

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

Fakulta elektrotechniky
a komunikačních technologií

DIPLOMOVÁ PRÁCE

Brno, 2017

Bc. Barbora Mlýnková



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA ELEKTROTECHNIKY

A KOMUNIKAČNÍCH TECHNOLOGIÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

ÚSTAV TELEKOMUNIKACÍ

DEPARTMENT OF TELECOMMUNICATIONS

ROZPOZNÁNÍ ODĚVU OSOB V OBRAZOVÉM SIGNÁLU

RECOGNITION OF PERSONS CLOTHING IN THE VIDEO SIGNAL

DIPLOMOVÁ PRÁCE

MASTER'S THESIS

AUTOR PRÁCE

AUTHOR

Bc. Barbora Mlýnková

VEDOUCÍ PRÁCE

SUPERVISOR

Ing. Jiří Přinosil, Ph.D.

BRNO 2017



Diplomová práce

magisterský navazující studijní obor **Telekomunikační a informační technika**

Ústav telekomunikací

Studentka: Bc. Barbora Mlýnková

ID: 144808

Ročník: 2

Akademický rok: 2016/17

NÁZEV TÉMATU:

Rozpoznání oděvu osob v obrazovém signálu

POKYNY PRO VYPRACOVÁNÍ:

Nastudujte teoretické možnosti automatické analýzy oděvu osob vyskytujících se v obrazovém signálu. Na základě získaných poznatků proveďte teoretický návrh vhodné metody pro daný účel a stanovte vstupní předpoklady a omezení jejího možného využití. Navrženou metodu implementujte ve vybraném programovacím jazyku a ověřte její spolehlivost na vhodné veřejně dostupné databázi.

DOPORUČENÁ LITERATURA:

[1] GALLAGHER, Andrew C.; CHEN, Tsuhan. Clothing cosegmentation for recognizing people. In: Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on. IEEE, 2008. p. 1-8.

[2] YANG, Ming; YU, Kai. Real-time clothing recognition in surveillance videos. In: 2011 18th IEEE International Conference on Image Processing. IEEE, 2011. p. 2937-2940.

Termín zadání: 1.2.2017

Termín odevzdání: 24.5.2017

Vedoucí práce: Ing. Jiří Přinosil, Ph.D.

Konzultant:

doc. Ing. Jiří Mišurec, CSc.
předseda oborové rady

UPOZORNĚNÍ:

Autor diplomové práce nesmí při vytváření diplomové práce porušit autorská práva třetích osob, zejména nesmí zasahovat nedovoleným způsobem do cizích autorských práv osobnostních a musí si být plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č. 40/2009 Sb.

ABSTRAKT

Tato práce se zabývá detekcí oděvu v obraze za využitím identifikace osob. Tyto znaky popisuje a kategorizuje. Dále se zabývá popisem struktury databáze, ve které pracuje s maskami a kategoriemi znaků pro jejich zpracování. Tato práce využívá haar kaskady k detekci obličeje a určení pozice oblečení, za účelem detekce barev.

KLÍČOVÁ SLOVA

Databáze, znaky, tagy, kategorizace, biometrické znaky, identifikace osob, RGB, HSV, Haar kaskady

ABSTRACT

This paper is dealing with the detection clothes characteristics in the picture, for the use of person identification. These characteristics are described and categorized. It also deals with the design of the database structure, which works with masks and categories of characteristics for their processing. This work uses haar cascades to detect face and to determine the position of clothing for the purpose of color detection.

KEYWORDS

Database, signs, tags, categorization, biometrics, person identification, RGB, HSV, Haar cascades

MLÝNKOVÁ, Barbora *Rozpoznání oděvu osob v obrazovém signále*: diplomová práce. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav telekomunikací, 2016. 47 s. Vedoucí práce byl Ing. Jiří Přinosil, Ph.D.

PROHLÁŠENÍ

Prohlašuji, že svou diplomovou práci na téma Rozpoznání oděvu osob v obrazovém signále jsem vypracoval samostatně pod vedením vedoucího diplomové práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

Jako autor uvedené diplomové práce dále prohlašuji, že v souvislosti s vytvořením této diplomové práce jsem neporušil autorská práva třetích osob, zejména jsem nezasáhl nedovoleným způsobem do cizích autorských práv osobnostních a/nebo majetkových a jsem si plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon), ve znění pozdějších předpisů, včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č. 40/2009 Sb.

Brno

.....

(podpis autora)

PODĚKOVÁNÍ

Ráda bych poděkovala vedoucímu semestrální práce panu Ing. Jiřímu Přinosilu Ph.D, za odborné vedení, konzultace, trpělivost a podnětné návrhy k práci.

Brno

.....

(podpis autora)

PODĚKOVÁNÍ

Výzkum popsáný v této diplomové práci byl realizován v laboratořích podpořených z projektu SIX; registrační číslo CZ.1.05/2.1.00/03.0072, operační program Výzkum a vývoj pro inovace.

Brno

.....
(podpis autora)

OBSAH

1	Biometrické systémy	12
1.1	Historie	12
1.2	Antropometrie	13
2	Identifikace osob	14
2.1	Identita	14
2.2	Identifikace	15
2.2.1	Pozitivní identifikace	15
2.2.2	Negativní identifikace	15
2.3	Verifikace	16
3	Databáze používaných fotografií a masek	17
4	Zpracování obrazu počítačem	18
5	Rozdělení, popis a sledované znaky nacházející se v obraze	20
6	Masky	22
6.1	Datová struktura	23
7	Struktura databáze	24
7.1	Vkládání dat do databáze	25
7.2	Navržený nástroj pro anotaci obrazu	26
8	Detekce v obraze pomocí klasifikátorů	28
8.1	Klasifikátory trénované neuronovými sítěmi	28
8.2	Klasifikátory trénované pomocí AdaBoostu	29
8.3	Haarovi příznaky	29
8.4	Detekce objektů	30
8.5	Kaskádový klasifikátor	30
9	Barevné modely	31
9.1	Rozdělení barev pro detekci barev	34
10	Detekce barev v obraze	35
11	Návrh aplikace	38
11.1	Knihovna OpenCV	38
11.2	Popis aplikace	38

12 Instalace a konfigurace	41
13 Závěr	42
Literatura	43
Seznam příloh	46
A Obsah přiloženého CD	47

SEZNAM OBRÁZKŮ

4.1	Ukázka pořízených fotografií ke zpracování.	18
4.2	Ukázka vyznačené polohy osoby na fotografii, její maska a původní fotografie.	19
6.1	Ukázka fotografie nesoucí informace o osobě a k ní náležící vytvořené bitmapové binární obrazy masek (tyto masky a originál na ukázkou jsou v nepoměru na rozdíl od těch s kterými v tomto projektu pracujeme.	22
7.1	Návrh databáze s názornou ukázkou vzájemných vazeb.	24
7.2	Ukázka vyznačené polohy osoby na fotografii, její maska a původní fotografie.	27
7.3	Ukázka vyznačené polohy osoby na fotografii, její maska a původní fotografie.	27
8.1	barevný model HSV[25].	29
9.1	Ukázka rozdílu aditivního a subtraktivního míchání barev[30][31] . . .	32
9.2	barevný model HSV[33].	32
9.3	barevný model HSV[34].	33
10.1	Fotografie použita v programu.	35
10.2	Znázornění geometrie osoby dle norem [37].	36
10.3	Vyznačená pole zájmu na příkladné fotografii.	37
11.1	Fotografie použita v programu s vyznačenými poli zájmu zpracování. . .	39
11.2	Výstupní obraz programu pro fotografii P1070458.	40
11.3	Výstupní hodnoty vypsané v konzoli programu pro fotografii P1070458. . .	40

SEZNAM TABULEK

2.1	Charakteristika rozdílů mezi pozitivní a negativní identifikací 2.1[2].	16
9.1	Přehled podmínek definujících jednotlivé barvy v modelu HSV	34
10.1	Hodnoty barev jednoho obrazu s použitím dvou různých masek.	35

ÚVOD

Hlavním cílem diplomové práce je bylo nastudování možnosti automatické detekce oděvu osob vyskytujících se v obrazovém signálu.

Je zde uvedena kategorizace informací, které můžeme získat z pohledu na obraz nesoucí informace o osobě. Zaměřuje se na nebiometrické znaky, které jsou pro nás v této diplomové práci nejdůležitější.

Při identifikaci osob jsou hlavní primární biometrické znaky. Pokud tyto znaky není možné detekovat, mohou pomoci sekundární biometrické znaky. Zaměřit se je možno právě na sekundární biometrické znaky¹. Pro nás jsou u identifikace osob z obrazu důležité zejména pohlaví člověka, věk, tělesné proporce, barva vlasů, silueta postavy. Z nebiometrických znaků tetování, doplňky jako např. brýle nebo hodinky, barva oblečení. Tato semestrální práce se zabývá právě barvou oblečení osob v obrazovém signálu.

Prvním krokem zpracování je všeobecné rozdělení sledovaných znaků. Každému bude přidána jeho charakteristika a masky popisující velikost a polohu. Základním prvkem je pro nás fotografie osoby, ze které budeme vycházet. Z fotografie budou vypracovány masky jednotlivých znaků pomocí bitmapového binárního obrazu. Bílá barva se bude nacházet na místě popisovaného znaku a černá okolo. Toto řešení bude umožňovat razantní snížení nároků na datové úložiště. Masky jsou vytvořeny pro všechny znaky, které se na fotografii u identifikované osoby nachází[1].

V diplomové práci budou využity polohy oděvů v obraze vycházející z rozložení geometrie člověka. Pro určení polohy oblečení osob v obrazovém signále musíme nejprve detekovat polohu tváře osob a z této definované oblasti budeme vycházet a zjišťovat podíl barev zastoupený v jednotlivých dílech oblečení osob.

Výsledky zastoupení barev budou porovnány s databází těchto fotografií a k nim přiřazenými vlastnostmi. Díky této databázi budeme moci zkontrolovat navržený program pro rozpoznání barev v obraze pro jednotlivé, námi testované masky. Tuto databázi jsem vytvořila v rámci bakalářské práce.

¹Biométrické znaky, které nejsou jedinečné pro každého člověka.

1 BIOMETRICKÉ SYSTÉMY

Než přejdeme k popisu identifikace osob, zaměříme se na ujasnění pojmu biometrické systémy[1] a pár zajímavostí z historie a jejich využití.

