

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ
ÚSTAV INFORMAČNÍCH SYSTÉMŮ

FACULTY OF INFORMATION TECHNOLOGY
DEPARTMENT OF INFORMATION SYSTEMS

ROZPOZNÁNÍ OBLIČEJE

DIPLOMOVÁ PRÁCE

MASTER'S THESIS

AUTOR PRÁCE

AUTHOR

Bc. ADAM KOPŘIVA

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

ROZPOZNÁNÍ OBLIČEJE

FACE RECOGNITION

DIPLOMOVÁ PRÁCE

MASTER'S THESIS

AUTOR PRÁCE

AUTHOR

Bc. ADAM KOPŘIVA

VEDOUCÍ PRÁCE

SUPERVISOR

Ing. PETR CHMELÁŘ

BRNO 2010

Abstrakt

Tato diplomová práce se věnuje metodám pro rozpoznání obličeje. Popisuje metody založené na různém přístupu: znalostní metody, metody založené na rysech tváře, metody založené na modelech a metody založené na vzhledu tváře. Nejvíce se však zaměřuje na metody deformačních modelů a statistické metody, kterými jsou analýza hlavních komponent (PCA) a lineární diskriminační analýza (LDA). Podrobně popisuje princip metod deformačních modelů jako je metoda active shape models (ASM) a metoda active appearance models (AAM).

Abstract

This master's thesis considers methods of face recognition. There are described methods with different approaches: knowledge-based methods, feature invariant approaches, template matching methods and appearance-based methods. This master's thesis is focused particularly on template matching method and statistical methods like a principal component analysis (PCA) and linear discriminant analysis (LDA). There are described in detail template matching methods like active shape models (ASM) and active appearance models (AAM).

Klíčová slova

Rozpoznání obličeje, metody deformačních modelů, PCA, LDA, ASM, AAM

Keywords

Face recognition, Template matching, PCA, LDA, ASM, AAM

Citace

Adam Kopřiva: Rozpoznání obličeje, diplomová práce, Brno, FIT VUT v Brně, 2010

Rozpoznání obličeje

Prohlášení

Prohlašuji, že jsem tuto diplomovou práci vypracoval samostatně pod vedením pana Ing. Petra Chmelaře. Uvedl jsem všechny literární prameny a publikace, ze kterých jsem čerpal.

.....

Adam Kopřiva

23. května 2010

© Adam Kopřiva, 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	4
1.1	Detekce obličejů	4
1.1.1	Lokalizace obličejů	4
1.1.2	Rozpoznání obličejů	5
1.1.3	Problémy při detekci obličejů	6
1.2	Struktura práce	7
2	Metody detekce obličejů	8
2.1	Metody založené na modelech	9
2.1.1	Předem definované šablony	10
2.1.2	Deformační modely	10
2.2	Metody založené na rysech tváře	10
2.2.1	Rysy tváře	11
2.2.2	Textura obličejů	11
2.2.3	Barva kůže	12
2.2.4	Kombinace metod	12
2.3	Metody založené na vzhledu tváře (Appearance-Based)	12
2.3.1	Neuronové sítě	12
2.3.2	Support Vector Machines (SVM)	13
2.4	Statistické metody	13
2.4.1	Principal Component Analysis	14
2.4.2	Linear Discriminant Analysis	16
2.5	Znalostní metody	18
3	Metody založené na modelech	20
3.1	Statistical Shape Models	20
3.1.1	Vhodné orientační body	21
3.1.2	Zarovnání trénovací sady	21
3.1.3	Deformování modelu	22
3.1.4	Určení počtu módů	23
3.1.5	Omezující podmínky	23

3.1.6	Příklady modelů	24
3.2	Active Shape Models	24
3.2.1	Získání nových bodů	25
3.2.2	Prizpůsobení modelu novým bodům	27
3.2.3	ASM s více úrovněmi rozlišení	29
3.2.4	Příklady ASM	30
3.3	Statistical Models of Appearance	31
3.3.1	Statistický model textury	31
3.3.2	Kombinace modelů	33
3.3.3	Příklady Statistical Models of Appearance	34
3.4	Active Appearance Models	34
3.4.1	Rozdíly mezi AAM a ASM	35
3.4.2	Příklady AAM	36
4	Návrh a implementace	37
4.1	Návrh aplikace	37
4.1.1	Návrh demonstračního a testovacího rozhraní	37
4.1.2	Úložiště dat	38
4.1.3	Implementační jazyk	38
4.2	Implementace	38
4.2.1	Popis jednotlivých tříd	39
4.2.2	Schéma aplikace	40
4.2.3	Použité knihovny	40
4.2.4	Určení počáteční aproximace modelu	41
5	Testování	42
5.1	Dostupné databáze obličejů	42
5.1.1	The Extended M2VTS Database	42
5.1.2	The IMM Face Database	43
5.1.3	Další databáze obličejů	45
5.2	Experiment 1 - lokalizace obličeje metodou ASM	45
5.2.1	IMM databáze	46
5.2.2	XM2VTS databáze	49
5.2.3	Vyhodnocení	51
5.3	Experiment 2 - rozpoznání obličeje	51
5.3.1	Vytvoření modelu	52
5.3.2	Získání šablon	52
5.3.3	Rozpoznání obličeje metodou LDA	52
5.3.4	Vyhodnocení	53

6	Závěr	54
6.1	Možná rozšíření	54
6.2	Využití metody ASM	55
A	Obsah DVD	59
B	Příručka k programu	60
B.1	Panel Statistical Shape Model	60
B.1.1	Popis panelu	60
B.1.2	Postup pro vytvoření modelu	61
B.2	Panel Active Shape Model	61
B.2.1	Popis panelu	61
B.2.2	Postup pro zahájení lokalizace pomocí ASM	62
B.3	Panel Localization - Euclidean distance	62
B.3.1	Popis panelu	62
B.3.2	Postup pro lokalizaci všech obličejů v testovací sadě	63
B.4	Panel Identification - LDA	63
B.4.1	Popis panelu	63
B.4.2	Postup pro rozpoznání obličejů pomocí LDA	64
B.5	Panel Help	65
C	Schéma aplikace	66
D	Experiment 2	67

Kapitola 1

Úvod

Rozpoznání obličeje v současné době zaujímá spolu s otisky prstů v oblasti vývoje identifikačních systémů dominantní podíl. Děje se tak hlavně díky vysoké akceptaci od uživatelů. Uživatelé jsou mnohem ochotnější k nasnímání obličeje než například k nasnímání sítnice oka či zanechání vzorku DNA. Aplikace pro rozpoznání obličeje se uplatní kromě bezpečnostních systémů také v kriminalistice [14] a stále rozšířenější jsou i komerční aplikace (např. detekce úsměvu u fotoaparátů). Metody pro rozpoznání obličeje se uplatňují i v aplikacích pro sdílení fotografií (picasa, facebook).

Detekce obličeje spadá do kategorie rozpoznání objektů. Vstupní informací pro metody v této kategorii jsou data získaná z fotografie či ze sekvence snímků (video). Obličej je velmi složitým objektem, ale má své zákonitosti, kterých využívá celá řada metod s odlišnými přístupy.

Ačkoliv je určování identity člověka na základě rozpoznání obličeje velmi přirozenou úlohou, setkáváme se s velkou řadou problémů, které musí algoritmy pro rozpoznávání obličeje řešit. Za jeden z nich lze považovat překrytí části obličeje (například vlasy, brýle). Mezi další problémy řadíme natočení tváře vůči snímácímu zařízení, tento problém se snaží řešit algoritmy pracující se snímky v 2,5D či 3D. K dalším problémům patří řešení stárnutí či zranění v obličeji.

1.1 Detekce obličeje

Detekci obličeje lze považovat za úlohu skládající se ze dvou fází. V první fázi se rozhoduje, zda se na vstupním obrázku nějaké oblasti s obličejem nalézají. Pokud je první fáze úspěšná, přistoupí se ke druhé fázi, ve které je snaha obličej identifikovat či verifikovat.

1.1.1 Lokalizace obličeje

První krok každého systému na detekci obličeje začíná určením oblasti, kde se tvář může nalézat (tzv. lokalizace). Nalezení obličeje na snímku nebo na sekvenci snímků lze považovat

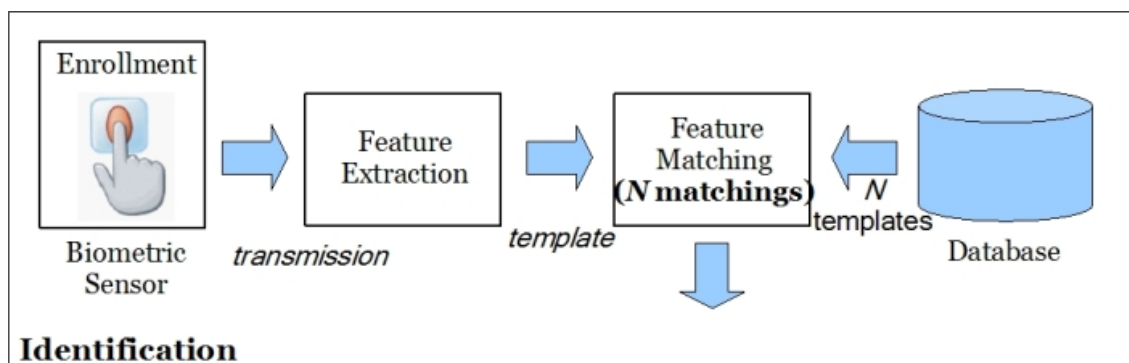
za relativně složitou úlohu [26]. Je totiž potřeba brát v úvahu odlišné vlastnosti na jednotlivých obrázcích. K těm patří například různá velikost a orientace obličeje, směr pohledu, barva kůže, výraz tváře nebo podmínky osvětlení při pořizování fotografie. Cílem metod pro lokalizaci obličeje je na základě daného snímku rozhodnout zda a v kterých místech obsahuje tvář či tváře.

1.1.2 Rozpoznání obličeje

V případě, že se na vstupním obrázku tváře nalézají, daný systém se tyto obličeje snaží buď identifikovat nebo verifikovat [19]. Speciálním případem rozpoznání je seznam sledovaných (watch list).

Identifikace

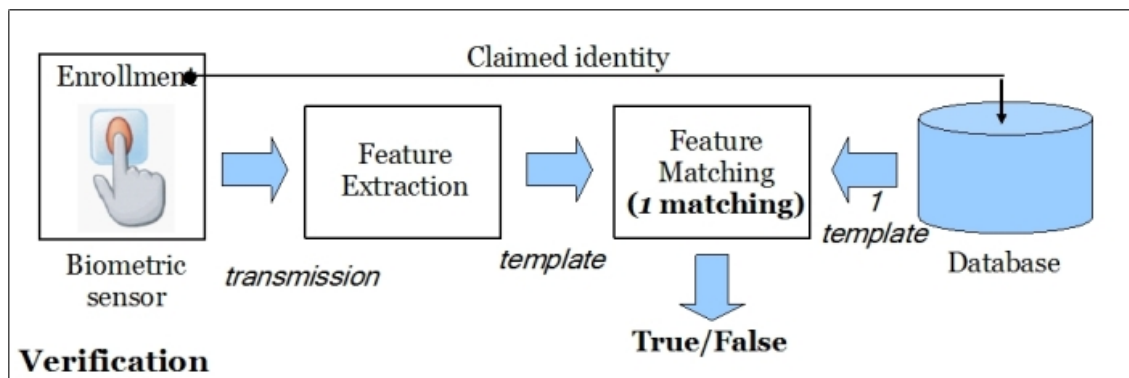
Při identifikaci se postupuje tak, že se vstupní data porovnávají se všemi dostupnými vzorky. Jedná se tedy o porovnání 1:N. Například při identifikaci na základě otisků prstů se porovnávají nasnímané či jinak získané otisky prstů se všemi dostupnými vzorky v databázi. Pokud je zjištěna shoda našeho vzorku s některým ze vzorků v databázi, došlo k jeho identifikaci [19]. Demonstrace funkce identifikačního systému je na obrázku 1.1.



Obrázek 1.1: Příklad identifikačního systému. Převzato z [10].

Verifikace

Na rozdíl od identifikace při verifikaci se porovnává získaný vzorek s jedním referenčním vzorkem, jedná se o porovnání 1:1. Například při rozpoznání uživatele na základě geometrie ruky se postupuje tak, že uživatel nejprve uloží pod svým identifikačním číslem do systému svůj vzorek ruky. Při verifikaci pak uživatel zadá své identifikační číslo a nasnímá vzorek ruky, který je porovnán právě s jeho jedním referenčním vzorkem [19]. Obrázek 1.2 prezentuje postup při verifikaci otisku prstu.



Obrázek 1.2: Příklad verifikačního systému. Převzato z [10].

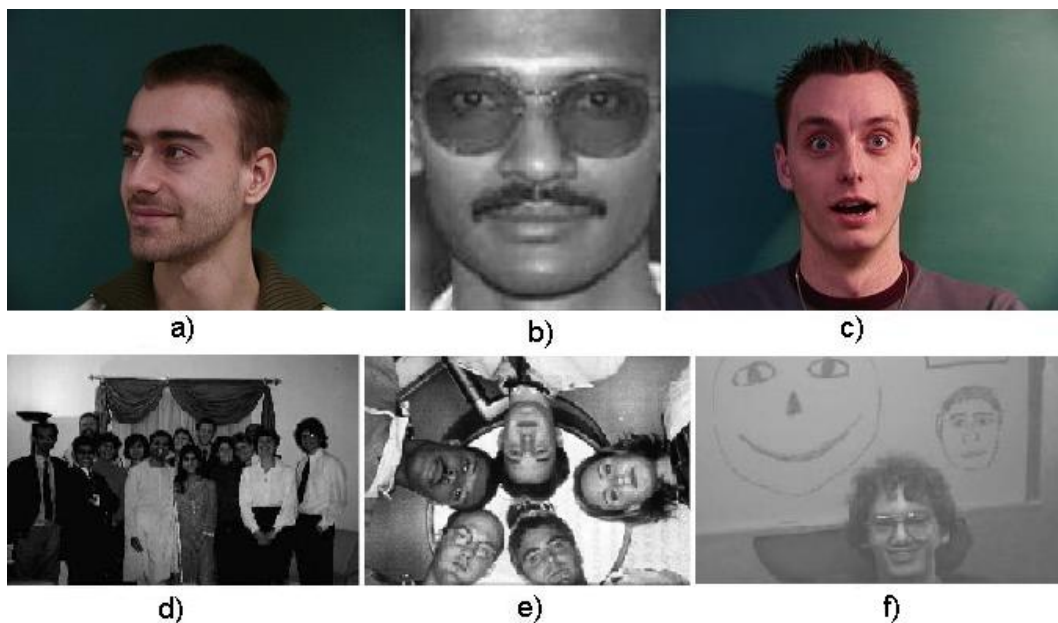
Seznam sledovaných

Speciálním případem identifikačního systému je seznam sledovaných (watch list). Systém musí nejprve na základě vstupní biometrické informace zjistit, zda osoba je, nebo není, na seznamu sledovaných. Pokud je osoba na seznamu, musí být systém schopen tuto osobu správně identifikovat [10], [18].

1.1.3 Problémy při detekci obličeje

Jak už bylo zmíněno, v systémech pro detekci obličeje se setkáváme s celou řadou komplikací [26], [27]:

- **Póza.** Snímky obličeje mohou být pořízeny v různém vzájemném vztahu mezi fotoaparátem (kamerou) a snímanou osobou (čelní pohled, natočení tváře vůči snímacímu zařízení, pohled z profilu, shora či naopak za sebou). Některé z rysů obličeje mohou být také částečně či úplně zakryty (přivření či zavření očí). Příklad natočení hlavy je vidět na obrázku 1.3 a).
- **Přítomnost či absence doplňků** jako jsou vousy, knírky a brýle vytváří velkou proměnlivost obličeje v souvislosti s jejich tvarem, barvou a velikostí. Příklad na obrázku 1.3 b).
- **Výraz obličeje.** Podobu tváře značně ovlivní i okamžitý výraz v obličeji při snímání, jak znázorňuje obrázek 1.3 c).
- **Překrytí tváře.** Při pořizování snímků obličeje se častou vyskytne situace, při které dochází k částečnému překrytí tváře cizím objektem, či jiným obličejem (typicky u skupinových fotografií) obrázek 1.3 d).
- **Natočení snímku.** Obrázek tváře může být pořízen s různou rotací snímacího zařízení. Příklad na obrázku 1.3 e)



Obrázek 1.3: Obrázek s příklady komplikací při detekci obličeje. Upraveno z [7], [21], [26].

- **Podmínky při pořizování obrázku.** Vytváření snímku ovlivňují faktory jako osvětlení (spektrum barev, odlesky, intenzita) a také parametry snímacího zařízení. Detekci obličeje ovlivňuje i pozadí snímku, příklad problematického pozadí fotografie je uveden na obrázku 1.3 f).

1.2 Struktura práce

V této práci se nejdříve zaměřím na přehled a příklady jednotlivých metod určených zejména pro rozpoznání obličeje. V následující kapitole budou podrobně rozebrány metody založené na deformačních modelech. V kapitole dostupné databáze obličejů budou představeny databáze, které byly využity v této diplomové práci. V dalších kapitolách je uveden postup návrhu, implementace a diskuze výsledků získaných z testování.

Kapitola 2

Metody detekce obličeje

V průběhu posledních let lze pozorovat velmi intenzivní výzkum v oblasti detekce lidské tváře [7], [26]. Díky tomu se vyskytlo mnoho přístupů k řešení tohoto problému. Metody pro detekci obličeje lze rozdělit do čtyř kategorií: metody založené na modelech, metody založené na rysech tváře, metody založené na vzhledu tváře (Appearance-Based) a znalostní metody.

- **Metody založené na modelech.** Tato metoda využívá sadu šablon, které popisují obličej jako celek nebo jeho rysy. Pomocí vypočtení korelačního koeficientu, který slouží k porovnání vstupního obrázku a uložených šablon, se rozhodne o případné shodě. Tyto metody se používají jak pro lokalizaci, tak identifikaci obličeje [4].
- **Metody založené na rysech tváře.** Tyto algoritmy mají za cíl najít tvář pomocí jednotlivých obličejových rysů a jejich vzájemnou polohou. Tyto metody se používají hlavně pro lokalizaci tváře [27], [7].
- **Metody založené na vzhledu tváře (Appearance-Based).** Tyto metody využívají sadu trénovacích obrázků, která by měla obsahovat zástupce s různým vzhledem a výrazem tváře. Na základě trénovací sady se vytváří předloha, která se používá zejména pro identifikaci tváře [26].
- **Statistické metody.** Metody, které se řadí do statistických metod, tvoří matematický základ pro mnoho technik rozpoznání obličeje. Mezi statistické metody lze zařadit metodu analýzy hlavních komponent (PCA) a metodou lineární diskriminační analýzy (LDA) [7].
- **Znalostní metody.** Tyto metody používají přístup shora dolů. Hledají obličej na obrázku jako celek. Provádějí nejprve rychlé testování na nízké úrovni detailu. Tyto metody jsou navrženy pro lokalizaci obličeje [26].

Tabulka 2.1: Tabulka s přehledem přístupů k řešení problému a metod, které do jednotlivých kategorií řadíme. Upraveno z [26].

Přehled metod detekce obličeje a přístupů k řešení problému
Metody založené na modelech - předem definovaná šablona tváře - deformační modely (ASM, AAM)
Metody založené na rysech tváře - obličejové rysy (seskupování hran) - textura (šedotónové rozložení) - barva kůže - kombinace metod (využití více vlastností - barva kůže, tvar obličeje, atd.)
Metody založené na vzhledu tváře (Appearance-Based) - Eigenface - Metody založené Gaussově rozložení - Neuronové sítě - SVM - Modelování rozložení pravděpodobnosti tříd (Bayes) - Skryté markovovy modely - Znalostně teoretický přístup (Kullback)
Statistické metody - PCA - LDA
Znalostní metody - metoda různých úrovní detailu

Základní souhrn metod je uveden v tabulce 2.1. Některé metody překračují hranice tohoto rozdělení a lze je zařadit do více kategorií. Většina metod v podstatě vychází ze statistických metod, tedy předpokládají, že lidskou tvář lze detekovat na základě rozdílu mezi obrázky.

2.1 Metody založené na modelech

Základem těchto metod je ručně označená sada snímků tváří. Příklad označené tváře z trénovací sady snímků je na obrázku 2.1. Při zpracování vstupního obrázku pak proběhne porovnání se šablonami uložených obrázků a vypočte se korelační koeficient (neboli vzájemný vztah mezi obrázkem a šablonou). Na základě tohoto koeficientu se spočítá minimální vzdálenost (např. pomocí euklidovské vzdálenosti). Existence tváře se pak určí na základě maximální hodnoty korelačního koeficientu a minimální vzdálenosti.



