

ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE

Fakulta elektrotechnická

Katedra radioelektroniky

Metody klíčování obrazu

Květen 2015

Bakalant: Martin Hlinovský
Vedoucí práce: Ing. Stanislav Vitek, Ph.D.

Čestné prohlášení

Prohlašuji, že jsem zadanou bakalářskou práci zpracoval sám s přispěním vedoucího práce a konzultanta a používal jsem pouze literaturu v práci uvedenou. Dále prohlašuji, že nemám námitek proti půjčování nebo zveřejňování mé bakalářské práce nebo její části se souhlasem katedry.

Datum: 22. 5. 2015

.....

podpis bakalanta

České vysoké učení technické v Praze
Fakulta elektrotechnická
katedra radioelektroniky

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

Student: **Martin Hlinovský**

Studijní program: Komunikace, multimédia a elektronika
Obor: Multimediální technika

Název tématu: **Metody klíčování obrazu**

Pokyny pro vypracování:

1. Prostudujte metody klíčování obrazu.
2. Vybrané metody implementujte v programovém prostředí Matlab tak, aby bylo možné vybrat libovolnou barvu pozadí.
3. Připravte vhodné videosekvence na testování algoritmů.
4. Diskutujte robustnost algoritmů s ohledem na světelné podmínky scény a výběr barvy pozadí.

Seznam odborné literatury:

- [1] Foster J., The Green Screen Handbook: Real-World Production Techniques, Sybex, 2010, ISBN 978-0470521076
[2] Gonzales R.C., Woods R.V., Digital Image Processing Using MATLAB, Gatesmark Publishing, 2009, ISBN 978-0982085400

Vedoucí: Ing. Stanislav Vitek, Ph.D.

Platnost zadání: do konce letního semestru 2015/2016

L.S.

doc. Mgr. Petr Páta, Ph.D.
vedoucí katedry

prof. Ing. Pavel Ripka, CSc.
děkan

V Praze dne 10. 2. 2015

Anotace:

Tato práce se věnuje principu klíčování obrazu a základům, které je potřeba dodržovat jako osvětlení, barva plátna nebo oblečení pro dosažení přirozeného výsledku. Je zde krátce zmíněn princip virtuálního studia. Popisuje k čemu je parametr alfa a jak ho lze využít pro klíčování obrazu. Dále se věnuje barevným prostorům RGB, HSV a LAB. Poslední kapitola se zabývá shlukovou analýzou a jejím využitím pro třídění dat, problémem k-means clustering, jeho základními vlastnostmi a uplatněním. Jsou zde také zmíněné některé závěrečné úpravy, které se používají pro dosažení lepších a opravdovějších výsledků, například rozmazání hran nebo gaussian mixture models.

Klíčová slova: klíčování, parametr alfa, RGB, HSV, LAB, shluková analýza, k-means clustering, lloydův algoritmus, gaussian mixture models

Summary:

This final project deals with principle of blue screen matting methods and basic features that need to be respected as lighting, color backdrop or clothing to achieve a natural result. There is described the principle of the virtual studio and use of the alpha parameter. In the third part are the basics of color spaces RGB and HSV lab. The last chapter deals with cluster analysis and its use for sorting data, the problem of k-means clustering, its application. There are also mentioned some final adjustments, which are used to achieve a better and more realistic results, for example, blurring the edges or mixture Gaussian models.

Index Terms: chroma keying, green screen, blue screen, alpha channel, RGB, HSV, LAB, cluster analysis, k-means clustering, lloyd's algorithm, gaussian mixture models

Obsah

1	Úvod.....	1
1.1	Virtuální studio	3
2	Klíčování	5
2.1	Parametr alfa	5
2.2	Oblečení.....	8
2.3	Typy pozadí	8
2.4	Osvětlení.....	9
2.5	Použití	10
2.6	Pre - klíčování	11
2.7	Výhody digitálního klíčování.....	12
3	Návrh metodiky	13
3.1	RGB	13
3.2	Vlastnosti RGB.....	14
3.3	Vztah mezi RGB a HSV	15
3.4	HSV.....	15
3.5	Vlastnosti HSV	16
3.6	LAB	17
4	Algoritmy	18
4.1	Shlukování	18
4.2	K-means clustering	19
4.3	Použití	20
4.4	Lloydův algoritmus	22
4.5	Voroného diagram.....	22
4.6	Binární alfa.....	23
4.7	Postprocessing	25
4.8	Rozmazání	26
4.9	Gaussian Mixture Models.....	27
4.10	Vyhodnocení efektivity algoritmu.....	278
5	Závěr.....	32
6	Zdroje	33
7	Přílohy	34

1 Úvod

Poslední dobou vidáme ve filmech spoustu zajímavých scén, o kterých si říkáme, jak by mohly být asi tak natočené. Navštěvujeme zde místa jako z jiného světa, obrovské chrámy, překrásné historické budovy. Na filmové scéně se pohybují všemožné stroje, které v našem světě vůbec neexistují. Roboti, kteří se mění v auta, vidíme létat různé super hrdiny mezi mrakodrapy nebo obrovská monstra, která ničí města.

Ve filmech jsou často zobrazena místa z celého světa od deštných pralesů po ledové póly, svět v období od pravěku až do budoucnosti. Dříve, když výpočetní technika nebyla ještě tak rozvinutá, bylo k natočení takovýchto snímků potřeba cestovat po celém světě, abychom získali správný záběr. Pokud chtěl český režisér natočit pár snímků u moře, musel se k němu vypravit i kvůli deseti sekundovému záběru. Bylo nutné použít spoustu triků k natočení scén pod mořem, ve vzduchu nebo ve vesmíru. Filmový štáb potřeboval k nafilmování hezkých záběrů mnoho rekvizit a speciálního vybavení, aby divák věřil, že jsou záběry skutečné. Natočení kvalitního filmu se tak stále více prodražovalo a jen velká filmová studia si mohla dovolit zrealizovat celovečerní film.

Další zajímavostí, které si každý jistě všiml, když kouká na televizní zprávy a vidí televizní zpravodajku, jak se prochází po mapě České republiky, pohybem ruky přivolává další mapy a ukazuje na nich aktuální počasí. Mapy se za ní postupně mění od zobrazení evropského kontinentu po lokální oblasti a ona vysvětluje, jaká bude zítra teplota vzduchu.

Při sportovních událostech vidíme sportovní komentátory, kteří sedí v televizním studiu v České republice. Při tom se baví se známým fotbalistou o poločase, který zrovna hraje na mistrovství světa v Brazílii. Všechny tyto speciální efekty by se jen těžko realizovaly bez pomoci triku zvaného klíčování.

Při použití tohoto efektu můžeme vzít libovolný objekt, člověka, zvíře klidně i celé budovy a umístit je do prostředí, které požadujeme. Prostředí může být fotka pořízená na dovolené u moře, v horách nebo jako u televizních novin mapa České republiky s jednotlivými údaji o počasí. Můžeme pracovat s libovolnou reálnou fotografií, další možností je vytvoření fiktivního prostředí pomocí různých 3D programů. V 3D programech, jsme omezeni jen naší vlastní fantazií a nic nám nebrání vyrobit libovolné pozadí. Můžeme si vymodelovat hluboká údolí, vysoké hory, podmořské hlubiny. Nebo vnitřní prostory výzkumných nebo vesmírných stanic. Prakticky všechny věci, které jsme

schopti vymyslet naší vlastní fantazií. Na všechna takto vytvořená pozadí pak mohou být umístěny jiné objekty.

Například, když chceme natočit historický film ve vnitřních budovách zámku a nemáme peníze na pronájem prostor. Nebo je také možné, že se v zámku nesmí natáčet z důvodu poškození památky, tak bez moderních technologií by nebylo možné tento film realizovat. Za použití techniky klíčování si nejdříve vyfotíme vnitřní prostory jiného zámku a poté je upravíme do námi požadovaného tvaru, nebo si vytvoříme úplně identické prostory na počítači a sloučíme je se snímky postav oblečených do dobových oděvů.

Jedním z hlavních požadavků pro používání klíčování je snaha sloučit dva snímky tak, abychom z prvního mohli vzít pouze postavy nebo různé objekty a vložit je na druhý snímek, který by sloužil jako pozadí.

V angličtině se můžeme setkat s názvy pro tuto techniku jako "chroma key compositing", "chroma keying" nebo také "green screen", "blue screen". Poslední dva názvy jsou vyjádřeny pomocí barvy plátna, na kterém se klíčování provádí.

Celý princip klíčování je velice jednoduchý. Pracuje se základní myšlenkou, že ze snímku kde chceme vyjmout jen určité objekty, tak ty necháme a odstraníme nebo zprůhledníme vše ostatní a vložíme to na jiný snímek.

Požadovaný objekt, je jedno jestli je to postava, zvíře nebo budova se natočí na jednobarevném pozadí, nejčastěji se používá modré nebo zelené pozadí, ale v některých případech se pracuje i s jinými barvami třeba červenou.

Na obrázku číslo 1 můžeme vidět, jak vypadá takové natáčení na zeleném pozadí s jednotlivým osvětlením a zaoblenými rohy zeleného plátna tak, aby se vyloučily stíny na okrajích. Takto se například můžou natáčet zprávy o počasí, kde se vše ostatní přidává v reálném čase pomocí počítače.



Obr. 1. ukázka klíčovacího studia

Po natočení tohoto snímku se tato barva takzvaně "vyklíčuje", neboli nahradí jiným námi zvoleným pozadím. Zaprvé se zvolí základní rozsah barev, který má být v celém snímku nahrazen. Tyto barvy se odstraní nebo zprůhlední, efekt je stejný tak, aby byl za nimi vidět snímek pozadí.

Efekt zprůhlednění funguje, že každému pixelu ve snímku se přiřadí určitá hodnota, která udává jeho průhlednost. Tato hodnota se může vypočítat podle vztahu.

$$F = \alpha \cdot A + (1 - \alpha) \cdot B \quad [4]$$

Kde A a B jsou snímky popředí a pozadí, parametr alfa udává, s jakou intenzitou budou oba obrázky zobrazeny. Po spojení snímků se oba prolnou, že tam kde má jeden největší intenzitu bude mít druhý nejmenší a naopak. Více se budu tomuto tématu věnovat v kapitole číslo 2.7 nazvané parametr alfa.

V druhé kapitole jsou popsány některé principy, které je nutné dodržet při natáčení scény, kterou chceme klíčovat a některé obecné vlastnosti o klíčování. V kapitole číslo tři se věnuji třem barevným prostorům RGB, HSV a LAB. Popisuji zde důvody, na co se jaký prostor hodí lépe. Ve čtvrté kapitole se zabývám algoritmy, které používám pro klíčování v matlabu.

1.1 Virtuální studio

Virtuální studio je televizní studio, které umožňuje v reálném čase zpracovávat dohromady kombinaci postav nebo objektů a počítačem vytvořené prostředí. Pro propojení těchto objektů se používá právě technologie klíčování. Rozdílem oproti

samotnému klíčování je to, že kamera se může pohybovat ve 3D prostoru a obraz pozadí je vykreslen ve stejném čase a ze stejného pohledu kamery. Scéna se musí přizpůsobit jakémukoli pohybu kamery (přiblížení, pan, úhel, pohyb ve všech osách). Tím se také liší od způsobu zpracování ve filmu, kde se výsledné záběry dotvářejí až po natočení v post produkci. Ve virtuálním studiu se nepoužívá žádná post produkce. Potřebujeme pouze 3D grafika a nějaký grafický program, který vytvoří vkládané pozadí třeba výše zmiňovanou mapu České republiky a nějakou grafiku do popředí, nejčastěji logo televize, čas a někdy titulky.

