



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ

FACULTY OF INFORMATION TECHNOLOGY

ÚSTAV INTELIGENTNÍCH SYSTÉMŮ

DEPARTMENT OF INTELLIGENT SYSTEMS

**METODY EXTRAKCE CHARAKTERISTICKÝCH RYSŮ
OBLIČEJE**

FACE FEATURES EXTRACTION METHODS

BAKALÁŘSKÁ PRÁCE

BACHELOR'S THESIS

AUTOR PRÁCE

AUTHOR

ĽUBOMÍR ADAMČEK

VEDOUcí PRÁCE

SUPERVISOR

Ing. TOMÁŠ GOLDMANN

BRNO 2017

Abstrakt

Tvár je dlhodobo jedna z najatraktívnejších dostupných biometrií človeka pre jej ľahké a pohodlné nasnímanie. Využitie je široké – od bezpečnosti, cez monitorovanie až po zábavný priemysel. Táto práca predstavuje doménu biometrie tváre a analyzuje 3 extrakčné metódy charakteristických rysov tváre – PCA, LBP a HOG. Súčasťou práce je tiež efektívna implementácia týchto algoritmov spolu s navrhnutým grafickým používateľským rozhraním využitým na experimentovanie s touto implementáciou a na vyhodnotenie jej úspešnosti na sade snímok zachytávajújúcich osoby v rôznych podmienkach.

Abstract

The face has been one of the most attractive available human biometries for a long time due to it's easy and convenient way to obtain it. Possibilities of it's utilization are broad – from security, through monitoring up to entertainment industry. This work presents the domain of face biometry and analyses 3 extraction methods of facial features – PCA, LBP and HOG. A part of this work is also an efficient implementation of these algorithms including GUI designed to experiment with this implementation and to evaluate its performance on a set of images capturing people in various conditions.

Klíčové slová

biometria, rozpoznávanie tváre, metódy extrakcie, charakteristické rysy, vektor príznačkov, PCA, analýza hlavných komponent, LBP, lokálne binárne vzory, HOG, histogramy orientovaných gradientov

Keywords

biometry, face recognition, extraction methods, features, feature vector, PCA, principal component analysis, LBP, local binary patterns, HOG, histograms of oriented gradients

Citácia

ADAMČEK, Ľubomír. *Metody extrakce charakteristických rysů obličeje*. Brno, 2017. Bakalářská práce. Vysoké učení technické v Brně, Fakulta informačních technologií. Vedoucí práce Goldmann Tomáš.

Metody extrakce charakteristických rysů obličeje

Prehlásenie

Prehlasujem, že som túto bakalársku prácu vypracoval samostatne pod vedením pána Ing. Tomáša Goldmanna. Uviedol som všetky literárne pramene a publikácie, z ktorých som čerpal.

.....
Ľubomír Adamček
17. mája 2017

Pod'akovanie

Rád by som poďakoval svojmu vedúcemu Ing. Tomášovi Goldmannovi za odborné vedenie práce. Taktiež by som chcel poďakovať svojej rodine a priateľom za podporu pri štúdiu a písaní tejto bakalárskej práce.

Obsah

Úvod	3
1 Rozpoznávanie tváre	4
1.1 Základná terminológia	4
1.2 Biometria tváre	5
1.3 Proces rozpoznávania tváre	5
1.4 Metódy extrakcie charakteristických rysov	6
1.4.1 Analýza hlavných komponent	7
1.4.2 Lokálne binárne vzory	10
1.4.3 Histogramy orientovaných gradientov	13
1.5 Metriky vyhodnocovania rozpoznávacích systémov	13
1.5.1 Vyhodnocovanie verifikácie	14
1.5.2 Vyhodnocovanie identifikácie na uzavrenej množine subjektov	15
2 Návrh riešenia	17
2.1 Porovnávanie vektorov charakteristických rysov	17
2.2 Realizácia extrakčných metód	18
2.2.1 Analýza hlavných komponent	18
2.2.2 Metódy lokálnych binárnych vzorov a histogramov orientovaných gradientov	19
2.3 Analýza vybraných podúloh	21
2.3.1 Detekcia tváre	21
2.3.2 Interpolácia intenzity pixelov	22
2.3.3 Efektívne počítanie zmien bitových úrovní	22
2.3.4 Aproximácia výpočtu gradientov	24
3 Implementácia	25
3.1 Použité technológie	25
3.1.1 Knihnica OpenCV	25
3.1.2 Qt framework	26
3.2 Architektúra aplikácie	26
3.3 Používateľské rozhranie implementovanej aplikácie	27
3.4 Konzolová aplikácia	29
4 Experimenty	30
4.1 Použité dátové sady	30
4.2 Metodika experimentovania	31
4.2.1 Nastavenie východných parametrov pri experimentovaní	31

4.2.2	Proces experimentovania a vyhodnocovania	32
4.3	Vyhodnotenie rozpoznávania pri odhadnutých parametroch	33
4.4	Závislosť konfigurácie jednotlivých metód na úspešnosti rozpoznávania	34
4.4.1	Počet vlastných tvárí	35
4.4.2	Množstvo regiónov u metód LBP a HOG	35
4.4.3	Počet bodov a polomer LBP operátora	36
4.4.4	Počet a rozsah košov HOG histogramu	36
4.4.5	Rozlíšenie analyzovaných snímok	38
4.5	Vyhodnotenie rozpoznávania pri použití optimálnych hodnôt parametrov	38
4.6	Analýza vzdialenosti šablón a rozpoznávanej snímky	40
4.7	Rýchlosť a pamäťové nároky implementácie	40
4.8	Vyhodnotenie výsledkov experimentovania	42
4.9	Zhodnotenie riešenia a možnosti ďalšieho napredovania	43
	Záver	44
	Literatúra	46
	Zoznam príloh	49
	Príloha A	50

Úvod

Tvár je ideálna biometrika vo viacerých aspektoch – na rozdiel od iných biometrií zhotovenie jej snímky nevyžaduje nadbytočnú aktivitu zo strany používateľa, je ju v dnešnej dobe možné nasnímať za pomoci bežných cenovo dostupných technológií a bez potenciálnych zdravotných rizík spojených s používaním iných rozpoznávacích systémov, ako môže byť napríklad prenos nečistôt pri dotyku kontaktných snímačov. Jej využitie v oblasti bezpečnosti alebo interakcie s technológiami je preto mimoriadne žiadané. Avšak, analýza snímok tváre o štandardne vysokej dimenzionalite prináša určité komplikácie.

Pre ľudské oko a mozog jednoduchá a bežná úloha rozpoznať osobu z pohľadu na fotografiu sa pri snahe o automatizované spracovanie modernými technológiami ukazuje ako riešenie naozaj netriviálneho problému, ku ktorému hľadalo prístup v uplynulých niekoľkých dekádach mnoho odborníkov a vedeckých tímov.

Táto bakalárska práca predstavuje v kapitole 1 doménu biometrie tváre a detailne popisuje princíp fungovania a vlastnosti troch extrakčných metód charakteristických rysov tváre – analýza hlavných komponent, lokálne binárne vzory a histogramy orientovaných gradientov. Ďalej sú v tejto kapitole popísané metriky používané na vyhodnocovanie rozpoznávacích systémov.

Predmetom kapitoly 2 je návrh realizácie jednotlivých extrakčných metód. Okrem uvedených ich vývojových diagramov a návrhu spôsobu porovnávania vektorov charakteristických rysov je jej súčasťou tiež analýza vybraných podúloh, ktorým je pri implementácii nutné čeliť, a ku každej podúlohe je navrhnuté efektívne riešenie problému.

Na obsah predchádzajúcej kapitoly nadväzuje kapitola 3, v ktorej je popísaná architektúra implementovanej aplikácie, použité technológie na jej realizáciu a predstavené grafické a konzolové rozhranie.

Kapitola 4 prezentuje experimenty prevedené na troch rôznych zvolených dátových sadách, ktorých cieľom bolo odmerať úspešnosť a rýchlosť rozpoznávania implementácie z tejto práce na identifikačných a verifikačných úlohách a bližšie skúmať závislosť úspešnosti identifikácie na použitej konfigurácii jednotlivých metód. Na jej záver je zhodnotené riešenie zrealizované v rámci tejto práce a sú tu tiež navrhnuté možnosti ďalšieho napredovania v oblasti výskumu metód zameraných na extrakciu charakteristických rysov tváre.

Záverečná kapitola sumarizuje výsledky tejto práce.

Kapitola 1

Rozpoznávanie tváre

Rozpoznávanie tváre je úloha výpočtového systému, ktorej cieľom je po zanalyzovaní predloženej digitálnej snímky tváre určiť, či je na nej osoba, ktorá sa očakáva, prípadne o akú osobu z databázy ide. Na vykonanie tejto pre ľudský mozog triviálnej úlohy sa musí v prípade vyhodnocovania technológiou vykonať značné množstvo výpočtov vo fázach od prípravy snímky do vhodného formátu v rámci fázy predspracovania, cez extrakciu významných črt tváre, ktoré umožnia danú osobu spoľahlivo odlíšiť od iných osôb, až po porovnávanie reprezentácie týchto vyextrahovaných črt so záznamami uloženými v databáze.

Obsahom tejto kapitoly je uvedenie do problematiky rozpoznávania tvári. Na začiatku je predstavená základná terminológia z domény biometrie a dôvody, ktoré motivujú výskum v tejto oblasti. Ďalej je predstavená tvár ako jeden z významných biometrických reprezentantov a úlohy z každodenného života, pri ktorých je bežne uplatňované rozpoznávanie tváre.

Druhá časť tejto kapitoly je venovaná popisu štandardného procesu rozpoznávania ľudskej tváre a konkrétne kroku extrakcie charakteristických rysov, ktorý je predmetom tejto bakalárskej práce. Následne sú podrobne predstavené tri významné metódy slúžiace na extrakciu charakteristických rysov.

Na záver kapitoly sú uvedené metriky užitočné pri vyhodnocovaní úspešnosti implementácií rôznych prístupov k problematike rozpoznávania tvári.

1.1 Základná terminológia

Charakteristický rys alebo *príznak* je merateľná vlastnosť pozorovaného objektu. Množina takýchto príznakov je potom označovaná ako *vektor charakteristických rysov* a je využívaný na popis pozorovaných objektov [13].

Biometrika je fyzická alebo psychologická črta, ktorá môže byť odmeraná, zaznamenaná a kvantifikovaná, a je možné ju využiť na vytvorenie *biometrického záznamu*. Takýto záznam potom umožňuje s istou mierou pravdepodobnosti určiť, či osoba zaznamenaná v budúcnosti je tá istá, ktorej patrí pôvodný záznam [26].

Technológie založené na *biometrii* zahŕňajú rozpoznávanie založené na fyziologických alebo behaviorálnych črtách. Príkladom prvej kategórie sú otlacky prstov, tvar prstov alebo ruky, rozloženie žíl v ruke alebo oku, alebo štruktúra dúhovky či sietnice. Reprezentantmi druhej kategórie je zasa chôdza alebo podpis osoby.

Na rozdiel od rôznych hesiel, kľúčov alebo čipov, ktoré sú v dnešnej dobe bežne využívané na zabezpečenie informácií alebo overenie identity a môžu byť relatívne ľahko ukradnuté,

skopírované alebo sú nepraktické na zapamätanie, má biometrický záznam niekoľko výhod. Daná osoba ho nemôže zabudnúť ani stratiť, a ak je vhodne zvolený je unikátny, stály a nepodvrhnutelný [16].

1.2 Biometria tváre

Ako je uvedené v [16], tvár je z hľadiska dôvodov popísaných v sekcii 1.1 ideálna biometrika vo viacerých aspektoch – na rozdiel od iných biometrik jej nasnímanie nevyžaduje nadbytočnú aktivitu zo strany používateľa (v porovnaní s napríklad snímaním otlaku prsta), čo je výhodné najmä pri bezpečnostných a monitorovacích účeloch. Nasnímanie tváre je v dnešnej dobe možné za pomoci bežných cenovo dostupných kamier alebo fotoaparátov, na rozdiel od špeciálnych technológií na snímanie, napríklad, sietnice oka. A keďže zaznamenanie snímky tváre je bezkontaktné, nehrozí napríklad prenos nečistôt od iných používateľov snímacieho zariadenia a s tým spojené zdravotné riziká.

Existujú 2 základné úlohy, na ktoré je rozpoznávanie tvári používané:

- **verifikácia (1:1)** – k dispozícii je nasnímaná tvár neznámej osoby a predložená identita; systém overuje, či zachytená osoba je tou, za ktorú sa vydáva,
- **identifikácia (1:N)** – nasnímaná tvár osoby je porovnaná s mnohými identitami v databáze s cieľom určiť osobu, ktorej táto tvár patrí.