Obrázek 2.1: Příklad označení tváře. Převzato z [4].

2.1.1 Předem definované šablony

V tomto nejjednodušší přístupu se využívá k sestavení modelu několik menších šablon pro oči, nos, ústa a konturu obličeje. Každá menší šablona je definována poměry hodnot jasu mezi jejími částmi. Rysy ze vstupního obrázku se získají na základě největších změn hodnot jasu a poté se porovnávají s menšími šablonami. Nejprve se porovnává vzájemný vztah získaných rysů s šablonou kontury obličeje. Čímž dojde k získání kandidátských oblastí (takové oblasti, kde by se mohl vyskytovat obličej). Poté se na tyto oblasti použijí další menší šablony [26].

2.1.2 Deformační modely

Tato technika využívá pro detekci rysů obličeje elastický model, který postupně tvaruje tak, aby kopíroval rysy tváře [4]. Obličejové rysy popisují šablony s parametry. Na vstupním obrázku se postupně přizpůsobují parametry šablony tak, aby odpovídala parametrům získaným z obrázku.

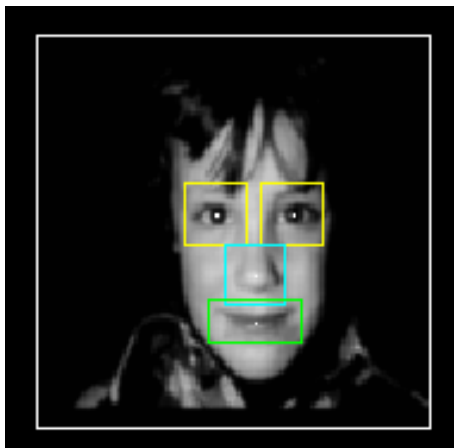
Za výhodu těchto metod lze považovat jejich jednoduchost a snadnou implementaci. Na druhou stranu výsledky rozpoznání závisí na vlastnostech šablon obrázků (velikost, orientace, tvar, atd.). V této práci se budeme podrobně zabývat metodami založenými na modelech v kapitole 3.

2.2 Metody založené na rysech tváře

Metody zahrnuté v této kategorii společně charakterizuje přístup k detekci obličeje tzv. zdola nahoru. Tento přístup funguje na principu nalezení nejprve částí objektu v obraze, které se pokouší v dalším úseku spojit dohromady. Existuje mnoho metod, které se zaměřují na detekci obličejových rysů (ústa, oči, obočí, rty, atd.) jednotlivě a podle jejich vzájemných

pozic určují obličej. Tento přístup využívají metody založené na detekci rysů tváře, textuře obličeje, barvě kůže a kombinací více metod. Příklad oblastí s vyznačenými rysy obličeje je demonstrován na obrázku 2.2. K zvýraznění rysů tváře, které vede k snadnějšímu nalezení oblasti s obličejem, se často využívá hranových detektorů.

Za největší nedostatek těchto metod lze považovat to, že si neporadí v případě, kdy rysy lidské tváře mohou být na snímku poničeny díky osvětlení, šumu nebo odleskům.



Obrázek 2.2: Příklad oblastí s vyznačenými rysy. Upraveno z [19].

2.2.1 Rysy tváře

Metod, které detekují obličej na základě jeho rysů, je nespočet [7], [26], [27]. Metody velmi často postupují ve dvou krocích. Nejprve dochází k lokalizaci tváře, respektive oblasti kde by se tvář mohla nacházet a poté k detekci jednotlivých částí.

Jako příklad těchto metod lze uvést techniku, která využívá k lokalizaci tváře mapu hran (získanou pomocí Cannyho hranového detektoru) a pomocí heuristiky odstraňuje nebo naopak slučuje hrany tak, že zachová pouze konturu tváře. Elipsa představující oblast s obličejem je pak umístěna na hranici obličeje a pozadí obrázku.

2.2.2 Textura obličeje

Základ této metody tvoří myšlenka, že lidská tvář má odlišnou texturu, která může být použita k rozlišení obličeje od ostatních objektů. Textura je spočítána na menším obrázku o rozměru 16x16 pixelů. Uvažují se tři typy rysů: kůže, vlasy a ostatní. Poté se použitím kaskádové neuronové sítě pro třídění textur vytváří třídy pro odlišné textury. Přítomnost tváře je z textury odvozena na základě výskytu vlasů a kůže [26].

2.2.3 Barva kůže

Barva lidské kůže se používá jako efektivní příznak lidské tváře v mnoha metodách. Ačkoliv různí lidé mají různou barvu kůže, mnoho studií prokázalo, že hlavní rozdíl tkví převážně v intenzitě než v barvě. Pro označení barvy kůže se využívá mnoho barevných modelů (RGB, normalizované RGB, HSV, YCrCb, ...).

Nejjednodušší metodou je definovat rozsah pixelů s odstínem kůže pomocí hodnot Cr , Cb ze vzorků pixelů lidské kůže. S opatrně zvoleným prahem $[Cr_1, Cr_2]$ a $[Cb_1, Cb_2]$ pixel se klasifikuje jako odstín kůže pokud hodnoty (Cr, Cb) skončí v rozsahu $Cr_1 \leq Cr \leq Cr_2$ a $Cb_1 \leq Cb \leq Cb_2$ [26].

2.2.4 Kombinace metod

Nespočet metod kombinuje několik technik pro detekci obličeje na základě rysů tváře. Nejčastěji se využívá obecných rysů jako je barva kůže, velikost a tvar k nalezení kandidátských oblastí. Poté se k ověření těchto oblastí používá detailnějších rysů jako je obočí, nos a vlasy.

Typický postup zahrnuje seskupování pixelů na základě barvy kůže. Dále prozkoumá tuto oblast vzhledem k jejímu tvaru, pokud je tvar oválný, označí ji za kandidátskou oblast. Pak v těchto oblastech hledá oči, obočí, nos. Pokud jsou hledané rysy nalezeny, prohlásí kandidátskou oblast za část obrázku, ve které se nachází obličej [26].

2.3 Metody založené na vzhledu tváře (Appearance-Based)

Metody v této kategorii potřebují dvě trénovací sady obrázků. Jedna sada obsahuje obrázky tváří, v druhé sadě se naopak obličeje nevyskytují. Na základě technik ze statistické analýzy a strojového učení se tyto metody snaží najít významné parametry v obou trénovacích sadách. Získané znalosti zpracovávají ve formě distribučních modelů (modelů rozložení) nebo diskriminačních funkcí, které se následně používají pro detekci obličeje [26].

Do této kategorie metod lze řadit metody založené na klasifikátorech, jako jsou neuronové sítě, AdaBoost, SVM, atd.

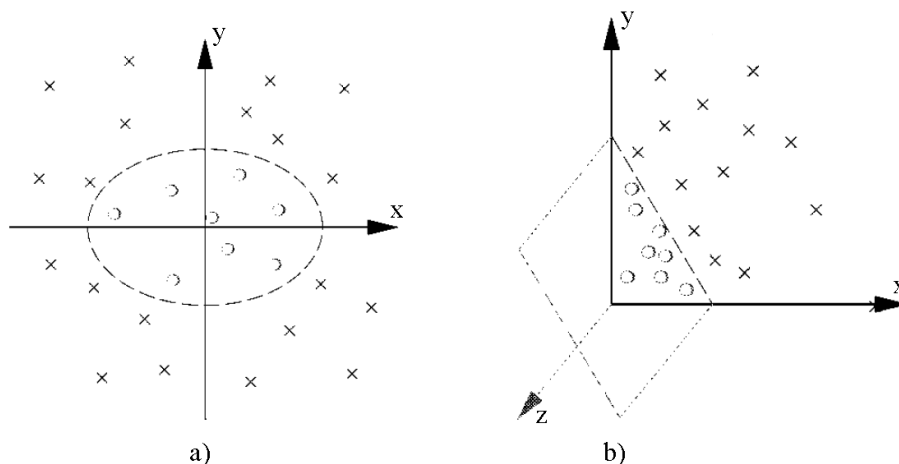
2.3.1 Neuronové sítě

Neuronové sítě byly úspěšně použity v řadě úloh pro rozpoznávání vzorů (například v aplikaci pro automatické řízení vozidla). Při detekci obličeje se využívá možnosti trénování neuronových sítí. Získá se tak celková informace o zákonitostech lidské tváře. Na druhou stranu k nevýhodám této metody patří, že neuronová síť musí být co nejlépe vyladěná (množství vrstev, uzlů, nastavení vah, atd.), aby mohla být použita [26].

2.3.2 Support Vector Machines (SVM)

Jedná se o lineární metodu strojového učení. Snaží se o nalezení nadroviny, která co nejlépe rozdělí trénovací data v prostoru příznaků. Data jsou odděleny do dvou množin. Výsledná rovina, která odděluje obě množiny dat, má k sousedním bodům obou tříd maximální vzdálenost [26].

Pokud data nelze rozdělit lineárně, lze využít jádrových metod [17]. Tyto metody transformují data do vícerozměrného prostoru, ve kterém již lze tato data separovat. Na obrázku 2.3 a) jsou zobrazena data ve dvojdimensionálním prostoru, která nelze lineárně rozdělit. Po přidání dimenze (obrázek 2.3 b) již lze obě množiny dat separovat.



Obrázek 2.3: a) Data, která nelze lineárně rozlišit. b) Po přidání dimenze již lze množiny dat separovat. Upraveno z [17].

2.4 Statistické metody

Mezi statistické metody lze zařadit metodu analýzy hlavních komponent (PCA - principal component analysis) a metodou lineární diskriminační analýzy (LDA - linear discriminant analysis). Statistické metody se používají zejména pro redukci vícerozměrných dat. V případě PCA se jedná o učení bez učitele a v případě LDA se mluví o metodě učení s učitelem [7].

Obě metody se využívají v mnoha oblastech, v souvislosti s rozpoznáním obličeje se pro metodu PCA vžil název *eigenfaces*, v případě metody LDA se jedná o *fisherfaces*. Každá z metod probíhá ve dvou fázích. Nejprve se uskuteční fáze trénování a poté proběhne fáze klasifikace neboli třídění (testování).

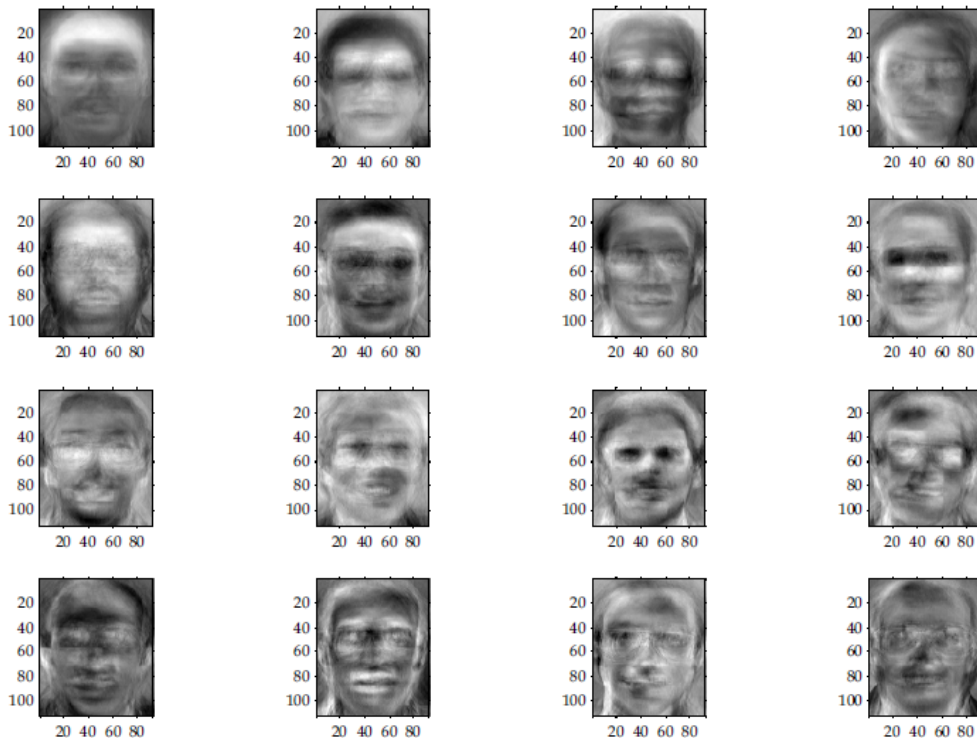
2.4.1 Principal Component Analysis

Analýza hlavních komponent je standardní technikou používanou ve statistickém rozpoznávání vzorů nebo ve zpracování signálů. Používá se pro redukci prostorových dat. Dochází tak k jejich zjednodušení při co nejmenší ztrátě informace. Vzory často obsahují přebytečné informace, jejich promítnutím na vektor rysů se lze těchto přebytečných dat zbavit a přitom zachovat většinu skutečně podstatných informací ze vzoru.

Obrázek obličeje s rozměry 720 x 576 (např. pro databázi M2VTS část 5.1.1) si lze také představit jako vektor z dimenze 414720. Do tohoto prostoru se promítne celá trénovací sada obrázků. Získá se tak sada bodů v tomto ohromném prostoru. Obrázky tváří jsou si podobné v celé řadě vlastností, proto nebudou body v prostoru rozmístěny náhodně. Díky tomu se dají popsat relativně méně dimenzionálním podprostorem.

Hlavní myšlenkou PCA je najít vektory, které nejlépe popisují rozmístění obrázků uvnitř tohoto podprostoru. Každý z těchto vektorů o velikosti 414720, popisuje 720 x 576 rozměrný obrázek a je lineární kombinací původního obrázku.

Protože tyto vektory jsou vlastními vektory (*eigenvectors*) kovarianční matice odpovídající originálnímu obrázku, a protože se pohybujeme v oblasti rozpoznání obličeje (*face recognition*) mluvíme o metodě *eigenfaces* [7]. Příklad tváří zpracovaných metodou *eigenfaces* prezentuje obrázek 2.4.



Obrázek 2.4: Příklad metody *eigenfaces*, převzato z [7].

Eigenfaces

Metoda, která byla jako první funkční a aplikovala PCA v oblasti rozpoznání obličejů, se nazývá *eigenfaces*. Byla navržena Turkem a Pentlandem [24]. Pomocí PCA se značně zjednoduší výpočet z řádu počtu pixelů v obrázku M (např. 414720, pro obrázek o rozměru 720x576 pixelů) do řádu počtu obrázků obsažených v sadě pro trénování N (např. 50 obrázků). V praxi je trénovací sada relativně malá $N \ll M$ (tj. $50 \ll 414720$), výpočty se tak stanou lépe řešitelnými. Přidružená vlastní čísla umožňují řadit vlastní vektory podle jejich použitelnosti v hledání rozdílu mezi obrázky. V praxi pak stačí podmonožina vektorů z M významných vlastních vektorů s největšími vlastními čísly, která jsou dostačující pro spolehlivou reprezentaci tváří v daném podprostoru.

Při identifikaci tváří se promítnou data z obrázku do tohoto podprostoru a získané vektory se porovnají s vektory připravenými z trénovací sady. Porovnání vektorů se provádí například pomocí euklidovské vzdálenosti. Průběh testovací fáze je uveden na obrázku 2.7.

Výpočet PCA pomocí kovarianční metody

Sada N vektorů x_i obsahuje informace o jednotlivých obrázcích (tedy počet vektorů odpovídá počtu obrázků v trénovací sadě, např. $N = 50$). Tyto vektory mají rozměr $M \times 1$, kde M je počet pixelů obrázku (např. u obrázku o rozměrech 720 x 576 je počet pixelů $M = 414720$). Vypočtením PCA dojde k zredukování dat z dimenze M do dimenze L přičemž platí, že $L \ll M$ [25].

1. Výpočet průměrného vektoru.

$$u = \frac{1}{s} \sum_{i=1}^s x_i \quad (2.1)$$

2. Výpočet odchylky jednotlivých vektorů.

Od každého vektoru x_i je odečten průměrný vektor u . Výsledné vektory s odchylkou jednotlivých vektorů jsou uloženy do matice \mathbf{B} o rozměrech $M \times N$.

3. Výpočet kovarianční matice C .

$$C = \frac{1}{N} \sum_{i=1}^N B \cdot B^T \quad (2.2)$$

4. Nalezení vlastních vektorů a vlastních čísel kovarianční matice.

Výpočet matice V obsahující vlastní vektory matice C :

$$V^{-1}CV = D, \quad (2.3)$$

kde D je matice vlastních čísel matice C

5. Seřazení vlastních čísel a jim odpovídajících vlastních vektorů.

Seřaď jednotlivá vlastní čísla a jim odpovídající vlastní vektory tak, aby

$$\lambda_i \geq \lambda_{i+1}, \quad (2.4)$$

kde λ_i je odpovídající vlastní číslo z matice D .

6. Výběr podmnožiny vlastních vektorů.

Vyber prvních L vlastních vektorů tak, že:

$$1 \leq L \leq M, \quad (2.5)$$

Počet vlastních vektorů lze zvolit na základě prahu tak, aby byla například zachována 90% varibilita dat:

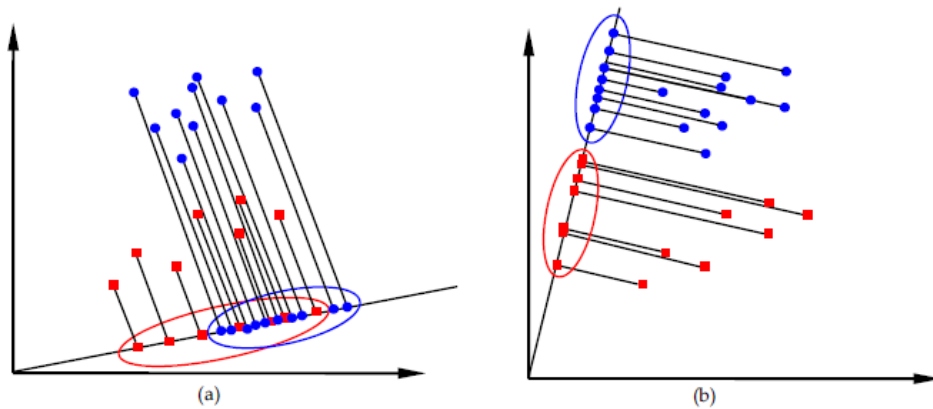
$$\sum_{i=1}^L \lambda_i \geq 0.9 \cdot \sum_{j=1}^M \lambda_j, \quad (2.6)$$

kde λ_i je odpovídající vlastní číslo z matice D .

Konkrétní využití metody PCA je uvedeno v části 3.1.3.

2.4.2 Linear Discriminant Analysis

Metoda LDA, nebo také v souvislosti s oblastí rozpoznávání obličejů lze hovořit o Fisherfaces, překonává omezení metody PCA použitím Fisherova lineárního diskriminantu (FLD). Na rozdíl od předchozí metody využívá LDA učení s učitelem. To spočívá v tom, že algoritmus bude poučen, které obrázky z trénovací sady patří do stejné třídy (různé snímky od stejné osoby).



Obrázek 2.5: Dvě sady bodů promítnuty na jednu přímku a) sady bodů nelze rozlišit, b) sady bodů jsou odděleny. Převzato z [7].

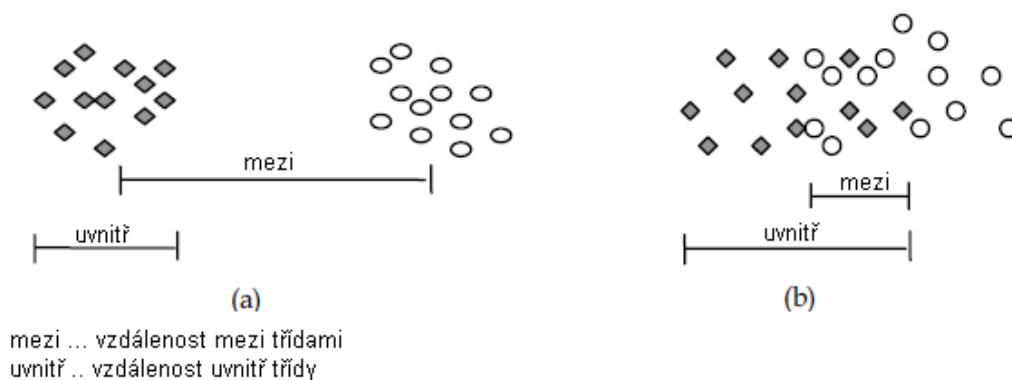
Fisherfaces

Fisherovi diskriminanty seskupují snímky obličeje stejné osoby do stejné třídy a oddělují obrázky různých osob do odlišných tříd. Uvnitř třídy se uchovávají informace o odlišnostech ve zjevu stejné osoby způsobené rozdílným osvětlením nebo výrazem tváře. Rozdílné třídy pak obsahují různé osoby. Použitím této metody se na jednu stranu maximalizuje vzdálenost mezi obrázky tváří z různých tříd, na druhou stranu se minimalizuje vzdálenost mezi obrázky ze stejné třídy.