Pro vytvoření virtuálního studia se používá hodně technik, nejčastější je hlavně tato "camera tracking". Při této technice se používá optické nebo mechanické získávání souřadnic, kde se právě kamera nachází. Zaznamenávají se souřadnice místa, přiblížení a úhel natočení kamery. Podle těchto údajů se vytvoří obraz, jak bude televizní studio vypadat z tohoto zaznamenaného pohledu a poté se vše zkombinuje do výsledného snímku. Hlavní rozdíl mezi virtuálním studiem a klíčováním je ten, že ve virtuálním studiu se vše provádí v reálném čase a je možné tuto techniku používat v živém vysílání v televizi. Na rozdíl od metody klíčování, která se využívá hlavně ve filmovém průmyslu.

2 Klíčování

Klíčování můžeme rozdělit na analogové a digitální. V analogové televizi je barevný tón reprezentován jako fáze přenášeného signálu. Klíčování spočívá v tom, že se porovnává fáze barvy, kterou chceme odstranit vůči fázi filmového signálu a když dojde ke shodě, tak se nahradí námi požadovaným signálem s určitým pozadím.

V digitálním světě máme práci podstatně jednodušší, protože barva je zde reprezentována barevnými souřadnicemi. Obraz, se projíždí bod po bodu, a když je nalezen pixel hledané barvy, tak se nahradí požadovaným bodem umístěným na stejném místě ve snímku, který chceme jako pozadí. V dalším textu se budu zabývat především digitálním klíčováním.

2.1 Parametr alfa

Jedním z prvních kdo přišel s nápadem průhlednosti u obrázku, neboli alfa kanálem byl Alvy Ray Smith na konci 70 let. Jeho nápad později rozvedli roku 1984 Thomas Porter a Tom Duff. Ti také definovali operátory pro práci s hodnotou alfa využívané v počítačích "over", "in", "out", "atop" a "xor". Tyto operátory jsou často používány v grafických programech pro práci s vrstvami.

V počítačové grafice se pomocí alfa kanálu dají sloučit dva snímky dohromady tak, že se budou buď úplně překrývat, nebo budou navzájem průsvitné a uvidíme je oba. Často se to používá, když se pracuje s jednotlivými vrstvami snímku, kdy nemůžeme zasahovat do ostatních vrstev, a na konci práce se všechny vrstvy sloučí do jedné. K tomu je důležité, abychom někde uchovávali hodnotu, jak moc je daný snímek průhledný a bude moc být překryt jiným snímkem.

K této práci byl přidán k barevným složkám ještě alfa kanál, který udává průhlednost jednotlivých obrazových bodů. Hodnota, která popisuje, s jakou intenzitou jsou jednotlivé body zobrazeny. Když jeden obrázek překrývá druhý, potom bude každý pixel zobrazen s intenzitou zadanou právě pomocí parametru alfa.

Alfa kanál nabývá hodnot od nuly do jedničky. Hodnota nula říká, že snímek nebude zobrazen vůbec, je úplně průsvitný. Při hodnotě jedna je snímek celý neprůhledný a nic jiného kromě něj nebude zobrazeno. Například když v RGBA máme hodnotu zadanou (0.0, 0.5, 0.0, 0.5), tak tento pixel bude mít 50% zelenou intenzitu a bude 50% průhledný.

Hodnota alfa kanálu může být součástí jednotlivých barevných formátů například PNG, RGBA nebo může být také zadána jako bitmapa. Pak mluvíme o takzvané bitmapové masce obrázku. Hodnota alfa jde vypočítat podle vzorce.

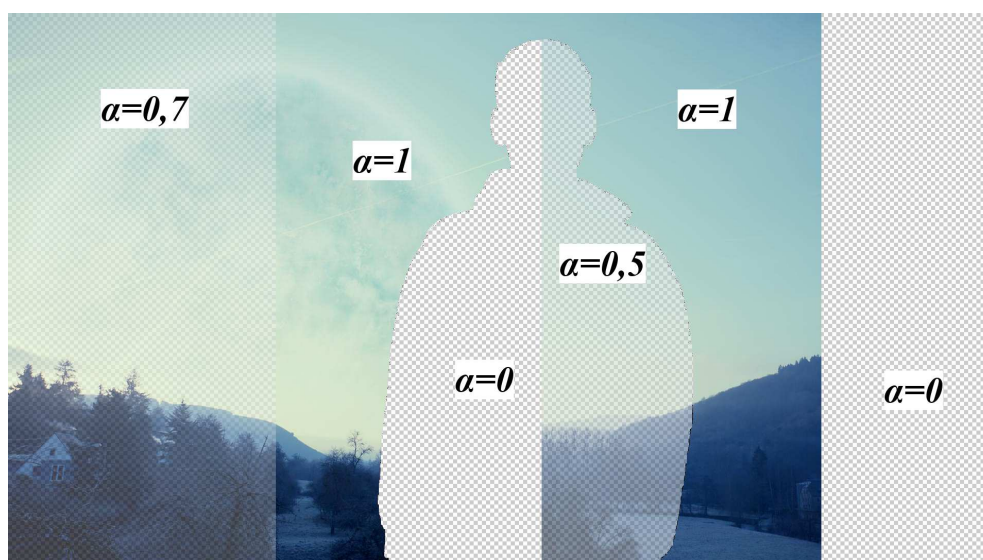
$$F = \alpha \cdot A + (1 - \alpha) \cdot B \quad [4]$$

Kde F je výsledný obraz, α je alfa hodnota. Když $\alpha = 1$ tak pixel obrazu A má 100% intenzitu a pixel B bude mít 0% intenzitu. Pro $\alpha = 0,5$ bude mít pixel A 50% intenzitu a pixel B také 50% intenzitu.

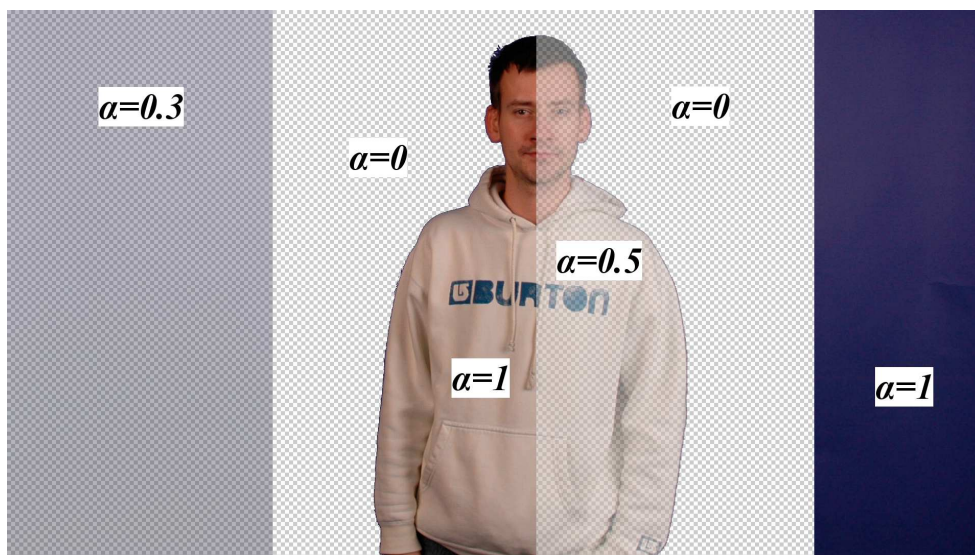
Když klíčováním určíme, která barva má být odstraněna, máme dvě možnosti, jak sloučit oba snímky dohromady. Buď, procházíme celý obrázek pixel po pixelu a místo bodů, které mají být odstraněny vložit body z pozadí.

Druhou možností je nastavit u jednotlivých bodů, které mají být odstraněny alfa kanál na různé hodnoty. Bod, který má být odstraněn, bude mít hodnotu $\alpha = 0$ a bod, který má zůstat bude mít hodnotu $\alpha = 1$. U snímku, který má být vložen jako pozadí to bude obdobné. Body, které vložíme, mají hodnotu $\alpha = 1$ a ostatní $\alpha = 0$. Tímto se vytvoří dvě masky, kde bude mít každý obrazový bod pomocí alfa parametru určeno, jestli bude zobrazen nebo ne. Po sloučení těchto snímků dohromady dostanu výsledný obraz.

Na těchto obrázcích je ukázáno, jak dochází k sloučení dvou obrázků při různém nastavení hodnot alfa. Obrázek číslo 2 zobrazuje obrázek pozadí, číslo 3 je obrázek popředí, který má být vložený na pozadí. Poslední obrázek číslo 4 je výsledný po spojení obou předchozích.



Obr. 2. pozadí s plochami s různou hodnotou alfa



Obr. 3. popředí s plochami s různou hodnotou alfa



Obr. 4. obrázek po sloučení předchozích snímků

2.2 Typy pozadí

Zadní stěna může být realizována jako jednobarevná látka nebo přímo nátěr na stěně. Ve velkých televizních studiích se preferují převážně natřené stěny. Kvůli lepšímu a rovnoměrnějšímu osvětlení a snazší instalaci. Ve venkovních a malých studiích se zadní stěna většinou skládá z barevných látek nebo barevných papírů.

Vlastně pro klíčování můžeme použít, jakoukoli barvu a tu poté z celého snímku odstraníme. Problém je, že nejčastěji chceme natáčet postavy a tak musíme používat barvy, které jsou nejméně obsažené v lidské pleti, očích a vlasech.

Z tohoto důvodu se nejvíce používají dvě barvy a to modrá a zelená, jak již bylo zmíněno, protože obsahují nejméně pleťové barvy. V poslední době se více používá zelené

pozadí. Hlavní důvod více využívání zeleného pozadí je, že zelený kanál obsahuje největší hodnotu jasu ze všech tří kanálů, více než modrý a červený kanál. Zelený kanál v HD kameře je dvakrát více navzorkovaný a ve výsledku dá méně šumu a větší citlivost než zbylé dva kanály. Dnešní HD kamery pracují s formátem 4:2:2, to znamená, že zelené je tam stejně jako modré a červené dohromady

Lidské oči jsou nejcitlivější na zelenou barvu. Zelená barva odráží daleko více světla a není nezbytné, jí moc osvětlovat, tedy uspoříme elektrickou energii.

Někdy jsou situace, že se zelená barva nehodí. Důvodem je zmíněná vysoká světelná odrazivost zelené barvy. Pokud budeme mít postavu se světlými vlasy nebo ve světlém oblečení, tak se zelená barva někdy odráží na okrajích siluety. To vede k nepříjemným efektům, které jsou ve výsledném snímku velice rušivé, protože dojde nejen k odstranění pozadí, ale i okrajů objektu. V takovém případě je lepší použít modré plátno.