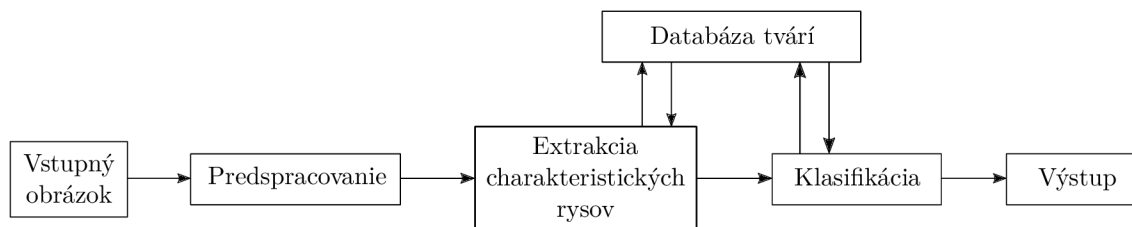
Táto funkcionálna nachádza uplatnenie v mnohých oblastiach, medzi ktoré patrí najmä bezpečnosť (kontrola vstupov do budov, letísk, používanie bankomatov), monitorovanie (bezpečnostné kamery upozornia na prítomnosť hľadaných zločincov, kontrola bdelosti vodiča v aute), kontrola identity občanov a vyhľadávanie podozrivých v policajných databázach. Najnovšie je rozpoznávanie tvári tiež populárne v zábavnom priemysle, či už ako súčasť herných systémov (napr. konzola Xbox One umožňuje snímať tvár pomocou zariadenia Kinect a projektovať výrazy tváre v sieťových hrách [20]), robotov schopných komunikovať s ľuďmi, alebo u sociálnych sietí umožňujúcich identifikovať známe osoby v nahraných fotografiách [38].

1.3 Proces rozpoznávania tváre

Pri rozpoznávaní ľubovoľného objektu v oblasti spracovania obrazu sa vo väčšine prístupov opakujú veľmi podobné kroky, ktoré je vždy potrebné vykonať. Podľa zjednodušenej schémy prevzatej z [10] a uvedenej na obrázku 1.1 má všeobecný systém určený na rozpoznávanie tvári niekoľko častí:

- **vstupný obrázok** – vstupná fotografia alebo snímka z videozáznamu určená na rozpoznávanie,
- **predspracovanie** – upravenie vstupného obrázka do požadovaného formátu pre extrakčné metódy,
- **extrakcia charakteristických rysov** – identifikácia a extrakcia príznakov s najväčšou informačnou hodnotou do *vektora príznakov*,
- **databáza tvári** – obsahuje uložené vektory príznakov známych tvári,

- **klasifikácia** – porovnáva vyextrahovaný vektor príznakov z rozpoznávanej tváre s príznakmi uloženými v databáze tvárí a
- **výstup** – je výstupom klasifikácie informujúcim o tom, či ide o neznámu tvár alebo tvár známej osoby z databázy.



Obr. 1.1: Základné kroky všeobecného systému na rozpoznávanie tvárí.

Vstupný obrázok je možné vyhotoviť za použitia fotoaparátu alebo videokamery, pričom v druhom prípade je potrebné identifikovať snímok obsahujúci tvár na rozpoznávanie.

V štúdií [39] publikovanej v roku 1969 bolo zistené, že ľudia majú oveľa väčší problém s rozpoznávaním prevrátených tvárí ako s rozpoznávaním akýchkoľvek iných prevrátených objektov. Rovnako aj metódy určené na rozpoznávanie tvárí fungujú najefektívnejšie ak sú im predkladané snímky tvárí v normalizovanom tvare. Medzi základné operácie kroku predspracovania preto patrí detekcia tváre, orezanie a natočenie vstupného obrázka do normalizovaného tvaru. Ďalej je to často úprava veľkosti slúžiaca prevažne na zrýchlenie spracovania a zníženie pamäťovej náročnosti analýzy, normalizácia jasu, odstránenie šumu, prevedenie do odtieňov šedej farby a iné.

Predspracovaná snímka tváre slúži ako vstup do kroku extrakcie charakteristických príznakov, ktorý vytvorí vektor príznakov špecifický pre použitú metódu. V prípade vytvárania databázy známych tvárí počas extrakcie charakteristických rysov je následne tento identifikátor tváre uložený do databázy s informáciou, ktorej osobe daná tvár patrí. Ak už prebieha fáza rozpoznávania, je vektor príznakov určený na klasifikáciu.

V kroku klasifikácie je vektor príznakov rozpoznávanej tváre porovnaný so záznamami v databáze tvárí a hľadá sa záznam tváre, ktorý je čo najpodobnejší rozpoznávanej tvári. V prípade, že je takýto záznam identifikovaný, ide o známu tvár konkrétnej osoby z databázy, v opačnom prípade ide o neznámu tvár. Táto informácia je výstupom jadra rozpoznávacieho systému, na základe ktorej sa môže systém využívajúci toto rozpoznávanie rozhodnúť, či podľa jeho účelu používateľovi napríklad povolí alebo zamietne vstup, zobrazí údaje o osobe, požiadajú o ďalší snímok tváre a podobne.

1.4 Metódy extrakcie charakteristických rysov

Jedna z vecí, ktorá rozpoznávanie aj známej tváre mimoriadne komplikuje, je to, že snímka tváre pri rozpoznávaní nikdy nie je rovnaká ako snímka tváre, z ktorej boli extrahované charakteristické rysy osoby pridávanej do databázy.

Väčšina metód má vysokú úspešnosť rozpoznávania, ak sú snímky zachytené v kontrolovaných podmienkach. Avšak, v prípade, že nastane zmena osvetlenia, pózy alebo výrazu tváre snímanej osoby, schopnosť rozpoznávania mnohých metód výrazne klesá. Medzi ďalšie komplikácie tiež patrí starnutie, kozmetické líčenie, módné doplnky ako sú napr. okuliare, alebo premenlivé ochlpenie tváre.

Ideálna metóda pracuje s takými vyextrahovanými rysmi tváre, ktoré sú odolné voči všetkým vyššie spomínaným problémom. Príkladom aparátu s nadštandardnými rozlišovacími schopnosťami s ohľadom na tieto komplikácie je ľudský zrak a vnímanie [1].

Jedna z možných klasifikácií rozdeľuje metódy na extrakciu charakteristických rysov do 3 základných kategórií, a to podľa toho, aké typy príznakov využívajú [31]:

1. **holistické** – na rozpoznávanie je ako vstup využívaná tvár ako celok,
2. **lokálne** – na rozpoznávanie sú využívané lokálne príznaky v rôznych oblastiach tváre,
3. **hybridné** – využívajú kombináciu globálnych a lokálnych príznakov na docielenie úspešnejšieho rozpoznávania.

Ako už bolo uvedené v predchádzajúcom rozdelení, holistické metódy sú význačne tým, že na rozdiel od ostatných metód tvár nesegmentujú na menšie časti, ale považujú ju celú za vstup do svojich algoritmov. Snímka tváre je štandardne transformovaná spojením šedých hodnôt všetkých pixelov na jediný viacrozmerný vektor, čo umožní zachovať celostné informácie o rozložení a textúre zachytenej tváre na rozdiel od lokálnych metód. V súvislosti s týmto prístupom však vzniká problém uchovávanía potenciálne redundantných dát a spracovávanía vektorov o vysokej dimenzionalite, čo môže navyšovať požiadavky na pamäťový a výpočtový výkon systémov. Medzi významných predstaviteľov tejto kategórie patrí analýza hlavných komponent popísaná ďalej v podsekcii 1.4.1 a analýza lineárnych diskriminantov (známa aj pod pojmom *Fisherove tváre*) [11].

Prístupy z kategórie lokálnych metód využívajú namiesto jedného vstupu pozostávajúceho z celej tváre len jej vhodne vybrané časti. Vďaka tomu je zredukovaný základný problém holistických metód týkajúci sa vysokej dimenzionality. Prinášajú väčšie možnosti rozpoznávania ak je na snímke zachytená len časť tváre a v prípade zvolených dostatočne rôznorodých rysov napomáhajú lepšej segmentácii priestoru, s ktorým pracuje klasifikátor. Pre ich úspešné fungovanie je však mimoriadne dôležité zakomponovanie informácie o vzťahoch spracovaných častí k tvári ako celku. V tejto práci sú popísané dvaja predstavitelia tejto kategórie – metóda lokálnych binárnych vzorov v podsekcii 1.4.2 a histogramov orientovaných gradientov v podsekcii 1.4.3.

Hybridné metódy využívajú kombináciu globálnych (holistických) a lokálnych rysov, pričom dôležitý je najmä správny výber spôsobu nakombinovania týchto príznakov. Spoliehajú sa na myšlienku, že tieto dve rozličné typy vlastností sú nositeľmi komplementárnych informácií pre klasifikátor. Zástupcov tejto kategórie je v súčasnosti najmenej a patrí medzi ne napríklad metóda lokálnych pravdepodobnostných podpriestorov [18]. Avšak, z určitého hľadiska je za hybridné metódy (resp. ich špecifickú podmnožinu) možné považovať aj lokálne metódy pre ich zakomponovanie globálnej informácie medzi lokálne deskriptory.

1.4.1 Analýza hlavných komponent

V roku 1987 bola v [29] po prvýkrát publikovaná myšlienka reprezentovať ľudskú tvár s využitím analýzy hlavných komponent, a neskôr v roku 1991 bola aplikovaná na detekciu a rozpoznávanie tvárí [32]. Išlo o jednu z prvých fungujúcich technológií v oblasti rozpoznávania, ktorá slúžila ako východisko pre mnoho komerčných riešení na automatizované rozpoznávanie ľudských tvárí a dodnes je považovaná za metódu určujúcu minimálnu očakávanú úspešnosť fungovania systémov v tejto problematike [17].

Analýza hlavných komponent, po anglicky označovaná ako *Principal Component Analysis (PCA)*, je holistická metóda z oblasti štatistiky, ktorá aplikuje princípy z teórie vektorov na zredukovanie vysokodimenzionálneho vstupu pri zachovávaní väčšiny informačnej

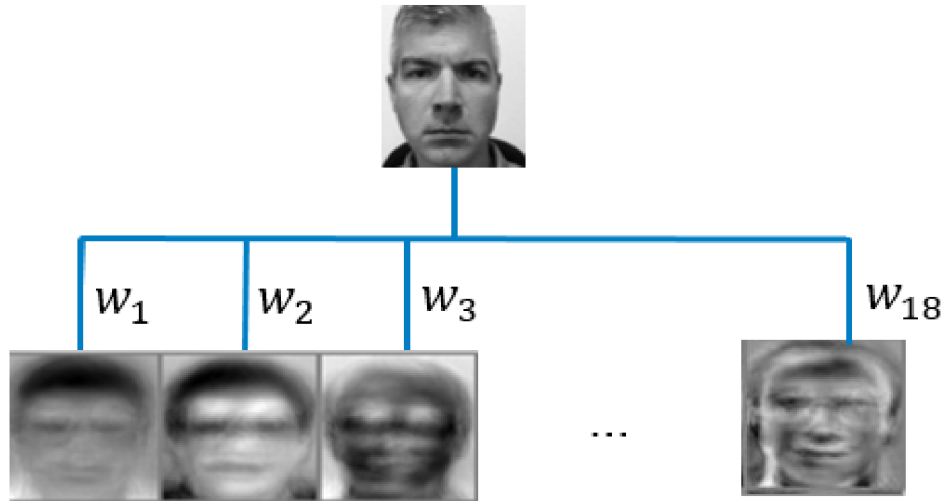
hodnoty. Na extrakciu relevantných informácií z tváre využíva zachytenie štatistických odchýlok medzi snímkami porovnávaných tvárí. Pretože táto technika pracuje tvármi reprezentovanými vektormi o nízkej dimenzionalite, rieši sa tiež základný problém na pamäťovú a výpočtovú náročnosť spomínaný v súvislosti holistickými metódami v sekcii 1.4. Medzi jej silné stránky patria scenáre, keď je dostupný malý počet tréningových snímkov niektorej z osôb, alebo v prípade, že sú tréningové dáta nerovnomerne rozložené v podpriestore osoby, ktorú reprezentujú [19].

Proces rozpoznávania sa v tejto metóde skladá z dvoch fáz – *inicializácie* a *rozpoznávania* [32].

