VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ
ÚSTAV INFORMAČNÍCH SYSTÉMŮ

FACULTY OF INFORMATION TECHNOLOGY
DEPARTMENT OF INFORMATION SYSTEMS

IDENTIFIKACE OBLIČEJE

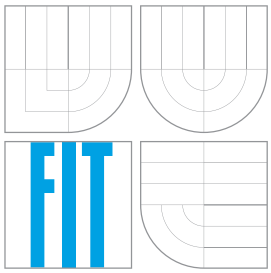
DIPLOMOVÁ PRÁCE
MASTER'S THESIS

AUTOR PRÁCE
AUTHOR

Bc. OTO MACENAUER

BRNO 2010

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY



FAKULTA INFORMAČNÍCH TECHNOLOGIÍ
ÚSTAV INFORMAČNÍCH SYSTÉMŮ

FACULTY OF INFORMATION TECHNOLOGY
DEPARTMENT OF INFORMATION SYSTEMS

IDENTIFIKACE OBLIČEJE

FACE RECOGNITION

DIPLOMOVÁ PRÁCE

MASTER'S THESIS

AUTOR PRÁCE — Bc. OTO MACENAUER

AUTHOR

VEDOUCÍ PRÁCE — Ing. PETR CHMELAŘ

SUPERVISOR

BRNO 2010

Abstrakt

Tato práce uvádí čtenáře do problematiky rozpoznání obličeje. Jsou zde zmíněny a rozřazeny různé metody sloužící k identifikaci obličeje. Hlavní důraz práce je kladen na problémy současných řešení problému rozpoznání obličeje a následně je vyvinuta snaha tyto problémy vyřešit. Vyřešení těchto nedostatků technologie je nutností pro masivní rozšíření implementací identifikace obličeje. Dále se práce zaměřuje na implementaci zvolených metod, což jsou lineární diskriminační analýza a analýza hlavních komponent. Tyto metody jsou vzájemně porovnány a výsledky jsou poskytnuty v závěru práce.

Abstract

This document introduces the reader to area of face recognition. Miscellaneous methods are mentioned and categorized to be able to understand the process of face recognition. Main focus of this document is on issues of current face recognition and possibilities do solve these inconveniences in order to be able to massively spread face recognition. The second part of this work is focused on implementation of selected methods, which are Linear Discriminant Analysis and Principal Component Analysis. Those methods are compared to each other and results are given at the end of work.

Klíčová slova

Identifikace obličeje, analýza hlavních komponent, lineární diskriminační analýza, rozdělení metod identifikace obličeje.

Keywords

Face recognition, Principal Component Analysis, Linear Discriminant Analysis, categorization of methods of face recognition.

Citace

Oto Macenauer: Identifikace obličeje, diplomová práce, Brno, FIT VUT v Brně, 2010

Identifikace obličeje

Prohlášení

Prohlašuji, že jsem tuto diplomovou práci vypracoval samostatně pod vedením pana Ing. Petra Chmelaře

. . . . . .

Oto Macenauer

26. května 2010

Poděkování

Tímto bych rád poděkoval panu Ing. Petru Chmelařovi za vedení diplomové práce a rodině za podporu během studia.

© Oto Macenauer, 2010.

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

Obsah

1 Úvod **2**

2 Popis identifikace obličeje **4**

2.1 Biometrie obličeje . . . 4

2.2 Aplikace . . . 6

2.2.1 Oblasti aplikací . . . 6

2.3 Budoucí rozvoj . . . 6

3 Proces identifikace obličeje **8**

3.1 Pre-processing . . . 8

3.1.1 Transformace . . . 9

3.1.2 Doplnění trénovací množiny dat . . . 10

3.1.3 Normalizace osvětlení . . . 10

3.2 Detekce částí obličeje . . . 10

3.2.1 Detekce hlavy . . . 12

3.2.2 Detekce hranice obličeje . . . 12

3.2.3 Detekce očí . . . 12

3.2.4 Detekce úst . . . 13

3.3 Extrakce rysů . . . 14

3.4 Učení . . . 15

3.5 Rozpoznání . . . 16

3.5.1 Vzdálenosti . . . 16

3.6 Ukládání dat pro identifikaci . . . 18

3.6.1 Indexování . . . 18

3.6.2 Hashování . . . 19

4 Metody identifikace **20**

4.1 Algebraické předpoklady identifikace obličeje . . . 20

4.1.1 Vlastní hodnoty, vektory a prostor . . . 21

4.1.2 Kovariance . . . 22

4.1.3 Singulární rozklad matice . . . 22

4.2 Analýza hlavních komponent . . . 23

4.2.1 Použití pro identifikaci obličeje . . . 23
4.2.2 Základní principy metody . . . 23
4.2.3 Popis metody . . . 24
4.2.4 Spojení PCA s jinými metodami . . . 26
4.3 Lineární diskriminační analýza . . . 26
4.3.1 Popis metody . . . 27
4.3.2 Redukce dimenzionality vstupních dat . . . 28
4.3.3 Výhody a nevýhody . . . 28
4.4 Analýza nezávislých komponent . . . 28
4.4.1 Formální definice . . . 28
4.5 Skrytý Markovův model . . . 30
4.5.1 Formální popis . . . 30
4.5.2 Užití pro rozpoznání obličeje . . . 31
4.5.3 Aplikace . . . 32
4.6 Elastický srovnávací diagram . . . 32
4.6.1 Preprocessing pomocí Gáborových filtrů . . . 32
4.6.2 Reprezentace obličeje . . . 32
4.7 Podpůrné vektory . . . 32
4.7.1 Binární klasifikace . . . 33
4.7.2 Klasifikace více tříd . . . 33
4.7.3 Užití pro identifikaci obličeje . . . 34
4.8 Propojení jednotlivých metod . . . 34

5 Implementace zvolených metod **35**
5.1 Použité zdroje a knihovny . . . 35
5.1.1 OpenCV . . . 35
5.1.2 Vstupní data . . . 37
5.2 Řešení . . . 38
5.2.1 Implementace Analýzy hlavních komponent . . . 38
5.2.2 Implementace Lineární diskriminační analýzy . . . 46
5.2.3 Porovnání metod . . . 49

6 Závěr **52**

Literatura **56**

A Obsah CD **57**
A.1 Aplikace . . . 57
A.2 Písemná zpráva . . . 57
A.3 Programová dokumentace . . . 57
A.4 Použité databáze obličejů . . . 57

Kapitola 1

Úvod

V poslední době se čím dál více setkáváme v různých médiích s tématy, která zmiňují rozpoznání obličeje. Ať už to jsou otázky bezpečnosti, které zahrnují snímání obličejů lidí cestujících do zahraničí, obličejů fanoušků fotbalových zápasů, kvůli kontrole lidí, kteří páchají často trestné činy, nebo užití v hazardním průmyslu při rozpoznávání obličejů hráčů, či dokonce použití v inteligentních hračkách, které dokážou rozpoznat svého majitele a reagovat na jeho náladu. Oblastí využití je nespočet a jen to dokazuje, že tato technologie má v budoucnosti dveře otevřené.

Dalo by se říct, že vývoj rozpoznání obličeje začal již někdy v 19. století prací Alphonse Bertillona na rozpoznávání osob pomocí jejich antropometrických rysů. Od té doby metody znatelně pokročily, ale stále platí, že užití v kriminalistice a bezpečnosti je jedno z hlavních využití této technologie. Co se však změnilo, je způsob jakým jsou metody prováděny. Zdlouhavé měření pomocí pravítek a jiných nástrojů vystřídaly metody založené na analýze fotografie, videa, 3D modelu či různých jiných elektronických interpretacích. Výpočetní výkon databází výrazně zrychlil vyhledávání příslušných osob a je zřejmé, že vývoj stále ještě není u kon ce. Nové objevy v informačních technologiích posouvají dále i oblast rozpoznání obličeje.

Jak již bylo zmíněno, vývoj rozpoznání započal na základě znalostí z anatomie, což však není jediná oblast, která se prolíná s metodologií rozpoznání obličeje. Tento obor, stejně jako jeho součásti (extrakce rysů, detekce v obraze) čerpá hodně z dalších vědních oborů tedy nejen z biologie a poznání přírody všeobecně, ale například i z psychologie. Co se týče psychologie, řada metod je založena na snaze napodobit práci lidské mysli, takže určitě nepřekvapí, že obor rozpoznání obličeje úzce souvisí s umělou inteligencí, resp. strojovým učením.

I když existuje celá řada postupů, stále ještě je spousta oblastí, které lze prozkoumávat a zlepšovat. Hlavní problém je totiž v odlišnosti lidí, je tedy nesmírně obtížné postihnout všechny případy. Vezměme si jako příklad barvu kůže. Různé algoritmy jsou založeny na detekci obličeje podle předpokladu, že kůže má určitou barvu, která je odlišná od okolního prostředí, když ale vezmeme v úvahu, že v jiných částech světa mají odlišnou barvu kůže zjistíme, že detekce podle tohoto rysu je problematická. Důležitým předpokladem tedy je

nejen tyto rysy hledat, ale normalizovat je tak, aby byly vzájemně porovnatelné.

Identifikace obličeje je tedy oblast zpracování obrazu, která má jistě slibnou budoucnost v různých odvětvích lidské činnosti. Jejímu masovému rozšíření však brání řada problémů, které omezují její použití. Tato práce se snaží tyto problémy charakterizovat a pokouší se poukázat na vhodná řešení. Předmětem této práce je také nastínění problematiky rozpoznání lidí podle obličeje, popis různých metod řešení této problematiky.

V kapitole 2, Popis identifikace obličeje, se popisuje podstata tohoto postupu. V úvodu kapitoly je stručný úvod do problematiky, dále v následující části - Biometrie obličeje - je vysvětlena vhodnost obličeje pro metody identifikace osob a také popsány části obličeje vzhledem k jejich aplikaci na identifikaci. V závěrečné části se pojednává o aplikacích identifikace obličeje v běžném životě, rozdělení odvětví, kde najde uplatnění a nakonec také budoucnost této technologie.

Kapitola č. 3 popisuje proces identifikace osob pomocí obličeje, který je rozdělen na preprocessing (normalizaci vstupních dat), detekci jednotlivých částí obličeje, extrakci rysů z nalezených oblastí, vysvětlení pojmu strojového učení, identifikaci a v závěru se kapitola zaměřuje na orientaci ve velkém objemu dat pomocí indexování a hashování.

Kapitola 4 je zaměřena na jednotlivé metody rozpoznání obličeje. V úvodu kapitoly je popsáno možné dělení metod a následný popis jednotlivých metod se zaměřením na výhody a nevýhody, případně situace, kdy je vhodné jednotlivé metody použít. Konkrétně jde o metody: Analýza hlavních komponent, Lineární diskriminační analýza, Analýza nezávislých komponent, Skrytý Markovův model, Elastický srovnávací diagram a Podpůrné vektory.

V poslední, 5. kapitole je popsána implementace dvou zvolených metod používaných při identifikaci obličeje a to Analýza hlavních komponent a Lineární diskriminační analýza. Je zde dále popsána knihovna OpenCV, která je jedním z nejpoužívanějších nástrojů v oblasti počítačových vidění a usnadňuje tvorbu aplikací. Dále jsou zmíněny existující zdroje dat pro testování funkce metod. U implementovaných metod je kladen důraz na zkoumání vlivu různých parametrů, jako jsou světelné podmínky, při kterých byl snímek pořizován, orientace obličeje ke snímacímu zařízení a různé další, které mohou pozitivně či negativně ovlivnit proces rozpoznání obličeje. Posléze jsou obě metody porovnány.

Kapitola 2

Popis identifikace obličeje

Identifikace obličeje je proces rozpoznání zadaného snímku na základě podobnosti a jeho přiřazení určité osobě. Tento postup lze algoritmizovat pomocí informačních technologií, kdy se automaticky pomocí přesně daných postupů rozpoznává nebo verifikuje obličej oproti digitalizovanému snímku nebo sérii snímků z videa [43]. Aplikace sloužící k tomuto účelu získá snímek, změří charakteristiky jako například vzdálenost mezi očima, délku nosu, úhel čelisti a vytvoří unikátní šablonu. S pomocí těchto šablon aplikace navzájem mezi sebou porovnává snímky a stanovuje ohodnocení, které určuje, jak jsou si jednotlivé snímky podobné. Nadále se bude mluvit pouze o automatických postupech identifikace obličeje.

Pokud nemáme ideálně zadaná vstupní data, skládá se proces jako takový většinou ze dvou částí a to detekce obličeje a jeho následná verifikace či identifikace [24]. Detekce obličeje slouží k určení oblasti, kde se obličej či skupina obličejů v obraze nachází. K tomu slouží různé metody jako například klasifikace pomocí šablon, klasifikace pomocí neuronových sítí, shluková analýza, rozpoznání obličeje podle barvy kůže, detekce obličeje z pohybu a další. Tyto algoritmy jsou však závislé na podmínkách, při kterých byl obraz pořizován jako např. světelné podmínky, orientace obličeje oproti objektivu kamery, pozadí apod., které mohou ovlivnit spolehlivost výsledku.

Výsledkem předchozí části je oblast, na kterou může být aplikován další krok a to identifikace či verifikace. Rozdíl mezi těmito dvěma možnostmi je v množství porovnávaných obrazů. Při verifikaci porovnáváme pouze dva obrazy, tedy řešíme, zda obličeje na dvou daných obrazech jsou shodné. Naproti tomu při identifikaci hledáme v databázi obrazů obraz, který by odpovídal námi danému obrazu, z toho vyplývá, že je potřeba rozsáhlý soubor dat, aby bylo možné identifikovat zadaný obraz jako obličej specifické osoby. Kromě množství materiálu jsou přístupy k identifikaci a verifikaci podobné.

2.1 Biometrie obličeje

Předpokladem úspěšné identifikace lidského obličeje je jeho jedinečnost. Počátky aplikace biometrie v identifikace osob sahají do 19. století, kdy Alphonse Bertillon, vedoucí oddělení

identifikace zločinců pařížské policie použil metriky těla (včetně obličeje) k identifikaci.

Existuje několik kritérií, které určují, zda je zvolená metoda vhodná k identifikaci osob [24]. Nejdůležitější z nich jsou:

Univerzálnost - každá osoba tyto vlastnosti má.

Rozlišitelnost - jakékoliv dvě osoby jsou dostatečně rozlišitelné pomocí těchto charakteristik.

Neměnnost - vlastnosti jsou dostatečně neměnné (s přihlédnutím na kritéria identifikace) během určitého období.

Možnost sběru informací - charakteristiky lze získat bez větších potíží a mohou být uchovávány.

V případě identifikace osob pomocí obličeje jsou splněny všechny body. Uvedená kritéria jsou u všech osob zaručena, vzhledem k tomu, že každá osoba má obličej, obličeje se liší, během života se obličej sice mění, ale tyto změny nezpůsobují výjimky v metodě. Co se týče možnosti sběru informací, lze ji uskutečnit pomocí běžně dostupných zařízení pro záznam obrazu (fotoaparáty, kamery), takže i tento bod je splněn.

Některé části obličeje jsou pro identifikaci důležitější než jiné. Podle důležitosti by se daly seřadit jako: oči, nos, ústa a za nimi další části obličeje. Se zaměřením na tyto oblasti se provádí určitá měření. Významnými metrikami z hlediska identifikace jsou například [8] :

- vzdálenost mezi očima
- délka nosu
- hloubka očních důlků
- tvar lícních kostí
- délka a tvar čelisti

Oproti ostatním biometrickým metodám identifikace má identifikace pomocí snímku obličeje výhodu v tom, že není potřeba přímý kontakt uživatele se snímacím zařízením, dokonce je možné vyhodnocovat shodu bez jeho vědomí.

Další nespornou výhodou identifikace pomocí obličeje je srozumitelnost metody obyčejnému člověku. Kdokoliv může porovnat snímky, které metoda vyhodnotila jako totožné obličeje, což lze použít například u identifikace osob při vstupu do objektů, kdy má ochranka možnost potvrdit správnost výsledku, což by například u otisků prstů bylo velice obtížné, až nemožné.

2.2 Aplikace

Jak již bylo zmíněno v kapitole Biometrie obličeje, byl původně spojen vývoj identifikace obličeje s organizacemi pro vymahatelnost práva, z identifikace osob se později vyvinul systém pro náhodné snímání osob v davu. Další aplikací byla snaha předejít podvodům během voleb. Jednou z posledních aplikací v praxi je systém US-VISIT (United States Visitor and Immigrant Status Indicator Technology), který je zaměřen na návštěvníky Spojených států amerických. V tomto případě je použito systémů identifikace obličeje k verifikaci při příletu na letišti, zda je člověk, který je kontrolován ta samá osoba, které bylo vydáno vízum [26]. Ohledně dalších států je typickým případem využití metod identifikace obličeje v rozsáhlém kamerovém systému města Londýna [2].

Kromě již zmíněných aplikací se systémy identifikace obličeje stále více prosazují do běžného života, díky zefektivňování a zlevňování technologie. Běžně se nyní implementují aplikace pro identifikaci obličejů do některých notebooků. Dalšími aplikacemi jsou např. identifikace zákazníků při opakovaném nákupu či podpora bezpečnosti na letištích.

2.2.1 Oblasti aplikací

Nejběžnější oblasti aplikace identifikace obličeje jsou znázorněny v tabulce 2.1 [43].