Na obrázku 2.5 se vyskytují dvě sady bodů, tyto body se promítnou na přímku. Podle polohy přímky se mohou skupiny bodů buď prolínat (obrázek 2.5 a) nebo se skupiny bodů oddělí (obrázek 2.5 b). Fisherův diskriminant si klade za úkol najít takovou hyperrovinu, která v tomto dvojdimensionálním prostoru nejlépe oddělí obě sady bodů.

Při rozpoznávání testovacích obrázků, je promítnutý testovací obrázek porovnán s každým obrázkem z trénovací sady, poté je přiřazen k nejbližšímu obrázku z trénovací sady.

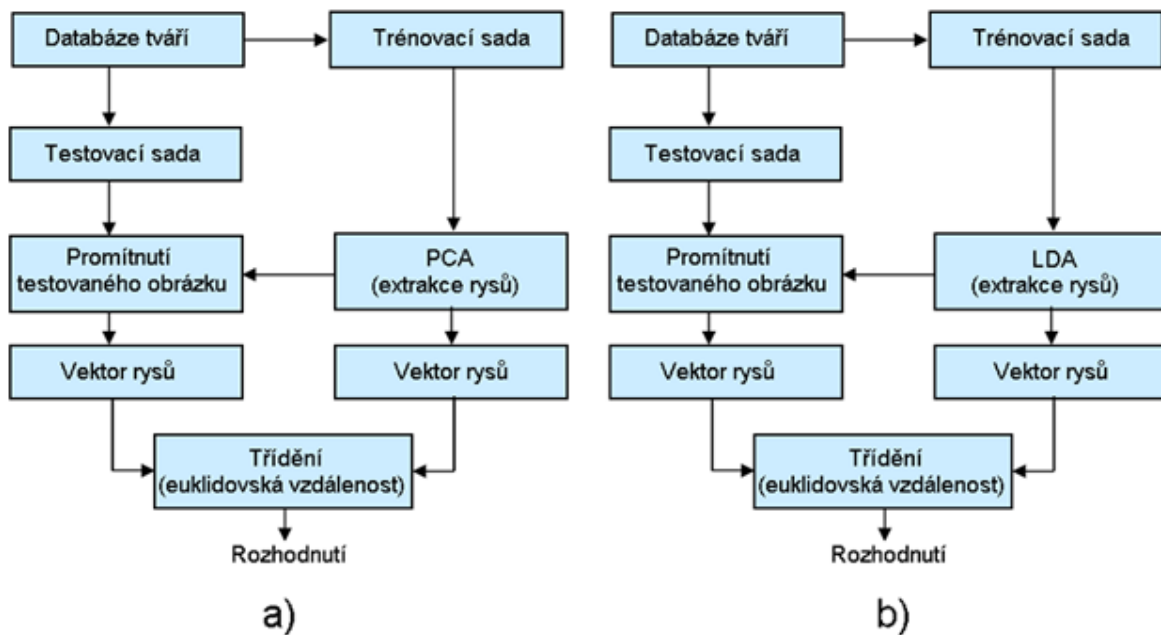
PCA se zaměřuje na extrakci rysů z obrázků tváří, které nejlépe charakterizují tyto obrázky s obličeji. Metoda LDA se na rozdíl od PCA pokouší nalézt podprostor, který nejlépe odlišuje různé třídy tváří, jak je ukázáno na obrázku 2.6.



Obrázek 2.6: a) vhodné rozdělení tříd b) nekvalitní rozdělení tříd. Upraveno z [7].

Na obrázku 2.7 lze vidět srovnání testovací fáze obou metod. Testovací fáze metody LDA se od metody PCA liší pouze ve výpočtu podprostoru. Při identifikaci tváří se testovaná data promítnou z obrázku do podprostoru. Získané vektory se porovnají s vektory připravenými ze sady obrázků určené pro trénování algoritmu. Porovnání vektorů je možno opět provést pomocí euklidovské vzdálenosti. Výsledkem testovací fáze je rozhodnutí, ke které osobě uložené v trénovací sadě patří testovaný obrázek.

Příklad tváří zpracovaných metodou fisherfaces lze vidět na obrázku 2.8.



Obrázek 2.7: a) Testovací fáze PCA. b) Testovací fáze LDA. Upraveno z [7].

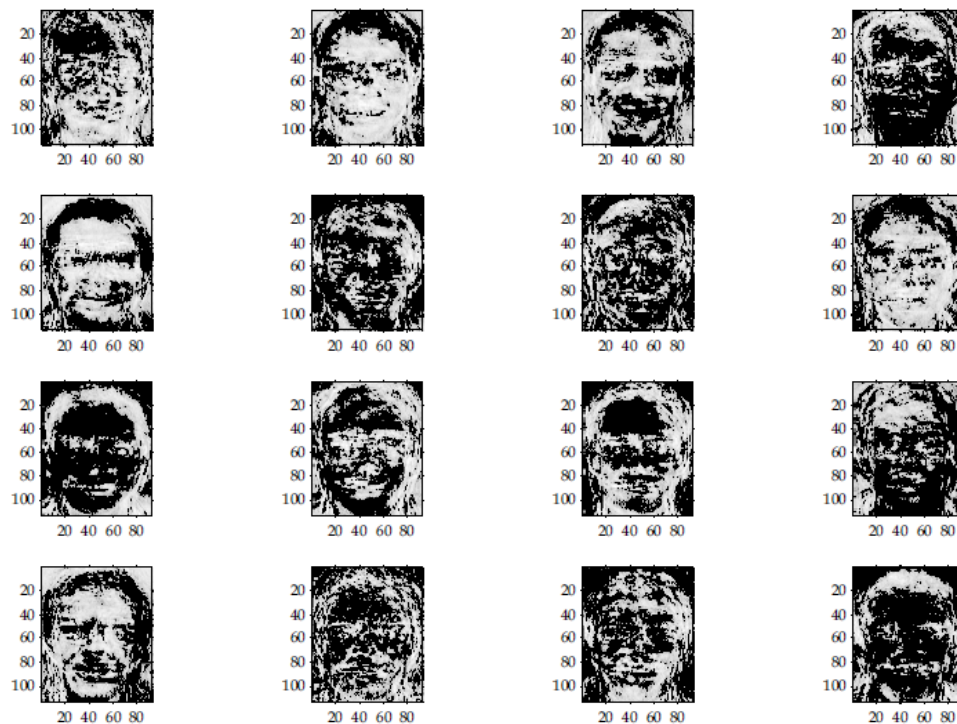
2.5 Znalostní metody

Tyto metody využívají přístupu shora dolů. Využívají pravidla, která popisují obličejové rysy jako celek. Například lidský obličej tvoří dvě oči, jeden nos a ústa. Pravidla pak popisují vztahy mezi těmito částmi obličeje a jejich vzájemnou polohu. V našem příkladě lidského obličeje, lze definovat tyto pravidla: středy očí leží na jedné přímce, středy očí a úst vytváří téměř rovnoramenný trojúhelník, atd [26].

S problémy se potýkají zejména při tvoření pravidel. Pokud se formulují pravidla detailní (přesná) algoritmus může selhat, protože nesplní všechna pravidla. Pokud se naopak vytvoří pravidla příliš obecná, algoritmus zahrne do výsledku i objekty, které nejsou tváří. Další problém spočívá v složitosti rozšířit tento přístup na detekci tváří v různých pózách. Pokud je požadováno tuto situaci v našem algoritmu zahrnout, zjistí se, že vyjádřit všechna případná natočení a výrazy obličeje lze pouze vytvořením velkého počtu pravidel. Na druhou stranu tato metoda pracuje dobře při detekci obličejů nasnímaných z čelního pohledu.

Jedna z možných modifikací metody používá hierarchické rozložení pravidel. Tento konkrétní systém se skládá ze tří stupňů pravidel. Základ tvoří obrázky s různou úrovní detailu. Obrázky lze vytvořit pomocí průměrování a podvzorkování snímků, jak je vidět na obrázku 2.9.

Na nejvyšším (prvním) stupni, probíhá prohledávání obrázku s nízkou úrovní detailu pomocí okna, ve kterém se postupně aplikuje první sada pravidel. Dochází tak k nalezení kandidátských oblastí s případnými tvářemi. Čím vyšší stupeň pravidel, tím jsou pravidla



Obrázek 2.8: Příklad metody fisherfaces převzato z [7].

obecnější. Díky tomu provede algoritmus nejprve rychlé testy na nízké úrovni detailu, poté prochází tyto oblasti podrobněji. Na stupni 2 se nejdříve upravují kandidátské oblasti pomocí histogramu a detektorů hran. Po úpravě se aplikuje další sada pravidel. Oblasti, které vyhověly obou sadám pravidel, prověří třetí (nejnižší) stupeň. Na tomto stupni se aplikuje poslední sada pravidel, která se zaměřuje na detailní rysy obličeje (oči, ústa, nos). Myšlenku hierarchického zpracování detekce obrazu je možno pozorovat i u další metod.



Obrázek 2.9: Příklad průměrování obrázku. Převzato z [26].

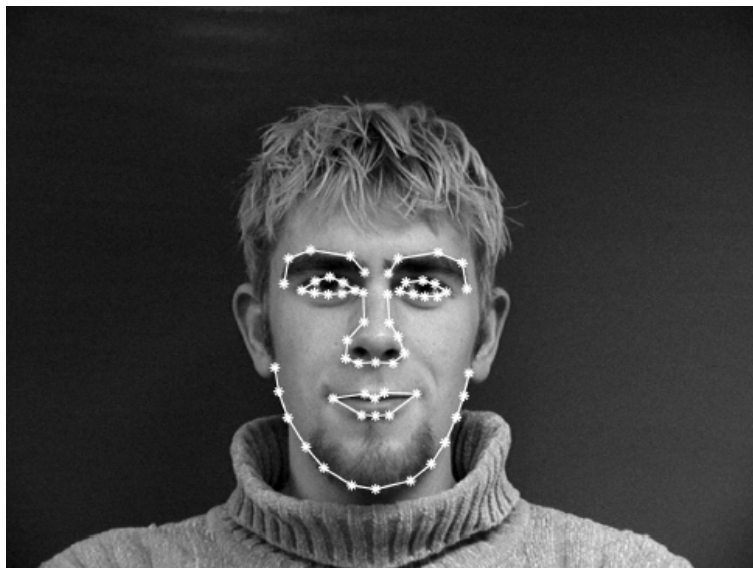
Kapitola 3

Metody založené na modelech

Následující kapitola pojednává o metodách založených na modelech. Detailně je rozebrána problematika metody pro vytvoření statistických modelů tvaru (*Statistical Shape Models*), textury a jejich spojení (*Statistical Models of Appearance*). Tyto metody podrobně popsali ve své práci *Statistical Models of Appearance for Computer Vision* T. F. Cootes a C. J. Taylor [4].

3.1 Statistical Shape Models

Jedná se o statistické modely tvaru, které budou použity ke znázornění objektů v obrázcích. Tvar objektu (ne nutně tváře) je označen sadou n bodů, které se mohou vyskytovat v libovolném množství dimenzí. Obvykle však jsou tyto body ve dvou či třech dimenzích. Tvar objektu zůstává neměnný při posunu, rotaci či změně měřítka.



Obrázek 3.1: Příklad vyznačených významných částí obličeje. Převzato z [20].

3.1.1 Vhodné orientační body

Pro přípravu trénovací sady obrázků je potřeba nejprve označit tvar objektu na těchto obrázcích. Vybírají se takové body, které jsou shodně umístěny na všech obrázcích. V praxi je tato činnost časově velmi náročná. Příklad označení tvaru obličeje je možno vidět na obrázku 3.1. Sadu obličejů [20], z které je jako příklad vybrán obrázek 3.1, označují šablony vytvořeny z 58 bodů.

Pokud je tvar obrázku popsán n body v d dimenzích, znázorňuje se tvar pomocí nd vektoru obsahujícího body modelu [2]. Jednotlivé dimenze jsou řazeny do vektoru za sebe. Například ve dvou dimenzích lze popsat tvar šablony tvořené n body vektorem x :

$$x = (x_1, \dots, x_n, y_1, \dots, y_n)^T.$$

Pro obrázek 3.1, který je popsán šablonou s 58 body by měl vektor x podobu:

$$x = (x_1, \dots, x_{58}, y_1, \dots, y_{58})^T.$$

Pokud obsahuje trénovací sada s obrázků, vytvoří se s vektorů x .

3.1.2 Zarovnání trénovací sady

Aby bylo možné s modelem dále pracovat, musí se sada určená pro trénování algoritmu nejprve zarovnat. K tomuto účelu existuje velké množství metod, nejpoužívanější je Prokrustova analýza [22]. Pomocí ní se posune každá šablona (x_i) s n body do počátku souřadnic. Vypočtení průměru bodů šablony:

$$x_m = \frac{x_1 + x_2 + \dots + x_n}{n}, y_m = \frac{y_1 + y_2 + \dots + y_n}{n}, \quad (3.1)$$

kde x_m a y_m je vypočtený průměr bodů.

Posunutí šablony do počátku souřadnic, pro každý bod šablony x_i platí:

$$x_n = (x_n - x_m), y_n = (y_n - y_m), \quad (3.2)$$

Zarovnání šablony posunuté do počátku souřadnic na pevně danou velikost:

$$s = \sqrt{(x_1)^2 + (y_1)^2 + \dots + (x_n)^2 + (y_n)^2} \quad (3.3)$$

$$x_n = \frac{x_n}{s}, y_n = \frac{y_n}{s}, \quad (3.4)$$

kde s upravuje měřítko šablony.

Rotaci o úhel θ a změnu měřítka s jednotlivých šablon zajišťuje transformace \mathbf{T} .

$$T_{s,\theta} \begin{pmatrix} x_n \\ y_n \end{pmatrix} = \begin{pmatrix} s \cos\theta & s \sin\theta \\ -s \sin\theta & s \cos\theta \end{pmatrix} \begin{pmatrix} x_n \\ y_n \end{pmatrix} \quad (3.5)$$

Iterativní algoritmus [4], popisující postupné zarovnání všech šablon, je uveden ve výpisu algoritmu 1.

Algoritmus 1 Zarovnání trénovací sady

- 1: Posuň každou šablonu tak, aby její těžiště bylo v počátku (rovnice 3.1 a 3.2).
 - 2: Vyber jednu šablonu za výchozí a označ ji jako \bar{x}_0 a uprav její měřítko na pevně danou velikost (rovnice 3.3 a 3.4).
 - 3: Zarovnej všechny šablony s šablonou \bar{x}_0 (rovnice 3.5).
 - 4: Vypočti průměrnou šablonu a označ ji \bar{x}_0 .
 - 5: Uprav měřítko průměrné šablony (\bar{x}_0) na pevně danou velikost (rovnice 3.3 a 3.4).
 - 6: Pokud se průměrná šablona významně mění, vrať se na 3.
-

Průměrná šablona se vypočte prostým průměrováním bodů všech šablon. Upravení měřítka na pevně danou velikost zajišťuje, aby se průměrná šablona během iterování například neustále nezmenšovala.

3.1.3 Deformování modelu

V dalším postupu se předpokládá, že je zarovnáno s šablon x_i . Pokud se vytvoří průměrný model a bude se tvarovat toto rozložení bodů, je možno vytvářet nové příklady, podobné šablonám z originální trénovací sady. Musí se však také rozhodnout, zda se může tento nový tvar modelu považovat za přijatelný.

Přesněji se tedy hledá parametrizovaný model $x = M(b)$, kde b je vektor s parametry modelu. Tento model lze použít pro generování nových tvarů, respektive nových vektorů x . Aby byl zjednodušen problém výpočtu nových tvarů modelu, provede se nejprve PCA. Díky tomu se zredukuje vícerozměrná data, jak je již uvedeno v kapitole 2.4.1.

Postup [4] pro získání deformovatelného modelu je uveden ve výpisu algoritmu 2.

Algoritmus 2 Získání deformovatelného modelu

- 1: Vypočti průměrný tvar ze zarovnané sady šablon.

$$\bar{x} = \frac{1}{s} \sum_{i=1}^s x_i \quad (3.6)$$

- 2: Vypočti kovarianci (odchylku) dat.

$$S = \frac{1}{s-1} \sum_{i=1}^s (x_i - \bar{x})(x_i - \bar{x})^T \quad (3.7)$$

- 3: Vypočti vlastní čísla λ a k nim odpovídající vlastní vektory ϕ z kovarianční matice S . Vlastní čísla seřaď tak, aby $\lambda_i \geq \lambda_{i+1}$.
-

Matice ϕ po provedení PCA obsahuje t vektorů odpovídajících seřazeným vlastním číslům. Libovolnou šablonu z trénovací sady lze získat deformací průměrného modelu \bar{x} :

$$x \approx \bar{x} + \phi b, \quad (3.8)$$

kde $\phi = (\phi_1 | \phi_2 | \dots | \phi_t)$ a b je t rozměrný vektor určený:

$$b = \phi^T (x - \bar{x}). \quad (3.9)$$

Vektor b , tvoří parametry deformačního modelu. Změnou členů vektoru b je možno měnit tvar modelu x , pomocí rovnice 3.8.

3.1.4 Určení počtu módů

Počet hodnot (t) vlastních čísel (λ) a jim odpovídajících vlastních vektorů (ϕ) (tj. módů), pomocí kterých se bude model tvarovat lze určit mnoha způsoby. Nejjednodušším způsobem, jak určit počet módů, je pravděpodobně zvolit procento (např. 98%) rozdílnosti dat v tréninkové sadě, které je potřeba zachovat [2]. Díky tomu se zahrnou do modelu módy, které mají na jeho tvar rozhodující vliv a navíc se zjednoduší výpočet tím, že se zanedbají módy, které na celkový tvar modelu mají jen nepatrný vliv.

Model tvoří λ_i vlastních čísel kovariační matice (S) určené z tréninkových dat. Každé vlastní číslo udává proměnlivost dat ve směru odpovídajícího vlastního vektoru (ϕ_i). Celková proměnlivost v trénovacích datech je určena součtem všech vlastních hodnot, $V_t = \sum_{i=1}^t \lambda_i$. Počet módů t , které se zachovávají, lze určit tak, že:

$$\sum_{i=1}^t \lambda_i \geq f_v V_T, \quad (3.10)$$

kde f_v určuje procento rozdílnosti, které se zachová (např. 98%). V konkrétním případě kdy 300 obličejů v trénovací sadě bylo označeno 133 body, výsledný model zahrnoval při zachování 98% rozdílnosti dat 36 módů [4].

3.1.5 Omezující podmínky

Během vytváření nových tvarů modelu se lze setkat s problémem, kdy se musí rozhodnout, zda je nový tvar modelu realistický (tzn. zda je daný tvar obličejem). Tvar modelu určuje sada parametrů b (rovnice 3.9).

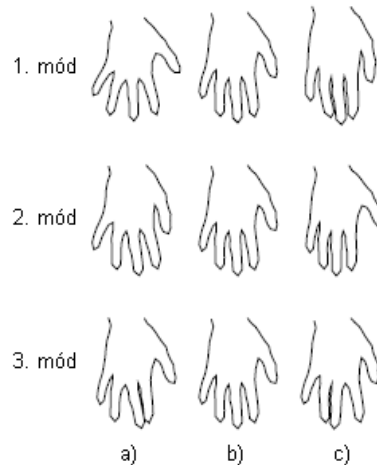
Sadu parametrů b je označena jako nový možný tvar modelu pokud $p(b) \geq p_t$, kde p_t je vhodný práh. Jestliže $p(b) < p_t$ upraví se tvar modelu x tak, aby splňoval omezující podmínky.

V našem případě se omezí sadu parametrů b pomocí vlastních čísel (λ) jednotlivých vektorů. Pro každý parametr platí:

$$|b_i| \leq 3\sqrt{\lambda_i}. \quad (3.11)$$



Obrázek 3.2: Ukázka šablon ruky. Upraveno z [4].



Obrázek 3.3: Ukázka tvarování prvních tří módů modelu ruky. Upraveno z [4].