Inicializácia metódy

Fáza inicializácie sa skladá z 3 krokov:

1. Získanie počiatočnej množiny tvárí (tzv. *tréningová množina*).
2. Výpočet *vlastných tvárí* (viď obr. 1.2) z tréningovej množiny definujúcej priestor tvárí, pričom je ponechaných len M vlastných tvárí podľa najvyšších *vlastných čísel*. Pri objavení nových tvárí sú tieto hodnoty aktualizované.
3. Výpočet váh pôvodnej tváre vzhľadom k jednotlivým vlastným tváram vo vzniknutom M -dimenzionálnom priestore projekciou pôvodnej tváre na tento tzv. *priestor vlastných tvárí*.



Obr. 1.2: Ukážka dekompozície fotografie tváre na množinu tzv. vlastných tvárí [14].

Nech $\Gamma_1, \Gamma_2 \dots \Gamma_M$ je tréningová množina snímkov tvárí, pričom tvár $\Gamma_m \in N^2$, kde N je šírka, resp. výška obrázka. Každý z týchto obrázkov potom predstavuje bod v N^2 -dimenzionálnom priestore tvárí. *Priemernú tvár* Ψ v tréningovej množine je možné vypočítať podľa nasledujúceho vzorca 1.1 a každá tvár Γ_i v tejto množine sa od priemernej tváre líši o vektor Φ_i podľa 1.2.

$$\Psi = \frac{1}{M} \sum_{n=1}^M \Gamma_n \quad (1.1)$$

$$\Phi_i = \Gamma_i - \Psi \quad (1.2)$$

Využitím analýzy hlavných komponent (tiež nazývanej aj *Karhunen-Loeveová transformácia*) je potom možné nájsť vektory, ktoré najviac prispievajú k variácii jednotlivých tvárí v celkovom priestore tvárí – *vlastné vektory* (resp. *vlastné tváre* keďže po zvizualizovaní pripomínajú ľudské tváre). Tieto vektory u_k , kde $|u_k| = N^2$, sú anglicky označované ako *eigenvectors* (resp. *eigenfaces*). Ide o ortonormálne vektory, ktoré vzniknú lineárnou kombináciou tvárí $\Gamma_1 \dots \Gamma_M$ a sú vlastnými vektormi kovariančnej matice C korešpondujúcich s obrázkami pôvodných tvárí. Kovariančná matica C značiaca závislosť medzi jednotlivými prvkami Φ_n vektora A normalizovaných tvárí, kde $A = [\Phi_1, \Phi_2 \dots \Phi_M]$, je zadenovaná podľa rovnice 1.3, pričom X^T označuje transponovaný vektor X .

$$C = A A^T = \frac{1}{M} \sum_{n=1}^M \Phi_n \Phi_n^T \quad (1.3)$$

Bohužiaľ má táto matica $N^2 \times N^2$ prvkov, a preto nie je výpočtovo prijateľné zisťovať N^2 vlastných vektorov a vlastných hodnôt. Avšak, keďže pre potreby tejto metódy je užitočných vlastných vektorov len M , čo je mnohonásobne nižší počet ako veľkosť priestoru tvárí, môžeme riešiť namiesto pôvodnej matice maticu $L = A^T A$ len o rozmeroch $M \times M$. Maticu L zostavíme podľa rovnice 1.4 a v nej nájdeme M vlastných vektorov v_l . Z týchto vlastných vektorov v_l matice L a normalizovaných tvárí Φ_n vypočítame ďalšou rovnicou 1.5 hľadané vlastné tváre u_l .

$$L_{mn} = \Phi_m^T \Phi_n \quad (1.4)$$

$$u_l = \sum_{k=1}^M v_{lk} \cdot \Phi_k, \quad \text{kde } l = 1 \dots M \quad (1.5)$$

Vlastné číslo λ_k vlastného vektora u_k určujúce významnosť, ktorou daná vlastná tvár prispieva k variácii tvárí z trénovacej množiny je nakoniec možné vypočítať podľa vzorca 1.6.

$$\lambda_l = \frac{1}{M} \sum_{n=1}^M (u_l^T \Phi_n)^2 \quad (1.6)$$

Po zoradení vlastných tvárí u_l podľa pridružených vlastných čísel λ_l môžeme pôvodný priestor tvárí zredukovať na taký počet dimenzií, koľko tvárí s najväčším vlastným číslom si zvolíme zachovať. Tieto budú slúžiť ako báza M' pre rozpoznávanie.

Výpočet váh, ktorými vlastné tváre prispievajú k variácii pôvodných tvárí z trénovacej množiny, je potom analogický ako výpočet váh pre ľubovoľnú novopredloženú tvár na rozpoznávanie a je uvedený v nasledujúcej časti.

Rozpoznávanie tváre

Po dokončení fáze inicializácie je možné rozpoznať predloženú tvár v nasledujúcich 3 krokoch:

1. Výpočet množiny váh predloženej tváre vzhľadom na M vlastných tvárí postupným projektovaním rozpoznávanej tváre na každú z vlastných tvárí v priestore tvárí.

2. Kontrola, či je predložená snímka skutočnou tvárou – overenie dostatočnej blízkosti k časti priestoru vlastných tvári, v ktorom sa uložené vlastné tváre nachádzajú.
3. Ak ide o tvár, klasifikuj množinu vypočítaných váh ako množinu patriacu *známej* alebo *neznámej* osobe.

Nech Γ je obrázok predkladaný na rozpoznanie. Na jeho projekciu do priestoru vlastných tvári, a teda na určenie váh ω_k , ktorými tieto zložky u_k prispievajú k variácii predkladanej tváre vzhľadom na priemernú tvár Ψ , je možné použiť vzorec 1.7.

$$\omega_k = u_k^T(\Gamma - \Psi) , \quad \text{kde } k = 1 .. M' \quad (1.7)$$

Výsledkom je potom vektor váh $\Omega^T = [\omega_1, \omega_2, \dots, \omega_{M'}]$, ktorý je možný použiť na klasifikáciu vzhľadom na vektory váh tvári z trénovacej množiny. Ako potrebnú metriku je možné využiť Euklidovskú vzdialenosť ϵ_k^2 uvedenú v 1.8, kde Ω predstavuje vektor váh rozpoznávanej tváre. Váhový vektor Ω_k vznikne spriemerovaním váhových vektorov tvári trénovacej množiny patriacej do k -tej triedy, kde každá trieda predstavuje jednu známou osobu.

$$\epsilon_k^2 = \|(\Omega - \Omega_k)\|^2 \quad (1.8)$$

Ak je pre rozpoznávanú tvár vzdialenosť od najbližšej triedy menšia ako zvolený prah θ_ϵ , je považovaná za známou tvár a je možné prepočítať váhový vektor tejto triedy, v opačnom prípade za neznámu.

Je však potrebné zaviesť ešte metriku ϵ_k uvedenú na 1.9, ktorá v prípade neznámych tvári dokáže podľa prahu θ_ϵ o danej rozpoznávanej snímke aspoň prehlásiť, že ide skutočne o snímku tváre. Jej úlohou je určiť veľkosť odchýlky pôvodnej normalizovanej tváre Φ od jej spätnej projekcie Φ_f vypočítanej z vektora váh Ω rozpoznávanej snímky, nachádzajúcom sa v podpriestore vlastných tvári, do pôvodného priestoru tvári.

$$\begin{aligned} \epsilon_k &= \|(\Phi - \Phi_f)\|^2 , \quad \text{kde} \\ \Phi &= \Gamma - \Psi \\ \Phi_f &= \sum_{i=1}^{M'} \omega_i u_i \end{aligned} \quad (1.9)$$

1.4.2 Lokálne binárne vzory

Metóda *lokálnych binárnych vzorov*, označovaná skratkou *LBP* (angl. *Local Binary Patterns*), je metóda na extrakciu charakteristických rysov predstavená v roku 1996 v [21] určená na analýzu textúry. Jednoduchosť tejto metódy umožňuje veľmi rýchlu extrakciu príznakov a vďaka modifikáciám predstaveným v tejto podsekcii si našla uplatnenie v mnohých ďalších oblastiach analýzy obrazu.

LBP operátor je definovaný ako popisná miera textúry, ktorá je invariantná voči monotónnym transformáciám v odtieňoch šedej farby. Kvôli svojej rozlišovacej sile a výpočtovej nenáročnosti bol tento operátor použitý v mnohých oblastiach vrátane vizuálnej kontroly, vyhľadávani obrázkov, analýze snímok v biomedicíne, analýze pohybu a ďalších [15].

Využitie LBP na reprezentáciu rysov tváre vychádza z myšlienky, že tvár môže byť videná ako kompozícia mikrovzorov (plochy, body, čiary alebo hrany) invariantných voči monotónnym transformáciám v odtieňoch šedej farby, ktoré sú veľmi dobre opísateľné takýmto operátorom.

V oblasti biometrie tváre priniesla táto metóda nadštandardné výsledky pri detekcii tváre alebo rozpoznávaní emócií [2]. Pri rozpoznávaní tváří je zasa vcelku odolná voči rôznym svetelným podmienkam, starnutiu subjektov a samotným výrazom tváre [25].

Princíp fungovania

Keďže deskriptor textúry zvykne priemerovať nad analyzovanými oblasťami, na rozdiel od popisu napr. tkaniny, kde je vhodný holistický prístup, je pri popise tváre dôležité zachovanie priestorových vzťahov medzi informáciami. Preto je zvolený spôsob, pri ktorom je vstupný obrázok rozdelený na malé oblasti, z ktorých sú nezávisle na sebe získané lokálne deskriptory textúry. Z nich je potom spojením vytvorený globálny deskriptor tváre [2].

Takéto deskriptory potom zachytávajú popis tváre na troch úrovniach:

- binárne vzory obsahujú informácie o vzoroch na úrovni pixelov,
- vzory agregované na malej ploche opisujú lokálnymi deskriptormi jednotlivé oblasti tváre a
- spojené lokálne deskriptory charakterizujú tvár na globálnej úrovni.

Lokálny binárny vzor je definovaný ako zoradená množina binárnych porovnaní intenzít pixelov medzi stredovým a susediacimi pixelmi. Práve vďaka zohľadňovaniu len znamienok rozdielov v odtieňoch pixelov sa dosahuje invariancia voči monotónnym transformáciám v odtieňoch šedej farby.

Úlohou LBP operátora je vypočítať binárny reťazec reprezentujúci zadaný pixel vzhľadom na zvolené okolie. Pri analýze tkaniny bolo v [22] využívané Moorove okolie, avšak v ďalších štúdiách bolo rozšírené na ľubovoľné kruhové okolie o P bodoch vo vzdialenosti R od stredového pixelu a takýto operátor nesie označenie $LBP_{P,R}$.

Pozícia p -teho pixelu v kruhovom okolí pixelu c sa dá vypočítať podľa vzorcov 1.10 [25]:

$$\begin{aligned}x_p &= x_c + R \cdot \cos\left(\frac{2\pi p}{P}\right), \\y_p &= y_c + R \cdot \sin\left(\frac{2\pi p}{P}\right)\end{aligned}\tag{1.10}$$

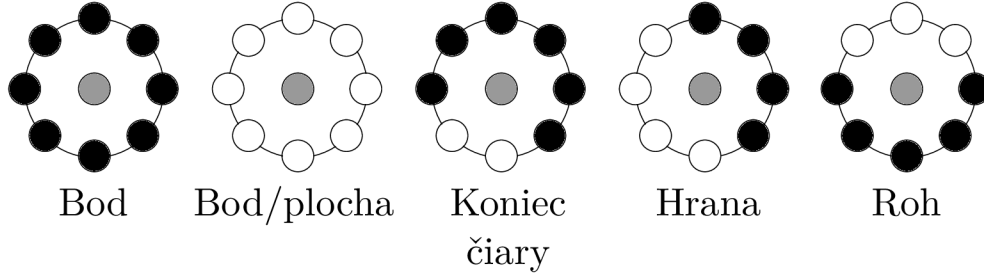
Na každý bod (x_i, y_i) obrázka o rozmeroch $M \times N$, kde $x_i \in \{R + 1..N - R\}$, resp. $y_i \in \{R + 1..M - R\}$, sa dá aplikovať operátor $LBP_{P,R}$ podľa vzorca 1.11, pričom g_j udáva hodnotu intenzity daného pixelu v odtieňoch šedej farby [25]:

$$\begin{aligned}LBP_{P,R}(x_c, y_c) &= \sum_{p=0}^{P-1} 2^p s(g_p - g_c), \quad \text{kde} \\s(v) &= \begin{cases} 1, & \text{pre } v \geq 0 \\ 0, & \text{pre } v < 0 \end{cases}\end{aligned}\tag{1.11}$$

Výsledný reťazec bitov je považovaný za cyklický. Ak takýto vzor obsahuje najviac dva bitové prechody medzi znakmi 0 a 1, je označovaný ako *uniformný*. Ukázalo sa, že uniformné vzory sú nositeľmi oveľa vyššej informačnej hodnoty ako ostatné typy vzorov a pri operátore $LBP_{8,2}$ predstavujú 90% vyskytujúcich sa vzorov.