Oblast	**Aplikace**
Bezpečnost	Bezpečnost počítačů a fyzický přístup na základě kontroly přístupu
Vláda	Kriminalistika/Mapování terorismu
	Sledování
Pasy/Občanské průkazy	Identifikace
	Detekce ilegální imigrace
Vozidla	Upozorňování řidiče na nebezpečí založené na pohybu očních víček
Kasína	Rozpoznání podezřelých gamblerů-VIP
Hračky	Inteligentní roboti

Tabulka 2.1: Oblasti aplikace identifikace obličeje

2.3 Budoucí rozvoj

Budoucnost aplikací identifikace obličeje lze spatřovat v prodeji zboží. Při zlepšení technologie bude možné identifikovat větší počet lidí v davu s větší přesností. Toho lze využít v různě, například ke tvorbě statistik nákupů častých zákazníků, identifikaci člověka při

nákupu, což dá personálu možnost připravit si informace o něm a zaměřit se na určitou věc. Dále se nabízí implementace do handheldů pro případy verifikace uživatele.

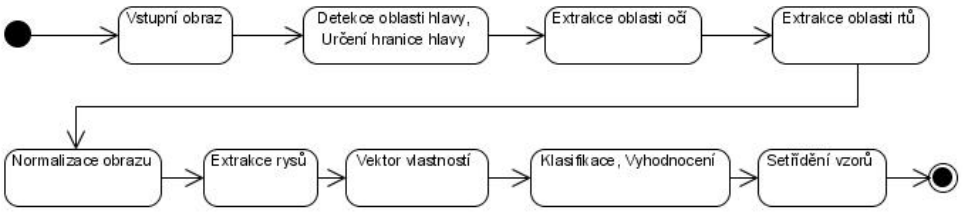
Co se týče technologie, nabízí se možnost použití 3D algoritmů, které eliminují problém nasvícení scény při klasické identifikaci obličeje. Tento přístup však vyžaduje vysoký výpočetní výkon a je rovněž náročný na rychlost databáze [7].

Kapitola 3

Proces identifikace obličeje

Proces identifikace obličeje lze rozdělit na více částí: pre-processing, který slouží k předpřipravení snímku pro detekci, detekci částí obličeje, kdy se pomocí algoritmu prochází obraz a hledají důležité oblasti pro extrakci rysů, která je následujícím krokem, následná identifikace a zpracování dat, tedy post-processing [21].

Celý proces lze vyjádřit následujícím obrázkem:



Obrázek 3.1: Diagram procesu rozpoznání obličeje [21].

3.1 Pre-processing

Preprocesing (předpřipravení) je proces, kdy se normalizuje zadaný snímek, aby byl co nejvhodnější pro účely identifikace obličeje.

Mnoho z problémů, které se objevují při procesu identifikace obličeje se dá odstranit ve fázi preprocessingu a tím významně ovlivnit spolehlivost celého procesu.

Preprocessing lze dělit na automatický a manuální. Plně automatické systémy identifikace obličeje jsou schopny algoritmicky řešit problémy zmíněné níže, naopak manuální preprocessing předpokládá zásah uživatele. Jedná se o pracnější metodu, ale dosahuje lepších výsledků vzhledem k tomu, že se každý snímek posuzuje zvlášť.

Detekce obličeje

Většina metod rozpoznání obličeje přepokládá jako vstupní data určitou část obličeje. Pokud by byly metodě předány data, která by se lišila oblastí, kde se obličej rozpoznává, vedlo by

to k výraznému zhoršení výsledků [29]. Tímto problémem trpí převážně holistické metody rozpoznání obličeje, které ukládají informace o celkovém obrázku, ale neuvažují možnost, že daný obličej je transformovaný.

Vysoké úspěšnosti detekce obličeje lze dosáhnout trénováním klasifikátoru na rozsáhlé množině trénovací množiny dat. Složitějším případem je detekce obličejů, které jsou sejmuté pod určitým úhlem. Za tímto účelem se používá trénovaní na množině snímků pořízených z více pohledů, což zvyšuje úspěšnost detekce obličejů, které jsou natočeny pod úhlem větším než 35° [18].

Příkladem normalizovaných oblastí pro rozpoznání obličeje mohou být databáze obličejů používané k testování algoritmů identifikace. Podrobněji jsou jednotlivé databáze popsány v části 5.1.2. Detekce hlavy je zmíněna v 3.2.1. Je ale vhodné, aby nebyla zachována příliš velká plocha pozadí, která nenese informace důležité k rozpoznání (naopak rozdílnost pozadí při stejném obličeji může způsobovat nesprávnou kategorizaci obličeje), a aby nebyly zanedbány některé důležité rysy obličeje.

Jak je uvedeno ve standardu MPEG-7 [31], normalizovaná oblast obličeje je rozdělena mřížkou o 56 řádcích a 46 sloupcích. Dále je zmíněna poloha očí a to 24. řádek a 16. sloupec pro levé oko a 31. sloupec pro pravé oko. Pomocí této znalosti lze určit oblast obličeje tak, že po detekci očí se změří jejich vzájemná vzdálenost a z ní se určí poloha jednotlivých hranic obličeje. Pokud je vzdálenost očí označena jako $dist_{eye}$, pak:

$$W = \frac{dist_{eye}}{15} \cdot 56 \text{ a} \tag{3.1}$$

$$H = \frac{dist_{eye}}{15} \cdot 46 \tag{3.2}$$

kde W je šířka oblasti a H je výška oblasti. Je potřeba si však uvědomit, že oči nemusí být zobrazeny vždy ve stejné rovině, proto je potřeba buď počítat se zkosením, nebo nanášet vzdálenosti od středu spojnice očí.

3.1.1 Transformace

Transformace je jeden ze způsobů, jak řešit změnu orientace obličeje. Problém je vhodné řešit pomocí učení na více snímcích obličeje z různých póz, ale pokud nelze tyto snímky získat, musí se pracovat pouze se zadanými daty. V [4] je vysvětlen způsob, kdy se ze zadaného snímku vytvoří množina virtuálních snímků, podle prototypů, které označují jednotlivá postavení obličeje vůči kameře. Pro větší úhly natočení obličeje je však tato metoda nevhodná a je výhodnější použít některou z metod založených na 3D reprezentaci obličeje [43].

Další problém, který se transformacemi řeší, je změna měřítka. Některé metody se nedokážou vypořádat s rozdílnou velikostí vstupního obrazu, proto je potřeba změnit velikost na velikost danou pro všechny vzorky a případně ořezat, pokud je výška a šířka v jiném poměru.

3.1.2 Doplnění trénovací množiny dat

Toto řešení je spojeno s více problémy. Je užitečné převážně pro holistické metody rozpoznání obličeje, protože zpřesňuje rozpoznání obličeje při rozdílných orientacích obličeje či různých světelných podmínkách.

Nejjednodušším případem doplnění trénovací množiny je použití zrcadlově otočených snímků. Toto vychází ze symetrie lidského obličeje. Ve většině případů je levá polovina velmi podobná pravé polovině a tak se tímto krokem dá napodobit orientace obličeje v opačném úhlu, ale také nasvícení obličeje z opačné strany. Tento případ však postrádá smysl, pokud je trénovací snímek pořízen v úhlu 0° oproti kameře při osvětlení zepředu.

Další možností doplnění množiny snímky, které byly spočítány ze vstupních dat. Například doplnění snímků, které mají různé osvětlení (viz 3.1.3) či transformovaný úhel natočení obličeje vůči kameře (viz 3.1.1 nebo [4]).

Tímto způsobem lze rovněž řešit nejednotnou oblast určenou k identifikaci obličeje. Tento problém nastává, pokud není oblast v množině vstupních dat normalizována. Jak již bylo zmíněno v 3.1, i malé rozdíly v zarovnání snímků mohou způsobovat chybu v rozpoznání. Řešením je doplnění množiny o částečně posunuté snímky [29].

3.1.3 Normalizace osvětlení

Problém osvětlení je jeden z největších problémů, který se týká identifikace obličeje. Hlavním důvodem je, že změny v osvětlení jsou často výraznější než rozdíly mezi jednotlivci, což vede k chybné klasifikaci dané osoby. Tato skutečnost byla experimentálně ověřena v [30].

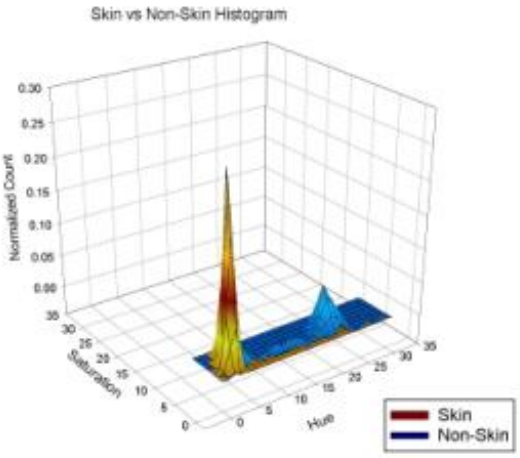
Osvětlení se týkají dva problémy. Jeden je nasvícení obličeje z různých úhlů a druhým je nejednoznačné nastavení jasu a kontrastu u vstupních dat. Druhý problém se řeší většinou normalizací pomocí histogramu, kdy jsou světelné podmínky upraveny tak, aby byly pro všechny snímky podobné.

Problém osvětlení z různých je složitější. Lze řešit stejným způsobem jako problém natočení obličeje, tedy rozšířením učicí množiny dat o snímky nafocené ze stejného úhlu, ale pomocí osvětlení z různých stran. To pomůže klasifikátoru identifikovat obličeje dané třídy při odlišných světelných podmínkách. Jiné řešení, týkající se úpravy kontrastu obrazu, je ukázáno v [17]. Jde o metodu založenou na vnímání obličeje jako symetrický objekt, což má podstatu v lidském vnímání, které je citlivé na symetrii a dokáže vnímat stejné obličeje při různých světelných podmínkách. Tato metoda dosáhnula výrazného zlepšení při rozpoznání obrazu za různého osvětlení.

3.2 Detekce částí obličeje

Detekce částí obličeje spočívá v nalezení oblastí, kde se nachází elementy obličeje důležité pro jeho identifikaci. Tato část není povinná pro všechny přístupy. Některé metody dokážou extrahovat rysy obličeje bez toho, aby znaly přímé umístění částí obličeje.

Některé části obličeje se vyznačují tmavějšími barvami (vlasy, oči, obočí a ústa), zatímco zbytek obličeje charakterizují světlejší odstíny barev. Tyto oblasti lze lokalizovat pomocí histogramu, který má v případě obličeje podobu obrázku 3.2.



Obrázek 3.2: Typický normalizovaný histogram lidského obličeje, zobrazený v barevném prostoru HSV (Hue, Saturation and Values). Červenou a žlutou barvou jsou zobrazeny oblasti obličeje a modrou barvou jsou v histogramu znázorněny oblasti, které do obličeje nepatří [27].

Položením prahu za první maximální hodnotu zleva získáme obraz, který bude obsahovat pouze tmavé regiony obličeje.

Další možností hledání částí obličeje vychází z geometrických vztahů mezi očima, nosem a ústy. Předpokládá, že u většiny tváří je vertikální vzdálenosti mezi očima, nosem a ústy jsou proporční. Pokud označíme horizontální vzdálenost mezi pravým a levým okem jako D, pak vztahy mezi vzdálenostmi ostatních částí obličeje lze definovat jako [37]:

1. Vertikální vzdálenost mezi středem spojnice očí a středem úst je rovna D.
2. Vertikální vzdálenost mezi středem spojnice očí a středem spojnice nosních dírek je rovna $0.6 \cdot D$.
3. Délka úst je rovna D.
4. Délka nosu je rovna $0.8 \cdot D$.
5. Vertikální vzdálenost mezi očima a obočím je $0.4 \cdot D$.

Graficky jsou tyto vzdálenosti zobrazeny na obrázku 3.4. Tento model je vhodný v případě, že obraz po extrakci rysů obsahuje příliš velké shluky [37], avšak vyznačuje se nepřesností. Předpokladem jeho použití je správná detekce očí, z které vychází.

3.2.1 Detekce hlavy

Detekce polohy hlavy je prvním krokem, který je prováděn při detekci obličeje [22]. V tomto kroku je přibližně určen region, kde se nachází hlava subjektu například pomocí prahování histogramu. Tento krok lze zjednodušit tím, že se předpokládá, že je v obrazu jen jeden obličej.

Existuje několik metod detekce obličeje, liší se například množstvím obličejů, které dokáže rozpoznat, použitým barevným spektrem vstupních dat a dalšími vlastnostmi. Právě barva kůže je nejcharakterističtější barvou pro detekci [41]. Pokud bychom neuvažovali jen jeden obličej, lze s přihlédnutím na tento fakt detekovat přibližně regiony k detekci jednotlivých obličejů. Ve většině případů jsou však vstupní data poskytnuta v barevném spektru stupňů šedi. V takovém případě se problém řeší pomocí klasifikace srovnáváním vzorů [22].

3.2.2 Detekce hranice obličeje

Po nalezení samotného obličeje, čehož lze docílit způsobem zmíněným v předchozí části (3.2.1), je možné přikročit k detekci hranice obličeje.

Tvar obličeje je apriori znám [21] a lze jej aproximovat na elipsu. Běžně používaná metoda je Houghova transformace [23], avšak ta nedokáže postihnout případy, kdy se tvar významně liší od elipsy. V tomto případě jsou výhodné genetické algoritmy [36].

3.2.3 Detekce očí

Detekce očí je velmi složitý úkol vzhledem k tomu, že kontrast oka je velmi slabý a je tedy obtížné získat hranice oka. Určité části oka však mají vyšší kontrast jako např. hranice bělma [13]. Přibližnou detekci regionu očí lze realizovat z binárního obrazu, jehož získání je popsáno v části 3.2.1. Tento postup lze rozdělit na následující části [36]:

1. Nalezení regionu vlasů, což je největší tmavá část v binárním obraze.

2. Nalezení regionu očí, které jsou první tmavou oblastí, kterou lze najít pod oblastí vlasu, ležící na její ose L. Zleva i zprava jsou oči omezeny světlejšími oblastmi hranic tváře. Získáme tak oblast E.

3. Extrakce regionu očí. V oblasti E se nachází dvojice tmavějších částí, které reprezentují polohu očí. Vzhledem k symetrii obličeje by měly být tyto části regionu souměrné podle osy L, ležet v jedné linii, která je kolmá na tuto osu a být položeny pod obočím.

Pomocí těchto poznatků lze zhruba určit polohu očí.

Další zpřesnění detekce polohy očí se provádí s přihlédnutím na detekci duhovky (kruhový tvar uprostřed regionu očí) a posléze hranic oka, které by měly protínat svislou a vodorovnou osu duhovky.

3.2.4 Detekce úst

Po získání polohy očí je nezbytné otestovat přítomnost úst, to nám zaručí nejen ověření správnosti detekce očí, ale umožní to detekci dalších části obličeje a umožní zpřesnit polohu a tvar obličeje samotného [25]. Je vhodné detekovat ústa dříve než nos, protože jejich lokalizace je snazší vzhledem k rozdílné barvě úst a kůže. Další výhodou oproti detekci nosu je fakt, že i při odlišné rotaci obličeje leží stále na ose očí.

Co se týče metod nalezení úst, nejpoužívanější je zřejmě detekce pomocí Fisherova lineárního diskriminantu (FLD, více o Lineární Diskriminační analýze viz 4.3) či detekce pomocí šablon.

FLD se používá ve spojení s předpokladem, že lidská ústa se vyznačují vysokou hodnotou červené barvy a velmi nízkou hodnotou modré barvy. Pomocí transformace a následného prahování je tedy možné získat polohu regionu úst. Pomocí rovnice (3.3) se obraz upraví tak, aby více vynikly regiony s vyšším obsahem červené barvy [32].

$$\mathbf{FLD} = \begin{bmatrix} -0.269 & 0.379 & 0.038 \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{3.3}$$

Pomocí další transformace

$$g = \frac{G}{R + G + B} \tag{3.4}$$

získáme obraz, který postihuje některé další případy barevných odchylek.

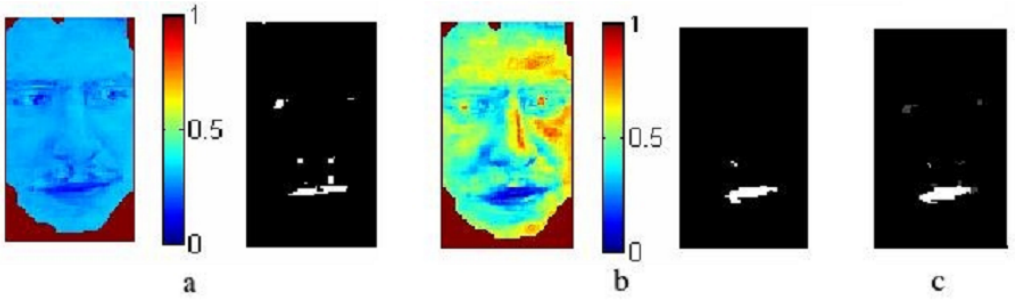
Pomocí prahování metodou mean-shift lze z takto transformovaných obrazů binární mapu [32], na které je patrný region úst. Zde však vyvstává několik problémů jako je obsah červené složky v pozadí (lze eliminovat pomocí preprocesingu - detekce hranic obličeje), změna červené složky v ústech vlivem líčení (obtížný problém pro tuto metodu) a změna barvy obličeje např. vlivem osvětlení (řešitelné pomocí kombinace této metody s některou metodou extrakce rysů, např. Gáborovou vlnkovou transformací).

Získané binární obrazy se sloučí v poměru daném rovnicí 3.5, s větším důrazem na obraz získaný pomocí FLD transformace.

$$\beta = 0.25 \cdot (\beta_{FLD} + \beta_g) + 0.75 \cdot \beta_{FLD} \cdot \beta_g, \tag{3.5}$$

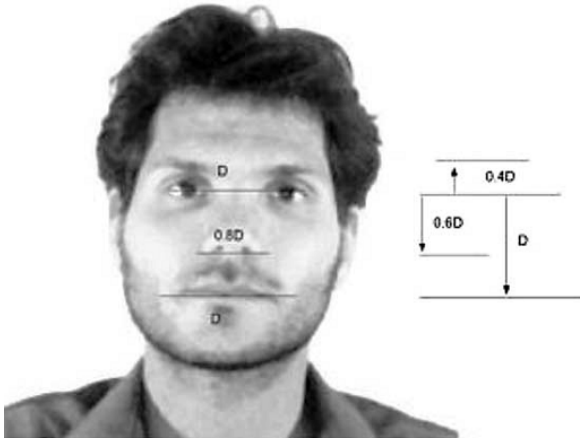
kde β značí výsledný binární obraz, β_{FLD} binární obraz získaný transformací podle Fisherova lineárního diskriminantu a β_g je binární obraz podle zelené složky. Výsledky transformací jsou ukázány na obrázku 3.3.