3.1.6 Příklady modelů

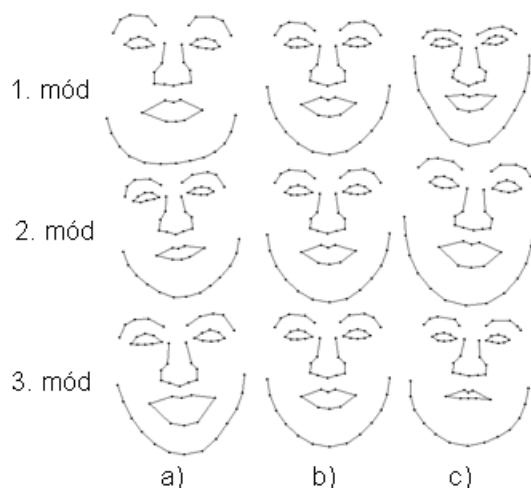
Sada určená k trénování algoritmu obsahuje šablony ruky, (obrázek 3.2). Po vytvoření statistického modelu x , lze tento model tvarovat pomocí rovnice 3.8. Obrázek 3.3 demonstruje deformaci tvaru modelu změnou prvních tří parametrů $b_{1,2,3}$ v rozmezí $b_{1,2,3} = -3\sqrt{\lambda_{1,2,3}}$ (obrázek 3.3 a) přes průměrný tvar modelu $b_{1,2,3} = 0$ (obrázek 3.3 b), až po $b_{1,2,3} = 3\sqrt{\lambda_{1,2,3}}$ (obrázek 3.3 c). Parametry vektoru b , kromě toho, který je zrovna měněn, jsou vždy rovny 0. První mód má na deformaci největší vliv - v důsledku jeho změny se mění tvar celého modelu. Třetí mód modelu určuje polohu prostředníčku.

Druhý statistický model byl vytvořen na základě trénovací sady obsahující 40 obrázků od 20 osob. Každá osoba poskytla jeden neutrální snímek a jeden snímek s úsměvem. Šablony k této trénovací sadě tvoří 58 bodů. Výsledný model je možno deformovat pomocí 21 módů. První 3 módy předvádí obrázek 3.4. Jednotlivé módy jsou opět nastavovány na maximální rozsah od $b_{1,2,3} = -3\sqrt{\lambda_{1,2,3}}$ (obrázek 3.4 a) přes průměrný tvar modelu $b_{1,2,3} = 0$ (obrázek 3.4 b), až po $b_{1,2,3} = 3\sqrt{\lambda_{1,2,3}}$ (obrázek 3.4 c). Pro všechny ostatní parametry b_i , které nejsou tvarovány, platí $b_i = 0$.

3.2 Active Shape Models

Aby bylo možné objekty na obrázku popisovat pomocí modelů, je potřeba najít sadu parametrů, při které model nejlépe odpovídá hledanému objektu v obrázku. Tato sada parametrů definuje tvar (b), pozici (X_t, Y_t), natočení (Θ) a změnu měřítka (s) modelu tvaru. Později mohou být takto získané informace použité k měření či rozpoznávání objektů [5].

Nejdříve je zapotřebí vytvořit instanci statistického modelu tvaru x (část 3.1) a určit jeho počáteční aproximaci. Iterační algoritmus mapující statistický model tvaru na objekt



Obrázek 3.4: Ukázka tvarování prvních tří módů modelu obličeje.

postupuje ve dvou fázích. V prvním kroku se hledají pro každý bod modelu nové, vhodnější body. V druhé fázi dochází k deformaci modelu tak, aby se nejlépe přizpůsobil nově určeným bodům.

Aby byla vylepšena shoda instance modelu x a objektu na obrázku, postupuje se na základě algoritmu 3.

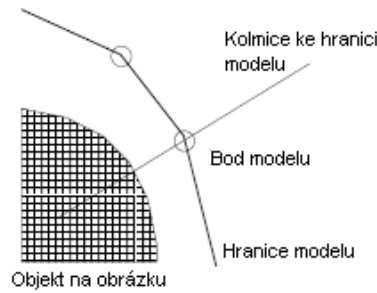
Algoritmus 3 Vylepšení shody instance modelu x a objektu na obrázku

- 1: Prozkoumej oblast na obrázku okolo každého bodu modelu X_n a urči nejlepší pozici pro nový bod X'_n (část 3.2.1)
 - 2: Uprav parametry modelu (X_t, Y_t, s, Θ, b) , aby se model co nejlépe shodoval s nově nalezenými body X'_n (část 3.2.2)
 - 3: Opakuj dokud dochází k významné změně parametrů modelu.
-

3.2.1 Získání nových bodů

K získání nových bodů modelu se prohledává okolí každého bodu šablony [3]. Jednotlivými body šablony jsou vedeny pomyslné kolmice k hranici modelu, jak je znázorněno na obrázku 3.5. Podél těchto kolmic se hledají nové body, které nejlépe splňují podmínky pro zvolený bod. Mezi tyto podmínky může například patřit informace, že se bod nachází na hraně. V takovém případě by se hledala podél kolmice k hranici modelu nejvýraznější hrana na obrázku. Tato pozice by pak byla navržena jak nová vhodná pozice pro nový bod.

Musí se však brát v potaz, že všechny body neodpovídají nejvýraznějším hranám, mohou také označovat slabší hrany nebo konstrukci objektu. V takovém případě se za nejlepší postup považuje, naučit se ze sady pro trénování modelu, jaké oblasti na cílovém obrázku přesně hledat.



Obrázek 3.5: Při hledání nových bodů se proloží každým bodem šablony kolmice ke hranici modelu. Upraveno z [3].

První možností je vytvoření statistických modelů struktury obrázku okolo každého bodu. Během určování hranice se hledají body, které nejlépe odpovídají těmto modelům. Tato možnost bude podrobně rozebrána v části *určení nové pozice bodu*.

Druhou možností je nakládat s tímto problémem jako s třídící (klasifikační) úlohou. Mohou se shromáždit příklady rysů ze správné oblasti, a příklady rysů poblíž nesprávných oblastí. Na základě těchto dat lze vytvořit klasifikátor o dvou třídách. Klasifikátor lze použít k rozhodnutí, který bod nejlépe odpovídá zvolené oblasti.

Určení nové pozice bodu

K definování nové pozice bodu modelu lze využít informací získaných ze sady určené pro trénování modelu. Za zdroj informací poslouží hodnoty intenzity jednotlivých pixelů v obrázku. V okolí bodu (X_i) modelu se nasnímají hodnoty intenzity (k pixelů na každou stranu od zvoleného bodu) a uloží se do profilu. Na obrázku 3.6 a) je naznačeno, jak probíhá vzorkování pixelů ($k = 3$) a jejich zarovnání do profilu je předvedeno na obrázku 3.6 b). Tento postup se provede pro každý obrázek obličeje ze sady určené pro trénování. Vygenerované profily s hodnotami intenzity v okolí bodu se uloží do vektoru g_i . Vektor g_i bude mít rozměr $j \times (2k + 1)$, kde j udává počet obrázků v trénovací sadě.

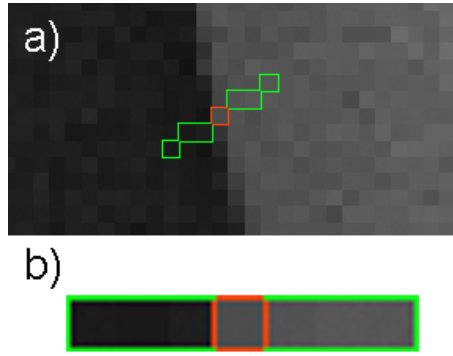
Aby hodnoty intenzity nebyli ovlivněny změnou osvětlení, ukládají se do profilu místo konkrétních hodnot intenzity, hodnoty změny intenzity:

$$I_i = I_{i-1} - I_{i+1}, \quad (3.12)$$

kde I_i je hodnota intenzity na pozici i . Takto upravené hodnoty se normalizují:

$$g_i \rightarrow \frac{1}{\sum_{s=1}^j |g_{is}|} g_i, \quad (3.13)$$

jednotlivé profily uložené ve vektoru g_i se podělí součtem všech profilů, kde j je počet obrázků v trénovací sadě.



Obrázek 3.6: a) Snímání hodnot intenzity podél kolmice k hranici obrázku ($k=3$), b) Vytvoření profilu.

Mahalanobisova vzdálenost

K určení nejlepší pozice pro nový bod (X'_i) lze použít Mahalanobisovu vzdálenost [1]. K výpočtu Mahalanobisovy vzdálenosti, je potřeba znát průměrný normalizovaný vzorek intenzity pro daný bod a kovarianční matici demonstrující odchylku dat. Průměrný vzorek g' je určen z vektoru g_i jako průměr jednotlivých sloupců, jedná se tak o vektor s $(2k + 1)$ hodnotami. Kovarianční matici S_g lze získat podobně jako v části 3.1.3. Mahalanobisova vzdálenost vzorku z obrázku ($f(g_s)$) od průměrné hodnoty (g') je určena jako

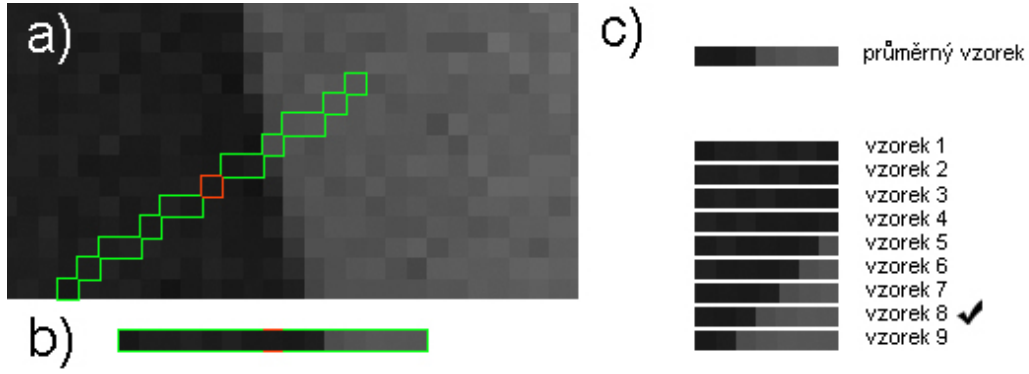
$$f(g_s) = (g_s - g')^T S_g^{-1} (g_s - g') \quad (3.14)$$

Během hledání nového bodu (X_i), se vytvoří profil o velikosti m pixelů ($m > k$) na každou stranu od aktuálního bodu (X_i) (obrázek 3.7 a), b). Poté se testuje jeho Mahalanobisova vzdálenost pro každý z $2(m - k) + 1$ možných vzorků. Na závěr se porovná Mahalanobisova vzdálenost jednotlivých vzorků a vybere se ten vzorek, pro který je $f(g_s)$ nejmenší (obrázek 3.7 c). Souřadnice středu nejlepšího vzorku určují nový bod (X'_i).

3.2.2 Přizpůsobení modelu novým bodům

Při aktivní deformaci modelu tvaru je potřeba znát algoritmus, který co nejlépe přizpůsobí model (x) zadaným bodům (Y). Parametry modelu (tvar b , pozice X_t, Y_t , natočení Θ a změna měřítka s) je potřeba upravit tak, aby nový model nejlépe odpovídal zadaným bodům (Y). Aby bylo možno model deformovat, musí se nejprve jeho instance převést (transformovat) do prostoru, ve kterém se nacházejí body Y . Pozice bodů modelu v obrázku je určena:

$$x = T_{X_t, Y_t, \Theta, s}(\bar{x} + \phi b), \quad (3.15)$$



Obrázek 3.7: a) Snímání hodnot intenzity podél kolmice k hranici obrázku ($m = 7$), b) Vytvoření profilu (9 vzorků), c) Vytvoření jednotlivých vektorů pro porovnání a výběr vektoru s nejmenší Mahalanobisovou vzdáleností

kde $\bar{x} + \phi b$ je instance modelu tvaru jak bylo uvedeno v 3.1.3. Funkce $T_{X_t, Y_t, \Theta, s}$ skládající se z transformačních matic pro rotaci, translaci a změnu měřítka je pro daný bod (x, y) určena:

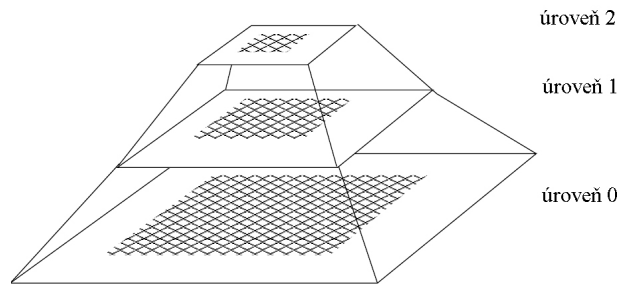
$$T_{X_t, Y_t, s, \Theta} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} X_t \\ Y_t \end{pmatrix} + \begin{pmatrix} s \cos \Theta & s \sin \Theta \\ -s \sin \Theta & s \cos \Theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \quad (3.16)$$

Funkce T provádí posunutí bodu o X_t, Y_t , rotaci o úhel Θ a změnu měřítka pomocí s .

Pro nalezení nejlepší shody instance modelu x a nových bodů Y , je třeba minimalizovat vzdálenost mezi body Y a instancí modelu x . Iterační algoritmus [4], který demonstruje jak toho dosáhnout, je uveden ve výpisu algoritmu 4.

Po ukončení předchozího algoritmu lze vytvořit instanci modelu x s novými parametry tvaru b pomocí:

$$x = (\bar{x} + \phi b) \quad (3.20)$$



Obrázek 3.8: Pyramida obrázků s různými úrovněmi rozlišení. Upraveno z [4].

Algoritmus 4 Nalezení nejlepší shody instance modelu x a nových bodů Y

- 1: Inicializuj parametry modelu tvaru, b , na nulu.
- 2: Vytvoř instanci modelu $x = \bar{x} + \phi b$.
- 3: Najdi parametry (X_t, Y_t, Θ, s) které nejlépe mapují x na Y .
- 4: Proveď inverzní transformaci T^{-1} pro všechny body Y . Tím se dosáhne transformace bodů Y do počátku souřadnic.

$$y = T_{X_t, Y_t, \Theta, s}^{-1}(Y), \quad (3.17)$$

- 5: Zarovnej y s instancí modelu x .

$$y' = \frac{y}{y \cdot \bar{x}} \quad (3.18)$$

- 6: Uprav parametry tvaru modelu b , aby odpovídaly y' . Deformuj model pomocí:

$$b = \phi^T(y' - \bar{x}) \quad (3.19)$$

- 7: Aplikuj omezující podmínky (část 3.1.5).
 - 8: Jestliže se parametry tvaru modelu b významně mění vrať se na 2.
-

3.2.3 ASM s více úrovněmi rozlišení

K vylepšení výkonnosti a robustnosti algoritmu ASM, lze tento algoritmus nasadit na obrázky s více úrovněmi rozlišení. Vytvořením více úrovní rozlišení obrázků se algoritmus ASM zrychlí a je odolnější proti uvíznutí v nesprávné pozici [4].

Pro každý obrázek z testovací i trénovací sady se vytvoří pyramida s různými úrovněmi rozlišení, jak je naznačeno na obrázku 3.8. Originální obrázek se nachází na úrovni 0, jeho rozmazáním a podvzorkováním lze získat obrázek s poloviční šířkou i výškou (úroveň 1). Opětovnou aplikací rozmazání a podvzorkování lze získat obrázek se čtvrtinovou šířkou a výškou oproti originálnímu obrázku (úroveň 2).



Obrázek 3.9: Obrázek obličeje se čtyřmi úrovněmi rozlišení. Upraveno z [16].

Při provádění algoritmu ASM se začíná na nejvyšší úrovni pyramidy (obrázek s nejmenším rozlišením), to dovoluje výrazné změny polohy a tvaru modelu v prvních iteracích algoritmu. Na nižších úrovních pyramidy se provádí detailnější úpravy polohy a zejména

tvaru modelu. Na obrázku 3.9 lze vidět snímek se čtyřmi úrovněmi rozlišení.

3.2.4 Příklady ASM

Pro první příklad byl vytvořen statistický model pomocí databáze IMM [20] z 20 obrázků s počtem 13 módů. Byla vytvořena pyramida rozlišení obrázku o čtyřech úrovních (80x60, 160x120, 320x240 a 640x480). Start algoritmu ASM začíná na obrázku s nejnižším rozlišením 80x60 (úroveň 3) a postupně pokračuje až po originální rozlišení obrázku (úroveň 0). Postupným iterováním algoritmu ASM dochází k přibližování modelu k obličeji na obrázku.



Obrázek 3.10: Hledání rysů tváře pomocí metody ASM. Demonstrace tvaru modelu při iterování algoritmu.

Obrázek 3.10 demonstruje průběh lokalizace obličeje. Na 3.10 a) lze vidět počáteční stav modelu, 3.10 b) zobrazuje stav přibližování po 2 iteracích, na 3.10 c) po 6 iteracích a na 3.10 d) konečná podoba lokalizace tváře po 24 iteracích algoritmu ASM. Na každé úrovni

rozlišení bylo provedeno 6 iterací. V prvních iteracích algoritmu dochází k nejradikálnějším změnám pozice a tvaru modelu.

Obrázek 3.11 demonstruje, jak může metoda ASM selhat, pokud se zvolí počáteční pozice příliš daleko od cíle. Při prohledávání okolí kolem hranice modelu se metodě ASM nepodařilo najít vhodné body a posunout tak šablonu do lepší pozice.



Obrázek 3.11: Demonstrace selhání metody ASM, způsobené nevhodnou počáteční inicializací. Upraveno z [4].

3.3 Statistical Models of Appearance

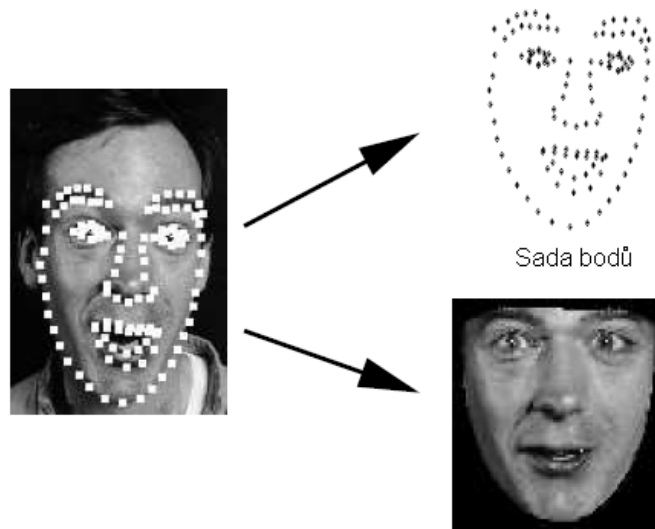
Abychom získali kompletní obrázek objektu, musí se modelovat jak tvar, tak textura. V předchozí části bylo popsáno jak vytvořit statistický tvar modelu (sekce 3.1). V této části si popíšeme, jak vytvořit statistický model textury a jeho propojení se statistickým modelem tvaru. Spojení těchto modelů může být použito k vytváření realistických obrázků.

3.3.1 Statistický model textury

K vytvoření statistického modelu textury se nejdříve zarovná každý obrázek z trénovací sady tak, aby kontrolní body seděly s tvarem obličeje. Tím se odstraní nežádoucí změny textury způsobené rozdíly ve tvaru obličejů. Pro vytvoření vektorů g_{im} , které uchovávají informaci o intenzitě jednotlivých obrázků, se musí postupně nasnímat hodnoty intenzity z jednotlivých obrázků. Vzorkování probíhá po celé oblasti ohraničené kontrolními body. Získané hodnoty intenzity však ještě obsahují nepřesnosti. Obrázek 3.12 obsahuje označený trénovací snímek, šablonu obličeje a část obličeje normalizovaného do průměrného tvaru.

Pro snížení efektu, který způsobuje změna osvětlení, se vzorky intenzity (g_{im}) normalizují pomocí α a β .

$$g = (g_{im} - \beta)/\alpha \quad (3.21)$$



Obrázek 3.12: Každý obrázek z trénovací sady se rozdělí na šablonu tvaru a na informaci o textuře. Upraveno z [4].

K získání průměru normalizovaných dat (vektor \bar{g}) slouží rekurzivní algoritmus. Po každé iteraci je vektor \bar{g} znovu vypočítán jako průměr vektorů g_{im} . Hodnoty α a β jsou určeny tak, aby vektor se vzorky intenzit (g_{im}) nejlépe odpovídal průměrnému vektoru (\bar{g}). Hodnoty α a β nutné pro získání normalizovaného g_{im} jsou dány:

$$\alpha = g_{im} \cdot \bar{g}, \beta = (g_{im})/n, \quad (3.22)$$

kde n je počet prvků vektoru g_{im} .