1.1 Historie

Již od pradávna je známo používání biometrických vlastností, denně je používáno biometrické rozpoznávání pro rozlišení lidí v okolí (podle hlasu, obličeje, chůze). Tyto lidské biometrické vlastnosti mohou být zaznamenány a zpracovány. Jsou to tedy signály nesoucí informace o biometrických vlastnostech.

Záznamy nejstarších biometrií ze 14. století n.l. byly kresby na skalních stěnách vyjadřující strukturu podobnou otiskům prstů nebo otisky prstů autora na keramice.

Historické doklady z carského Ruska byly zachovány. Jako nejstarší zachovaný dokument se udává nařízení velkoknížete Vasileje Dmitrijeviče z roku 1397 povinně cejchovat každého zloděje. Jiné nařízení z roku 1533 přikazuje usekávat ruce všem zlodějům. Od roku 1637 bylo zlodějům na tvář vypáleno znamení *VOR*¹.

Petr Veliký nařídil v roce 1694 cejchovat všechny omilostněné osoby, které byly předtím odsouzeny k smrti, písmenem B. Později přibyla další písmena, označující osoby posílané na nucené práce, do vyhnanství apod.[2].

První průkazné materiály o použití biometrie pocházejí z 19. století, tehdy se začaly používat otisky prstů v kriminalistice. Z dosud dochovaných písemných materiálů se konkrétně jedná o následující[3]:

- **William James Herschel**

1858 – anglický guvernér v Indii, který začal používat otisky prstů pro stvrzení negramotných dělníků, čímž bylo vyřešeno právoplatné převzetí peněz. Zároveň začal sbírat a zkoumat otisky prstů, což vedlo k sepsání díla o původu otisku prstů[4].

- **Francis Galton**

1865 – přišel se studií o dědičnosti fyzických vlastností[5], ve které rozebírá skutečnost, že děti dědí od rodičů některé vlastnosti (fyzické charakteristiky, vlastnosti jednání nebo chování).

1869 – spoluzakladatelem vědy *eugenika*, nauky o dědičných chorobách a vad u plodu.

1875 – zakladatelem výzkumu dvojčat.

¹Zloděj

1880 – zakládá vědní obor zabývající se měřením lidských tělesných rozměrů – *antropometrie*[6].

1892 – vydává ucelené dílo *Fingerprints*[7], z nějž vychází zavedení *daktyloskopie*[8] do praxe v roce 1900. V roce 1893 porovnává Francis Galton daktyloskopii s antropometrií a roku 1894 dochází k závěru, že obě metody jsou dobré a spolehlivé-obě metody jsou doporučeny k praktickému používání. Od roku 1896 je v Argentině zaveden identifikační systém na základě daktyloskopie.

1900 – daktyloskopie je prosazena pro identifikační a verifikační účely. Byla prokázána neměnnost a jedinečnost reliéfů kůže ². Daktyloskopie je zavedena do policejní praxe.

- **Alphonse Bertillon**

1882 – zabývá se postupem zvaným *Bertillonáž*[9]³ již od roku 1879. Dochází ke kolizi s postupem Francise Galtona.

- 1924 – založení oddělení identifikace otisků prstů u FBI (*Federal Bureau of Investigation*).
- 1965 – poprvé byl použit daktyloskopický systém AFIS (*Automated Fingerprint Identification System*)[10] s 810 tisíci otisků prstů.
- 2000 – systém AFIS u FBI obsahuje celkem 47 miliónů desetic otisků prstů, denně jsou průměrně 50 tisíckrát prohledány. Reakce na vzdálené vyhledávání v databázi činí přibližně dvě hodiny[3].
- 2010 – systém AFIS u FBI obsahuje celkem 66 milionů desetic otisků prstů, denně jsou průměrně 162 tisíckrát prohledány. Reakce na vzdálené vyhledávání v databázi činí přibližně jedna hodina a deset minut, v urgentním případě trvá deset minut[3].

1.2 Antropometrie

Obor, zabývající se měřením, popisem a rozbořem tělesných znaků charakterizujících růst a stavbu těla. Mezi základní znaky tělesného vývoje patří výška, hmotnost, objem hlavy a hrudníku. Výsledky získané prostřednictvím hromadných studií umožňují posoudit tělesný vývoj a v určitém smyslu i zdravotní stav.

²Papilárních linií na prstech

³Antropometrie a Bertillonáž[9] jsou shodné. Jedná se o metody měření a záznamu různých lidských rozměrů a jejich použití k identifikaci nebo verifikaci osoby.

2 IDENTIFIKACE OSOB

Identifikace osob je téma, které je velice aktuální, rychle se rozvíjející a využívané v mnoha odvětvích. V praxi se využívají hlavně primární biometrické znaky, u nichž je ovšem hodně důležitá kvalita dat ke zpracování. Například přesné rysy obličeje získáte jen při detailní a kvalitní fotografii. Pokud není možnost, aby snímání bylo možné na takovéto úrovni kvality, je potřeba zaměřit se na znaky, pomocí kterých můžeme eliminovat identifikovanou osobu pomocí dostupných dat[1].

U identifikace osob je nejdůležitější co nejpřesnější popis identifikované osoby. Za tímto účelem jsme museli rozvrhnout postup naší práce při zpracovávání fotografií osob. Z každé fotografie můžeme získat jen omezený počet informací o dané osobě, které musíme co nejlépe využít.

Nejprve si určíme, kde na obraze se daná osoba nachází. Tato poloha pro nás bude východiskem. Pro přesnou identifikaci hledíme na to, abychom z fotografie získali co největší počet znaků, které nám danou osobu co nejpřesněji popíší. Tyto znaky hledáme již pouze v oblasti vyznačené polohou osoby. Nejdůležitější informace u osob jsou pro nás znaky s menší pravděpodobností výskytu u osob s podobnými znaky.

Při zpětném vyhledávání poté jen zadáváme hledané znaky a vlastnosti, které námi hledaná osoba má mít. Pro přesné dohledání a identifikaci osoby je potřeba zadat co nejvíce znaků, pokud zadáme jen jeden máme velký počet lidí splňujících toto kritérium. Kombinací dalších znaků nám bude vycházet menší množina lidí splňujících hledanou kombinaci znaků.

2.1 Identita

Identita je jednoznačná charakteristika každého z nás. Je však třeba rozlišovat fyzickou a elektronickou identitu. Fyzickou identitu máme pouze jednu, tato identita je definována naším vzhledem a chováním. Na světě neexistuje člověk, který má shodnou fyzickou identitu s někým jiným (např. DNA je i u jednovaječných dvojčat odlišná). U elektronické identity to ovšem neplatí. V elektronickém světě si můžeme vytvořit identit tolik, kolik chceme – Jedná se např. o freemailové portály či různé identifikační karty[3].

2.2 Identifikace

Identifikace je popsána jako využití jedinečných, měřitelných, fyzikálních nebo fyziologických znaků (tzv. markantů¹) nebo projevů² člověka k jednoznačnému zjištění identity.

Pro identifikační účely se používají anatomické nebo fyziologické charakteristiky, které jsou pro každého člověka unikátní a časově neměnné[2].

Identifikace je proces porovnávání, ztotožnění (v anglosaské literatuře nazývaný i *One-To-Many-Matching*[11], jeden k mnoha, *1:n*, rekognice³) nasnímaného biometrického vzorku se všemi referenčními šablonami, uloženými v databázi (nebo seznamu apod.) vedoucího ke zjištění, která referenční šablona (existuje-li v databázi) odpovídá nasnímanému vzorku. Identifikující biometrická aplikace pak rozpozná totožnost prověřované osoby. Identifikace odpovídá na otázku *Kdo to je?*[2].

2.2.1 Pozitivní identifikace

Cílem pozitivní identifikace je zabránit používání identity jedné osoby dalšími osobami (tzn. musíme zabránit, aby se ostatní osoby prokazovaly mou identitou). Jestliže biometrická aplikace využívající princip pozitivní identifikace v procesu porovnávání nenajde shodu mezi šablonou předkládaného biometrického vzorku s žádnou referenční šablonou uloženou v databázi, výsledkem je odmítnutí přístupu (oprávnění) uživatele do objektu, počítačové sítě apod. Ztotožnění obou šablon naopak znamená přijetí uživatele[2].

2.2.2 Negativní identifikace

Cílem negativní identifikace je vyloučení nežádoucího stavu, kdy jedna osoba využívá identitu více osob (tzn. musíme zabránit osobě vydávat se za jinou osobu). Jestliže biometrická aplikace, která využívá princip negativní identifikace v procesu porovnání, nenajde shodu mezi šablonou uloženou v databázi, výsledkem je přijetí přístupu (oprávnění) uživatele. Ztotožnění obou šablon naopak znamená odmítnutí uživatele.

Rozdíl mezi pozitivní a negativní identifikací lze stručně charakterizovat pomocí tabulky Tab. 2.1 [2].

¹Charakteristické a významné příznaky, specifické znaky.

²*Chování člověka* můžeme chápat jako množinu jeho projevů.

³Rekognice je v kriminalistice chápána jako vizuální, zpravidla svědecká identifikace laického charakteru.

Tab. 2.1: Charakteristika rozdílu mezi pozitivní a negativní identifikací 2.1[2].

Pozitivní identifikace	Negativní identifikace
Cílem je prokázat, že já jsem již registrován v systému, databázi apod.	Cílem je dokázat, že já nejsem registrován v systému, databázi (a není tam registrován ani nikdo jiný s touto identitou).
Porovnání mnou předložené šablony s jedinou referenční šablonou. Jedná se o verifikaci.	Mnohonásobné porovnání mnou předložené šablony se všemi podobnými s cílem vyloučit nalezení případného duplikátu šablony.

2.3 Verifikace

Verifikace je ověření identity člověka. Proces porovnávání (*One-To-One Matching*, jeden ku jedné, 1:1, autentizace) jediné šablony vytvořené z nasnímaného biometrického vzorku s jedinou referenční šablonou, patřící prověřované osobě. Cílem je zjistit, zda prověřovaná osoba je opravdu tou osobou, za kterou se vydává nebo navenek jinak jeví. Biometrická aplikace potvrzuje nebo vyvrací identitu prověřované osoby[11].

