

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ
ÚSTAV INFORMAČNÍCH SYSTÉMŮ

FACULTY OF INFORMATION TECHNOLOGY
DEPARTMENT OF INFORMATION SYSTEMS

NALEZENÍ A ROZPOZNÁNÍ DOMINANTNÍCH RYSŮ OBLIČEJE

DIPLOMOVÁ PRÁCE

MASTER'S THESIS

AUTOR PRÁCE

AUTHOR

Bc. HYNEK ŠVÁBEK

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

NALEZENÍ A ROZPOZNÁNÍ DOMINANTNÍCH RYSŮ OBLIČEJE

DETECTION AND RECOGNITION OF DOMINANT FACE FEATURES

DIPLOMOVÁ PRÁCE

MASTER'S THESIS

AUTOR PRÁCE

AUTHOR

Bc. HYNEK ŠVÁBEK

VEDOUCÍ PRÁCE

SUPERVISOR

Ing. PETR CHMELAŘ

BRNO 2010

Abstrakt

Tato práce pojednává o stále více se rozvíjející oblasti biometrických systémů, kterou je rozpoznání obličejů. Dokument se zabývá možnostmi lokalizace tváře na obrázcích a jejich normalizací, která je nutná z důvodu vnějších vlivů i vlivem různých snímacích technik. Jsou zde popsány různé techniky lokalizace dominantních rysů obličeje, kterými jsou např. oči, ústa, nos. V neposlední řadě práce popisuje různé přístupy k rozpoznání obličejů. Dále je zpracován návrh s implementací aplikace pro rozpoznání dominantních rysů obličeje demonstrující zvolené metody pro nalezení dominantních rysů (Houghova transformace pro detekci kružnic, lokalizace úst ze znalosti umístění očí) a identifikaci obličeje (lineární diskriminační analýza, jádrová diskriminační analýza). Poslední část diplomové práce obsahuje shrnutí dosažených výsledků a závěr.

Abstract

This thesis deals with the increasingly developing field of biometric systems which is the identification of faces. The thesis deals with the possibilities of face localization in pictures and their normalization, which is necessary due to external influences and the influence of different scanning techniques. It describes various techniques of localization of dominant features of the face such as eyes, mouth or nose. Not least, it describes different approaches to the identification of faces. Furthermore it deals with an implementation of the Dominant Face Features Recognition application, which demonstrates chosen methods for localization of the dominant features (Hough Transform for Circles, localization of mouth using the location of the eyes) and for identification of a face (Linear Discriminant Analysis, Kernel Discriminant Analysis). The last part of the thesis contains a summary of achieved results and a discussion.

Klíčová slova

Dominantní rysy obličeje, rozpoznání obličeje, analýza hlavních komponent (PCA), lineární diskriminační analýza (LDA), jádrová diskriminační analýza (KDA).

Keywords

Dominant face features, face recognition, Principal Component Analysis, Linear Discriminant Analysis, Kernel Discriminant Analysis.

Citace

Hynek Švábek: Nalezení a rozpoznání dominantních rysů obličeje, diplomová práce, Brno, FIT VUT v Brně, 2010

Nalezení a rozpoznání dominantních rysů 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.

.....

Hynek Švábek
24. května 2010

© Hynek Švábek, 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	3
2	Využití v biometrii	5
2.1	Identifikace obličeje	5
2.2	Verifikace obličeje	5
2.3	Hledání osob	5
2.4	Spolehlivost biometrických systému	6
2.4.1	Míra chybného přijetí	6
2.4.2	Míra chybného odmítnutí	6
2.4.3	Míra chybné shody	6
2.4.4	Míra chybné neshody	7
2.4.5	Míra vyrovnání chyb	7
2.4.6	ROC křivky	8
3	Lokalizace obličeje	9
3.1	Strukturální metody	9
3.2	Lokalizace obličeje pomocí barvy kůže	11
3.3	Detekce obrysu tváře	12
3.4	Porovnávání šablon	12
3.4.1	Předdefinované šablony	12
3.4.2	Tvarovatelné šablony	13
4	Nalezení dominantních rysů obličeje	14
4.1	Detekce očí	14
4.1.1	Snímání osvětleného bodu oka	14
4.1.2	Robustní algoritmus pro detekci očí ve tváři ve stupních šedi bez brýlí	15
4.1.3	Houghova transformace pro detekci kružnic	17
4.2	Detekce úst	18
4.2.1	Ze znalosti umístění očí	18
4.2.2	Na základě barevných modifikací obrazu	18
4.3	Detekce nosu	20
4.4	Detekce obočí podle umístění očí	20
5	Normalizace obrazu obličeje	21
5.1	Roztažení kontrastu	22
5.2	Vyrovnání histogramu	23

6	Rozpoznání obličeje	26
6.1	Lineární analýza	26
6.1.1	Analýza hlavních komponent	27
6.1.2	Analýza nezávislých komponent	28
6.1.3	Lineární diskriminační analýza	29
6.1.4	Komponentní lineární diskriminační analýza	31
6.2	Jádrové metody	33
6.2.1	Jádrová analýza hlavních komponent	35
6.2.2	Jádrová diskriminační analýza	36
6.3	Elastické shlukové grafy	36
6.4	Aktivní vzhledový model	38
6.5	Shrnutí	38
7	Implementace a validace	39
7.1	Reprezentace dat	39
7.2	Načtení informací o obrázcích v databázi	39
7.3	Rozpoznání dominantních rysů	39
7.4	Identifikace	40
7.5	Umístění obličeje	41
7.6	Testy	43
7.6.1	IMM Face Database	43
7.6.2	Databáze XM2VTSDB	46
7.7	Shrnutí	49
8	Rozšíření a možnosti budoucího vývoje	51
8.1	Rozpoznání rysů a identifikace více osob na jednom obrázku	51
8.2	Vlastní detektor obličeje	51
8.3	Automatizovaná volba vhodných parametrů pro KDA – cross-validace	52
8.4	Možnost ručního označení dominantních rysů	52
8.5	Zvýšení úspěšnosti detekce očí mnohonásobnou detekcí kružnic o různém průměru	53
8.6	Uložení jednou detekovaných dominantních rysů do souboru	53
8.7	Kombinace více metod	53
9	Závěr	54
A	Obsah DVD	63
B	Parametry testů	64
B.1	Identifikace nad daty databáze IMM	64
B.2	Identifikace nad daty databáze XM2VTSDB	64
C	Databáze obličejů, nápověda a diagram tříd aplikace	66
C.1	Databáze obličejů	66
C.2	Nápověda	66
C.3	Diagram tříd aplikace	66

Kapitola 1

Úvod

Rozpoznání obličeje je velice důležitou oblastí v počítačovém vidění i v biometrii. V posledních letech se stále více vědeckých pracovníků zabývá touto problematikou. V biometrii patří rozpoznání obličeje spolu s otisky prstů k nejkoumanějším oblastem, jelikož pořizovací náklady jsou díky rozšíření výpočetní techniky v posledních letech nízké. Tento způsob identifikace osob je navíc většinou dobře snášen samotnými uživateli. Díky akceptovatelnosti uživatelů existuje předpoklad rozšíření biometrických systémů pro rozpoznávání obličeje do komerční sféry, ať už pro systémy docházky či jiné přístupové systémy. Z hlediska bezpečnosti je možné využití pro hledání pohřešovaných a hledaných osob na místech, kde se shromažďuje mnoho lidí (sportovní utkání, demonstrace) nebo pro pouhé snímání venkovních prostor městským kamerovým systémem napojeným na databázi hledaných osob.

Cílem této práce je poskytnout ucelený přehled procesu identifikace osob, ať už se jedná o různé přístupy k lokalizaci obličeje a dominantních rysů či o samostatnou identifikaci obličeje. Výstupem je dále aplikace demonstrující zvolenou metodu lokalizace dominantních rysů a identifikace obličeje.

Následující kapitola 2 stručně shrnuje základní pojmy využití v biometrii, kterými jsou identifikace a verifikace.

Další kapitola 3 se zabývá metodami detekce lidské tváře v obraze. Existující metody využívají např. apriorní znalosti tvaru lidského obličeje či barvy kůže. V praxi se využívá kombinace více přístupů. Rychlejší, ale méně přesná metoda určí kandidátské oblasti, kde by se mohl nalézat obličej, a následně je aplikována pomalejší a přesnější metoda pouze na tuto oblast.

Kapitola 4 pojednává o detekci dominantních částí obličeje jako jsou oči, ústa, nos či obočí. Detekovat tyto dominantní rysy je možné buď v oblasti již lokalizovaného obličeje, nebo také v kandidátské oblasti obličeje, kdy může dojít k potvrzení, že tato oblast obličejem opravdu je.

Následující část textu v kapitole 5 pojednává o nutnosti obraz normalizovat a to z toho důvodu, že obraz tváře není vždy pořízen kvalitní technikou či za vhodného osvětlení.

Dále je problémem také situace, kdy je tvář na zpracovávaném obrázku v jiném než čelním pohledu. Normalizace se snaží všechny tyto problémy řešit.

Následující kapitola 6 popisuje různé přístupy k rozpoznání obličeje. Jsou zde popsány linerní metody, jako jsou např. analýza hlavních komponent, analýza nezávislých komponent, lineární diskriminační analýza, dále také jádrové metody (kernel methods), kterými jsou jádrová diskriminační analýza a jádrová analýza hlavních komponent. Je zde dále popsána metoda elastických shlukových grafů popisující topologickou strukturu lidské tváře a aktivní vzhledové modely.

Kapitola 7 se zabývá implementací a validací aplikace. Podkapitoly 7.3 a 7.4 popisují zvolené metody a přístupy k rozpoznání dominantních rysů a identifikaci obličeje. Jsou zde prezentovány výsledky testů a nachází se zde shrnutí výsledků dosažených při testování aplikace.

V předposlední kapitole 8 jsou zmíněna některá budoucí rozšíření a směr vývoje, jakým by se aplikace mohla ubírat.

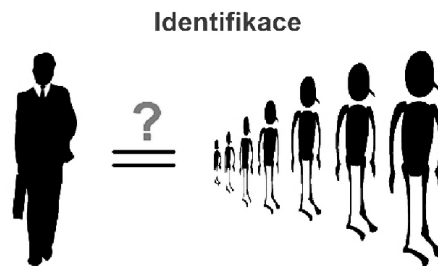
Závěrečná kapitola 9 zhodnocuje dosažené výsledky.

Kapitola 2

Využití v biometrii

2.1 Identifikace obličeje

Identifikace obličeje je proces zpracování obrázku neznámé tváře a zjištění informace, které osobě náleží, porovnáváním tohoto obrázku s databází známých tvář. Jedná se vlastně o porovnávání 1:N viz obrázek 2.1, kdy se hledá největší podobnost se vzorem uloženým v databázi. Před vlastní identifikací musí být nejříve hledaný obraz normalizován.



Obrázek 2.1: Identifikace. Převzato z [7].

2.2 Verifikace obličeje

Verifikace je proces potvrzení identity osoby na obrázku. Jedná se o porovnání pouze s jednou referenční šablonou v databázi, tedy o porovnávání 1:1 jak je zobrazeno na obrázku 2.2. Před vlastním procesem verifikace je opět nutné provést normalizaci.

2.3 Hledání osob

Pomocí hledání osob je možné hledat obraz tváře, který nemusí být uložen v databázi, se kterou jej porovnáváme. Obraz obličeje je postupně porovnáván s obličejí uloženými v databázi a výsledkem je seznam osob seřazený dle míry podobnosti. Tato metoda se dá



Obrázek 2.2: Verifikace. Převzato z [7].

využít s databází hledaných osob s využitím běžného kamerového systému města či na sportovním stadionu.

2.4 Spolehlivost biometrických systému

Spolehlivost biometrického systému udává kvalitu a efektivitu systému. Měří se pomocí pravděpodobnosti s jakou systém odmítne neoprávněnou osobu, přijme oprávněnou osobu, přijme neoprávněnou osobu či odmítne osobu oprávněnou.

2.4.1 Míra chybného přijetí

Míra chybného přijetí [7] (FAR – False Acceptance Rate) je pravděpodobnost, že biometrický systém klasifikuje dva různé biometrické vzory jako shodné. Pokud se tak stane, systém selže při odmítnutí potenciálního útočníka. FAR je možné vypočítat jako

$$FAR = \frac{\text{pocet shodných porovnaní rozdílných vzorů}}{\text{celkový počet porovnaní rozdílných vzorů}}. \quad (2.1)$$

2.4.2 Míra chybného odmítnutí

Míra chybného odmítnutí [7] (FRR – False Rejection Rate) určuje pravděpodobnost, že biometrický systém klasifikuje dva různé biometrické vzory od stejné osoby jako různé a tím selže při přijetí oprávněného uživatele. FRR je spočítáno pomocí rovnice

$$FRR = \frac{\text{pocet porovnaní vzorů osoby A vedoucích k neshodě}}{\text{celkový počet porovnaní vzorů osoby A}}. \quad (2.2)$$

2.4.3 Míra chybné shody

Míra chybné shody [7] (FMR – False Match Rate) udává podíl chybně akceptovaných osob. Oproti FAR neuvažuje v celkovém součtu pokusy, které byly neúspěšné ještě před samotným porovnáváním. FMR je možno spočítat pomocí

$$FNR(T) = \int_T^1 p_i(s|H_i) ds, \quad (2.3)$$

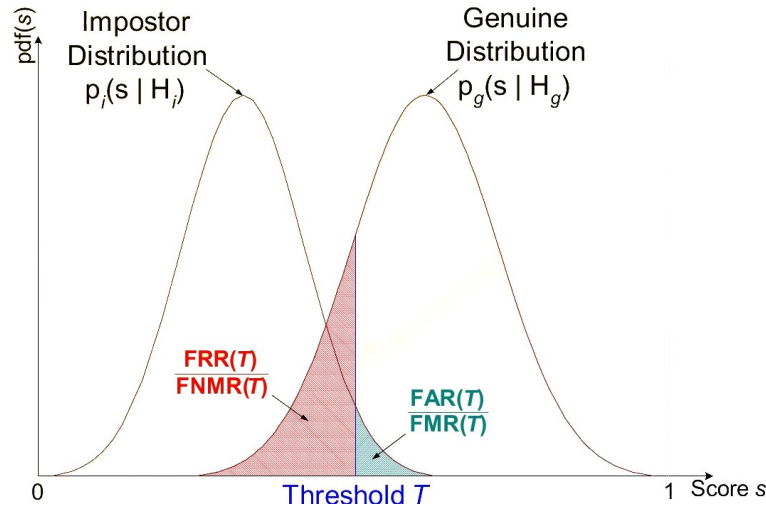
kde T je rozhodovací práh, H_i je výrok, že vzor a šablona pochází od různých osob, p_i je pravděpodobnostní hustota, že výrok v závorce je pravdivý, a s je skóre porovnání.

2.4.4 Míra chybné neshody

Míra chybné neshody [7] (FNMR – False Non-Match Rate) vyjadřuje podíl osob, které byly chybně odmítnuty biometrickým systémem. Narozdíl od FRR nejsou do celkových součtů započítány neúspěšné pokusy ještě před vlastním porováním. FNMR lze získat pomocí vztahu

$$FNMR(T) = \int_0^T p_g(S|H_g)ds, \quad (2.4)$$

kde T je rozhodovací práh, H_g je výrok, že vzor a šablona pochází od stejné osoby, p_g je pravděpodobnostní hustota, že výrok v závorce je pravdivý a s je skóre porovnání. Na

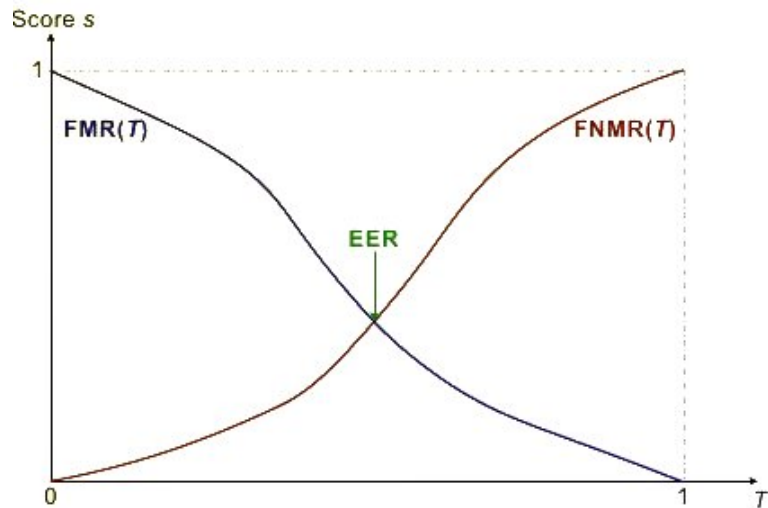


Obrázek 2.3: Vztah FMR/FNMR a FAR/FRR. Převzato z [7].

obrázku 2.3 křivka Impostor Distribution udává podíl neoprávněných osob a křivka Genuine Distribution udává podíl osob oprávněných pro vstup do systému. Podle rozhodovacího prahu T se podíl chybně přijatých/odmítnutých osob může měnit. Pokud práh T bude nastaven na hodnotu, kde se obě křivky protínají, bude podíl chybně přijatých/odmítnutých osob stejný a bude označen jako míra vyrovnaní chyb.

2.4.5 Míra vyrovnaní chyb

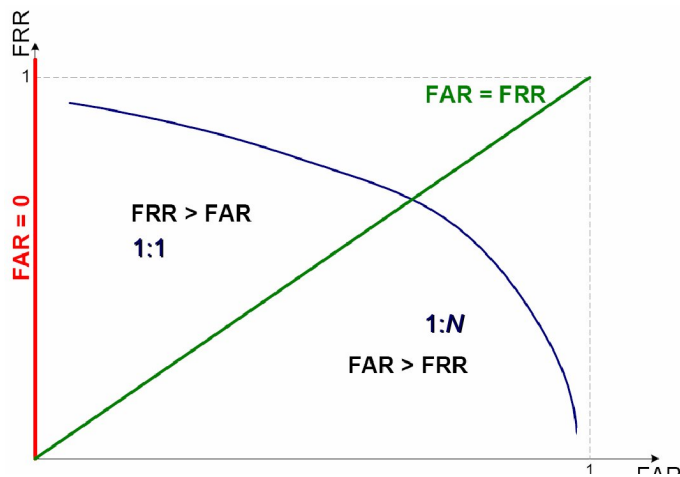
Míra vyrovnaní chyb [7] (EER) reprezentuje oblast, ve které se FMR(T) a FNMR(T) shodují a která je zobrazena na obrázku 2.4. Pokud nastavíme porovnávací práh T na hodnotu EER, bude se počet chybně akceptovaných osob rovnat počtu osob chybně přijatých. Díky tomuto je možné nastavit práh tak, aby hodnoty FMR a FNMR odpovídaly požadavkům na biometrický systém.



Obrázek 2.4: Míra vyrovnání chyb. Převzato z [7].

2.4.6 ROC křivky

ROC křivky [7] (Receiver Operating Curve) jsou v dnešní době standardem při popisu vlastností daného systému. V biometrii představují detekční schopnost funkce FMR k FNMR eventuelně FAR k FRR, tj. $FNMR=f(FMR)$. Příklady některých ROC křivek jsou zobrazeny na obrázku 2.5.



Obrázek 2.5: ROC křivky. Převzato z [7].

Kapitola 3

Lokalizace obličeje

Lokalizace obličeje je prvním a velmi důležitým krokem při rozpoznávání obličeje. Správné určení umístění kandidátských oblastí obličeje v obraze může ušetřit nutná rozhodování při kontrole toho, zda kandidátská oblast je či není obličejem.

3.1 Strukturální metody

Přístup k detekci lidské tváře založený na strukturálních metodách je vyvinut na základě znalostí vzhledu lidské tváře. Je nutné vytvořit jednoduchá pravidla popisující rysy obličeje a vazby mezi nimi. Obličej se obvykle vyskytuje v pohledu, kdy jsou viditelné obě oči, které jsou vzájemně symetrické, nos a ústa. Vazby mezi rysy mohou být reprezentovány jejich relativní vzdáleností a polohou. Rysy obličeje jsou ze vstupního obrazu extrahovány jako první a kandidátské oblasti obličejů jsou rozpoznány na základě kódovacích pravidel. Je ještě možné použít verifikační proces ke snížení špatných detekcí obličeje. Problémem tohoto přístupu je obtížné vytváření dobře definovaných pravidel na základě znalostí vzhledu lidského obličeje. Pokud budou pravidla vytvořena příliš podrobně, mohou být špatně detekovány některé tváře z důvodu, že nebudou vyhovovat všem pravidlům. Naopak, pokud budou pravidla příliš obecná může dojít ke špatné detekci obličeje. Dále je také obtížné použít tento přístup pro detekci obličejů pod různými úhly, protože je téměř nemožné popsat všechny možné případy natočení obličeje. Ale přesto při detekci tváří při čelním pohledu tento přístup funguje dobře [32].

Například metoda založená na hierarchických znalostech podoby lidské tváře popsána v [32] se skládá ze tří úrovní pravidel. Na nejvyšší úrovni jsou prohledávány všechny kandidátské oblasti obličejů. Kandidátské oblasti jsou prohledávány aplikováním okénka na vstupní obraz a aplikací sady pravidel na každém místě. Pravidla na nejvyšší úrovni popisují obecný vzhled tváře, pravidla na nižších vrstvách se spoléhají na detaily obličejových rysů.

Například hierarchie obrázků o různém rozlišení je vytvořena průměrováním a podvzrokováním 3.1. Příklady pravidel používaných k detekci kandidátských oblastí obličejů

v nejnižším rozlišení jsou například – střední část obličeje (tmavé části na obrázku 3.2 složené z oblastí s rovnoměrnou intenzitou), horní část obličeje (světlé části na obrázku 3.2 s rovnoměrným rozložením intenzity) a významný je také rozdíl průměrných hodnot intenzity mezi střední a horní částí obličeje.