Použitie uniformných vzorov na popis obrázka tiež prináša dve zásadné výhody:

1. šetrí pamäť – pri uniformnom operátore $LBP_{P,R}^{u2}$, predstavenom ďalej v texte, je možných len $P(P - 1) + 2$ vzorov, pričom všetkých možných vzorov pri bežnom operátore je až 2^P ,
2. deteguje len dôležité lokálne textúry ako body, konce čiar, hrany alebo rohy (viď obrázok 1.3).



Obr. 1.3: Príklady uniformných lokálnych vzorov detegovaných operátorom $LBP_{P,R}^{u2}$ [25].

Deskriptory lokálnej úrovne sú reprezentované vo forme histogramov. Všetky neuniformné vzory sú v histograme zaradené pod spoločným označením a pre každý uniformný vzor je vytvorené samostatné označenie. Počet vzorov s dvoma prechodmi je $P(P - 1)$, pričom s dvoma homogénnymi vzormi a jedným spoločným označením pre neuniformné vzory činí celkový počet rôznych možných označení histogramu $P(P - 1) + 3$ [22]. LBP operátor, ktorý dokáže takto priradovať označenia pixelov sa nazýva *uniformný* a je v texte značený ako $LBP_{P,R}^{u2}$.

V každej oblasti obrázka je možné agregovať výsledky operátora do lokálneho histogramu podľa vzorca 1.12, kde $i = 0..(|H| - 1)$ a $|H|$ je počet možných označení operátora $LBP_{P,R}^{u2}$:

$$H_i = \sum_{x,y} I \{ LBP_{P,R}^{u2}(x,y) = i \} , \quad \text{kde} \quad (1.12)$$

$$I\{A\} = \begin{cases} 1, & \text{ak } A \text{ je pravda} \\ 0, & \text{ak } A \text{ je nepravda} \end{cases}$$

Porovnávanie globálnych deskriptorov

Aj keď sa pri analýze tkanín pomocou LBP operátora štandardne využíva logaritmickej miera vzdialenosti, podľa [1] je jej výkonnosť pri malých oblastiach využívaných pri analýze tváre slabá. Oveľa lepšie výsledky prináša χ^2 -rozdelenie pravdepodobnosti, ktoré je definované podľa vzorca 1.13, kde S označuje obrázok určený na rozpoznanie, M tvár v databáze, H je vektor lokálneho histogramu, K počet oblastí a $S_{i,j}$ (resp. $M_{i,j}$) sú počty výskytov i -teho vzoru v oblasti j rozpoznávaného (resp. uloženého) obrázka. Pretože oblasti okolo očí alebo úst môžu mať významnejšiu popisnú hodnotu ako napr. oblasti čela, w_j značia váhy jednotlivých regiónov. Čím je výsledná vzdialenosť menšia, tým sú porovnávané obrázky podobnejšie [25].

$$\chi^2(S, M) = \sum_{j=1}^K w_j \left(\sum_{i=1}^{|H|} \frac{(S_{i,j} - M_{i,j})^2}{S_{i,j} + M_{i,j}} \right) \quad (1.13)$$

1.4.3 Histogramy orientovaných gradientov

Histogramy orientovaných gradientov, skrátene *HOG*, je lokálna metóda slúžiaca na extrakciu charakteristických rysov rozpoznávaného objektu podobná predchádzajúcej metóde lokálnych binárnych vzorov. Bola predstavená v roku 2005 N. Dalalom a B. Triggom za účelom detekcie chodcov na statických fotografiách [8] a neskôr rozšírená aj na ďalšie typy objektov. Vlastnosti tejto metódy sú podobné ako u metódy LBP, takže rovnako prináša nadštandardné výsledky pri rozpoznávaní tvárí osôb v rôznych svetelných podmienkach, pri ich postupnom stárnutí, alebo pri premenlivých výrazoch tváre [9].

HOG deskriptor je lokálna popisná jednotka orientácií gradientov v časti obrázku invariantná voči pootočeniu a zmenám v osvetlení. Jeho funkčnosť je založená na myšlienke, že vzhľad a tvar časti objektu môže byť charakterizovaný smerovaním hrán a intenzitami gradientov rozloženými v danej oblasti [27].

Ako je uvedené v [34], gradient je vektor, ktorého zložky merajú prudkosť zmeny intenzity pixelov v dimenzii x a y . Po vypočítaní gradientov G_x a G_y v horizontálnom, resp. vertikálnom smere, je následne možné z vektora gradientov G určiť pre každý bod obrázka veľkosť $|\nabla I|$ a uhol θ gradientu v danom bode podľa vzorcov 1.14 a 1.15:

$$|\nabla I(x, y)| = \sqrt{G_x^2 + G_y^2} \quad (1.14)$$

$$\theta(x, y) = \arctan\left(\frac{G_y}{G_x}\right) \quad (1.15)$$

Podobne ako v prípade lokálnych binárnych vzorov z podsekcie 1.4.2 používame aj v prípade tejto metódy na reprezentáciu deskriptora lokálnej oblasti histogram. V tomto prípade je histogram 1.16 zložený z $|H|$ košov predstavujúcich intervaly uhlov rovnomerne rozdeleného rozsahu $\langle -\pi, \pi \rangle$, pričom do každého koša H_i pre $i = 0..(|H| - 1)$ zaraďujeme veľkosti gradientov tých pixelov, ktoré majú uhol v očakávanom rozsahu. Na to, aby bol histogram invariantný k osvetleniu a kontrastu, je na záver potrebné normalizovať vektor daného histogramu H na \hat{H} podľa 1.17, kde ϵ je malá kladná hodnota.

$$H_i = \sum_{x,y} |\nabla I(x, y)| \{\theta(x, y) \in \langle \varphi_i, \varphi_{i+1} \rangle\} \quad (1.16)$$

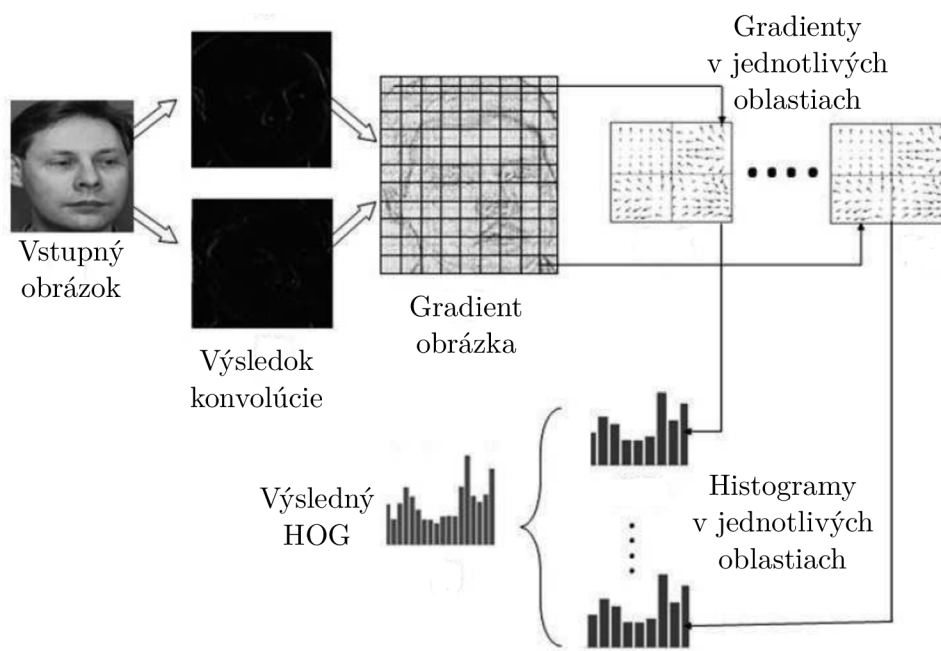
$$\hat{H} = \frac{H}{\sqrt{\|H\|^2 + \epsilon^2}} \quad (1.17)$$

Nakoniec sú histogramy lokálnych oblastí spojené do globálneho deskriptora tváre. Celý proces je zosumarizovaný v obrázku 1.4 [27].

Jedným z možných prístupov na porovnanie podobnosti globálnych deskriptorov rozpoznávanej tváre a známej tváre z databázy je napríklad χ^2 -rozdelenie pravdepodobnosti uvedené vo vzorci 1.13 predchádzajúcej podsekcii.

1.5 Metriky vyhodnocovania rozpoznávacích systémov

Ako už bolo spomínané na začiatku tejto kapitoly, rozpoznávanie tvárí zahŕňa dve odlišné úlohy, a to verifikáciu a identifikáciu subjektov. Aj keď systémy slúžiace na rozpoznávanie môžu implementovať rovnaký algoritmus na extrakciu charakteristických rysov rozpoznávanej osoby a následné porovnávanie so záznamami v databáze, kvôli rozličným prípustným



Obr. 1.4: Jednotlivé kroky procesu extrakcie charakteristických rysov pomocou metódy HOG [27].

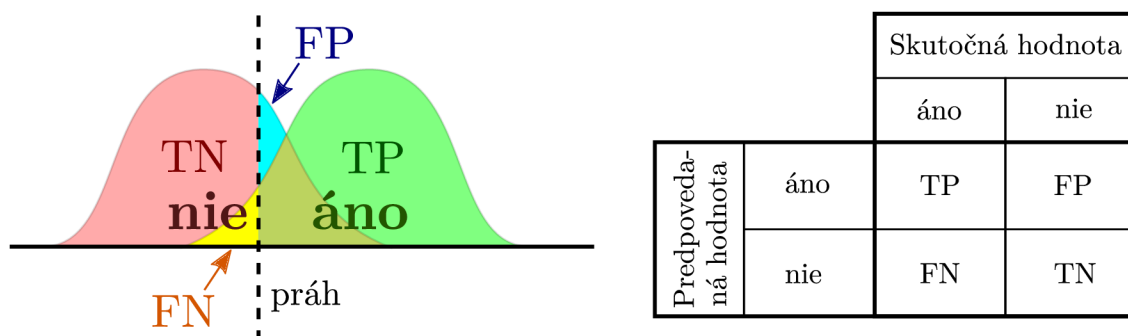
hodnotám na výstupe je potrebné zdefinovať samostatné formy na reprezentáciu úspešnosti systémov slúžiacich na dané úlohy.

1.5.1 Vyhodnocovanie verifikácie

Vo verifikácii môže systém odpovedať jednou z dvoch možností – *áno*, ak je systém presvedčený, že predkladaná identita patrí danej osobe, alebo *nie* v opačnom prípade. Keďže ide o binárnu klasifikáciu, môžeme na zápis výsledkov pozorovania použiť tzv. **kontingenčnú tabuľku** znázornenú na obrázku 1.5, v ktorej význam jednotlivých polí je nasledovný:

- **Skutočne pozitívne** (angl. *True positive*, **TP**) – ak predkladaná identita patrí osobe a systém správne odpovedá *áno*
- **Skutočne negatívne** (angl. *True negative*, **TN**) – ak predkladaná identita nepatrí osobe a systém správne odpovedá *nie*
- **Falošne pozitívne** (angl. *False positive*, **FP**) – tzv. *Chyba I. typu*, ak predkladaná identita nepatrí osobe a systém nesprávne odpovedá *áno*
- **Falošne negatívne** (angl. *False negative*, **FN**) – tzv. *Chyba II. typu*, ak predkladaná identita patrí osobe a systém nesprávne odpovedá *nie*

Z týchto empiricky získaných hodnôt je možné podľa vzťahov 1.18 vypočítať ďalšie metriky, z ktorých pre naše experimenty budú prínosné nasledujúce:



Obr. 1.5: Kontingenčná tabuľka a vzťah jej hodnôt k nastaveniu prahu.

- **Miera skutočne pozitívnych detekcií** (angl. *True positive rate*, **TPR**) – označuje podiel nameraných skutočne pozitívnych odpovedí z celkového možného počtu verifikácií, ktoré mali byť označené ako *áno*
- **Miera falošne pozitívnych detekcií** (angl. *False positive rate*, **FPR**) – označuje podiel meraní nesprávne označených ako *áno* spomedzi všetkých možných verifikácií, ktoré mali byť označené ako *nie*
- **Presnosť** (angl. *Accuracy*) – podiel správnych odpovedí z celkového počtu meraní

$$\begin{aligned}
 TPR &= \frac{TP}{TP + FN} & FPR &= \frac{FP}{FP + TN} \\
 presnosť &= \frac{TP + TN}{TP + FN + FP + TN}
 \end{aligned}
 \tag{1.18}$$

Získavanie odpovede v procese verifikácie je založené na vypočítaní skóre, ktoré je prahovaním klasifikované do jednej z dvoch možných odpovedí. Je zrejme, že posúvaním prahu sa menia počty hodnôt nameraných v kontingenčnej tabuľke. Tieto zmeny je možné vizualizovať pomocou tzv. **ROC krivky** (angl. *Receiver Operating Characteristic*), ktorá zachytáva vzťah medzi TPR a FPR pri všetkých prípustných hodnotách prahu. Krivka presne odpovedajúceho systému potom inklinuje k ľavému hornému rohu priestoru grafu, takže vyhodnocovanie je možné vykonávať na základe obsahu plochy pod grafom (tzv. *AUC – Area Under Curve*).