Ale velkou výhodou je že, velice málo lidí nosí na sobě zelenou barvu a tak pozadí nekoresponduje s oblečením ani s barvou kůže a očí. To u modrého plátna je nevýhoda, třeba při nošení džínových kalhot. Červené pozadí se využívá při natáčení domů nebo neživých předmětů. Důležitou věcí je, aby se plátno nemačkalo a nevytvářelo na látce záhyby, těch se pak ve výsledku velice těžko zbavuje.

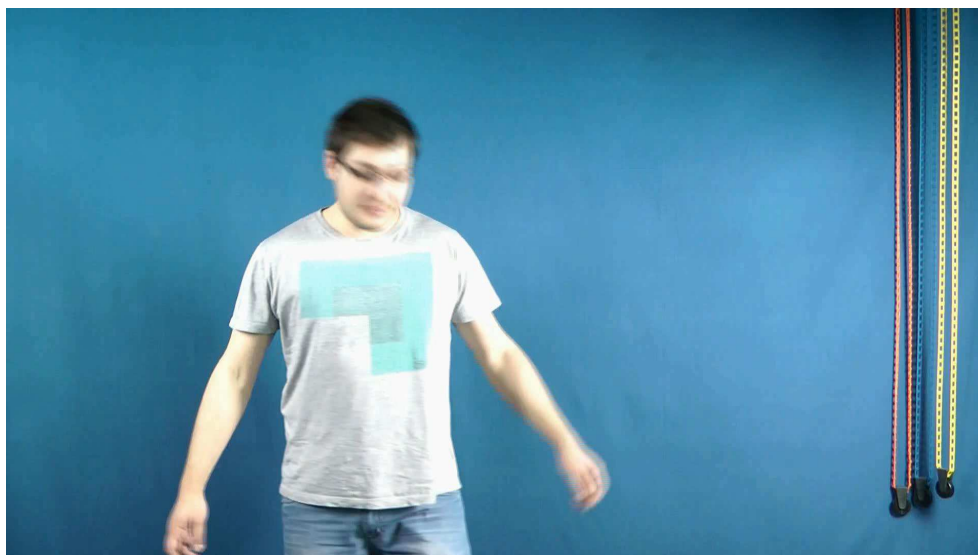
2.3 Oblečení

Základním požadavkem je, že by člověk na sobě neměl mít oblečení se stejnou barvou jakou má pozadí. Důvod, jak byl popsán, je, že se odstraňuje z obrazu jedna barva a oblečení stejné barvy, jakou má klíčovací plocha, by bylo také vyjmuto. Proto se nejčastěji využívá zelené plátno. Tedy zelená barva není v oblečení tak běžná. Kdyby měl hlasatel při modrém plátně na sobě například modré sako, tak by ho to odstranilo také a na jeho místo by se vložila třeba mapa počasí.

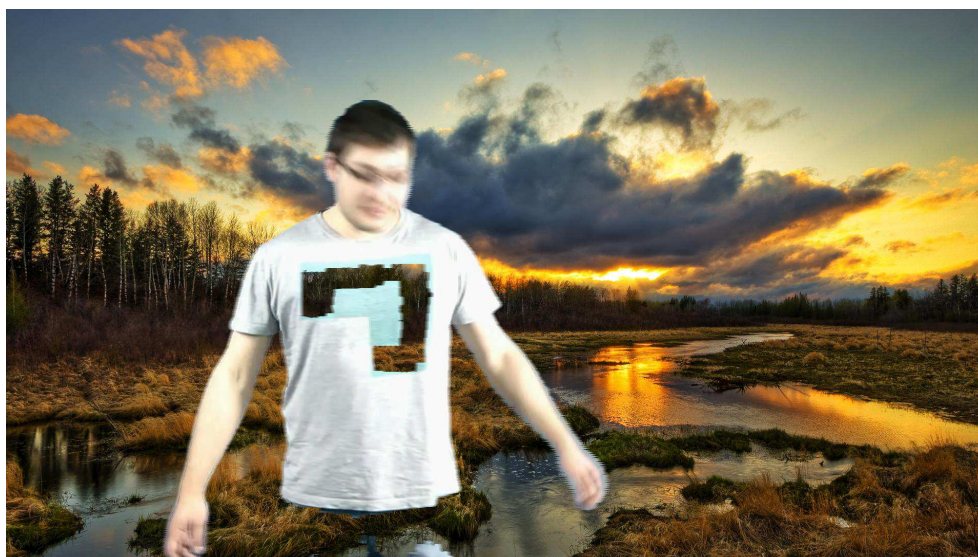
Když se tak stane tato část těla je ve filmu neviditelná, tedy nahrazena barvou pozadí. Tohoto efektu je někdy využíváno ve filmech.

Jedním z takových filmů je Harry Potter a kámen mudrců. Když Harry dostane od Brumbála neviditelný plášť, který používá. Pokud si Harry oblékal plášť, byl to vlastně kus látky, která měla zelenou barvu stejně jako pozadí, na kterém byl díl natáčen. Po oblečení zeleného pláště, byla zelená barva z celého snímku odstraněna a dáno místo ní jiné pozadí. A divák viděl zajímavý efekt neviditelného pláště.

Na obrázcích 5 a 6 je vidět co se stane, když je část oblečení stejná nebo podobná jako pozadí.



Obr. 5. klíčovací stěna, podobná barva oblečení jako pozadí



Obr. 6. co se stane po odstranění pozadí

2.4 Osvětlení

Nejdůležitější věcí je správně nasvítit celou scénu tak, aby byl co největší barevný rozdíl mezi plochou, která má být odstraněna a objektem v popředí. Velice důležitý je ostrý přechod na hranici barevné plátno a objekt. Tmavá místa na plátně nemusí být identifikována ještě jako barva pozadí a poté nebudou odstraněna a výsledný snímek bude znehodnocen. Správně osvětlit celou scénu je velice těžké a vyžaduje velké zkušenosti.

Pro správné osvětlení musí být vybrán dobrý materiál, který odráží světlo tak jak je požadováno. Materiál velice ovlivňuje výslednou kvalitu. Například lesklá plocha odráží hodně světla a její nasvícení bude velice komplikované.

Hodně nasvícená plocha budou vypadat světle a málo nasvícená naopak velice tmavě a to zvětší celkový barevný rozsah plochy. Proto je velice důležité nastavení správného osvětlení tak, aby na zadním plátně nebyly žádné stíny a aby pozadí mělo co nejvíce jednotou barvu. Na přímé osvětlení se používají takzvaná měkká světla, proto aby za hercem nevytvářely moc velké stíny. Další světla se používají k bočnímu nebo zadnímu osvětlení. Kdyby byla některá místa tmavší, tak se pomocí klíčování nemusí tyto místa odstranit a ve výsledném snímku zůstanou špatně zbarvená místa. Práci si můžeme zjednodušit tím, že stačí nasvítit jen herce a jeho bezprostřední okolí a zbytek scény odřízneme.

Někdy mohou být stíny na zadním plátně žádoucí a používají se pro speciální efekty. Oblasti se stínem na plátně se můžou nahradit tmavší verzí daného pozadí. Tím se docílí toho, že to vypadá jako by osoba vytvářela skutečný stín.

2.5 Použití

V dnešní době existuje spousta hardwaru i softwaru použitelného ke klíčování obrazu. Tato zařízení můžeme rozdělit na dvě skupiny. První skupina poskytuje výsledky současně s natáčením bez zdlouhavé post produkce, ta se používají ve virtuálním studiu nebo v televizních přenosech například ve zprávách. Druhou možností je upravovat scénu po natočení v postprodukčních studiích

Velkou výhodou je pro běžné uživatele spousta dostupných programů. Na trhu je hodně programů, pomocí kterých se může dosáhnout velice hezkých výsledků. Mezi profesionální nebo poloprofesionální softwarové programy pro ukázkou patří například Autodesk Smoke, Final Cut Pro, Pinnacle Studio, Adobe After Effects. V těchto programech je většinou hodně možností, jak nastavit jednotlivé vlastnosti pro klíčování. Zde jsou vypsány některé pro program Adobe After Effects.

"Color Difference Key effect" zde dojde k rozdělení obraz do dvou masek. Do masky A s barvou pozadí a B s ostatními barvami. U masek A a B se nastaví rozdílná hodnota alfa kanálu. Obě vrstvy se sloučí do třetího snímku, takže dostaneme vrstvu s průsvitným pozadím a neprůsvitným objektem.

"Color Key effect" tento efekt je nejjednodušší, zprůhlední se pouze body, které mají podobnou barvu jako je barva zadaná uživatelem.

"Color Range effect" u tohoto efektu dojde zprůhlednění barev z určitého rozsahu zadaného uživatelem. Je možné vybrat různé barevné prostory RGB, YUV nebo LAB.

"Difference Matte effect" tuto metodu můžeme použít, když máme dva snímky jeden statický a jeden pohybující se. Porovnávají se změny, co se mnění ve snímku zůstane, ostatní se zprůhlední.

"Extract effect" u tohoto efektu dojde k zprůhlednění bodů na základě rozdílného jasů. Možné použití je pokud natáčíme postavu na černém, bílém pozadí nebo pokud máme světlé, tmavé pozadí rozdílných barev.

"Linear Color Key effect" tento efekt porovnává jednotlivé body se zadanou barvou, když jsou barvy shodné, tak se pixel zprůhlední, čím menší je shoda, tím méně je pixel průhledný.

"Luma Key effect" zde jsou odstraněny pixely se specifickým jasnem.

U ostatních programů jsou volby podobné. Volně ke stažení je třeba program VSDC Free Video Editor. Ať už používáme hardwarové nebo softwarové zařízení výsledky budou stejně kvalitní.

2.6 Pre – klíčování

Je to proces, který slouží k sloučení dvou jednotlivých filmových pásek a vytvoření z nich pouze jednoho, který bude nahrán na finální filmovou pásku. Nejdříve je vyexponován první snímek a poté je přidán druhý. Vzhledem k tomu, že film nemůže být exponován dvakrát bez vytvoření dvojité expozice. První oblast musí být maskována, zatímco druhá je exponována a naopak. Každé maskování je děláno takzvaně za pomoci pohyblivé masky.

Stejně jako u digitálního klíčování je používáno jednobarevné pozadí nejčastěji modré nebo zelené. Na čočku kamery je umístěn filtr, který nepropustí pouze barvu pozadí a bere ho jako černé, takže na negativu bude oblast celá bílá, tato oblast bude průhledná. První expozice negativu se provede na vysoce kontrastní snímek, takže na negativu bude jen bílá a černá oblast. Pozadí bude neprůhledné a objekt v popředí bude průhledný. Další snímek bude opět vysoce kontrastní akorát s převracenými oblastmi. Snímek bude inverzní k druhému. Poté se vytvoří tři vrstvy. Ve spodu bude neexponovaný film, další bude maskovací snímek. Maskovací snímek zakryje vše kromě objektů v popředí a jen to bude exponováno. Na vrchu bude původní nahrávka. Toto se zopakuje ještě jednou akorát s inverzní maskou a místo obrázku postav zde bude obrázek pozadí. Výsledek bude kombinace těchto dvou snímků.