Obrázek 3.1: a) $n=1$, původní obrázek, b) $n=4$, c) $n=8$, d) $n=16$. Původní a odpovídající obrázky v nízkém rozlišení. Každá čtvercová oblast se skládá z $n \times n$ bodů u kterých je hodnota intenzity každého bodu nahrazena průměrnou hodnotou intenzity bodů čtvercové oblasti. Převzato z [32].



Obrázek 3.2: Typický obličej používaný ve strukturálních metodách. Kódující pravidla jsou založena na znalosti vlastností obličejových oblastí (rozložení intenzity a jejich rozdíl). Převzato z [32].

Na obrázku s nejnižším rozlišením (úroveň 1) jsou hledány kandidátské oblasti obličejů a tyto oblasti jsou dále zpracovávány ve vyšším rozlišení. Na druhé úrovni je provedeno lokální vyrovnání histogramu na kandidátních oblastech obličeje získaných z první úrovně. Poté následuje detekce hran. Kandidátské oblasti, které zůstaly po aplikaci pravidel na druhé úrovni, jsou zkoumány na třetí úrovni další sadou pravidel odpovídající obličejovým rysům jako jsou oči a ústa.

Výhodou tohoto hierarchického přístupu je, že postupné zaměření od hrubé oblasti obličeje po jemné detaily vede ke snížení výpočetní náročnosti.

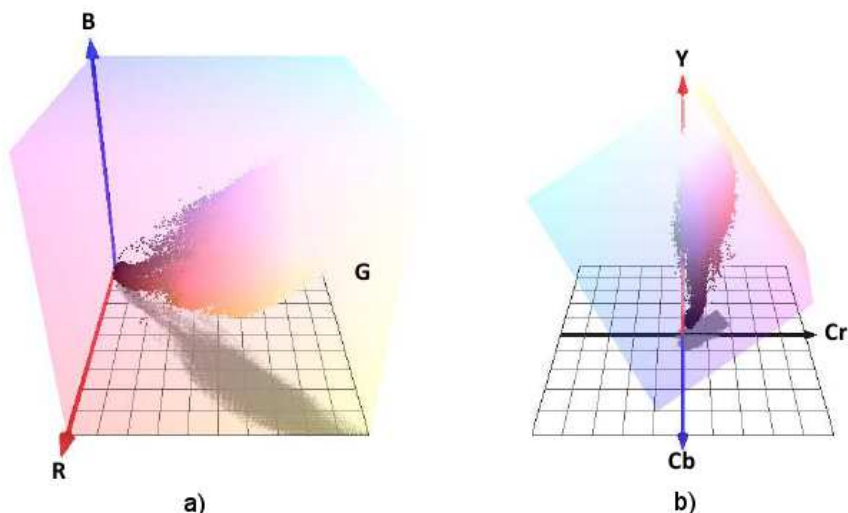
3.2 Lokalizace obličeje pomocí barvy kůže

Detekce obličeje pomocí barvy kůže [30], [23] je základní a velmi úspěšnou metodou pro detekci obličeje. Pomocí této metody jsou nalezeny kandidátské oblasti, v nichž by se mohl nalézat obličej, a pak se na tuto omezenou oblast aplikují další metody, které by byly pro detekci obličeje v celém obraze velmi pomalé. Tyto metody potvrzují, zda je kandidátská oblast obličejem. U lokalizace pomocí barvy kůže se využívá toho, že lidská kůže sdílí určitý malý podprostor v barevném prostoru. Vymezením tohoto prostoru jsme schopni modelovat barvu lidské kůže. Z této znalosti pak může být klasifikována barva jednotlivých pixelů obrazu.

Nejdříve je vhodné zvolit barevný model. U základního barevného modelu RGB je barva definována třemi složkami (červená, zelená a modrá) o různé intenzitě. Dalším barevným modelem, který je možné použít, je barevný model oddělující jasovou a barevnou složku YCbCr. Na tento barevný model může být velice jednoduše převeden model RGB. Převodní vztah je následující:

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 65.481 & 128.553 & 24.966 \\ -37.797 & -74.203 & 112.00 \\ 112.000 & -93.786 & -18.214 \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (3.1)$$

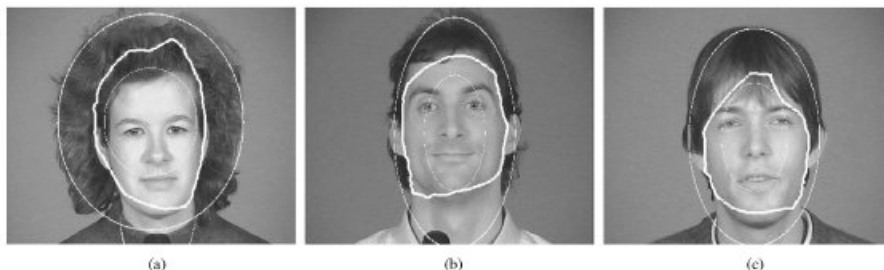
Dále je nutné zvolit podprostor pro barvu kůže 3.3. Toto rozložení je získáno experimentálně srovnáním různých testovacích vzorů. Velkou výhodou je snadná implementace a rychlá klasifikace. Úspěšnost výsledku je velmi závislá na vybraném modelu a přesnosti pravidel, která jsou určena experimentálně [30], [23].



Obrázek 3.3: a) Podprostor barvy kůže v RGB. b) Podprostor barvy kůže v YCbCr. Převzato z [30].

3.3 Detekce obrysu tváře

Další možností, jak lokalizovat obličej v obraze, je pomocí Houghovy transformace [20], [18]. Z apriorní znalosti toho, že lidský obličej má přibližně tvar elipsy, se pomocí Houghovy transformace hledají elipsy v obraze. Nalezené oblasti jsou prohlášeny za kandidátní oblasti a na ně jsou aplikovány další metody, které rozhodnou, zda se jedná o obličej či nikoliv. Tento způsob lokalizace obličeje je velice časově náročný v porovnání například s metodou založenou na detekci barvy kůže.



Obrázek 3.4: Nalezené obličejové pomoci Houghovy transformace a zvýraznění přesné oblasti obličeje pomocí aktivních kontur. Převzato z [18].

3.4 Porovnávání šablon

Metoda porovnávající šablony neboli Template Matching využívá standardní vzor obličeje, nejčastěji v čelním pohledu, který je ručně předdefinován nebo parametrizován funkcí. Na vstupním obraze je korelace hodnot standardních vzorů počítána nezávisle pro kontury obličeje, očí, nos a ústa. Existence obličeje je stanovena na základě korelace hodnot. Tento přístup má tu výhodu, že je jednoduchý na implementaci. Byl však prokázán za nevhodný pro rozpoznání obličeje, protože nedokáže efektivně pracovat s rozdíly v měřítku, pozici a tvaru obličeje. K docílení neměnnosti měřítka a tvaru obličeje bylo následně navrženo multirozlišení, multiměřítka, dílčí šablony a deformovatelné šablony [32].

3.4.1 Předdefinované šablony

Možností použití několika předdefinovaných dílčích šablon je několik. Jednou z nich je použití samostatné šablony pro oči, nos, ústa a obrys tváře. Každá šablona je definována samostatnými pravidly. V prvním kroku jsou porovnávány oblasti se šablonou obrysu obličeje a jsou detekovány kandidátní oblasti obličejů. Poté je na tyto oblasti kandidátských obličejů použito porovnávání s ostatními dílčími šablonami [32].

Dalším způsobem lokalizace obličeje pomocí předdefinovaných šablon je metoda založená na šabloně tvaru obličeje při čelním pohledu. Nejprve je použit hranový detektor

pro extrakci hran, například Sobelův hranový detektor. Tyto hrany jsou seskupeny dohromady pro vyhledávání šablony obličeje s několika omezeními. Po detekci umístění obrysu obličeje je stejný proces aplikován s různými měřítky k nalezení rysů jako jsou oči, obočí a ústa [32].

3.4.2 Tvarovatelné šablony

Metody založené na tvarovatelných šablonách k extrakci obličejových rysů jsou založeny na vhodném elastickém modelu, který využívají většinou k detekování hran metodou energetických křivek, zvaných též hadi či aktivní kontury, kdy se křivky deformují pomocí tzv. vnitřních, vnějších a obrazových sil. Vnitřní síly kontrolují hladkost průběhu, obrazové síly směřují tvarování kontury směrem ke hraně objektu a vnější síly jsou výsledkem počátečního umístění kontury. Aktivní kontura se postupně tvaruje ke hraně objektu do pozice s nejmenší celkovou energií [24].

Mějme konturu reprezentovanou diskrétně [24]:

$$v_n = [x_n, y_n], \quad n = 0, 1, \dots, N. \quad (3.2)$$

Výsledná pozice aktivní kontury E_S odpovídá lokálnímu minimu energie:

$$E_S = \sum_{n=1}^N E_N v_n + \sum_{n=1}^N E_I v_n + \sum_{n=1}^N E_T v_n, \quad (3.3)$$

kde E_N je vnitřní energie kontury, E_I je energie obrazu a E_T je energie omezení. Vnitřní energie kontury může být definována jako energie spojitosti E_C a zakřivení E_K :

$$E_N = \alpha(n)E_C(v_n) + \beta(n)E_K(v_n), \quad (3.4)$$

$$E_C = \frac{d - |v_n - v_{n-1}|}{\max d - |v_n(j) - v_{n-1}|}, \quad (3.5)$$

$$E_K = \frac{|v_{n-1} - 2v_n + v_{n+1}|^2}{\max |v_{n-1} - 2v_n(j) + v_{n+1}|^2}, \quad (3.6)$$

kde $\alpha(n)$ a $\beta(n)$ jsou váhové funkce definující elasticnost a tuhost kontury, d je průměrný obvod kontury a $v_n(j)$ reprezentuje osmiokolí bodu v_n , $j = 1, 2, \dots, 8$.

Metodou založenou na tvarovatelných šablonách je metoda popsána v [12] využívající aktivních kontur [11], [13]. Na obraz je nejdříve aplikován filtr k rozostření a pak morfologický operátor ke zvýraznění hran. Upravených n bodů hada je použito k nalezení a eliminaci malých rušivých segmentů zakřivení. Obličej je u této metody aproximován elipsou a k nalezení je možné použít Houghovu transformaci aplikovanou na oblasti, které jsou zvýrazněny hady, a nalézt dominantní elipsu reprezentující obličej. Tímto je získána sada 4 parametrů popisujících elipsu a oblast reprezentovaná touto sadou je prohlášena za kandidátskou oblast obličeje. Na každou z těchto kandidátských oblastí je aplikována podrobná metoda tvarovatelných šablon, tentokrát k nalezení detailních rysů. Pokud je nalezena většina obličejových rysů a jejich vzájemná poloha je ve správném poměru, je kandidátská oblast považována za obličej [32], [33].

Kapitola 4

Nalezení dominantních rysů obličeje

Nalezení dominantních rysů obličeje je nedílnou součástí strukturovaného přístupu rozpoznání obličeje. Postupně jsou lokalizovány jednotlivé dominantní části obličeje, jako jsou oči, ústa, nos a obočí. Po detekci je nutná normalizace kvůli předpokládaným rušivým vlivům. Poté je možné detekované dominantní části porovnávat s údaji uloženými v databázi pomocí různých porovnávacích algoritmů a na základě statistické podobnosti rozhodnout o míře jejich shody.

Pokud jsou nalezeny všechny tyto dominantní rysy, například pomocí metod popsanych v následující části textu, na jednom obrázku či na předem daném výřezu obrázku, může být tato oblast prohlášena za obličej.

4.1 Detekce očí

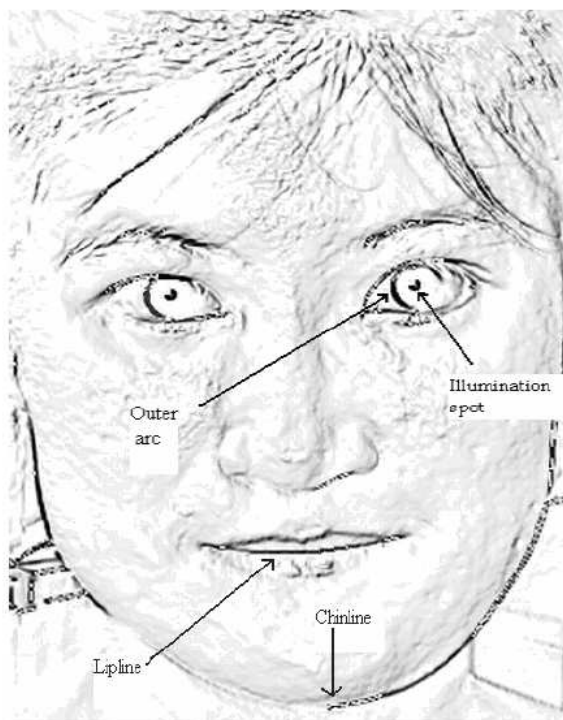
Nejdůležitějším dominantním rysem obličeje jsou bezesporu oči. Existují různé metody pro jejich lokalizaci. Některé metody jsou založeny na šablonovém přístupu či na příznamech obrazu, jiné využívají apriorních znalostí například tvaru duhovky a v neposlední řadě existují i metody kombinující různé typy přístupů.

4.1.1 Snímání osvětleného bodu oka

První metodou pro detekci a lokalizaci očí je metoda Snímání osvětleného bodu oka (Capturing Eye Illumination Spot) popsaná v [22], která využívá znalosti toho, že střed pupily má vlivem osvětlení v obraze velmi vysokou intenzitu. Po provedení nutných úprav obrazu, jakými jsou převedení obrazu do stupňů šedi a zvýraznění hran, získáme obraz, kde části reprezentující vnější okraj duhovky a zornici jsou velmi tmavé a obraz mezi nimi je naopak velmi světlý.

Detekce očí může být reprezentována jako množina $TEye = \{G1, D, G2\}$, kde $G1$ je práh ve stupních šedi určující barvu vnějšího okraje duhovky a zornice. Tento práh bude nabývat

velmi malých hodnot reprezentujících téměř černou barvu. $G2$ naproti tomu reprezentuje práh ve stupních šedi pro body, které jsou mezi vnějším okrajem duhovky a zornicí. Prah zde naopak nabývá vysokých hodnot reprezentujících téměř bílou barvu. D je minimální práh vzdálenosti mezi vnějším okrajem duhovky a zornicí. V této oblasti se musí vyskytovat pouze body barvy prahu $G2$. Pokud bude vhodně zvolen práh D , dá se díky jeho přítomnosti předpokládat, že metoda bude fungovat i na obrázky tváří, které jsou různě vzdáleny. Tato šablona bude aplikována na upravený obrázek se zvýrazněnými hranama 4.1.



Obrázek 4.1: Snímání osvětleného bodu oka. Outer arc – vnější okraj duhovky, Illumination spot – osvětlený bod zornice. Převzato z [22].

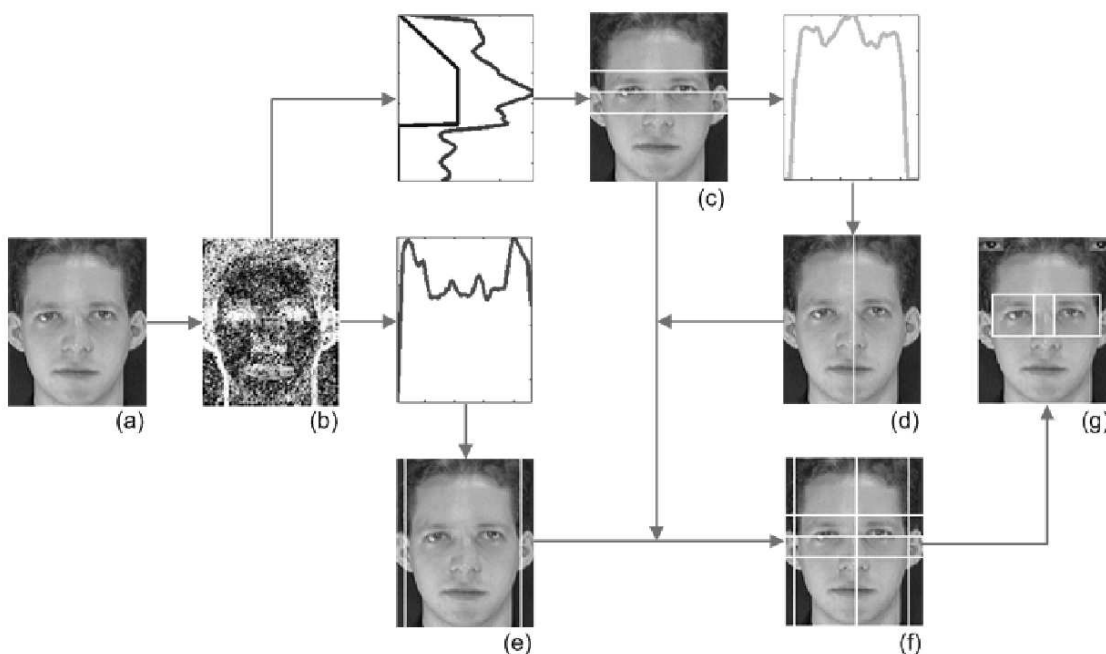
4.1.2 Robustní algoritmus pro detekci očí ve tváři ve stupních šedi bez brýlí

Tato metoda popsaná v [19] je robustním algoritmem pro detekci očí v obrazu tváře bez brýlí (A Robust Algorithm for Eye Detection on Gray Intensity Face without Spectacles), který je ve stupních šedi, založeným na kombinaci metod využívajících charakteristických vlastností a přístupu na základě šablon. Kombinací různých technik a překonáním jejich nedostatků tento algoritmus používá nejprve metodu založenou na příznacích k přibližnému určení oblasti očí. Poté metoda založená na šablonách přesně určí umístění oka a nalezne střed duhovky. Tento přístup je nezbytný z toho důvodu, že postup založený na

porovnávání se šablonou je časově velice náročný. Kdyby se oči měly detekovat pouze šablonovým přístupem, musel by se procházet celý obraz bod po bodu. A jelikož by nebyla známa ani přibližná velikost očí, musel by obraz být vícekrát porovnáván bod po bodu se šablonami různých velikostí. Z tohoto důvodu je nejdříve omezena oblast pravděpodobného výskytu očí a díky tomu se rapidně sníží časová náročnost porovnávání se šablonou. Jelikož ještě známe přibližnou velikost oblastí, kde by se mohly nacházet oči, je možné odhadnout velikost šablony a tím ještě více snížit časovou náročnost.

Postup detekce je popsán na obrázku 4.2. Nejdříve je z původního obrázku (a) vypočítán gradient obrazu, který je zobrazen na obrázku (b). Pomocí gradientu je určena souřadnice Y pravděpodobného výskytu očí (obrázek (c)). Průmětem všech bodů v horizontální oblasti v obrázku (c) je získán vertikální střed obličeje (obrázek (d)). Svislým průmětem gradientu obrázku je získán obrázek (e) a z něj je možno nalézt hranice tváře a odhadnout šířku obličeje. Spojením výsledků (c), (d) a (e) získáme segmentovaný obraz (f) na jehož základě a na základě odhadu šířky obličeje jsou zvýrazněny oblasti očí (obrázek (g)).

Na tyto oblasti očí je již možné aplikovat pomalejší, ale zato přesnější metodu založenou na šablonovém přístupu, například metodu Snímání osvětleného bodu oka zmíněnou v 4.1.1. Po aplikaci šablonové metody je možné rozhodnout, zda se o oči jedná či nikoliv.



Obrázek 4.2: Detekce oblastí očí. Převzato z [19].

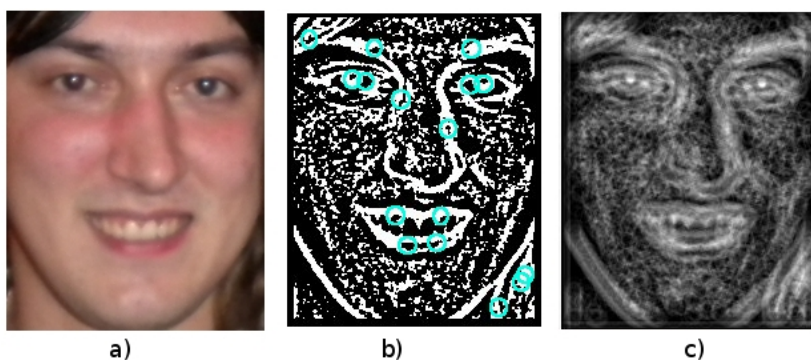
4.1.3 Houghova transformace pro detekci kružnic

Houghovu transformaci pro detekci kružnic je možné také využít k nalezení očí. U této metody se využívá znalosti, že duhovka má kulatý tvar. Na základě velikosti obličeje je určen přibližný průměr duhovky oka. Poté je na obraz aplikována Houghova transformace hledající kružnice zadaného průměru. Tato metoda má tu výhodu, že dokáže detekovat i přerušené, překryté či jinak poškozené kružnice. Nalezené kružnice jsou poté prohlášeny za kandidátské oblasti očí (viz obrázek 4.4). Tyto kandidátské oblasti jsou dále zpracovávány, a pokud budou vyhovovat zadaným podmínkám, budou prohlášeny za oblast očí (viz obrázek 4.5). Mezi tyto podmínky mimo jiné patří, že v určité vzdálenosti nad okem musí existovat obočí a dále musí být mezi očima ústa.