1.5.2 Vyhodnocovanie identifikácie na uzavrenej množine subjektov

V prípade identifikácie, na rozdiel od verifikácie klasifikujeme do ľubovoľného počtu tried. Konkrétne, v prípade systému s tréningovou množinou o počte N rôznych osôb je možné klasifikovať buď do N tried, ak predpokladáme, že na vstupe sa objavia len známe osoby z tréningovej množiny, alebo do $N + 1$ tried, kde nadbytočná trieda zastrešuje potenciálne neznáme osoby. Prvý prípad sa označuje aj ako identifikácia na **uzavrenej množine** a druhý prípad na **otvorenej množine** subjektov.

Ako je uvedené v [23], na vyhodnotenie systému identifikujúceho na uzavrenej množine subjektov je najvhodnejšie zodpovedanie otázky „*Nachádza sa správna odpoveď medzi k-najlepšími porovnaniami?*“ Na identifikáciu osoby musí totiž systém porovnať vstupnú

snímku so všetkými osobami v databáze, pričom vyberie tú najpodobnejšiu vzhľadom na jeho implementáciu.

Ak zoradíme všetky osoby $b \in \mathcal{B}$ v databáze podľa skóre podobnosti od najlepšej zhody po najhoršiu, na prvej pozícii $pos(1)$ je umiestnená identita osoby uvedenej vo výsledku identifikácie, po nej na pozícii $pos(2)$ nasleduje druhá najpravdepodobnejšia identita, atď. **Identifikačná miera pre stupeň k** (angl. *Identification Rank*) je potom možné vypočítať podľa 1.19 ako podiel počtu správne identifikovaných osôb umiestnených na pozícii k alebo lepšej vzhľadom na celkový počet identifikácií $|\mathcal{U}|$:

$$IR(k) = \frac{|(b|\forall b \in \mathcal{B} : pos(k) \leq k)|}{|\mathcal{U}|} \quad (1.19)$$

Zaznačenie identifikačnej miery $IR(k)$ do grafu vzhľadom na k potom vytvorí neklesajúcu kumulatívnu funkciu s názvom **CMC krivka** (angl. *Cumulative Match Characteristic*). Je zrejmé, že rovnako ako v prípade ROC krivky má úspešnejší systém väčšiu plochu pod grafom.

Kapitola 2

Návrh riešenia

Táto kapitola využíva teoretické poznatky prezentované v predchádzajúcej kapitole s cieľom dekomponovania jednotlivých metód na konkrétne postupy, ktoré bude možné použiť pri implementácii aplikácie zameranej na experimentovanie s identifikačnými úlohami na uzavrenej množine osôb. Tento obsah je uvedený v prvých dvoch sekciách 2.1 a 2.2. V druhej časti 2.3 je analýza problémov, ktoré prináša realizácia jednotlivých krokov, a návrh ich efektívnych riešení.

2.1 Porovnávanie vektorov charakteristických rysov

Kľúčovým prvkom v procese rozpoznávania je spôsob porovnania už vyextrahovaných charakteristických rysov, resp. ich vektorov. Od jeho vhodného výberu závisí následne v identifikačnej úlohe schopnosť vybrať najpodobnejšiu šablónu z databázy alebo v prípade verifikačnej úlohy zasa to, či je verifikovaná snímka dostatočne podobná zvolenej šablóne.

Naskytá sa teda otázka, akým spôsobom je vhodné vyjadriť podobnosť dvoch vektorov charakteristických rysov. Pri hľadaní odpovede sa môžeme inšpirovať teóriou metrických priestorov a za spôsob vyjadrenia podobnosti môžeme považovať vzdialenosť týchto matematických objektov v nejakom priestore. Len v prípade, že sú dva porovnávané objekty identické, má ich vzdialenosť hodnotu rovnú číslu 0 a s narastajúcimi rozdielmi medzi nimi sa zväčšuje aj hodnota vzdialenosti medzi týmito objektami.

Princíp identifikácie je teda zrejmý – na určenie najpodobnejšej šablóny je potrebné porovnať vyextrahovaný vektor charakteristických rysov rozpoznávanej snímky s každou šablónou v databáze a zvoliť šablónu s najmenšou vzdialenosťou od vyextrahovaného vektora. V prípade verifikácie je potrebné zasa vypočítať vzdialenosť vyextrahovaného vektora k určenej šablóne reprezentujúcej konkrétnu identitu a zistiť, či je táto vzdialenosť menšia ako hodnota stanovená nejakým prahom.

V prípade analýzy hlavných komponent je reprezentantom vektora charakteristických rysov jednorozmerný vektor váh. Je tvorený reálnymi číslami, ktorých počet je závislý na použitom počte vlastných tvárí určených na rozpoznávanie. Ako už bolo uvedené v podsekcii 1.4.1 predchádzajúcej kapitoly, analýza hlavných komponent pracuje s ideou priestoru definovaného detegovanými vlastnými tvármi, pričom váhový vektor označuje umiestnenie bodu v tomto priestore. Na výpočet vzdialenosti medzi dvoma váhovými vektormi vlastných tvárí sa preto ponúka ako najvhodnejšia Euklidovská metrika, ktorá pomocou Pytagorovského teorému určuje priamu vzdialenosť medzi dvoma bodmi v n -rozmernom

priestore s ortogonálnou bázou. Výpočet tejto vzdialenosti je vyjadrený vzorcom 2.1, kde ω_{ik} a ω_{jk} sú k -te prvky dvoch váhových vektorov Ω_i a Ω_j o dĺžke n prvkov.

$$d(\Omega_i, \Omega_j) = \sqrt{\sum_{k=1}^n (\omega_{ik} - \omega_{jk})^2} \quad (2.1)$$

Pri metódach LBP a HOG je vektor charakteristických rysov reprezentovaný histogramom. Ten už nepredstavuje konkrétny bod v multidimenzionálnom priestore, ale distribúciu dát, konkrétne vzorov identifikovaných v konkrétnom regióne analyzovaného obrázka. Je preto vhodné zvoliť metriku na pravdepodobnostnej báze, akou je napríklad vzdialenosť χ^2 popisovaná na konci podsekcie 1.4.2 [6] a používať ju v kombinácii s normalizovanými histogramami.

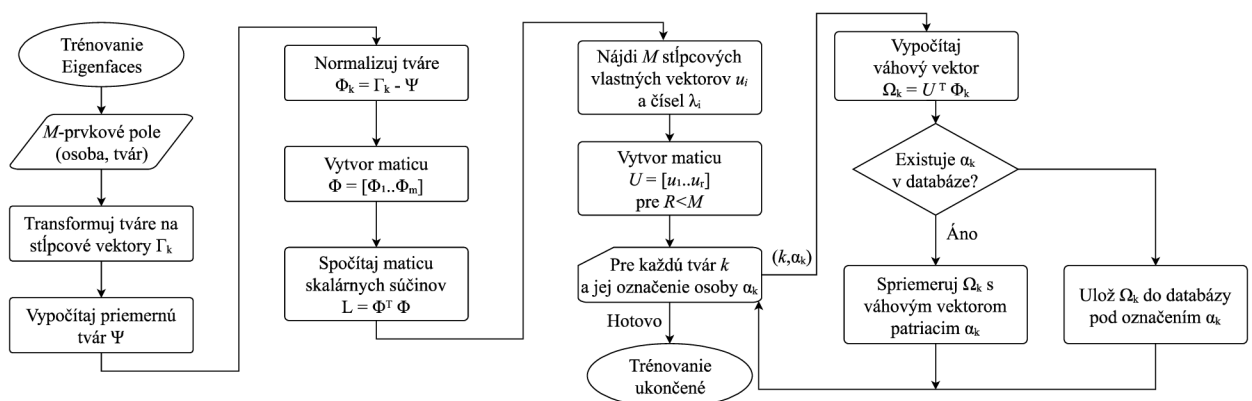
2.2 Realizácia extrakčných metód

V tejto sekcii sú zachytené konkrétne kroky potrebné pre realizáciu všetkých 3 metód, ktoré boli predstavené v predchádzajúcej kapitole, vo forme vývojových diagramov spolu so stručným slovným popisom. Pre každú metódu je prezentovaná fáza extrakcie charakteristických rysov pomocou danej metódy a aj samotný proces rozpoznávania predloženej snímky.

2.2.1 Analýza hlavných komponent

Problém rozpoznávania tváří metódou analýzy hlavných komponent bol predstavený v podsekcii 1.4.1 predchádzajúcej kapitoly 1. Obe jeho fázy je možné previesť do postupov jednoznačne vyjadrených vývojovými diagramami.

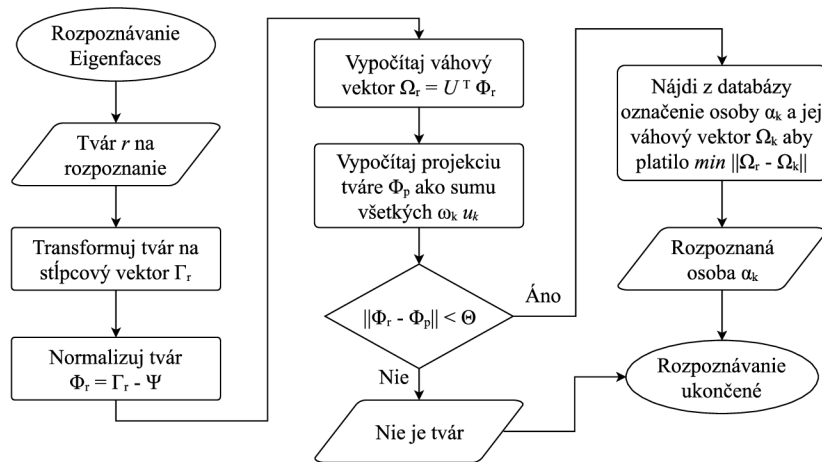
Fáza extrakcie charakteristických rysov, ktorej výsledkom je vytvorenie šablón známych osôb v databáze, je zobrazená vo vývojovom diagrame na obrázku 2.1. Na vstupe očakáva M dvojíc (*osoba*, *tvár*), kde *osoba* je označenie osoby, ktorej patrí už predspracovaná snímka tváre *tvár*. Pre každú osobu môže existovať na vstupe ľubovoľný počet snímok tváří.



Obr. 2.1: Vývojový diagram algoritmu na vytvorenie šablón v databáze metódou analýzy hlavných komponent zo vstupnej množiny tváří a korešpondujúcich označení osôb.

Každá tvár je zo vstupnej matice transformovaná na jeden stĺpcový vektor Γ_k , z ktorých je následne vypočítaná priemerná tvár Ψ . Tváre sú ďalej normalizované na Φ_k odpočítaním priemernej tváre, spojené do matice Φ , nad ktorou je spočítaná matica skalárnych súčinov L . Keďže vzniknutá matica je symetrická, je možné použiť napríklad Jacobiho iteračnú metódu na výpočet M vlastných vektorov u_i a vlastných čísel λ_i . Prvých R vlastných vektorov s najvyššími vlastnými číslami sa uloží do matice U . Následne je pre každú tvár k vypočítaný jej váhový vektor Ω_k a ten je buď vložený do databázy ako nový záznam osoby s označením α_k , alebo spriemerovaný s už existujúcim váhovým vektorom danej osoby v databáze.

Pri rozpoznávaní je potom predložená neznáma snímka a podľa algoritmu na vývojovom diagrame 2.2 sa spôsobom obdobným ako v prípade fázy extrakcie charakteristických rysov vypočíta váhový vektor Ω_r predloženej snímky s využitím už predpočítaného vektora vlastných tvárí U . Tento vektor váh je ďalej možné opätovne aplikovať na vektory vlastných tvárí za účelom získania spätnej projekcie tváre Φ_p . Ak sa normalizovaná tvár Φ_r líši od jej spätnej projekcie pomocou vlastných tvárí o viac ako je heuristicky zvolený prah Θ , na vstupnej snímke nie je tvár a rozpoznávanie končí. V prípade, že podobnosť je dostatočná, hľadá sa váhový vektor Ω_k , ktorý je najbližšie váhovému vektoru rozpoznávanej osoby a po jeho identifikovaní je na vstupnej snímke rozpoznaná osoba s označením α_k .