2.7 Výhody digitálního klíčování

V dnešní době už se výhradně používá digitální klíčování oproti dřívějšímu analogovému klíčování, které je popsáno výše. Ve starém systému jsme museli mít pět filmových pásek. Pásku s popředím, pozadím negativ a pozitiv a jednu pásku na kopírování. To vedlo k různým nežádoucím výsledkům. Oproti tomu je digitální klíčování daleko výhodnější. Také při kopírování filmů z pásky na pásku, při výrobě finálního snímku docházelo ke ztrátě kvality na rozdíl od digitálního kopírování, kde ke ztrátě nedochází.

U digitálního klíčování jde skládat více vrstev na sebe, aniž by došlo ke snížení kvality záznamu. Každou vrstvu můžeme natočit samostatně a pak je složit dohromady. S dřívějšími metodami klíčování se vše muselo znovu exponovat, spojovat a docházelo ke vzniku různých artefaktů a snižování kvality celého záznamu.

3 Návrh metodiky

Pro klíčování je hlavní nalézt odpovídající barvu ve snímku, která má být nahrazena pozadím. S největší pravděpodobností to bude barva, která má největší procentuální zastoupení v celém obraze. Jedno jaká barva to bude. K identifikaci barvy, se používají různé metody.

V první stačí určit hledanou barvu, procházet snímek pixel po pixelu a vyhledávat barvu podle přesně zadaných podmínek. Pokud je nalezen pixel, který splňuje tyto podmínky, můžu ho nahradit pixelem z jiného snímku na odpovídající pozici.

Další z možností je rozdělit snímek na dvě oblasti podle barev, tak aby tyto skupiny byli přibližně stejně velké. Rozdělím snímek na barvu, která má být nahrazena a na ostatní barvy, které ve snímku zůstanou.

V obou dvou principech je důležité, aby pozadí mělo podobnou barevnost, nebyly zde hodně tmavá nebo příliš světlá místa nebo dokonce úplně jiné barvy. V této části používám dva různé barevné prostory RGB a HSV proto je zde trochu více popíši.

3.1 RGB

RGB model patří nejspíše mezi nejnámější barevné prostory vůbec. Historie tohoto modelu má základy v teorii Thomase Younga a Herman von Helmholtze a jejich trojbarevném vidění a Maxvellově barevném trojúhelníku. Každou existující barvu můžeme získat kombinací tří jiných základních barev. Před použitím RGB modelu v moderní technice měl využití právě v popisu teorie lidského vidění.

K popsání barev se využívá kolorimetrické soustavy. K získání údaje o barvě jsou minimálně zapotřebí tři základní souřadnice, proto jsou také někdy jmenovány jako trichromatické souřadnice. Každou barvu je možné nahradit aditivním mícháním tří základních barev. Tyto barvy tvoří vrcholy trichromatického trojúhelníka, kde souřadnice R, G a B leží na vrcholech trojúhelníku, který má výšku jedna. Těmto vrcholům přísluší vlnové délky o velikosti $\lambda_R = 700\text{nm}$, $\lambda_G = 546,1\text{nm}$ a $\lambda_B = 435,8\text{nm}$. Nevýhodou je, že některé barvy leží mimo tento trojúhelník a tak je nemůžeme získat. Proto se někdy zavádí soustava X, Y a Z, která tento problém řeší.

RGB je barevný model, který je založený na aditivním míchání tří základních barev červené, zelené a modré. Při aditivním míchání barev se jednotlivé složky sčítají a vytvářejí odpovídající barvu. Je to podobné jako rozsvěcování zelené, modré a červené

žárovky. Když žádná nesvítí, vnímáme tmu, tedy černou barvu, když naopak svítí všechny, vidíme světlo bílé. Tento model nepotřebuje žádné přídavné světlo. Tedy pomocí množství světla, které vyzaří jednotlivý činitel RGB, lidský mozek vytvoří výsledné barevné světlo.

Jako rozšíření RGB modelu se zavedl model RGBA. RGBA nemá odlišný barevný prostor, písmeno A reprezentuje alfa kanál, který udává průhlednost jednotlivých pixelů. RGB prostor může mít podle různých standardů trochu odlišně definované vlastnosti, všechny ale pracují s prostorem RGB. Nejznámější jsou Adobe RGB, Apple RGB a sRGB

V počítači je tento model reprezentován různými způsoby. Mají, ale stejný význam a udávají jak moc je daná barva v pixelu obrazu obsažena.

Zobrazení v číslech od 0 do 1 se většinou používá v teoretické analýze, kde se pracuje s desetinou čárkou. Například hodnota (1, 0, 0) udává plně červenou barvu.

Další zobrazením je procentuálním vyjádřením od 0% do 100%. Maximální intenzitu červené představuje toto zadání (100%, 0%, 0%).

Nejčastější reprezentace v počítači a v grafických programech je pomocí hodnot od 0 do 255, což představuje 8 bitové číslo, 2^8 pozic. Červenou barvu představuje hodnota (255, 0, 0). Bitová hloubka může být různá, existují například 4, 8, 16, 36 nebo i 48 mód.

3.2 Vlastnosti RGB

Model RGB můžeme také zobrazit jako krychli, kde všechny zobrazitelné barvy jsou rozloženy v celém objemu. Je to vizualizace RGB modelu. Na osách x, y a z leží hodnoty základních barev modré, červené a zelené, čím dále jdeme od počátku souřadnicové soustavy, tím větší hodnotu jednotlivých barev získáváme. V počátku souřadnic je barva černá, modrá, červená a zelená barva má hodnotu nula. Na konci je barva bílá, všechny barvy mají maximální intenzitu.

Výhodou RGB modelu je, že v tomto režimu pracuje většina zařízení, takže se po úpravách výsledky musí převést zpět do RGB. RGB se používá při konstrukci některých zařízení například monitorů, skenerů nebo fotoaparátů.

RGB model se používá pro snímače kamer, fotoaparátů, kde získáváme z intenzity světla tři barevné kanály. Jelikož jsou zde všechny tři souřadnice závislé na jasu, tedy na množství dopadajícího světla, tak tento model není moc intuitivní pro práci. Když například chceme snímku ubrat pouze jas, snímek ztmavit, musíme pracovat se všemi souřadnicemi naráz. To vedlo k zavedení jiných barevných prostorů. Kde jsou souřadnice barvy a jasu od sebe oddělené a na sobě nezávislé. Jedním z takových prostorů je HSV.

3.3 Vztah mezi RGB a HSV

Má povahu algoritmu kde r , g , b jsou barevné složky, hodnota \max je největší hodnota a \min je nejmenší hodnota. Odstín nabývá hodnot $h \in (0,360]$ a saturace a světlost jsou hodnoty $s, l \in [0,1]$.

$$h = \left\{ \begin{array}{l} \text{nedefinován.....jestliže } \max = \min \\ 60^\circ \times \frac{g-b}{\max-\min} + 0^\circ, \text{.....jestliže } \max = r \text{ a } g \geq b \\ 60^\circ \times \frac{g-b}{\max-\min} + 360^\circ, \text{.....jestliže } \max = r \text{ a } g < b \\ 60^\circ \times \frac{b-r}{\max-\min} + 120^\circ, \text{.....jestliže } \max = g \\ 60^\circ \times \frac{r-g}{\max-\min} + 240^\circ, \text{.....jestliže } \max = b \end{array} \right.$$

$$l = \frac{1}{2}(\max + \min)$$

$$s = \left\{ \begin{array}{l} 0, \text{.....jestliže } l = 0 \text{ nebo } \max = \min \\ \frac{\max-\min}{\max+\min} = \frac{\max-\min}{2l}, \text{.....jestliže } 0 < l \leq \frac{1}{2} \\ \frac{\max-\min}{2-\max+\min} = \frac{\max-\min}{2-2l}, \text{.....jestliže } l > \frac{1}{2} \end{array} \right.$$

3.4 HSV

Nejdříve bych zde popsal barevný prostor HVC, který s prostorem HSV úzce souvisí. HVC prostor vytvořil Albert Henry Munsell (1858–1918). Albert Henry Munsell byl americký malíř, který se přímo nezabýval barevným vnímáním a fyzikální podstatou vzniku barev, ale o jednotlivé vztahy mezi nimi. Vytvořil ucelený systém barev a zavedl pojmy jako je doplňková barva, vyvážení barev a barevná kombinace.

Munsell v roce 1917 vyvinul systém zvaný "Munsell Color System", kdy sloučil každou barvu pomocí tří parametrů odstínu "hue", hodnoty "value", a barevnosti "chroma".

Je to tří rozměrný prostor, kde každý rozměr je nezávislý na ostatních. Na vertikální ose jsou umístěny neutrální barvy, úplně dole je černá a nahoře bílá, uprostřed jsou odstíny šedi. Vzdálenost od vertikální osy udává barevnou sytost, Čím dále jsme od vertikální osy, tím je sytost větší. Odstín barvy udává úhel okolo vertikální osy.

V současné době se používají různé modifikace Munsellova barevného prostoru. Například prostor HSB je kombinace odstínu "hue", sytosti "saturation" a jasu "brightness". Nebo HSV ten je určen pomocí parametrů odstín "hue", sytost "saturation" a hodnota "value". Tento model vytvořil v roce 1978 Alvy Ray Smith a použil ho na jednom z prvních kreslicích programů SuperPaint.

Odstín udává tón barvy a může nabývat hodnot od 0° do 360° . Sytost představuje množství šedi obsažené v barvě a udává se v procentech od 0% do 100%. Hodnota je množství jasu, tedy množství bílého světla, které barva obsahuje. Tento model není možné změřit, protože jsme schopni měřit pouze množství dopadajícího světla, je odvozen z modelu RGB pomocí pevně definovaných matematických vztahů. To znamená, že HSV model má stejná barevná omezení jako RGB model a když budou RGB barvy vypadat špatně tak i HSV barvy.

3.5 Vlastnosti HSV

Tyto modely jsou zobrazovány pomocí kužele nebo kruhu. U kužele se na obvodu nastavuje barevný odstín pomocí stupňů. Vzdálenost středu do kraje představuje sytost a výška kužele hodnotu. U kruhu je uvnitř trojúhelník kde, jedna strana reprezentuje sytost a druhá hodnotu, na obvodu kruhu je odstín ve stupních.

Jednou z výhod HSV prostoru je, že pro člověka, který často nepracuje s barevnými modely v grafických programech, představuje lepší a intuitivnější model pro práci s barvami než model RGB. Barevné modely, ať už HSV nebo jeho modifikace, jsou často používány v grafických programech, protože umožňují snazší nastavování barev než model RGB, kde se musí nastavovat množství barvy ve všech kanálech současně.

Pokud chceme upravit hodně světlý nebo tmavý obraz, bude trvat dlouho, než správně nastavíme rozložení barev v RGB hodnotách, tak abychom se zbavili světlých nebo tmavých míst pro správné nastavení scény. Nikdy bychom se také nemuseli dobrat konce. Když pracujeme v HSV prostoru, stačí jen jeden ovládací prvek, pomocí kterého můžeme snadno opravit špatně osvětlenou scénu.