Další používanou metodou detekce je detekce pomocí šablon (template matching). Tato metoda je založena na porovnávání předem definovaných parametrizovaných šablon s obrázkem, který obsahuje hledaný region. K hledání úst jsou použity stejně jako v případě očí dvě šablony. Je definována funkce, která spojuje hrany, maxima a minima v intenzitě obrazu s příslušnými částmi šablony. Pomocí změny parametrů funkce při každé iteraci se snaží algoritmus tuto funkci minimalizovat.



Obrázek 3.3: Obličej zpracovaný pomocí různých transformací jeho barevných složek, aby vyniknul region úst. (a) zobrazuje obličej po transformaci pomocí rovnice (3.3) a jeho barevnou mapu, (b) je obličej po transformaci pomocí (3.4) s příslušným binárním obrazem a (c) zobrazuje kombinaci předchozích binárních obrazu z (a) a (b) [32]

Poslední zde zmíněnou metodou je metoda založená na geometrické poloze jednotlivých částí obličeje. Předpokládáme, že vzdálenost mezi pravým a levým okem je přibližně stejná jako vzdálenost mezi ústy a středu spojnice obou očí, viz obrázek 3.4. Ústa tedy hledáme přibližně v tomto místě na průsečíku vertikální osy očí v místě největší intenzity šedé barvy [37].



Obrázek 3.4: Rozložení jednotlivých částí obličeje podle jejich geometrické polohy [37]

3.3 Extrakce rysů

Extrakce rysů (tj. vlastností) je v případě zpracování obrazu proces, kdy jsou vstupní data redukována kvůli nadbytečným informacím nebo pro usnadnění algoritmizace následné práce s nimi. Rysy jsou jisté charakteristické vlastnosti obsahu média. Rysy člověka mohou být například barva kůže, vlasů, očí, jejich vzdálenost, velikost nosu, hlavy, drsnost kůže atd. [9] Vzhledem k potřebě redukce dat nastává otázka, které rysy lze z obrazu bez problémů vypustit a jakým způsobem se naopak mají určit ty vlastnosti, které jsou pro danou úlohu

relevantní. Tato otázka naráží na problematiku lidského vnímání a umělé inteligence, která ho napodobuje [28]. V podstatě nelze globálně určit, které rysy budou potřeba pro veškeré případy.

Druhy extrakce rysů lze rozdělit do několika skupin [9], [15]:

- Nízká úroveň popisu, což je detekce bodů, rohů, hran apod. Rysy jsou společné pro celý zkoumaný subjekt. Jsou relativně jednoduché.
- Střední úroveň popisu se soustředí na přechod od ikonického popisu ke strukturálnímu. Jeho reprezentací může být sada geometrických vlastností jako body, čáry nebo oblasti, které popisují významné struktury.
- Vysoká úroveň popisu oproti předchozím úrovním zavádí navíc sémantické znalosti o zkoumaném subjektu. Příkladem může být získání kartografických hodnot z digitálních snímků terénu.

Co se týče identifikace obličeje, je tato fáze specifická pro různé metody (viz 4). Metody se rozdělují do třech základních skupin:

Generické metody - založené na nízké úrovni popisu, tzn. body, hrany, čáry a křivky.

Metody založené na šablonách - určené převážně k detekci částí obličeje.

Strukturální metody - založené na geometrických vztazích mezi jednotlivými rysy.

3.4 Učení

Poslední fáze procesu se dělí na učení a rozpoznání. Učicí část identifikace obličeje se provádí zpravidla jen jednou, aby se vytvořila pravidla pro identifikaci, která jsou poté použita ve fázi rozpoznání. Pravidla jsou uložena pro pozdější použití, takže není nutné provádět tuto operaci vícekrát.

Učení (v tomto případě strojové učení) je oblast, jejímž účelem je stanovit pomocí daných algoritmů a technik způsob, kterým se budou zpracovávat empirická data. Cílem je dosáhnout takového chování, které může být považováno za inteligentní. Se strojovým učením jsou spojeny oblasti umělé inteligence, teorie pravděpodobnosti, statistika, dolování dat, rozpoznání vzorů apod.

Strojové učení se rozděluje do následujících kategorií

Učení s učitelem se snaží nalézt spojitost mezi zadanými vstupními a výstupními daty. V případě identifikace obličeje jsou vstupní obličeje označeny podle příslušnosti k jednotlivým osobám.

Učení bez učitele modeluje zadaná data tak, že mezi nimi hledá podobnost. Příkladem učení bez učitele je shlukování a co se týče rozpoznání obličeje, lze jako příklad zmínit metodu PCA (viz 4.2).

Kombinace učení s učitelem a bez učitele využívá jak označená, tak neoznačená vstupní data k vytvoření žádané funkce nebo klasifikátoru.

Zpětnovazebné učení, též učení posilováním je založeno na upravování řešení podle analýzy výsledku.

3.5 Rozpoznání

V této fázi jsou využity principy dané učením na vstupní množině dat. Běžně používaným způsobem je projekce vstupních dat do jiného prostoru. Takto získaná data jsou poté porovnávána pomocí různých algoritmů (například algoritmus k-nejbližších sousedů) s daty získanými z učicí množiny.

3.5.1 Vzdálenosti

Pomocí vzdálenosti se měří rozdílnost jednotlivých vektorů, které obsahují informace o obličejích. Vzdálenost je numerický popis, který vyjadřuje, jak daleko jsou jednotlivé objekty od sebe. Zde se používá matematického významu vzdálenosti, který je zobecněním vzdálenosti fyzikální. Je definována jako funkce [12, 5]

$$d : X \times X \to \mathbb{R} \,, \tag{3.6}$$

pro kterou platí:

$$d(x, y) \geq 0 \text{ (pozitivnost)} \tag{3.7}$$

$$d(x, y) = 0 \text{ , právě když } x = y \text{ (identita nerozlišitelných)} \tag{3.8}$$

$$d(x, y) = d(y, x) \text{ (symetrie)} \tag{3.9}$$

$$d(x, z) \leq d(x, y) + d(y, z) \text{ (trojúhelníková nerovnost)} \tag{3.10}$$

Vzdálenost lze měřit v libovolném n-rozměrném prostoru, což je výhodné pro měření odlišnosti obličejů popsaných n-rozměrnými vektory.

Eukleidova vzdálenost

Eukleidovská vzdálenost je metrika, která popisuje klasickou vzdálenost mezi dvěma body, která je dána Pythagorovou větou. Je dána vzorcem [5]:

$$|\mathbf{p} - \mathbf{q}\| = \sqrt{(\mathbf{p} - \mathbf{q}) \cdot (\mathbf{p} - \mathbf{q})} = \sqrt{\|\mathbf{p}\|^2 + \|\mathbf{q}\|^2 - 2\mathbf{p} \cdot \mathbf{q}} \tag{3.11}$$

Jedná se o nejběžněji používanou vzdálenost, protože je jednoduše představitelná v reálném světě.

Manhattan vzdálenost

Manhattan vzdálenost je metrika, která měří vzdálenost v n-rozměrném prostoru, ve kterém se lze pohybovat pouze ve směru jeho os. Délka dráhy z bodu do bodu lze tedy vypočítat pomocí eukleidovské vzdálenosti (nejedná se však o eukleidovskou vzdálenost bodů), ale vhodnější je použít vztahu [5]:

$$d_1(\mathbf{p}, \mathbf{q}) = \|\mathbf{p} - \mathbf{q}\|_1 = \sum_{i=1}^{n} |p_i - q_i| \,, \tag{3.12}$$

kde

$$\mathbf{p} = (p_1, p_2, \ldots, p_n) \quad \text{a} \quad \mathbf{q} = (q_1, q_2, \ldots, q_n) \text{ jsou vektory.} \tag{3.13}$$

Minkovského vzdálenost

Minkovského vzdálenost je zobecněním předchozích příkladů vzdáleností. Je dána vzdahem

$$d_1(\mathbf{p}, \mathbf{q}) = \left(\sum_{i=1}^{n} |x_i - y_i|^p \right)^{1/p} \tag{3.14}$$

Mahalanobisova vzdálenost

Mahalanobisova vzdálenost je zajímavá, protože její hodnota je závislá na hustotě rozložení dat a bodu v prostoru, z kterého je pozorována. Je dána vztahem:

$$d(\vec{x}, \vec{y}) = \sqrt{(\vec{x} - \vec{y})^T S^{-1} (\vec{x} - \vec{y})}, \tag{3.15}$$

kde $\sum^{-1}$ je kovariační matice distribuce. Pozorovací vzdálenost je stejná pro regiony se stejnou hustotou, na rozdíl od Eukelidovy vzdálenosti. Body v daných Mahalanobisových vzdálenostech leží na elipsoidu se středem ve výchozím bodě pozorování.

Chebyshevova vzdálenost

Chebyshevova vzdálenost je definována jako největší rozdíl vzdáleností v jakémkoliv prostoru [5]:

$$d(\vec{x}, \vec{y}) = \max_i (|p_i - q_i|). \tag{3.16}$$

což je rovno limitě

$$\lim_{k \to \infty} \left(\sum_{i=1}^{k} |p_i - q_i|^k \right)^{1/k} \tag{3.17}$$

Pro vysvětlení Chebyshevovy vzdálenosti se používá paralela s šachovnicí. Vzdálenost mezi dvěma poli na šachovnici je dána počtem tahů, které musí vykonat král aby se mezi těmito poli přemístil (Král se může pohybovat vodorovně, svisle a po diagonálách, ale vždy jen o jedno pole).

3.6 Ukládání dat pro identifikaci

V této prácí jsou použity vzorky dat, které obsahují maximálně 295 subjektů. Reálné systémy identifikace obličejů, však teoreticky mohou obsahovat až 6 miliard různých tříd lidí, které je potřeba nějakým způsobem procházet. Zde narážíme na problém efektivního uložení dat, aby bylo umožněno co nejrychlejší vyhledávání a přístup k datům. Nejpoužívanější přístupy k tomuto problému jsou hashování a indexování.

3.6.1 Indexování

Indexování je pojem, který označuje obor, který se zabývá přístupem k datům pomocí unikátního klíče (indexu). Kromě primárního klíče mohou být databáze doplněny od sekundární klíč, který napomáhá orientaci ve struktuře databáze.

Bez indexu by musel vyhledávací systém databáze projít každý záznam a vyhodnotit shodu se zadaným vyhledávacím dotazem. Databáze se sice zmenší o prostor nutný k uložení indexovací struktury a každý zápis do databáze se zpomalí o práci s indexem, avšak toto je naprosto vyváženo urychlením operací vyhledávání v datech.

Pro orientaci a rychlost úprav indexů jsou používány různé druhy stromů. Přístupy se rozlišují na Bodové přístupové metody, které vnímají data v databázi jako body v prostorů o dané dimenzi, a Prostorové přístupové metody, které uvažují uložená data jako množinu geometrických tvarů v prostoru.

Nejpoužívanějšími metodami indexování jsou [5]:

B-strom - generalizace binárního stromu. Je to acyklický graf s jedním kořenem. Pro strom stupně n platí, že uzel má nejméně dva potomky, pokud není koncový, uzel má $\frac{n}{2} - 1$ až n potomků, dále všechny cesty od kořene k potomkům jsou stejně dlouhé a data v uzlech jsou logicky organizována do posloupností.

Mřížkový soubor - rozdělení prostoru podle mřížky, která je vytvořena v závislosti na hustotě dat v tomto prostoru. Části vzniklé rozdělením mřížkou mohou být dále děleny tak, že osa dělení je kolmá na již existující mřížku.

KD-strom - postupné dělení plochy tak, že se dělí vždy delší hrana obdélníkové plochy. Plocha se dělí do té doby, než je v každé ploše obsažen jen jeden prvek. Vznikne tak binární strom.

Quad-strom - algoritmus je podobný KD-stromu, ale plochy dělí ve všech osách a v přesné polovině.

R-strom - strom složený z obdélníkových struktur, které mohou obsahovat buď další obdélníkové struktury, nebo objekt.

SS-strom - podobně jako R-strom, jen jsou užity kruhové (příp. kulové) struktury.

VP-strom - dělení prostoru podle vzdálenosti od daného výchozího bodu.

Hashovací techniky (viz níže) nejsou používány pro indexování vzhledem k problémům, které ve vícedimenzionálních prostorech nastávají [5]. Jde o to, že důležité parametry jako je obsah (objem) oblasti závisí exponenciálně na dimenzi datového prostoru, tím pádem jsou tyto algoritmy vhodné jen pro prostory s malými dimenzemi.

3.6.2 Hashování

Hashování je proces, kdy jsou s použitím hashovací funkce převedena vstupní data na relativně malé číslo. Tyto funkce lze použít např. k porovnávání rozměrných dat, rychlému vyhledávání, indexování nebo kontrole integrity dat [3].

Výhodou hashování je, že z libovolných vstupních dat generuje řetězec konstantní délky, dále jakákoliv změna na vstupních datech se projeví ve výsledku funkce. Tím, že výsledný řetězec má menší délku než vstupní data, však může nastat problém, že pro různé vstupy vygeneruje funkce stejné hashe. Tento problém nelze z principu zcela odstranit.

Mezi používané metody hashování patří hashování dělením, metoda středních čtverců, hashování násobením, metoda rozkladu, Fibonacciho hashování a Cormackovo hashování.

Kapitola 4

Metody identifikace

K identifikaci obličeje lze přistupovat mnoha způsoby, které se liší již ve fázi extrakce rysů (viz 3.3). Některé algoritmy identifikují obličej pomocí extrakce hran nebo rysů obličeje, což spočívá v měření relativní pozice, velikosti a tvaru očí, nosu, úst, brady a lícních kostí. Jiné přístupy se zaměřují na normalizaci galerie obličejů a poté komprimují takto získaná data tím, že zachovávají jen ta data, která jsou důležitá pro identifikaci obličeje.

Mezi významné algoritmy patří: Analýza hlavních komponent (PCA - Principal Component Analysis), Lineární diskriminační analýza (LDA - Linear Discriminat Analysis), Elastický srovnávací diagram (EBGM - Elastic Bunch Graph Matching), Skrytý Markovův model (HMM - Hidden Markov model), Podpůrné vektory (SVM - Support Vector Machines) a grafové algoritmy pomocí neuronových sítí.

Algoritmy identifikace obličeje mohou být rozděleny na tři hlavní přístupy, holistický, strukturální a hybridní přístup, který je kombinací předchozích [21].

Holistický přístup zpracovává obraz jako celek. Jako vstup používá abstrahované informace z celé scény například eigenfaces u metody PCA.

Strukturální přístup rozděluje obraz na části, které popisují daný subjekt (př. obočí, oči, ústa). Při hledání shody se jako vstup klasifikátoru používá reprezentace jednotlivých částí.

Hybridní přístup je kombinací obou předchozích metod.

4.1 Algebraické předpoklady identifikace obličeje

I když jednotlivé metody identifikace obličeje přistupují více či méně různými způsoby k řešení stejného problému, určité termíny mají společné. Jde hlavně o pojmy eigen hodnota, eigen vektor a eigen prostor (vlastní číslo, vlastní vektor, vlastní prostor). Jedná se o principy lineární algebry. Zmíněné pojmy jsou vlastnosti, které poskytují důležité informace o matici a mohou být použity pro její faktorizaci.

4.1.1 Vlastní hodnoty, vektory a prostor

Nechť A je čtvercová matice řádu n

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

Jestliže pro určitý skalár λ existuje vektor $u \in \mathbb{R}_n$, $u \neq o$ takový, že

$$Au = \lambda u \tag{4.1}$$

potom λ se nazývá vlastní číslo, resp. charakteristická hodnota matice A a vektor u se nazývá vlastní vektor matice A. Podmínku (4.1)

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} = \lambda \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} \tag{4.2}$$

lze zapsat ve tvaru

$$\begin{pmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{pmatrix} \cdot \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{4.3}$$

Soustava (4.3) je homogenní a lze zapsat jako

$$(A - \lambda E)u = o \tag{4.4}$$

kde E je jednotková matice řádu n.

Homogenní soustava má netriviální řešení, právě když je determinant matice soustavy roven nule, tj. v případě soustavy (4.3), resp. (4.4) platí

$$|A - \lambda E| = 0 \tag{4.5}$$

Determinant $A(\lambda) = |A - \lambda E|$ nazýváme charakteristický polynom matice A — jedná se o polynom stupně n v proměnné λ, který má v oboru komplexních čísel n kořenů.Rovnici $A(\lambda) = 0$ nazýváme charakteristická rovnice matice A — jejími kořeny jsou charakteristické hodnoty (resp. vlastní čísla) matice A.

4.1.2 Kovariance

Kromě již zmíněných pravidel lineární algebry v části 4.1, jsou některé metody založeny na kovarianci. Kovariance je vlastnost, která vyjadřuje závislost dvou veličin. Je vyjádřena jako střední hodnota součinu odchylek dvou veličin (X, Y) od jejich středních hodnot. Je dána vztahem

$$\mathrm{Cov}(X,Y) = \mathrm{E}\left[(X - \mathrm{E}[X])(Y - \mathrm{E}[Y])\right] = \mathrm{E}\left[XY\right] - \mathrm{E}[X] \cdot \mathrm{E}[Y],$$

kde $E(X)$ je očekávaná hodnota X.