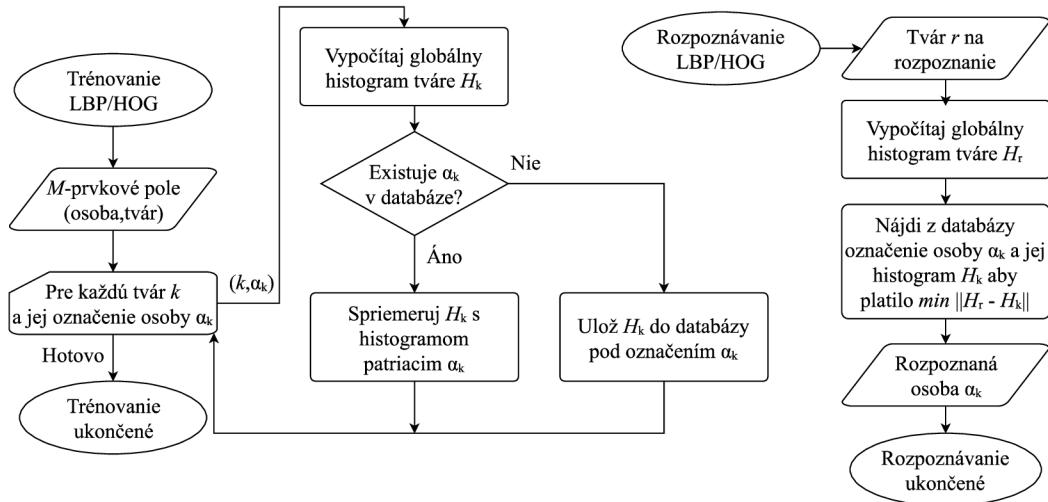


Obr. 2.2: Vývojový diagram algoritmu na rozpoznávanie zadanej tváre metódou analýzy hlavných komponent.

2.2.2 Metódy lokálnych binárnych vzorov a histogramov orientovaných gradientov

Zvyšné dve metódy predstavené v podsekciami 1.4.2 a 1.4.3 teoretickej kapitole 1 zdieľajú spoločný postup znázornený na obrázku 2.3 na extrakciu charakteristických rysov aj použitie na rozpoznávanie zadanej tváre. Jediný krok, v ktorom sa odlišujú, je spôsob, akým vytvárajú lokálne histogramy zo spracovávanej snímky, ktoré potom slúžia ako zložky globálneho histogramu danej snímky.

Fáza rozpoznávania začína rovnako ako v prípade analýzy hlavných komponent predložením poľa snímok tvárí a k nim priradeným označeniam osôb. Pre každú snímku tváre je vytvorený globálny histogram popisujúci túto snímku a ten je buď vložený do databázy ako



Obr. 2.3: Vývojové diagramy algoritmov na vytvorenie šablón v databáze pomocou metód lokálnych binárnych vzorov a histogramov orientovaných gradientov a ich následné použitie na rozpoznanie zadanej tváre.

nový záznam pod označením korešpondujúcej osoby, alebo spriemerovaný s už existujúcim záznamom v databáze.

Pri aplikácii týchto metód na rozpoznanie predloženej tváre je z tváre vytvorený globálny histogram H_r , ktorý je porovnávaný s histogramami v databáze a hľadá sa histogram H_k , ktorý je najbližšie histogramu rozpoznávanej osoby. Označenie osoby α_k , ktorej patrí nájdený histogram, je potom identifikovanou osobou k na rozpoznávanom snímku.

Vytvorenie globálneho histogramu je v prípade oboch metód zrealizované spojením lokálnych histogramov $L_{1..N}$ postupne získaných z N regiónov obrázka. Spôsob získania lokálnych histogramov jednotlivými metódami je zosumarizovaný v nadchádzajúcej časti práce.

Pri metóde lokálnych binárnych vzorov získame lokálny histogram daného regiónu v týchto 3 krokoch:

1. Vytvoríme lokálny histogram L_i s $|L|$ košmi inicializovanými na hodnotu 0, kde podľa 1.4.2:
 - Jeden koš reprezentuje všetky neuniformné binárne reťazce,
 - zvyšných $|L| - 1$ košov reprezentuje jednotlivé uniformné reťazce.
2. Pre každý bod (x, y) regiónu i :
 - (a) Získame vzor bodu reprezentovaný binárnym reťazcom postupným porovnávaním hodnoty intenzity stredového bodu s hodnotami intenzít v jednotlivých bodoch kruhového okolia, kde 1 reprezentuje stredovú intenzitu menšiu ako intenzita suseda a 0 opačný výsledok.
 - (b) Získaný reťazec považujeme za cyklický a k hodnote zodpovedajúceho košu histogramu L_i pripočítame 1.
3. Histogram L_i normalizujeme.

V prípade histogramov orientovaných gradientov môžeme pre každý región zrealizovať získavanie lokálneho histogramu v nasledujúcich 3 krokoch:

1. Vytvoríme lokálny histogram L_i s $|L|$ košmi inicializovanými na hodnotu 0, kde každý koš reprezentuje rozsah:
 - (a) $2\pi/|L|$ radiánov v prípade zatried'ovania nezmenených hodnôt uhlov,
 - (b) alebo $\pi/|L|$ radiánov v prípade zatried'ovania len absolútnych hodnôt uhlov.
2. Pre každý bod (x, y) regiónu i :
 - (a) Určíme horizontálnu zložku G_x a vertikálnu zložku G_y gradientu intenzity G_{xy} .
 - (b) Vypočítame veľkosť $|\nabla I_{x,y}|$ a smer $\Theta_{x,y}$ gradientu intenzity $G_{x,y}$.
 - (c) Identifikujeme koš histogramu L_i , do ktorého patrí uhol $\Theta_{x,y}$ a k jeho aktuálnej hodnote pripočítame hodnotu $|\nabla I_{x,y}|$.
3. Histogram L_i normalizujeme.

2.3 Analýza vybraných podúloh

Aj keď po dekomponovaní jednotlivých metód do krokov vo vývojových diagramoch môže realizácia týchto metód pôsobiť priamočiara, existuje niekoľko problémov, ktoré sa objavajú až pri podrobnejšej analýze. V tejto sekcii sú preto analyzované niektoré podúlohy nevyhnutné pre fungovanie týchto metód a sú k nim navrhované čo najefektívnejšie riešenia.

2.3.1 Detekcia tváre

Pre úspešné fungovanie jednotlivých metód na rozpoznávanie tvári predstavených v predchádzajúcej kapitole 1 je potrebné, aby tváre zaberali na predkladaných obrázkoch čo najväčšiu plochu a v ideálnom prípade boli zarovnané tak, aby význačné časti tvári (oči, nos, ústa, ...) boli umiestnené vždy približne na rovnakých pozíciách.

Jednou z možností je samozrejme manuálne orezávanie vstupných fotografií, čo však je v prípade veľkého množstva tréningových dát nepraktické a vôbec to neumožňuje nasadenie takto zrealizovaného systému do ostrej prevádzky.

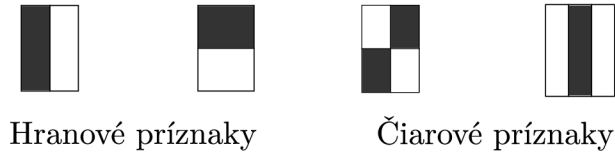
Existujú rôzne algoritmy na detekciu tvári v obraze na základe rozličných prístupov, ako napríklad geometrického či farebného popisu tváre, analýzy pohybu vo videu a mnohé iné. Keďže táto problematika nie je predmetom tejto práce, bude priblížený len princíp fungovania pravdepodobne najbežnejšie používanej metódy – *Haarovych príznakov*.

Haarove príznaky

Metóda predstavená v roku 1997 Paulom Violom a Michaelom Jonesom v [35] využíva masky tvarom podobným Haarovým vlnkám na natréňovanie viacerých slabých klasifikátorov pomocou algoritmu AdaBoost, ktoré sú pri rozpoznávaní zapojené do kaskády.

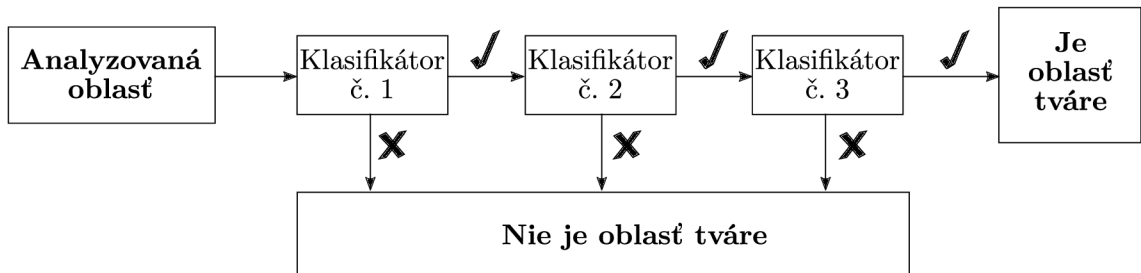
Princípom fungovania jednotlivých klasifikátorov je analýza obrazu rôznymi typmi Haarových príznakov, ktoré v analyzovanej časti obrazu vypočítavajú rozdiel medzi sumami intenzít pixelov nachádzajúcich sa v bielej a čiernej oblasti aktuálne aplikovanej masky. Ukážka príkladov Haarových príznakov je na obrázku 2.4.

Ako je schematicky znázornené na obrázku 2.5, aktuálne analyzovaná oblasť obrázku je najprv vyhodnotená prvým klasifikátorom, ktorý rozhodne či sú splnené kritériá tváre,



Obr. 2.4: Príklady Haarových príznakov [35].

následne druhým, atď'. V prípade, že postupne všetky klasifikátory prehlásia, že ide o oblasť tváre, je táto oblasť považovaná za oblasť obsahujúcu tvár. Ak ktorýkoľvek klasifikátor prehlási, že v oblasti tvár nie je, analýza končí a začína odznova v inej oblasti.



Obr. 2.5: Schematické znázornenie kaskádového vyhodnocovania detekcie tváre v analyzovanej oblasti [35].

2.3.2 Interpolácia intenzity pixelov

Pri aplikácii operátora $LBP_{P,R}$ v metóde lokálnych binárnych vzorov je bežné, že koordináty niektorého z P bodov kruhového okolia vo vzdialenosti R od stredového bodu budú ukazovať na rozhranie viacerých pixelov. Na čo najpresnejšie zistenie intenzity v takomto bode je vhodné použiť interpoláciu medzi intenzitami okolitých pixelov.

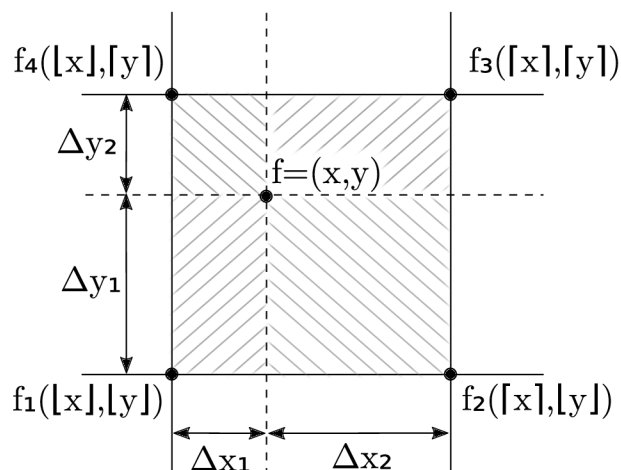
Najjednoduchším riešením problému interpolácie v dvoch rozmeroch Kartézskej mriežky je tzv. *bilinéárna interpolácia* [24], ktorá získava výslednú intenzitu váženým zložením intenzít susediacich pixelov podľa vzťahov uvedených v 2.2. Funkcia $f(x, y)$ označuje intenzitu v bode (x, y) , funkcia $\lfloor x \rfloor$ a $\lceil x \rceil$ značí zaokrúhlenie čísla x na najbližšie celé nižšie, resp. vyššie číslo:

$$\begin{aligned}
 f(x, y) &= f_1(1-t)(1-u) + f_2t(1-u) + f_3tu + f_4(1-t)u, \quad \text{kde} \\
 f_1 &= f(\lfloor x \rfloor, \lfloor y \rfloor), & f_2 &= f(\lceil x \rceil, \lfloor y \rfloor), \\
 f_3 &= f(\lfloor x \rfloor, \lceil y \rceil), & f_4 &= f(\lceil x \rceil, \lceil y \rceil), \\
 t &= \frac{x - \lfloor x \rfloor}{\lceil x \rceil - \lfloor x \rfloor}, & u &= \frac{y - \lfloor y \rfloor}{\lceil y \rceil - \lfloor y \rfloor}
 \end{aligned}
 \tag{2.2}$$

Na schematickom znázornení bilinéarnej interpolácie 2.6 sú váhy reprezentované znázornenými plochami.