Velké využití má tento nebo příbuzné modely například v počítačovém vidění, analýze obrazu, rozdělení obrazu na části podle podobných vlastností. Důvodem je, že máme oddělené jednotlivé složky a jsou na sobě nezávislé. Můžeme získat některé informace z obrazu na základě stejných barevných dat.

Další aplikací zahrnujících detekci objektů na základě podobných vlastností je rozpoznání například, tváří, textu, registračních značek aut nebo analýza lékařských snímků

Jednou z hlavních částí používané k analýze barevných obrázků je přímé rozšíření algoritmů vyvinutých nejprve pro černobílé obrázky. Jsou to třeba algoritmy "k-means", "fuzzy clustering " nebo "canny edge detection".

V principu můžeme barevné složky oddělit od ostatních stejným algoritmem. Je proto důležité, aby tento algoritmus dokázal pracovat jen s barvou a dokázal rozlišit jednotlivé barvy jen v jednom kanálu. Podobné barvy leží v prostoru vedle sebe. Proto je třeba nevhodné používat R, G a B složky, protože ty jsou vypočítávané z množství světla, které dopadne na povrch senzoru. Může se stát, že třeba tmavě modrá leží úplně na jiném místě než světle modrá, protože každá z těchto barev může obsahovat úplně jiné množství světla. Přestože jsou obě dvě barvy modré. Proto se stává výpočet pomocí těchto koeficientů nevhodný.

Další z jednoduchých důvodů pro použití HSV je, že simuluje, jak naše oko zpracovává barevný vjem a jak chápeme vlastnosti barev. Často určujeme barvu podle toho, jak hodně jsou objekty tmavé nebo světlé, jakou sytost a jas mají.

Když pracujeme se scénou pořízenou pomocí RGB souřadnic, tak můžeme přejít do HSV souřadnic a tam nastavit parametry díky jeho intuitivnosti, tak jak se nám líbí a jaké požadavky na výsledek máme a poté zpět přejít do RGB prostoru.

3.6 LAB

Barevný prostor LAB je odvozen z prostoru CIE XYZ a je jistou obdobou barevného prostoru HSV. Je zde také oddělena jasová složka od barevných. Jasová složka se zobrazuje na ose L a může nabývat hodnot od 0 do 100, kde 0 neprezentuje nejnižší a 100 nejvyšší jas. Osa a představuje barvu zelenou a červenou a osa b modrou a žlutou.

Tento barevný prostor lze snadno použít pro úpravu fotografií, jelikož zde také můžeme upravovat jednotlivé barvy, aniž bychom zasahovali do jasové složky. Můžeme ho tedy s výhodou využít i pro klíčování, protože podobné barvy leží blízko sebe a snadno jdou oddělit

Prostor LAB je navržen, aby simulovat lidské vidění. Jednotlivé barevné složky a množství jasů. Je to nezávislý barevný prostor a jeho zobrazení nezávisí na konkrétním zobrazovacím zařízení.

4 Algoritmy

4.1 Shluková analýza

Na základě podobných vlastností vstupních dat, můžeme tyto data roztřídit do jednotlivých skupin takzvaných shluků. V literatuře, která se zabývá statistikou, existuje spousta možností a přístupů jak tímto způsobem roztřídit data. Jenom část z nich je založena na statistických principech a využívá statistickou interpretaci dat.

V této analýze je důležité, aby data se stejnými nebo podobnými vlastnostmi měly co nejmenší vzdálenost od sebe a naopak jednotlivé shluky měly rozptýl od sebe co největší.

Objekty v stejném shluku by měly mít co nejvíce podobné vlastnosti a objekty v ostatních shlucích zase co nejvíce rozdílné vlastnosti. Každý objekt můžeme popsat skupinou proměnných. Celkově se shlukování může popsat tím, že máme počet objektů o počtu n , které roztřídíme do m shluků. Teoreticky se počet objektů může rovnat počtu shluků, ale ve výsledku to nemá smysl, protože u této metody chceme většinou dosáhnout komprese dat.

Pro určení podobnosti se většinou používá míra vzdálenosti jednotlivých objektů, třeba eukleidovská, Hammingova, Čebyševova nebo Mahalanobisova. Podobnost se pak může určovat třeba jako minimální, maximální vzdálenost, proměnná vzdálenost atd.

Celkový počet shluků může vycházet jak z teoretických tak i praktických požadavků, například jak moc chceme celou skupinu zredukovat.

Shluková analýza je spíše způsob řešení než přímo konkrétní algoritmus. Jednotlivé metody popisují co to shluk je a jak takový shluk nalézt. Podle daného algoritmu se objekty vypočítají a rozdělí do jednotlivých shluků. Výsledek nemusí být vždy stejný, záleží na námi zadaných požadavcích. Shluková analýza je obecný název pro spousta různých metod. Ty nejběžnější jsou metoda nejbližšího souseda, metoda nejvzdálenějšího souseda, metoda těžiště, párová vzdálenost nebo Wardova metoda. Tyto metody mají široké využití v praxi například při data dolování nebo při statistickém zpracování dat.

Metoda těžiště využívá vzdálenost jednotlivých těžišť. "Centroid" nebo také těžiště se vypočítává jako průměr souřadnic všech objektů v daném shluku. Objekty, které mají nejmenší vzdálenost od těžiště, se sloučí. Metoda těžiště se využívá pro třídící problém nazvaný k-means clustering.

4.2 K-means clustering

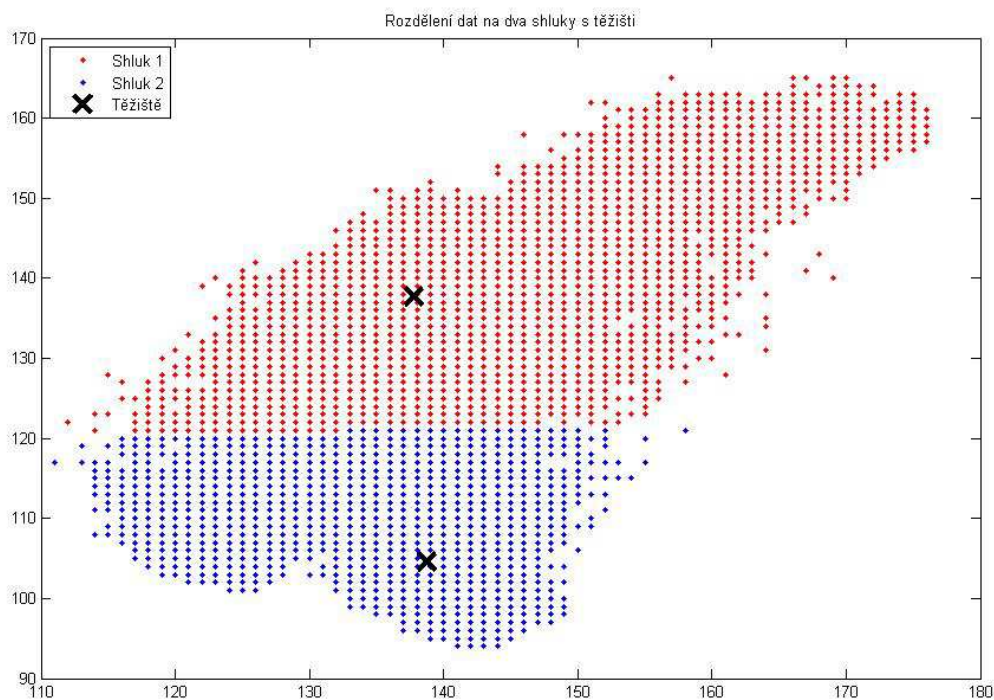
Jeden z problémů pomocí kterého můžeme třídit jednotlivé informace založený na shlukování, se nazývá k-means clustering. Je to takzvaná vektorová kvantizace, která byla původně vyvinuta pro teorii zpracování signálů. Poprvé byla uvedena Johnem MacQueenem v roce 1967, ačkoli s tímto nápadem přišel Hugo Steinhaus již roku 1957. Základní algoritmus použitý pro řešení k-means clustering původně navrhl v roce 1957 Stuart Lloyd pracující v Bellových laboratořích jako výpočet pro pulsně kódovou modulaci, ale do roku 1982 nebyla tato metoda publikována. V roce 1965, E. W. Forgy publikoval podobnou metodu, která je známá jako Lloyd-Forgy. Efektivnější metoda byla navržena a publikována v Fortranu Hartiganem a Wongem v letech 1975 až 1979.

Pomocí tohoto problému rozdělíme skupinu vstupních dat o určitém počtu do výstupních skupin nebo také shluků, kde každý vstupní objekt patří do nejbližšího možného shluku tak, aby od jeho středu měl menší vzdálenost než od středů ostatních shluků. Nejmenší vzdálenost většinou představuje to, že body ležící u sebe mají podobné vlastnosti. Je tedy velice výhodné je sloučit dohromady. Výsledné rozdělení se podobá takzvanému Voroného diagramu.

Objekty (x_1, x_2, \dots, x_n) , kde každý objekt je reálný vektor, který může mít určitý počet dimenzí. Tyto vektory jsou rozděleny do shluků, pro který platí, že (počet shluků \leq počet objektů) a nastaví se body těžiště shluků $S = \{S_1, S_2, \dots, S_k\}$, tak aby vzdálenost jednotlivých objektů byla, co nejmenší od středů S kam byly přiřazeny.

V K-means clustering jsou jednotlivé středy shluků reprezentovány vektory, které se nutně nemusí shodovat s vektory vstupních dat.

Na obrázku číslo 7 je vidět rozdělení obrazových bodů na dva shluky, křížky označují těžiště shluků. Body by měly být rozděleny tak, aby měly nejbližší právo ke svému těžišti, u kterého leží.



Obr. 7. ukázka rozdělení dat na dva shluky s těžišti

Algoritmus použitý pro k-means clustering má polynomiální výpočetní složitost a výsledek je vždy pouze přibližné řešení. Při opakování běhu algoritmu se může dojít pokaždé k trochu odlišným řešením.

Pro mechanismus výpočtu se používá Lloydův algoritmus, který se také někdy nazývá k-means algoritmus. Pomocí něho najdeme pouze jedno z možných řešení a často se spouští vícekrát s náhodnými iniciacemi.

Na k-means clustering existují různé obměny například omezení těžiště souboru dat "k-medoids", výběr mediánu "k-medians", méně náhodný výběr počátečních podmínek "k-means++" nebo počáteční přiřazení každému vstupnímu objektu jeden shluk "fuzzy k-means".

Jednou z nevýhod tohoto algoritmu je, že počet výsledných shluků se musí zadávat předem. Dále se spíše vytvářejí shluky podobných velikostí, protože se přiřazuje objekt nejbližší těžišti. To někdy vede k problémům se špatným rozpoznáváním hranic mezi shluky.

4.3 Použití

Tento algoritmus používá dvě fáze iterativního algoritmu (*To je algoritmus, který se v volání programu snaží nalézt řešení výsledku, tím, že se program opakovaně volá a řešení se čím dál tím více zpřesňuje. Začíná se prvním náhodným odhadem. Výsledkem*

opakovaných výpočtů jsou posloupnosti, které se přibližují správnému výsledku.) k nalezení minimální vzdálenosti od jednotlivých bodů k centru jednotlivých shluků a přiřazení bodů do takovýchto shluků.