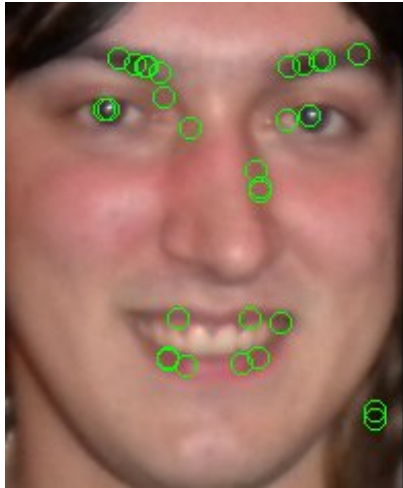
Při detekci kružnic pomocí Houghovy transformace se postupuje tak, že se hledají kružnice popsané rovnicí

$$(x - a)^2 + (y - b)^2 = r^2, \quad (4.1)$$

kde a , b a r jsou hledané parametry. Jelikož jsou hledány tři parametry bude Houghův prostor dimenze 3. Pro účely detekce očí se hledají pouze parametry a a b , jelikož parametr r je explicitně určen. Postupuje se tak, že nejdříve je obrázek upraven prahováním na binární (4.3 b). Poté je pro každý bod určený souřadnicemi $[x, y]$ aplikována rovnice (4.1). Pokud se na testovaném místě nalézá zvolená hodnota bodu (dle binarizace – buď 0 nebo 255), je na zadané místo v akumulátoru přičtena jednička. Nakonec jsou hodnoty akumulátoru seřazeny vzestupně, kdy nejvyšší hodnotu mají nejpravděpodobnější místa výskytu kružnic v obraze. Pro zvýšení přesnosti Houghovy transformace, hlavně na okrajích obrázku, může mít Houghův prostor, a tedy i akumulátor, větší rozměr než obrázek. Na obrázku 4.3 c) je možné vidět Houghův prostor. Body s nejvyšší hodnotou intenzity jsou středy hledaných kružnic. Na obrázku 4.4 jsou zvýrazněny všechny nalezené kružnice o zadaném průměru a na obrázku 4.5 jsou zvýrazněny pouze kandidátské oblasti očí.



Obrázek 4.3: Houghova transformace. a) původní obrázek, b) binarizovaný obrázek se zvýrazněním nalezených kružnic, c) Houghův prostor (hodnoty jsou přepočítány do intervalu $[0, 255]$).



Obrázek 4.4: Kandidátské oblasti očí. Všechny kružnice detekované Houghovou transformací o zadaném průměru.



Obrázek 4.5: Oblast očí. Kružnice prohlášené za oblast očí. Zelenou barvou pravé oko, červenou levé.

4.2 Detekce úst

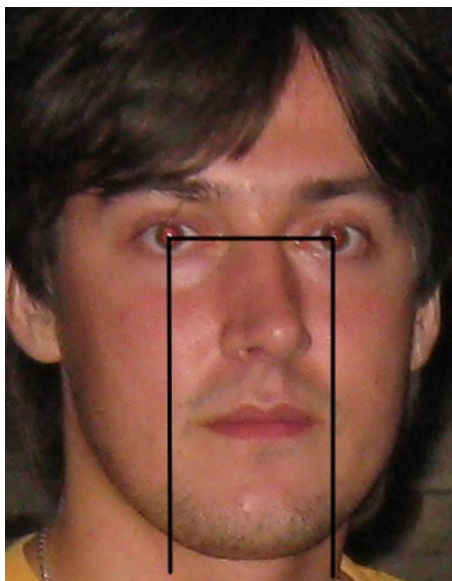
Ústa také patří mezi důležité dominantní rysy. Detekce úst může být založena na znalostních metodách, jako jsou například metody založené na znalosti tvaru či barvy rtů, kdy rty obsahují vysokou hodnotu barvy červené a naopak velmi nízkou hodnotu barvy modré barevného modelu RGB. Při použití metod založených na znalosti tvaru musí být bráno v potaz i to, že tvar rtů může být ovlivňován mimickými svaly.

4.2.1 Ze znalosti umístění očí

Tato metoda [22] vychází ze znalosti umístění očí, pomocí níž lze určit obdelníkovou oblast, kde se rty budou nalézat. Tato oblast je zobrazena na obrázku 4.6. U této metody je nutné zvolit práh hodnoty intenzity pro rty a také minimální velikost rtů. Pokud se na jedné linii minimální délky L bude nalézat pouze hrana požadovaných hodnot intenzity, tak se s největší pravděpodobností jedná o rty. Šablona pro lokalizaci rtů může být definována například jako množina $TL = \{x_1, x_2, y, G, L\}$, kde x_1 je souřadnice X středu levého oka, x_2 je souřadnice X středu pravého oka, y je největší Y souřadnice ze středů obou očí, G je práh v hodnotách intenzity bodů ležících na linii rtů a L je práh minimální délky linie rtů.

4.2.2 Na základě barevných modifikací obrazu

Lokalizace úst na základě barevných modifikací obrazu podle [30] spočívá v tom, že je nejdříve provedena barevná transformace do F-barevného prostoru, který je vytvořen pomocí Fisherovy lineární diskriminační analýzy (viz obrázek 4.7). Tato transformace je založena na předpokladu, že rty obsahují vysokou hodnotu červené barvy a velmi nízkou hod-

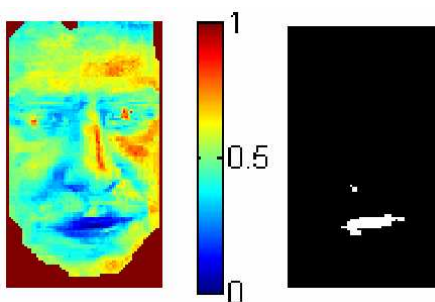


Obrázek 4.6: Vymezená oblast pro hledání rtů.

notu barvy modré modelu RGB.

$$F_{prostor} = [-0.289 \ 0.379 \ 0.038] \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix}. \quad (4.2)$$

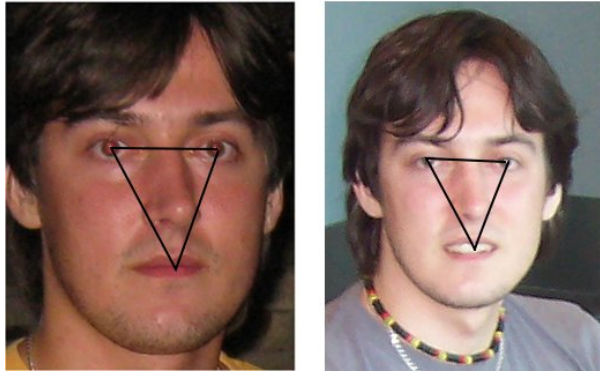
Jelikož u této barevné transformace existuje možnost výskytu části pozadí s velmi vysokou hodnotou červené a nízkou hodnotou modré barvy, aplikujeme barevnou transformaci pouze na kandidátské oblasti obličejů. Pak provedeme prahování 4.7 a poté pouze stačí provést segmentaci, například pomocí algoritmu mean shift [25]. Shluk, který se nachází v oblasti mezi očima v určité vzdálenosti, kterou je možné odhadnout, je prohlášen za ústa.



Obrázek 4.7: Obraz v F-barevném prostoru a její binární obraz. Převzato z [30].

4.3 Detekce nosu

Detekovat nos je možné například podle umístění očí a úst. Díky této apriorní znalosti je možné vymežit oblast, kde se nachází nos. Na obrázku 4.8 je zobrazena vymezená oblast, kde se nalézá nos. Dále oblastí zájmu může být přesné umístění špičky nosu. Tu je možné lokalizovat z poznatku, že špička nosu je obvykle světlejší než okolí [2].



Obrázek 4.8: Vymezená oblast nosu.

4.4 Detekce obočí podle umístění očí

Tato metoda [22] je založena na znalosti umístění očí podle níž je možné lokalizovat i polohu obočí. Šablona pro obočí může být definována jako množina $TEb = \{G, L, C\}$, kde G je práhová hodnota intenzity bodů obočí, L je práh vzdálenosti obočí od středu zornice, C je počítadlo prahových hodnot, které vymezuje šířku obočí. Tato šablona je aplikována pouze na oblast určenou z umístění očí, kdy v prvním kroku je z umístění očí určena minimální a maximální vzdálenost obočí od oka a také je podle velikosti oka určena práhová hodnota L . Prahová hodnota G je určena po zpracování obrázku hranovým detektorem, u obrázku, kde bod může nabývat hodnot intenzity od 0 do 255 je tato práhová hodnota většinou kolem 230 a byla zjištěna experimentálně. Na konci je vyhodnocena šířka obočí pomocí počítadla prahových hodnot C , a pokud splňuje podmínku minimální šířky obočí, tak je možné zadanou oblast prohlásit za obočí. Podmínku minimální šířky obočí je možné zvolit více způsoby, například poměrem ze vzdálenosti mezi levým a pravým okem.

Kapitola 5

Normalizace obrazu obličeje

U vzájemného porovnávání obrazů obličejů je zapotřebí počítat s tím, že každá osoba může být na obrázku pod různým úhlem, v různé vzdálenosti a různě osvětlená, a z tohoto důvodu je velice vhodná normalizace.

Umístění a orientaci obličeje je možné určit šesti parametry: třemi rotacemi a třemi translacemi. Shodná transformace obličeje nebo jiného objektu je určena těmito šesti parametry. Shodná transformace obličeje představuje pro velké množství rozdílů vzhled 2D obrazu. Mimo jiné i změny osvětlení způsobené libovolně umístěnými světelnými zdroji s různou intenzitou také představují značné množství rozdílů. Současně nestabilní deformace obličeje (v důsledku ovládní mimických svalů a rozdílů identity) způsobují ještě jemné rozdílly ve 2D obraze. Identita jedince, ačkoliv je zachycena s těmito malými odchylkami, nemusí být určena v důsledku mnoha dalších pohybů těla a osvětlení obličeje, proto je nezbytné tyto rušivé vlivy kompenzovat nebo normalizovat obličej na určitou polohu a osvětlení tak, aby byly odchylky díky tomu minimalizovány. V důsledku toho, že se malé odchylky v obraze, způsobené identitou jedince a ovládním svalů, stanou dominantním zdrojem intenzity odchylek v obraze, mohou být tyto malé odchylky analyzovány pro účely rozpoznávání [10].

Připomeňme výstup fáze detekce obličeje a lokalizace. Oči, nos a ústa byla určena použitím přímých metod zpracování obrazu. Předpokládejme nyní, že horizontální pozice nosu byla také určena a přesná poloha pro špičku nosu je taktéž dostupná. Detekce umístění těchto příznakových bodů (oči, nos a ústa) dává odhad umístění jednotlivých tváří. Jakmile je známá pozice nebo 3D pozice a orientace obličeje, je možné invertovat účinek translací a rotací a vytvořit standardizovaný čelní pohled jedince. Navíc pozice příznakových bodů dovoluje zhruba segmentovat obrys tváře k odstranění rušivých informací pozadí. Po segmentaci může být vypočítán histogram obličeje k vyrovnání světelných změn v obraze [10].

5.1 Roztažení kontrastu

Roztažení kontrastu [9] (často nazývané normalizace) je jednoduchou metodou pro vylepšení obrazové kvality. Kontrast obrazu je zlepšen „natažením“ rozsahu hodnot intenzity, tak aby obrázek obsahoval rozsah požadovaných hodnot, například celou škálu hodnot intenzity, které obrázek příslušného typu obrazu umožňuje. Narozdíl od techniky vyrovnání histogramu může roztažení kontrastu použít pouze lineární škálovací funkci na hodnoty obrazových bodů. V důsledku toho není výsledek příliš hrubý. Ve většině případů se roztažení kontrastu aplikuje na obrázek v odstínech šedi a výstupem je opět obrázek ve stupních šedi.

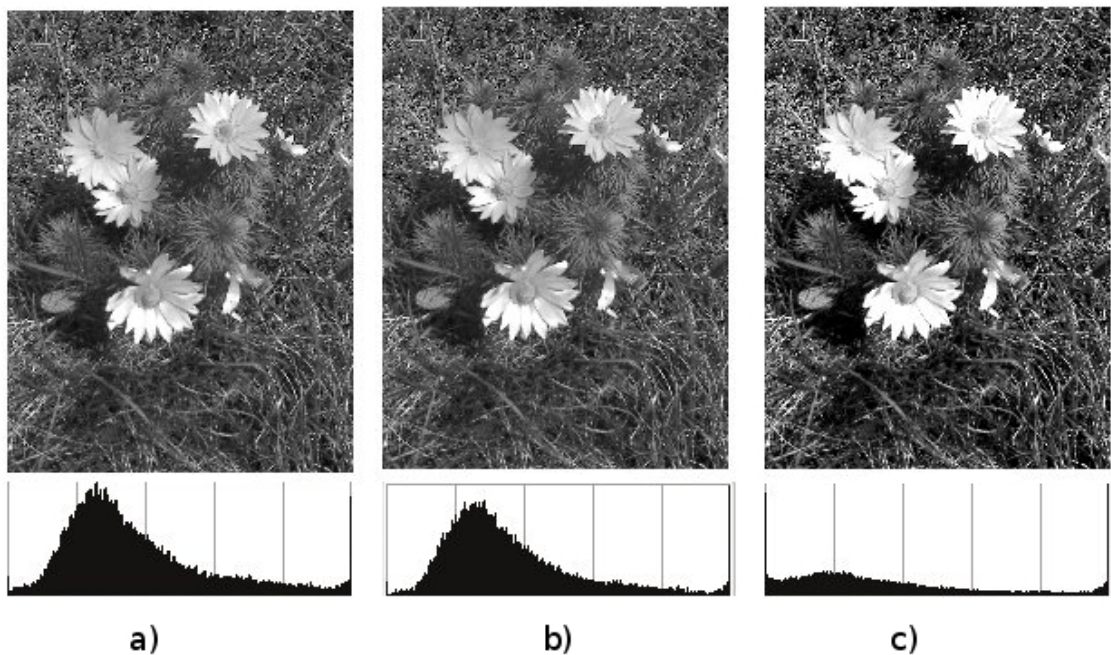
Před samotným provedením roztažení kontrastu je nezbytné určit mezní horní a dolní hodnoty bodů podle kterých má být obraz normalizován. Těmito mezními hodnotami jsou často minimální a maximální hodnoty bodů, kterých může daný typ obrázku nabývat. Pro obrázek v odstínech šedi může tato dolní a horní hranice být 0 a 255, jak je zobrazeno na obrázku 5.1 b). Dále je v aktuálním obraze nalezena nejnižší a nejvyšší hodnota bodu. Teď již jen stačí aplikovat pro každý bod obrazu následující funkci a získat tak novou hodnotu bodu

$$P_{out} = (P_{in} - c) \left(\frac{b-a}{d-c} \right) + a, \quad (5.1)$$

kde P_{in} je původní hodnota bodu, c je minimální hodnota bodu v aktuálním obrázku, a je zvolená minimální hodnota bodu, b je maximální zvolená hodnota bodu, d je maximální hodnota bodu v aktuálním obrázku. Hodnoty menší než a jsou nastaveny na a , hodnoty větší než b jsou nastaveny na b . Problém může nastat, pokud jediný bod má velmi nízkou, či naopak vysokou hodnotu intenzity. Tím může být vážně ovlivněna hodnota c a d , což může vést k velice zkreslené a neodpovídající změně hodnoty ostatních bodů. Proto je vhodné nejprve vzít histogram obrázku a z něj teprve vybrat hodnoty c a d . Pak je možné například zvolit, že 5 % hodnot bodů bude nižších než c a 5 % hodnot bude vyšších než d jak je možno vidět na obrázku 5.1 c). Tímto se zabrání negativní změně výsledného kontrastu vlivem odhlehých a málo se vyskytujících hodnot intenzity bodů.

Další obvyklou technikou je použití histogramu intenzit k nalezení nejvíce zastoupených hodnot intenzity v obrázku a následně definování mezní hodnoty, pod kterou budou data ignorována. Histogram intenzit se postupně zpracovává od 0 až do prvního výskytu hodnoty vyšší než mezní hodnoty. Této hodnoty bude nabývat c . Podobně je histogram intenzity zpracováván od 255, dokud první hodnota není vyšší než mezní hodnota. Tuto hodnotu reprezentuje d .

Výše popsané způsoby roztažení kontrastu lze aplikovat také na barevné obrázky. V tomto případě budou všechny barevné kanály roztaženy použitím stejných změn z důvodu zachování správného poměru barev.



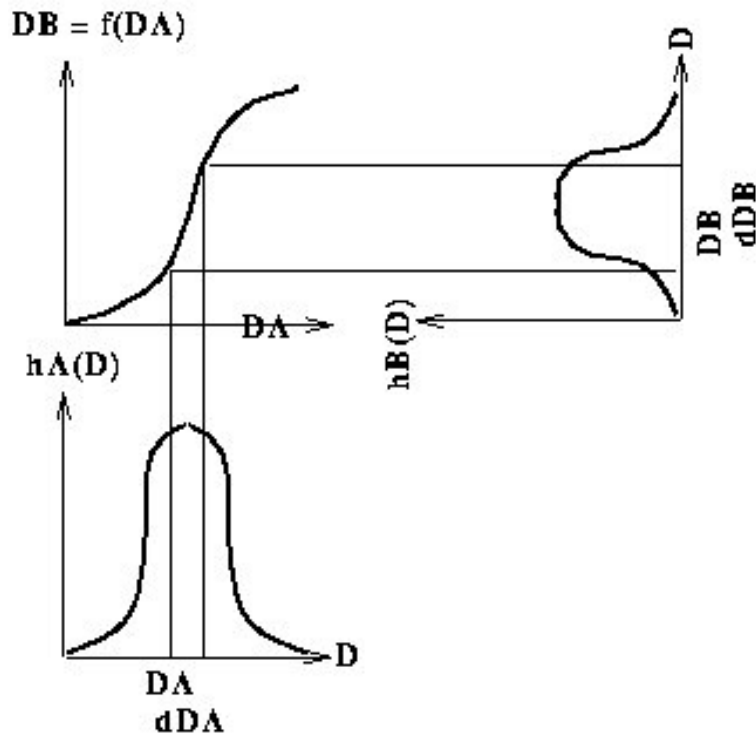
Obrázek 5.1: Roztažení kontrastu. a) původní obrázek, b) roztažení kontrastu zvolením dolní a horní hranice 0 a 255, c) roztažení kontrastu, dolní a horní hranice určena pomocí procent.

5.2 Vyrovnání histogramu

Techniky založené na vyrovnání histogramu [9] poskytují sofistikovanější metody pro úpravu dynamického rozsahu a kontrastu obrazu tím, že upraví histogram intenzit na požadovaný tvar. Na rozdíl od metody vyrovnání kontrastu, modelování histogramu může využít nelineární a nemonotónní přenosové funkce k namapování hodnot intenzity bodů ze vstupních na výstupní obrázky. Vyrovnání histogramu využívá monotónní, nelineární mapování, které znovu určí hodnoty intenzit bodů ve vstupním obraze tak, aby výstupní obraz obsahoval rovnoměrné rozložení intenzit. Tato technika se používá v obrazových srovnávacích procesech, protože je efektivní ve zvyšování detailů a pro opravu nelineárních efektů.

Modelovaný histogram je obvykle reprezentován použitím spojitých, spíše než diskretních průběhových funkcí. Proto je možné se domnívat, že obrazy zájmu obsahují spojitě úrovně intenzity (v intervalu $[0, 1]$) a transformační funkce f , která zobrazuje vstupní obraz $A(x, y)$ na výstupní obraz $B(x, y)$, je v tomto intervalu spojitá. Dále se předpokládá, že přenosové pravidlo $D_B = f(D_A)$ je jednoznačné a monotónně rostoucí tak, že je možné definovat inverzní pravidlo $D_A = f^{-1}(D_B)$. Příklad takové přenosové funkce je znázorněn na obrázku 5.2.

Všechny body vstupního obrazu s hustotami v oblasti D_A na $D_A + dD_A$ budou mít své hodnoty změněny tak, že je předpokládána hodnota hustoty výstupního bodu v rozmezí



Obrázek 5.2: Transformační funkce histogramu. Převzato z [9].

D_B až $D_B + dD_B$. Plocha $h_A(D_A)dD_A$ a $h_B(D_B)dD_B$ bude tedy stejná, což odpovídá

$$h_B(D_B) = \frac{h_A(D_A)}{d(D_A)}, \quad (5.2)$$

kde $d(x) = \frac{df(x)}{dx}$.

Tento výsledek může být zapsán v jazyce pravděpodobnostní teorie jestliže histogram h je považován za spojitou funkci hustoty pravděpodobnosti p popisující distribuci hodnot intenzity:

$$p_B(D_B) = \frac{p_A(D_A)}{d(D_A)}. \quad (5.3)$$

V případě ekvalizace histogramu by všechny výstupní hustoty pravděpodobnosti měly mít stejný podíl maximální hodnoty intenzity vstupního obrazu D_M . Přenosová funkce nezbytná k dosažení tohoto výsledku je

$$d(D_A) = D_M \times p_A(D_A). \quad (5.4)$$

Z toho vyplývá vztah

$$f(D_A) = D_M \int_0^{D_A} p_A(u) du = D_M \times F_A(D_A), \quad (5.5)$$

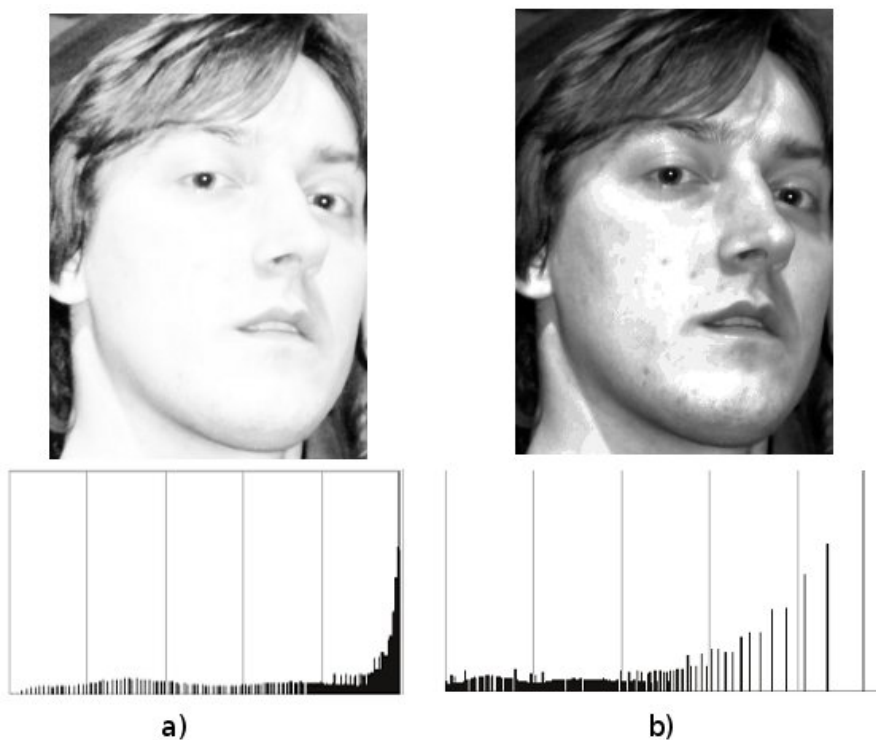
kde $F_A(D_A)$ je prostá kumulativní distribuce pravděpodobnosti původního obrazu.

