337	3.46%	13.02%

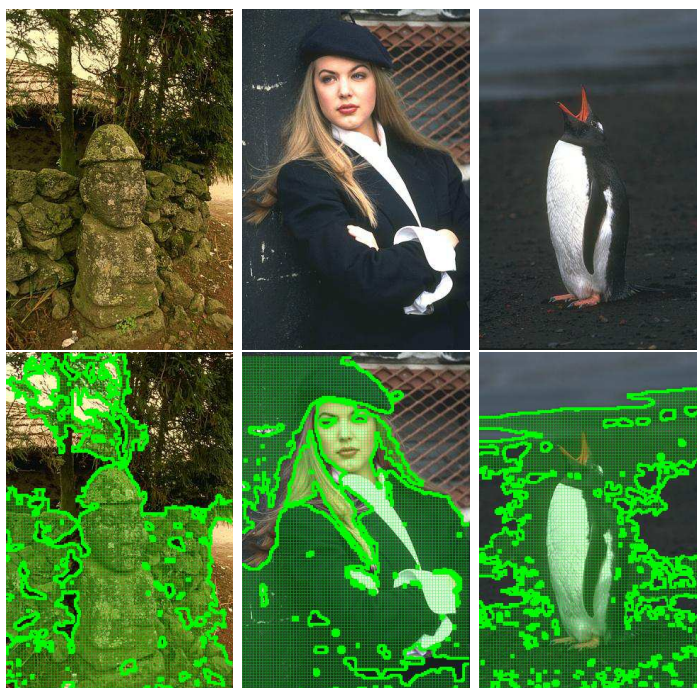
Tabulka 6.2: Naměřené hodnoty chyb ve výstupu vytvořené segmentační metody oproti ruční anotaci. Stejnomeně pojmenované snímky a veškeré obrázky jsou na příloženém CD. O značí poměr velikosti objektu proti celému snímku. $ErrC_{FP}$ je počet chyb typu I. $ErrC_{FN}$ je počet chyb typu II. $ErrC_{Total}$ je celkový počet chyb. Err_{image} udává poměr celkového počtu chyb proti velikosti snímku. Err_{object} udává poměr celkového počtu chyb proti velikosti objektu.

Samotný návrh vytvořené metody, založený na sestavení barevných modelů popředí a pozadí, neumožňuje správnou segmentaci fotografií, kde objekt není oproti pozadí dostatečně kontrastní. Proto byla pro testování uvažována pouze první zmíněná kategorie fotografií z datové sady. Pro každý testovací obrázek byla vytvořena ruční anotace vzorové masky pro výřez – binární obraz. Ten byl srovnávám se surovou nevyhlazenou výstupní maskou vytvořené metody a byly zaznamenávány rozdíly mezi těmito dvěma maskami (srovnání je na obrázku 6.2). Výsledkem srovnání je počet chyb segmentace, které podle [15] rozdělujeme na dva typy:

- **Chyba typu I** (nebo také *false positive*) je chybné zamítnutí pravdivé nulové hypotézy, tedy pixel je metodou označený jako popředí, přestože v referenční masce je označený jako pozadí.
- **Chyba typu II** (nebo také *false negative*) je selhání v zamítnutí nepravdivé nulové hypotézy, tedy pixel je metodou označený jako pozadí, přestože v referenční masce je označený jako popředí.

V testech provázání vytvořené metody (podkapitola 4.5) s jinými metodami se ukázalo, že přítomnost většího množství chyb typu I má závažnější následky na výslednou kvalitu navázané metody. Velké množství chyb typu I snižuje použitelnost výsledku segmentace pro další účely.

V tabulce 6.2 jsou uvedeny výsledky testů – chyby ve výstupech segmentace a jejich porovnání s velikostí oblastí celého snímku a oblastí segmentovaného objektu. V sloupci tabulky Err_{object} je uveden poměr, který nejvíce vypovídá o celkové úspěšnosti segmentace daného objektu. Průměrná úspěšnost vytvořeného algoritmu dosahuje na této sadě hodnoty 80.5% (medián 88.9%). Vysoké úspěšnosti je dosaženo zejména výběrem pouze té podmnožiny snímků z testovací sady, pro kterou je metoda navržena. Vysoké míry neúspěchu dosahuje metoda u snímků, ve kterých je objekt nedostatečně kontrastní oproti pozadí (viz obrázek 6.3).