První fáze používá to, co se v literatuře často nazývá jako "batch", kde průběh sestává z přiřazení každého bodu k nejbližšímu centru těžiště. Tato fáze je velice rychlá, ale poskytuje pouze přibližné řešení jako startovní bod pro druhou fázi.

Druhá fáze využívá toho, co se v literatuře často jmenuje jako "on-line" vylepšení, kde jsou vypočítána jednotlivá těžiště shluků a body jsou individuálně znovu přerozděleny tak, aby se minimalizovala velikost vzdáleností od bodů k těžišti a jednotlivá těžiště jsou znovu po každém přiřazení bodu do shluku přepočítána. Každé volání v druhé fázi sestává z průchodu všemi body.

K-means clustering se často jen přibližuje k lokálnímu optimu, které je v tomto případě přerozdělení bodů, kde se každým jednotlivým bodem pohybuje zvlášť a snižuje se tak součet vzdáleností v každém shluku. Počáteční stav může být řešen náhodným nebo předem určeným výběrem výchozích bodů.

K-means clustering představuje heuristický přístup k problému, snaží se najít přibližné nejlepší řešení daného problému. K tomu se využívá Lloydův algoritmus, který je snadné aplikovat a je možné pracovat i s velkým množstvím dat. Takovéto řešení je velice úspěšné a praktikuje se v hodně odvětvích například v programech segmentace trhu, počítačové vidění, geostatistice, astronomii.

Je také často používán k počátečnímu rozřídění dat, jako předstupeň pro jiné algoritmy, k nalezení počáteční konfigurace.

K-means clustering se stále hodně využívá v zpracování signálů, pro které byl původně navrhnut. Velké využití našel tento algoritmus v počítačové grafice ke kvantizaci barev a tedy k zjednodušení celkového počtu barev na množství dané počtem clusterů. K-means se často právě pro toto používá a dosahuje velice dobrých výsledků.

Další použití tento algoritmus našel v shlukové analýze, kde se množství vstupních dat rozdělí na určitý počet výstupních dat. Nicméně původní k-means algoritmus není moc flexibilní pro tento typ úloh, hlavně proto, že parametr počtu shluků musíme znát předem a je velice těžké ho správně určit. Dalším omezením tohoto algoritmu je, že nemůže být použit pro nečíselné funkce nebo funkce vzdáleností.

Dva hlavní rysy tohoto algoritmu, které jsou někdy velice účinné, se také někdy označují za jeho největší nedostatky. Počet shluků je vstupní parametr a jeho nevhodný výběr může vést ke špatným výsledkům. Proto při průběhu k-means algoritmu, je někdy

nezbytné sem tam provádět diagnostické kontroly pro určení počtu shluků v daném souboru. Přiblížení se k lokálnímu minimu, může někdy vytvářet špatné nebo neintuitivní výsledky.

4.4 Lloydův algoritmus

Lloydův algoritmus, je pojmenovaný po Stuartovi P. Lloydovi. Tento algoritmus hledá rovnoměrné rozdělení bodů v euklidovském prostoru. Tyto části bodů vkládá do tvarovaných a stejně velkých buněk.

Je podobný k-means clusteringu, který opakovaně hledá těžiště jednotlivých shluků a rozděluje výstup podle toho který bod je k jakému nejbližší. Lloydův algoritmus zobrazuje výsledek ve formě Voroného diagramu.

Tento algoritmus nejčastěji pracuje s euklidovským prostorem, ale jemu podobné algoritmy mohou pracovat i s vícedimenzionálními prostory nebo i s jinými než euklidovskými prostory.

Nejběžnější druhy iniciace, zadání počátečních podmínek, které se provádějí, jsou "forgy" a "random partition". U "forgy" metody se náhodně určí nějaké vstupní hodnoty a ty jsou vybrány jako počáteční středy shluků. "Random partition" metoda nejdříve náhodně přiřadí každé vstupní hodnotě jeden cluster a pak pokračuje aktualizací krok, kdy se vypočítá těžiště takto náhodně přiřazených bodů.

"Random partition" metoda je výhodnější pro algoritmy jako harmonic k-means a fuzzy k-means zatímco metoda "forgy" se spíše používá pro k-means clustering.

Při výpočtu se opakovaně provádějí dva kroky, První krok je přiřazení bodů do jednotlivých buněk, jehož vzdálenost na druhou je nejmenší od daného centra. Druhý krok, který se nazývá aktualizace, se vypočítá nové těžiště jednotlivých buněk a přerozdělí se znovu body tak, aby byla vzdálenost na druhou jednotlivých bodů opět nejmenší.

4.5 Voroného diagram

Tyto diagramy jsou pojmenovány po ukrajinském matematikovi Georgij Feodosjevičovi Voronejovi. Voroného diagramy, také se někdy můžeme setkat s názvem Voroného teselace, Dirichletova teselace nebo Voronojova dekompozice.

Pomocí těchto diagramů můžeme přerozdělit euklidovský prostor do určitého množství výstupní ploch. Kdy každému vstupnímu bodu se přidělí taková rovina neboli plocha, aby jeho vzdálenost byla do středu roviny menší než k jakékoli jiné rovině v prostoru. Pro takovéto rozdělení můžeme použít například Fortunův algoritmus.

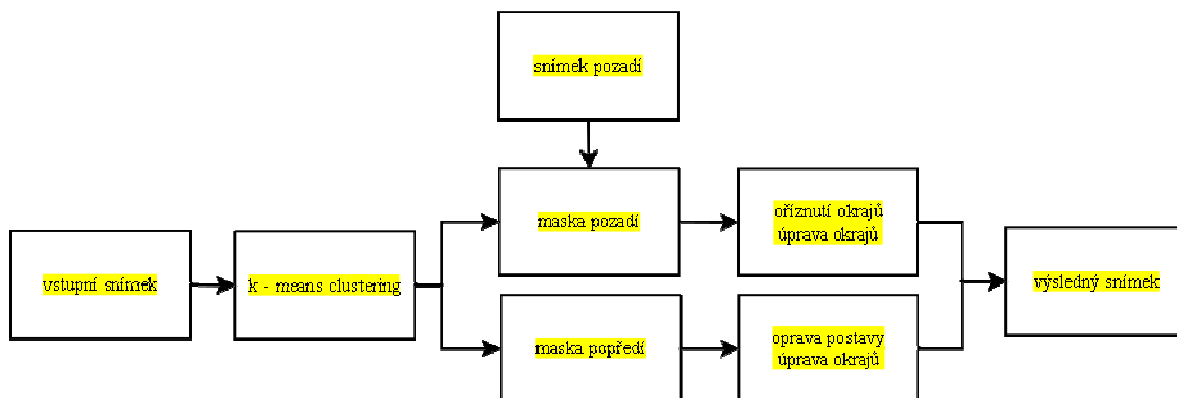
S těmito diagramy se také můžeme setkat v meteorologii nebo v hydrologii, kde se ale nazývají Thiessenovy polygony pojmenované podle amerického meteorologa Alfreda Thiessenova. A slouží k práci s prostorovými daty.

4.6 Binární alfa

Algoritmus k-means clustering má využití pro klíčování snímků, že se nastaví počet jednotlivých shluků na počet dva. Při použití prostoru HSV leží podobné barvy na místě vedle sebe. Tedy vzdálenosti barev s podobnými vlastnostmi budou menší než vzdálenosti barev nepodobných. Snímek se rozdělí na dvě podobně velké skupiny barev, které budou mít vlastnosti co nejvíce shodné.

Jedna skupina bude obsahovat barvu, kterou chceme odstranit, tedy barvu pozadí. Tuto barvu obsahuje přibližně polovina snímku a její souřadnice leží vedle sebe, protože mají totožné vlastnosti. Jeden shluk bude tvořen touto barvou. Druhý shluk bude tvořen zbylými barvami.

Pomocí k-means clustering dostaneme dvě barevné masky. Jednu můžeme odstranit a nahradit libovolným snímkem a druhou, představující popředí, které chceme nechat. Po sloučení obou snímků dostaneme výsledný snímek s vyklíčovaným pozadím. Na obrázku číslo 8 je vidět blokové schéma, které jsem používal pro klíčování.

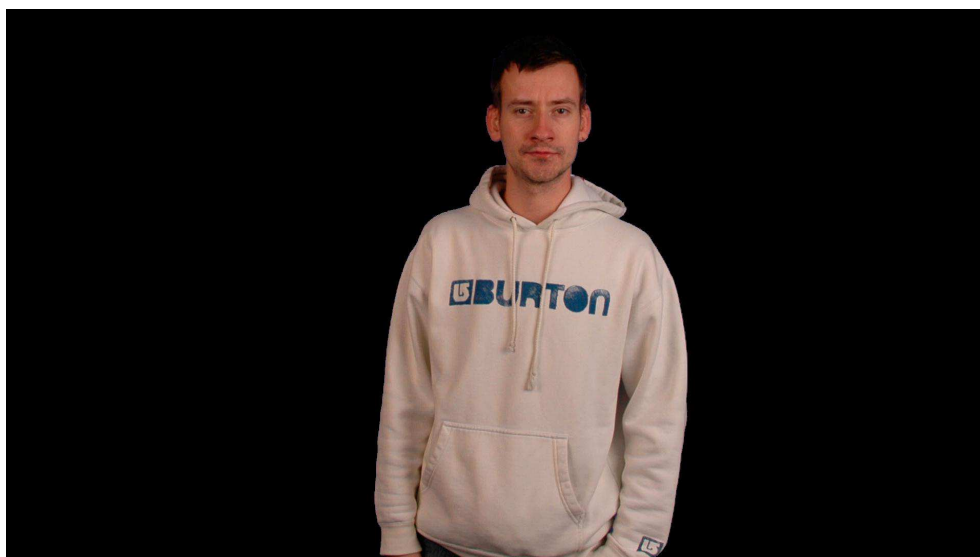


Obr. 8. blokové schéma použité pro klíčování

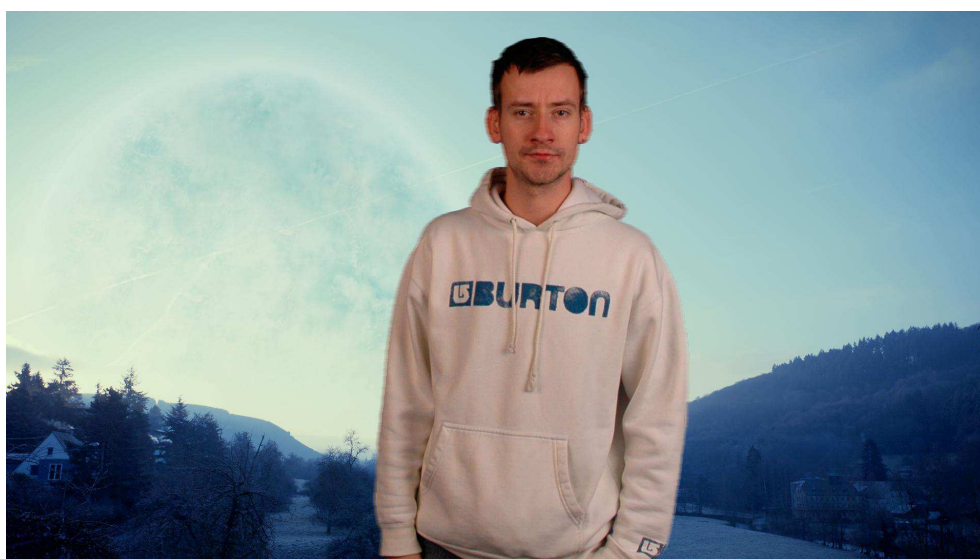
Na těchto třech obrázcích můžeme vidět dva clustery a výsledný snímek. Na snímku číslo 8 vidíme cluster pozadí s výřezem na postavu, na snímku 9 je cluster převrácený, nechaná postava a odstraněné okolí. Po sloučení těchto dvou snímků dostaneme výsledný 10 snímek, spojené popředí a pozadí.