Digitální provedení vyrovnání histogramu je obvykle vykonáno definováním přenosové funkce ve tvaru:

$$f(D_A) = \max\left(0, \text{round}\left[\frac{D_M \times n_k}{N^2}\right] - 1\right), \quad (5.6)$$

kde N je počet obrazových bodů a n_k je počet bodů s hodnotou intenzity k nebo nižší.

V digitálním provedení nemusí být výstup obrazu nutně plně vyrovnán, proto mohou vznikat tzv. díry v histogramu. Tento jev je naštěstí snižován s tím, jak počet bodů a intenzita kvantování hodnot intenzity ve vstupním obrázku roste. Na obrázku 5.3 je zobrazen původní a upravený obrázek pomocí vyrovnání histogramu.



Obrázek 5.3: Vyrovnání histogramu. a) původní obrázek, b) obrázek po vyrovnání histogramu.

Kapitola 6

Rozpoznání obličeje

Po lokalizaci a normalizaci obličeje lze přistoupit k dalšímu kroku, jímž je rozpoznání obličeje. Existuje mnoho různých metod a přístupů. Jsou to například metody založené na lineární analýze. Z nich je možné jmenovat například analýzu hlavních komponent (PCA), analýzu nezávislých komponent (ICA) a lineární diskriminační analýzu (LDA). Někdy lineární přístup nestačí, tento nedostatek však řeší jádrové metody, například jádrová analýza hlavních komponent (KPCA) a jádrová diskriminační analýza (KDA). Jiným přístupem může být metoda založená na elastických shlukových grafech (EBGM), která se snaží o kopírování topologické struktury lidské tváře, či aktivní vzhledový model (AAM) využívající statistického modelu a mnoho dalších.

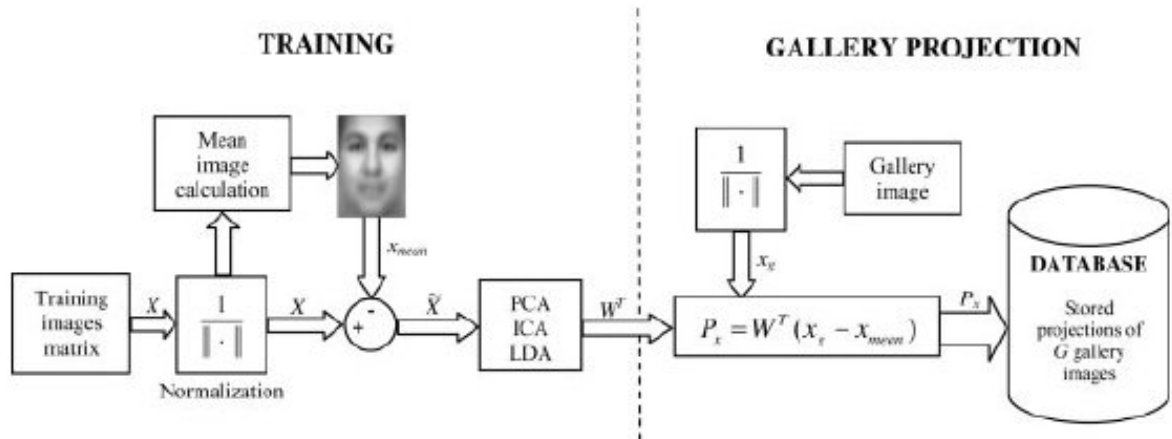
6.1 Lineární analýza

Metody lineární analýzy se snaží lineárně transformovat původní obrazová data reprezentovaná vektorově do projekčního vektoru a tím současně nalézt vhodné lineární rozdělení prostoru, aby data jednotlivých tříd byla co nejlépe separovatelná. Mezi tři základní metody lineární analýzy patří PCA, ICA a LDA. Každá metoda má svou vlastní reprezentaci vysokodimenzionálního vektorového prostoru založenou na statistických metodách.

Na obrázku 6.1 je znázorněn princip transformace dat do projekčního vektoru. V levé části obrázku je zobrazeno trénování systému s redukcí prostoru a v pravé části je možné vidět proces projekce původních obrazových dat do prostoru s nižší dimenzí. X je matice obsahující obrázky vyjádřené jako vektory v jednotlivých sloupcích matice, x_{mean} je průměrný obrázek (ve formě vektoru), \tilde{X} je matice obsahující odečtené průměrné obrázky ve sloupcích, W^T je projekční matice získaná procesem lineární analýzy nad daty \tilde{X} a x_g je obrázek z galerie reprezentovaný vektorem. Během trénovací fáze je vypočítána projekční matice W^T a poté je obrazová galerie promítnuta do podprostoru $P_x = W^T(x_g - x_{mean})$, kde P_x je projekční vektor, který je následně uložen do databáze.

Při procesu identifikace je nejdříve výše uvedeným postupem vypočítán projekční vektor P_x a v následujícím kroku je porovnáván s projekčními vektory uloženými v databázi.

Výsledkem je posloupnost vzdáleností, kdy výsledek je vybrán dle nejmenší vzdálenosti.



Obrázek 6.1: Proces lineární analýzy. Převzato z [5].

6.1.1 Analýza hlavních komponent

Analýza hlavních komponent (Principal Component Analysis), nebo též Karhunen-Loéve transformace popsaná v [1], [6], [8], je standard techniky pro statistické (strukturální) rozpoznávání a zpracování signálu pro redukci dat a extrakci příznaků. Protože vzor v obraze většinou obsahuje nadbytečné informace, je možné se jeho mapováním na vektor příznaků zbavit redundance, a přesto zachovat většinu důležitého informačního obsahu vzoru. Tyto extrahované příznaky mají velkou úlohu při rozlišení vstupních vzorů.

Obraz tváře ve 2D s rozměry $X \times Y$ může být také považován za jednorozměrný vektor o rozměru XY . Například obraz tváře z databáze Olivettin Research Labs o rozměrech 112×92 může být považován za vektor o rozměru 10 304 nebo za bod v 10 304 rozměrném prostoru. Obrazy tváří, které si jsou celkově podobné, nebudou náhodně distribuovány v tomto obrovském prostoru, a proto mohou být popsány v relativně rozměrově malém podprostoru. Hlavní myšlenkou analýzy hlavních komponent je nalezení vektorů, které nejlépe představují rozdělení obrazů tváří uvnitř celého obrazového prostoru. Tyto vektory definují podprostor obrazů tváří, který nazýváme prostor obličejů. Každý z těchto vektorů má délku XY , popisující obraz s rozměry $X \times Y$, a je lineární kombinací původního obrazu tváře.

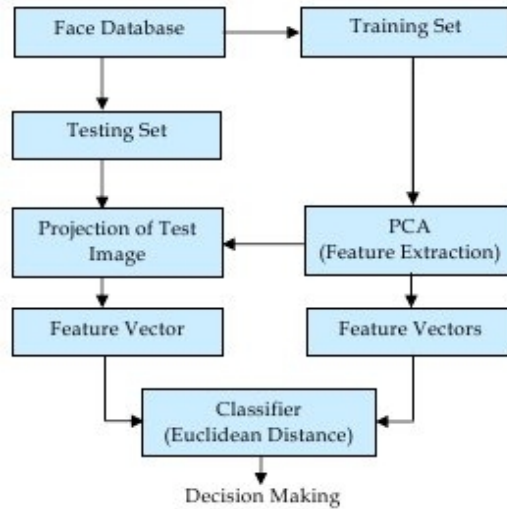
Při procesu identifikace pomocí PCA se může postupovat tak, že se nejdříve spočítá hodnota μ podle vztahu

$$\mu = \frac{1}{N} \sum_{i=1}^N x_i, \quad (6.1)$$

kde N je počet obrázků, x_i je i tý obrázek trénovací sady a μ je střední hodnota, která je poté odečtena od každého bodu obrázku reprezentovaného vektorem. Dále je vypočítána kovariační matice C

$$C = \frac{1}{N} \sum_{i=1}^N (x_i - \mu)(x_i - \mu)^T. \quad (6.2)$$

Z kovariační matice C jsou určeny vlastní vektory a čísla. Podle vlastních čísel jsou vzestupně seřazeny odpovídající vlastní vektory. Ze seřazené posloupnosti je vybráno prvních k vlastních vektorů, kde k odpovídá počtu hlavních komponent PCA. Z vybraných vlastních vektorů je generována množina dat v nové reprezentaci. Při vlastním procesu identifikace je nejdříve od testovacího obrázku odečtena střední hodnota μ , poté je obrázek porovnáván s trénovacími obrázky v nové reprezentaci pomocí k hlavních komponent za pomoci měřítka vzdálenosti. Výsledkem je obrázek z trénovací sady, který je nejbližší testovacímu obrázku. Celý postup je znázorněn na obrázku 6.2.



Obrázek 6.2: PCA přístup k rozpoznání obličeje. Převzato z [6].

6.1.2 Analýza nezávislých komponent

Analýza nezávislých komponent (Independent Component Analysis) [15] vyššího řádu statisticky určuje nezávislé zdrojové komponenty z jejich lineárního smíchání. ICA tak poskytuje vhodnější reprezentaci dat než metoda PCA a jejím cílem je poskytovat nezávislé dekompozice a reprezentace obrazových dat.

ICA náhodného vektoru hledá lineární transformaci s nejmenší lineární závislostí mezi komponentami, a to zejména pokud $X \in \mathbb{R}^2$ bude náhodný vektor reprezentující obraz, kde D je rozměr obrazového prostoru. Vektor je tvořen zřetěžením řádků a sloupců obrazu, který může být normalizován na jednotkový model nebo převeden na histogram. Kovarianční matice C je definována vztahem

$$C = E\{[X - E(X)][X - E(X)]^T\}, \quad (6.3)$$

kde E je očekávaný operátor, T udává transponovanou operaci a $X \in \mathbb{R}^{D \times D}$. Metoda ICA pomocí X rozkládá kovarianční matici C do následujícího tvaru

$$\sum_X = F\Delta F^T, \quad (6.4)$$

kde Δ je diagonální reálná pozitivní matice a F transformuje původní data X na Z

$$X = FZ \quad (6.5)$$

tak, že komponenty nových dat Z jsou nezávislé nebo alespoň co možná nejvíce nezávislé na ostatních komponentách.

Odvození ICA transformace F je možné pomocí algoritmu P. Comona, který se skládá ze tří operací: bělení, rotace a normalizace.

$$X = \Phi\Lambda^{1/2}U, \quad (6.6)$$

kde Φ a Λ jsou odvozeny z řešení následující rovnice vlastních hodnot.

$$\sum_X = \Phi\Lambda\Phi^t, \quad (6.7)$$

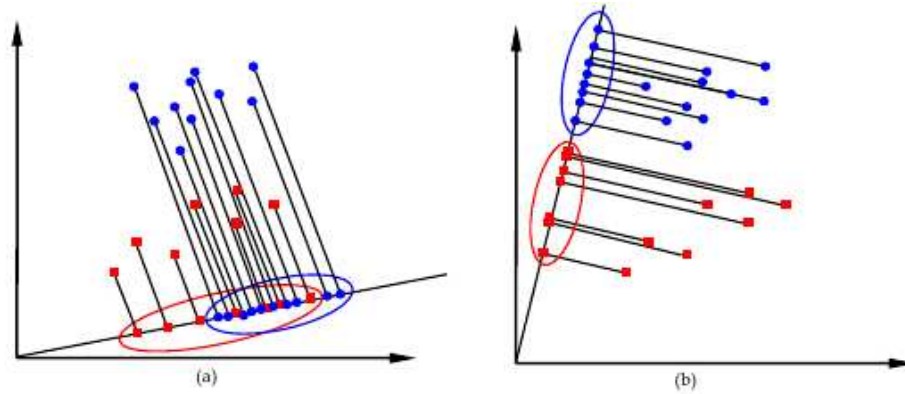
kde $\Phi = [\Phi_1, \Phi_2, \dots, \Phi_N]$ je ortonormální vlastní vektor matice a $\Lambda = \text{diag}[\lambda_1, \lambda_2, \dots, \lambda_N]$ je diagonála vlastních čísel matice C .

6.1.3 Lineární diskriminační analýza

Lineární diskriminační analýza (Linear Discriminant Analysis) popsaná v [1], [6], [8] překonává omezení metod založených na vlastních vektorech za použití Fisherova lineárního diskriminačního kritéria. Toto kritérium se snaží maximalizovat poměr mezi determinantem mezitřídní matice rozptylu promítnutých vzorků a determinantem vnitřní matice rozptylu promítnutých vzorků.

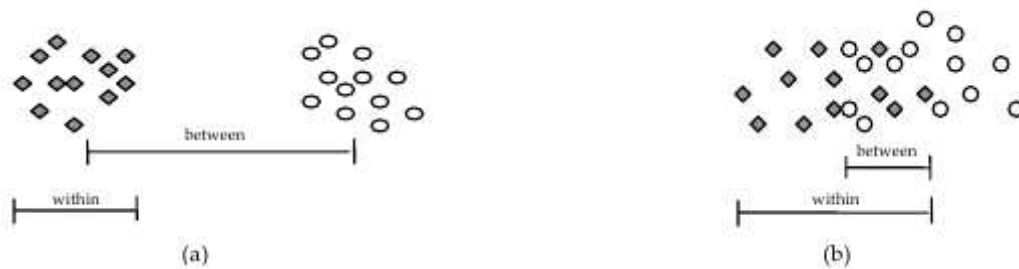
Fisherfaces seskupuje do skupin obrázky ze stejné třídy a odděluje obrázky z různých tříd. Obrázky jsou promítány z M^2 rozměrného prostoru do D rozměrného prostoru, kde D je počet tříd obrázků. Například uvažujme dvě sady bodů ve dvojrozměrném prostoru, které jsou promítány do jedné linie. V závislosti na směru linie mohou být body buď smíchány (obrázek 6.3 (a)) nebo odděleny (obrázek 6.3 (b)). Fisherův diskriminant nalezne linii, která nejlépe rozděluje body. Pro rozpoznání vstupního testovacího obrazu je promítnutý testovací obraz porovnán s každým promítnutým trénovacím obrazem a testovací obraz je určen jako nejbližší trénovací obraz.

Stejně jako vlastní prostor projekce jsou trénovací obrazy promítnuty do podprostoru. Testovací obrazy jsou promítnuty do stejného podprostoru a jsou určeny za použití porovnávacích měřítek. Rozdíl je v počítání podprostoru. Na rozdíl od metody PCA, která extrahuje příznaky nejlepší reprezentace obrazu obličeje, zkouší metoda LDA nalézt podprostor, který nejlépe popisuje rozdílné třídy obličejů, jak je možné vidět na obrázku 6.3. V rámci vnitřní rozptylové matice, také nazývané intra-personal, jsou reprezentovány rozdíly v podobě stejného jedince v důsledku rozdílného osvětlení a výrazu obličeje, zatímco v mezitřídní rozptylové matici, nazývané též extra-personal, jsou reprezentovány rozdíly v podobě odlišné totožnosti. Použitím této metody nalezneme směry projekce, které na jedné straně maximalizují vzdálenost mezi obrazy obličejů různých tříd a na straně druhé



Obrázek 6.3: (a) Smíchané body při promítnutí na linii. (b) Oddělené body při promítnutí na jinou linii. Převzato z [6].

minimalizují vzdálenost mezi obrazy obličejů stejné třídy. Jinými slovy se jedná o maximalizaci mezitřídní rozptylové matice S_b a zároveň minimalizaci vnitřní rozptýlené matice S_w v promítnutém podprostoru. Obrázek 6.4 ukazuje dobré a špatné rozdělení tříd.



Obrázek 6.4: (a) Dobré rozdělení tříd. (b) Špatné rozdělení tříd. Převzato z [6].

Mezitřídní rozptýlená matice S_b je definována vztahem

$$S_b = \sum_{i=1}^D N_i (\mu_i - \mu)(\mu_i - \mu)^T, \quad (6.8)$$

kde N_i je počet trénovacích vzorků třídy i , μ_i je střední hodnota třídy i a μ představuje střední hodnotu všech tříd.

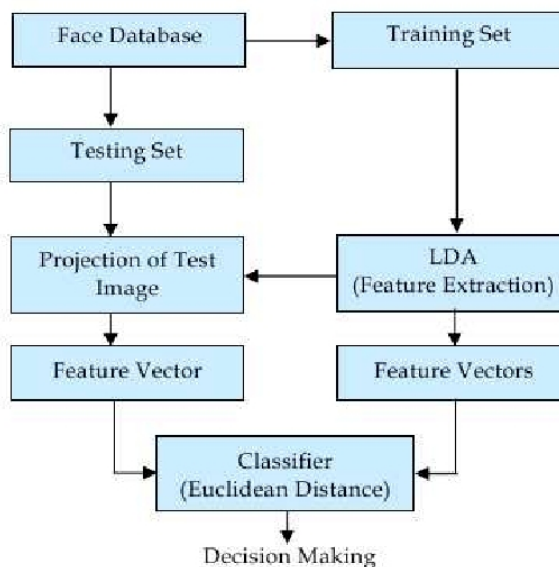
Vnitřní rozptylová matice S_w je definována jako

$$S_w = \sum_{i=1}^D \sum_{x_k \in X_i} (x_k - \mu_i)(x_k - \mu_i)^T, \quad (6.9)$$

kde x_k je i tý vzorek třídy i , μ_i je střední hodnota třídy i , D je počet tříd a X_i obsahuje vzorky patřící do třídy i .

Podprostor LDA zahrnuje množinou vektorů $W = [W_1, W_2, \dots, W_i]$

$$W_{arg \max(W_i)} = \left| \frac{W^T S_b W}{W^T S_w W} \right|. \quad (6.10)$$



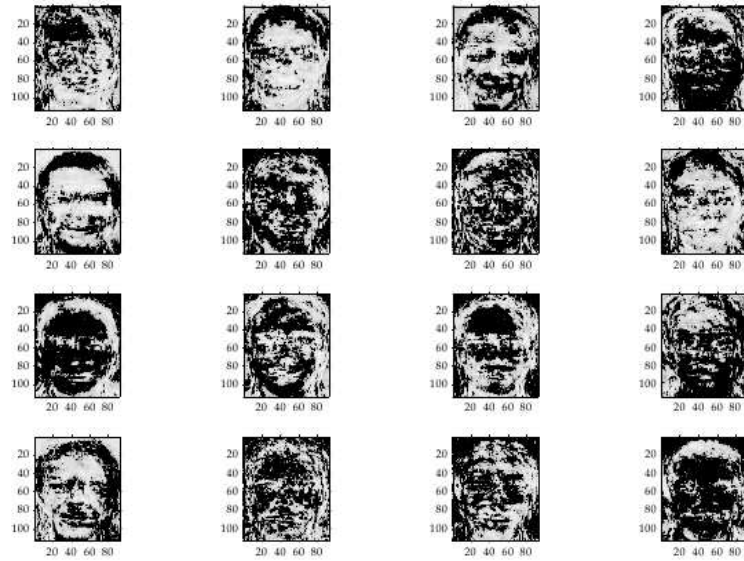
Obrázek 6.5: LDA přístup k rozpoznání obličeje. Převzato z [6].

Vnitřní rozptylová matice reprezentuje, jak jsou obrazy tváří těsně rozděleny uvnitř tříd, a mezitřídní rozptylová matice popisuje, jak jsou třídy odděleny od sebe navzájem. Jsou-li obrazy tváří promítány na diskriminační vektory W , obrazy tváří mohou být rozděleny těsněji do vnitřních tříd a více odděleny ve vnějších třídách tak, jak je to jen možné. Jinými slovy tyto diskriminační vektory minimalizují jmenovatele a maximalizují čitatele v rovnici (6.10), z níž mohou být sestaveny vlastní vektory $S_W^{-1}S_b$. Obrázek 6.6 ukazuje prvních 16 vlastních vektorů s nejvyšší přidruženou vlastní hodnotou $S_W^{-1}S_b$. Tyto vlastní vektory jsou také známy jako Fisherovy obličeje. Existují různé metody řešící problémy LDA, například pseudoinverzní metoda, podprostorová metoda nebo metoda nulového prostoru [6].

LDA přístup je podobný jako u metody vlastních vektorů, která využívá projekce trénovacích obrazů do podprostoru. Testovací obrazy jsou promítány do stejného podprostoru a rozpoznány pomocí porovnávacích metrik. Jediným rozdílem mezi metodami je počítání podprostoru charakterizujícího prostor obličeje. Obličej, který má nejmenší vzdálenost od testovacího obrazu obličeje, je vybrán jako identita obrazu. Nejmenší vzdálenost může být spočítána pomocí Euklidovy vzdálenosti. Obrázek 6.5 zobrazuje testovací fázi v LDA přístupu k rozpoznání obličeje.