2.3.3 Efektívne počítanie zmien bitových úrovní

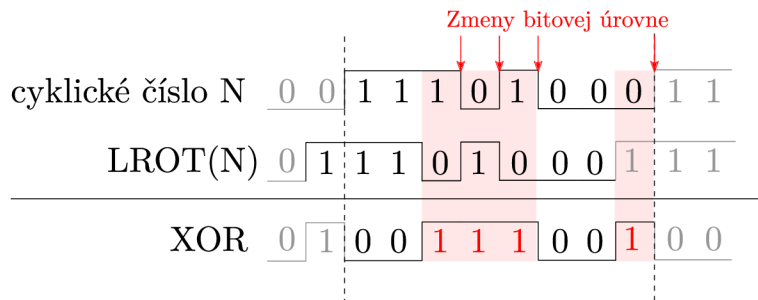
Jedným z ďalších potrebných úkonov v metóde lokálnych binárnych vzorov je tiež klasifikácia cyklických binárnych reťazcov na uniformné, príp. neuniformné. Ako bolo vysvetlené



Obr. 2.6: Schematické znázornenie bilineárnej interpolácie [24].

v podsekcii 1.4.2, uniformné reťazce obsahujú najviac dve zmeny hodnoty v postupnosti bitov. Pri naivnom prístupe je na analýzu reťazca o dĺžke N bitov potrebný presne rovnaký počet krokov na postupný prechod jeho jednotlivými prvkami.

Po zakomponovaní dvoch binárnych operácií je tento problém možné znížiť na najviac $N + 2$ krokov. Vzťahom $x \oplus ROT(x)$, teda aplikáciou nonekvivalencie (binárna operácia XOR) na pôvodný reťazec a produkt jeho rotácie o jedno miesto v ľubovoľnom smere, dostaneme reťazec označujúci jednotkovými bitmi miesta prechodov medzi dvoma logickými úrovňami. Ako je názorne zobrazené na príklade 2.7, počet týchto logických jednotiek teda udáva hľadaný počet zmien v postupnosti bitov.



Obr. 2.7: Príklad počítania zmien bitových úrovní.

Na rýchle spočítanie jednotových bitov je následne možné použiť metódu prezývanú ako *Kernighanová metóda*, ktorá bola spopularizovaná v učebnici jazyka C Brianom Kernighanom, avšak prvýkrát publikovaná už o 28 rokov skôr v americkom mesačníku *Communications of the ACM* [37]. Jej princípom je postupné nulovanie jednotlivých bitov zadaného bitového vektora a počítanie počtu potrebných iterácií na jeho úplné vynulovanie. Tento krátky ale mimoriadne efektívny algoritmus je uvedený v kóde 2.1.

```

|   unsigned long bitvector; //analyzovany bitovy vektor
|   int count = 0;          //vysledny pocet jednotkovych bitov

```

```

for(; bitvector; bitvector &= bitvector - 1)
    count++;

```

Kód 2.1: Algoritmus efektívneho spočítania jednotkových bitov v bitovom vektore.

2.3.4 Aproximácia výpočtu gradientov

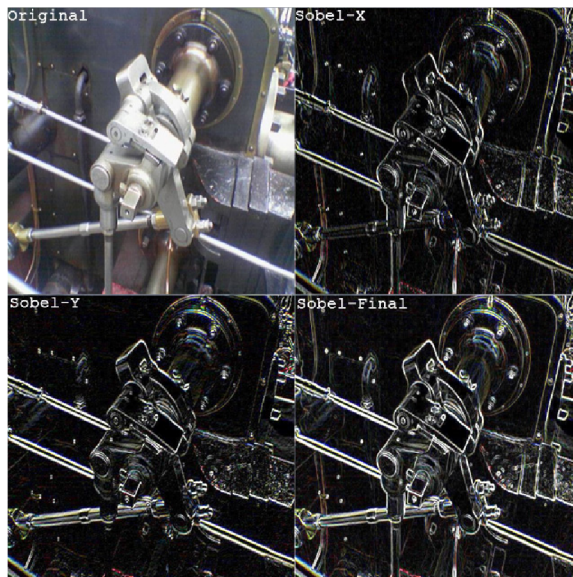
Ako už jej názov napovedá, metóda histogramov orientovaných gradientov pracuje vo svojom algoritme s orientovanými zmenami intenzity v obraze. Presný výpočet gradientov možný nie je, pretože funkcia intenzity nie je spojitá, ale je známa iba v určitých bodoch. Na ich vypočítanie je preto potrebné predpokladať, že body v obrázku sú diskretizovanou realizáciou spojitaj funkcie intenzity.

V počítačovej grafike je bežne používaným spôsobom na aproximovanie gradientov intenzity v jednotlivých bodoch snímky konvolúcia so špeciálne navrhnutým operátorom. Jedným z najznámejších je tzv. *Sobelov operátor* [30], ktorý je realizovaný ako dvojica matíc o veľkostiach 3×3 prvkov na aproximáciu horizontálneho gradientu s_x a vertikálneho gradientu s_y ako je uvedené v 2.3. Vzťah 2.4 potom označuje operáciu konvolúcie $*$ niektorého operátora s aplikovaného na obrázok I . Výpočet prebieha pre každý bod obrázku tak, že po zarovnaní stredového prvku operátora s na analyzovaný bod obrázku I sa operácia konvolúcie $*$ počíta ako súčet hodnôt prvkov operátora vynásobených s intenzitami korešpondujúcich bodov obrázku I . Chýbajúce intenzity po okrajoch obrázka sú zvyčajne nahradené hodnotou 0 alebo najbližšie susediacou intenzitou.

Ukážka aplikovania Sobelovho filtra na fotografiu je zachytená na obrázku 2.8.

$$s_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad s_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & 2 & -1 \end{bmatrix} \quad (2.3)$$

$$G = s * I \quad (2.4)$$



Obr. 2.8: Ukážka aplikácie Sobelovho filtra na fotografiu [7].

Kapitola 3

Implementácia

Na začiatku tejto kapitoly sú v sekcii 3.1 predstavené technológie zvolené na implementáciu troch metód, ktoré sú predmetom tejto práce, spolu s uvedením dôvodov, pre ktoré boli vybrané. Ďalej nasleduje popis architektúry implementovanej aplikácie a predstavenie funkcií vytvoreného grafického používateľského rozhrania, ako aj dodatočne naprogramovanej konzolovej alternatívy pre potreby vyhodnocovania experimentov.

3.1 Použité technológie

Na implementáciu riešenia som zvolil imperatívny programovací jazyk *C++* v štandarde z roku 2011 z dôvodu predchádzajúcich skúseností s programovaním v tomto jazyku ale aj z viacerých ďalších dôvodov. Keďže rýchlosť implementovaných metód je jedným z významných kritérií pri problematike identifikácie osôb z fotografií, použitie staticky kompilovaného jazyka považujem za mimoriadne vhodné. Taktiež podpora objektovo orientovaného paradigma prispieva k prehľadnejšiemu kódu implementácie a vďaka podpore štandardu *C++11*, ktorého implementácia je už dostatočne rozšírená a stabilná na všetkých hlavných podporovaných platformách, je možné používať v niektorých vhodných prípadoch elegantnejšie konštrukcie.

Na organizáciu súborov, prehľadné sledovanie zmien a zálohovanie práce som využíval verzovací systém Git, ktorý v prostredí cloudovej služby BitBucket [3] od spoločnosti Atlassian poskytoval bezpečnú zálohu počas celého písania textu práce a vývoja aplikácie.

Samotná implementácia tiež využíva 2 knižnice, ktoré sú popísané v nasledujúcich podsekciami. Tie boli okrem rýchlej krivky učenia na začatie ich používania a univerzálnosti volené aj s prihliadaním na podporu jazyka *C++*.

3.1.1 Knižnica OpenCV

OpenCV (skratka pre *Open Source Computer Vision*) je populárna multiplatformná otvorená knižnica poskytujúca funkcie prevažne zamerané na spracovanie obrazu v reálnom čase. Je šírená pod licenciou *BSD* umožňujúcou voľné použitie na akademické aj komerčné účely. Táto práca využíva verziu 2.4.9 skompilovanú pre 64-bitovú architektúru.

Implementácia realizovaná ako súčasť tejto práce primárne využíva jeden z elementárnych objektov, ktoré *OpenCV* poskytuje, a to univerzálny objekt *N*-rozmerného poľa `cv::Mat` na reprezentáciu matematických vektorov, obrázkových matíc a histogramov, a nad ním pracujúce efektívne implementované matematické operácie.

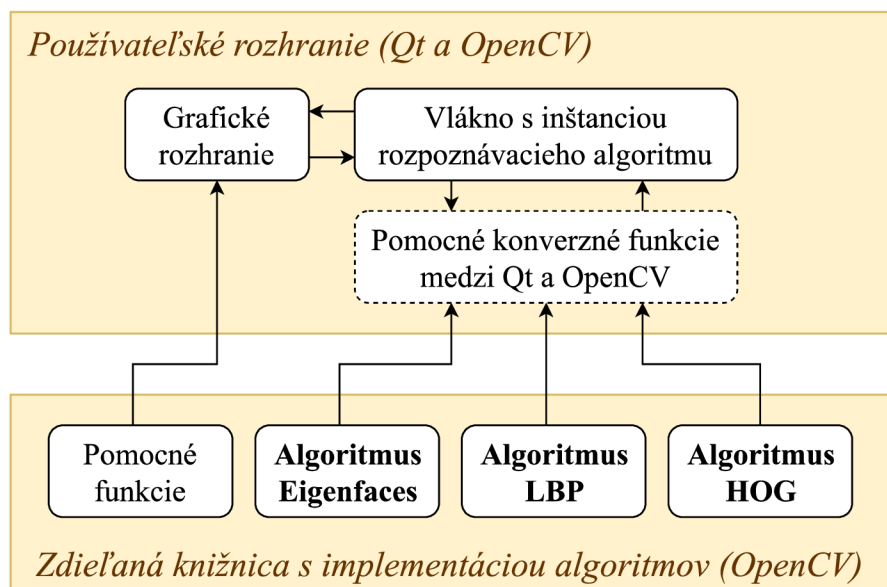
3.1.2 Qt framework

Na vytvorenie grafického používateľského rozhrania bola zvolená multiplatformná knižnica *Qt* vo verzii 5.8.0 šírená mimo iného aj pod využitou licenciou *GNU LGPL* verzie 3. Poskytuje veľké množstvo ľahko prispôsobiteľných komponent a mechanizmov, ktoré sprehľadnili a zjednodušili vývoj aplikácie a tiež výrazne urýchlili prototypovanie a experimentovanie pri jej návrhu. Znáмым problémom tejto

Samotná aplikácia, ktorá je predmetom tejto práce, bola vyvíjaná a ladená v prostredí pridruženej aplikácie *Qt Creator*.

3.2 Architektúra aplikácie

Od začiatku vývoja bol kladený dôraz na minimalizovanie množstva externých závislostí jadra implementovaného systému – rozpoznávacích algoritmov. Preto bola aplikácia dekomponovaná na dva hlavné celky, a to jej používateľské rozhranie a samotnú implementáciu algoritmov rozpoznávacích metód.



Obr. 3.1: Architektúra implementovanej aplikácie.

Ako je ukázané na obrázku 3.1, algoritmy a ďalšie pomocné funkcie na prácu s knižnicou *OpenCV* boli implementované v samostatne skompilovateľnom projekte nezávislom na *Qt* a umožňujúcim vytvorenie zdieľaných alebo staticky linkovaných knižníc. Tento prístup umožňuje v budúcnosti využitie nami implementovanej knižnice v ďalších projektoch nezávislých nielen na programovacom jazyku, keďže väčšina jazykov podporuje volania z externých knižníc, ale aj na samotnej platforme vďaka multiplatformnej podpore zabezpečenej závislosťou len na projekte *OpenCV*.

Na realizáciu grafického rozhrania aplikácie a podporu spúšťania metód v samostatných vláknach boli využité funkcie, ktoré poskytuje knižnica *Qt*.