Obr. 9. maska s pozadím



Obr. 10. maska s popředím



Obr. 11. sloučené snímky

Jedním z problémů, který vzniká u spojení dvou snímků do jednoho, je hranice mezi nimi. Hrany přechodu jsou velice ostré a nepřechází plynule. Výsledný snímek pak působí nerealisticky, jako by oba nepatřily k sobě.

Druhou věcí je, že na hranách klíčovaného objektu je odražené světlo od barevného plátna. To se zobrazuje ve formě barevného okraje na objektu a působí to velice rušivým dojmem. Tyto negativní skutečnosti je nutné odstranit, existuje na to hodně způsobů. Většinou se odstraňují při konečných úpravách takzvaném postprocessingu.

4.7 Postprocessing

Postprocessing můžeme volně přeložit jako finální úpravy. Je to úprava zdrojového snímku k dosažení lepších výsledků. Tyto úpravy probíhají těsně před tím, než dojde k promítání snímku divákům. Bez těchto úprav se neobejdou žádné velké filmy, ale i krátké snímky, které uživatel například nahrává na nějaký video server, se musí před poskytnutím ke shlédnutí ostatním uživatelům upravit.

Mezi nejznámější úpravy patří třeba změna jasu, zaostření, rozostření nebo korekce barev. K finálním úpravám snímků dochází ještě před promítáním. Postprocessing je možné provádět také v reálném čase, při přímém přehrávání snímku. Takovou nejběžnější úpravou v reálném čase, s kterou se každý z nás setkal, je zobrazení loga televize, přidání hodin nebo titulků k filmům a seriálům, která se provádí přímo při televizním vysílání.

Při klíčování videa tím, že dochází k vyjmutí části obrázku, se okraje tohoto vyjmutého objektu zdají okousané a působí velice rušivě. Každý asi zažil, když chtěl v grafickém programu z fotografie vyjmout nějakou část, třeba postavu nebo objekt a vložit ho na jiné pozadí, že na okrajích vznikly nerovnosti, které bylo potřeba zahladit.

Dalším velkým problémem je prosvítání barvy plátna na okraje osoby, ta vypadá jako by byla orámovaná. Toto jsou věci, které je potřeba zohlednit při finálních úpravách klíčovaného snímku a všechny tyto věci se musí opravit, před tím než se snímek promítne divákům.

Jednotlivými úpravami musíme zajistit to, aby výsledek byl co nejpřirozenější. Zlepšení přechodu mezi objekty lze dosáhnout šikovným nastavením parametru alfa a nebo rozmazáním hran přechodu.

Jedním ze způsobů je nastavení alfa masky tak, aby na okrajích docházelo k postupnému přechodu mezi pozadím a popředím. Kde má zůstat hlavní objekt popředí je hodnota alfa kanálu jedna, na ostatních pozicích je parametr alfa roven nule. U snímku

pozadí je to naopak. Na hranách můžeme nastavit parametr $\alpha = 0,5$, tak dojde k zprůsvitnění bodů a k prolnutí přechodu.

Nebo lze také postupně snižovat hodnotu alfa z jedničky například po hodnotách 0,1 do nuly a taktéž u obrázku popředí, jednotlivé hodnoty se budou postupně měnit, tím se docílí plynulého přechodu mezi popředím a pozadím a nedojde k rušivému jevu. Výsledný snímek bude vypadat velice přirozeně.

Další možností je aplikovat na hranu přechodu různá rozmazání jednotlivých bodů. Rozmazáním dojde k vyloučení ostrého přechodu ve snímku a smazání jednotlivých hran a nerovností na okrajích postavy, které vzniknou při použití k-means clustering, při vyjmutí objektu z jeho pozadí.

4.8 Rozmazání

U obrázku se rozmazání myslím tím, že každý pixel je smíchán s hodnotami jeho sousedních bodů, tedy že výsledný pixel je tvořena ze směsi okolních bodů, z obrázku odkud body bereme. Podle množství bodů z jaké vzdálenosti je bereme a jakou důležitost jim přiřadíme, se snímek může jevit více či méně rozmazaný.

Nejjednodušší způsob provedení je, že máme posuvné čtvercové okénko, kterým projíždíme po celém obraze. Pro každý pixel vezmeme hodnotu z tohoto posuvného okénka. Sečteme jednotlivé hodnoty bodů a podělíme je celkovým počtem pixelů v tomto okénku. Matice může vypadat například takto.

$$\begin{pmatrix} 1/9 & 1/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \end{pmatrix} \quad [7]$$

Vzorec pro průměrování a vytvoření nového pixelu z matice můžeme psát ve tvaru.

$$g = [x, y] \frac{1}{9} \sum_{r=-1}^1 \sum_{s=-1}^1 f[x+h, y+r] \quad [7]$$

Hledaný bod leží na pozici uprostřed matice, všechny body v okolí i s prostředním hledaným pixelem se podělí hodnotou 1/9 a vše se sečte dohromady a vytvoří se výsledný pixel. Operace, kde nějakou vstupní hodnotu násobíme přenosovou funkcí, se nazývá konvoluce. Násobením originálního snímku maticí s určitými hodnotami dostaneme výsledek. Když se toto provede na celý snímek, tak bude vypadat rozmazaný. Čím větší bude vyhledávací matice, tím rozmazanější snímek bude. Používají se i jiné konvoluční matice s jinými hodnotami.

Další možností je použití gaussovo průměrování. Body blíže ke středu mají větší váhu než body na okraji matice. Matice může vypadat například takto.

$$\begin{pmatrix} 1/16 & 1/8 & 1/16 \\ 1/8 & 1/4 & 1/8 \\ 1/16 & 1/8 & 1/16 \end{pmatrix} \quad [7]$$

Mezi další metody patří výpočet šumu pomocí mediánu, kde nový jas je vypočítán jako medián okolních bodů. Hodnoty v matici se seřadí vzestupně a nový jas je hodnota ležící uprostřed. Pro práci s mediánem se proto většinou používají matice, které mají lichý počet prvků, aby byl střed určen jednoznačně, například matice 3x3 nebo 5x5.

Takových to metod a různých matic pro výpočet šumu existuje velké množství, uvedl jsem jen tři nejběžnější a nejjednodušší.

4.9 Gaussian Mixture Models

Další možností opravy hran při klíčování je využití "Gaussian Mixture Models". Tento model pracuje na podobném principu jako k-means clustering. Je to model, který zobrazuje u jednotlivých prvků pravděpodobnosti, do jaké množiny patří. Není nutné přímo přiřazovat prvky do skupiny. Vrací pravděpodobnosti rozdělení prvků, podle podobnosti s jinými členy skupiny.

Gaussian mixture model je často využíván pro třídění dat. Nejdříve se nastaví počet clusterů, do kolika skupin chceme data rozdělit. Data jsou do clusterů přiřazovány na základě podobnosti, aby byla co největší pravděpodobnost, že do daného clusteru patří. Stejně jako k-means clustering i gaussian mixture model používá iterační algoritmus, který se spouští opakovaně a snaží se přiblížit co nejvíce k nejlepšímu možnému výsledku. Tento model může být někdy vhodnější než k-means clustering, protože jednotlivé shluky mohou mít různé velikosti a vztahy v nich. Clustering, který používá gaussian mixture model se někdy nazývá jako měkká metoda shlukování.

Největší rozdíl mezi k-means clustering a gaussian mixture model je ten, že zatímco u prvního zmiňovaného dojde k jednoznačnému přiřazení všech bodů do jednotlivých shluků, takže každý bod má jednoznačně dáno kam patří. U druhého modelu je každému bodu dána pravděpodobnost, do jakého shluku patří. Každý bod má určitou pravděpodobnost, že patří k danému shluku, čím dále je bod od centra shluku, tím menší má pravděpodobnost podobnosti. Body, které jsou nejpodobnější, leží nejbližší centru shluku, mají hodnotu 100%. Body, které nemají, žádnou podobnost, mají hodnotu 0%.

Vstupní parametry jsou stejné jako u k-means clustering, máme určitý počet vstupních parametrů a ty se rozdělí podle podobných vztahů do určitého námi zadaného počtu shluků. Akorát místo jednoznačného přiřazení bodů do shluků, bod buď patří do clusteru, nebo ne, se nám vrátí hodnoty pravděpodobnosti.

Této vlastnosti je možné využívat pro klíčování obrazu a opravu okrajů. Když bude obraz, ze kterého má být vyjmut nějaký objekt rozdělen do dvou clusterů, bodům které mají nejpodobnější barvu, která má v obrázku zůstat se nastaví parametr alfa na hodnotu jedné. Body s největší podobností s centrem druhého clusteru, budou mít hodnotu parametru alfa rovnou nule. Bodům, které nebudou mít 100% pravděpodobnost, že někam patří, se přiřadí odpovídající parametr alfa. Čím menší bude podobnost jednotlivých bodů se středy clusterů, tím menší bude hodnota alfy, nebo naopak větší. Tímto rozdělením dojde k automatickému nastavení parametru alfa v celém snímku a plynulému přechodu na hranici mezi snímky.

4.10 Vyhodnocení efektivity algoritmu

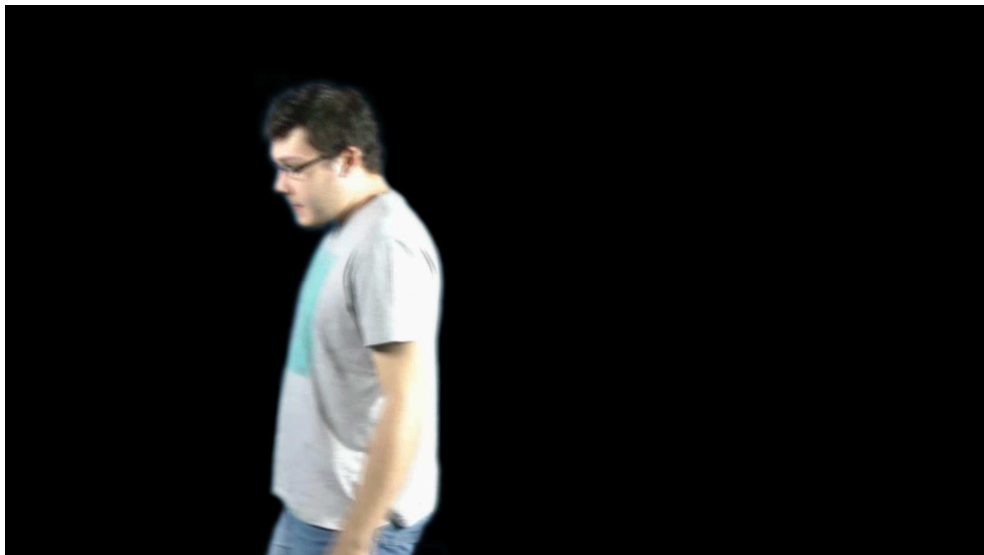
Pro vyhodnocení výsledků jsem pracoval s metodou popsanou na internetové stránce[10]. Vychází z toho, že máme původní snímek a z něho vyjmuté popředí, které se může považovat za referenční, bezchybné. Poté se na stejný snímek aplikuje klíčovací algoritmus a dostaneme druhé popředí. Oba snímky popředí se porovnají, referenční a získaný algoritmem. Rozdíl těchto dvou obrázků se považuje za chybu daného algoritmu.