3 DATABÁZE POUŽÍVANÝCH FOTOGRAFIÍ A MASEK

Databáze, kterou v této diplomové práci používám byla vytvořena v mé bakalářské práci[1], slouží k zařazení fotografií a k nim vytvořených masek. Databáze obsahuje přiřazené vlastností pro jednotlivé masky, které nesou informace o námi detekovaných sekundárních biometrických a nebiometrických znacích. Z této databáze osob bude práce dále vycházet, i zvolený postup pro detekci barvy oděvu.

U vytváření bitmapových masek znaků osob na fotografiích byl kladen důraz na stejný postup a styl při jejich tvorbě, aby bylo možné tyto data dále použít.

V této diplomové práci slouží jako vstupní data k vyhodnocení a ke kontrole správnosti vyhodnocení barvy oblečení na fotografiích v programu.

Použité fotografie byly vytvořeny autorkou této práce pomocí fotoaparátu Panasonic DMC-FZ28. Všichni lidé na těchto fotografiích dali souhlas k zachycení, zobrazení a zveřejnění své podoby (fotografie) podle ust. § 84 a násl. zákona č. 89/2012 Sb., občanský zákoník, v platném znění a podle zákona č. 101/2000 Sb., o ochraně osobních údajů, v platném znění udělili vlastnoručním podpisem souhlas.

4 ZPRACOVÁNÍ OBRAZU POČÍTAČEM

Je třeba uvažovat, že počítač zpracovává obraz jiným způsobem než lidské oko. Lidé vyhodnocují informace o obrazu na základě minulých zkušeností a přirovnávání ze vzpomínek, zkušeností, citů a okolních vlivů, kterých se v minulosti naučili. Počítače, na rozdíl od nás lidí, nemají tuto možnost a nedokáží takto jednoduše vyhodnotit informace z obrazu.

Ukázka fotografií ke zpracování informací o osobě z obrazu viz obr. 4.1.



Obr. 4.1: Ukázka pořízených fotografií ke zpracování.

Při vyhodnocování obrazu je nutné dané informace vyznačit a rozkategorizovat. Každému prvku přiřadit vlastnosti (barva a velikost, souřadnice místa, kde se nachází).

Určení polohy zkoumaného znaku je velice důležité, jelikož se často nemění a lze z něj odvodit polohu ostatních částí (vlasy budeme mít vždy na hlavě), z toho

můžeme usoudit, že oblečení bude ve spodnější části.

Začneme od největší oblasti fotografie – identifikace postavy 4.2, kde přesně se nachází, na kterou část obrazu je třeba se zaměřit.

V této semestrální práci jsou pro nás sledované znaky v následující části: trup a popis oblečení (poloha, barva, délka), současně oblečení spodní části těla. Podrobné rozdělení znaků je popsáno v kategorii 5.

Pro strojové učení je důležité učení realizovatelných pravidel[12].



Obr. 4.2: Ukázka vyznačené polohy osoby na fotografii, její maska a původní fotografie.

5 ROZDĚLENÍ, POPIS A SLEDOVANÉ ZNAKY NACHÁZEJÍCÍ SE V OBRAZE

diplomová práce využívá a zabývá se pouze kategorií oblečení trupu a nohou, ale pro přehled je zde uveden popis pro všechny kategorie vytvořených znaků, které jsou obsaženy v mé databázi zohledněny. Znak pro zpětné vyhledávání, byly zařazeny do kategorií pro snadnější orientaci a vyhledávání. Sledované znaky byly po úvaze rozděleny do 14 kategorií. Kategorie byly voleny tak, aby zahrnuly všechny masky a intuitivně si souhlasily[1].

Identifikace demografických atributů člověka, jako je věk, pohlaví a etnický původ s využitím počítačového vidění, byla věnována zvýšená pozornost v posledních letech. Tyto znaky mohou hrát důležitou roli v mnoha aplikacích, jako je interakce člověka s počítačem, podle obsahu indexování a vyhledávání, biometrii, demografickým studiích a cílené reklamě[13].

Rozdělení, popis a sledované znaky nacházející se v kategoriích. Tyto znaky jsou pro nás velice důležité, bude se na ně směřovat práce s databází a popis masek fotografií. **Pohlaví** První kategorií je pohlaví, obsahující pouze dva sledované znaky:

- Žena
- Muž

Tyto znaky jsou pro nás ovšem velice důležitým ukazatelem. Jde o uznání lidské rovnosti žen a mužů v počítačovém vidění.

K rozhodování dochází po přezkoumání využití informací z obličeje a celého těla (z pořízeného dosavadního obrazu nebo sekvence chůze), kde jsou prezentovány[13]. Tyto dva znaky popíší celou množinu potenciálních kandidátů na správné vyhledání.

Pokud zpětně v databázi vyhledáváme, již výběrem tohoto znaku, eliminujeme polovinu možných kandidátů.

Věk Tento znak se nám nepodaří nikdy dokonale určit, vždy bude pouze odhadem, ale i přes tuto nevýhodu je pro nás zásadní. Tato kategorie má právě tři znaky: Dítě, Dospělý, Stařec

Výška Tato kategorie je rozšířena a zdokonalena až za pomoci praktického použití databáze k vyhledávání osob. K určení přesné výšky je zapotřebí více obrazů pro výpočet (mohla by být použita například rychlost chůze, nebo porovnání se známým objektem jehož vlastnosti bychom znali).

Postava Proporce postavy jsou pro nás velice důležitým klasifikátorem, jejich rozdělení bude s přibývajícím počtem dat a vzorků zdokonaleno.

Rasa Koncept rasy (neboli plemena) je mezi laiky i v odborné literatuře natolik ustálen, že nějaký čas potrvá, než se tohoto termínu ve smyslu *skupina lidstva lišící se od jiné různými typickými znaky, zejména barvou pleti nebo skupina lidstva*

s podobnými tělesnými znaky zbavíme[14].

Vlasy a vousy Vlasy a vousy jsou velmi proměnlivým prvkem lidského vzhledu, neboť zde existuje mnoho variant. Jeho automatická detekce je velice náročná, jsou ale důležitým prvkem lidského vzhledu, ale i detekce, reprezentace, analýzy a nebyla dostatečně studována v oblasti počítačového vidění. Rozsáhlá diskuse vlastností vlasů pro počítačové vidění, vlasy mohou být zastoupeny v souladu s následujícími rozměry: délka, objem, plocha, dominantní barvy, čelo, hustota, souměrnost, rozdělení umístění, lesk[15].

Oblečení hlavy a doplňky Doplňky jsou dobře sledovatelným znakem, jelikož bývají hodně velké a pozorné. Oblečení hlavy je převážně sezónní záležitostí. Díky těmto rysům se velice dobře značí a vyhledávají. Nevýhoda doplňků je obsažena v tom, že hledaná osoba, kterou se pokoušíme identifikovat, se jich může velice rychle a snadno zbavit.

Obuv Přestože je obuv velice snadno rozdělitelná podle stylu (společenská – sportovní, dámská – pánská, zimní – letní). Tato vlastnosti ovšem nelze rozpoznat a přesně identifikovat u fotografie osoby kde třeba ani nevidíme jestli má obuv podpatky (pomocí něj bychom snadno přiřadily že se jedná o dámskou obuv)[17].

Kategorie znaků na které se zaměříme v této diplomové práci jsou oblečení trupu, nohou a jejich barva.

Oblečení trupu a nohou U oblečení nastává problém s velkou variabilitou možností oblečení, pokud bychom měli každý styl kategorizovat, přiřazení vytvořených znaků by nebylo přesné pro vrstvení oblečení jež nelze z obrazu rozpoznat. Proto se v této diplomové práci zaměříme na jeho barvu a umístění jeho polohy na osobě, která se nemění a můžeme z ní tedy dále vycházet při vyhodnocování.

Barva Tato kategorie nám slouží k přesnějšímu popsání vlastností u oblečení trupu, oblečení nohou, obuv i doplňků.

Pro rozdělení sledovaných vlastností a popis znaků nám vystačí zjednodušená paleta barev z těch nejčastěji používaných[16][17]:

Černá	Hnědá	Žlutá	Modrá
Bílá	Červená	Zelená	Fialová
Šedá	Oranžová	Azurová	Růžová

6 MASKY

Masky jsou vytvořené jako bitmapové binární obrazy. Bílá barva se nachází na místě popisovaného znaku a černá okolo viz. Obr. 6.1¹. Toto řešení bude umožňovat razantní snížení nároků na datové úložiště. Tyto masky budou vytvořeny pro všechny znaky, které se na fotografii u identifikované osoby nachází.



Obr. 6.1: Ukázka fotografie nesoucí informace o osobě a k ní náležící vytvořené bitmapové binární obrazy masek (tyto masky a originál na ukázkou jsou v nepoměru na rozdíl od těch s kterými v tomto projektu pracujeme).

¹Pro znázornění jsou zde masky pro jednu danou fotografii, maska polohy osoby, vlasů, trupu, maska s doplňky – hodinky, oblečení noh a maska bot.

Ve skutečnosti je rozměr fotografie i jejích masek totožný pouze na tomto obraze pro názorost jsou upraveny.

Přemapováním obrazu dostaneme pomocí spojení originální fotografie a masky fotografii s vyznačeným obrysem popisovaného prvku.

6.1 Datová struktura

Ukládání masek muselo být rozděleno, aby bylo možné při jejich editaci a přiřazování znaků určit, pro kterou fotografii osoby jsou masky vytvořeny.

Nejjednodušší je pro nás struktura složek. Máme hlavní složku, ve které se nachází podsložky. Každá z nich obsahuje právě jednu originální fotografii (fotografie je ve formátu .JPG) nesoucí informace o námi popisované osobě. Tato složka je pojmenována podle této fotografie, kterou obsahuje. Dále jsou v této složce masky dané fotografie popsány v kap. 6 a jsou ve formátu .png [18].