Zobecněním pro n veličin získáme kovarianční matici

$$\Sigma_{ij} = \mathrm{cov}(X_i, X_j) = \mathrm{E}\left[(X_i - \mu_i)(X_j - \mu_j)\right] \text{ , } 1 \leq i \leq n \text{ a } 1 \leq j \leq n,$$

kde

$$\mu_i = \mathrm{E}(X_i)$$

je střední hodnota i-té hodnoty vektoru X. Matici lze rovněž vyjádřit jako

$$\Sigma = \begin{bmatrix} \mathrm{E}[(X_1 - \mu_1)(X_1 - \mu_1)] & \mathrm{E}[(X_1 - \mu_1)(X_2 - \mu_2)] & \cdots & \mathrm{E}[(X_1 - \mu_1)(X_n - \mu_n)] \\ \mathrm{E}[(X_2 - \mu_2)(X_1 - \mu_1)] & \mathrm{E}[(X_2 - \mu_2)(X_2 - \mu_2)] & \cdots & \mathrm{E}[(X_2 - \mu_2)(X_n - \mu_n)] \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{E}[(X_n - \mu_n)(X_1 - \mu_1)] & \mathrm{E}[(X_n - \mu_n)(X_2 - \mu_2)] & \cdots & \mathrm{E}[(X_n - \mu_n)(X_n - \mu_n)] \end{bmatrix}$$

4.1.3 Singulární rozklad matice

Singulární rozklad matice (SVD - Singular value decomposition) je poměrně nová metoda určená při různých maticových výpočtech a analýze dat, která poskytuje informace o geometrické struktuře dané matice.

Každá matice $A \in \mathbb{R}^{m \times n}$ lze rozdělit na součin

$$A = UWV^T \text{ ,} \tag{4.6}$$

kde $U \in \mathbb{R}^{m \times m}$ a $V \in \mathbb{R}^{n \times n}$ jsou ortonormální matice a $W \in \mathbb{R}^{m \times n}$ je diagonální matice s kladnými reálnými prvky uspořádanými v sestupném pořadí [11].

SVD lze použít např. k řešení soustav lineárních rovnic, filtrování zašuměných signálů, analýze časových řad apod. V našem případě je jednou z metod, pomocí které lze v knihovně OpenCV počítat např. inverzní matice nebo řešit vlastní systémy.

4.2 Analýza hlavních komponent

Analýza hlavních komponent (PCA - Principal component analysis) je technika, která transformuje určitý počet korelovaných hodnot na menší počet dekorelovaných hodnot zvaných hlavní komponenty [33]. Redukuje se tedy dimenze příznakového prostoru, to se dá použít také v případě extrakce rysů (viz 3.3).

Redukci dimenzionality lze vysvětlit na speciálním případu PCA, metodě nejmenších čtverců. Jde o nalezení přímky, která by nejlépe popisovala body ve 2D scéně. Dochází tedy k redukci dimenzionality 2D prostoru na 1D. Tato přímka je položena tak že suma druhých mocnin vzdáleností k jednotlivým bodům je minimální.I když je přímka 1D objekt, nachází se ve větším, 2D prostoru a má tedy určitou orientaci (směrnici). Směrnice přímky zachycuje důležitou informaci o bodech, které popisuje a to jejich rozložení v prostoru. Přímka lze tedy jednoduše popsat jako

$$y = mx \ , \tag{4.7}$$

kde je směrnice přímky

$$m = \frac{\mathrm{d}x}{\mathrm{d}y} \ . \tag{4.8}$$

Tento popis zachycuje pohled na data, který je z pohledu PCA zajímavý, tedy směr který nejlépe separuje jednotlivá data od sebe.

Směr nejlepší separace v prostoru se nazývá první hlavní komponenta datové množiny. Druhý nejlepší směr separace je závislý na první, nazývá se druhá hlavní komponenta. Pro 2D prostor lze nalézt jen 2 hlavní komponenty. Dále záleží na počtu prvků. Například pro jeden bod neexistuje žádná hlavní komponenta, protože není co separovat.

4.2.1 Použití pro identifikaci obličeje

Co se týká rozpoznání obličeje, PCA využívá vektorů tváře odvozených z kovarianční matice pravděpodobnostní distribuční funkce k vytvoření šablony vhodné pro srovnávání.

Redukce dimenzionality je v tomto případě důležitá pro redukci šumu v obraze, který by negativně ovlivňoval výsledky rozpoznání obličeje. Například při použití vzorku 64x64 je v obrazu obsaženo 4096 dimenzí, které by byly při identifikaci obličeje uvažovány. Obraz lze popsat ale daleko menším počtem údajů, k čemuž slouží data získaná pomocí PCA.

Každá tvář lze rozdělit na tzv. eigenfaces (vzory tváří - matice jasových úrovní) a poté jde opět složit. Každá eigenface je reprezentována pouze číslem, takže se namísto obrázku ukládá pouze číslo.

4.2.2 Základní principy metody

Uvažujme trénovací množinu obrázků P, kdy všechny prvky této sady mají stejné rozměry WxH. Z těchto obrázků sestavíme matici o počtu řádků rovném počtu obrázků P a počtu

sloupců rovnému WxH. Obrázky tváří jsou si podobné, ať už ve více či méně rysech, a proto nejsou náhodně distribuovány (existuje mezi nimi nenulová korelace). Hlavní myšlenkou PCA pro námi zamýšlené účely je nalézt vektory, které nejlépe popisují distribuci obrázků tváří v celém prostoru obrázků. Tyto nalezené vektory dimenze n definují podprostor, který nazýváme prostor obrázků tváří (eigenspace, též facespace), jsou lineární kombinací původních vektorů a tvoří bázi eigenspace. Tyto vektory jsou vlastními vektory kovarianční matice korespondující k původním obrázkům, a protože jsou podobné obrázkům tváří, nazýváme je eigenfaces [14], [38].

4.2.3 Popis metody

Obrázek 4.1: Vstupní onličejová data získaná z databáze AT&T

2D obraz obličeje lze reprezentovat jako matici o jednom sloupci tak, že jednotlivé řádky uvedeme pod sebe. Pro M obrazů použijeme tedy M vektorů délky N, kde N je počet řádků obrazu x počet sloupců obrazu a p_j reprezentuje hodnoty jednotlivých pixelů.

$$x_i = [p_1 ... p_N]^T, i = 1, ..., M \tag{4.9}$$

Pro reprezentaci všech obličejů potřebujeme průměrný obraz, který odečteme od ostatních obrazů. Průměrný obraz je definován jako:

$$m = \frac{1}{M} \sum_{i=1}^{M} x_i \tag{4.10}$$

A střední průměrné obrazy získáme odečtením

$$w_i = x_i - m \tag{4.11}$$

Naším cílem je nalezení množiny hodnot e_i, které mají největší možnou projekci do hodnot w_i. Hledáme tedy množinu ortonormálních vektorů e_i, kde hodnota

$$\lambda_i = \frac{1}{M} \sum_{n=1}^{M} (e_i^T w_i)^2 \tag{4.12}$$

je maximální s ortonormální konstantou

$$e_l^T e_k = \sigma_{lk} \tag{4.13}$$

Obrázek 4.2: Průměrný obraz získaný kombinací vstupních dat

Bylo dokázáno, že hodnoty e_i a λ_i jsou dané eigenvektory a eigenhodnotami kovarianční matice

$$C = WW^T \,, \tag{4.14}$$

kde W je matice vytvořená ze sloupcových vektorů w_i položených za sebou. Velikost matice C je N x N, což může být příliš velké číslo. Například obraz velikosti 64 x 64 dává kovarianční matici o 4096 x 4096 prvcích. Není praktické řešit eigenvektory matice C přímo. Jednoduchý teorém lineární algebry ukazuje, že vektory e_i a skaláry λ_i lze získat řešením pro eigenvektory a eigenhodnoty M x M matice W^TW. Nechť d_i a μ_i jsou eigenvektory a eigenhodnoty matice W^TW, tedy

$$W^TWd_i = \pi_i d_i \tag{4.15}$$

Po vynásobení obou stran rovnice W získáme

$$WW^T(Wd_i) = \pi_i(Wd_i) \tag{4.16}$$

což znamená, že prvních $M-1$ eigenvektorů e_i a eigenhodnot λ_i matice WW^T jsou dané pomocí Wd_i respektive π_i. Wd_i je potřeba normalizovat aby se shodovalo s e_i. Díky tomu, že počítáme jen s konečným počtem vektorů obrazu M, hodnost kovarianční matice nemůže přesáhnout $M-1$ (-1 je z odečítání od průměrného vektoru m).

Obrázek 4.3: Vlastní vektory (eigenfaces) získané z vlastního systému

Eigenvektory odpovídající nenulovým eigenhodnotám kovarianční matice vytváří ortonormální bázi pro podprostor, v němž většina obrazových dat může být reprezentována s jen malou chybou. Eigenvektory jsou seřazeny od největších po nejmenších podle korespondujících eigenhodnot. Eigenvektor asociovaný s největší eigenhodnotou je ten, který odráží největší odchylku v obrazu, naopak nejmenší eigenhodnota označuje vektor s nejmenší odchylkou. Hodnoty se zmenšují exponenciálně, což znamená, že přibližně 90% variance je obsaženo v prvních 5-10% prostorech.

Obraz tváře může být promítnut do M' ($<< M$) spočtením

$$\Omega = [v_1 v_2 ... v'_M]^T \tag{4.17}$$

kde $v_i = e_i^T w_i$. v_i je i-tá souřadnice obrazu obličeje v novém prostoru, tedy hlavní komponenty. Vektory e_i jsou také obrazy, takzvané eigenobrazy nebo v našem případě eigenfaces. Ω tedy udává příspěvek jednotlivých eigenfaců do reprezentace obrazů obličejů pomocí množiny eigenfaců. Nejjednodušší metoda jak zjistit, která třída obličejů nejlépe popisuje vstupní obraz je nalezení třídy obličejů, která minimalizuje eukleidovskou vzdálenost

$$\epsilon = ||(\Omega - \Omega_k)|| \ , \tag{4.18}$$

kde Ω_k je vektor, který popisuje k-tou třídu obličejů. Pokud je ϵ_k menší než zvolený práh θ_ϵ, je obličej zařazen do třídy k.

4.2.4 Spojení PCA s jinými metodami

Analýza hlavních komponent dosahuje dobrých výsledků, ale v některých případech selhává. Jedním ze způsobů snížení chybovosti metody je spojení s jinou metodou.

Metoda je nejčastěji doplněná metodou ICA (viz 4.4). Podle [34] se při spojení těchto metod dosahuje přesnosti až 99.3%. Metody spolupracují tak, že nejprve je obraz rozpoznáván metodou PCA a pokud není nalezena shoda, postoupí se dále a rozpoznává se metodou ICA pro zpřesnění výsledku.

4.3 Lineární diskriminační analýza

Lineární diskriminační analýza (LDA - Linear Discriminant analysis) je metoda, kdy se třídí pořízené obrazy tváří do skupin. Cílem je maximalizovat rozdíly mezi jednotlivými skupinami a minimalizovat rozdíly v každé skupině, každý blok snímků reprezentuje jednu třídu [42].

V běžných aplikacích se používají dva až čtyři trénovací vzorky na osobu. Stejně jako u metody PCA (viz 4.2) je pro rozlišení vstupního obrazu 64 x 64 počet dimenzí rysů 4096. I pro projekci do podprostoru je dimenzí rysů přibližně 300. Při poměru dimenze/vzorek 150 nebo 75 dohází k exponenciálnímu zvětšení objemu informací, avšak i při takovýchto nevhodných situacích je možné použít LDA jako klasifikátor citeEtemad1997.

Vzhledem k tomu, že příslušnost jednotlivých obličejů trénovací množiny je u této metody popsána, je vhodné použít metodu, která redukuje dimenzionalitu prostoru rysů na základě rozdílnosti těchto tříd. Fisherův lineární diskriminant je metoda, která se snaží co nejvíce přizpůsobit rozptyl hodnot tak, aby byl co nejlépe využitelný pro klasifikaci.

4.3.1 Popis metody

Metoda LDA je založena na analýze rozptylové matice. Hledá se projekční matice W taková, že poměr mezi rozptylovou maticí mezi třídami a rozptylovou maticí uvnitř tříd je maximální.

Rozptylová matice mezi třídami je definována jako

$$S_B = \sum_{i=1}^{c} N_i(\mu_i - \mu)(\mu_i - \mu)^T \tag{4.19}$$

a rozptylová matice uvnitř tříd je definována vztahem

$$S_W = \sum_{i=1}^{c} \sum_{x_k \in X_i} (x_k - \mu_i)(x_k - \mu_i)^T \ , \tag{4.20}$$

kde μ_i je průměrný obraz třídy X_i a N_i je počet vzorků třídy X_i. Pokud je S_W regulární, pak optimální projekce W_{opt} je volena jako matice, jejíž sloupce jsou ortonormální, což maximalizuje poměr determinantu rozptylové matice mezi třídami projektovaných vzorků k determinantu rozptylové matice vně tříd projektovaných vzorků. Například

$$W_{opt} = \arg\max_W \frac{|W^T S_B W|}{|W^T S_W W|} = [w_1 w_2 \dots w_m] \ , \tag{4.21}$$

kde $\{w_i | i = 1, 2, \dots, m\}$ je množina generalizovaných eigenvektorů matic S_B a S_W k m největším odpovídajícím eigenhodnotám $\{\lambda | i = 1, 2, \dots, m\}$. Sloupcové vektory w_i matice W jsou označovány jako fisherfaces.

$$S_B w_i = \lambda_i S_W w_i, \ i = 1, 2, \dots, m \tag{4.22}$$

Maximální počet nenulových generalizovaných eigenvektorů je $c-1$, tedy i maximální hodnota m je $c-1$, kde c je počet tříd.

Systém daný rovnicí (4.22) lze upravit vynásobením obou stran inverzní maticí k S_W, čímž získáme:

$$S_W^{-1} S_B w_i = \lambda_i w_i, \ i = 1, 2, \dots, m \ , \tag{4.23}$$

což je již běžný vlastní systém, který lze řešit například pomocí singulárního rozkladu matice (SVD).

Výslednou transformační matici W lze poté použít k projekci vstupních dat do prostoru, kde je výhodnější jejich porovnáváni.

4.3.2 Redukce dimenzionality vstupních dat

Ve většině případů je nutná redukce dimenzionality vstupních dat vzhledem k tomu, že počet prvků M trénovací množiny dat je mnohem menší než dimenzionalita množiny D, to způsobuje to, že matice S_W je téměř vždy singulární. K tomuto účelu je běžně používaná metoda Analýza hlavních komponent, která je popsána v části 4.2. Prvních $(M - c)$ eigen-projekcí je použito k reprezentaci vektorů x_i. To mimo jiné umožňuje, aby byly matice S_W a S_B spočítatelné pomocí běžné velikosti paměti počítače. Optimální projekce W_{opt} je potom definována jako

$$W_{opt} = W_{fld} \cdot W_{pca} \ , \tag{4.24}$$

kde W_{pca} je projekční matice metody PCA a W_{fld} je optimální projekce získaná maximalizací

$$W_{fld} = \arg\max_W \frac{|W^T W_{pca}^T S_W W_{pca} W}{W^T W_{pca}^T S_B W_{pca} W|} \tag{4.25}$$

4.3.3 Výhody a nevýhody

- Na rozdíl od např. PCA, která kóduje informaci, která se vztahuje ke kompresi, LDA kóduje informace, které rozlišují vzorky od sebe.
- LDA je zvláště vhodná pro vzorky, které jsou lineárně separovatelné.
- Jednoduchá implementace.

4.4 Analýza nezávislých komponent

Analýza nezávislých komponent (ICA - Independent Component Analysis) je metoda, která se snaží nalézt zdrojové signály ze znalosti jejich kombinace. V literatuře se můžeme setkat s několika odlišnými definicemi ICA, ale všechny mají něco společného. Všechny předpokládají lineární kombinaci zdrojových dat (též označovaných jako komponenty) [35].

4.4.1 Formální definice

V případě ICA jsou data reprezentována vektorem m-rozměrným vektorem $X = x_1, x_2, \ldots, x_m$, nezávislé komponenty jsou reprezentovány n-rozměrným vektorem $S = s_1, s_2, \ldots, s_n$. Úkolem je nalezení nezávislých komponent, které byly transformovány pomocí směšovací matice rozměru $m \times n$. Platí

$$X = A \cdot S \tag{4.26}$$

a naopak

$$S = W \cdot X \ , \tag{4.27}$$

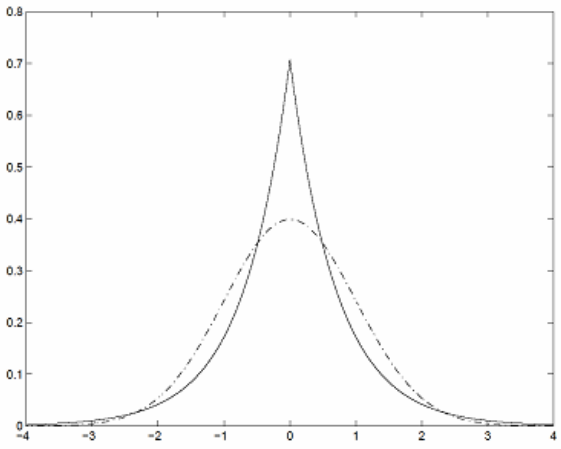
kde W je inverzní matice je směšovací matici A. Nalezení W je tedy řešení nalezení nezávislých komponent.