6.1.4 Komponentní lineární diskriminační analýza

Přestože reprezentace obrazových dat obličeje poskytovaná metodou LDA popsanou v 6.1.3 má mnoho užitečných vlastností, nedokáže se příliš dobře vyrovnat s různým osvětlením představujícím změny, které nebyly dobře zachyceny v trénovací sadě obrazových dat. Proto byla vyvinuta, s cílem posílit obecné schopnosti LDA, které nejsou zastoupeny v trénovací



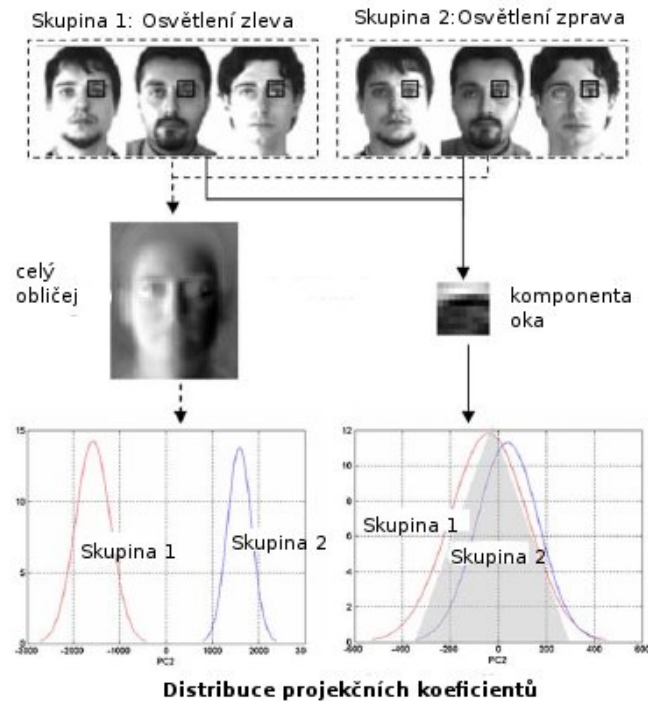
Obrázek 6.6: Prvních 16 Fisherových obličejů s nejvyššími vlastními hodnotami. Číselné hodnoty reprezentují souřadnice na ose X a Y . Převzato z [6].

sadě, komponentní LDA [28], kdy je tvář rozdělena na několik obličejových oblastí a pro každou tuto oblast je použita metoda LDA. Důvodem pro vznik komponentní LDA bylo, že vzhledem k 3D charakteru lidské tváře, budou mít změny osvětlení za následek globální nelineární změny v intenzitě obrazu. Nicméně místní účinek osvětlení může být aproximován lineární funkcí, což je mnohem snadnější než aproximovat nelineárně. Vliv osvětlení na změny v obraze je možné vidět na obrázku 6.7.

Reprezentace komponentní lineární diskriminační analýzy

Obraz může být rozdělen na komponenty LDA dvěma způsoby, jak je uvedeno na obrázku 6.8. Pozice a velikost komponent jsou relativní k umístění očí. První využívá 14 malých či větších komponent, které přesahují přes více menších. Menší komponenty, které jsou umístěny na fyzicky významných místech jako jsou oči, nos či ústa, jsou určeny k účinné reprezentaci místních statistik pro klasifikaci obličejů. Větší komponenty kolem hlavy, tváře a krku představují další důležité informace o obličejí. Přesah mezi sousedními komponentami podporuje zachování vazeb přilehlých komponent.

Další způsob komponent LDA je podobný prvnímu, avšak využívá menší počet komponent. A to tak, že jsou spojeny dvě či tři komponenty v jednu. Tímto způsobem je počet komponent snížen na nejvýše pět důkladným hledáním nejlepší kombinace na základě provedení vyhledávání na trénovací sadě. Výsledkem toho je, že přesnost vyhledávání ve srovnání s původním 14komponentním rozdělením není ovlivněna. Navíc popisovaná velikost a výpočetní složitost jsou podstatně sníženy.



Obrázek 6.7: Vliv osvětlení na LDA a komponentní LDA
Převzato a upraveno z [28].

Pokud je dána sada N trénovacích obrázků $\{x_1, x_2, \dots, x_N\}$, je možné získat sadu LDA transformačních matic.

Všechny obrázky jsou rozděleny na L složek obličeje. Obrazová data komponent jsou reprezentována pomocí vektoru. k tá komponenta bude výjádřena jako $\{c_1^k, \dots, c_N^k\}$. Pak je pro k tou obličejovou komponentu vypočítána odpovídající LDA transformační matice W^k .

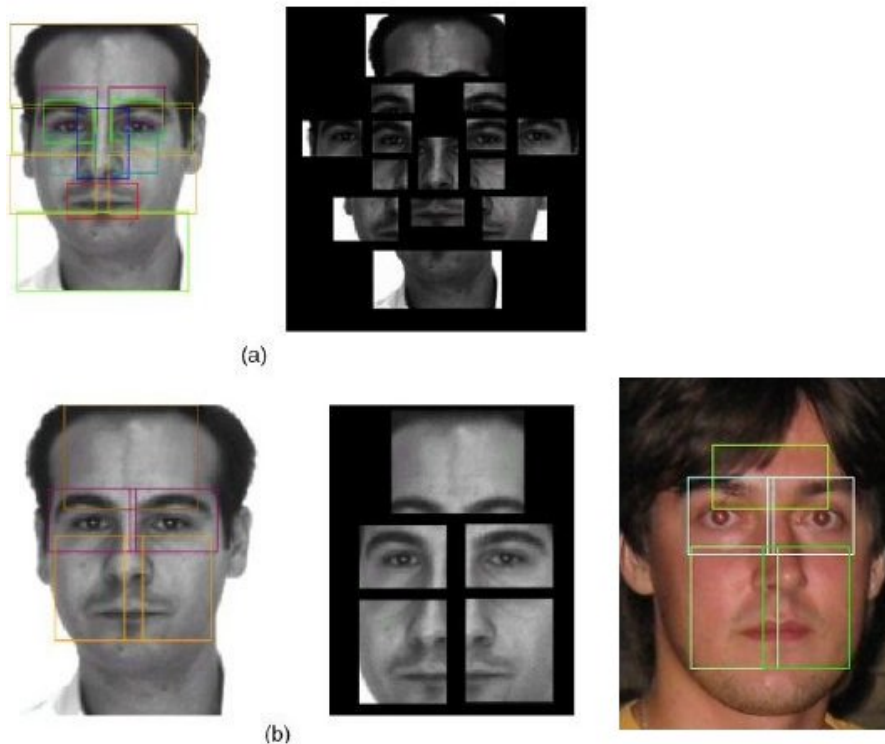
Během testování je získáno L vektorů $\{c^1, \dots, c^L\}$ odpovídajících obličejových komponent z obrázku obličeje x z testovací sady. Množina LDA příznakových vektorů $y = \{y^1, \dots, y^L\}$ je získána transformováním komponentních vektorů odpovídající LDA transformační maticí jako

$$y = (W^k)^T c^k, \quad k = 1, 2, \dots, L. \quad (6.11)$$

Poté je obrázek obličeje x reprezentován množinou LDA příznakových vektorů $\{y^1, \dots, y^L\}$.

6.2 Jádrové metody

Různé obrázky obličejů v podprostoru nemusí být lineárně separovatelné. Jádrové metody jsou zobecněním lineárních metod. Z důvodu výskytu nelinearity se stále více pozornosti výzkumu věnuje nelineárním metodám, protože mají co nabídnout z praktického hlediska.



Obrázek 6.8: Definice komponent LDA
a) 14 komponent b) 5 komponent. Převzato z [28].

Některá běžně používaná jádra pro linearizaci nelineárních dat jsou:

- Gaussovo: $\mathcal{K}(\vec{x}, \vec{y}) = \exp(-\frac{\|\vec{x}-\vec{y}\|}{2\sigma^2})$
- Jádra skalárního součinu:
 - Sigmoid: $\mathcal{K}(\vec{x}, \vec{y}) = \tanh(\alpha x^T y + c)$
 - Polynomiální: $\mathcal{K}(\vec{x}, \vec{y}) = (\vec{x} \cdot \vec{y} + c)^d$

Gaussovo jádro má jeden parametr σ , který je velice důležitý a měl by být pro daný problém pečlivě vyladěn. Pokud bude nadhodnocen, jádro se bude místo exponenciálního chování chovat téměř lineárně a vyšší dimenze projekce začnou ztrácet výhodu nelinearity. Na druhou stranu pokud bude podceněn, funkce bude postrádat jakákoli pravidla a rozhodovací hranice bude velmi citlivá na zašuměná trénovací data. U jádra sigmoid jsou dva nastavitelné parametry α a c , kde $\alpha = \frac{1}{N}$, N je dimenze dat. Polynomiální jádro má nastavitelné parametry c a d , c je konstanta a d je stupeň polynomu.

Ještě poznamenejme, že každý algoritmus, který může být vyjádřen pouze ve tvaru skalárního součinu, může být implementován v prostoru příznaků \mathcal{F} . Uplatnění klasifikátoru nadroviny nebo lineární funkce LDA rozdělení vede k obvyklým SVM klasifikátorům. Použití lineární PCA v prostoru příznaků vede k metodě KPCA ve vstupním prostoru.

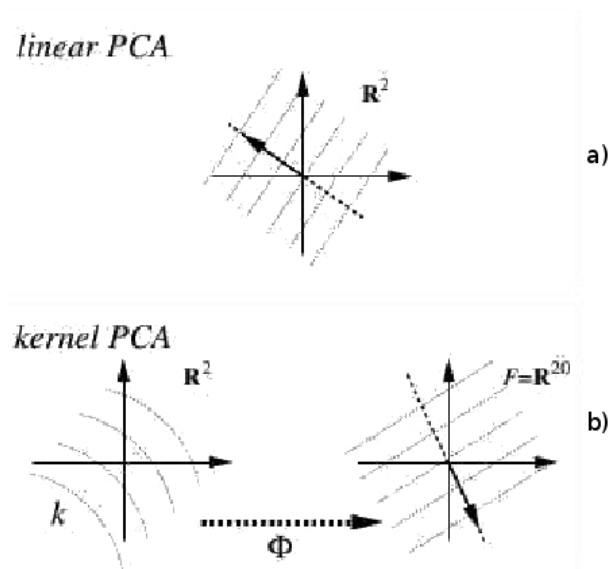
6.2.1 Jádrová analýza hlavních komponent

Metoda KPCA (Kernel Principal Component Analysis) popsaná v [29] je nelineární metoda extrakce příznaků, která úzce souvisí s metodami používanými v Support vector machine (SVM). Předpokládejme, že máme množinu vstupních dat $\{\vec{x}_i \in \mathbb{R}^n : i = 1 - N\}$, kde rozložení dat je nelineární. Jedním způsobem, jak řešit takové rozložení, je linearizace nelineárního mapování dat ze vstupního prostoru $\vec{x} \in \mathbb{R}^n$ do nového prostoru příznaků $\Phi(\vec{x}) \in \mathcal{F}$. Metody KPCA a SVM používají toto mapování. Zobrazení Φ je definováno implicitně určením tvaru skalárního součinu v prostoru příznaků. Takže pro libovolné dvojice mapovaných bodů dat je skalární součin definován v podmínkách funkce jádra takto:

$$\Phi(\vec{x}) \bullet \Phi(\vec{y}) \equiv \mathcal{K}(\vec{x}, \vec{y}) . \quad (6.12)$$

Může se použít Gaussovo jádro či sigmoid a polynomiální jádro.

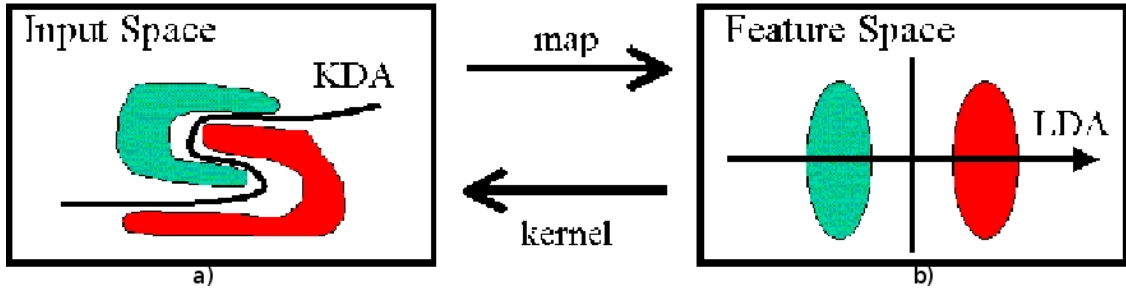
Základní myšlenka KPCA je zobrazena na obrázku 6.9. V některých vysokorozměrných prostorech příznaků \mathcal{F} (obrázek 6.9 b)) je provedena lineární PCA pouze jako PCA ve vstupním prostoru (obrázek 6.9 a)). Jelikož \mathcal{F} je nelineárně spojeno se vstupním prostorem (pomocí Φ), kontura čar z neměnné projekce na základní vlastní vektor (zakresleno jako šipka) se stává nelineární ve vstupním prostoru. Je možné si všimnout, že nelze zakreslit předešlý obrázek vlastních vektorů do vstupního prostoru, protože nemusí existovat. Rozhodující pro KPCA je fakt, že vlastně neprobíhá mapování do \mathcal{F} , ale namísto toho jsou prováděny všechny nezbytné výpočty za použití jádrové funkce \mathcal{K} ve vstupním prostoru (zde \mathbb{R}^2).



Obrázek 6.9: Základní myšlenka KPCA. Převzato z [21].

6.2.2 Jádrová diskriminační analýza

Princip metody KDA (Kernel Discriminant Analysis) popsaný v [14] může být ilustrován na obrázku 6.10. Vzhledem k velké nelinearitě je obtížné přímo spočítat diskriminační příznaky mezi dvěma třídami vzorů v původním vstupním prostoru (obrázek 6.10 a)). Z určení nelineárního mapování ze vstupního prostoru na vysokodimenzionální prostor příznaků (obrázek 6.10 b)) je očekáván zisk lineární rozdělující distribuce v prostoru příznaků. Pak LDA, která je lineární, může být použita v příznakovém prostoru k extrakci důležitějších diskriminačních příznaků. Spočítání může být problematické nebo dokonce nemožné v prostoru příznaků v důsledku velkého rozměru. Zavedením jádrové metody odpovídající nelineárnímu mapování lze všechny výpočty pohodlně přenést mimo vstupní prostor. Problém může být nakonec řešen jako vlastní dekompoziční problém pomocí PCA, LDA a KPCA. Podobně jako v podkapitole 6.1.4 může existovat i komponentní jádrová diskriminační analýza, kdy se na linearizaci a redukci prostoru každé komponenty použije jádrová metoda a problém je nakonec řešen pomocí LDA.



Obrázek 6.10: Jádrová diskriminační analýza. Převzato z [14].

6.3 Elastické shlukové grafy

Všechny lidské tváře sdílejí podobnou topologickou strukturu. V elastických shlukových grafech (Elastic Bunch Graph Matching) [31] jsou tváře reprezentovány jako grafy, kde jsou uzly umístěny ve výchozích bodech (obrázek 6.11). Hrany nesou informaci o vzdálenosti a uzly jsou označeny wavelet koeficienty daného pixelu. Tento model grafu lze poté použít pro generování obrazového grafu. Model může být posunut, otočen, deformován při porovnávání nebo se může měnit měřítko. Díky tomu tato metoda funguje, i v případě, že jsou velké rozdíly v obraze. Například vlivem osvětlení či jiného výrazu tváře.

Jety popisují malé kousky hodnot v hodnotách intenzity v obraze $\mathcal{I}(\vec{x})$ v okolí daného bodu. Jsou založeny na vlnkové transformaci, která je definována jako konvoluce

$$\mathcal{J}_j(\vec{x}) = \int \mathcal{I}(\vec{x}) \psi_j(\vec{x} - \vec{x}') d^2 \vec{x}'. \quad (6.13)$$

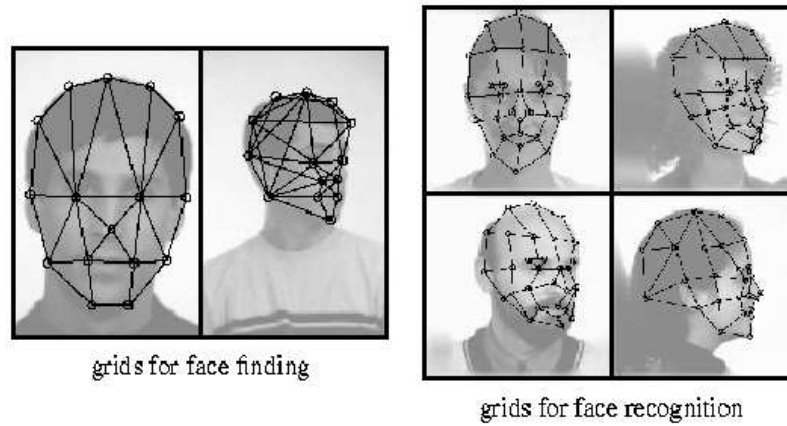
Jety jsou určeny Gaborovými jádry. Jedno z nich může vypadat například takto:

$$\psi_j(\vec{x}) = \frac{k_j^2}{\sigma^2} \exp\left(-\frac{k_j^2 x^2}{2\sigma^2}\right) \left[\exp\left(i\vec{k}_j \vec{x}\right) - \exp\left(-\frac{\sigma^2}{2}\right) \right], \quad (6.14)$$

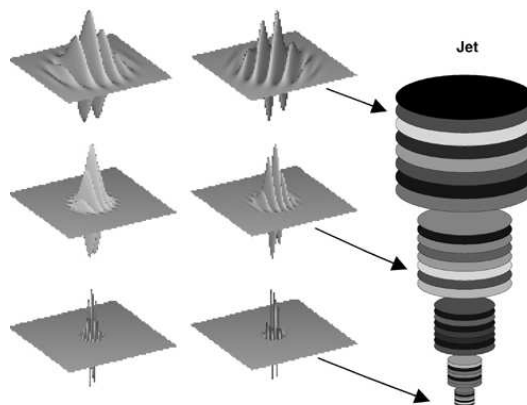
ve tvaru rovinných vln s vlnovým vektorem \vec{k}_j omezeném Gaussovou obalovací funkcí. Je využita diskrétní sada 5 frekvencí s indexy $v = 0, \dots, 4$ a 8 orientací s indexy $\mu = 0, \dots, 7$,

$$\vec{k}_j = \begin{pmatrix} k_{jx} \\ k_{jy} \end{pmatrix} = \begin{pmatrix} k_v \cos \varphi_\mu \\ k_v \sin \varphi_\mu \end{pmatrix}, \quad k_v = 2^{-\frac{v+2}{2}} \pi, \quad \varphi_\mu = \mu \frac{\pi}{8}. \quad (6.15)$$

Na obrázku 6.12 jsou zobrazeny příklady Gaborovy vlnkové transformace o různé velikosti a prostorové orientaci, která se používá k analýze digitálních fotografií. Pro všechny body jsou výsledky velikosti a orientace uspořádány do vektoru rysů zvaného jet.



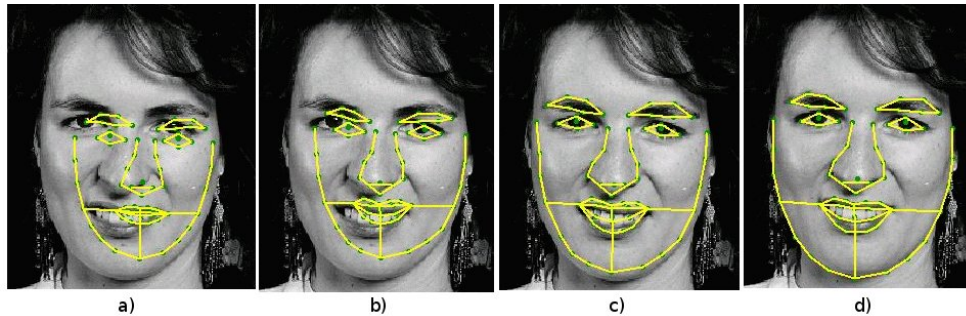
Obrázek 6.11: Elastický srovnávací diagram aplikován na obrázek tváře. Převzato z [31].



Obrázek 6.12: Příklady Gaborovy vlnkové transformace. Převzato z [16].

6.4 Aktivní vzhledový model

Aktivní vzhledový model (Active Appearance Models) [3] je statistický model, který kombinuje model kombinací tvarových vlastností s modelem vzhledových změn normalizovaného tvaru obličeje. AMM obsahuje statistický model, pokud existuje zobrazení tvaru a hodnot intenzity objektu zájmu, který se může zobecnit na téměř jakýkoliv validní vzor. Odpovídající obraz vyžaduje nalezení parametrů modelu s nejmenší odchylkou mezi obrazem a modelovým vzorem zobrazeným do obrazu, viz obrázek 6.13.



Obrázek 6.13: Active Appearance Models. a) inicializace b) pozice po první iteraci c) pozice po druhé iteraci d) konvergováno. Převzato z [3].

6.5 Shrnutí

V této kapitole byly popsány jednotlivé přístupy k identifikaci obličeje. Byly zmíněny metody lineární analýzy i jejich nelineární rozšíření ve formě jádrových metod, které se jeví jako velice účinné a efektivní a je na ně soustředěna stále větší pozornost jako na metody velmi perspektivní. Dále se kapitola zabývala metodou elastických shlukových grafů založenou na topologické struktuře tváře reprezentovanou grafy a statistickým modelem aktivního vzhledového modelu, který kombinuje model tvarových vlastností obličeje s modelem vzhledových změn obličeje, jehož tvar je normalizován.

Kapitola 7

Implementace a validace

Tato kapitola se zabývá implementací a validací aplikace, konkrétně volbou vhodné reprezentace obrazových dat a výběrem metod pro detekci dominantních rysů a identifikaci obličejů. Dále popisuje proces rozpoznání dominantních rysů a identifikace. Kapitola se také věnuje validaci aplikace nad zadanými testovacími daty.