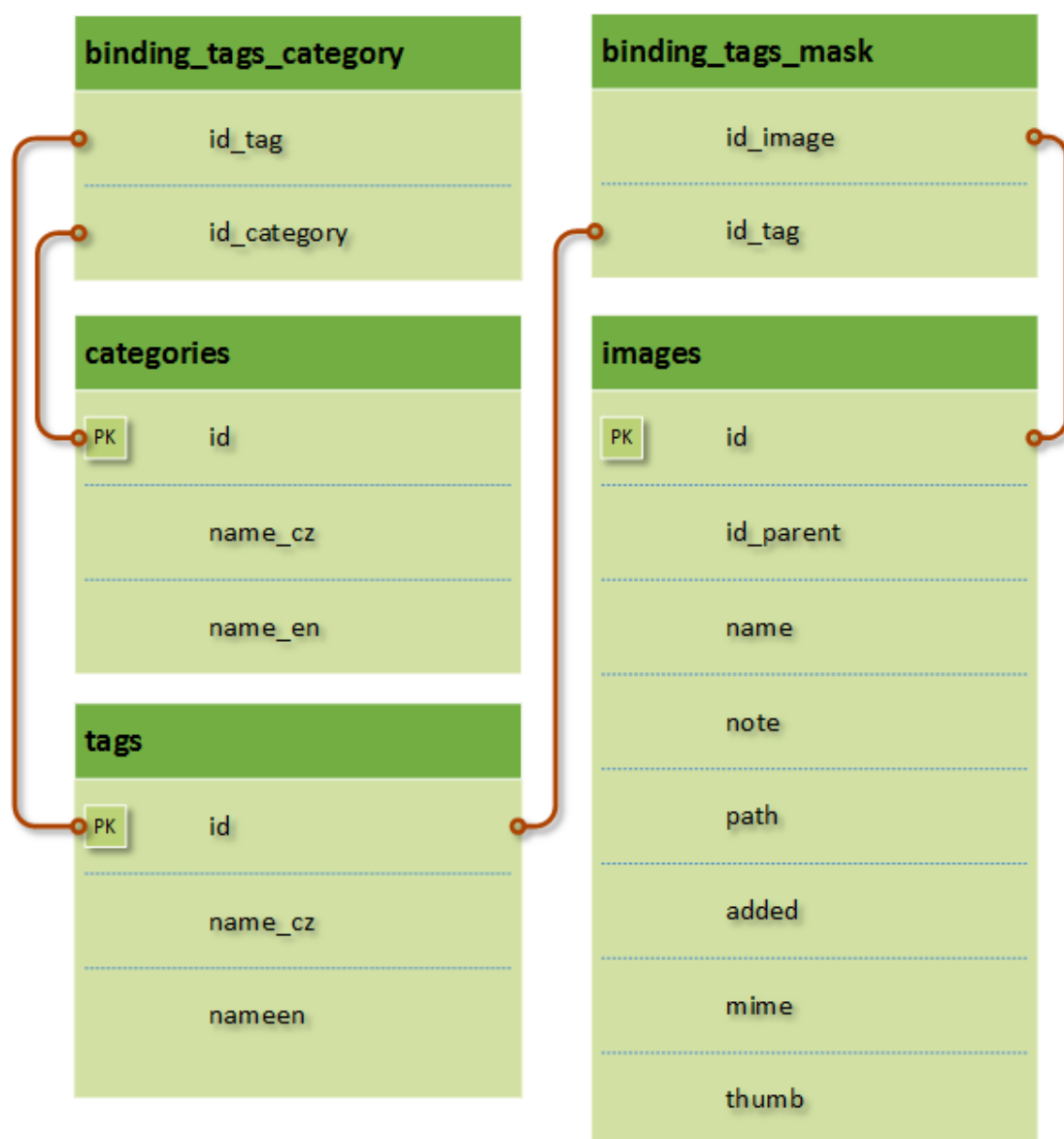
Tento formát je pro nás u dvoubarevných masek výhodnější než .jpg, který používá ztrátovou kompresi a dochází k velkému zkreslení ostrých hran a je datově náročnější [19]. Takže díky použitému formátu .png zmenšíme datové nároky masek, což vede k rychlejšímu zpracování v praktickém použití.

7 STRUKTURA DATABÁZE

Databáze byla vytvořena se zaměřením na možnost zpracování velkého množství dat, která hierarchicky rozdělíme a postupně zde popíšeme jejich rozdělení a vzájemné vztahy v databázi viz. obr. 7.1.

Databáze je vytvořena relačním databázovým systémem SQLite. Na rozdíl od MYSQL použité v semestrálním projektu zde není potřeba databázový server což bylo důvodem pro volbu této databáze [20].

Databáze se nachází v jednom souboru, to nám usnadní práci vytvořeného nástroje pro anotaci databáze.



Obr. 7.1: Návrh databáze s názornou ukázkou vzájemných vazeb.

Jak je z obrázku 7.1 patrné, všechny tabulky databáze jsou spolu provázány. Databáze je rozdělena do tří tabulek (images, tags a category). Popis funkce jednotlivých tabulek:

- Images – zde jsou vloženy fotografie a masky nesoucí informace o identifikovaných osobách. Každá nová fotografie je zařazena do tabulky a je jí přiřazeno jedinečné ID číslo. V databázi je třeba přiřadit fotografii masky k ní vytvořené, abychom tohoto docílili nastavujeme k maskám nový parametr ID parent což je právě ID číslo fotografie pro kterou byla přiřazovaná maska vytvořena.
- Category – zde jsou v tabulce uspořádány kategorie pro rozdělení znaků (rozpis kategorií 5). Editace kategorií je možné realizovat i pomocí vytvořeného nástroje pro anotaci databáze.
- Tags – v této tabulce jsou vypsány všechny námi používané znaky. Každému znaku je přiřazeno jedinečné ID číslo. Databáze je navržena tak, aby bylo možné, kdykoli bude potřeba, přidat nový parametr znaku 5. Editace znaků je možná provádět i pomocí vytvořeného nástroje pro anotaci databáze.

V databázi jsou použity dva cizí klíče [21].

První na místě svázání znaků a kategorií 5, tato varianta nám umožňuje přiřazení znaků přímo k dané kategorii, to považuji za efektivnější vzhledem k budoucímu praktickému použití, jedna kategorie může mít přiřazeno více znaků (například délku+barvu).

Druhý cizí klíč je použit mezi znaky (popisovače hledaných tagů) a maskami. Tímto řešením máme navrženou databázi pro fotografie a jejich masky. V budoucnu můžeme přiřazovat libovolný počet tagů, jelikož s použitím 1:N (kategorie : znaky) lze přiřadit kategorii nové znaky dle aktuální potřeby bez nutnosti změny schématu databáze.

Jako datové úložiště byla zvolena SQLite databáze.

7.1 Vkládání dat do databáze

Plnění dat v databázi probíhá tak, že nejprve naimportujeme strukturu znaků a kategorií, tyto prvky vkládáme do databáze ručně nebo pomocí vytvořeného nástroje pro anotaci databáze. Tento proces neprobíhá pouze na počátku, dále je možné upravit nebo přidat další znaky či kategorie pokud v průběhu práce zjistíme, že by se nám hodil další parametr.

Import fotografií probíhá pomocí umístění ve složkách. Každá fotografie nesoucí informaci o osobě je vložena v jedinečné složce pojmenované podle názvu originální fotografie. V každé složce jsou k originální fotografii s osobou přidány i masky, které jsou vytvořeny pro danou fotografii a soubor, obsahující fotografii s maskami

ve formě vrstev, kde je stále možnost je editovat v programu GIMP, zjistíme-li v průběhu práce chybu. Do databáze jsou nahrány pouze soubory ve formátu .JPG a .png (fotografie osoby + masky).

Celkově jsem zpracovala 338 fotografií nesoucích informace o osobách a k nim vytvořila 2120 černobílých masek.

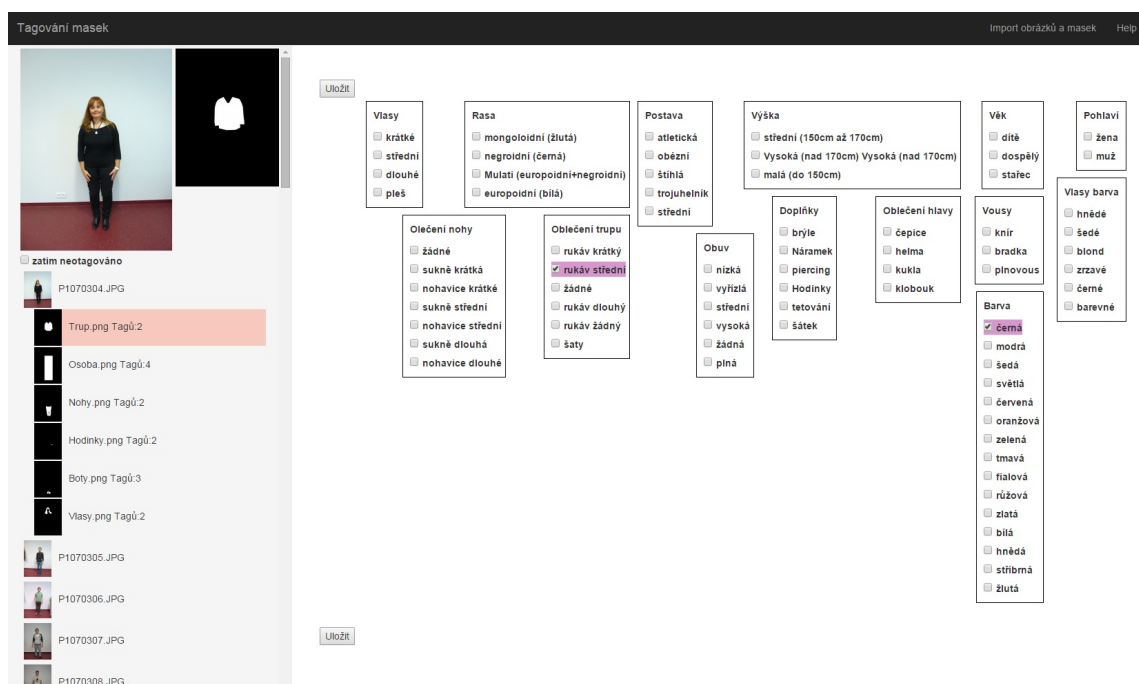
7.2 Navržený nástroj pro anotaci obrazu

Nástroj pro anotaci obrazu je vytvořený v .php a javascriptu, je provozován na webovém serveru apache, pro představu je dostačující. Při importu fotografií je nutné pracovat s miniaturami obrazů, s originály je práce příliš pomalá.