Rekurzivní algoritmus k získání průměru normalizovaných dat \bar{g} je uveden ve výpisu algoritmu 5.

Algoritmus 5 Získání průměru normalizovaných dat

- 1: Vyber jeden z vektorů g_{im} jako první odhad \bar{g} .
 - 2: Zarovnej vektory g_{im} s \bar{g} pomocí rovnice 3.21 a rovnice 3.22.
 - 3: Vypočti nový průměr dat \bar{g} ze zarovnané sady vektorů g_{im} .
 - 4: Pokud se \bar{g} významně mění, pokračuj bodem 2.
-

Aplikováním PCA (kapitola 2.4.1) na normalizovaná data lze získat model textury (podobně jako v části 3.1.3, kde se jednalo o model tvaru):

$$g = \bar{g} + \phi_g b_g, \quad (3.23)$$

kde \bar{g} je průměr normalizovaných vektorů intenzity, ϕ_g je matice vlastních vektorů a b_g je vektor s parametry modelu (podobně jako v části 3.1.3).

Texturu obrázku lze vytvořit pomocí parametrů textury, b_g , a parametrů normalizace α, β . Tyto parametry se uvádí ve vektoru $u = (u_1, u_2)^T = (\alpha - 1, \beta)^T$. Textura je pak definována:

$$g_{im} = T_u(\bar{g} + \phi_g b_g) = (1 + u_1)(\bar{g} + \phi_g b_g) + u_2. \quad (3.24)$$

3.3.2 Kombinace modelů

Model tvaru a model textury lze sloučit do jednoho modelu vzhledu pomocí vektorů b_s (3.1.3) a b_g (3.3.1). Aby se vytvořil vztah mezi tvarem a texturou, použije se další PCA na získaná data. Pro každou instanci modelu se vygeneruje vektor, který propojuje oba modely pomocí:

$$b = \begin{pmatrix} W_s b_s \\ b_g \end{pmatrix} = \begin{pmatrix} W_s \phi_s^T (x - \bar{x}) \\ \phi_g^T (g - \bar{g}) \end{pmatrix}, \quad (3.25)$$

kde W_s je matice s hodnotami vah pro každý parametr vektoru b_s (ovlivňuje model tvaru), dovolující rozdíly mezi modelem tvaru a textury. Po aplikování PCA na tyto vektory se vypočte celkový model vzhledu b :

$$b = \phi_c c, \quad (3.26)$$

kde ϕ_c jsou vlastní vektory a c je vector, jehož hodnoty ovlivňují jak tvar, tak úroveň šedi (texturu) modelu.

Statistický model vzhledu lze deformovat pomocí rovnic:

$$x = \bar{x} + \phi_s W_s^{-1} \phi_{cs} c, \quad g = \bar{g} + \phi_g \phi_{cg} c, \quad (3.27)$$

kde \bar{x} je průměrný tvar modelu (3.1.3) a \bar{g} je průměr normalizovaných vektorů intenzity (3.3.1). ϕ_c je určeno:

$$\phi_c = \begin{pmatrix} \phi_{cs} \\ \phi_{cg} \end{pmatrix}, \quad (3.28)$$

po zavedení substituce:

$$Q_s = \phi_s W_s^{-1} \phi_{cs}, \quad Q_g = \phi_g \phi_{cg}, \quad (3.29)$$

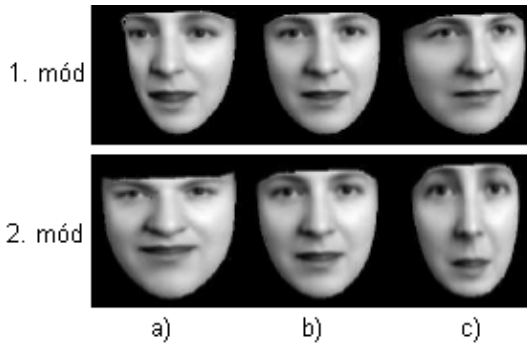
lze model deformovat pomocí rovnic:

$$x = \bar{x} + Q_s c, \quad g = \bar{g} + Q_g c. \quad (3.30)$$

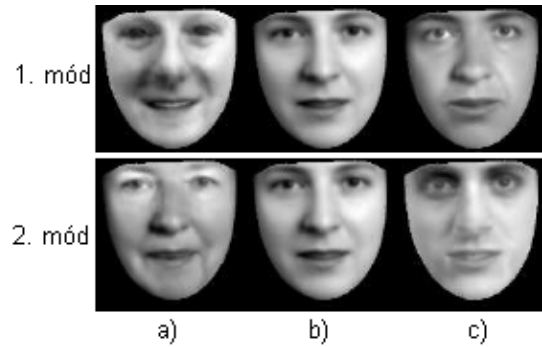
3.3.3 Příklady Statistical Models of Appearance

Pro vytvoření modelu vzhledu [4] umožňující deformovat jak tvar, tak texturu, byla použita trénovací sada o celkovém počtu 400 tváří. Každý obrázek byl označen šablonou s 122 body (obrázek 3.12). Z této sady byl vytvořen statistický model tvaru s 23 módy (část 3.1) a statistický model textury se 114 módy (část 3.3.1). Sloučením těchto modelů vznikl model vzhledu (část 3.3.2) pouze s 80 módy, které zajišťují 98% variabilitu modelu (část 3.1.4).

Obrázky 3.13 a 3.14 ukazují vliv změny prvních dvou módů na tvar (obrázek 3.13) a vliv změny prvních dvou módů na texturu (obrázek 3.14). Velikost změny jednotlivých parametrů modelů $b_{1,2}$ je prováděna v rozmezí od $b_{1,2} = -3\sqrt{\lambda_{1,2}}$ (obrázek 3.13, 3.14 a), přes průměrný tvar modelu $b_{1,2} = 0$ (obrázek 3.13, b), až po $b_{1,2} = 3\sqrt{\lambda_{1,2}}$ (obrázek 3.13, 3.14 c). První mód (část 3.1.4) odpovídá největší hodnotě vlastního čísla kovarianční matice a má tedy největší vliv na deformaci tvaru případně textury. Vliv na deformaci modelu se zvyšujícím se umístěním klesá - osmdesátý mód má na deformaci modelu vliv nejmenší.



Obrázek 3.13: Vliv změny prvních dvou módů na tvar modelu. Převzato z [4].



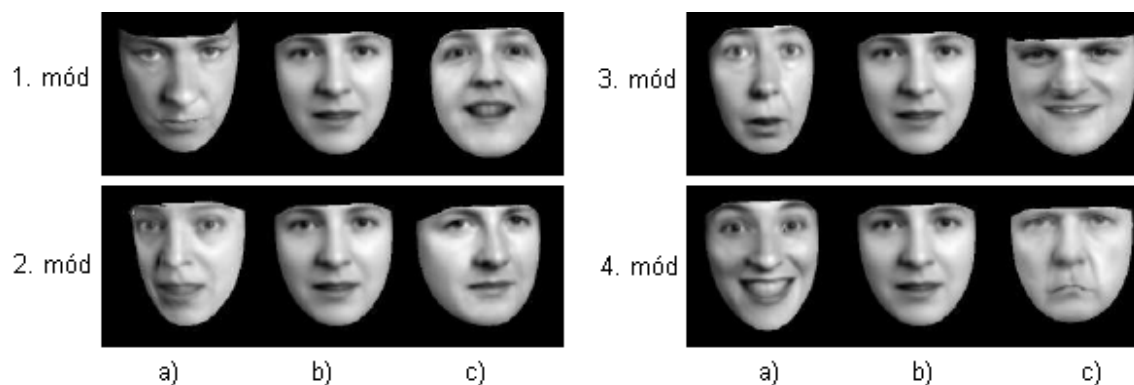
Obrázek 3.14: Vliv změny prvních dvou módů na texturu modelu. Převzato z [4].

Obrázek 3.15 demonstruje jednotlivě vliv změny prvních čtyř módů společného modelu vzhledu. Opět jsou jednotlivé parametry měněny v rozmezí od $b_{1,2,3,4} = -3\sqrt{\lambda_{1,2,3,4}}$ (obrázek 3.15 a), přes průměrný tvar modelu $b_{1,2,3,4} = 0$ (obrázek 3.14 b), až po $b_{1,2,3,4} = 3\sqrt{\lambda_{1,2,3,4}}$ (obrázek 3.15 c).

3.4 Active Appearance Models

V části 3.2 byl představen algoritmus ASM, který dokáže deformovat model tvaru tak, že se přizpůsobí objektu (obličej) na obrázku. Tento algoritmus využívá šablonu bodů obrázku a informace o intenzitě v jejich blízkém okolí. Metoda Active Appearance Models (AAM) [4] využívá k lokalizaci a případnému rozpoznání veškeré dostupné informace o intenzitě z obrázku, tedy jak tvar, tak i texturu celého objektu. Díky tomu lze vytvořit realistický obrázek co nejpodobnější originálnímu obrázku.

Metoda AAM využívá statistického modelu vzhledu popsaného v části 3.3. Změnou jeho parametrů dochází k postupnému zpřesňování modelu a k vytvoření syntetického obrázku.



Obrázek 3.15: Vliv změny prvních čtyř módu na model vzhledu, změna se týká jak tvaru tak textury modelu. Převzato z [4].

Algoritmus AAM je oproti ASM složitější a náročnější, ale jeho použitím lze získat více informací o hledaném objektu. Tyto informace je možno využít k zlepšení výsledků například u rozpoznání obličeje. Kde vektor s hodnotami intenzity celého obličeje dosahuje při identifikaci výrazně lepších výsledků než vektor s body charakterizující tvar objektu.

3.4.1 Rozdíly mezi AAM a ASM

Algoritmus AAM je detailně rozebrán v práci T. F. Cootese a C. J. Taylora - Statistical Models of Appearance for Computer Vision [4]. Zde jsou uvedeny nejpodstatnější rozdíly oproti ASM (3.2):

1. ASM využívá informace o textuře obrázku pouze v malé oblasti okolo každého bodu šablony (3.2.1).

AAM má k dispozici kompletní data o textuře obrázku v celé oblasti ohraničené obálkou bodů šablony.

2. ASM určuje novou pozici modelu na základě prohledávání obrázku podél kolmic ke hranici modelu (3.2.1).

AAM využívá informace o textuře pod aktuální pozicí modelu.

3. ASM se snaží minimalizovat vzdálenost mezi body modelu a novými vhodnějšími body nalezenými na obrázku (3.2.2).

AAM má za úkol minimalizovat rozdíl mezi vytvořeným umělým modelem a cílovým objektem na obrázku.

3.4.2 Příklady AAM

Algoritmus AAM byl využit pro lokalizaci tváří [4]. Obrázek 3.16 demonstruje nejlepší shodu modelu AAM a tváře na obrázku. Pro tento pokus byla využita anotovaná šablona s body ohraničujícími tvar obličeje.



Obrázek 3.16: Zrekonstruovaný obličej pomocí metody AAM (vlevo) a původní obrázek (vpravo). Převzato z [4].

Obrázek 3.17 znázorňuje postupné přibližování během jednotlivých kroků algoritmu AAM. Počáteční pozice modelu byla nastavena na střed tváře.



Obrázek 3.17: Postupné iterování algoritmu AAM. Upraveno z [4].

Kapitola 4

Návrh a implementace

Pro demonstrační aplikaci, která tvoří programovou část této diplomové práce, byla vybrána metoda Active Shapes Models. Tato metoda, jak již bylo uvedeno v části 3.2 využívá statistického deformačního modelu (sekce 3.1) a přizpůsobuje jej obličej na obrázku.

4.1 Návrh aplikace

Aplikace implementující metodu ASM má za úkol demonstrovat lokalizaci obličeje pomocí této metody. Vytvořená aplikace musí umožnit provedení experimentů pro vyhodnocení úspěšnosti lokalizace obličeje a úspěšnosti rozpoznání obličeje pomocí metody ASM. Návrh aplikace lze rozdělit na několik částí:

- návrh SSM
- návrh ASM
- návrh demonstračního a testovacího rozhraní

Návrh metod SSM a ASM vychází z jejich popisu v částech 3.1 (SSM) a 3.2 (ASM).

4.1.1 Návrh demonstračního a testovacího rozhraní

Testovací rozhraní aplikace tvoří GUI vytvořené pomocí Java knihovny JFC Swing [15]. Skládá se ze čtyř panelů umožňující demonstraci či testování metod SSM, ASM a jednoho panelu pro nápovědu.

Demonstrace metody SSM

Umožňuje vytvoření statistického modelu na základě tréninkové sady obrázků a jeho manuální deformaci změnou hodnot jednotlivých módů modelu (3.1.4).

Demonstrace metody ASM

Slouží k demonstraci lokalizace obličeje pomocí metody ASM. Tento proces zahrnuje vytvoření testovacích obrázků ve čtyřech úrovních rozlišení, vytvoření statistického modelu pro jednotlivé body šablony a samotné hledání nových vhodnějších bodů a deformaci modelu.

Testování úspěšnosti lokalizace pomocí metody ASM

Testování probíhá ve dvou fázích. V první fázi je vytvořen model. V druhé fázi dochází k přizpůsobení (deformaci) modelu na základě vstupního obrázku. Následuje vyhodnocení úspěšnosti lokalizace obličeje neboli vypočtení euklidovské vzdálenosti jednotlivých bodů modelu od referenčních bodů získaných z přiloženého datového souboru. Druhá fáze se provede pro všechny testovací obrázky.

Testování rozpoznávání obličejů pomocí LDA

Slouží k vytvoření dat pro rozpoznávání obličejů pomocí LDA a k samotnému rozpoznání obličejů metodou LDA.

Nápověda k programu

Nápověda k jednotlivým částem programu.

4.1.2 Úložiště dat

Data pro demonstrační aplikaci jsou uložena v souborech. Jedná se o obrázky obličejů a k nim odpovídající soubory s informacemi o referenčních bodech definující tvar šablony. Formát dat v souborech s body šablony se liší podle dané databáze (více v 5.1).

4.1.3 Implementační jazyk

Za implementační jazyk byl vybrán jazyk Java [12], [11], [13]. Mezi výhody tohoto jazyka lze zařadit jeho snadnou přenositelnost mezi různými platformami a také snadné vytváření grafického uživatelského rozhraní pomocí knihovny JFC Swing. Další výhodou je množství dostupných knihoven.

Důležitou součástí volby programovacího jazyka je i volba vývojového prostředí. Programová část byla vytvářena v NetBeans 6.8.

4.2 Implementace

Z návrhu aplikace vychází její implementace. Program se skládá ze dvou balíčků *template-matching* a *dataprocessing*. Balík *templatematching* obsahuje jednotlivé třídy potřebné pro vytvoření statistického modelu a jeho deformaci. Balík *dataprocessing* slouží k načítání dat ze souborů a zpracování obrázků.

4.2.1 Popis jednotlivých tříd

V následném výpisu je uveden stručný popis jednotlivých tříd programu. Jsou vynechány třídy, které nemají přímou souvislost s popisovanou problematikou deformačních modelů (například třída zajišťující chybové výpisy, třídy rozšiřující rozhraní `AbstractTableModel`, atd.).

Balík `templatematching`

TemplateMatchingView Třída definující grafické uživatelské rozhraní vytvořené pomocí knihovny JFC Swing.

TrainingData Třída, která slouží k vytvoření statistického modelu tvaru. Tato třída je implementována na základě popisu metody SSM (3.1).

FittingModel Třída, pomocí které lze model přizpůsobovat (deformovat) objektům na obrázku. Třída je implementována dle popisu metody ASM (3.2).

CovarianceMatrix Tato třída slouží k výpočtu kovarianční matice. Kovarianční matice se používá při výpočtu PCA (3.1.3) a Mahalanobisovy vzdálenosti (3.2.1).

GaussianPyramid Třída, která vytváří obrázky se čtyřmi úrovněmi rozlišení (3.2.3). Obrázky jsou vytvořeny změnou rozlišení z originálního obrázku.

PointsStatisticalModel Třída, pomocí které lze vytvořit statistický model pro jednotlivé body šablony modelu. Tato třída je implementována na základě popisu v části 3.2.1.

ImagePanel Tato třída se stará o vykreslování modelu na základě jednotlivých souřadnic bodů.

EuclideanDistance Třída, která umožňuje testování úspěšnosti lokalizace obličeje. Úspěšnost lokalizace je určena na základě výpočtu euklidovské vzdálenosti mezi modelem přizpůsobeným obličeji na obrázku a referenční sadou bodů.

LDA Třída sloužící k přípravě dat pro knihovnu LDA (4.2.3). Třída také vyhodnotí výsledky, které knihovna LDA poskytuje.

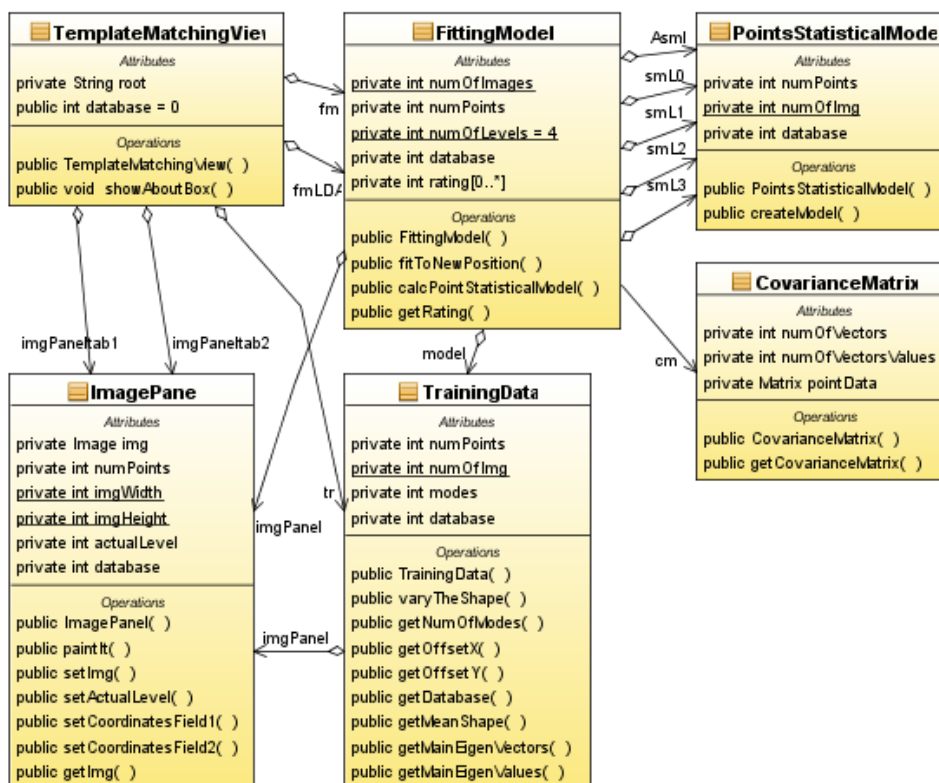
Balík `dataprocessing`

DataParser Třída sloužící ke zpracování souborů s informacemi o referenčních bodech pro databáze IMM i M2VTS (5.1).

ImageProcessing Třída, která se stará o zpracování obrázků pro účely metod deformačních modelů.

4.2.2 Schéma aplikace

Na obrázku 4.1 je pomocí UML diagramu tříd znázorněna implementace části programu. Schéma obsahuje šest tříd a demonstruje jejich vzájemný vztah. V příloze C je na obrázku C.1 schéma programu rozšířeno o další tři třídy.



Obrázek 4.1: Schéma části aplikace modelované pomocí UML diagramu tříd.

4.2.3 Použité knihovny

Pro implementaci programu byly využity dvě externí knihovny. Jedná se o knihovnu implementující metodu LDA a knihovnu pro práci s maticemi JAMA.

JavaStat

Knihovna pro statistické zpracování dat JavaStat [8] umožňuje zpracovávat a vyhodnocovat data metodou LDA. Princip této metody byl podrobně popsán v části 2.4.2. V této práci je metoda LDA využita pro rozpoznání obličejů na základě tvaru šablony deformačního

modelu. Podrobný postup rozpoznání obličejů bude uveden v kapitole Testování v části 5.3.

A Java Matrix Package (JAMA)

Jedná se o knihovnu napsanou v jazyce Java, která byla vytvořena spoluprací společnosti The MathWorks a institutu The National Institute of Standards and Technology (NIST) [23]. Tato knihovna umožňuje provádět základní operace s maticemi a nabízí i několik složitějších funkcí. Operace s maticemi, které knihovna umožňuje, a současně byly využity při implementaci, jsou uvedeny v tabulce 4.1.