7.1 Reprezentace dat

Reprezentace obrazových dat může být vyjádřena pomocí bodů ve vysokorozměrném prostoru [4]. Souřadnice takových bodů mohou být také vyjádřeny pomocí vektoru. Například vektor $X_i = \{x_1, x_2, x_3, \dots, x_{N-1}, x_N\}$ reprezentuje matici $1 \times N$, kde x_i může být bodem ve vysokorozměrném prostoru s velikostí $n = X \times Y$ reprezentujícím obrázek o rozměrech $X \times Y$ nebo vektorem dimenze n . Hodnota n je počet všech pixelů v obraze, N udává počet všech obrázků ke zpracování. Pro další zpracování je vhodné použít metodu pro redukci vysokorozměrných dat, například jádrovou diskriminační analýzu nebo jádrovou analýzu hlavních komponent.

7.2 Načtení informací o obrázcích v databázi

Data ze zadaného adresáře jsou načtena do seznamu objektů typu `FaceFeatures`, který obsahuje informace o cestě k obrázku, příslušnost ke třídě obličejů, v případě trénovacích dat a ukončené identifikace také predikovanou příslušnost ke třídě obličejů. Dále zde mohou být nastaveny souřadnice detekovaných dominantních rysů a další různé podpůrné informace.

7.3 Rozpoznání dominantních rysů

Aplikace pro rozpoznání dominantních rysů obličejů využívá různé metody. Pro lokalizaci očí v zadané oblasti obličejů je použita metoda založená na detekci kružnic pomocí Houghovy transformace popsaná v podkapitole 4.1.3, kdy je aplikován hranový detektor Laplace of

Gaussian (LoG)¹ pouze na oblast obličeje a ostatní části obrázku jsou vymaskovány na nulovou hodnotu intenzity. LoG byl zvolen z důvodu dobré odolnosti proti šumu. Třídy a metody implementující rozpoznání dominantních rysů se nachází v balíku `recognition`.

Poté je na obraz aplikována Houghova transformace pro detekci kružnic nalézající se v balíku `dip` ve třídě `HoughTransf`. V tomto případě se hledají kružnice o zadaných průměrech. Po nalezení kružnic je provedena jednoduchá filtrace kružnic, které určitě nemohou být oblastmi očí. Filtrace probíhá tak, že pokud kružnice nemá do páru druhou kružnici reprezentující druhé oko v určité pozici vůči první, tak je odstraněna. Zbývající kružnice jsou prohlášeny za kandidátské oblasti očí a jsou rozděleny do seznamu všech možných dvojic reprezentujících kandidátské oblasti levého a pravého oka.

V dalším kroku je hledáno obočí. Pro jeho lokalizaci je použita metoda založená na šablonovém přístupu popsaná v podkapitole 4.4. Pokud není nalezeno obočí u obou oblastí očí, je tato kandidátská oblast vyřazena ze seznamu kandidátských oblastí očí.

Posledním detekovaným rysem jsou ústa, kdy je na zbývající kandidátské oblasti aplikována metoda popsaná v podkapitole 4.2.1 využívající šablonového přístupu k rozpoznání. Pokud nejsou nalezena ústa, je dvojice kandidátských oblastí očí vymazána ze seznamu.

V tomto kroku již máme seznam kandidátských oblastí, které splňují všechny požadavky, aby mohly být prohlášeny za oblasti očí. Problém nastává, pokud je těchto oblastí mnoho a musí být vybrána pouze jedna jediná. V tomto případě je za oblast očí prohlášena ta, která má nejvyšší součet hodnot akumulátoru Houghovy transformace pro potenciální oblast levého a pravého oka, a tudíž by měla být nejideálnější kružnicí a tím i oblastí oka. Pokud má více oblastí shodnou hodnotu akumulátoru, je vybrána oblast s nejmenším rozdílem souřadnic y .

Souřadnice detekovaných rysů jsou následně nastaveny každému obličeji v seznamu objektů typu `FaceFeatures` pro jejich pozdější využití při procesu identifikace.

7.4 Identifikace

Pro proces identifikace byla vybrána lineární diskriminační analýza. Jelikož prostor obličejů není lineárně separovatelný bylo nutné tento prostor linearizovat a k tomuto účelu byla použita nelineární jádrová diskriminační analýza. Tato metoda zlinearizuje nelineární vstupní prostor tak, aby bylo možno dále pracovat s klasickou lineární metodou LDA. Mimo jiné KDA redukuje rozměr transformovaného vstupního prostoru.

V aplikaci je implementována metoda KDA ve třídě `Kda` a vlastní dekompoziční problém je již řešen pomocí LDA za pomoci knihovny `JavaStat`² ve třídě `LinearDiscriminantAnalysis`. Všechny třídy zabývající se identifikací jsou dostupné v balíku `kda`.

Samotný proces identifikace v prvním kroku, který probíhá již při načítání obrázků do

¹Dostupný online na adrese <http://homepages.inf.ed.ac.uk/rbf/HIPR2/flatjavasrc/Log.java>

²Dostupný online na adrese http://www2.thu.edu.tw/~wenwei/javastat/javastat_beta1.4.zip

seznamu, rozdělí všechny obrázky obličejů v trénovací sadě do tříd, kdy obrázky ve stejných adresářích patří do stejné třídy. Po rozdělení trénovací sady do tříd je rozdělena i testovací sada, kdy testovací vzorek patří do stejné třídy jako trénovací pokud název adresáře ve kterém jsou uloženy je stejný.

Druhým krokem je dříve popsaná detekce dominantních rysů a krokem následujícím je normalizace všech detekovaných oblastí obličejů na stejnou velikost a následné provedení procesu identifikace. Nejdříve musí být data upravena na vektorovou reprezentaci podle podkapitoly 7.1. Poté jsou trénovací data předána k redukci a linearizaci prostoru metodě KDA. Výstupem metody KDA jsou již lineárně separovatelná data, která jsou předána ke zpracování metodě LDA. Teď je možné přistoupit k procesu identifikace na testovacích datech, která jsou nejdříve předána metodě KDA k úpravě na lineárně separovatelná data. Tato úprava využívá stejné parametry jako při linearizaci trénovacích dat spolu s vlastními vektory spočtenými nad trénovacími daty. Upravená data jsou předána k identifikaci metodě LDA, která má již spočteny diskriminační vektory nad trénovací sadou, a teď bude pouze určovat příslušnost jednotlivých testovacích vzorků k té které třídě.

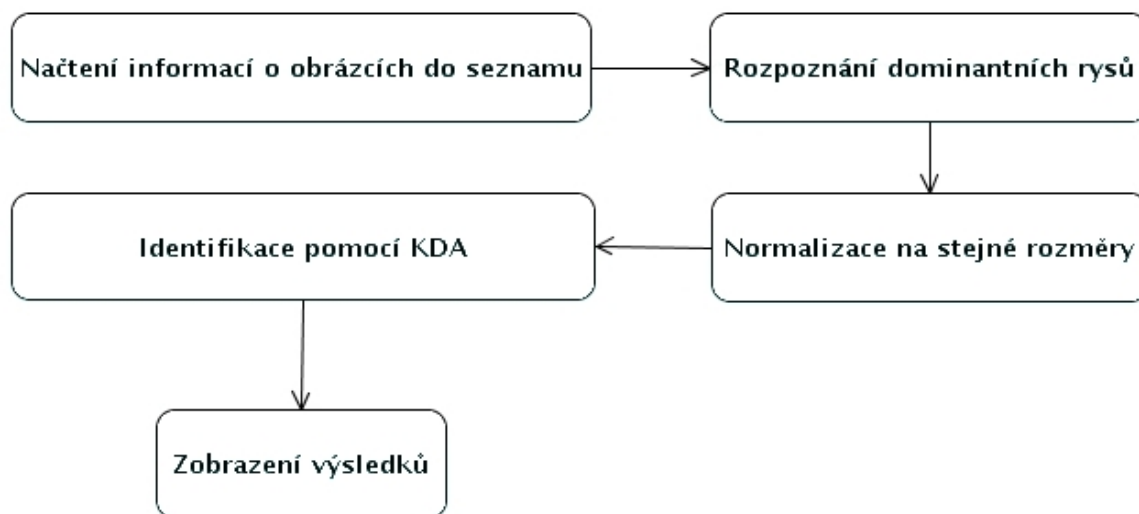
Posledním krokem je zobrazení výsledků do přehledné tabulky a výpočet výsledné úspěšnosti rozpoznání nad zadanými daty.

Obdobně by se postupovalo při identifikaci pomocí komponentní LDA s tím rozdílem, že by proces identifikace byl aplikován pro každou komponentu samostatně a výsledná příslušnost k dané třídě by byla rozhodnuta podle třídy, která by byla určena pro jednotlivé komponenty nejčastěji. V případě shody počtu příslušnosti do dvou a více tříd je vybrána ta, která má dle příslušných komponent vyšší váhu. Komponenty od nejnižší váhy jsou seřazeny takto – čelo, pravé oko, levé oko, pravá část úst, levá část úst.

Cílem této práce je vytvořit nástroj pro nalezení dominantních rysů obličeje a následnou identifikaci osob, jehož princip činnosti je zobrazen na obrázku 7.1 a na obrázku C.1 je zobrazen diagram tříd potřebných k nalezení dominantních rysů a následné identifikaci obličeje. Pro testovací účely jsou použity dvě různé databáze obličejů na kterých bude demonstrována úspěšnost metody. Nejdříve je nutné rozpoznat dominantní rysy obličeje, což je velice důležitý krok, jelikož špatně lokalizované dominantní rysy mohou negativně ovlivnit proces identifikace. Nástroj obsahuje mimo jiné možnost provést identifikaci přímo nad upravenými a normalizovanými vstupními daty a tímto přeskočit proces rozpoznání dominantních rysů a normalizaci. Tato možnost může být velice vhodná, pokud jsou data předpřipravena jiným nástrojem. Nejdříve bude testována úspěšnost detekce dominantních rysů na různých vzorcích dat. Dále bude nad těmito daty proveden proces identifikace. Data pro všechny testy jsou přítomna na příloženém DVD A.

7.5 Umístění obličeje

Nástroj neobsahuje vlastní detektor pro lokalizaci obličeje, proto je nutné u každého souboru s obličejem dodat speciální soubor obsahující souřadnice oblasti obličeje. Tento soubor je



Obrázek 7.1: Schéma aplikace.

pojmenován jako `obrazek.face` a kód určující umístění obličeje je následující

```
#komentář
x1 y1 x2 y2
```

x_1, y_1 odpovídá levému hornímu rohu a x_2, y_2 pravému dolnímu rohu obdélníkové oblasti obličeje.

Oblast obličeje je možné získat pomocí různých nástrojů, například nástrojem Detektor obličejů, akcelerovaný pomocí SIMD³. Pro testovací účely byla zvolena možnost, že pro obrázky do maximálního rozlišení 320×240 včetně bude oblast obličeje očekávána mezi body $[x_1, y_1], [x_2, y_1], [x_1, y_2], [x_2, y_2]$, kde

$$x_1 = \text{sirka_obrazku}/5, \quad (7.1)$$

$$x_2 = \text{sirka_obrazku} - x_1, \quad (7.2)$$

$$y_1 = \text{vyska_obrazku}/6, \quad (7.3)$$

$$y_2 = \text{vyska_obrazku} - y_1. \quad (7.4)$$

Pro testovací účely budou použity také databáze obsahující vzorky dat o rozlišení 720×576 a 640×480 . U obrázků o těchto rozlišeních je oblast obličeje očekávána mezi body $[x_1, y_1], [x_2, y_1], [x_1, y_2], [x_2, y_2]$, kde

$$x_1 = \text{sirka_obrazku}/4, \quad (7.5)$$

³dostupný na FIT VUT na adrese <http://www.fit.vutbr.cz/research/prod/index.php.cs?id=68¬itle=1>

$$x_2 = sirka_obrazku - x_1, \quad (7.6)$$

$$y_1 = vyska_obrazku/4, \quad (7.7)$$

$$y_2 = (vyska_obrazku - y_1) * 0.6. \quad (7.8)$$

Samozřejmě i u těchto dat může být ve speciálním souboru `obrazek.face` určena přesná oblast obličeje. Výše uvedené rovnice pro zjištění umístění obličeje v obraze byly zjištěny experimentálně nad testovacími daty.

7.6 Testy

Pro testovací účely byly použity dvě databáze obličejů, první z nich je IMM Face Database [27] obsahující 240 obrázků obličejů 40 různých osob. Druhou databází je databáze XM2VTSDB [17] obsahující 295 různých osob, která od každé osoby obsahuje osm vzorků. Obě databáze obsahují vzorky dat různě osvětlené a natočené pod různým úhlem – maximálně však o 45°.

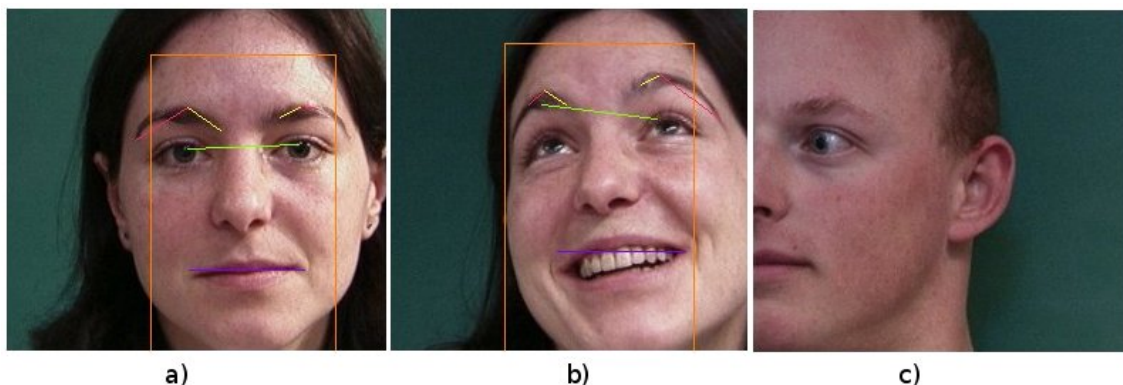
7.6.1 IMM Face Database

Detekce dominantních rysů

Byly provedeny celkem čtyři testy detekce dominantních rysů nad databází IMM. První test probíhal nad celou databází bez dodatečných úprav testovacích dat. Výsledek je zobrazen v tabulce 7.1 ve sloupci **test 1**. V tomto testu nebyly rysy vůbec detekovány u jednoho obličeje z důvodu příliš velkého náklonu, u dalších tří z důvodu umístění obličeje mimo předpokládanou oblast. Zcela chybně byly rysy detekovány u jednoho obličeje z důvodu umístění mimo předpokládanou oblast a na dalších pěti bylo detekováno správně jedno oko a druhé mělo větší či menší odchylku od správného umístění. Příklady lokalizace dominantních rysů jsou zobrazeny na obrázku 7.2.

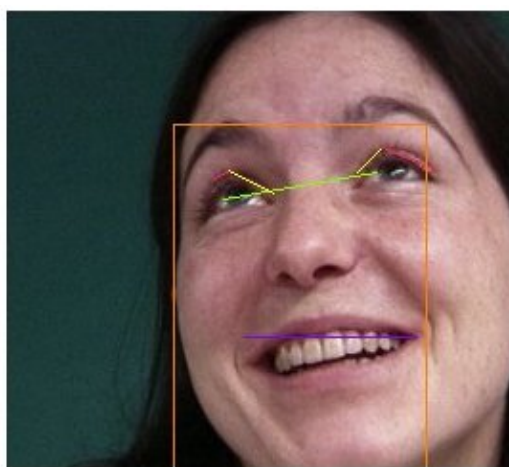
Další test byl proveden nad stejnou datovou sadou jako test první, s tím rozdílem, že obličeje, které byly mimo předpokládané umístění, byly doplněny o pomocný soubor `obrazek.face` obsahující přesné souřadnice obličeje. V tomto případě byly detekovány dominantní rysy správně na třech ze čtyř vzorků obsahujících pomocný soubor s umístěním obličeje a detekce rysů u ostatních vzorků zůstala neměnná. Výsledek testu je možné vidět v tabulce 7.1 ve sloupci **test 2**.

Aplikace umožňuje i ruční volbu velikosti oka a následnou detekci s touto velikostí. Ruční volba průměru kružnic reprezentující oči je v tomto testu aplikována na všechny vzorky obličejů u kterých byly chybně detekovány dominantní rysy. Vzorky dat obsahující obličeje mimo předpokládanou oblast jsou v tomto testu přítomny s pomocným souborem `obrazek.face`. Na ostatní data je použito automatické rozpoznání jako v předešlých testech. Výsledek testu je zobrazen v tabulce 7.1 ve sloupci **test 3**. Důvodem zvýšení úspěšnosti při ručním zadání velikosti očí je, že oblast očí je z kandidátských oblastí očí vybrána chybně.



Obrázek 7.2: Lokalizace dominantních rysů. a) nalezení správných dominantních rysů, b) nalezení správně pouze části dominantních rysů, c) nenalezení rysů z důvodu špatného umístění obličeje v obraze. Upraveno z [27].

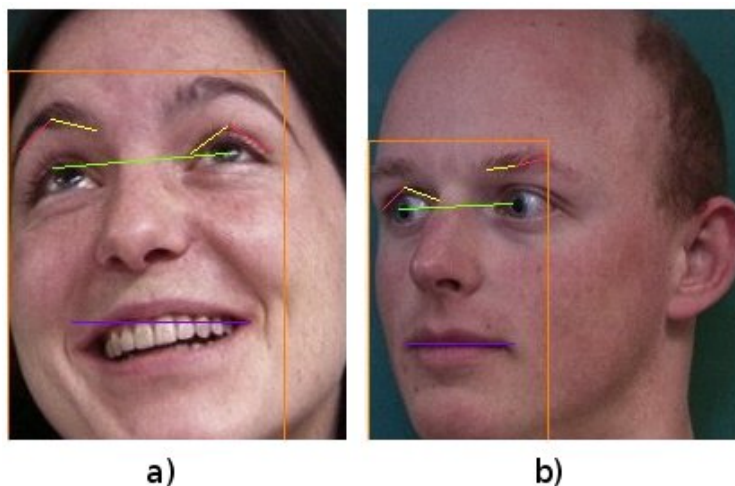
Způsobů, jak zvýšit úspěšnost správného výběru oblasti očí z kandidátských oblastí, existuje více a bude se jimi zabývat kapitola pojednávající o budoucím rozšíření a vývoji aplikace. Výsledek ručního zadání rozměru je možné vidět na obrázku 7.3 kdy ruční rozpoznání bylo aplikováno na obrázek 7.2 b).



Obrázek 7.3: Nalezení správných dominantních rysů při zadání velikosti kružnic reprezentující oči. Upraveno z [27].

Poslední výsledek testu zobrazený v tabulce 7.1 ve sloupci **test 4** je test, kdy ke všem obrázkům tváří, kde byly chbně automaticky rozpoznány dominantní rysy obličeje je přidán soubor **obrázek.face** se souřadnicemi přesného umístění obličeje. Je možné si všimnout, že oproti **testu 1** se zvýšila úspěšnost správného rozpoznání dominantních rysů a jen minimálně zaostává za výsledkem **testu 3**. Pokud by bylo aplikováno ruční zadání velikosti

kružnic na zbývající špatně rozpoznané rysy, mohla by být očekávána výsledná úspěšnost podobná **testu 3**. Relativně vysoká úspěšnost **testu 4** je způsobena nižším vlivem rušivého pozadí, které mělo negativní vliv na detekci hran pomocí hranového detektoru Laplace of Gaussian. Výsledné detekce jsou zobrazeny na obrázku 7.4, kdy `obrázek.face` byl přidán k obrázkům 7.2 b) a 7.2 c).



Obrázek 7.4: Detekce dominantních rysů u souborů s přidáním souborem `obrázek.face` obsahující souřadnice obličeje. Upraveno z [27].

Úspěšnost rozpoznání dominantních rysů				
lokalizace	test 1	test 2	test 3	test 4
očí	95,83 %	97,08 %	99,16 %	98,33 %
úst	97,08 %	98,33 %	99,16 %	98,75 %

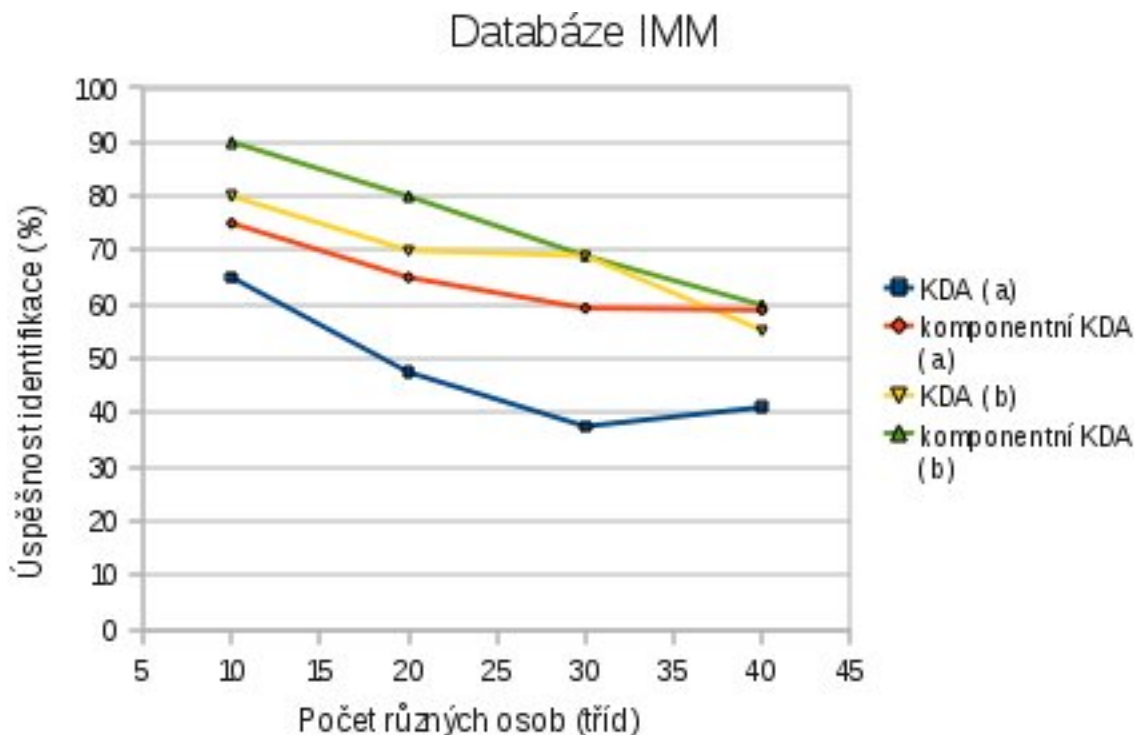
Tabulka 7.1: Úspěšnost lokalizace dominantních rysů nad daty databáze IMM

Identifikace obličeje

Identifikace obličeje nad databází IMM je provedena nad dvěma testovacími sadami. První testovací sada obsahuje čtyři vzorky trénovacích dat a dva vzorky dat testovacích. Druhá testovací sada pak obsahuje pět trénovacích vzorků a jeden vzorek testovací. Identifikace byla postupně provedena pro 10, 20, 30 a 40 různých osob (tříd) pomocí KDA a komponentní KDA, kde je vlastní dekompoziční problém řešen pomocí metody LDA. Parametry pro testy nad první testovací sadou jsou zobrazeny v tabulce B.1 a nad druhou testovací sadou v tabulce B.2.