Obrázek 6.3: Obrázky nevhodné pro segmentaci vytvořenou metodou. Objekt popředí je oproti pozadí příliš málo kontrastní.

Kapitola 7

Závěr

Jedním z cílů diplomové práce bylo studium existujících technik segmentace popředí a pozadí, zaměření se zejména na metody, které by umožňovaly segmentaci videa v reálném čase na mobilním zařízení a metody, které umožňují získání kvalitních alfa masek v nezávislosti na výpočetní náročnosti. V práci jsou zmíněné všechny metody, které významně přispěly k vytvoření konečného řešení. Pro všechny metody, které jsou v práci diskutovány je navrženo jejich uplatnění v tomto projektu.

Je navržena a implementována výpočetně nenáročná metoda pro výřez objektu z videa v reálném čase, která vyžaduje výjimečně malé množství uživatelských vstupů. Jsou představeny prvky metody, které vedou k vysoké rychlosti segmentace a možnost provázání s jinými metodami pro zajištění vysoké kvality segmentace. Pro testování metody byla vytvořena ručně anotovaná testovací sada, na které dosahovala metoda průměrné úspěšnosti 80% při měření odlišností proti ručně anotovaným datům. Naměřený čas v nachystané testovací datové sadě na zpracování jednoho snímku na PC (procesor Intel i5) v rozlišení 481×321 nabývá průměrné hodnoty 5,46 ms (zpracování 188 snímků za sekundu). V použitém mobilním zařízení (Samsung S4 mini 9095i) dosahuje metoda snímkové frekvence 10 až 30 snímků za sekundu v závislosti na velikosti segmentovaného objektu z náhledu kamery v rozlišení 1280×720 . Použití metody je z důvodu samotného principu omezené pouze na obrázky s dostatečně kontrastním objektem oproti pozadí.

Vytvořená metoda je zasazena do demonstrační aplikace pro mobilní telefony i pro PC. Demonstrační aplikace ukazuje návrh uživatelského rozhraní pro rychlou segmentaci náhledu z kamery telefonu, umožňující jednoduché a intuitivní zadávání uživatelských vstupů pro segmentační metodu. Implementovaná metoda ve variantě pro PC demonstruje možnost jednoho případu užití – záměna pozadí pro videokonferenci. Je implementována varianta segmentace v reálném čase založená na metodě rozvodí (pouze ve variantě pro PC), která ovšem z důvodů temporální nestability nebyla implementována pro mobilní zařízení. Demonstrační aplikace pro mobilní zařízení byla využita pro získání uživatelské zpětné vazby k návrhu uživatelského rozhraní a způsobu zadávání vstupů vytvořené metodě. Mobilní aplikace je z důvodů použitých knihoven cílena na mobilní zařízení s platformou Android.

Možných pokračování nebo rozšíření této práce je několik. V návrhu vytvořené metody je zmíněn blok pro zanesení výsledků segmentace. V současné implementaci pro mobilní zařízení je toto zapsání výsledků segmentace prováděno v procesoru. Do vstupního snímku je při průchodu přes pixely zapisován výsledek segmentace a poté je snímek odeslán na grafickou kartu k vykreslení. Zapisování výsledků by mohla provádět přímo grafická karta. Profilování aplikace ukázalo, že tento krok je úzkým hrdlem algoritmu, proto by bylo vhodné jeho odstranění při dalším úsilí pro zrychlení metody.

Významným krokem je vytvoření aplikace pro veřejnost, která umožňuje rychlou a snadnou tvorbu fotomontáží. Uživatel by mohl měnit polohu, velikost a otočení vkládaného objektu do jiné fotografie pouhou změnou polohy svého mobilního zařízení. Technologie zavedené a vyvinuté v této práci jsou dostačující na vytvoření takové aplikace. To by umožnilo získání masivní uživatelské zpětné vazby na návrh uživatelského rozhraní a funkci vytvořené rychlé segmentační metody. Nabízí se ovšem úvaha, zda v aplikaci pro veřejnost je vhodné, aby probíhala tvorba konečné fotomontáže (použití metody pro získání kvalitní alfa masky) přímo v mobilním zařízení potenciálního zákazníka, nebo byla zpracována na vzdáleném serveru.