Tabulka 4.1: Přehled operací, které umožňuje knihovna pro práci s maticemi JAMA.

Základní operace s maticemi	sčítání, odčítání, násobení (dvou matic i matice prvkem), dělení matice prvkem, transponování matice
Rozšířené operace s maticemi	výpočet vlastních čísel a vlastních vektorů, výpočet determinantu, výpočet inverzní matice

4.2.4 Určení počáteční aproximace modelu

Pro metodu ASM je důležitá počáteční pozice modelu. Při implementaci byly zvoleny dva způsoby. První možností je určení průměrné pozice středu obličeje podle pozice obličejů v trénovací sadě. Získaný průměr středů obličejů je pak při testování použit jako počáteční pozice modelu. Vzhledem k tomu, že některé obličeje jsou od vypočteného průměru výrazně vzdáleny, může u některých testovacích obrázků dojít k selhání metody ASM.

Druhým, přesnějším způsobem je využít přiloženého souboru s daty testovaného obrázku. Tento soubor se využívá pro porovnání úspěšnosti metody. Lze však z něj získat i geometrický střed obličeje. Oba dva způsoby jsou implementovány u demonstrace metody. U rozpoznávání obličejů, kde je podstatné, aby byla oblast s obličejem správně určena, se používá druhý způsob.

Kapitola 5

Testování

V následující kapitole budou nejprve představeny dostupné databáze obličejů. V další části budou provedeny dvě varianty experimentů. V první variantě se bude testovat úspěšnost lokalizace obličeje pomocí metody Active Shape Models na základě euklidovské vzdálenosti mezi deformovanými modely a anotovanou sadou snímků. V druhé fázi experimentů bude provedeno rozpoznávání obličejů pomocí metody LDA.

5.1 Dostupné databáze obličejů

5.1.1 The Extended M2VTS Database

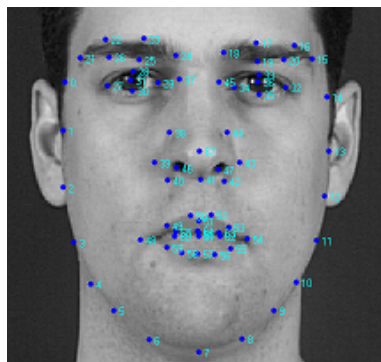
Databáze obličejů The Extended M2VTS Database [16] obsahuje 2360 snímků obličejů od 295 osob. Rozlišení těchto obrázků je 720x576 pixelů. Tyto snímky byly pořízeny během čtyř sezení, během kterých byly pořízeny vždy dvě fotografie. Příklad snímků obličejů databáze M2VTS je na obrázku 5.1.



Obrázek 5.1: Příklad snímků obličejů z databáze M2VTS. Upraveno z [16].

Snímky této databáze jsou označeny pomocí 68 bodů, aby mohly být využity například pro vytváření modelů nebo ověřování přesnosti při lokalizaci obličeje pomocí metod založených na rysech tváře. Příklad označení tváře je na obrázku 5.2.

Souřadnice bodů jsou uloženy v souboru, který má stejný název jako obrázek. Jen jeho příponu tvoří „.pts“. Soubor má následující formát:



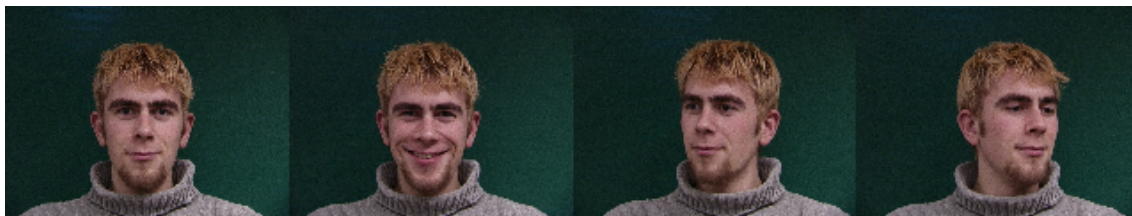
Obrázek 5.2: Příklad označení snímku obličeje z databáze M2VTS. Upraveno z [16].

```
version: 1          (This points file version no. can be ignored)
image_size_x: 720   (The image width in pixels)
image_size_y: 576   (The image height in pixels)
n_points: 68        (The number of labelled points on the image)
{
x1  y1
x2  y2
.   .
.   .
}
```

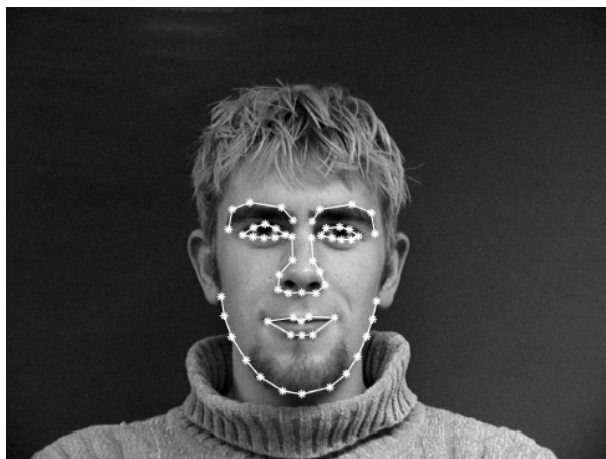
Pro každý bod je uložena souřadnice jeho polohy vzhledem k levému hornímu rohu obrázku.

5.1.2 The IMM Face Database

Na rozdíl od M2VTS databáze obličejů The IMM Face Database [21] obsahuje i snímky s různým výrazem (úsměv), obrázky s odlišným osvětlením a snímky s natočením hlavy na levou i pravou stranu. Databáze je však méně obsáhlá tvoří ji 240 fotografií od 40 osob. Příklad snímků z databáze IMM je na obrázku 5.3. Rozlišení jednotlivých obrázků je 640x480 pixelů a jsou označeny pomocí 58 bodů. Příklad označené tváře můžete vidět na obrázku 5.4.



Obrázek 5.3: Příklad snímků z databáze IMM. Upraveno z [21].



Obrázek 5.4: Příklad označené tváře z databáze IMM. Převzato z [20].

Souřadnice jednotlivých bodů jsou uloženy v souborech s příponou „.asf“. Soubory mají následující formát:

```
#####
#
# AAM Shape File - written: Monday May 08 - 2000 [15:22]
#
#####
#
# number of model points
#
58
#
# model points
#
# format: <path#> <type> <x rel.> <y rel.> <point#> <connects from> <connects to>
#
0 0 0.07690192 0.44500541 0 82 1
0 0 0.09916206 0.42914406 1 0 2
0 0 0.12925033 0.39573063 2 1 3
...
0 0 0.07579006 0.52910086 80 79 81
0 0 0.06128729 0.49762829 81 80 82
0 0 0.05858913 0.46610570 82 81 0
#
# host image
#
F1011flb.bmp
```

Význam jednotlivých zkratk je následující:

<path#> je množina bodů, do které daný bod patří

<type> je typ bodu

<x rel.> určuje relativní souřadnici x daného bodu

<y rel.> určuje relativní souřadnici y daného bodu

<point#> označuje číslo bodu

<connects from> je číslo předchozího bodu v dané smyčce

<connects to> je číslo následující bodu v dané smyčce

Souřadnice bodů jsou uvedeny v relativních souřadnicích. To znamená, že hodnota souřadnice je podělena šířkou (u souřadnice x) nebo výškou (u souřadnice y) obrázku. Například pro souřadnice $x = 30$, $y = 50$ v obrázku o rozměrech 640×480 platí, že jejich relativní souřadnice jsou pro $x = 30/640 = 0,046875$ a pro $y = 50/480 = 0,104167$. Tento způsob uložení souřadnic je pak nezávislý na změnách velikosti obrázku.

5.1.3 Další databáze obličejů

Existuje mnoho dalších databází obličejů [9]. Důvodem proč byly vybrány pro testování právě databáze IMM a M2VTS je zejména to, že jejich součástí jsou data s popisem obličejů pomocí množiny bodů. Tyto data jsou nezbytná pro vytvoření deformačních modelů. Dalším důvodem je také jejich dostupnost (v případě IMM) a rozsáhlost (M2VTS). Tím, že M2VTS obsahuje od každé osoby 8 snímků z čelního pohledu, je tato databáze vhodná i pro rozpoznání obličeje.

AT&T „The Database of Faces“ (formerly „The ORL Database of Faces“)

Tuto databázi tvoří 10 rozdílných snímků od každé ze 40 osob. Snímky byly pořizovány v několika sezeních. Osoby na obrázcích mají také různý výraz (otevřené / zavřené oči, usmívají se). Všechny snímky byly pořizovány proti tmavému pozadí [9].

PUT Face Database

Databáze PUT Face Database obsahuje téměř 10000 snímků od 100 osob. Osoby na obrázcích mají neutrální výraz. Od každé osoby byly pořízeny snímky s rotací hlavy zprava do leva (až 30 snímků), otáčením hlavy dokola (až 20 snímků), rotace se zvednutou hlavou (až 20 snímků), rotace se skloněnou hlavou (až 20 snímků) a snímky bez omezení výrazu obličeje (až 10 snímků). Všechny obrázky mají k dispozici ručně zadanou pozici rysů, jako jsou nos, oči, ústa [9].

5.2 Experiment 1 - lokalizace obličeje metodou ASM

Testování lokalizace obličeje metodou ASM probíhá ve dvou krocích. V první fázi je vytvořen model ze sady určené pro trénování metody. V dalším kroku je tento model pomocí metody ASM přizpůsoben obrázkům v testovací sadě. Pro každý testovací obrázek je vyhodnocena

úspěšnost lokalizace mezi zdeformovaným modelem a referenční šablonou bodů získanou z příloženého datového souboru.

Vyhodnocení je provedeno na základě euklidovské vzdálenosti mezi jednotlivými body modelu a body z referenční šablony. U jednotlivých testů je v tabulkách uvedena celková euklidovská vzdálenost bodů (celková ED), což je součet euklidovské vzdálenosti mezi všemi body. Dále je uvedena průměrná euklidovská vzdálenost (průměrná ED). To je celková euklidovská vzdálenost podělená počtem bodů v šabloně. Jedná se tedy o průměrnou vzdálenost jednoho bodu modelu od referenčního bodu. Poslední hodnotou pro porovnání přesnosti lokalizace obličejů je procento bodů, jejichž euklidovská vzdálenost od referenčního bodu je menší než 10 pixelů (v tabulkách uvedeno jako „hit rate“).

Pro tento experiment jsou využity databáze IMM 5.1.2 a XM2VTS 5.1.1. Osoby v testovací sadě jsou vyloučeny ze sady pro trénování modelu.

5.2.1 IMM databáze

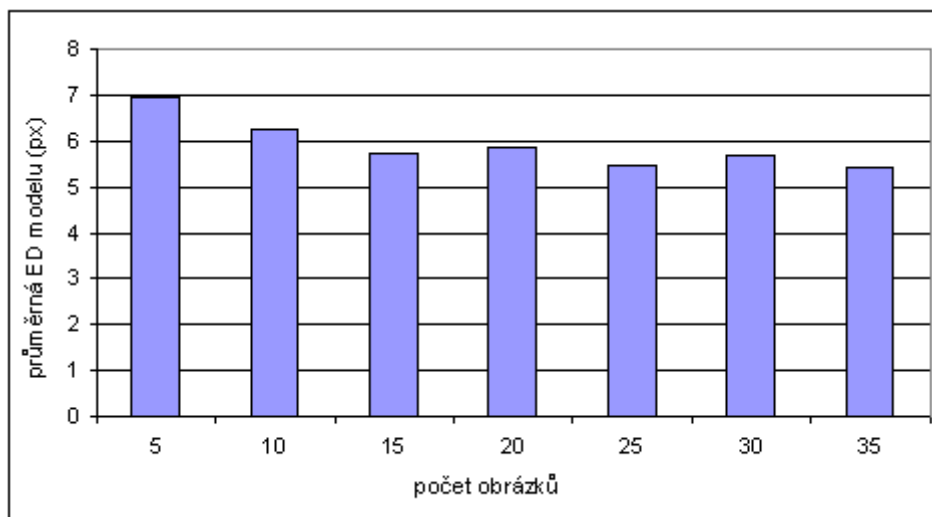
Vliv počtu obrázků

Při následujícím experimentu testovací sada obsahovala 5 obrázků od 5 osob. Bylo vytvořeno 7 trénovacích sad obrázků. Počet obrázků pro trénování modelu začal na 5 obrázcích a postupně se o 5 obrázků navyšoval až do celkového počtu 35 obrázků. Výsledky experimentu obsahuje tabulka 5.1 a závislost počtu obrázku na průměrné euklidovské vzdálenosti bodů modelu od referenční šablony se nachází v grafu na obrázku 5.5.

Tabulka 5.1: IMM - Vliv počtu obrázků v trénovací sadě na úspěšnost lokalizace pomocí deformačního modelu.

počet osob / obrázků	celková ED (px)	průměrná ED (px)	hit rate (%)	počet módů
5 / 5	405	6,97	82,07	3
10 / 10	363	6,26	87,24	7
15 / 15	333	5,75	87,93	10
20 / 20	340	5,86	85,52	13
25 / 25	316	5,45	89,66	15
30 / 30	331	5,70	88,28	17
35 / 35	313	5,40	88,62	19

Počet obrázků v tréninkové sadě má na úspěšnost modelu zásadní vliv. Vzhledem k tomu, že se jedná o obrázky pouze s čelním pohledem, dosáhla průměrné euklidovské vzdálenosti pod 6 pixelů již sada s 15 tréninkovými obrázky.



Obrázek 5.5: IMM - Vliv počtu obrázků v trénovací sadě na úspěšnost lokalizace pomocí deformačního modelu.

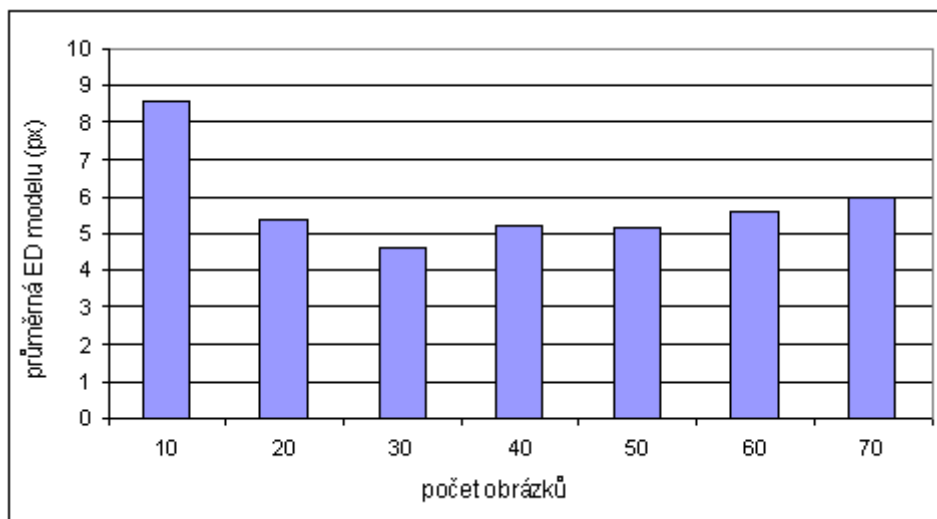
Vliv výrazu obličeje

V dalším experimentu testovací sadu tvoří obrázky s úsměvem od 5 osob. Tréninková sada je tvořena obrázky s čelním pohledem a obrázky s úsměvem, vždy jsou zahrnuty dva obrázky od jedné osoby. Výsledky experimentu s navyšováním počtu obrázků se nachází v tabulce 5.2 a v grafu na obrázku 5.6.

Tabulka 5.2: IMM - Vliv výrazu obličeje a počtu obrázků v trénovací sadě na úspěšnost lokalizace pomocí deformačního modelu.

počet osob / obrázků	celková ED (px)	průměrná ED (px)	hit rate (%)	počet módů
5 / 10	498	8,59	73,79	6
10 / 20	312	5,38	92,76	12
15 / 30	269	4,64	93,79	16
20 / 40	302	5,20	90,69	19
25 / 50	298	5,14	92,07	24
30 / 60	323	5,58	88,62	26
35 / 70	347	5,99	84,48	27

Nejvýraznější vzestup úspěšnosti lokalizace je mezi 10 a 20 obrázky. Navyšování počtu obrázku není přímo úměrné úspěšnosti lokalizace obličeje. Pro model je podstatné, aby tréninkové obrázky byly co nejvíce odlišné, v opačném případě dochází k přetrénování modelu.



Obrázek 5.6: IMM - Vliv výrazu obličeje a počtu obrázků v trénovací sadě na úspěšnost lokalizace pomocí deformačního modelu.

Pro kvalitu modelu je podstatné zahrnout do tréninkové sady takový počet obrázků, který pokrývá proměnlivost výrazů jednotlivých obličejů. Toto tvrzení lze ověřit pokud se do porovnání přidá i model v jehož trénovací sadě nejsou zahrnuty obrázky s úsměvem (tabulka 5.3). Pokud je třeba lokalizovat obličeje s různým výrazem, je nutné zahrnout

Tabulka 5.3: IMM - Porovnání úspěšnosti modelů obsahující obrázky s úsměvem a bez nich.

počet osob / obrázků	celková ED (px)	průměrná ED (px)	hit rate (%)	počet módů
15 / 30	269	4,64	93,79	16
30 / 30 - bez úsměvu	338	5,82	85,17	17

proměnlivost ve výrazu také do tréninkové sady modelu.

Vliv natočení obličeje

V následujícím experimentu budou vytvořeny dva modely. Tréninková sada modelu **A** obsahuje 60 obrázků od 20 osob. Od každé osoby byl zařazen kromě snímku z čelního pohledu také snímek s natočením obličeje na levou a pravou stranu. Model **B** bude vytvořen z tréninkové sady 30 obrázků pořízených z čelního pohledu od 30 osob. Testovací sadu tvoří 10 obrázků od pěti osob s natočením hlavy na obě strany. Výsledek experimentu obsahuje tabulka 5.4. Absence obrázků s natočením hlavy v trénovací sadě modelu **A** způsobila rozdíl ve vzdálenosti mezi jednotlivými body modelů v průměru o 4,92 px.

Tabulka 5.4: IMM - Vlivu natočení obličeje na úspěšnost lokalizace.

Model	celková ED (px)	průměrná ED (px)	hit rate (%)	počet módů
model A	497	8,57	73,79	19
model B	783	13,49	47,76	17

Vliv změny osvětlení

V následujícím experimentu se zkoumá vliv osvětlení obličeje na úspěšnost jeho lokalizace. V trénovací sadě modelu **C** jsou obsaženy obrázky z čelního pohledu i obrázky se změnou osvětlení. Trénovací sada modelu **D** obsahuje pouze obrázky z čelního pohledu. Testovací sada obsahuje obrázky od 5 osob, na kterých je obličej osvětlen z různých stran. Jak vyplývá

Tabulka 5.5: IMM - Vlivu změny osvětlení na úspěšnost lokalizace.

Model	celková ED (px)	průměrná ED (px)	hit rate (%)	počet módů
model C	355	6,12	86,90	23
model D	356	6,15	87,93	19

z tabulky 5.5 rozdíl průměrné euklidovské vzdálenosti modelů je 0,03 px. To znamená, že změna osvětlení nemá na kvalitu modelu podstatný vliv.

5.2.2 XM2VTS databáze

Vliv barvy kůže

Následující experiment prověří vliv barvy kůže na úspěšnost modelu. Testovací sada bude obsahovat 3 obrázky od 3 osob s tmavou barvou kůže. První model **E** bude vytvořen z trénovací sady obsahující pouze osoby se světlou barvou kůže. Trénovací sada druhého modelu **F** obsahuje jak osoby se světlou tak tmavou barvou kůže. Třetí model **G** je sestaven na základě trénovací sady obsahující pouze osoby s tmavou barvou kůže. Výsledek experimentu

Tabulka 5.6: IMM - Vliv barvy kůže na kvalitu modelu.

Model	celková ED (px)	průměrná ED (px)	hit rate (%)	počet módů
model E	359	5,28	95,59	13
model F	410	6,03	86,77	12
model G	417	6,14	85,29	13

jsou uvedeny v tabulce 5.6. Nejúspěšnějším modelem se ukázal být model vytvořen pomocí obrázků se světlou barvou kůže. Tím je vyvrácen vliv barvy kůže na úspěšnost modelu.