Výsledky testů jsou zobrazeny v grafu 7.5, kde KDA (a) a komponentní KDA (a) jsou výsledné úspěšnosti první testovací sady a KDA (b) a komponentní KDA (b) jsou výsledné

úspěšnosti identifikace nad daty druhé testovací sady pomocí metod KDA a komponentní KDA s využitím metody LDA pro řešení vlastního dekompozičního problému. Z grafu je patrné, že druhá testovací sada má na shodném počtu tříd obličejů o něco vyšší úspěšnost rozpoznání, což je s největší pravděpodobností způsobeno větším počtem trénovacích dat pro metody KDA a LDA, kdy rozdělení do tříd lépe reprezentuje stejné tváře stejné osoby v prostoru.



Obrázek 7.5: Identifikace nad daty databáze IMM.

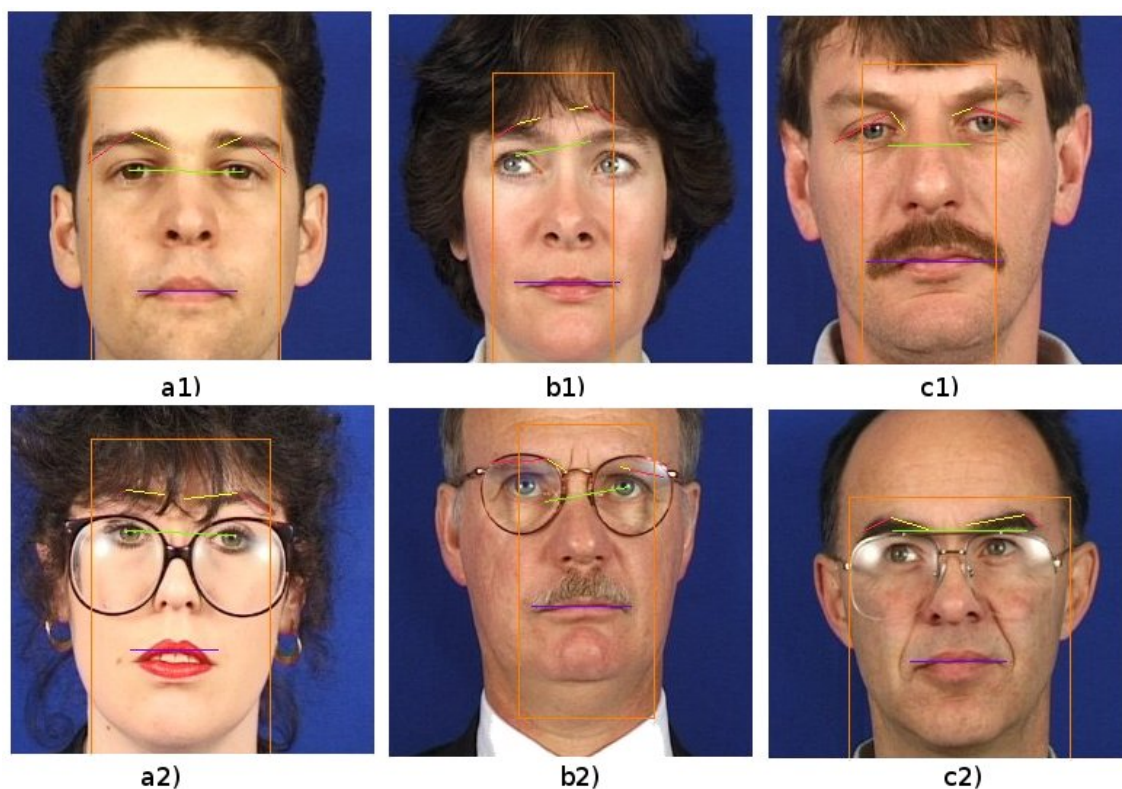
7.6.2 Databáze XM2VTSDB

Tato databáze obsahuje datovou sadu od 295 různých osob a všech různých ras. Narozdíl od IMM Face Database obsahuje také vysoký podíl osob s brýlemi, rozmazaných obrázků a osob s přivřenýma očima.

Detekce dominantních rysů

Nad touto databází byly provedeny čtyři testy podobně jako nad databází IMM. První test byl proveden nad celou databází a výsledná lokalizace dominantních rysů je zobrazena na obrázku 7.6, kde obrázky 7.6 a1) a 7.6 a2) zobrazují správnou lokalizaci, obrázky 7.6 b1) a 7.6 b2) chybnou lokalizaci jednoho oka a obrázky 7.6 c1) a 7.6 c2) chybnou lokalizaci obou očí. Ve druhém testu byl přidán soubor `obrazek.face` k obrázkům obličejů, které byly

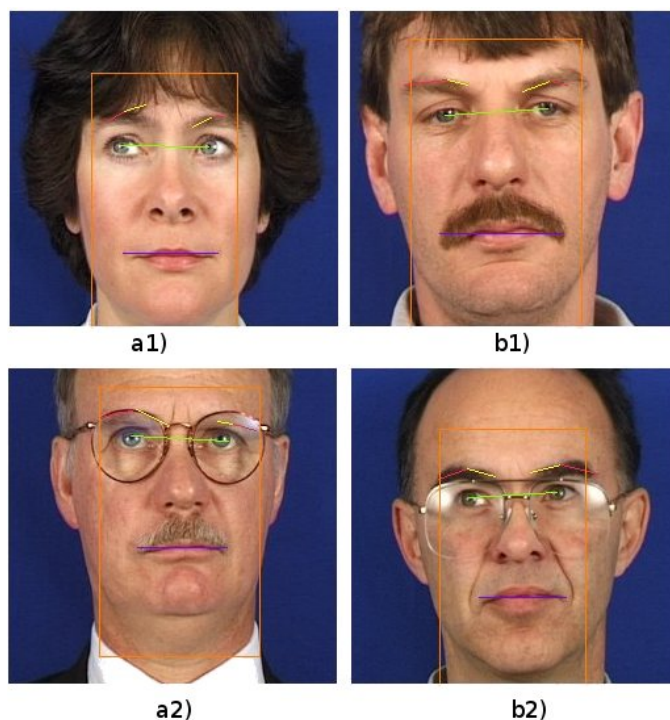
umístěny mimo očekávanou oblast. Ve třetím testu se u všech obličejů s chybně lokalizovanými dominantními rysy zkoušelo správně lokalizovat dominantní rysy ručním zadáním průměru kružnic reprezentujících oči. Výsledek testu pro obrázky 7.6 b1), 7.6 b2), 7.6 c1) a 7.6 c2) je zobrazen na obrázku 7.7. V posledním testu byly dodány soubory se souřadnicemi oblasti obličeje ke každému obrázku, na kterém v testu číslo 1 nebyly správně detekovány dominantní rysy. Výsledky všech čtyř testů je možné vidět v tabulce 7.2. Nejzajímavější výsledkem je výstup testu číslo 4, kdy stejně jako při testech nad databází IMM je úspěšnost blížící se testu číslo 3, ale narozdíl od testu číslo 3 se vše děje automaticky bez jakéhokoli zásahu uživatele. Na obrázku 7.8 je možné vidět výsledek pro tváře se špatnou lokalizací dominantních rysů zobrazených na obrázku 7.6.



Obrázek 7.6: Nalezení dominantních rysů nad daty databáze XM2VTSDB. Upraveno z [17].

Úspěšnost rozpoznání dominantních rysů				
lokalizace	test 1	test 2	test 3	test 4
očí	87,28 %	87,45 %	97,11 %	94,91 %
úst	96,94 %	97,11 %	98,13 %	98,64 %

Tabulka 7.2: Úspěšnost lokalizace dominantních rysů nad daty databáze XM2VTSDB.

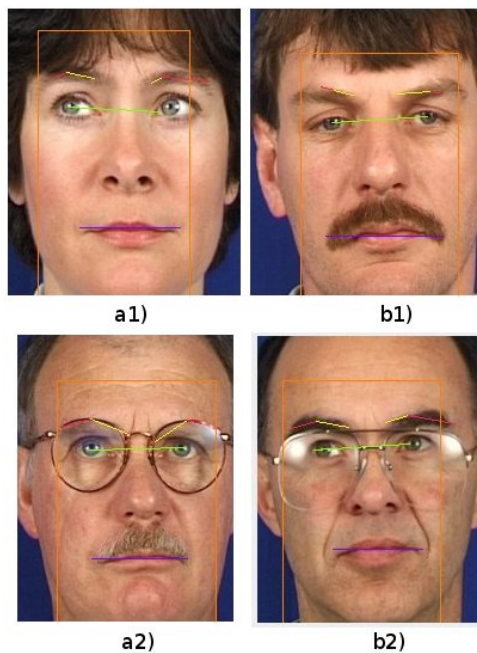


Obrázek 7.7: Nalezení dominantních rysů nad daty databáze XM2VTSDB pomocí ručního zadání velikosti kružnic reprezentujících oblast očí. Upraveno z [17].

Identifikace obličeje

Identifikace obličeje nad databází XM2VTSDB je provedena pro 295 různých osob. Od každé osoby je k dispozici osm vzorků obličejů s různým osvětlením a jinou pozicí. Pro trenovací sadu je od každé osoby použito šest obrázků obličeje a pro sadu testovací jsou použity dva obrázky obličeje. Identifikace nad zadanými daty je provedena postupně pro 10 až 200 tříd s krokem 10 a dále od 220 do 295 s krokem 20 pomocí metod KDA a komponentní KDA, kdy je opět použita metoda LDA pro vlastní dekompoziční problém. V tabulce B.3 jsou uvedeny nastavené parametry použité při testování.

Výsledky procesu identifikace je možné vidět v grafu 7.9, kde KDA zobrazena modrou barvou reprezentuje výslednou úspěšnost procesu identifikace pro zadaný počet tříd pomocí metody KDA a oranžová barva znázorňuje komponentní KDA a její úspěšnost rozpoznání osob pomocí komponentní KDA. Komponentní metoda má vyšší úspěšnost než klasická KDA, což je způsobeno lepším vyrovnáním komponentní metody s odchylkami obrazu způsobenými různým osvětlením.

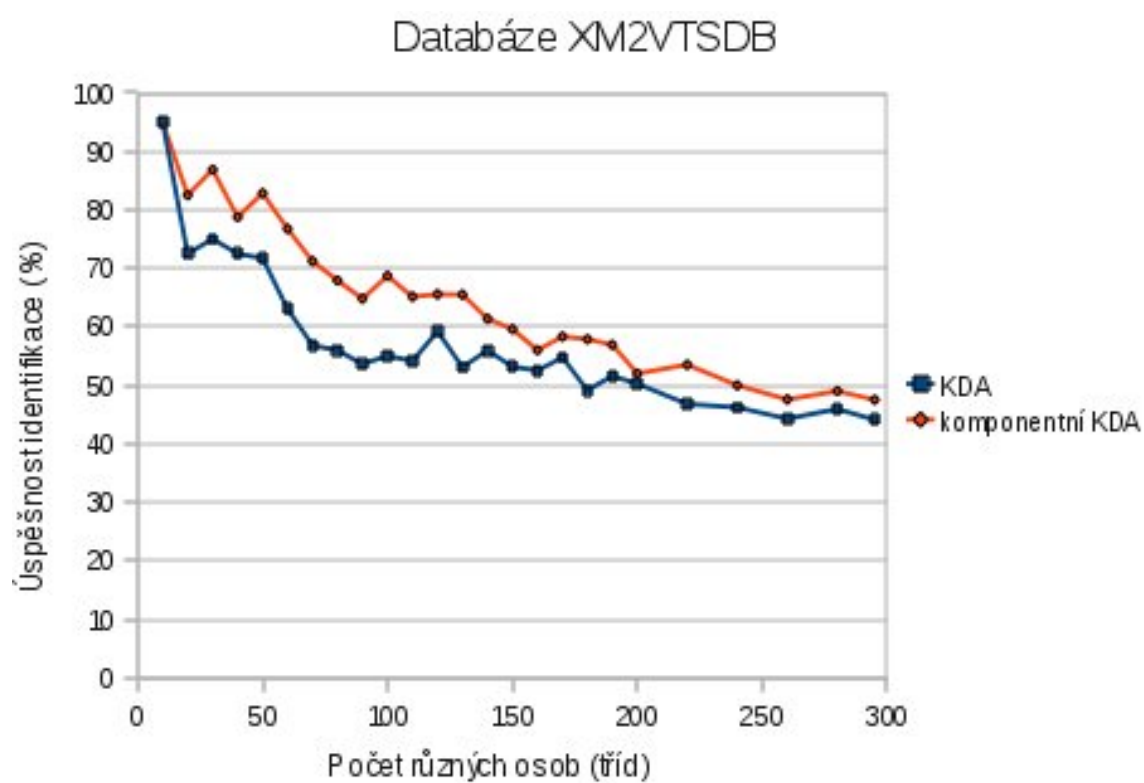


Obrázek 7.8: Detekce dominantních rysů u souborů s přidaným souborem `obrazek.face` obsahující souřadnice obličeje. Upraveno z [17].

7.7 Shrnutí

Z výsledků testů detekce dominantních rysů plyne, že pokud existuje relativně přesná oblast obličeje bez rušivých vlivů pozadí, vzroste úspěšnost správné detekce dominantních rysů až téměř k hodnotě kdy uživatel zadává průměr kružnic reprezentujících oči. Dále je patrné, že i když je oblast očí chybně lokalizovaná, ústa jsou lokalizována správně.

Z úspěšnosti testů identifikace obličeje vyplývá, že velký vliv na proces identifikace má počet trénovacích dat každé testovací třídy obličejů. Všechny výše reprezentované výsledky testů je možné si ověřit pomocí testovacích datových sad a aplikace přítomné na příloženém DVD A.



Obrázek 7.9: Identifikace nad daty databáze XM2VTSDb.

Kapitola 8

Rozšíření a možnosti budoucího vývoje

Tato kapitola se zabývá budoucím vývojem aplikace a možnostmi jejího rozšíření. Vývoj by se měl ubírat směrem přidávání různých rozšiřujících funkcí a vylepšením funkcí stávajících.

8.1 Rozpoznání rysů a identifikace více osob na jednom obrázku

Jedním z budoucích rozšíření je možnost v pomocném souboru `obrazek.face` definovat souřadnice více obličejů a všechny tyto obličeje následně zpracovat v procesu rozpoznání rysů a identifikace. Pro proces identifikace by v tomto případě bylo nutné explicitně definovat náležitost ke třídě obličejů, jelikož by nebylo možné automaticky rozpoznat náležitost k té které třídě. Explicitní zadání třídy obličejů by se týkalo datové sady k trénování. Pro testovací data by toto nebylo nezbytně nutné, pouze by selhalo automatické počítání úspěšnosti a uživatel by sám ze zobrazených výsledků musel rozhodnout, zda identifikace obličeje proběhla v pořádku.

8.2 Vlastní detektor obličeje

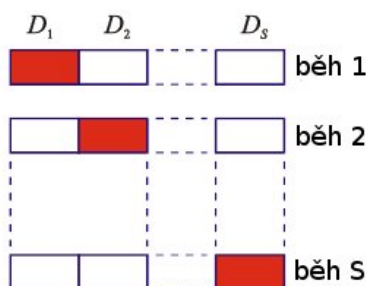
Rozšíření ve formě vlastního detektoru obličeje v aplikaci by bylo vítaným rozšířením, odpadla by nutnost použití dalšího samostatného nástroje¹, který by našel obličeje v obrázcích a vyexportoval do pomocných souborů `obrazek.face` souřadnice obličeje. Tento detektor by mohl být založen například na detekci barvy lidské kůže, kdy detektory založené na tomto způsobu lokalizace dosahují uspokojivých výsledků a složitost implementace není příliš vysoká. Jelikož aplikace má detektor rysů, stačilo by lokalizovat pouze kandidátské

¹například dříve zmíněný Detektor obličejů, akcelerovaný pomocí SIMD dostupný na FIT VUT na adrese <http://www.fit.vutbr.cz/research/prod/index.php.cs?id=68¬itle=1>

oblasti a na ně aplikovat detekci rysů. Pokud by byly nalezeny oči, ústa a obočí byla by tato oblast prohlášena za obličej a předána k dalšímu zpracování.

8.3 Automatizovaná volba vhodných parametrů pro KDA – cross-validace

Aplikace vyžaduje při procesu identifikace metodou KDA ruční zadání parametrů **sigma** a **dimenze prostoru**. Řešením a zautomatizováním by mohla být cross-validace [26], jejíž princip znázorněný na obrázku 8.1 je následující: V prvním kroku je množina dat rozdělena do S skupin. Ve druhém kroku se provede trénovací část na $S - 1$ skupinách a testování na skupinách zbývajících. Dokud nejsou vyčerpány všechny možnosti či není splněna jiná omezující podmínka, je proces od druhého kroku opakován. V posledním kroku jsou výsledky cross-validace vyhodnoceny. Proces cross-validace je velice časově náročný.



Obrázek 8.1: Cross-validace a rozdělení na S skupin a postupné trénování a testování. Převzato z [26].

8.4 Možnost ručního označení dominantních rysů

Proces identifikace je do určité míry ovlivněn přesností detekce dominantních rysů. Pokud se zaměříme hlavně na demonstraci účinnosti metody KDA pro identifikaci osob, může být výhodné mít možnost přesně označit umístění dominantních rysů obličeje. Tím je dosažen co nejpřesnější proces identifikace, který není téměř ovlivněn chybnou lokalizací dominantních rysů. Ručně označené dominantní rysy pak mohou být vyexportovány do speciálního souboru `obrazek.features` pro pozdější použití.

8.5 Zvýšení úspěšnosti detekce očí mnohonásobnou detekcí kružnic o různém průměru

Dalším možným rozšířením zvyšujícím úspěšnost nalezení očí je jejich vícenásobná detekce pomocí kružnic o různém průměru, kdy v oblasti oka, narozdíl od ostatních oblastí obličeje, jsou detekovány kružnice o různých průměrech se středem v téměř stejném místě. Poté by bylo určeno, kolik kružnic o různém průměru má střed v bodě $[x, y]$ s určitou tolerancí a na získané hodnoty by byl brán zřetel při konečném výběru oblasti očí z kandidátských oblastí.

8.6 Uložení jednou detekovaných dominantních rysů do souboru

Aplikace v tomto stádiu jednou detekované dominantní rysy po ukončení aplikace nikam neukládá, tudíž může docházet ke zbytečné opakované detekci spotřebovávající výpočetní výkon. Řešením této situace by byla možnost ukládat jednou nalezené dominantní rysy do speciálního souboru `obrazek.features`. Tyto uložené dominantní rysy by mohly být místo opětovné detekce volitelně načteny.

8.7 Kombinace více metod

Rozšíření a budoucí vylepšení pomocí kombinace více metod by mohlo být použito pro detekci úst a očí, kdy by ústa byla detekována pomocí kombinace metod popsaných v podkapitolách 4.2.1 a 4.2.2. Oblast prohlášená za ústa by musela vyhovovat podmínce obou metod, aby byla prohlášena za oblast úst. Oči by mohly být detekovány pomocí metod popsaných v podkapitolách 4.1.1, 4.1.2 a 4.1.3, kdy by nejdříve byla určena přibližná oblast očí pomocí metody popsané v podkapitole 4.1.2 a až poté by na tuto oblast byly aplikovány metody zbývající.

Kapitola 9

Závěr

Cílem práce je seznámení s problematikou identifikace obličeje a určení a řešení problémů spojených s rozpoznáním obličeje. Jedná se např. o problémy spojené s pořízením obrazu tváře za různých podmínek či nelinearitu lidských tváří ve vysokorozměrném prostoru.

Práce se zaměřuje převážně na nalezení význačných bodů a lokalizaci dominantních rysů obličeje. Pro nalezení dominantních rysů jsou použity metody popsané v podkapitolách 4.1.3, 4.2.1 a 4.4 a podle výsledků testů v podkapitole 7.6 tento přístup dosahuje relativně dobrých výsledků.

Kapitola 5 popisuje jak je důležitá normalizace, jelikož obraz tváře je pořízen z různé vzdálenosti, a pod různým úhlem a v neposlední řadě s různým osvětlením. S odlišným osvětlením se lépe vyrovnávají komponentní varianty metod.

Co se týká identifikace obličeje, práce se převážně zaměřuje na metody lineární analýzy, konkrétně na lineární diskriminační analýzu (LDA) a nelineární jádrovou diskriminační analýzu (KDA), která řeší největší nedostatky lineárních metod. Největším nedostatkem lineárních metod je především nemožnost správně separovat nelineární prostor, ve kterém jsou obrázky obličejů reprezentovány.