Dalším možným pokračováním je změna návrhu vytvořené segmentační metoda ve fázi klasifikace pixelů. Použité histogramy na reprezentaci barevných modelů popředí a pozadí nejsou dostatečně robustní v situacích, kdy objekt popředí sestává ze stejných nebo podobných barev jako jeho pozadí. Zařízení, která byla použita na testování a měření doby zpracování jednoho snímku, jsou již dva roky na trhu. Můžeme očekávat, že zařízení mladší, nebo zařízení, která v blízké budoucnosti budou na trh uvedena, budou disponovat vyšším výpočetním výkonem. Tím vznikají jisté výkonové rezervy dovolující použití výpočetně náročnějšího klasifikačního modulu.

Literatura

- [1] Beucher, S.: The Watershed Transformation [online].
<http://cmm.enscm.fr/~beucher/wtshed.html>, 2010-05-11 [cit. 2014-12-29].
- [2] Board, I. S. T. Q.: Standard glossary of terms used in Software Testing. October 2012.
- [3] Boykov, Y.; Veksler, O.; Zabih, R.: Fast Approximate Energy Minimization via Graph Cuts. *IEEE Trans. Pattern Anal. Mach. Intell.*, Listopad 2001: s. 1222–1239, ISSN 0162-8828, doi:10.1109/34.969114.
- [4] Canny, J.: A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1986.
- [5] Celebi, M. E.: A simple and efficient algorithm for connected component labeling in color images. 2012, doi:10.1117/12.911047.
URL <<http://dx.doi.org/10.1117/12.911047>>
- [6] Cormen, T. H.; Stein, C.; Rivest, R. L.; aj.: *Introduction to Algorithms*. McGraw-Hill Higher Education, 2001, ISBN 0070131511.
- [7] Gomes, R. B.; Aroca, R. V.; de Carvalho, B. M.; aj.: Real Time Interactive Image Segmentation Using User Indicated Real-World Seeds. In *SIBGRAPI*, IEEE Computer Society, 2012, ISBN 978-1-4673-2802-9, s. 182–189.
URL <<http://dblp.uni-trier.de/db/conf/sibgrapi/sibgrapi2012.html#GomesACG12>>
- [8] He, K.; Rhemann, C.; Rother, C.; aj.: A global sampling method for alpha matting. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, June 2011, ISSN 1063-6919, s. 2049–2056, doi:10.1109/CVPR.2011.5995495.
- [9] Huang, T.; Yang, G.; Tang, G.: A fast two-dimensional median filtering algorithm. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, ročník 27, č. 1, Feb 1979: s. 13–18, ISSN 0096-3518, doi:10.1109/TASSP.1979.1163188.
- [10] Jiřík, M.; Ryba, T.: Texture Based Segmentation Using Graph Cut and Gabor Filters. 2010.
URL <http://www.kky.zcu.cz/en/publications/JirikM_2010_TextureBased>
- [11] Kass, M.; Witkin, A.; Terzopoulos, D.: Snakes: Active contour models. *INTERNATIONAL JOURNAL OF COMPUTER VISION*, ročník 1, č. 4, 1988: s. 321–331.