Vliv doplňků

V následujícím experimentu bude ověřen vliv doplňků jako jsou vousy, knírky a brýle. Aby byl ověřen vliv doplňků na úspěšnost lokalizace obličeje budou vytvořeny dva modely a jejich úspěšnost bude ověřena na dvou testovacích sadách. Model **H** bude vytvořen pomocí trénovací sady obrázků obličejů obsahující doplňky. Trénovací sada modelu **I** nebude obsahovat obrázky s osobami nosící vousy, knírky či používající brýle. Testovací sada A obsahuje 5 obrázků obličejů bez vousů, knírků a brýlí. Testovací sada B je vytvořena z 5-ti obrázků obličejů obsahující doplňky. Z tabulky 5.7 vyplývá, že model jehož trénovací sada

Tabulka 5.7: IMM - Vliv doplňků obličeje na kvalitu modelu a úspěšnost lokalizace.

Model	celková ED (px)	průměrná ED (px)	hit rate (%)	počet módů
Testovací sada A - obrázky obličejů bez vousů, knírků a brýlí.				
model H	310	4,56	93,53	16
model I	368	5,41	92,94	16
Testovací sada B - obrázky obličejů s vousy, knírky nebo brýlemi.				
model H	430	6.32	89,71	16
model I	475	6.99	78,24	16

obsahuje obrázky s doplňky obličeje, vytváří kvalitnější model. Avšak vliv doplňků obličeje na úspěšnost lokalizace je patrný u obou modelů. Při testování obrázků obsahující vousy, knírky či brýle došlo ke zhoršení úspěšnosti lokalizace u obou modelů.

Vliv počtu bodů modelu

Pro ověření vlivu počtu bodů modelu na kvalitu lokalizace obličeje byl použit model **H** z předchozího pokusu. Testovací sada A také zůstala stejná. Programově byl snížen počet bodů stejného modelu zařazených do výpočtu o 10 na celkových 58. Byly odebrány dva body z okolí očí, dva body z oblasti obočí, dva body z kontury obličeje, dva body v okolí nosu a dva body z oblasti úst. Výsledky lokalizace obličeje pomocí modelů tvořených 58 a 68 body jsou uvedeny v tabulce 5.8 Se snížením počtu bodů modelu došlo k zhoršení

Tabulka 5.8: IMM - experiment vlivu počtu bodů modelu na kvalitu lokalizace.

model	celková ED (px)	průměrná ED (px)	hit rate (%)	počet módů
model - 68 bodů	310	4,56	93,53	16
model - 58 bodů	339	5,85	86,55	16

úspěšnosti lokalizace modelu v průměru o 1,29 px na jeden bod modelu.

5.2.3 Vyhodnocení

Na základě předchozích experimentů lze říci, že pro získání kvalitního modelu je potřeba vytvořit trénovací sadu obsahující:

- takové obrázky obličejů, které zaručují co největší variabilitu obličeje
- změny ve výrazu obličeje (úsměv, překvapení, zamračený výraz, ..)
- natočení hlavy (pokud je předmětem testování)
- obrázky obličejů obsahující různé doplňky (vousy, knírky, brýle)
- dostatečný počet bodů šablony

Naopak vliv na kvalitu modelu nemá:

- změna osvětlení obličeje
- barva kůže

Změna osvětlení nebo rozdíl v barvě kůže nevytváří výrazný rozdíl v kvalitě modelu. To je způsobeno principem algoritmu ASM, který místo absolutních hodnot intenzity, snímá rozdíl mezi hodnotami intenzity. Zatím co absence snímků natočení hlavy v tréninkové sadě způsobí, že model není schopen tuto deformaci napodobit. Vhodná trénovací sada obsahuje obrázky, které společně pokrývají rozdíly mezi obrázky, nad kterými bude algoritmus testován.

Na úspěšnost lokalizace obličeje pomocí deformačních modelů má vliv zejména přítomnost či absence doplňků v obličeji. Největším problémem jsou vousy, které znesnadňují lokalizaci úst či kontury obličeje. Některé brýle znemožňují přesnou lokalizaci očí či obočí.

5.3 Experiment 2 - rozpoznání obličeje

Metoda LDA byla podrobně popsána v části 2.4.2. V následujícím experimentu bude provedeno rozpoznání šablon obličejů právě pomocí metody LDA. Šablony obličejů budou vytvořeny metodou ASM. Obrázky obličejů poskytne databáze M2VTS, která je pro rozpoznání obličejů vhodná tím, že obsahuje 8 snímků od jedné osoby ze čtyř časově oddělených sezení. Celkově obsahuje 295 osob. Samotné rozpoznání pak probíhá v několika krocích:

1. vytvoření modelu
2. získání šablon
3. rozpoznání obličeje metodou LDA

5.3.1 Vytvoření modelu

Při tvorbě modelu byly využity znalosti z předchozího experimentu. Na základě těchto znalostí byl sestaven model z 66 obrázků od 11 osob.

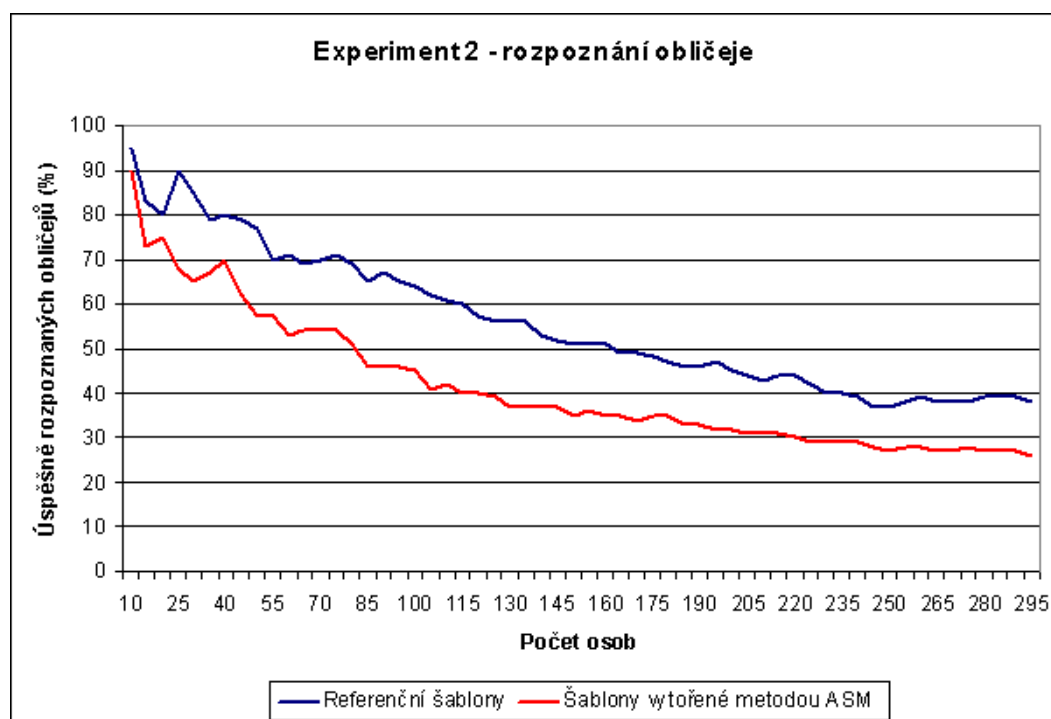
5.3.2 Získání šablon

Aby bylo možné provést rozpoznání šablon obličeje pomocí LDA, je potřeba získat šablony všech obličejů v databázi. Tyto šablony vzniknou lokalizací obličeje pomocí metody ASM. Tato část testu je časově velmi náročná. Vzhledem k tomu, že cílem testu je vyhodnotit rozpoznání obličejů od 295 osob a od každé osoby bylo pořízeno 8 snímků, musí algoritmus ASM provést 2360 lokalizací obličeje a výsledné šablony uložit pro zpracování metodou LDA.

5.3.3 Rozpoznání obličeje metodou LDA

Pro rozpoznání obličejů, jsou výsledné šablony rozděleny tak, že 6 šablon od každé osoby je určeno pro trénování a 2 šablony jsou určeny pro testování metody LDA. Experiment začíná od rozpoznání 10 osob a postupně je přidáváno po 5 dalších osobách až do celkového počtu 295 osob.

Pro srovnání byl experiment proveden i s referenčními šablonami obličejů. Tyto šablony představují ideální lokalizaci, ke které se metoda ASM snaží co nejvíce přiblížit.



Obrázek 5.7: Rozpoznání obličeje pomocí metody ASM.

Výsledek rozpoznání obličejů je uveden v grafu na obrázku 5.7. Experiment začíná od 10 tříd (10 osob s 20 obrázky) a v každém kroku je do rozpoznání přidáno dalších 5 tříd. Tabulky se všemi hodnotami jsou uvedeny v příloze D.

5.3.4 Vyhodnocení

Metoda ASM neposkytuje dostatečné množství informací o rozdílech v obličejích jednotlivých osob a není tak pro rozpoznání obličejů vhodná. V případě použití ideálních dat dokáže metoda založená na rozpoznávání jednotlivých šablon obličejů při 80 snímcích určit pouze 80% z nich přesně. V případě použití šablon vytvořených metodou ASM je při 80 snímcích určeno přesně pouze 70% obličejů. Rozdíl 10% vznikl pravděpodobně nepřesnou lokalizací obličeje. S narůstajícím počtem snímků se rozpoznání obličeje na základě této metody stává nepoužitelným.

Kapitola 6

Závěr

Tato diplomová práce je zaměřena na rozpoznání obličeje za pomoci metod deformačních modelů. V teoretické části byla rozebrána metoda *Active Shape Models* (ASM) a metoda *Active Appearance Models* (AAM). Metoda ASM byla pak také implementována a testována v demonstrační aplikaci. Po zvážení výsledků testování metody ASM lze konstatovat následující dva závěry.

Metodu ASM lze využít pro lokalizaci obličeje, pokud jsou splněny požadavky na data určená pro trénování modelu. Především je zapotřebí poskytnout dostatečnou skupinu obrázků s daty obsahující souřadnice bodů šablony.

K rozpoznání obličejů však metoda ASM příliš vhodná není a to už ze své podstaty. Tím, že je šablona tvořena omezeným počtem bodů (např. 68), není toto množství informace o jednotlivých obličejích k jejich přesnému rozpoznání dostatečné. To se projevuje při testech s větším množstvím osob.

6.1 Možná rozšíření

Pro vylepšení úspěšnosti lokalizace obličeje metodou ASM lze navrhnout zlepšení určování nových bodů modelu. U současného modelu se využívá profilu o šířce jednoho pixelu. Pokud by byl profil zvětšen na šířku například tří pixelů lze předpokládat přesnější určení nového bodu šablony a tím i úspěšnější lokalizaci celého obličeje.

Nejschůdnějším řešením, jak vylepšit lokalizaci i rozpoznání obličeje, je zvýšení počtu bodů šablony. V případě dodržení vstupního formátu dat není zapotřebí měnit implementovaný program, který umí pracovat s šablonami tvořenými různým počtem bodů. Nevýhodou a značnou komplikací je však časová náročnost anotace jednotlivých snímků.

Aby bylo možné využít metod deformačních modelů pro rozpoznání obličeje je zapotřebí rozšířit metodu ASM o informaci o celé textuře obličeje. Na tomto principu funguje metoda AAM. V případě, že místo omezeného počtu bodů šablony bude do rozpoznání zapojena celá textura obličeje, lze předpokládat, že bude dosaženo výrazně uspokojivějších výsledků, než u samotné metody ASM.

6.2 Využití metody ASM

Metodu ASM lze obecně využít k lokalizaci objektů. Využívá se však nejen k již mnohokrát zmíněné lokalizaci obličeje, ale testuje se její nasazení také například v lékařství [6]. Kde vstupní data tvoří snímky z magnetické rezonance a metoda slouží k lokalizaci různých orgánů (srdce, prostata).

Kromě samotné lokalizace se nabízí využít metody deformačních modelů pro detekci změny výrazu (například úsměvu, či pohybu rtů). Metoda by také mohla být úspěšná pro detekci natočení hlavy.

Literatura

- [1] AI ACCESS: Mahalanobis distance [online]. Dostupný z WWW:
http://www.aiaccess.net/English/Glossaries/GlosMod/e_gm_mahalanobis.htm,
[cit. 2009-04-30].
- [2] Cootes T. F., Baldock R.: Model-based methods in analysis of Biomedical images.
Image Processing and Analysis: A Practical approach, 1999: s. 223–248.
- [3] Cootes T. F., Taylor C. J.: *Statistical models of appearance for medical image analysis and computer vision*. University of Manchester, 2001, p. 14.
- [4] Cootes T. F., Taylor C. J.: *Statistical Models of Appearance for Computer Vision 1*. University of Manchester, 2004, 125 p.
- [5] Cootes T. F., Taylor C. J., Cooper D. H., Graham J.: Active Shape Models-Their Training and Application. *Computer Vision and Image Understanding*, ročník 61, č. 1, 1995: s. 38–59.
- [6] Cootes T. F., Taylor C. J., Hill A., Haslam J.: The use of active shape models for locating structures in medical images. *Image and Vision Computing*, ročník 12, č. 6, 1994: s. 355–366.
- [7] Delac K., Grgic M.(editors): *Face Recognition*. I-Tech: Vienna, Austria, 2007, p. 558. ISBN 978-3-902613-03-5.
- [8] Department of Statistics, Tung Hai University: JavaStatSoft [online]. Dostupný z WWW: <http://www2.thu.edu.tw/~wenwei/>, [cit. 2009-05-14].
- [9] face-rec.org: Face recognition - Databases [online]. Dostupný z WWW:
<http://www.face-rec.org/databases/>, [cit. 2009-05-10].
- [10] Griaule Biometrics: Understanding Biometrics [online]. Dostupný z WWW:
<http://www.griaulebiometrics.com/page/book/understanding-biometrics>,
[cit. 2009-04-25].
- [11] Herout, P.: *Java - bohatství knihoven*. České Budějovice: Kopp, 2006, 244 s. ISBN 978-80-7232-288-5.

- [12] Herout, P.: *Učebnice jazyka Java*. České Budějovice: Kopp, 2006, 320 s. ISBN 978-80-7232-323-4.
- [13] Herout, P.: *JAVA - grafické uživatelské prostředí a čeština*. České Budějovice: Kopp, 2007, 352 s. ISBN 978-80-7232-328-9.
- [14] Hinner J.: Detekce a rozpoznávání obličejů osob a jejich identifikační význam. 2003 [online]. Dostupný z WWW: http://web.mvcr.cz/archiv2008/casopisy/kriminalistika/2003/03_01/hinner.html, [cit. 2009-12-20].
- [15] Kiszka Bogdan: *1001 tipů a triků pro jazyk Java*. Computer Press, 2009, p. 542.
- [16] Messer, K.; Matas, J.; Kittler, J.; aj.: XM2VTSDB: The Extended M2VTS Database. In *Second International Conference on Audio and Video-based Biometric Person Authentication*, March 1999.
- [17] Muller K., Mika S., Ratsch G., Tsuda K. and Scholkopf B.: An Introduction to Kernel-Based Learning Algorithms. *IEEE Transactions on Neural Networks*, ročník 12, č. 2, 2001: s. 181–201.
- [18] Phillips P. J., Grother P, Micheals R. J., Blackburn D. M., Tabassi E., Bone M.: Face recognition vendor test 2002: Overview and Summary [online]. Dostupný z WWW: http://www.frvt.org/DLs/FRVT_2002_Overview_and_Summary.pdf, [cit. 2009-04-25].
- [19] Rak Roman, Matyáš Václav, Říha Zdeněk: *Biometrie a identita člověka - ve forenzních a komerčních aplikacích*. Grada, 2008, p. 664. ISBN-13: 978-80-247-2365-5.
- [20] Stegmann M. B.: *Generative Interpretation of Medical Images*. Technical University of Denmark, 2004, p. 248.
- [21] Stegmann M. B., Ersbøll B. K., Larsen R.: FAME – A Flexible Appearance Modelling Environment. *IEEE Trans. on Medical Imaging*, ročník 22, č. 10, 2003: s. 1319–1331.
- [22] Stegmann M. B., Gomez D. D.: *A Brief Introduction to Statistical Shape Analysis*. Technical University of Denmark, 2004, p. 15.
- [23] The MathWorks and NIST: JAMA : A Java Matrix Package [online]. Dostupný z WWW: <http://math.nist.gov/javanumerics/jama/>, [cit. 2009-05-14].
- [24] Turk M., Pentland A.: Eigenfaces for Recognition. *Journal of Cognitive Neuroscience*, ročník 3, č. 1, 1991: s. 71–86.
- [25] Wikipedia: Principal component analysis [online]. Dostupný z WWW: http://en.wikipedia.org/wiki/Principal_component_analysis, [cit. 2009-04-28].

- [26] Yang Ming-Hsuan, Kriegman David J., Ahuja Narendra: Detecting Faces in Images: A Survey. *IEEE Trans. on pattern analysis and machine intelligence*, ročník 24, č. 1, 2002: s. 34–58.
- [27] Zhao W.: Face Recognition: A Literature Survey. *IACM Computing Surveys*, ročník 35, č. 4, 2003: s. 399–458.

Dodatek A

Obsah DVD

Význačné soubory a adresáře:

\src - zdrojové kódy programu

\dist - distribuce projektu ve formě JAR archivu včetně potřebných externích knihoven

\text - text této diplomové práce ve formátu pdf

\databases - databáze IMM a M2VTS

\experiment1 - data pro experiment 1

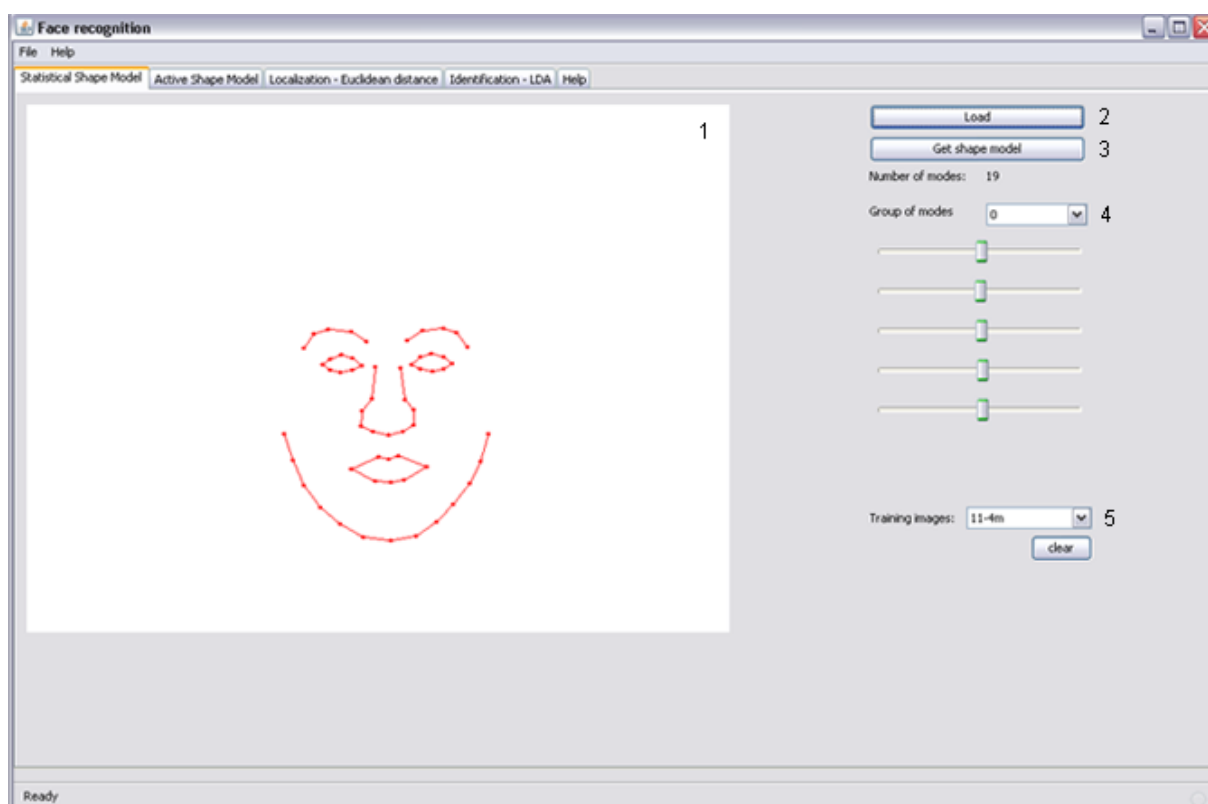
\experiment2 - data pro experiment 2

\models - několik předem vytvořených modelů

Dodatek B

Příručka k programu

B.1 Panel Statistical Shape Model



Obrázek B.1: Popis panelu Statistical Shape Model.