Prvky x_i vektoru X jsou tvořeny jako součet nezávislých komponent s_k, kde $k = 1, 2, \ldots, n$

$$x_i = a_{i,1}s_1 + \ldots + a_{i,k}s_k + \ldots + a_{i,n}s_n, \tag{4.28}$$

kde $a_{i,k}$ jsou váhy jednotlivých nezávislých komponent. Separace nezáviských komponent se tedy provádí tak, aby bylo dosaženo co největší nezávislosti a zároveň platilo, že lineární kombinace nezávislých komponent tvoří původní data. To je dosaženo maximalizací funkce F, která označuje nezávislost komponent.

Nejdůležitějším faktem pro odhad ICA je, že proměnné Gaussova rozložení nelze rozseparovat. Vzhledem k tomu, že řešíme separaci proměnných, hledáme tedy takové komponenty, které se nejvíce liší od normálního Gaussova rozdělení.



Obrázek 4.4: Hustota pravěpodobnosti Laplaceova rozdělení, které je příkladem supergausian ve srovnání (plná čára) s Gaussovým rozdělením (čerchovaná čára)

K odhadu těchto komponent se používají následující funkce:

Špičatost - vlastnost, která měří hustotu rozdělení oproti normálnímu.

$$\text{kurt}(y) = E(y^4) - 3 \cdot E(y^2)^2 \tag{4.29}$$

Proměnné, které mají zápornou špičatost jsou označovány jako subgaussian a ty, které mají špičatost kladnou, jsou označovány jako supergaussian. Špičatost je velice jednoduchá na výpočet, ale citlivá na extrémy a chyby při měření dat. Špičatost není dostatečně robustní pro ICA.

Negativní entropie - je veličina založena na informačně-teoretickém množství entropie. Entropie náhodné veličiny může být interpretována jako střední počet bitů, pomocí nichž lze informaci zakódovat. Entropie H pro diskrétní signál je definována jako:

$$H(y) = -\sum_i P(Y = a_i) \log P(Y = a_i)$$

Nevýhodou entropie je její výpočetní náročnost.

4.5 Skrytý Markovův model

Skrytý Markovův model (HMM - Hidden Markov model) je pravděpodobnostní model, který je tvořen uzly a přechody mezi nimi. Vnitřní vazby mezi uzly jsou pomocí matice pravděpodobnosti přechodů A. Přechody mezi jednotlivými uzly vyjadřují pravděpodobnost, s jakou se může model dostat z jednoho uzlu do jiného. Hodnoty jednotlivých uzlů jsou vyjádřeny pomocí matice generovaných vzorů B. Počáteční stav je určen vektorem pravděpodobností počátečních stavu Π.

4.5.1 Formální popis

Matematicky je model popsán souborem parametrů $\lambda = (A, B, \Pi)$. kde jednotlivé parametry mají následující význam:

- N je počet stavů v modelu. Pokud S je množina stavů, pak $S = S_1, S_2, \ldots, S_N$. Stav modelu v čase t je dán vztahem $q_t \in S, 1 \leq t \leq T$, kde T je délka pozorované posloupnosti (počet snímků).
- M je počet různých symbolů pozorování. Pokud V je množina všech možných pozorování, pak $V = v_1, v_2, \ldots, v_n$.
- A je matice pravděpodobnosti přechodu, např. $A = \{a_{ij}\}$, kde

  $$a_{ij} = P[q_t = S_j | q_{t-1} = S_i] 1 \leq i, j \leq N$$

  s podmínkou

  $$0 \leq a_{ij} \leq 1 \tag{4.30}$$

  a

  $$\sum_{j=1}^{N} a_{ij} = 1, 1 \leq i \leq N \tag{4.31}$$

- B je matice generovaných vzorů (hodnot jednotlivých uzlů), např $B = \{b_j(k)\}$, kde

  $$b_j(k) = P[O_t = v_k | q_t = S_j] \tag{4.32}$$

  $$1 \leq j \leq N, 1 \leq k \leq M \tag{4.33}$$

  a O_t je pozorovaný symbol v čase t.

- Π je počáteční stav, např. $\Pi = \{\pi_i\}$, kde

$$\pi_i = P[q_1 = S_i], 1 \leq i \leq N \tag{4.34}$$

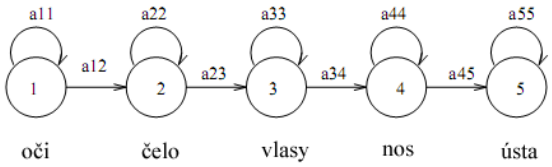
Výše zmíněná charakteristika odpovídá diskrétnímu HMM, kde jednotlivá pozorování jsou diskrétní symboly z konečné abecedy $V = \{v_1, v_2, \ldots, v_n\}$. V HMM se spojitou hustotou jsou stavy chrakterizovány spojitou funkcí hustoty. Nejobecnější reprezentace funkce hustoty pravděpodobnosti je konečná směs tvaru:

$$b_i(O) = \sum_{k=1}^{M} c_{ik} N(O, \pi_{ik}, U_{ik}), 1 \leq i \leq N \tag{4.35}$$

kde c_{ik} je koeficient směsi pro k-tou směs ve stavu i. Bez ztráty obecnosti je $N(O, \pi_{ik}, U_{ik})$ považováno za gausovskou funkci hustoty pravděpodobnosti s průměrným vektorem π_{ik} a kovarianční maticí U_{ik}

4.5.2 Užití pro rozpoznání obličeje

Běžně se používá jednodimenzionální HMM metoda (rozpoznání řeči). Dvou dimenzionální alternativa HMM sice existuje, ale ukázala se jako velmi komplexní. Avšak pokud uvážíme regiony tváře odshora dolů, vystačíme si s 1D alternativou, jak je znázorněno na 4.5. Fáze jsou v přirozeném pořadí a korespondují s pořadím detekce jednotlivých částí obličeje, jak bylo zmíněno v 3.2.



Obrázek 4.5: Diagram zobrazující fáze identifikace obličeje pomocí HMM

Proces se dělí jako u ostatních metod na trénování a vlastní rozpoznávání.

Při trénování se daný obraz dělí na bloky, z kterých se extrahují rysy, poté se inicializuje model, kdy jsou odhadnuty hodnoty matice B. Hodnoty matice A a počátečního stavu Π jsou dány strukturou modelu. V dalších krocích jsou hodnoty přepočítány pomocí metody Expectation-Maximization, k dosažení co nejvyšší $P(O|\lambda)$. Iterace se ukončí, když je dosaženo cílené konvergence, např. když je dosaženo rozdílů pravděpodobnosti menší než zadaný práh C:

$$|P(O|\lambda^{k+1}) - P(O|\lambda^k)| < C \tag{4.36}$$

Při rozpoznání se vstupní obraz opět rozdělí a extrahují se z něj rysy je spočítána pravděpodobnost pozorovaného vektoru podle příslušného HMM modelu obličeje. Obličej t je rozeznán jako obličej k pokud platí:

$$P(O^{(t)}|\lambda_k) = \max_n P(O^{(t)}|\lambda_n) \tag{4.37}$$

4.5.3 Aplikace

Kromě zmíněného užití při identifikaci obličeje jsou nejběžnějšími aplikacemi HMM rozpoznání řeči, segmentace obrazu, analýza biologických dat apod.

4.6 Elastický srovnávací diagram

Metoda Elastický srovnávací diagram (EBGM - Elastic bunch graph matching) byla vyvinuta, jelikož předešlé metody nemohou uvažovat nelineární charakteristiky jako je osvětlení okolí, pozice hlavy anebo výraz tváře (úsměv, zamračení). Na obličeji se definuji uzlové body, které se poté propojí a tím definují linie tváře v prostoru, vznikne tím souřadnicová síť obličeje. Samotné rozpoznávání pak probíhá tak, že systém pomocí filtru uzlových bodů reaguje na jednotlivé snímané tváře a může je pak porovnávat a vyhodnocovat. Problémem je přesnost lokalizace orientačních bodů na tváři, řešením může být kombinace s PCA nebo LDA metodou [39].

4.6.1 Preprocessing pomocí Gáborových filtrů

Reprezentace lokálních dat je založena na transformaci obrazu pomocí banky gáborových filtrů. Pro preprocessing obrazu se používá množina orientaci a frekvencí Gáborových filtrů pro každý pixel označovaná jako jety. Jet popisuje malý prostor šedých hodnot v obraze. Je dán vlnkovou transformací, jejíž definice lze nalézt v [40].

Gáborovy filtry jsou použity pro jejich robustnost a jejich souvislost s biologickými daty. Po normalizaci sad filtrů lze dosáhnout dobrých výsledků i při různém kontrastu obrazu.

4.6.2 Reprezentace obličeje

Pro obličej jsou definována zajímavá místa, která tvoří uzly grafu (př. sítnice, koutky úst atd.). Uzly jsou popsány jednotlivými jety a hrany vzdálenostmi mezi jednotlivými uzly.

Grafy pro různé obličeje se liší jednak ve tvaru grafu, ale také v hodnotách uzlů.

4.7 Podpůrné vektory

Podpůrné vektory (Support Vector Machine - SVM) je metoda která spadá do kategorie tzv. jádrových algoritmů (kernel machines). Jedním ze základních principů této metody je převedení problému to vícedimenzionálního prostoru, kde lze oddělit jednotlivé třídy lineárně namísto složité nelineární funkce [10].

K vysvětlení problému je vhodné nejprve zmínit binární klasifikaci, od které se přejde ke klasifikaci více tříd.

4.7.1 Binární klasifikace

Úlohou je nalezení maximální hranice (margin) mezi dvěma třídami, která je definována podpůrnými vektory. Je dána trénovací množina prvků $x \in \mathbb{R}^n$, $i = 1, 2, \ldots, l$, kde každý z prvků náleží do jedné ze dvou tříd označených identifikátorem (label) $y = \{-1, 1\}$. Metoda se snaží oddělit tyto třídy pomocí nadroviny tak, aby vzdálenost mezi nimi byla maximální. Tato nadrovina se nazývá optimální lineární oddělovač a je definován pomocí parametrů α_i daných řešením úlohy:

$$\alpha = \arg\max_{\alpha} \left(\sum_{i=1}^{l} \alpha_i - \frac{1}{2} \sum_{i=1}^{l} \sum_{j=1}^{l} \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) \right),$$

za podmínek

$$\alpha_i \geq 0 \quad , \quad i = 1, 2, \ldots, l$$
$$\sum_{i=1}^{l} = 0 \quad .$$

Koeficienty α a y se hledají pomocí metody kvadratického programování.

Po spočtení optimálních parametrů je je oddělovač dán rovnicí

$$h(x) = \text{sign} \left(\sum_{i} \alpha_i y_i \left(x \cdot x_i \right) \right)$$

Vzhledem k omezením daným při hledání optimálního lineárního oddělovače platí, že body, které jsou k němu nejblíže mají nulové váhy α_i. Takto určené body jsou nazývány podpůrné vektory a definují maximální vzdálenost, která odděluje jednotlivé třídy. Ve většině příkladů je těchto bodů mnohem méně než ostatních bodů trénovací množiny a díky tomu dochází ke zmenšení objemu dat potřebného pro klasifikaci.

Data x_i se vyskytují pouze ve formě skalárního součinu, což umožňuje použít některou z jádrových funkcí $K(x_i, x_j)$. Důležitou jádrovou funkcí je

$$K(x_i, x_j) = (1 + x_i \cdot x_j)^d,$$

která odpovídá prostoru atributů, jehož dimenze roste exponenciálně s d. S použitím takových funkcí je možné nalézt lineární oddělovače efektivně i v prostoru, který má velký počet rozměrů.

Nalezený oddělovač poté lze mapovat do původního prostoru, čímž lze získat nelineární hranici mezi jednotlivými třídami.

4.7.2 Klasifikace více tříd

Pro klasifikaci N tříd se běžně používají dva způsoby [20]:

- Porovnání jedné třídy se všemi pomocí binární klasifikace. V tomto případě je natrénováno N SVM.

- Vytvoření stromové struktury. Jednotlivé klasifikátory jsou uspořádány do stromu, kde každý uzel je SVM. Takto vzniklý strom obsahuje $q(q-1)/2$ SVM.

Běžněji se používá první přístup vzhledem k tomu, že počet SVN je lineární.

4.7.3 Užití pro identifikaci obličeje

Algoritmus lze bez větších obměn použít pro rozpoznání obličeje. Trénovací množina $x \in R^n$ je složena z n-rozměrných vektorů, kde n je počet jednotlivých pixelů obrazu (výška $\times$ šířka), nebo v lepším případě hodnoty eigenfaces. Počet tříd je roven počtu jednotlivých druhů obličejů trénovací množiny.

Co se týče přesnosti klasifikace, podle [19] dosahovala metoda SVM průměrné chyby 3.0% na databázi Cambridge ORL, pro srovnání standardní klasifikační metoda pro eigenfaces, Nearest Center Clasification (NCC), dosahovala průměrné chyby 5.85% .

4.8 Propojení jednotlivých metod

Vzhledem k nedostatkům jednotlivých metod jsou řešení navzájem spojována k dosažení lepších výsledků. Běžným řešením je například kombinace s metodou PCA, která je použita k redukci dimenzionality a tím zmenšení objemu používaných dat (např. u metody LDA).

Kapitola 5

Implementace zvolených metod

Cílem implementační části této práce je získat praktické zkušenosti se zvolenými metodami rozpoznání obličeje, které byly popsány v části 4. Jako zkoumané metody byly zvoleny Analýza hlavních komponent a Lineární diskriminační analýza vzhledem k jejich rozšířenosti a snadné implementaci. Kromě jejich použití je zkoumáno chování metod při nastavení různých parametrů a při různých vstupních datech. Jako vstupní data jsou použity standardní sady obličejů, které jsou volně dostupné (viz 5.1.2). Po prozkoumání chování jednotlivých metod jsou následné obě metody porovnány co do úspěšnosti rozpoznání zadaných obličejů.

Metody jsou implementovány s použitím jazyka C++ a v tomto oboru často používané knihovny OpenCV, která usnadňuje základní operace s obrazovými daty. Jako vývojové prostředí je použito Visual C++ 2008 Express Edition.

5.1 Použité zdroje a knihovny

5.1.1 OpenCV

OpenCV je open source knihovna učená k zjednodušení práce v oboru počítačového vidění. Knihovna je napsána v jazyce C a C++ a je použitelná pod operačními systémy Linux, Windows a Mac OS X. Ve vývoji jsou rozhraní pro Ruby, Matlab, Python a další jazyky.

Knihovna je vyvíjena pod licencí BSD, což znamená, že knihovna může být zdarma používána v neziskových i komerčních aplikacích bez dodatečných podmínek (např. zveřejnění kódu či poskytnutí aplikace zdarma) za předpokladu uvedení autora a informace o licenci, spolu s upozorněním na zřeknutí se odpovědnosti za dílo.

OpenCV je zaměřena na výpočetní efektivnost se silným zaměřením na real-time aplikace. Knihovna je napsána v jako optimalizovaný kód C, tak aby mohla využít výhod vícejádrových procesorů. Vzhledem k tomu, že byla knihovna původně vyvíjená společností Intel, dokáže pracovat s Intel Performance Primitives (IPP), pokud jsou přítomny v systému.

OpenCV se dělí na čtyři části, které rozlišují jednotlivé funkce knihovny. Jde o:

cv - zpracování obrazu a počítačové vidění. Funkce této části jsou filtrování obrazu, geometrické a další transformace obrazu, histogramy, detekce rysů, analýza pohybu a sledování objektů, planární dělení, detekce objektů, kalibrace snímacího zařízení a 3D rekonstrukce.

cxcore - jádro funkcionality. Tato část má na starosti práci se základními strukturami, operace s poli, dynamickými strukturami, kreslicí funkce, ukládání dat jako XML/YAML, shlukování a hledání ve vícedimenzionálním prostoru.

highgui - práce s uživatelským rozhraním a vstupem/výstupem dat, tedy čtení a zápis statických obrázků a videa.

ml - podknihovna strojového učení.

Informace o všech obsažených funkcích knihovny a způsob jejich implementace lze nalézt v [1] nebo [6].

Počítačové vidění je většinou spojeno s metodami strojového učení, které jsou součástí mnoha funkcí OpenCV. K jejich podpoře slouží již zmíněná podknihovna Machine Learning Library (MLL), která mimo jiné podporuje:

Boosting - strojové učení založené na spojování slabých klasifikátorů za účelem vytvoření silnějšího klasifikátoru.

Rozhodovací strom - metoda strojového učení, která dělí vstupní data do tříd podle daných atributů.

Expectation-maximization algoritmus - metoda pro odhad nepozorovaných veličin.

Algoritmus k-nejbližších sousedů - algoritmus určený pro rozpoznání vzorů.

Naivní Bayesovský klasifikátor - pravděpodobnostní klasifikace založená na užití Bayesova teorému.

Neuronové sítě - výpočetní model napodobující biologické struktury určený pro distribuované paralelní zpracování dat.

Náhodný les - soubor mnoha rozhodovacích stromů, jejichž výstupem je modus tříd, které jsou výstupem jednotlivých stromů.

Podpůrné vektory (SVM) - nástroje pro klasifikaci a regresi, viz 4.7

Detekce obličeje

Co se týče detekce obličeje v obraze, knihovna OpenCV obsahuje již existující, dobře dokumentované řešení, které je zveřejněno na internetových stránkách knihovny[1].

[1] Face Detection using OpenCV - http://opencv.willowgarage.com/wiki/FaceDetection

Toto řešení je v souboru facedetect.c, který je distribuován společně s knihovnou. Detekce je založena na přístupu pomocí kaskády Haarových filtrů. Tento přístup je založen na kódování orientovaných kontrastů mezi regiony v obraze. Množina těchto rysů umožňuje zaznamenat zvláštnosti lidského obličeje.