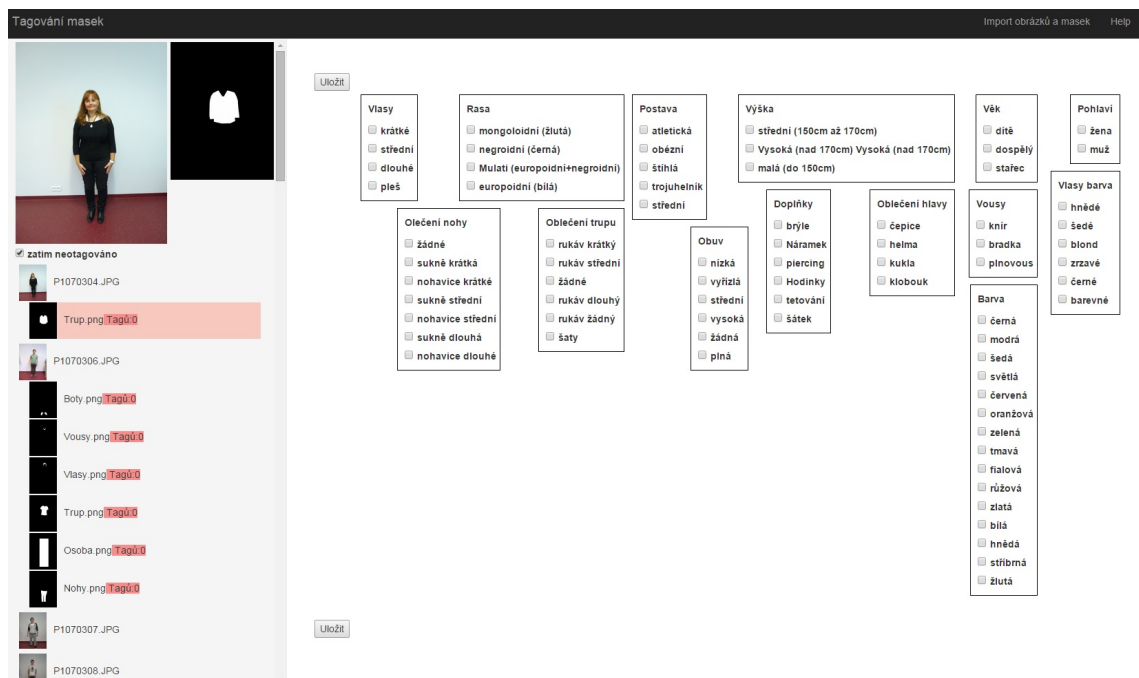
Na obrázku 7.2 je vidět, jak je rozděleno a navrženo uživatelské rozhraní. Vrchní ovládací lišta obsahuje nástroj pro import fotografií a masek do databáze, odkud s nimi budeme pracovat. Pro import je potřebné nové fotografie, které chceme přidat do databáze, vložit do složky import.

Fotografie jsou umístěny u levého okraje obrazu. Pro vybranou fotografii se vždy v levém horním rohu zobrazí editovaná fotografie a vybraná maska. Ve zbylém prostoru se nám zobrazí kategorie se znaky. Každé masce můžeme přiřadit znaky (tagy). Pro uložení zvolených tagů vybrané masce je nutné kliknout na tlačítko uložit. Poté se nám u masky zbarví vybrané tagy fialově pro rychlejší přehled viz. Obr. 7.2.

Je zde také další funkce, která má zabránit zapomenutí přiřazení tagů maskám. Vlevo mezi miniaturami a zvětšenou vybranou fotografií se nachází výběr zatím neotagováno, pokud ho zaklikneme zobrazí se nám pouze ty fotografie, u kterých je maska, která nemá přiřazený znak. Díky tomuto si můžeme zkontrolovat, zdali máme všechny masky a fotografie otagovány viz. Obr.: 7.3.



Obr. 7.2: Ukázka vyznačené polohy osoby na fotografii, její maska a původní fotografie.



Obr. 7.3: Ukázka vyznačené polohy osoby na fotografii, její maska a původní fotografie.

8 DETEKCE V OBRAZE POMOCÍ KLASIFIKÁTORŮ

Klasifikátory jsou šablony trénované na určité skupiny vzorů (auta, tváře, ovoce...). Podle jejich názvu lze odvodit jejich hlavní úlohu klasifikace (třídění, hodnocení a rozdělování). Procházejí obrázek jako celek nebo jen jeho část a vyhodnocují, zda se v této oblasti nachází nebo nenachází hledaný objekt. Klasifikátory se podle jejich hodnoty, kterou vrací, rozdělují na binární a vícehodnotové. Pokud je klasifikátor binární, jeho zpětná vazba je pozitivní nebo negativní, u vícehodnotových se vrací pravděpodobnost s kterou se v této oblasti hledaný objekt nachází.

8.1 Klasifikátory trénované neuronovými sítěmi

Neuronová síť je struktura používaná na distribuované paralelní zpracování dat. Je vytvořena na základě biologických nervových systémů. Skládá se z umělých neuronů, které také vycházejí z reálných biologických neuronů. Každý umělý neuron má libovolný počet vstupů, ale pouze jeden výstup. Podle typu neuronu a typu neuronové sítě se použije vhodná přenosová funkce. Neurony jsou vzájemně propojeny a navzájem si předávají signály a transformují je pomocí určitých přenosových funkcí. Umělé neurony jsou vytvořené na základě určitých modelů, kterých je popsáno mnoho, od jednoduchých až po velmi složité. Jedním z nejpoužívanějších je model McCullochem a Pittsem[22].

Učení neuronové sítě

Cílem učení neuronové sítě je nastavit síť tak, aby dávala přesné výsledky. Umělé neuronové sítě ukládají své zkušenosti v matematickém ekvivalentu (vahách). Učení neuronové sítě rozlišujeme na učení s učitelem a učení bez učitele. Fáze učení neuronové sítě bývá nazývána adaptivní. Po naučení neuronové sítě je síť ve fázi vybavování[22].

- **Učení s učitelem** – při tomto učení potřebujeme mít sadu vzorů a k nim výsledky, které jsou správné. Tyto správné výsledky porovnáváme s výsledky, které jsme dostali na výstupu neuronové sítě, určíme výslednou chybovost. Podle typu neuronové sítě se vypočítá korekce a upraví se hodnoty prahů a vah tak, aby jsme zmenšili chybovost, což opakujeme dokud neklesne pod určenou hodnotu maximální chyby.
- **Učení bez učitele** – toto učení je ideální v případech kdy máme sadu vzorů ale nemáme k nim správné výsledky. Neuronová síť si celou sadu sama roztrídí do skupin a sleduje podobnosti nebo typické zástupce.

Výhody – paralelní zpracování informací, když se použije učení s učitelem dokáže se přizpůsobit změně parametrů, když se použije učení bez učitele dokáže být neuronová síť velmi rychlá.

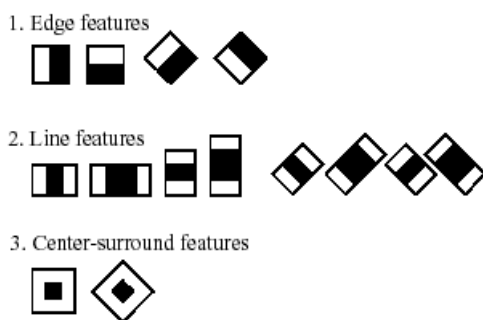
Nevýhody – Učení složitých neuronových sítí trvá dlouho, nelze použít tak, kde jsou vyžadovány přesné výsledky, bez univerzální metody pro návrh architektury sítí proto se postupuje metodou pokus – omyl a to zvyšuje časovou náročnost.

8.2 Klasifikátory trénované pomocí AdaBoostu

AdaBoost je lineární klasifikátor, jeho výstup konverguje na logaritmus poměru pravděpodobnosti. Je to algoritmus pro tvorbu silného klasifikátoru na základě lineární kombinace slabých klasifikátorů. Každému slabému klasifikátoru přiřadí určitou váhu na základě jeho chyby a podle ní se může v následujícím trénování rozhodnutí opravit. Hlavní vlastností AdaBoostu je odolnost k přetrénování (situace, kdy při trénování s velkým množstvím dat dochází k horším výsledkům). Pomocí tohoto algoritmu se dají v poměrně krátkém čase natrénovat klasifikátory s malou chybovostí, používají se často při detekci tváře v reálném čase[23].

8.3 Haarovi příznaky

Na začátku trénovacího procesu jsou vygenerovány Haarovy příznaky. Tyto příznaky jsou odvozeny od několika obdélníků, které se dělí podle typu informace, která má být s jejich pomocí detekována na: hranové, čárové a středové příznaky viz. Obr. 8.1. Každý z příznaků je generován pro detekci určitého charakteristického znaku nacházejícího se v detekčním okně, což je část (výřez) analyzovaného obrazu nejčastěji čtvercového tvaru. Pro danou velikost detekčního okna jsou příznaky generovány pro každý typ, velikost a pozici Haarova příznaku, jejich počet je tedy velmi vysoký (až stovky tisíc)[24].



Obr. 8.1: barevný model HSV[25].

Haarové příznaky jsou rychlé pomocí sčítání pixelů, na které se používá integrální obraz, popsany v [26]. Integrální počet slouží k urychlení výpočtu sum. Můžeme si ho představit jako transformaci obrazu do dvourozměrného pole s rozměry o jeden pixel většími velikost původního obrazu. Každý jednotlivý prvek poté obsahuje součet všech prvků v oblasti nalevo a nahoře od tohoto bodu, z čehož vyplývá, že první řádek a sloupec mají hodnotu nula, pixel vpravo dole obsahuje sumu všech pixelů.

8.4 Detekce objektů

Detekce objektů v obraze se vykonává porovnáním určitého výřezu obrázku se vzorem. Výřez porovnání mění každým krokem svou pozici a velikost, dokud neprojde celý obrázek. Velikost výřezu se volí vynásobením originální velikosti a zvoleného kroku zvětšení. Zvětšování výřezu probíhá dokud některá hrana nepřekročí velikost obrázku. Do klasifikátoru vstupují příznaky získané z těchto oblastí a ten rozhodne, jestli se příznaky rovnají se vzorovými příznaky. Významného zrychlení se dosáhne pomocí kaskádového zapojení klasifikátorů[27].