Vlastní porovnání jsem prováděl na dvou videích. První sekvenci tvoří 8 a druhou 10 obrázků. V grafickém programu jsem z těchto 18 obrázků vyřízl popředí. Tyto snímky jsem měl za referenční. Z těch samých 18 snímků jsem vyklíčoval popředí i algoritmem. Poté jsem srovnával, jaký je rozdíl mezi referenčními a snímky klíčovanými algoritmem.

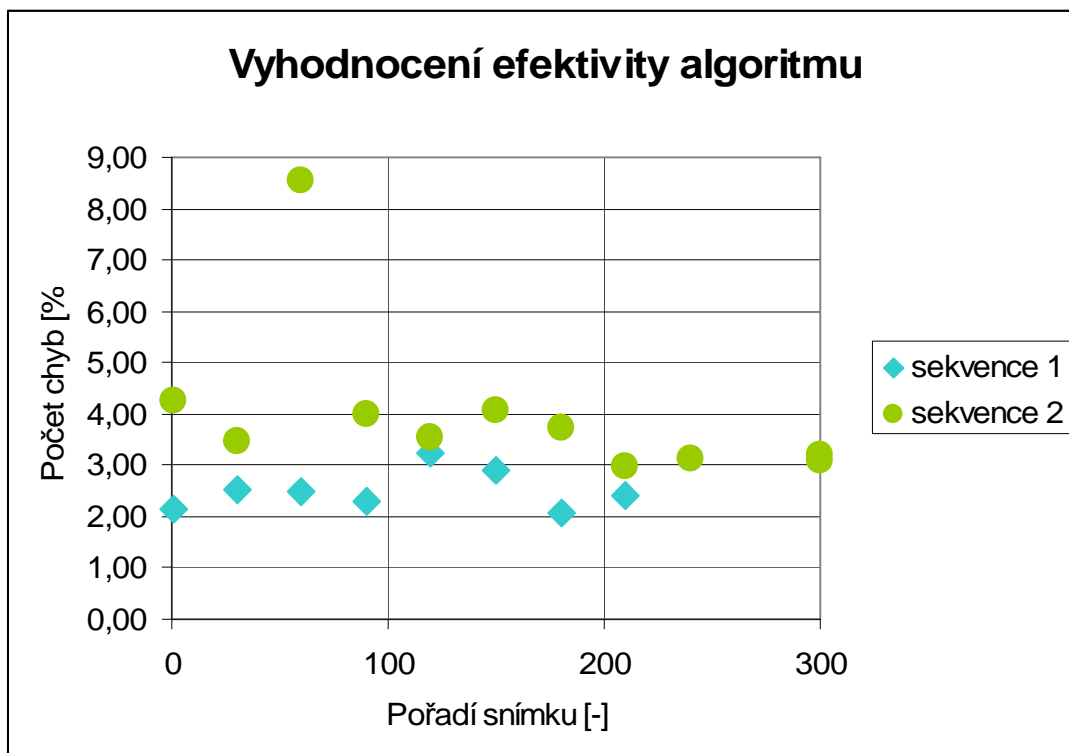
Na obrázku 12 je vidět popředí klíčované algoritmem a na obrázku číslo 13 je popředí vyjmuté pomocí grafického programu brané jako referenční.



Obr. 12. popředí klíčované algoritmem



Obr. 13. prostředí vyříznuté grafickým programem



Obr. 14. grafické znázornění počtu chyb na pořadí snímků

-	sekvence 1		sekvence 2		
	Pořadí snímku	Shoda pixelů	Chyby v %	Shoda pixelů	Chyby v %
1	2029303	2,14	1985243	4,26	1,83
30	2021410	2,52	2001823	3,46	1,03
60	2022297	2,47	1896685	8,53	6,10
90	2026175	2,29	1990878	3,99	1,56
120	2006389	3,24	2000482	3,53	1,09
150	2013849	2,88	1989273	4,07	1,63
180	2030969	2,06	1996166	3,73	1,30
210	2024004	2,39	2011556	2,99	0,56
240	-	-	2008811	3,12	0,69
300	-	-	2009704	3,08	0,65
300	-	-	2007354	3,19	0,76
medián	2023150	2,43	2000482	3,53	1,09
průměr	2021800	2,50	1990725	4,00	1,56

Celkový počet pixelů = 2073600

Tabulka výsledků je rozdělená na sekvenci videí 1 a 2. U první sekvence byly testované snímky skoro úplně shodné s referenčními. Většinou se zde nevyskytovaly žádné velké chyby, pozadí bylo odstraněno celé. Chyba oproti referenčním obrázkům je průměrně 2,50%. Tato chyba se dá, ale spíše považovat za odlišnost vzniklou na okrajích referenčního snímku při ukládání a ořezu v grafickém programu. Pokud se bude tato nepřesnost považovat za chybu vzniklou na okrajích, tak v tomto případě se blíží úspěšnost algoritmu k 100%.

U druhé video sekvence se chyba pohybovala v průměru okolo 4,00%. Pokud od této chyby odečteme počet chybných pixelů na okrajích, tak průměrná chyba bude 1,56%.

Tato odchylka je již dána nepřesností algoritmu v případě, že není pozadí dostatečně jednobarevné nebo osoba má stejné oblečení jako pozadí, popřípadě jiné chyby.

Úspěšnost tohoto algoritmu je velice závislá na kvalitě pořízené nahrávky. V případě kvalitní nahrávky se blíží ke 100%, V ostatních případech může být kvalita různá, stále se zhoršující až algoritmus úplně selže.

5 Závěr

V mé práci jsem se zabýval principy klíčování obrazu. Krátce zde popisuji, co je nutné dodržet, aby byl výsledek co nejlepší, jaké můžeme použít pozadí a osvětlení a co je to virtuální studio.

Pro klíčování obrazu je důležité určit barvu, která se má odstranit a tu nahradit svým vlastním pozadím. Pro detekci barvy pozadí využívám problém k-means clustering, který je založen na shlukové analýze. Vstupní obrazové body se rozdělí do dvou shluků. Jeden shluk je tvořen pixely pozadí a druhý shluk je z ostatních pixelů. Poté se body popředí nechají a pixely pozadí se nahradí jiným snímkem.

V prostředí matlab pracuji s metodou nazvanou kmeans. Metoda sloučí vstupní data do dvou podobně velkých skupin. V ideálním případě bude v jedné skupině jen barva pozadí a v druhé skupině budou barvy, které mají v obrazu zůstat. Dostanu dvě masky, se kterými poté pracuji. Dále zde mám napsané metody na vyříznutí části obrázku, tím se zbavím okrajů, které bývají často špatně osvětlené. Opravu okrajů, která vytváří plynulý přechod mezi vyjmutým objektem a vloženým pozadím a opravu postavy. Tato metoda opravuje některé chyby vzniklé uvnitř postavy.

Algoritmus je velice závislý na jednoduše pozadí. Druhou důležitou věcí je, aby osoba v popředí neměla věci ve stejné barvě jakou má pozadí.

Algoritmus jsem testoval na třech video sekvencích. U prvního snímku mělo pozadí dostatečně stejnou barvu, takže došlo k odstranění celého pozadí a nahrazení jiným. Dále zde byly vyhlazeny okraje mezi popředím a pozadím. V tomto případě měl algoritmus úspěšnost jdoucí ke 100%.

U druhého snímku mělo pozadí stejně homogenní barvu jako v prvním případě. Postava byla, ale oblečená do trička a kalhot podobné barvy jako pozadí, takže došlo k nahrazení části oděvu, jak bylo popsáno v úvodní kapitole. K opravě jsem používal dvě metody, opravu postavy a opravu středu. Částečně byl tento problém odstraněn, ale někdy při otáčení nebo nahýbání jsou některé části průsvitné. Zde měl algoritmus úspěšnost okolo 90%.

V třetím snímku vzhledem k nedostatečnému osvětlení zadní plochy došlo k odstranění různých částí obrázku, protože tento algoritmus sice rozdělí obraz na dvě masky, ale nedokáže identifikovat správné barvy. V tomto případě algoritmus selhává.

6 Literatura

- [1] Foster J., The Green Screen Handbook: Real-World Production Techniques, Sybex, 2010, ISBN 978-0470521076
- [2] Gonzales R.C., Woods R.V., Digital Image Processing Using MATLAB, Gatesmark Publishing, 2009, ISBN 978-0982085400
- [3] Kubíček, Ondřej. Analýza Metody pro detekci obrazového pozadí. Brno, 2007. Bakalářská práce. Masarykova univerzita v Brně. Fakulta informatiky. Vedoucí práce doc. Ing. Jiří Sochor, CSc..
- [4] Petr Klímek. E+M Ekonomie a Management. <http://www.ekonomie-management.cz>. [online]. 7.4.2008 [cit. 2015-05-18]. Dostupné z: http://www.ekonomie-management.cz/download/1331826675_2e7a/11_klimek.pdf
- [5] MathWorks. <http://www.mathworks.com>. [online]. 1994-2015 [cit. 2015-05-18]. Dostupné z: <http://www.mathworks.com/help/stats/gaussian-mixture-models.html>
- [6] MathWorks. <http://www.mathworks.com>. [online]. 1994-2015 [cit. 2015-05-18]. Dostupné z: <http://www.mathworks.com/help/stats/kmeans.html>
- [7] MathWorks. <http://www.mathworks.com>. [online]. 1994-2015 [cit. 2015-05-18]. Dostupné z: <http://www.mathworks.com/help/images/examples/color-based-segmentation-using-k-means-clustering.html>
- [8] JH LABS Java Stuff. <http://www.jhlab.com>. [online]. [cit. 2015-05-18]. Dostupné z: <http://www.jhlab.com/ip/blurring.html>
- [9] Frankies Carnival Time Inc.. green-screen-8.gif. [obrázek]. <http://www.frankiescarnival.com>. [online]. 2012 [cit. 2015-05-18]. Dostupné z: <http://www.frankiescarnival.com/greenscreen.html>. Formát: 589 × 363.
- [10] VideoMatting. videomattng.com. [online]. 2014 [cit. 2015-05-18]. Dostupné z: <http://videomattng.com/>

7 Přílohy

- skript s programem a metodami
- video a pozadí pro vyzkoušení algoritmu
- tři ukázky videí a videa s vyklíčovaným popředím