V kapitole 7 jsou popsány implementované metody a je provedeno testování aplikace nad daty různých obličejových databází. Výsledky jsou uspokojivé a úspěšnost identifikace 295 různých osob dosahuje téměř 50 % při použití komponentní KDA. Úspěšnost u lokalizace dominantních rysů se nachází v rozmezí 87–97 %. Aplikace by mohla mít reálné využití jako vstupní biometrický systém menší či střední firmy, kde by nebyl kladen velký důraz na bezpečnost, pokud by byl dostatečný počet trénovacích dat a uživatelé systému by dodržovali určitá pravidla (jako například vzdálenost od kamery a úhel natočení obličeje) a pokud by ještě ke všemu bylo zajištěno konstantní osvětlení, úspěšnost rozpoznání by se zajisté zvýšila. Na druhou stranu využití aplikace je nevhodné na místech, kde je důležitá co nejvyšší bezpečnost, jelikož přelstít systém založený pouze na 2D identifikaci obličeje není obtížné a zvládne to průměrně technicky zdatný jedinec. Vhodná je kombinace s jiným biometrickým systémem využívající např. otisků prstů či 3D identifikace obličeje.

Předposlední kapitola 8 pojednává o budoucím směru vývoje aplikace a jeho možných rozšířeních, jako jsou především vlastní detektor obličeje (podkapitola 8.2), automatizovaná volba vhodných parametrů (podkapitola 8.3) a zvýšení úspěšnosti detekce očí mnohonásobnou detekcí kružnic o různém průměru (podkapitola 8.5).

Literatura

- [1] Bishop, C. M.: *Pattern Recognition and Machine Learning*. Springer, 2006, ISBN 978-0-387-31073-2.
- [2] Chew, W. J.; Seng, K. P.; Ang, L.-M.: Nose Tip Detection on a Three-Dimensional Face Range Image Invariant to Head Pose. *IMECS*, ročník 1, 2009: s. 1–5, [Online], [cit. 2010-01-06].
URL <http://www.iaeng.org/publication/IMECS2009/IMECS2009_pp858-862.pdf>
- [3] Cootes, T.: Active Appearance Models. [Online], [cit. 2010-01-06].
URL <<http://personalpages.manchester.ac.uk/staff/timothy.f.cootes/Models/aam.html>>
- [4] Částek, P.: *Identifikace Obličejů*. Diplomová práce, FIT VUT v Brně, 2008.
- [5] Delac, K.; Grgic, M.; Grgic, S.:
Independent comparative study of PCA, ICA and LDA on the FERET data set. *Wiley Periodicals*, ročník 15, 2005: s. 252–260, [Online], [cit. 2010-05-04].
URL <http://www.face-rec.org/algorithms/Comparisons/delac_03.pdf>
- [6] Delac, K.; Grgic, M.: *Face Recognition*. I-TECH Education and Publishing, 2007, ISBN 3-86611-283-1, 94–101 s.
- [7] Drahanský, M.: Přehled biometrických systémů a testování jejich spolehlivosti. [Online], [cit. 2010-01-06].
URL <http://data.security-portal.cz/clanky/113/odborne_prednasky/Prezentace.pdf>
- [8] Duda, R. O.; Hart, P. E.; Stork, D. G.: *Pattern Classification*. Wiley-Interscience, 2001, ISBN 0-471-05669-3.
- [9] Fisher, R.; Perikns, S.; Walker, A.; aj.: HIPR. [Online], [cit. 2010-05-06].
URL <<http://homepages.inf.ed.ac.uk/rbf/HIPR2/index.htm>>

- [10] Jebara, T.: Face Normalization and Recognition. [Online], [cit. 2010-01-06].
URL <<http://www1.cs.columbia.edu/~jebara/htmlpapers/UTHESES/node46.html>>
- [11] Kass, M.; Witkin, A.; Terzopoulos, D.: Snakes: Active Contour Models. *International Journal of Computer Vision*, 1988: s. 321–331, [Online], [cit. 2010-05-09].
URL <<http://www.cs.ucla.edu/~dt/papers/ijcv88/ijcv88.pdf>>
- [12] Kwon, Y. H.; da Vitoria Lobo, N.: Face Detection Using Templates. *Pattern Recognition*, 1994: s. 764–767, [Online], [cit. 2010-05-09].
URL <<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=576435>>
- [13] Leymarie, F.; Levine, M. D.: Tracking Deformable Objects in the Plan Using an Active Contour Model. *IEEE Transactions Pattern Analysis and Machine Intelligence*, ročník 15, č. 6, 1993: s. 617–634, [Online], [cit. 2010-05-09].
URL <<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=216733>>
- [14] Li, Y.; Gong, S.; Liddell, H.: Kernel Discriminant Analysis. [Online], [cit. 2010-01-06].
URL <http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/LI1/kda/index.html>
- [15] Liu, C.; Wechsler, H.: Comparative Assessment of Independent Component Analysis (ICA) for Face Recognition. *Second International Conference on Audio- and Video-based Biometric Person Authentication*, 1999: s. 1–6, [Online], [cit. 2010-01-06].
URL <<http://www.face-rec.org/algorithms/ICA/liu99comparative.pdf>>
- [16] Loos, H. S.; Wieczorek, D.; Würtz1, R. P.; aj.: Computer-based recognition of dysmorphic faces. [Online], [cit. 2010-01-06].
URL <<http://www.nature.com/ejhg/journal/v11/n8/full/5200997a.html>>
- [17] Messer, K.; Kittler, J.; Luettkrusain, J.; aj.: XM2VTSDB. [Online], [cit. 2010-05-10].
URL <<http://www.ee.surrey.ac.uk/CVSSP/xm2vtsdb/>>
- [18] Nikolaidis, A.; Pitas, I.: Facial Feature Extraction and Determination of Pose. *Pattern Recognition*, 2000: s. 1783–1791, [Online], [cit. 2010-05-09].
URL <<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.48.1852&rep=rep1&type=pdf>>
- [19] Peng, K.; Chen, L.; Ruan, S.; aj.: A Robust Algorithm for Eye Detection on Gray Intensity Face without Spectacles. *JCS&T*, ročník 5, č. 3, 2005: s. 1–6, [Online], [cit. 2010-01-06].
URL <<http://journal.info.unlp.edu.ar/Journal/journal15/papers/JCST-Oct05-3.pdf>>

- [20] Pietrowcew, A.: Face detection in colour images using fuzzy Hough transform. *OPTO-ELECTRONICS REVIEW*, 2003: s. 247–252, [Online], [cit. 2010-05-09]. URL <<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.97.13&rep=rep1&type=pdf>>
- [21] Schölkopf, B.; Smola, A.; Müller, K.-R.: Nonlinear component analysis as a kernel eigenvalue problem. *Neural Comput.*, ročník 10, č. 5, 1998: s. 1299–1319, ISSN 0899-7667, [Online], [cit. 2010-01-06]. URL <http://www.face-rec.org/algorithms/Kernel/kernelPCA_scholkopf.pdf>
- [22] Singh, N.: Face Recognition by Capturing Eye Illumination Spot. *Applied Imagery Pattern Recognition Workshop*, ročník 33, 2004: s. 153–158, ISSN 0-7695-2250-5, [Online], [cit. 2010-01-06]. URL <<http://ieeexplore.ieee.org/Xplore/login.jsp?url=http%3A%2F%2Fieeexplore.ieee.org%2Fiel5%2F9674%2F30558%2F01409691.pdf%3Farnumber%3D1409691&authDecision=-203>>
- [23] Singh, S. K.; Chauhan, D. S.; Vatsa, M.; aj.: A Robust Skin Color Based Face Detection Algorithm. *Tamkang Journal of Science and Engineering*, ročník 6, č. 4, 2003: s. 227–234, [Online], [cit. 2010-05-06]. URL <<http://www2.tku.edu.tw/~tkjse/6-4/6-4-6.pdf>>
- [24] Španěl, M.: Aktivní kontury. [Online], [cit. 2010-05-08]. URL <<http://www.fit.vutbr.cz/~spanel/segmentace/>>
- [25] Španěl, M.: Segmentace obrazu. [Online], [cit. 2010-05-08]. URL <https://www.fit.vutbr.cz/study/courses/POV/private/lectures/pov_04_segmentace_obrazu.pdf>
- [26] Španěl, M.: Statistické rozpoznání. [Online], [cit. 2010-05-08]. URL <https://www.fit.vutbr.cz/study/courses/POV/private/lectures/pov_02_statisticke_rozpoznavani.pdf>
- [27] 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.
- [28] Taekyun, K.; Hyunwoo, K.; Wonjun, H.: Componentbased LDA face description for image retrieval and MPEG7 standardisation. *Image and Vision Computing*, ročník 23, 2005: s. 631–642, [Online], [cit. 2010-04-24]. URL <<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.105.7840&rep=rep1&type=pdf>>

- [29] Twining, C.: Kernel Principal Component Analysis. [Online], [cit. 2010-01-06].
URL <http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/TWINING1/node1.html>
- [30] Vlach, J.; Přinosil, J.: Lokalizace obličej v obraze s komplexním pozadím. *Elektrorevue*, 2007: s. 1–12, ISSN 1213-1539.
- [31] Wiskott, L.; Fellous, J.-M.; Krüger, N.; aj.: Face Recognition by Elastic Bunch Graph Matching. *IEEE Trans. Pattern Anal. Mach. Intell.*, ročník 19, č. 7, 1997: s. 775–779, ISSN 0162-8828, [Online], [cit. 2010-01-06].
URL <<http://www.face-rec.org/algorithms/EBGM/WisFelKrue99-FaceRecognition-JainBook.pdf>>
- [32] Yang, M.-H.; Kriegman, D. J.; Ahuja, N.: IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE. *Image and Vision Computing*, ročník 24, č. 1, 2002: s. 34–58, [Online], [cit. 2010-05-08].
URL <<http://vision.ai.uiuc.edu/mhyang/papers/pami02a.pdf>>
- [33] Yuille, A.; Hallinan, P.; Cohen, D.: Feature Extraction from Faces Using Deformable Templates. *Computer Vision and Pattern Recognition*, ročník 8, č. 2, 1992: s. 99–11, [Online], [cit. 2010-05-09].
URL <<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=37836>>

Seznam obrázků

2.1	Identifikace. Převzato z [7].	5
2.2	Verifikace. Převzato z [7].	6
2.3	Vztah FMR/FNMR a FAR/FRR. Převzato z [7].	7
2.4	Míra vyrovnání chyb. Převzato z [7].	8
2.5	ROC křivky. Převzato z [7].	8
3.1	a) $n=1$, původní obrázek, b) $n=4$, c) $n=8$, d) $n=16$. Původní a odpovídající obrázky v nízkém rozlišení. Každá čtvercová oblast se skládá z $n \times n$ bodů u kterých je hodnota intenzity každého bodu nahrazena průměrnou hodnotou intenzity bodů čtvercové oblasti. Převzato z [32].	10
3.2	Typický obličej používaný ve strukturálních metodách. Kódující pravidla jsou založena na znalosti vlastností obličejových oblastí (rozložení intenzity a jejich rozdíl). Převzato z [32].	10
3.3	a) Podprostor barvy kůže v RGB. b) Podprostor barvy kůže v YCbCr. Převzato z [30].	11
3.4	Nalezené obličej pomocí Houghovy transformace a zvýraznění přesné oblasti obličej pomocí aktivních kontur. Převzato z [18].	12
4.1	Snímání osvětleného bodu oka. Outer arc – vnější okraj duhovky, Illumination spot – osvětlený bod zornice. Převzato z [22].	15
4.2	Detekce oblastí očí. Převzato z [19].	16
4.3	Houghova transformace. a) původní obrázek, b) binarizovaný obrázek se zvýrazněním nalezených kružnic, c) Houghův prostor (hodnoty jsou přepočítány do intervalu $[0, 255]$).	17
4.4	Kandidátské oblasti očí. Všechny kružnice detekované Houghovou transformací o zadaném průměru.	18
4.5	Oblast očí. Kružnice prohlášené za oblast očí. Zelenou barvou pravé oko, červenou levé.	18
4.6	Vymezená oblast pro hledání rtů.	19
4.7	Obraz v F-barevném prostoru a její binární obraz. Převzato z [30].	19
4.8	Vymezená oblast nosu.	20

5.1	Roztažení kontrastu. a) původní obrázek, b) roztažení kontrastu zvolením dolní a horní hranice 0 a 255, c) roztažení kontrastu, dolní a horní hranice určena pomocí procent.	23
5.2	Transformační funkce histogramu. Převzato z [9].	24
5.3	Vyrovnnání histogramu. a) původní obrázek, b) obrázek po vyrovnnání histogramu.	25
6.1	Proces lineární analýzy. Převzato z [5].	27
6.2	PCA přístup k rozpoznání obličeje. Převzato z [6].	28
6.3	(a) Smíchané body při promítnutí na linii. (b) Oddělené body při promítnutí na jinou linii. Převzato z [6].	30
6.4	(a) Dobré rozdělení tříd. (b) Špatně rozdělení tříd. Převzato z [6].	30
6.5	LDA přístup k rozpoznání obličeje. Převzato z [6].	31
6.6	Prvních 16 Fisherových obličejů s nejvyššími vlastními hodnotami. Číselné hodnoty reprezentují souřadnice na ose X a Y . Převzato z [6].	32
6.7	Vliv osvětlení na LDA a komponentní LDA	33
6.8	Definice komponent LDA	34
6.9	Základní myšlenka KPCA. Převzato z [21].	35
6.10	Jádrová diskriminační analýza. Převzato z [14].	36
6.11	Elastický srovnávací diagram aplikován na obrázek tváře. Převzato z [31].	37
6.12	Příklady Gaborovy vlnkové transformace. Převzato z [16].	37
6.13	Active Appearance Models. a) inicializace b) pozice po první iteraci c) pozice po druhé iteraci d) konvergováno. Převzato z [3].	38
7.1	Schéma aplikace.	42
7.2	Lokalizace dominantních rysů. a) nalezení správných dominantních rysů, b) nalezení správně pouze části dominantních rysů, c) nenalezení rysů z důvodu špatného umístění obličeje v obraze. Upraveno z [27].	44
7.3	Nalezení správných dominantních rysů při zadání velikosti kružnic reprezentující oči. Upraveno z [27].	44
7.4	Detekce dominantních rysů u souborů s přidaným souborem <code>obrázek.face</code> obsahující souřadnice obličeje. Upraveno z [27].	45
7.5	Identifikace nad daty databáze IMM.	46
7.6	Nalezení dominantních rysů nad daty databáze XM2VTSDB. Upraveno z [17].	47
7.7	Nalezení dominantních rysů nad daty databáze XM2VTSDB pomocí ručního zadání velikosti kružnic reprezentujících oblast očí. Upraveno z [17].	48
7.8	Detekce dominantních rysů u souborů s přidaným souborem <code>obrázek.face</code> obsahující souřadnice obličeje. Upraveno z [17].	49
7.9	Identifikace nad daty databáze XM2VTSDB.	50

8.1	Cross-validace a rozdělení na S skupin a postupné trénování a testování. Převzato z [26].	52
C.1	Diagram tříd aplikace.	67

Příloha A

Obsah DVD

Na DVD se v kořenovém adresáři nalézají technická zpráva v souboru `techzprava.pdf` obsahující text diplomové práce. Adresář `zdrojove_soubory_zprava` obsahuje zdrojové soubory technické zpravy. Dále ve složce `/src` jsou umístěny zdrojové soubory aplikace. Přeložená aplikace je uložena v adresáři `/dist`. Nápověda k aplikaci je uložena v adresáři `napoveda`. Programová dokumentace je umístěna v adresáři `/doc`. Adresář `/FaceDatabase` obsahuje databáze obličejů včetně nachystaných testovacích dat.

- `dist`
- `doc`
- `FaceDatabase`
- `napoveda`
- `src`
- `zdrojove_soubory_zprava`

Příloha B

Parametry testů

B.1 Identifikace nad daty databáze IMM

počet tříd	KDA		komponentní KDA	
	sigma	dimenze prostoru	sigma	dimenze prostoru
10	35	15	100	15
20	50	15	100	15
30	50	20	130	30
40	100	25	130	30

Tabulka B.1: Parametry použité při testování identifikace nad daty databáze IMM, při použití 4 vzorků dat jako trénovacích a 2 jako testovacích.

počet tříd	KDA		komponentní KDA	
	sigma	dimenze prostoru	sigma	dimenze prostoru
10	1000	20	100	20
20	100	20	100	25
30	100	25	100	30
40	100	30	1000	30

Tabulka B.2: Parametry použité při testování identifikace nad daty databáze IMM, při použití 5 vzorků dat jako trénovacích a 1 testovacího.

B.2 Identifikace nad daty databáze XM2VTSDB

počet tříd	KDA		komponentní KDA	
	sigma	dimenze prostoru	sigma	dimenze prostoru
10	50	30	50	25
20	100	40	10000	30
30	10	40	1000	35
40	1000	45	1000	40
50	100	40	100	40
60	100	45	1000	45
70	1000	50	1000	50
80	1000	60	1000	50
90	1000	60	1000	55
100	10000	80	10000	70
110	10000	80	10000	80
120	10000	80	10000	80
130	10000	80	10000	80
140	10000	80	10000	100
150	10000	120	10000	150
160	10000	120	10000	150
170	10000	140	10000	140
180	10000	140	10000	140
190	10000	140	10000	160
200	10000	160	10000	160
210	-	-	-	-
220	10000	180	10000	180
230	-	-	-	-
240	10000	180	10000	180
250	-	-	-	-
260	10000	180	10000	180
270	-	-	-	-
280	10000	180	10000	180
290	-	-	-	-
295	10000	180	10000	180

Tabulka B.3: Parametry použité při testování identifikace nad daty databáze XM2VTSDB, při použití 6 vzorků dat jako trénovacích a 2 jako testovacích.

Příloha C

Databáze obličejů, nápověda a diagram tříd aplikace

C.1 Databáze obličejů

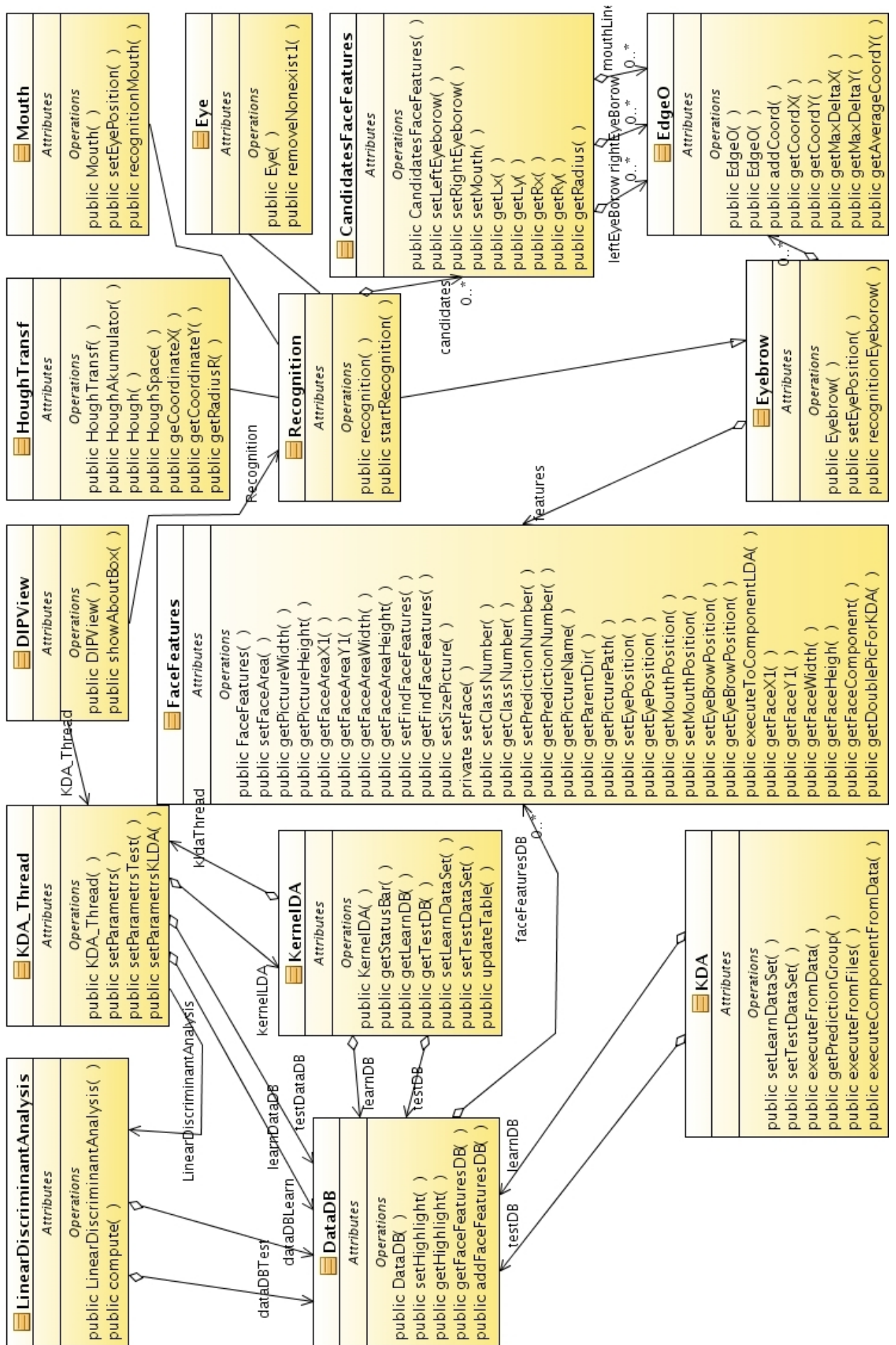
IMM Face Database <<http://www2.imm.dtu.dk/~aam/>>

XM2VTSDB databáze <<http://www.ee.surrey.ac.uk/CVSSP/xm2vtsdb/order/datasets.html>>

C.2 Nápověda

Nápověda je dostupná na přiloženém DVD A a aktuální verze nápovědy je dostupná také na adrese <<http://www.stud.fit.vutbr.cz/~xsvabe02/dip/>>.

C.3 Diagram tříd aplikace



Obrázek C.1: Diagram tříd aplikace.