Kaskádovým zapojením klasifikátorů se získá sada klasifikátorů o X stupních, kde poslední je nejsložitější a první je nejjednodušší. Úkolem jednotlivých klasifikátorů je vyřadit co největší počet negativních oblastí a tím označit co nejmenší oblast pro složitelnější klasifikátory.

8.5 Kaskádový klasifikátor

Kaskáda v klasifikátoru znamená, že výsledný klasifikátor se skládá z několika jednodušších slabých klasifikátorů, které jsou následně aplikovány na oblast zájmu. Význam kaskády je, aby byla doba hledání objektu co nejkratší. Každý klasifikátor v kaskádě má určený svůj práh, podle kterého rozhodne, zda-li je daný sub-obrázek pozitivní nebo negativní. Protože počítat např. všech 180 000 příznaků obsažených v jednom sub-obrázku o rozměru $X \times X$ je velmi nepraktické. Postup při detekci je tedy takový, že první klasifikátor odstraní co nejvíce negativních sub-obrázků a k dalšímu klasifikátoru pošle pozitivní sub-obrázky. Druhý klasifikátor na řadě by měl být nejlépe složen ze všech slabých klasifikátorů z předchozího stupně. Správně upravená kaskáda by měla mít co nejlepší úspěšnost a co nejméně falešných detekcí[28].

9 BAREVNÉ MODELÝ

Po fyzikální stránce je barva definována vlnovou délkou světla. Barvy které jsme schopni vnímat (takzvané viditelné spektrum) jsou tvořeny směsí světla různých vlnových délek v rozmezí cca 380 – 800 nm[29].

Pro digitální zpracování je ale třeba barvu nějak reprezentovat (kódovat). Tato reprezentace musí být co nejjednodušší (typicky 3 složky) a zároveň co nejrobustnější, neboli co nejlépe popisovat barevné spektrum. Taková reprezentace barev je vždy určitým zjednodušením reality, nemůže reálné spektrum barev popisovat naprosto přesně. Z tohoto důvodu vznikly mnohé způsoby digitální reprezentace barev, které se zaměřují na určitou oblast využití. Některé jsou zaměřeny na fyzickou reprodukci tiskem (CMYK), jiné na elektronické zobrazování (RGB) a některé se s výhodou využívají pro zpracování obrazu (HSV)[36].

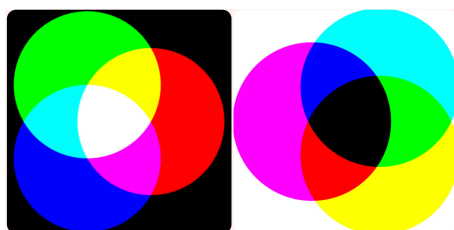
Barevný model je abstraktní matematický model popisující způsob, jakým barvy mohou být reprezentovány jako kombinace čísel, typicky tři nebo čtyři hodnoty popřípadě kombinace barevných složek.

Barevný model používá základní barvy a mísení těchto základních barev do výsledné barvy. Barva světla je totiž dána v nejjednodušších případech i jedinou vlnovou délkou viditelného záření, běžně v přírodě však jde o směs mnoha záření různých vlnových délek, někdy dokonce o směsi celých spojitých spekter. Možných kombinací je mnoho, lidské oko z nich navíc vnímá pouze některé. Matematicky lze definovat jeho antropomorfní barevný prostor (nebo alespoň jeho aproximaci), navíc absolutně měřitelný v reálných fyzikálních jednotkách. Barevné modely pak popisují v rámci daného prostoru především relativní poměry jednotlivých bodů snímku mezi sebou vzájemně, než aby bylo cílem měření jednoho bodu jako takového. Různé barevné modely se navíc sice snaží napodobit barvy co nejvěrněji, v praxi se však používají modely, u kterých je zvolen vhodný kompromis mezi přesností podání barevného dojmu na jedné straně a složitostí konkrétního modelu a obtížností jeho použití na straně druhé.

Barevné modely se dají podle způsobu skládání barev rozdělit do dvou skupin podle způsobu míchání barev na aditivní a subtraktivní.

Aditivní modely pracují na principu sčítání barev (vlnových délek), výsledkem je světlo větší intenzity než jednotlivé složky. Podobá se skládání barevného světla - odpovídá vzájemnému prolínání tří světelných kuželů, které mají filtr odpovídající základní barvě viz. Obr. 9.1. Aditivní skládání barev pracuje se třemi základními barvami: modrou, červenou a zelenou. Základním zástupcem aditivního barevného modelu a zároveň nejpoužívanějším modelem je RGB. Aditivní skládání barev se používá při zobrazování na barevných monitorech, displejích a projektorech.

Subtraktivní modely pracují naopak na principu odčítání barev. Každá barevná složka „ubere“ určité vlnové délky světla a zbylé vlnové délky tedy určují výslednou barvu viz. Obr. 9.1. Tohoto principu se používá hlavně při tisku barevnými pigmenty. Nejznámější zástupce subtraktivního skládání barev je model CMY (CMYK). Tyto modely ale nemají v mé další práci využití.

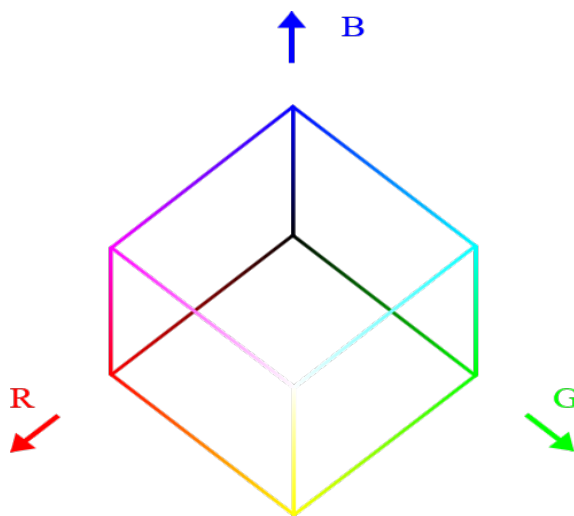


Obr. 9.1: Ukázka rozdílu aditivního a subtraktivního míchání barev[30][31]

RGB

Barevný model RGB je nejbližší spojen s lidským okem, které je složeno ze tří složek citlivých na elektromagnetické záření vlnových délek. Tyto složky využívají základních barev červená, zelená a modrá viz. Obr. 9.2. Tento model se nazývá aditivní, protože se jedná o míchání barev. Výsledná barva odpovídá sloučením tří základních složek, každá v rozmezí 0-255. Pro získání černé barvy musí být zastoupeny všechny intenzity složky na nule (0, 0, 0). Naopak bílá barva má všechny složky s plnou intenzitou (255, 255, 255).

Použití RGB modelu je hlavně v monitorech nebo projektorech. Jelikož se jedná o míchání vyzařovaného světla, tak tyto zařízení nepotřebují žádné okolní světlo[32].

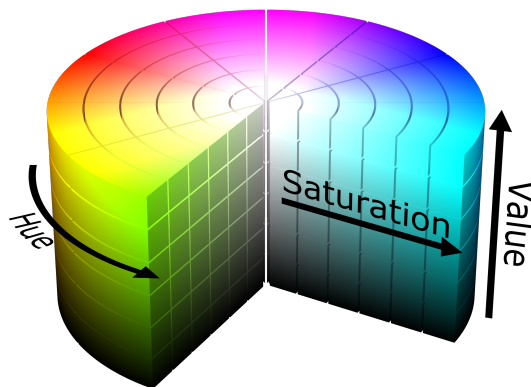


Obr. 9.2: barevný model HSV[33].

HSV

Barevný model HSV je skládá ze složek, které jsou reprezentovány třemi vlastnostmi základních barev viz. Obr. 9.3. Tyto vlastnosti jsou:

- **Hue** (barevný tón) – zde se používá tzv. barevný kruh, kde je odstín barvy určen polohou (0° až 360°). Pro tři základní barvy platí tyto zlomky úhlu (červená 350° – 360° a 0° – 8° , zelená 75° – 146° a modrá 205° – 254°).
- **Saturation** (sytost barvy) – Někdy též chroma, čistota nebo síla barvy, představuje množství šedi v poměru k odstínu, měří se v procentech od 0% (šedá) do 100% (plně sytá barva). Na barevném kruhu vzrůstá sytost od středu k okrajům.
- **Value** (hodnota jasu) – Relativní světlost nebo tmavost barvy. Jas vyjadřuje kolik světla je barva schopna odrazit, dalo by se také označit jako přidávání černé do základní barvy[32].



Obr. 9.3: barevný model HSV[34].

V této práci využíváme právě těchto dvou barevných modelů RGB a HSV. Fotografie jsou pořízeny a digitálně zpracovány pomocí modelu RGB, při jejich zpracování a vyhodnocování jsou barvy převedeny do HSV viz. v Tab. 9.1.

9.1 Rozdělení barev pro detekci barev

Tab. 9.1: Přehled podmínek definujících jednotlivé barvy v modelu HSV

HSV	barevný tón (H)	sytost barvy (S)	hodnota jasu (V)
Černá		0–100	0–70
Bílá		0–45	180–255
Šedá		0–45	70–130
Hnědá	10–48		40–75
Červená	0–4 a 175–180		
Oranžová	5–18		
Žlutá	19–37		
Zelená	38–73		
Azurová	74–102		
Modrá	103–127		
Fialová	128–149		
Růžová	150–175		

V Tab. 9.1 jsou uvedeny podmínky, kterými v této práci určujeme jednotlivé barvy. Vycházíme nejprve od hodnot nízké sytosti barvy, kterou mohou obsahovat pouze černá, šedá a bílá. Tyto barvy se mezi sebou rozlišují hodnotou jasu, od nejnižší pro černou přes šedou k nejvyšším hodnotám popisující barvu bílou.

Jakmile máme tyto hodnoty z množiny barevného modelu definovány, můžeme dále za pomoci barevného tónu definovat ostatní barvy, tak jak je uvedeno v Tab. 9.1, kromě barvy hnědé, která je zapotřebí specifikovat i hodnotou jasu.

10 DETEKCE BAREV V OBRAZE

Hlavním úkolem projektu, bylo zvládnutí rozpoznání zastoupení barev v obrazovém signále jednotlivých fotografií. Originální fotografie s obrazovou informací o osobách a jejich oblečení jsou primárně pořízeny v barevném modelu RGB [35].