B.1.1 Popis panelu

- (1) Plocha pro vykreslení modelu.

- (2) Tlačítko Load.
- (2) Tlačítko Get shape model.
- (4) ComboBox pro výběr skupiny módů modelu.
- (5) ComboBox pro prohlížení tréninkových obrázků a tlačítko clear.

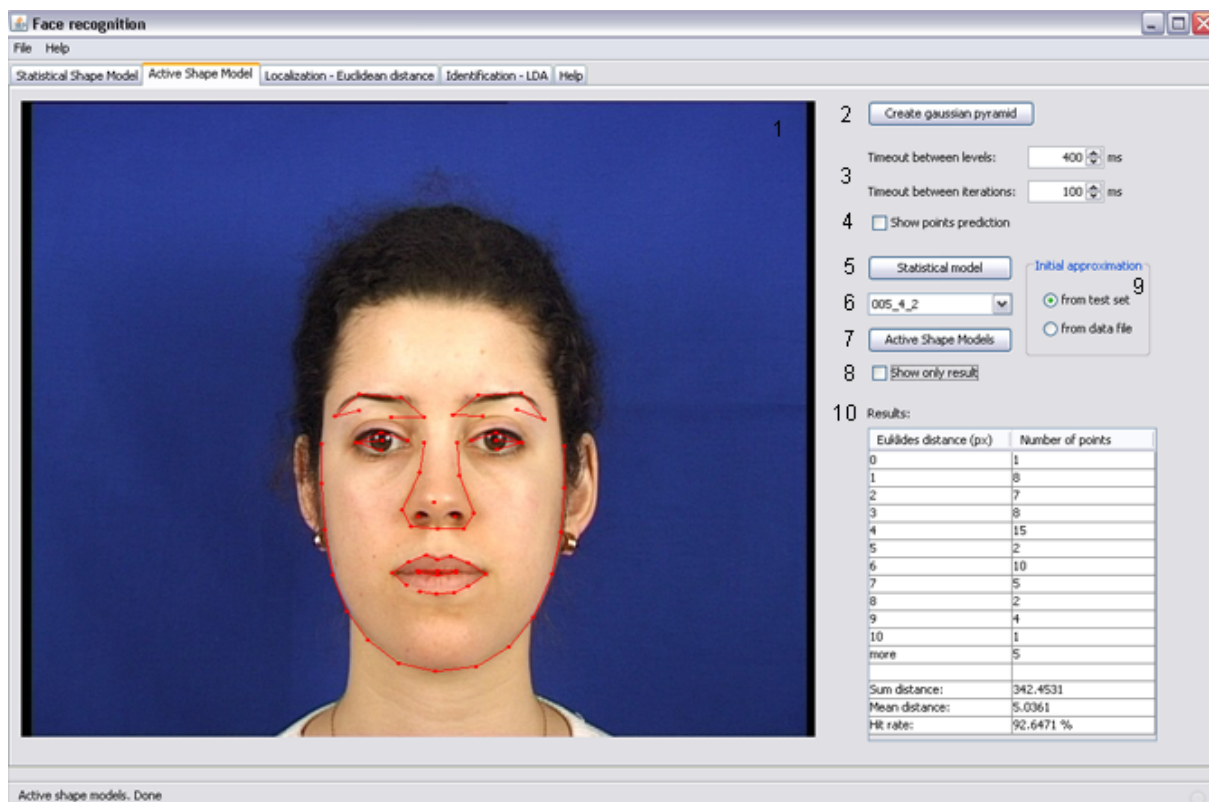
B.1.2 Postup pro vytvoření modelu

1. Po kliknutí na tlačítko (1) se pomocí nabídky vybere složka s modelem (např. E:\models\model01).
2. Stisknutím na tlačítko (2) se vytvoří statistický model.
3. Pomocí ComboBoxu (4) lze vybírat skupiny módů a pomocí sliderů pod ním lze model deformovat. V první skupině módů lze tvarovat módy 1-5 v druhé skupině 6-10 atd. až do počtu módů modelu.
4. V ComboBoxu (5) lze vybírat tréninkové obrázky, které se zobrazí na ploše (1) pod modelem. Tlačítko clear slouží pro vyčištění plochy (1).

B.2 Panel Active Shape Model

B.2.1 Popis panelu

- (1) Plocha pro demonstraci lokalizace obličeje metodou ASM.
- (2) Tlačítko Create gaussian pyramid.
- (3) Spinner pro nastavení časové prodlevy mezi přechodem z jedné úrovně rozlišení na druhou a mezi jednotlivými iteracemi algoritmu.
- (4) Tlačítko k aktivaci a deaktivaci zobrazení bodů, které algoritmus vybral jako nové body pro model.
- (5) Tlačítko Statistical model pro vytvoření statistických modelů pro jednotlivé body šablony.
- (6) ComboBox umožňující výběr obrázku k testování.
- (7) Tlačítko Active Shape Models, které zahájí lokalizaci pomocí algoritmu ASM.
- (8) Tlačítko, po jehož zatržení bude zobrazen rovnou výsledek lokalizace obličeje.
- (9) Nastavení počáteční pozice modelu. Buď jako výpočet z trénovací sady nebo na základě přiloženého datového souboru.
- (10) Tabulka s výsledky lokalizace.



Obrázek B.2: Popis panelu Active Shape Model.

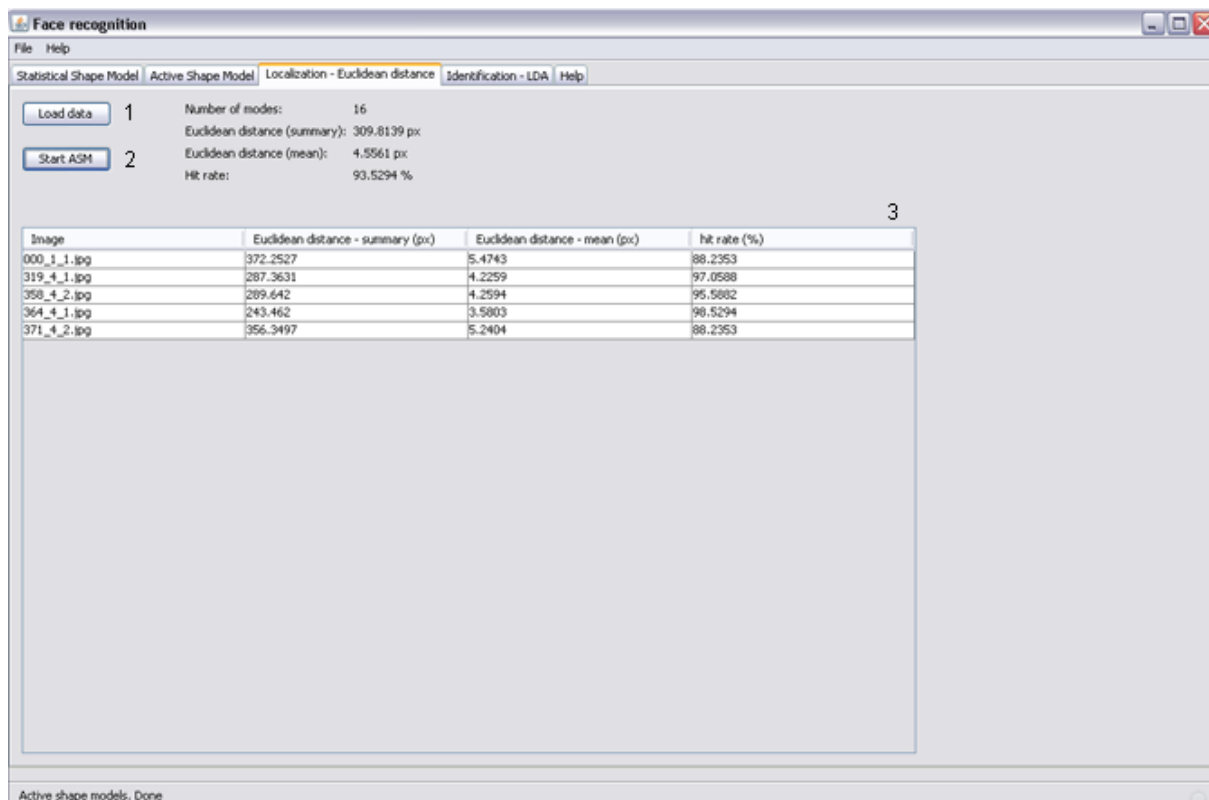
B.2.2 Postup pro zahájení lokalizace pomocí ASM

1. Pokud je třeba vytvořit více úrovní rozlišení, stiskne se tlačítko (2). U všech modelů na přiloženém DVD jsou již úrovně rozlišení pro každý model vytvořeny.
2. Stisknutím tlačítka (5) se vytvoří statistický model.
3. Pomocí ComboBoxu (6) se vybere obrázek k testování.
4. Stiskem tlačítka (7) se zahájí lokalizace.

B.3 Panel Localization - Euclidean distance

B.3.1 Popis panelu

- (1) Tlačítko pro načtení modelu.
- (2) Tlačítko pro zahájení lokalizace.
- (3) Tabulka s výsledky lokalizace.



Obrázek B.3: Popis panelu Localization - Euclidean distance.

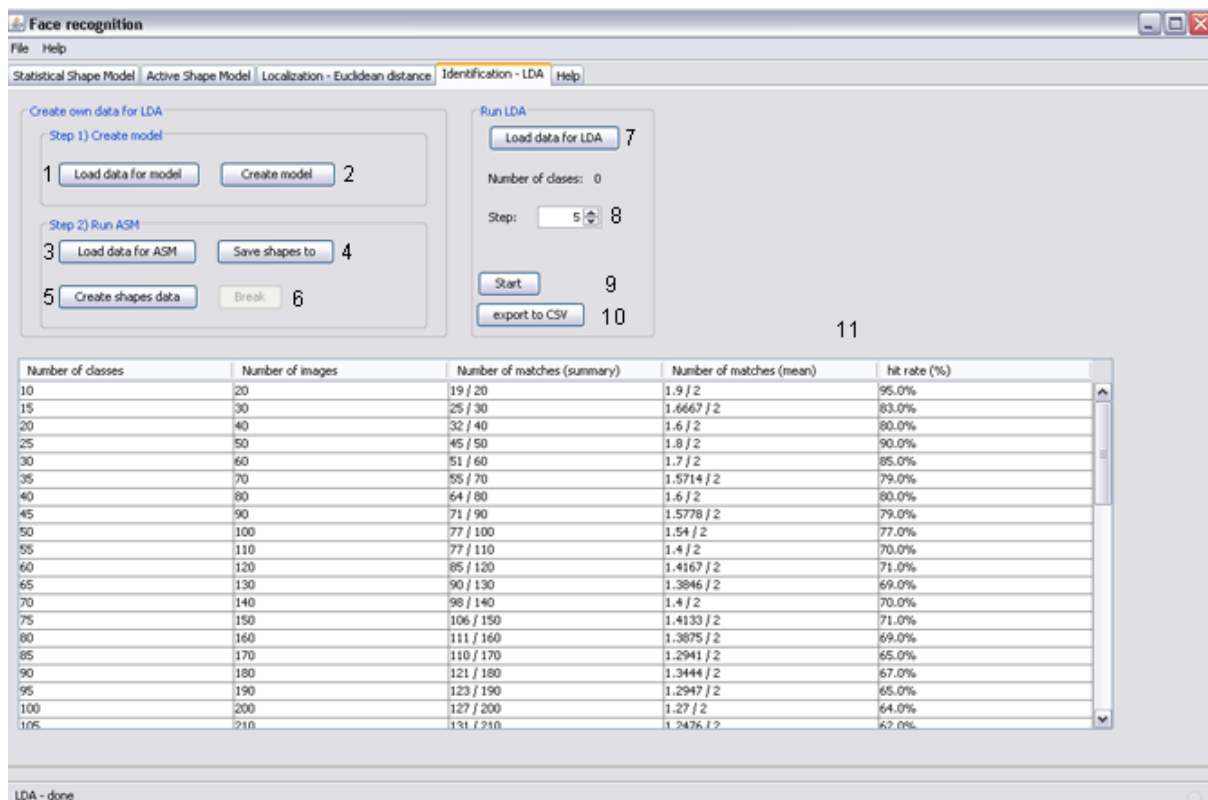
B.3.2 Postup pro lokalizaci všech obličejů v testovací sadě

1. Tlačítkem (1) je vybrána složka se sadou pro lokalizaci všech obličejů v testovací sadě (např. E:\experiment1\07_vliv_poctu_bodu\model_68).
2. Tlačítkem (2) je zahájena lokalizace.
3. V tabulce (3) jsou vypsány výsledky lokalizace pro jednotlivé obrázky. Souhrn výsledků pro daný model je uveden ve výpisu nad tabulkou.

B.4 Panel Identification - LDA

B.4.1 Popis panelu

- (1) Tlačítko Load data for model.
- (2) Tlačítko Create Model.
- (3) Tlačítko Load data for ASM.
- (4) Tlačítko Save shapes to.



Obrázek B.4: Popis panelu Identification - LDA.

- (5) Tlačítko Create shapes data.
- (6) Tlačítko Break.
- (7) Tlačítko Load data for LDA.
- (8) Spinner s nastavením kroku počtu tříd.
- (9) Tlačítko Start.
- (10) Tlačítko export to CSV.
- (11) Tabulka s výsledky rozpoznání pomocí LDA.

B.4.2 Postup pro rozpoznání obličejů pomocí LDA

a) příprava dat

1. Natrénování modelu - pomocí tlačítka (1) se vybere složka s modelem (např. E:\experiment2\model_LDA) a tlačítkem (2) se vytvoří model.

2. Pomocí tlačítka (3) se vybere složka s daty pro LDA (E:\experiment2\M2VTS_LDA) a pomocí tlačítka (4) se vybere složka pro uložení lokalizovaných šablon. Tlačítkem (5) se spustí lokalizace pro všechny osoby (sada M2VTS_LDA obsahuje 2360 osob).
3. Protože tato část je časově náročná lze ji tlačítkem (6) přerušit a využít šablon uložených na přiloženém DVD.

b) zahájení rozpoznávání

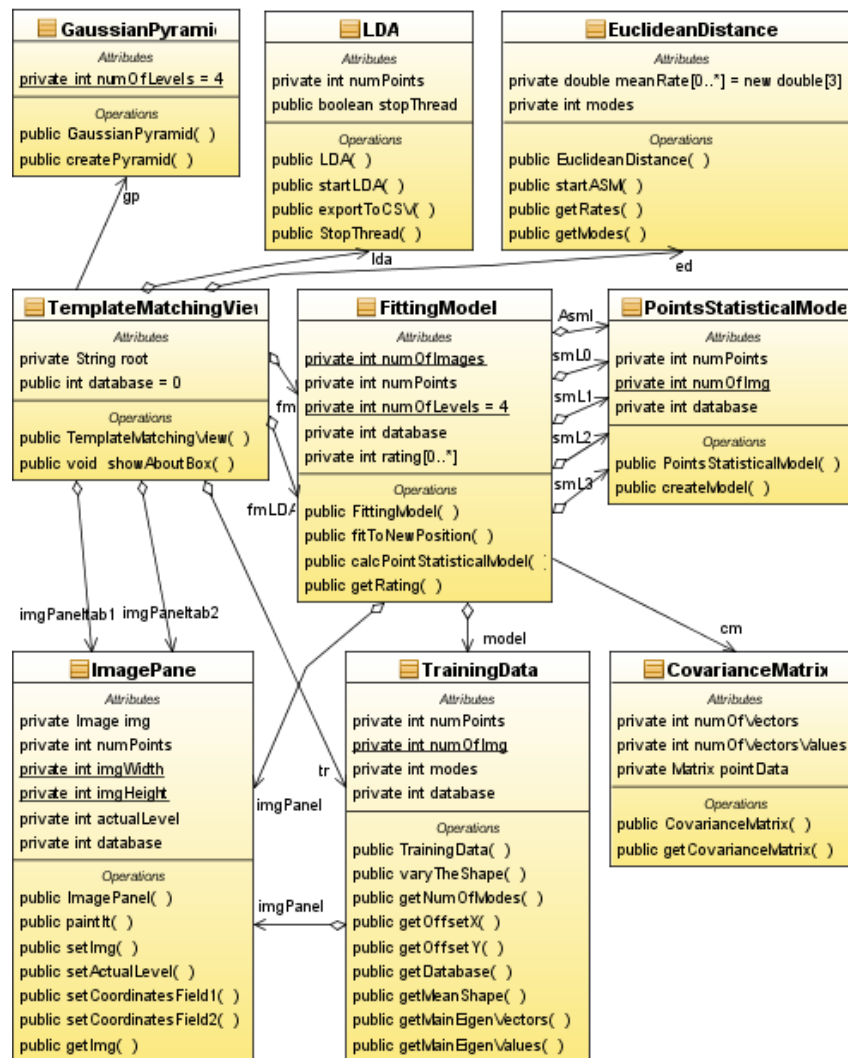
1. Tlačítkem (7) lze načíst šablony pro rozpoznání pomocí LDA (E:\experiment2\M2VTS_MANUALLY).
2. Tlačítkem (9) se zahájí rozpoznávání.
3. V tabuce (11) jsou vypsány hodnoty pro daný počet tříd a tlačítkem (10) se provede export dat z tabulky do souboru.

B.5 Panel Help

Nápověda k programu.

Dodatek C

Schéma aplikace



Obrázek C.1: Schéma aplikace modelované pomocí UML diagramu tříd.

Dodatek D

Experiment 2

Tabulka D.1: Experiment 2 - testování rozpoznání obličejů, část 1.

počet tříd	referenční šablony		šablony vytvořené metodou ASM	
	rozpoznaných obličejů		rozpoznaných obličejů	
	počet	procent (%)	počet	procent (%)
10	19 / 20	95	18 / 20	90
15	25 / 30	83	22 / 30	73
20	32 / 40	80	30 / 40	75
25	45 / 50	90	34 / 50	68
30	51 / 60	85	39 / 60	65
35	55 / 70	79	47 / 70	67
40	64 / 80	80	56 / 80	70
45	71 / 90	79	56 / 90	62
50	77 / 100	77	57 / 100	57
55	77 / 110	70	63 / 110	57
60	85 / 120	71	64 / 120	53
65	90 / 130	69	70 / 130	54
70	98 / 140	70	76 / 140	54
75	106 / 150	71	81 / 150	54
80	111 / 160	69	82 / 160	51
85	110 / 170	65	79 / 170	46
90	121 / 180	67	82 / 180	46
95	123 / 190	65	88 / 190	46
100	127 / 200	64	89 / 200	45
105	131 / 210	62	87 / 210	41
110	134 / 220	61	92 / 220	42
115	139 / 230	60	93 / 230	40
120	138 / 240	57	95 / 240	40

Tabulka D.2: Experiment 2 - testování rozpoznání obličejů, část 2.

počet tříd	referenční šablony		šablony vytvořené metodou ASM	
	rozpoznaných obličejů		rozpoznaných obličejů	
	počet	procent (%)	počet	procent (%)
125	141 / 250	56	98 / 250	39
130	145 / 260	56	95 / 260	37
135	150 / 270	56	101 / 270	37
140	149 / 280	53	103 / 280	37
145	152 / 290	52	106 / 290	37
150	152 / 300	51	106 / 300	35
155	157 / 310	51	113 / 310	36
160	162 / 320	51	112 / 320	35
165	161 / 330	49	114 / 330	35
170	166 / 340	49	114 / 340	34
175	167 / 350	48	124 / 350	35
180	170 / 360	47	125 / 360	35
185	171 / 370	46	123 / 370	33
190	174 / 380	46	127 / 380	33
195	183 / 390	47	125 / 390	32
200	181 / 400	45	129 / 400	32
205	179 / 410	44	126 / 410	31
210	182 / 420	43	130 / 420	31
215	188 / 430	44	134 / 430	31
220	194 / 440	44	132 / 440	30
225	191 / 450	42	132 / 450	29
230	186 / 460	40	133 / 460	29
235	189 / 470	40	136 / 470	29
240	189 / 480	39	139 / 480	29
245	183 / 490	37	138 / 490	28
250	185 / 500	37	136 / 500	27
255	193 / 510	38	144 / 510	28
260	202 / 520	39	145 / 520	28
265	202 / 530	38	143 / 530	27
270	207 / 540	38	147 / 540	27
275	209 / 550	38	152 / 550	28
280	216 / 560	39	152 / 560	27
285	221 / 570	39	152 / 570	27
290	225 / 580	39	154 / 580	27
295	225 / 590	38	152 / 590	26