5.1.2 Vstupní data

Za účelem vývoje a testování funkčnosti rozpoznání obličeje existuje několik zdrojů, které poskytují data vhodná pro dané požadavky. Výhodou užití těchto databází je možnost porovnání výsledků s dalšími pracemi podobného zaměření. Před použitím některé databáze je vhodné uvážit, které vlastnosti budou zkoumány, protože zdroje se liší nejen počtem zaznamenaných obličejů, ale i podmínkami při kterých byly obličeje vyfoceny, tzn. různý úhel záběru, světelné podmínky, výřez obličeje apod.

Mezi nejznámější zdroje patří [16]:

AT&T "The Database of Faces" - databáze používaná v této práci. Zahrnuje 40 různých subjektů, každý o deseti různých pózách oproti kameře. Obličeje se liší úhlem pohledu, různými emocemi, světelnými podmínkami a různými detaily (brýle na obličeji). Všechny obličeje jsou snímané proti stejnému pozadí.

Extended M2TVS Database - soubor obrazů obličejů pořízených na University of Surrey ve Velké Británii. Obsahuje snímky 295 subjektů, které byly foceny v průběhu 4 měsíců. Každý subjekt v několika variacích a několika polohách. Kromě barevných obrázků obličejů ve vysoké kvalitě databáze obsahuje zvukové soubory, video sekvence a 3D modely.

The Yale Face Database B - tato databáze byla pořízena za účelem testování algoritmů rozpoznávání obličeje za rozdílných světelných podmínek a pozic oproti kameře. Každý z deseti subjektů byl nasvícen za 64 různých světelných podmínek při devíti úhlech natočení oproti kameře, což dohromady znamená 5850 různých snímků.

The Color FERET Database, USA - databáze programu FERET je rozsáhlým zdrojem obličejových dat, která byla sesbírána nezávisle na vývoji algoritmů rozpoznání obličeje. Snímky byly pořizovány v průběhu několika let, což znamená různé vlastnosti okolního prostředí v době pořizování. Databáze se skládá z 1564 sad snímků, které dohromady dávají 14126 snímků obličejů 1199 různých osob.

Mezi další používané databáze obličejů patří SCface - Surveillance Cameras Face Database, PIE Database, Project - Face In Action (FIA) Face Video Database, Cohn-Kanade AU Coded Facial Expression Database a další.

5.2 Řešení

Předmětem řešení je implementace zvolených metod a experimentální prokázání jejich silných a slabých vlastností. Výsledky jsou poté porovnány, tak aby bylo patrné v jakých oblastech která metoda vyniká.

5.2.1 Implementace Analýzy hlavních komponent

Použité funkce

Obecně je proces definován v části 4.2, co se týče implementace, Analýza hlavních komponent je v knihovně OpenCV od verze 2.0 podporována jako taková. K implementaci slouží třída PCA, která zajišťuje nalezení vlastních hodnot a vektorů při trénování. Poté je možné pomocí funkcí této třídy převádět zadané obrazy do prostoru s redukovanou dimenzionalitou a zde je zkoumat a poté pomocí zpětné projekce opět převést na obraz, který se původnímu v závislosti na počtu použitých hlavních komponent více či méně podobá.

Objekt, který podporuje práci s metodou PCA, se vytváří pomocí konstruktoru:

```
PCA pca(Mat pcaset, Mat mean, int flags, int maxComponents);
```

kde pcaset je trénovací množina dat, mean je průměrný obraz (je možno jej nechat spočítat, pokud zadáme obraz jako prázdnou matici Mat()), flags je definice způsobu uložení dat ve vstupní množině dat a maxComponents je maximální počet hlavních komponent, který bude při výpočtu použit.

Po vytvoření objektu je možné pracovat například s použitím funkce

```
void PCA::project(Mat source, Mat destination);
```

která promítá danou matici do prostoru určeného pomocí extrakce rysů metody PCA. source označuje matici se vstupními daty, tedy daný obraz (uložený jako řádkový vektor) a destination označuje matici, která ukládá výsledek redukce dimenzionality. Tato matice musí mít stejný počet sloupců jako je počet vlastních hodnot. Další potřebnou funkcí je

```
void PCA::backProject(Mat source, Mat destination);
```

který promítá zvolený zpět do původního prostoru a vytváří z něj tedy opět obraz. Věrnost rekonstruovaného obrazu závisí na parametrech metody PCA.

Zkoumané vlastnosti

Počet hlavních komponent

Poté co je vytvořena množina vlastních vektorů. je možné každý obraz promítnout do prostoru s redukovanou dimenzionalitou. Tím se zadaný obraz transformuje na vektor násobků vlastních vektorů (v tomto případě eigenfaces). Tyto vektory lze poté zpětnou projekcí

transformovat opět do původního prostoru tím, že jsou složeny průměrný obraz a vlastní vektory v určitém poměru.

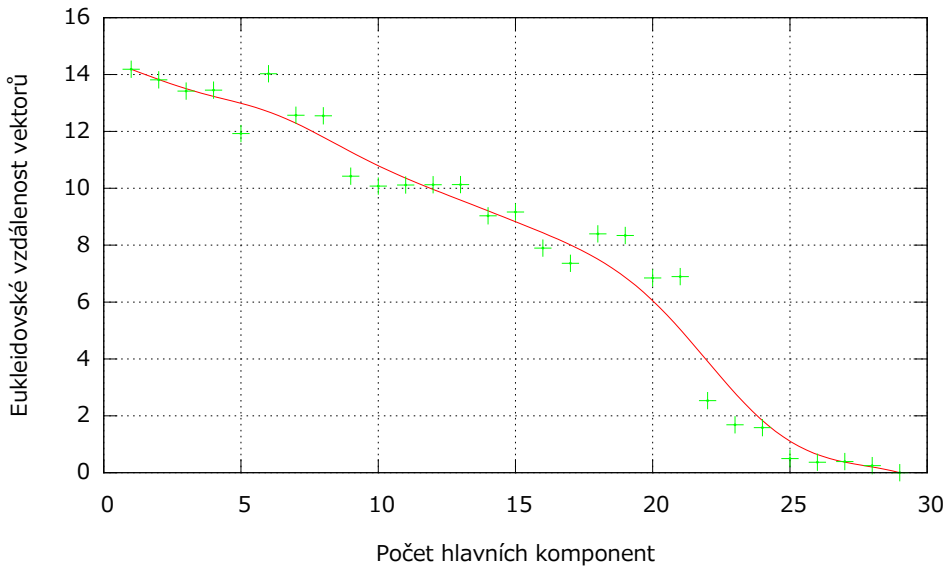
Obraz získaný zpětnou projekcí je závislý na počtu hlavních komponent, což bylo předmětem tohoto zkoumání. Trénovací množina pro PCA byla složena ze 30ti snímků čtyř různých lidí a jednoho snímku, který byl testován. Daný snímek byl zobrazen pomocí projekcí s různým počtem hlavních komponent (1 až 29) a byl zkoumán součet eukleidovských vzdáleností odpovídajících si hodnot vektorů (resp. eukleidovská norma) daný jako

$$\mathrm{d}(\mathbf{p},\mathbf{q}) = \sqrt{(p_1-q_1)^2+(p_2-q_2)^2+\cdots+(p_n-q_n)^2} = \sqrt{\sum_{i=1}^{n}(p_i-q_i)^2},$$

nebo jinak

$$\|\mathbf{p}-\mathbf{q}\| = \sqrt{(\mathbf{p}-\mathbf{q})\cdot(\mathbf{p}-\mathbf{q})} = \sqrt{\|\mathbf{p}\|^2+\|\mathbf{q}\|^2-2\mathbf{p}\cdot\mathbf{q}},$$

kde $p = \{p_1, p_2, \ldots, p_n\}$ je obraz před projekcí a $q = \{q_1, q_2, \ldots, q_n\}$ je obraz po zpětné projekci.



Obrázek 5.1: Závislost počtu komponent na přesnosti zpětné projekce obličeje

Z grafu je viditelná závislost počtu hlavních komponent na přesnosti zpětného zobrazení vstupního obrazu. Se zvyšujícím se počtem hlavních komponent se zmenšoval rozdíl mezi vstupním a výstupním obrazem až na hodnotu $8.94809 \cdot 10^{-6}$, což lze téměř považovat za totožné obrazy. Dále je zřejmé, že ne vždy zvětšení počtu hlavních komponent znamenalo zvýšení přesnosti. Ze subjektivního pohledu, při 22 hlavních komponentách byl výstupní obličej věrný vstupnímu obličeji a při 14 hlavních komponentách byly ještě patrné rysy vstupního obrazu u výstupního. Otázkou dalšího zkoumání je závislost počtu hlavních komponent na přesnosti rozlišení jednotlivých tříd obličejů.

Rozpoznání

Co se týče rozpoznání pomocí metody PCA, je problém podrobně popsán v části 4.2. Vstupní data jsou promítnuta do prostoru rysů, který je vytvořen pomocí trénovacích dat, a zde jsou porovnána pomocí metody nejbližšího souseda, což znamená, že záznam z trénovací množiny, který má nejmenší vzdálenost od rozpoznávaného obrazu v prostoru rysů je klasifikován jako prvek dané třídy obrazů.

Jak již bylo popsáno v 5.2.1, největší důraz je kladen na užití třídy

```
void PCA::project(Mat source, Mat destination); ,
```

která obstarává zapouzdření proměnných používaných pro příslušné operace (vlastní vektory - eigenfaces a vlastní hodnoty). Zobrazení vstupních dat do prostoru daného objektem vytvořeného podle třídy `PCA` získáme pomocí funkce

```
void PCA::project(Mat source, Mat destination);
```

tak vytvoříme z trénovací množiny dat množinu, která bude reprezentovat jednotlivé třídy obličejů. Posléze je stejná funkce použita na testovací data a spočítána vzdálenost promítnutého záznamu z testovací množiny oproti všem záznamům reprezentujících třídy obličejů. Eukleidovská vzdálenost záznamů se počítá pomocí funkce

```
double norm(Mat v1, Mat v2, int flags); ,
```

kde v1 a v2 jsou porovnávané vektory a flags jsou volby, podle kterých se bude vzdálenosti počítat. V této aplikaci je flags nastaveno jako `NORM_L2`, což je označení pro Eukleidovskou normu (vzdálenost). Nejmenší vzdálenost, která je menší než předem zvolený práh je označena jako určitá třída obličejů. Pokud je nejmenší hodnota vzdálenosti větší než práh, je vstupní obličej označen jako neznámý.

Vzhledem k tomu, že rozpoznání vyžaduje různorodější množinu dat, byla pro testování metody zvolena kompletní množina obličejů z databáze AT&T "The Database of Faces", tedy 400 obličejů 40ti různých osob. Jako trénovací množina pro rozpoznání byl volen různý počet obličejů z dané třídy.

Na daných datech byla testována závislost na složení trénovací množiny a závislost rozpoznání na úhlu natočení. Pro pozorování vlivu osvětlení na přesnost rozpoznání není databáze AT&T vhodná vzhledem k tomu, že všechny obličeje jsou pořízeny při totožném osvětlení.

Pro zjištění vhodných parametrů, bylo nejprve rozpoznání vyzkoušeno na třech třídách a dvou trénovacích obličejích. U této malé množiny se u úkolu zjištění závislosti orientace a velikosti obličeje se projevily některé slabiny PCA a to vysoká citlivost na zvolená trénovací data. Pokud byla zvolena trénovací data, která byla sice od totožné osoby, ale byla od ostatních prvků třídy výrazně odlišná v orientaci či postavení obličeje, nebyl systém schopný identifikovat shodu.

Konkrétní příklad problému metody PCA při rozpoznání obličeje je ukázán na obrázku 5.2, kde je obličej natrénován na obrázcích, kde je pohled na kameru v opačném úhlu než testovací snímky. V teoretické části bylo v tomto případě navrženo jako řešení doplnění zrcadlové obráceného snímku do trénovací množiny.

Obrázek 5.2: Případ vzorků, které nebyly správně klasifikovány pomocí metody PCA. Obrázky (a), (b) byly určeny jako trénovací množina. Obrázky (c) a (d) se liší od trénovací množiny natočením obličeje vůči kameře, ale částečně také emocemi v obličeji, což jsou jedny z problémů, které jsou řešeny v případě potřeby zpřesnění rozpoznání.

K dosažení přesnějších výsledků bylo zvoleno více trénovacích a testovacích množin. Postupně byla metoda PCA testována na pěti, deseti a dvaceti třídách obličejů. Trénovací množina obsahovala 3 nebo 5 prvků. Pro dosažení lepších výsledků při klasifikaci obličejů lišících se postavením oproti kameře byla jako úprava obrázků zvolena možnost zrcadlového otočení snímků, což bylo aplikováno na polovinu případů.

Počet tříd	N_L	N_T	Úpravy	Přesnost (%)
5	15	50	-	83.67%
5	30	50	Zrcadlení	87.75%
5	25	50	-	93.87%
5	50	50	Zrcadlení	95.91%
10	50	100	-	94.94%
10	100	100	Zrcadlení	96.96%
20	100	200	-	83.91%
20	200	200	Zrcadlení	84.92%
40	200	400	-	73.68%
40	400	400	Zrcadlení	75.93%

Tabulka 5.1: Tabulka srovnávající přesnost metody PCA v závislosti na vstupních datech a jejich úpravě. Počet tříd - počet tříd obličejů, se kterými se pracuje, N_L - Počet prvků trénovací množiny, N_T - Počet prvků testovací množiny, Úpravy - Změny na trénovací množině před samotným rozpoznáním.

Výsledky testování lze vidět v tabulce 5.1. Je zřejmé, že rozsáhlejší trénovací množina

vede k větší přesnosti u identifikace obličeje. Nejlepší přesnost, která byla dosažena, byla 95.91% správně klasifikovaných obličejů, což bylo dosaženo u identifikace pěti tříd obličejů. Průměrná přesnost bez úpravy vstupních dat byla 86.01%. Doplnění vstupních dat o zrcadlově otočené obličeje vedlo u většiny případů k mírnému zlepšení.

Dalším krokem bylo otestování metody PCA vzhledem ke schopnosti rozpoznávat obličeje v různých světelných podmínkách. Výše je zmíněno, že databáze AT&T "The Database of Faces"není k tomuto účelu vhodná, proto byla použita databáze The Yale Face Database B. Tato databáze obsahuje snímky, které se liší nejen světelnými podmínkami, ale i úhlem natočení ke kameře (8 různých pohledů). Vzhledem k tom, že se v tomto kroku testuje jen vliv světelných podmínek, byl použit jen pohled přímo do snímacího zařízení. Ostatních 7 pohledů nebylo do testování zahrnuto. Co se týče úhlů nasvícení, každý z osmi pohledů je nafocen při 65ti různých podmínkách. Všechny snímky subjektu jsou ve stejné poloze a mají stejný výraz.

Vzhledem k tomu, že snímky jsou foceny v rozlišení 640x480, bylo nutné snímky ořezat a zmenšit, což bylo provedeno manuálně. Jako cílené rozlišení bylo zvoleno rozlišení databáze AT&T, tedy 92x112 pixelů. Vzhledem k tomu, že snímané subjekty jsou vyfoceny vždy tak, že obličej leží ve stejném místě fotografie, nebyla nutná detekce, ale každá třída snímků byla ořezána podle stejného vzoru.

Z počtu 65 snímků pro každý subjekt bylo náhodně zvoleno 12, které byly pro každou osobu testovány. Trénovací množiny obsahovaly 1, 3 a 6 obličejů. Obličeje byly v některých případech doplněny jejich zrcadlovými obrazy podobně jako u předchozí úlohy. Výsledné hodnoty jsou zobrazeny v tabulce 5.2.

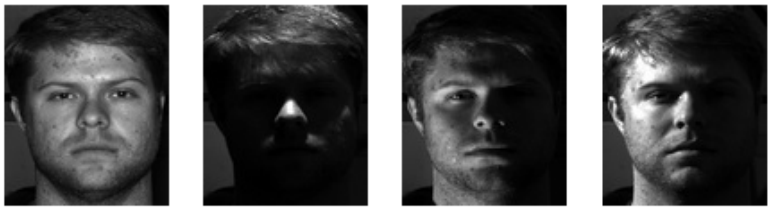
Počet tříd	N_L	N_T	Úpravy	Přesnost (%)
5	5	60	-	83.05
5	10	60	Zrcadlení	88.13
5	15	60	-	89.83
5	30	60	Zrcadlení	88.13
5	30	60	-	96.61
5	60	60	Zrcadlení	96.61

Tabulka 5.2: Tabulka srovnávající přesnost metody PCA při rozpoznávání obličejů databáze The Yale Face Database B při různých světelných podmínkách. Označení sloupců (Počet tříd, N_T, N_U) má stejný význam jako u tabulky 5.1.

Z výsledku je patrná vzrůstající úspěšnost správného rozlišení tříd obličejů při větší trénovací množině. Tato skutečnost potvrzuje návrhy k řešení problému nejednotného osvětlení vznesení v teoretické části, konkrétně 3.1.3. Je zřejmé, že čím více různých úhlů osvětlení trénovací množina obsahuje, tím je rozpoznání přesnější.

Konkrétní příklady záznamů použité databáze jsou zobrazeny na obrázku 5.3. Obličej

5.3(a) je pohled zepředu použitý v trénovací množině, 5.3(b) byl obsažen pouze ve třetí testovací množině, pro daný případ bylo dosaženo nejpřesnější hodnoty rozpoznání. Obličeje 5.3(c) a 5.3(d) nebyly součástí žádné trénovací množiny a rovněž nebyly pro žádné měření rozpoznány vzhledem k odlišnosti světelných podmínek u ostatních obličejů třídy.



Obrázek 5.3: Příklady obličejů z databáze The Yale Face Database B za různých světelných podmínek.

Co se týče srovnání vlivu natočení obličeje vůči kameře a osvětlení, z měření vyplývá, že metoda PCA je citlivější na postavení obličeje ke kameře než k osvětlení. Z hlediska zvýšení přesnosti, je podle tohoto měření vhodnější zaměřit se na postavení obličeje, protože má větší vliv.