Program pro určení barvy za pomoci masky je napsán v Programu Visual Studio 2015 C++ s využitím OpenCV verze 3.1.

Jelikož v barevném modelu RGB je definice jednotlivých barev složitá, využila jsem v tomto programu barevného modelu HSV a definování barev právě v tomto modelu [36].

Dále je uvedena fotografie viz obr. 10.1 na které byla použita detekce barev v oblasti určené maskou. a hodnoty vyhodnocené k této fotografii se nacházejí v Tab. 10.1 :

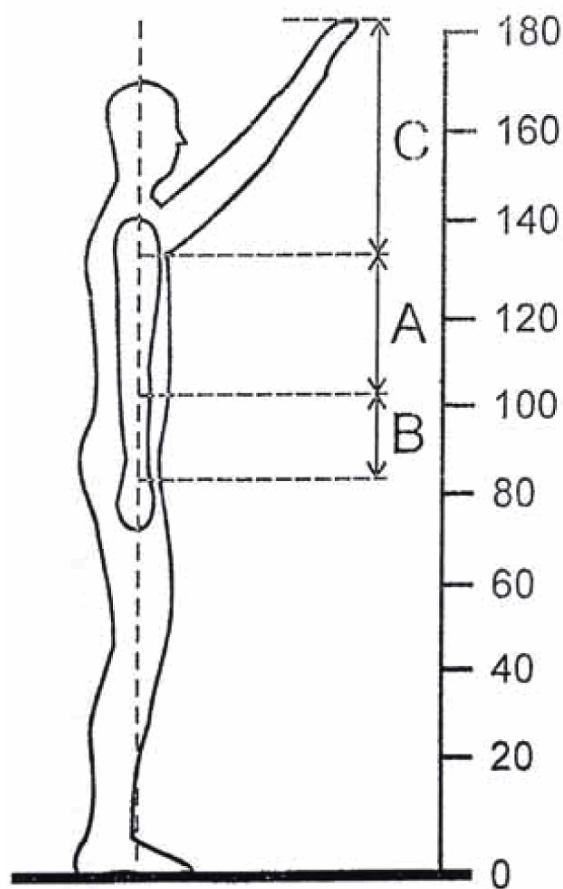


Obr. 10.1: Fotografie použita v programu.

Tab. 10.1: Hodnoty barev jednoho obrazu s použitím dvou různých masek.

Barva:	Trup	Nohy
Černá	0,10 %	80,92%
Šedá	0,08%	3,19%
Červená	58,78%	0,94%
Oranžová	5,45%	0,00%
Modrá	30,80%	13,82%
Růžová	3,02%	0,25%

Diplomová práce vychází při určování barev oděvu osob z fotografií bude vycházet z geometrie lidského těla viz. Obr. 10.2.



Obr. 10.2: Znázornění geometrie osoby dle norem [37].

V práci je využito dobře rozpoznatelného popisu a detekce obličeje osob. Pomocí detekce umístění obličeje v obraze a geometrie lidského těla můžeme vycházet při popisu polohy určení barvy oděvu trupu a nohou viz. Obr. 10.3.

Červený čtverec značí oblast detekovaného obličeje pomocí haarcascade, která je natrénován na osobách s čelním pohledem, což naší databázi a práci odpovídá, jelikož všechny zachycené osoby v databázi jsou foceny čelně.

Zelený čtverec na Obr. 10.3 znázorňuje pole ve kterém budeme určovat barvu oblečení trupu osob. Velikost a poloha se odvozuje od velikosti a polohy detekovaného obličeje. Stejně tak i u barvy oděvu nohou, kterou nám značí čtverec žlutý.



Obr. 10.3: Vyznačená pole zájmu na příkladné fotografii.

11 NÁVRH APLIKACE

Celý program byl implementovaný ve vývojovém prostředí Visual Studio Enterprise 2015 v jazyku C++ za pomoci knihovny opencv-3.1 od společnosti Intel. Knihovna obsahuje mnoho funkcí pro práci s obrazem a videem.

Celý program je ve formě konzolové aplikace a je vyvíjen pod operačním systémem Windows.

11.1 Knihovna OpenCV

Celá tato aplikace je založená na knihovně OpenCV verze 3.1. Tato knihovna byla zvolena proto, že je určena právě pro počítačové vidění a zpracování obrazu. Jako součást knihovny jsou například funkce pro detekce a rozpoznávání v obrazovém signále. OpenCv je volně dostupná knihovna pro počítačové vidění (open computer vision library). Knihovna obsahuje velké množství funkcí, které patří pod OpenCV. Nejhlavnější funkce, které z této knihovny využíváme jsou funkce pro zpracování obrazu z fotografie, videa a webové kamery.

11.2 Popis aplikace

Výstupem této diplomové práce je konzolová aplikace "Program pro detekci barvy vrchní nebo spodní části oblečení z videa nebo fotky".

Můžeme si zde vybrat a pracovat se čtyřmi rozdílnými vstupy, fotografií, složkou s fotografiemi, videem nebo vestavěnou webovou kamerou. Vybírat můžeme tyto typy zdroje: image|images|video|webcam (*parametr o*), poté co si vybereme s kterým zdrojem chceme dále pracovat, musíme mu určit jako vstupní proměnnou (*parametr d*). Název vstupního souboru (video, fotky) a cesty k nim (obrazky musí být v jedné složce).

Jako další parametr si musíme určit jestli chceme jako výstup obrázky, video nebo csv soubor v závislosti na vstupní proměnné, poté přejdeme k načtení xml souboru s Haarovými kaskádovými klasifikátory pro detekci obličejů (*parametr a*) používáme `haarcascade_frontalface_alt.xml` která měla nejlepší úspěšnost při detekci pro naši databázi proto je použita naši práci. *Parametr w* slouží pro zadání šířky výstupního formátu.

Parametr P nám udává procentuální hodnotu nastavené hranice pro barvy, které když budou v menším zastoupení než je zadaná hodnota percentilu, budou vynechány. Hodnotu máme nastavenou na 10%.

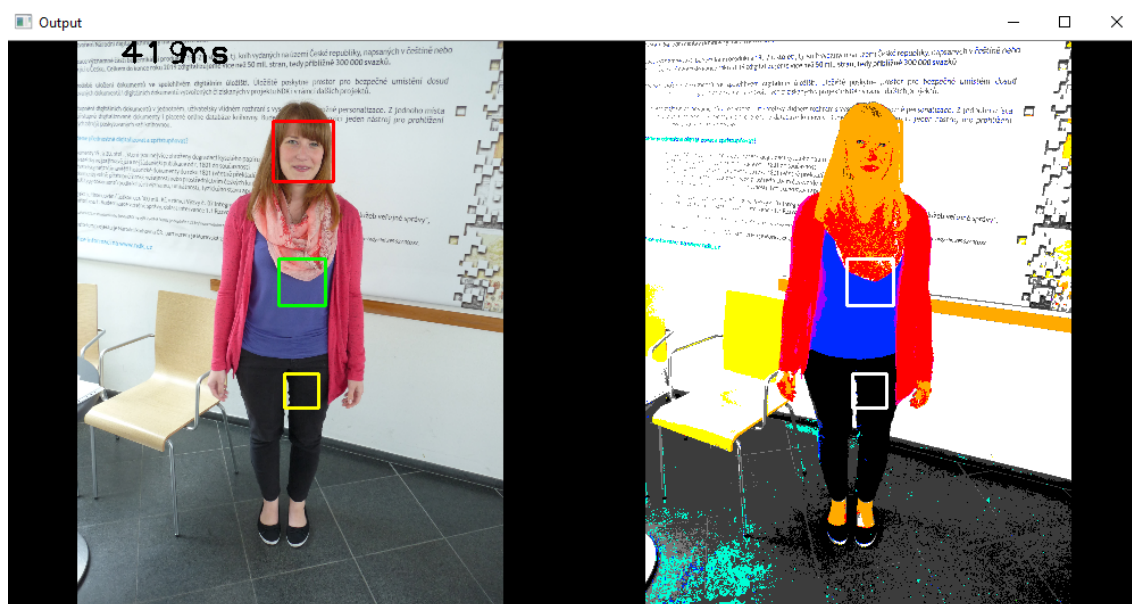
Parametr D udává vzdálenost od vrchní hrany detekovaného obličeje po horní hranu čtverce trupu. Hodnota je určena jako násobek výšky detekované oblasti obličeje z jehož detekce vycházíme při popisu oděvu.

Parametr X udává Šířku testované oblasti trupu v závislosti na šířce detekovaného obličeje (*Parametr Y* udává výšku) . *Parametr L* udává vzdálenost od vrchní hrany detekovaného obličeje po horní hranu čtverce noh. Hodnota je určena jako násobek výšky detekované oblasti obličeje. *Parametr x* udává šířku a *Parametr y* udává výšku testované oblasti noh v závislosti na šířce a výšce detekovaného obličeje. *Parametr g* nám odstraní z výstupu přepočítaný obrázek v pravé části a *parametr v* nám zobrazí informace o zastoupených barvách.

Program vychází z databáze fotografií, které jsou focené čelně a proto využívá možnosti detekce obličeje z jehož polohy v obraze vychází viz obr. 11.1. Červené ohraničení ve fotografii nám značí polohu obličeje na fotografii, z této polohy jsou odvozena pole zelené pro označení oděvu trupu a pole žluté pro nohy. Při vyhodnocování barev se zaměříme na tyto pole viz obr. 11.2 a obr. 11.3.



Obr. 11.1: Fotografie použita v programu s vyznačenými poli zájmu zpracování.



Obr. 11.2: Výstupní obraz programu pro fotografii P1070458.

```
C:\dp>C:\dp\colordetect\x64\Release\colordetect.exe -a C:\dp\lib\opencv_3_1\source\
070440.JPG
Nacitam soubor: C:\dp\data\ fotky\P1070440.JPG
Rozmery obrazku y:500x500
Nalezen 1 oblicej(u) 419ms
Horni cast tela zacina v239, 191 do 280, 232
Barva horni casti: Red (36% procent).
Obrazek/snimek P1070440.JPG trup Red procent 36
Barva horni casti: Blue (45% procent).
Obrazek/snimek P1070440.JPG trup Blue procent 45
Spodni cast tela zacina v244, 292 do 274, 322
Barva spodni casti: White (28% procent).
Obrazek/snimek P1070440.JPG nohy White procent 28
Barva spodni casti: Red (71% procent).
Obrazek/snimek P1070440.JPG nohy Red procent 71
```

Obr. 11.3: Výstupní hodnoty vypsané v konzoli programu pro fotografii P1070458.