- [12] Levin, A.; Lischinski, D.; Weiss, Y.: A Closed-Form Solution to Natural Image Matting. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, ročník 30, č. 2, Feb 2008: s. 228–242, ISSN 0162-8828, doi:10.1109/TPAMI.2007.1177.
- [13] Li, Y.; Sun, J.; Tang, C.-K.; aj.: Lazy Snapping. *ACM Trans. Graph.*, August 2004: s. 303–308, ISSN 0730-0301.
- [14] Martin, D.; Fowlkes, C.; Tal, D.; aj.: A Database of Human Segmented Natural Images and its Application to Evaluating Segmentation Algorithms and Measuring Ecological Statistics. In *Proc. 8th Int'l Conf. Computer Vision*, ročník 2, July 2001, s. 416–423.
- [15] Peck, R.; Devore, J.: *Statistics: The Exploration & Analysis of Data*. Available Titles Aplia Series, Cengage Learning, 2011, ISBN 9780840058010.
URL <<https://books.google.cz/books?id=NsAh3P-WrswC>>
- [16] Pukelsheim, F.: The Three Sigma Rule. *The American Statistician*, ročník 48, č. 2, may 1994: s. 88–91, doi:10.1080/00031305.1994.10476030.
URL <<http://dx.doi.org/10.1080/00031305.1994.10476030>>
- [17] Qualcomm Technologies, I.: Computer Vision (FastCV) [online].
<https://developer.qualcomm.com/mobile-development/add-advanced-features/computer-vision-fastcv>, 2010-2015 [cit. 2015-05-20].
- [18] Radke, R. J.: *Computer Vision for Visual Effects*. New York, NY, USA: Cambridge University Press, 2012, ISBN 0521766877, 9780521766876.
- [19] Roerdink, J. B.; Meijster, A.: The Watershed Transform: Definitions, Algorithms and Parallelization Strategies. *Fundam. Inf.*, Duben 2000: s. 187–228, ISSN 0169-2968.
- [20] Rother, C.; Kolmogorov, V.; Blake, A.: GrabCut -Interactive Foreground Extraction using Iterated Graph Cuts. *ACM Transactions on Graphics (SIGGRAPH)*, August 2004.
URL <<http://research.microsoft.com/apps/pubs/default.aspx?id=67890>>
- [21] Ruzon, M. A.; Tomasi, C.: Alpha Estimation in Natural Images. In *CVPR*, IEEE Computer Society, 2000, ISBN 0-7695-0662-3, s. 1018–1025.
URL <<http://dblp.uni-trier.de/db/conf/cvpr/cvpr2000.html#RuzonT00>>
- [22] Serra, J.: *Image Analysis and Mathematical Morphology*. Orlando, FL, USA: Academic Press, Inc., 1983, ISBN 0126372403.
- [23] Shapiro, L.; Stockman, G.: *Computer Vision*. Prentice Hall, 2001, ISBN 9780130307965.
- [24] Siddharth Srivastava, J. S. K. S. B., Aditya Rastogi: Study and Implementation of Image and Video Matting Techniquess. University School of Information Technology, GGS Indraprastha University, Kashmere Gate, Delhi 110006, 2013.
- [25] di Stefano, L.; Bulgarelli, A.: A Simple and Efficient Connected Components Labeling Algorithm. In *Proceedings of the 10th International Conference on Image Analysis and Processing, ICIAP '99*, Washington, DC, USA: IEEE Computer Society, 1999, ISBN 0-7695-0040-4, s. 322–.

- [26] Sun, J.; Jia, J.; Tang, C.-K.; aj.: Poisson Matting. *ACM Trans. Graph.*, ročník 23, č. 3, Srpen 2004: s. 315–321, ISSN 0730-0301, doi:10.1145/1015706.1015721.
- [27] Wang, D.; Lu, H.; Xiao, Z.; aj.: Fast and effective color-based object tracking by boosted color distribution. *Pattern Analysis and Applications*, ročník 16, č. 4, 2013: s. 647–661, ISSN 1433-7541, doi:10.1007/s10044-013-0347-5.
URL <<http://dx.doi.org/10.1007/s10044-013-0347-5>>
- [28] Zhang, H.: The Optimality of Naive Bayes. In *Proceedings of the Seventeenth International Florida Artificial Intelligence Research Society Conference (FLAIRS 2004)*, editace V. Barr; Z. Markov, AAAI Press, 2004.

Příloha A

Obsah CD

Příložený datový nosič obsahuje následující:

- Adresář `app`: obsahuje zdrojové texty vytvořených aplikací v průběhu řešení práce.
- Adresář `tex`: obsahuje zdrojové texty a obrázky textové části práce pro systém \LaTeX .
- Adresář `video`: obsahuje pořízená videa.
- Adresář `benchmark`: obsahuje výsledky testů vytvořené metody na kvalitu výstupu segmentace.
- Adresář `alpha_test`: obsahuje výsledky uživatelské zpětné vazby
- `dokumentace.pdf`: tento dokument
- `README.txt`: textový soubor s popisem obsahu CD