Vliv doplnění trénovací množiny o zrcadlově otočené snímky se ukázal pozitivní pouze u malé trénovací množiny.

Poslední testovací množinou, která byla použita je databáze XM2VTSDB (The Extended M2VTS Database). Tato databáze je vhodná pro testování velkého počtu tříd obličejů, vzhledem k tomu, že obsahuje 296 subjektů. Obličeje jsou zachyceny ve vysokém rozlišení (720x576) v barvě.

Oblast obličeje v databázi není normalizovaná, takže je nutné před samotnou identifikací nalézt region, kde se obličej nachází. K tomu lze použít některé z algoritmů detekce obličeje, které jsou součástí OpenCV. Součástí databáze jsou však i data, která obsahují již detekované souřadnice očí, pomocí kterých lze region obličeje ze snímku extrahovat. Za tímto účelem bylo použito znalostí ze standardu MPEG-7, který udává proporce, které má normalizovaná mřížka, která popisuje obličej.

Souřadnice očí jsou v souborech uložena v následujícím formátu:

```
aaa_b_c ddd eee fff ggg
```

kde jednotlivé položky mají význam:

aaa - identifikátor třídy obličejů, který označuje příslušný subjekt.

b - pořadí (od 1 do 4), v kterém byly snímky nafoceny.

c - druh snímku.

ddd - x-ová souřadnice levého oka (l_x).

eee - y-ová souřadnice levého oka (l_y).

fff - x-ová souřadnice pravého oka (r_x).

ggg - y-ová souřadnice pravého oka (l_y).

Pomocí těchto znalostí byl soubor zanalyzován a byly z něj extrahovány příslušné hodnoty použité dále k určení regionu.

Důležitou znalostí, jak již bylo zmíněno v teoretickém popisu detekce regionu obličeje je vzdálenost očí, která se určí jako:

$$dist = \sqrt{(r_x - l_x)^2 + (r_y - l_y)^2} \tag{5.1}$$

Podle standardu pro obličej určíme délku a šířku:

$$W = \frac{dist}{15} \cdot 56 \quad \text{a} \quad H = \frac{dist}{15} \cdot 46 \tag{5.2}$$

Dále je potřeba znát střed spojnice očí, podle kterého se nalezne poloha levého horního rohu regionu obličeje.

$$S = L + \frac{R - L}{2} \quad \text{, kde} \quad L = [l_x, l_y] \; a \; R = [r_x, r_y] \tag{5.3}$$

Odtud získáme levý horní roh (LT):

$$LT_x = S_x - \frac{W}{2} \quad \text{a} \quad LT_y = S_y - \frac{H \cdot 24}{56} \tag{5.4}$$

Z matice, do které načteme obrázek vyřízneme oblast danou rovnicemi (5.2) a (5.4) pomocí určení ROI[2] a jeho zkopírování do jiné matice. Problém zde tvoří obličeje, u kterých nejsou oči dostatečně vysoko oproti spodnímu okraji obrazu. U takových se y-ová pozice levého horního rohu oblasti upravuje tak, aby byla alespoň ve vzdálenosti H od spodního okraje obrazu.

Takto získaný obraz se poté zmenší na rozměry 92x112px, což je rozměr, s kterým se pracovalo u předchozích testů. Tímto lze získat vzorky s oblastí, která nebude zahrnovat příliš velkou oblast pozadí, které by negativně působilo na úspěšnost identifikace.

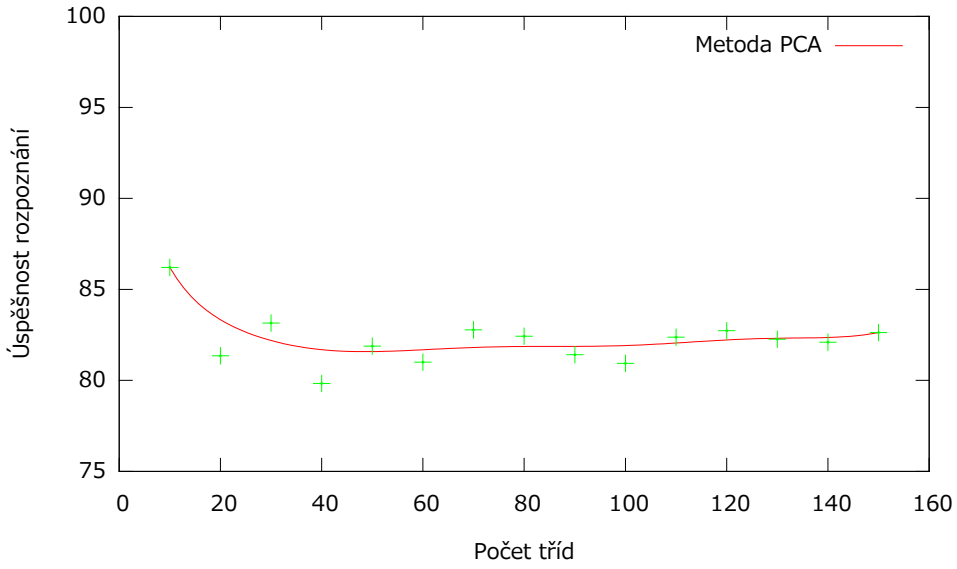
Výsledná množina obrázků byla poté rozdělena na trénovací a testovací množinu, z nichž každá obsahovala 3 obličeje stejné osoby s tím, že snímky v trénovací a testovací množině byly různé. Poté byla testována úspěšnost rozpoznání na 10 až 150 třídách obličejů z celkových 195 (pro zbylých 45 nebyly dostupné souřadnice očí). Výsledky jsou uvedeny v tabulce 5.3.

Z výsledků je patrná klesající úspěšnost identifikace se vzrůstajícím počtem rozpoznávaných tříd obličejů. Závislost je zobrazena v grafu 5.4. Je rovněž viditelná malá úspěšnost při relativně malých velikostech trénovací a testovací množiny. Při přibližně u 60ti rozpoznávaných třídách začala úspěšnost kolísat kolem hodnoty 82%.

[2]Region of Interest - oblast zájmu, tj. podoblast daného objektu, se kterou se provádí další operace

Počet tříd	N_L	N_T	Úpravy	Přesnost (%)
10	30	30	-	86.20
20	60	60	-	81.35
30	90	90	-	83.14
40	120	120	-	79.83
50	150	150	-	81.87
60	180	180	-	81.00
70	210	210	-	82.77
80	240	240	-	82.42
90	270	270	-	81.41
100	300	300	-	80.93
110	330	330	-	82.37
120	360	360	-	82.72
130	390	390	-	82.26
140	420	420	-	82.10
150	450	450	-	82.62

Tabulka 5.3: Tabulka srovnávající přesnost metody PCA při rozpoznávání obličejů databáze Extended M2TVS Database. Označení sloupců (Počet tříd, N_T, N_U) má stejný význam jako u tabulky 5.1.



Obrázek 5.4: Zobrazení klesající úspěšnosti identifikace použitím metody PCA při zvyšujícím se počtu tříd. Metoda je testována na databázi Extended M2TVS Database.

Návrhy pro dosažení vyšší přesnosti

Z výše zmíněných výsledků je patrných několik oblastí, na které je vhodné se při snaze o zpřesnění identifikace obličeje zaměřit.

Metoda není příliš robustní, pokud jde o světelné podmínky, je proto vhodné zahrnout od trénovací množiny co největší počet obličejů stejné třídy nasvícených z různých úhlů, použití některé metody normalizace osvětlení uvedené v 3.1.3, nebo se používá možnost odstranění třech nejvýraznějších hlavních komponent, které obsahují nejvíce informací osvětlení.

A druhým problémem je zpracování snímků obličejů, které nejsou přímo proti kameře. V případě menších úhlů je možné obraz transformovat, u větších úhlů je vhodné použít větší trénovací množinu.

5.2.2 Implementace Lineární diskriminační analýzy

Metoda rozpoznání obličeje pomocí Lineární diskriminační analýzy není na rozdíl od Analýzy hlavních komponent implementována v samotné knihovně, proto bylo nutné je napsat.

Byla použita varianta LDA, kdy je u vstupních dat redukována dimenzionalita. Za tímto účelem byla použita metoda PCA. Výpočet se takto výrazně urychlil.

Použité funkce

Dle teoretického rozboru (viz 4.3) se vycházelo z použití dvou matic rozptylu, které popisovaly vztahy mezi prvky uvnitř třídy (S_W) resp. mezi jednotlivými třídami (S_B).

Funkce pro rozpoznání obličeje byla definována jako:

```
void doLDA(Mat pcaset, Mat testset, int type, Size imsize,
                       Mat correctTrain, Mat correctTest);
```

Vstupem tedy jsou 2 matice, určené pro trénovací a testovací množinu, jejichž řádky tvoří řádkové vektory obsahující data jednotlivých snímků obličejů. Dále jsou vstupem rozměry původního snímku, typ dat matice a hodnoty příslušnosti řádků matic s daty k jednotlivým třídám.

Jako příprava se nejprve vytvoří struktury, které jsou použity pro výpočet jednotlivých matic:

```
std::set<int> classes;
std::map<int, Mat> classMeans;
```

Množina se naplní pomocí matice `correctTrain`, tím se zjistí počet tříd (je vhodné použít množinu, protože nemusí platit, že v testovací množině jsou třídy zastoupeny vždy ve stejném počtu). Poté se k příslušné třídě přiřadí její průměr, což se uloží jako mapa `classMeans` kvůli snadnějšímu přístupu.

S takto připravenými daty lze přistoupit k výpočtu matic S_W a S_B. Nejprve je vypočtena matice rozptylu mezi třídami tak, že pro všechny prvky v množině tříd je matice vypočtena jako:

```
mulTransposed(classMeans[pos], tmtx, 1, mev, 1, -1);
tmtx = elemCount[pos]* tmtx;
```

`pos` v tomto případě určuje aktuální třídu, tedy `classMeans[pos]` je průměr této třídy, dále `elemCount[pos]` označuje počet prvků třídy, mev je průměr všech prvků a `tmtx` je dočasná matice, která se posléze přičte k hodnotě matice S_B. Funkce mulTransposed označuje násobení matic takové, že jeden z prvků je před vynásobením transponován.

```
add(Sb, tmtx, Sb);
```

Po dokončení průchodu množinou je vypočtená výsledná matice S_B.

Matice S_W je vypočítána průchodem všech prvků trénovací množiny. Pro každý prvek je spočtena matice daná funkcí:

```
mulTransposed(inp.row(i), tmtx, 1, classMeans[curClass], 1, -1);,
```

kde `i` je pořadí průchodu (od 1 do počtu prvků trénovací množiny), `inp.row(i)` je jeden z vstupních snímků vyjádřený řádkovým vektorem a ostatní prvky mají stejný význam jako u matice S_B. Tyto dílčí matice jsou sečteny a tím vznikne požadovaná matice.

Vzniklý eigen systém daný rovnicí (4.22) je řešitelný několika způsoby. Jedním z nich je vynásobení rovnice inverzní maticí k S_W tak, že vznikne systém

$$S_W^{-1} S_B w_i = \lambda_i w_i, i = 1, 2, \ldots, m \tag{5.5}$$

Řešení lze nalézt pomocí funkce knihovny OpenCV

```
bool eigen(const Mat& src, Mat& eigenvalues, Mat& eigenvectors,
                           int lowindex=-1, int highindex=-1)
```

kde vstupem (src) je matice získaná vynásobením $S_W^{-1} S_B$. Výsledkem jsou vlastní vektory a hodnoty, které jsou uloženy do matic daných jako parametry eigenvalues a eigenvectors. Tímto jsme získali projekční matici, kterou lze použít pro rozdělení jednotlivých vstupních obrazů do tříd.

Problém nastává, pokud je matice S_W singulární, protože v tomto případě k ní neexistuje inverzní matice. Pokud je tomu tak, matice S_W se neuvažuje a hledá se řešení pouze pomocí matice S_B.

Zkoumané vlastnosti

Co se týče pozorovaných vlastností, byly u LDA aplikovány podobné testy, jako u předchozí metody, tedy PCA. Byla zkoumána přesnost rozpoznání v závislosti na osvětlení, pozici obličeje ke kameře a počtu rozpoznávaných tříd. Výsledné přesnosti jsou zobrazeny ve stejných tabulkách, aby bylo snadnější jejich porovnání v části 5.2.3. Stejně jako u PCA, byla metoda testována na databázích AT&T "The Database of Faces", Extended M2TVS Database a The Yale Face Database B.

První test byl proveden na databázi AT&T "The Database of Faces". Zde byla předmětem detekce obličejů lidí za stejných světelných podmínek, při různém úhlu pohledu (přibližně -5° až +5°) a u některých subjektů s nasazenými brýlemi a bez nich. Snímky v této databázi rovněž nejsou příliš dobře zarovnané (pozice očí nebo nosu nejsou vždy na stejném místě) a není zde jednotný region obličeje, což může některým metodám způsobovat problémy. Výsledkem testování je tabulka 5.4. Z tabulky je patrné, že metoda má při rela-

Počet tříd	N_L	N_T	Úpravy	Přesnost (%)
5	15	50	-	98.00
5	30	50	Zrcadlení	98.00
5	25	50	-	100.00
5	50	50	Zrcadlení	100.00
10	50	100	-	99.00
10	100	100	Zrcadlení	99.00
20	100	200	-	96.50
20	200	200	Zrcadlení	96.50
40	200	400	-	94.50
40	400	400	Zrcadlení	95.50

Tabulka 5.4: Tabulka srovnávající přesnost metody LDA při rozpoznávání obličejů databáze AT&T "The Database of Faces". Označení sloupců (Počet tříd, N_T, N_U) má stejný význam jako u tabulky 5.1.

tivně malém počtu tříd dobrou úspěšnost a zrcadlení má jen malý vliv na zlepšení přesnosti rozpoznání.

Na databázi The Yale Face Database B bylo testován vliv osvětlení na úspěšnost rozpoznání. Byl zvolen jen malý vzorek pěti tříd. Některé množiny trénovacích dat byly doplněny o zrcadlové obrazy. Výsledky jsou uvedeny v tabulce 5.5.

Počet tříd	N_L	N_T	Úpravy	Přesnost (%)
5	5	60	-	88.33
5	10	60	Zrcadlení	90.00
5	15	60	-	88.33
5	30	60	Zrcadlení	88.33
5	30	60	-	96.66
5	60	60	Zrcadlení	96.66

Tabulka 5.5: Tabulka srovnávající přesnost metody LDA při rozpoznávání obličejů databáze The Yale Face Database B při různých světelných podmínkách. Označení sloupců (Počet tříd, N_T, N_U) má stejný význam jako u tabulky 5.1.

Z tabulky vyplývá vzrůstající úspěšnost při zvětšování množiny trénovacích dat. Malé trénovací množiny nedokázaly dokonale postihnout variace v osvětlení. Co se týče doplnění množiny o zrcadlové obrazy, nebyl pozorován výrazný vliv. Při trénovací testovací množiny došlo ke zlepšení přesnosti až na 96.66%.

Poslední množinou dat na které byla tato metoda identifikace obličeje testována byla XM2VTSDB. Tento test byl stejně jako u PCA určen k testování vlivu počtu tříd na přesnost rozpoznání. Postupně byl tedy zvětšován počet tříd trénovací a testovací množiny. Ty byly stejně jako z PCA složeny každá ze třech snímků obličejů, s tím, že obličeje z trénovací množiny nebyly obsaženy v testovací.

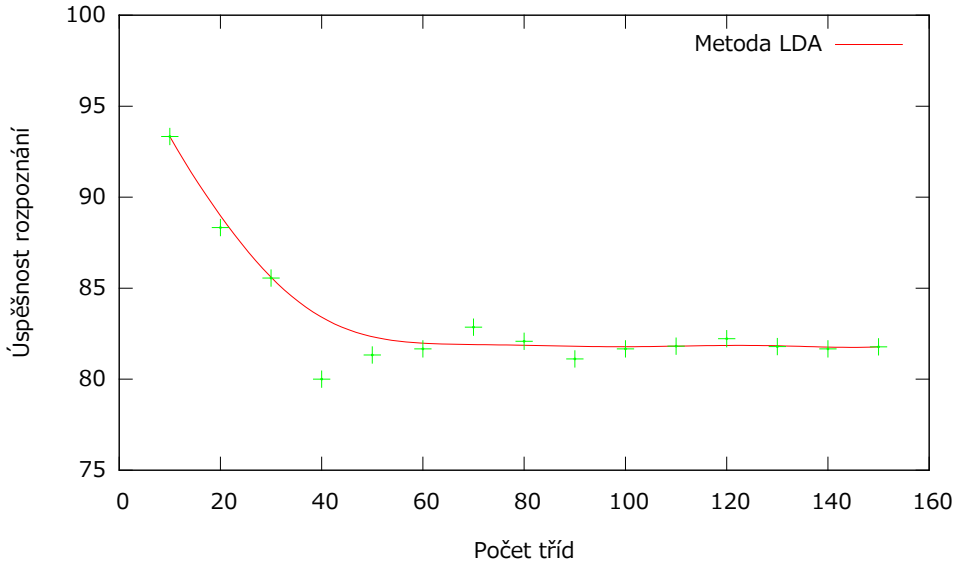
Počet tříd	N_L	N_T	Úpravy	Přesnost (%)
10	30	30	-	93.33
20	60	60	-	88.33
30	90	90	-	85.55
40	120	120	-	80.00
50	150	150	-	81.33
60	180	180	-	81.67
70	210	210	-	82.85
80	240	240	-	82.08
90	270	270	-	81.11
100	300	300	-	81.67
110	330	330	-	81.81
120	360	360	-	82.22
130	390	390	-	81.79
140	420	420	-	81.66
150	450	450	-	81.77

Tabulka 5.6: Tabulka srovnávající přesnost metody LDA při rozpoznávání obličejů databáze Extended M2TVS Database. Označení sloupců (Počet tříd, N_T, N_U) má stejný význam jako u tabulky 5.1.