12 INSTALACE A KONFIGURACE

Aplikace je napsaná pro platformu Windows. Pro spuštění zkompilevané aplikace je potřeba mít staženou knihovnu openCV ve verzi 3.1 a v prostředí windows nastaveny cesty k této knihovně.

Do systémového prostředí (v systému v upřesnění nastavení systému) definujte systémovou proměnnou `OPENCV_DIR` `C:\dp\lib\opencv_3_1\build\x64\vc14`.

K uživatelské proměnné `Path` je nutné přidat `C:\dp\lib\opencv_3_1\build\x64\vc14\bin`, aby se změny projevily, restartujte počítač.

Pro zkompilování aplikace ze zdrojových kódu je na přiloženém CD projekt v programu Visual Studio Enterprise 2015 s platformou pro C++. Pro správnou funkčnost je doporučeno ponechat strukturu složek, které by měli být na lokálním disku ve složce `C:\dp\`.

13 ZÁVĚR

Cílem práce bylo nastudování možnosti automatické detekce oděvu osob vyskytujících se v obrazovém signálu.

Jako podklad mi v diplomové práci sloužila databáze fotografií, kterou jsem vytvořila ve své bakalářské práci, z těchto fotografií jsem vycházela při návrhu a zpracování programu. Celkem jde o 338 fotografií osob, které jsou všechny foceny čelně a je jim vidět do obličeje. Vždy je zachycena celá postava osoby.

Po nastudování o otestování několika způsobů možností a metod detekce byla vybrána metoda detekce pomocí Haar klasifikátorů, ostatní metody, které jsou založené na machine learningu v knihovně OpenCV nebylo v této práci možné použít, protože vyžadují velkou databázi trénovacích vzorků v řádech tisíců až deseti tisíců.

Moje metoda je založená na detekci lidského obličeje za pomoci HAAR kaskád, která se ukázala jako velmi přesná a spolehlivá. Na naší databázi dosahovala úspěšnost detekce obličeje 98%. Pokud jsme schopni detekovat obličej můžeme využít tuto pozici v obrazu. Na základě geometrie lidského těla a pozice obličeje můžeme stanovit přibližné pozice jednotlivých částí postavy. Já jsem se zaměřila na určení barvy vrchní a spodní částí oděvů osob. Podle detekované velikosti a polohy čtverce obličeje odvozují velikosti a polohy oblastí vrchní a spodní části osob, ve kterých následně detekují barvu oděvů.

Z popisu metody vyplývá, že má metoda nebude spolehlivě detekovat ty obrazy, na kterých není obličej nebo je neúplný či zakrytý. Dále pokud je postava na fotografii natočená, zakrytá nebo pokud není ve vzpřímené poloze. V takovém případě metoda určí špatné hodnoty.

Metodu jsem implementovala v prostředí C++ s knihovnou OpenCV. Vytvořila jsem konzolovou aplikaci, která je schopná detekovat jednotlivé části postavy a vrátit jednotlivé procentuální zastoupení barev, které jsou převedeny podle mnou definovaných podmínek pro jednotlivé barvy.

Metodu jsem testovala na výše uvedené databázi. Vyhodnocení probíhalo na základě hodnot a znaků přiřazených v databázi. Úspěšnost detekce byla pro vrchní část oblečení 83% a pro spodní část nám vyšlo 67%. Důvodem horší úspěšnosti detekce barvy pro spodní část je fakt, že je tato oblast menší a se složitějším zastoupením v obraze.

LITERATURA

- [1] Mlýnková, B. *Databáze nebiometrických a sekundárních biometrických znaků osob* Brno: Vysoké učení technické v Brně. Fakulta elektrotechniky a komunikačních technologií. 2015.
- [2] Rak, R., Matyáš, V., Říha Z. a kolektiv: *Biometrie a identita člověka ve forenzních a komerčních aplikacích* Grada Publishing, a.s., 2008. ISBN 978–80–247–2365–5.
- [3] Dražanský, M., Orság, F. a kolektiv: *Biometrie*. Computer Press a. s., Brno 2011. ISBN 978–80–254–8979–6.
- [4] Herschel, W. J.: *The Origin of finger-Printing*. Humphrey Milford, 1916.
- [5] Galton, F. Hereditary Talent and Character.: *Macmillan's Magazine*. 1865, ročník 12, číslo 1865, s.157-166.
- [6] Krishan, K.: *Anthropometry in forensic medicine and forensic science–‘Forensic Anthropometry’*. Internet, J. Forensic Sci., 2 (2007)
- [7] Galton, F.: *Fingerprints*. Londýn? Macmillan and Co., 1892.
- [8] Straus, J. a. k.: *Kriminalistická daktyloskopie*. Praha: Policejní akademie ČR, 2005. ISBN 80–7251–192–0.
- [9] Hauptvogel, K. H., Ritzeschke, M.: *Biometrie um die Jahrhundertwende*. 2004.
- [10] Komarinski, P.: *Automated Fingerprint Identification System (AFIS)*. :Academic Press, 2004. ISBN 978–01–241–8351–3.
- [11] Jain, A. K., Flynn, P., Ross, A.: *Handbook of Biometrics*, Springer, 2008, ISBN 978-0-387-71041-9.
- [12] S. Seung, H. Sompolinsky, and N. Tishby,: *Statistical mechanics of learning from examples* Phys. Rev. A, vol. 45, pp.6056–6091, 1992.
- [13] Ng, Ch. B., Tay, Y. H., Goi, B. M.: *Recognizing human gender in computer vision: a survey*. Proceedings of 12th Pacific Rim International Conference on Artificial Intelligence. pp. 335–346. Springer, 2012.
- [14] Salzmann, Z.: *Český lid – Nový pohled na fyzickou romanitost lidstva*. Institute of Ethnology, Czech Academy of Sciences, 2003.
- [15] Yacoob, Y., Davis, L.: *Detection, analysis and matching of hair*. The tenth IEEE International Conference on Computer Vision 1 741–748, 2005.

- [16] Bossard, L., Dantone, M., Leistner, C., Wengert, C., Quack, T., Van Gool, L.: *Apparel Classification with Style*. in ACCV, 2013.
- [17] Chen, H., Gallagher, A., Girod, B.: *Describing clothing by semantic attributes*. In ECCV, 2012.
- [18] Rannells, J.: *PNG: A Fact Book on Modern Papua New Guinea*. Oxford University Press, 1995.
- [19] Pennebaker, W. B., Mitchell, J. L.: *JPEG Still Image Data Compression Standard*. Nostrand Reinhold, New York, 1993.
- [20] Kreibich, J. A.: *Using SQLite*. O'Reilly Media, Inc., 2010.
- [21] Owens, M.: *Embedding an SQL database with SQLite*. Linux Journal, 2., 2003.
- [22] Wikipedie: *Neuronová síť* / *Wikipedie, Svobodná encyklopedie* [online]. 2017 [cit.2017-04-18]. Dostupné na URL <http://cs.wikipedia.org/wiki/Neuronová_síť>
- [23] Baluja, Shumeet, and Henry A. Rowley. *Boosting sex identification performance*. International Journal of computer vision 71.1 (2007): 111-119.
- [24] Lienhart R., Maydt J.: An extended set of Haar-like features for rapid object detection. In: *Proceedings of International Conference on Image Processing*. Rochester, New York, 2002. pp. 900–903.
- [25] *Cascade Classification* 2017 [cit.2017-04-18]. <http://docs.opencv.org/2.4.13.2/modules/objdetect/doc/cascade_classification.html>
- [26] Paul Viola, Michael Jones.: *Rapid Object Detection using a Boosted Cascade of Simple Features*, cvpr, vol. 1, pp.511, 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'01) - Volume 1, 2001
- [27] Tom Fawcett : *ROC Analysis in Pattern Recognition*. Pattern Recognition Letters. Vol. 27, 2006. Pages 882-891
- [28] WILSON, P., FERNANDEZ, J., *Facial feature detection using Haar classifiers.*, Consortium for Computing Sciences in Colleges, 2006.
- [29] Stratton J. A. "Electromagnetic theory" *McGraw-Hill, New York 1949*. Český překlad "Teorie elektromagnetického pole", SNTL, Praha 1961.
- [30] Autor: Quark67(Modified color by Monami)–Image:Synthese+.svg, *CC BY-SA 3.0*, <<https://commons.wikimedia.org/w/index.php?curid=4798169>>

- [31] *CC BY-SA 3.0*, <<https://commons.wikimedia.org/w/index.php?curid=113383>>
- [32] JANOUGH, Michal.: *Zpracování obrazu jednočipovým mikroprocesorem*,. Diplomová práce. Praha: ČVUT Fakulta elektrotechnická, 2008, 72 s.
- [33] Autor: DaBler-(myself), *Volné dílo*, <<https://commons.wikimedia.org/w/index.php?curid=2869337>>
- [34] Autor: SharkD-Vlastní dílo, *CC BY-SA 3.0*, <<https://commons.wikimedia.org/w/index.php?curid=8421538>>
- [35] Süsstrunk, Sabine, Robert Buckley, and Steve Swen. "Standard RGB color spaces." *Color and Imaging Conference*. Vol. 1999. No. 1. Society for Imaging Science and Technology, 1999.
- [36] Yuan, Shuai, YingLi Tian, and Aries Ardit. "Clothing matching for visually impaired persons." *Technology and disability*. 23.2 (2011): 75-85.
- [37] *NAŘÍZENÍ VLÁDY 361/2007 Sb.* ze dne 12. prosince 2007

SEZNAM PŘÍLOH

A Obsah přiloženého CD

47

A OBSAH PŘILOŽENÉHO CD

Nezapomeňte uvést, co čtenář najde na přiloženém médiu. Je vhodné okomentovat obsah každého adresáře, specifikovat, který soubor obsahuje důležitá nastavení, který soubor je určen ke spuštění atd. Také je dobře napsat, v jaké verzi software byl kód testován (např. Matlab 2010b).