Z tabulky 5.6 je patrná klesající přesnost identifikace obličeje se zvyšujícím se počtem rozpoznávaných tříd. Přibližně od 50 rozpoznávaných tříd obličejů začne úspěšnost kolísat kolem hodnoty 81.7%. Hodnoty z tabulky jsou zobrazeny v grafu na obrázku 5.5.

5.2.3 Porovnání metod

U obou implementovaných metod (PCA, LDA) byla provedena série testů s důrazem otestování reakce na nenormalizované snímky (postavení obličeje, různý region obličeje, osvětlení) a větší počet tříd. Co se týče průměrných úspěšností v těchto testech, vycházela celkově lépe lépe metoda LDA.

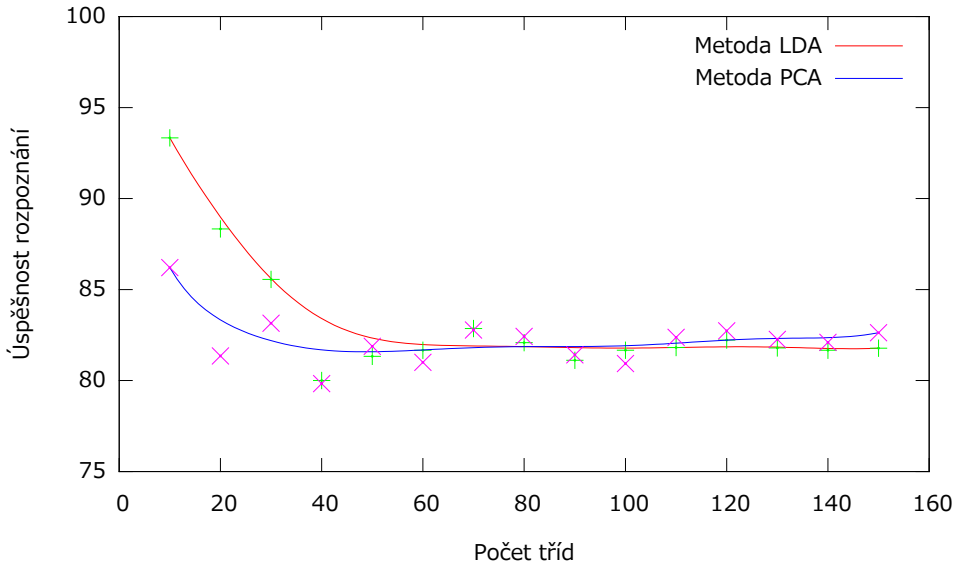


Obrázek 5.5: Zobrazení klesající úspěšnosti identifikace použitím metody LDA při zvyšujícím se počtu tříd. Metoda je testována na databázi Extended M2TVS Database.

Při testech na databázi AT&T "The Database of Faces"dokázala metoda LDA výrazně lépe rozpoznávat data, která nebyla správně normalizována. Metoda PCA měla problémy rozpoznat obličeje, které byly v jiném úhlu než ty, obsažené v trénovací množině. Pomocí LDA v tomto případě bylo dosaženo průměrné úspěšnosti 97.70%, kdežto pomocí metody PCA probíhalo rozpoznávání s úspěšností 87.15%, což není u takto malého počtu tříd nijak velká hodnota.

Co se týče databáze The Yale Face Database B a rozpoznávání obličejů s proměnlivým úhlem osvětlení, bylo rozpoznávání s pomocí metody LDA opět úspěšnější s tím, že dosáhnula průměrné úspěšnosti 91.39% oproti 90.39% u PCA.

Rozpoznávání snímků obličejů z databáze Extended M2TVS Database mělo u obou metod přibližně stejnou průměrnou úspěšnost (83,14% u LDA, 82,20% u PCA) s tím, že LDA dosahovala lepší úspěšnosti u menšího počtu tříd (až 93.33% u LDA oproti 86.20% u PCA). Porovnání implementovaných metod na rozpoznávání snímků z této databáze je zobrazeno v grafu na obrázku 5.6.



Obrázek 5.6: Porovnání metod PCA a LDA na rozpoznávání dat z databáze Extended M2TVS Database. Z grafu je patrná větší úspěšnost metody LDA při menším počtu rozpoznávaných tříd. Při větším počtu jsou obě metody přibližně stejně efektivní.

Kapitola 6

Závěr

V práci byly vysvětleny základní principy identifikace obličeje včetně částečného popisu metod. Výsledná podoba čerpá z informací uvedených v literatuře a odráží současný stav tohoto oboru. Identifikace obličeje je slibný obor, který nabízí uplatnění v mnoha oblastech. Zmíněny jsou i možnosti rozvoje technologie rozpoznání obličeje v budoucnosti.

Co se týče rozpoznání obličeje, byly popsány nejvýraznější úskalí tohoto oboru, tedy rozpoznávání obličejů snímaných za různého osvětlení, rozpoznávání obličejů snímaných z různého úhlu a rozpoznávání obličejů, které nemají normalizovanou oblast obličeje. S přihlédnutím na tyto zkušenosti se vycházelo při implementaci zvolených metod identifikace obličeje, tedy Analýzy hlavních komponent (PCA) a Lineární diskriminační analýzy (LDA).

Ke zjištění chování daných metod byla s použitím knihovny OpenCV vytvořena aplikace která implementovala metody na zvolených datech. Jednalo se o data, která pokrývala veškeré problémové případy, tedy různé výřezy obličejů, různě osvětlené obličeje a různě velké množiny dat. Výchozími databázemi pro učení a testování byly AT&T "The Database of Faces", Extended M2TVS Database a The Yale Face Database B.

Tato aplikace prokázala vyšší průměrnou míru úspěšnosti rozpoznání u metody LDA na všech typech vstupních dat. Metoda PCA byla méně efektivní při použití špatně normalizovaných dat. Co se týče vhodně připravených dat, obě metody byly podobně úspěšné při vyšším počtu rozpoznávaných tříd (50 a více). Obě metody dosahovaly při 150 rozpoznávaných třídách hodnoty přibližně 82% identifikovaných obličejů. Co se týče nerovnoměrného osvětlení obličeje, byl ověřen vliv na úspěšnost identifikace, s tím, že horší výsledky byly pozorovány u PCA. U rozpoznávání obličejů, které byly různě orientované, nebo rozpoznávané na různých výřezech bylo dosaženo horších výsledků než u snímků zachycujících obličeje s různým osvětlením.

Budoucí vývoj v tomto směru by vedl na praktičtější uplatnění identifikace obličeje, spočívající v řešení zmíněných problémů a zaměření na situace v reálném životě, jako rozpoznávání osob z videozáznamů, nebo osobních dokladů. Tyto oblasti uplatnění však provází další problémy, které je potřeba řešit.

Literatura

[1] OpenCV 2.0 C Reference.
http://opencv.willowgarage.com/documentation/index.html.

[2] How CCTV systems are used. *postnote*, ročník 175, 2002: s. 1–4.
URL http://www.parliament.uk/post/pn175.pdf

[3] Bakhtiari, S.; Safavi-naini, R.; Pieprzyk, J.; aj.: Cryptographic Hash Functions: A Survey. Technická zpráva, 1995.

[4] Beymer, D.; Poggio, T.: Face Recognition From One Example View. Technická zpráva, Cambridge, MA, USA, 1995.

[5] Böhm, C.; Berchtold, S.; Keim, D. A.: Searching in high-dimensional spaces: Index structures for improving the performance of multimedia databases. *ACM Comput. Surv.*, ročník 33, č. 3, 2001: s. 322–373, ISSN 0360-0300, doi:http://doi.acm.org/10.1145/502807.502809.

[6] Bradski, G.; Kaehler, A.: *Learning OpenCV*. O'Reilly Media Inc., 2008.
URL http://oreilly.com/catalog/9780596516130

[7] Brecht, B.: Facing the future with 3D facial recognition technology. *Biometric Technology Today*, ročník 17, č. 1, 2009: s. 8 – 9, ISSN 0969-4765, doi:DOI:10.1016/S0969-4765(09)70018-9.
URL http://www.sciencedirect.com/science/article/B6W70-4VND8DS-M/2/0e8c1441d96ace9ae169c9e0be8c2756

[8] Chee Khin George Loo, S.: Method and system for verifying the identity of a user. http://www.google.cz/patents?hl=cs&lr=&vid=USPATAPP11304171&id=1JScAAAAEBAJ&oi=fnd&printsec=abstract#v=onepage&q=&f=false, Prosinec 2005.

[9] Chmelař, P.: *Multimediální databáze*. VUT Brno, FIT, Ústav informačních systémů, 2006.

[10] Cortes, C.; Vapnik, V.: Support-Vector Networks. In *Machine Learning*, 1995, s. 273–297.

[11] Department, S. L.; Leach, S.: Singular Value Decomposition – A Primer.

[12] Deza, M.-M.; Deza, E.: *Dictionary of Distances; electronic version.* Burlington, MA: Elsevier, 2006.

[13] Feng, G. C.; Yuen, P. C.: Variance projection function and its application to eye detection for human face recognition. *Pattern Recognition Letters*, ročník 19, č. 9, 1998: s. 899 – 906, ISSN 0167-8655, doi:DOI:10.1016/S0167-8655(98)00065-8.
URL http://www.sciencedirect.com/science/article/B6V15-3V8TPN7-G/2/33dbff916d8a1a2d089d44c8fea48ccf

[14] Fritsch, L.: *Metoda PCA a její implementace v jazyce C++*. ČVUT v Praze, Fakulta elektrotechnická, Katedra radioelektroniky, 2008.

[15] Fuchs, C.; Heuel, S.: Feature Extraction. 1998.

[16] Gross, R.: Face Databases. In *Handbook of Face Recognition*, editace A. S.Li, New York: Springer, February 2005.

[17] Gross, R.; Brajovic, V.: An Image Preprocessing Algorithm for Illumination Invariant Face Recognition. In *Audio- and Video-Based Biometric Person Authentication*, *Lecture Notes in Computer Science*, ročník 2688, Springer Berlin / Heidelberg, 2003, ISBN 978-3-540-40302-9, ISSN 0302-9743 (Print) 1611-3349 (Online), doi:10.1007/3-540-44887-X_2.
URL http://www.springerlink.com/content/17a5cedhj0ptg6nc/

[18] Gu, L.; Li, S. Z.; Zhang, H.-J.: Learning Probabilistic Distribution Model for Multi-View Face Detection. *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on*, ročník 2, 2001: str. 116, ISSN 1063-6919, doi:http://doi.ieeecomputersociety.org/10.1109/CVPR.2001.990934.

[19] Guo, G.; Li, S. Z.; Chan, K.: Face Recognition by Support Vector Machines. 2000, s. 196–201.

[20] Heisele, B.; Ho, P.; Poggio, T.; a j.: Face Recognition with Support Vector Machines: Global versus Component-based Approach. In *In Proc. 8th International Conference on Computer Vision*, 2001, s. 688–694.

[21] Hinner, J.: Detekce a rozpoznávání obličejů osob a jejich identifikační význam. http://web.mvcr.cz/archiv2008/casopisy/kriminalistika/2003/03_01/hinner.html, 2003.

[22] Hjelmas, E.; Low, B. K.: Face Detection: A Survey. *Computer Vision and Image Understanding*, September 2001: s. 236–274, ISSN 1077-3142.
URL http://www.ingentaconnect.com/content/ap/iv/2001/00000083/00000003/art00921

[23] Hsu, R.; Abdel-Mottaleb, M.; Jain, A.: Face Detection in Color Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, ročník 24, 2002: s. 696–706, ISSN 0162-8828, doi:http://doi.ieeecomputersociety.org/10.1109/34.1000242.

[24] Jain, A. K.; Ross, A.; Prabhakar, S.: An introduction to biometric recognition. *IEEE Trans. on Circuits and Systems for Video Technology*, ročník 14, 2004: s. 4–20.

[25] Jebara, T. S.; Jebara, T. S.: 3d Pose Estimation And Normalization For Face Recognition. Technická zpráva, 1996.

[26] Juels, A.; Molnar, D.; Wagner, D.: Security and Privacy Issues in E-passports. In *Security and Privacy for Emerging Areas in Communications Networks, 2005. SecureComm 2005. First International Conference on*, Sept. 2005, s. 74–88, doi:10.1109/SECURECOMM.2005.59.

[27] Kapralos, B.; Jenkin, M.; Tsotsos, J.; aj.: Eyes 'n ears: A system of attentive teleconferencing. *The Journal of the Acoustical Society of America*, ročník 107, č. 5, 2000: s. 2790–2790, doi:10.1121/1.428975.
URL http://link.aip.org/link/?JAS/107/2790/1

[28] Levine, M.: Feature extraction: A survey. *Proceedings of the IEEE*, ročník 57, č. 8, Aug. 1969: s. 1391–1407, ISSN 0018-9219.

[29] Martinez, A. M.; Mart'inez, A. M.: Recognizing Imprecisely Localized, Partially Occluded and Expression Variant Faces from a Single Sample per Class. 2002.

[30] Moses, Y.; Adini, Y.; Ullman, S.: Face recognition: The problem of compensating for changes in illumination direction. In *Computer Vision ÔÇö ECCV '94*, *Lecture Notes in Computer Science*, ročník 800, Springer Berlin / Heidelberg, 1994, ISBN 978-3-540-57956-4, ISSN 0302-9743 (Print) 1611-3349 (Online), s. 286–296, doi:10.1007/3-540-57956-7_33.
URL http://www.springerlink.com/content/p4772g7t10n33833/

[31] Moving Picture Experts Group (MPEG): MPEG-7 Overview (version 10). http://mpeg.chiariglione.org/standards/mpeg-7/mpeg-7.htm, 2004.

[32] Prinosil, J.; Vlach, J.: Face detection in image with complex background. In *Personal Wireless Communications, IFIP International Federation for Information Processing*, ročník 245, Springer Boston, 2007, ISBN 978-0-387-74158-1, ISSN 1571-5736 (Print) 1861-2288 (Online), s. 533–544, doi:10.1007/978-0-387-74159-8_53.
URL http://www.springerlink.com/content/v871l6331t565m48/

[33] Riaz, Z.; Beetz, M.; Radig, B.: Shape invariant recognition of segmented human face images using eigenfaces. In *Multitopic Conference, 2008. INMIC 2008. IEEE International*, Dec. 2008, s. 587–591, doi:10.1109/INMIC.2008.4777807.

[34] Riaz, Z.; Gilgiti, A.; Mirza, S.: Face recognition: a review and comparison of HMM, PCA, ICA and neural networks. In *E-Tech 2004*, 31 2004, s. 41 – 46, doi:10.1109/ETECH.2004.1353842.

[35] Ručkay, L.: *Analýza nezávislých komponent*. 2005.

[36] Sato, H.; Sakamoto, K.; Mitsukura, Y.; aj.: Face Edge Detection System by Using the GAs. In *Knowledge-Based Intelligent Information and Engineering Systems*, *Lecture Notes in Computer Science*, ročník 3213, Springer Berlin / Heidelberg, 2004, ISBN 978-3-540-23318-3, ISSN 0302-9743 (Print) 1611-3349 (Online), s. 847–852, doi:10.1007/b100909.
URL http://www.springerlink.com/content/e85r0l6w2fxa97lh/

[37] Shih, F. Y.; Chuang, C.-F.: Automatic extraction of head and face boundaries and facial features. *Inf. Sci. Inf. Comput. Sci.*, ročník 158, č. 1, 2004: s. 117–130, ISSN 0020-0255.

[38] Smith, L. I.: A Tutorial on Principal Component Analysis.
URL http://kybele.psych.cornell.edu/~edelman/Psych-465-Spring-2003/PCA-tutorial.pdf

[39] Ščurek, R.: *Biometrické metody identifikace osob v bezpečnostní praxi*. 2008.

[40] Wiskott, L.; Fellous, J.-M.; Krüger, N.; aj.: Face recognition by elastic bunch graph matching. In *Computer Analysis of Images and Patterns*, *Lecture Notes in Computer Science*, ročník 1296, Springer Berlin / Heidelberg, 1997, ISBN 978-3-540-63460-7, ISSN 0302-9743 (Print) 1611-3349 (Online), s. 456–463, doi:10.1007/3-540-63460-6_150.
URL http://www.springerlink.com/content/2402786834342881/

[41] Zhang, H.; Zhao, D.: Spatial Histogram Features for Face Detection in Color Images. In *Advances in Multimedia Information Processing - PCM 2004*, *Lecture Notes in Computer Science*, ročník 3331, Springer Berlin / Heidelberg, 2005, ISBN 978-3-540-23974-1, ISSN 0302-9743 (Print) 1611-3349 (Online), s. 377–384, doi:10.1007/b104114.
URL http://www.springerlink.com/content/qamg56gx3l09m8d3/

[42] Zhao, W.; Chellappa, R.; Phillips, P.: Subspace Linear Discriminant Analysis for Face Recognition. Technická zpráva, 1999.

[43] Zhao, W.; Chellappa, R.; Phillips, P. J.; aj.: Face recognition: A literature survey. *ACM Comput. Surv.*, ročník 35, č. 4, December 2003: s. 399–458, ISSN 0360-0300, doi:10.1145/954339.954342.
URL http://dx.doi.org/10.1145/954339.954342

Příloha A

Obsah CD

A.1 Aplikace

Zdrojový kód aplikace a aplikace ve spustitelné formě.

A.2 Písemná zpráva

Zdrojový kód písemné zprávy a zpráva ve formátu PDF.

A.3 Programová dokumentace

Uživatelský manuál.

A.4 Použité databáze obličejů

Použité snímky z databází AT&T "The Database of Faces", Extended M2TVS Database a The Yale Face Database B upravené pro účely aplikace.