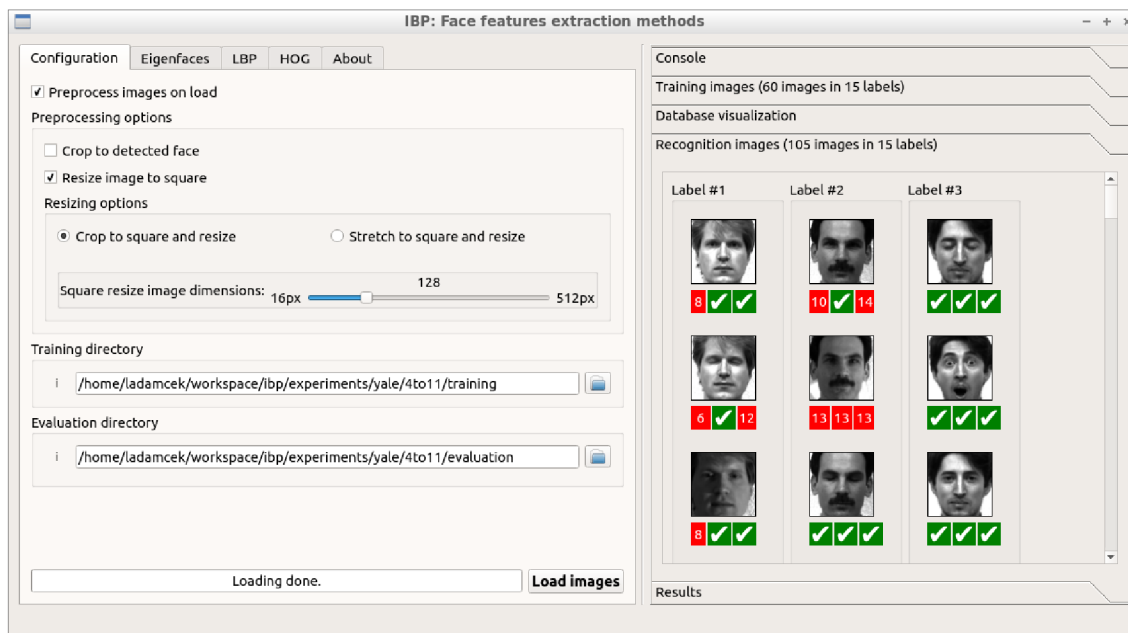
3.3 Používateľské rozhranie implementovanej aplikácie

Na pohodlné experimentovanie s implementovanými algoritmami bolo vytvorené používateľské rozhranie, ktoré umožňuje spúšťať jednotlivé metódy nad ľubovoľnými sadami obrázkov. Pri jeho návrhu bol kladený dôraz na univerzálnosť konfigurovania rozpoznávacích metód a zároveň prehľadnosť. Aj keď nami implementovaná knižnica s metódami ponúka rozšírenú funkcionalitu, vytvorené používateľské rozhranie bolo navrhnuté na experimentovanie s identifikačnými úlohami na uzavrenej množine osôb, takže neobsahuje možnosť nastavenia práhu pre detekciu neznámej osoby ani funkcionalitu súvisiacu s verifikačnými úlohami.

Výsledok realizácie je zachytený na snímkach 3.2 a 3.3. Možeme vidieť, že aplikácia je členená do dvoch hlavných častí: na ľavej strane je komponenta, ktorá združuje karty s nastaveniami, a na pravej strane sú panely s výstupmi aplikácie.

Na prvej snímke 3.2 sú zobrazené možnosti prvej karty s konfiguráciou spoločných nastavení pre rozpoznávanie:

- povolenie predspracovania obrázkov pri ich načítaní a vlastnosti predspracovania:
 - orezanie obrázka na detegovanú tvár pomocou Haarových príznakov s využitím funkcionality knižnice OpenCV,
 - povolenie zmeny veľkosti obrázka na štvorcový tvar o zadanej veľkosti hrany v pixeloch – buď orezaním dlhšieho rozmeru alebo rozťahnutím kratšieho rozmeru, a následne zväčšenie alebo zmenšenie na požadovanú veľkosť,
- výber „trénovacieho priechinka“, z ktorého snímkov sú vytvorené šablóny do databázy, a „vyhodnocovacieho priechinka“ so snímkami použitými pri vyhodnocovaní.

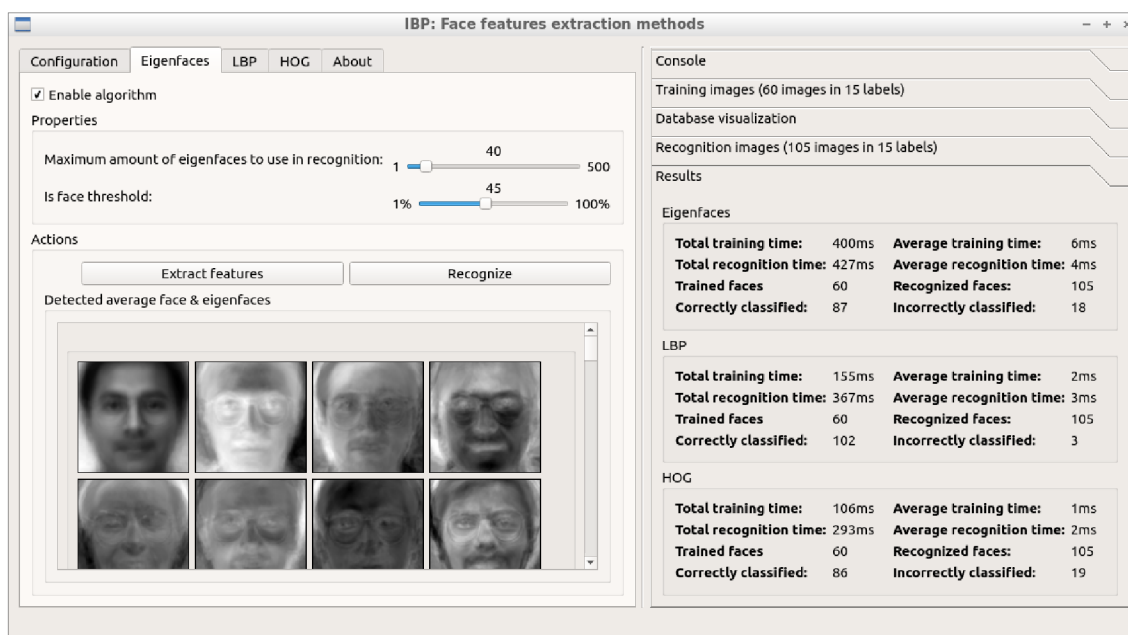


Obr. 3.2: Vzhľad konfiguračného panelu implementovanej aplikácie a výsledkov rozpoznávania.

Úplne dole sa nachádza indikátor priebehu načítavania a tlačidlo spúšťajúce, resp. prerušujúce proces načítavania obrázkov. Načítané obrázky sa zobrazia v zodpovedajúcich

paneloch napravo, kde sú zoskupené podľa označenia osoby ku ktorej patria. V prípade vyhodnocovacej množiny obrázkov sú pod každou snímku zobrazené tri indikátory vizualizujúce (ne)úspešnosť identifikácie jednej z troch implementovaných metód (po ukázaní kurzorom myši zobrazia názov svojej metódy), pričom v prípade neúspešnej identifikácie zobrazia chybné priradené označenie.

Každá z metód má svoju vlastnú kartu s konfigurovateľnými parametrami. Na snímke 3.3 sú zachytené možné nastavenia metódy analýza hlavných komponent. Keďže metóda na snímke už vyextrahovala charakteristické rysy z trénovacej množiny obrázkov, zobrazuje sa tiež vypočítaná priemerná tvár a detegované vlastné tváre. Po spustení rozpoznávania sú postupne identifikované všetky snímky z „vyhodnocovacej množiny“ načítaných obrázkov a nastavené zodpovedajúce indikátory informujúce o správnom alebo nesprávnom identifikovaní každého obrázka v tejto množine.

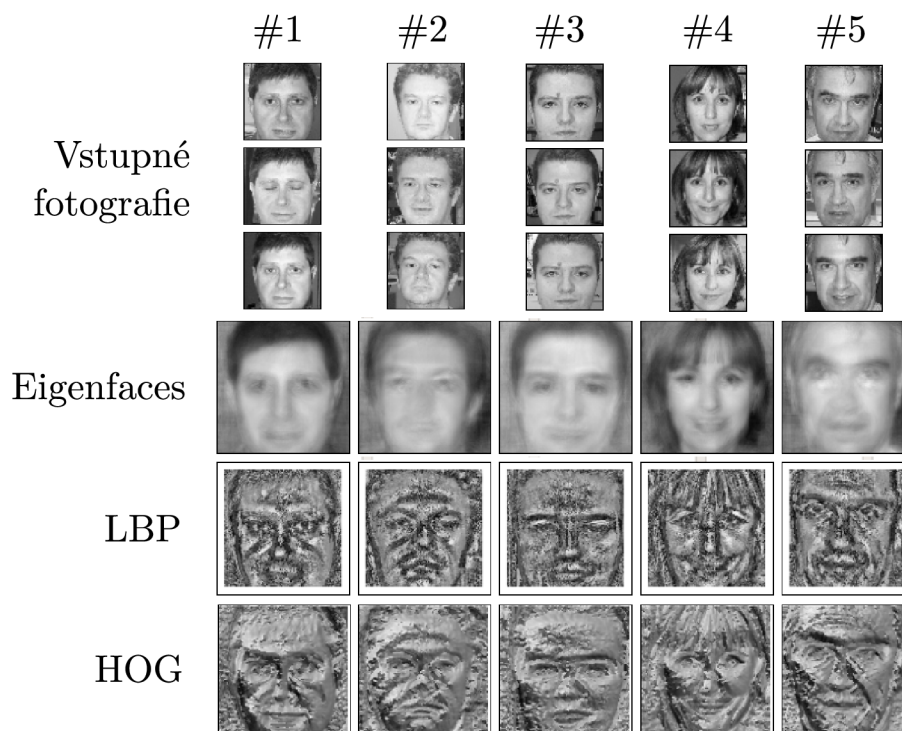


Obr. 3.3: Ukážka konfigurácie metódy pripravenej na rozpoznávanie a sumarizácie výsledkov identifikácie.

Na pravej strane môžeme vidieť obsah panelu s nameranými výsledkami, ktorý zobrazuje štatistiky súvisiace s jednotlivými metódami. Medzi ne patrí počet obrázkov trénovacej a vyhodnocovacej množiny, pri ktorom bola daná metóda spustená, celkový čas fázy prípravy šablón a rozpoznávania, priemerný čas pripadajúci na spracovanie jednej snímky a počet správne a nesprávne identifikovaných snímok.

Jednou zo zaujímavých funkcií implementovaného rozhrania je možnosť zobrazit vizuálnu reprezentáciu šablón jednotlivých osôb, ktoré sú uložené v databáze. Ukážka týchto výstupov je zachytená na obrázku 3.4. V prípade metódy analýza hlavných komponent (označená ako „Eigenfaces“) je vizualizácia vytvorená projekciou tváre Φ_f zo vzorca 1.9. V prípade metódy LBP ide o namapovanie normalizovaných hodnôt lokálnych binárnych vzorov do rozsahu 0 až 255, a v prípade HOG je zasa na rovnaký rozsah namapovaný výsledok $\nabla I + 2 * \theta$, kde ∇I je veľkosť gradientu intenzity v bode obrázka a θ je jeho uhol. Tretí menovaný vzťah bol objavený experimentálne a zvolený z dôvodu, že prinášal najatraktívnejšie vizuálne reprezentácie spomedzi všetkých pokusov o zvizualizovanie HOG šablón. Zo sekcie 2.1 dru-

hej kapitoly však vieme, že sa pri porovnávaní využívajú pôvodné reprezentácie vektorov charakteristických rysov, ktorými sú vektory váh a globálne histogramy.



Obr. 3.4: Vizualizácia šablón 5 rôznych osôb vytvorených pomocou implementovaných metód na extrakciu charakteristických rysov.

3.4 Konzolová aplikácia

Vo fáze návrhu a merania experimentov bola dodatočne naprogramovaná aj konzolová verzia aplikácie, ktorej účelom je podpora automatizovaného vyhodnocovania experimentov vďaka parametrizovateľnému spúšťaniu z príkazovej riadky a výstupu nameraných hodnôt v CSV formáte.

Pre identifikačné úlohy predstavuje výstup s nameranými hodnotami úspešnosti identifikácie pre jednotlivé pozície CMC krivky a tiež hodnotu minimálnej, priemernej a maximálnej vzdialenosti medzi rozpoznávanými snímkami a šablónami z databázy. Medzi ďalšiu funkcionality tejto verzie patrí, na rozdiel od grafického rozhrania, tiež vyhodnocovanie verifikácie obmedzené na výstup obsahujúci TPR a FPR hodnoty v jednotlivých úrovniach nastavenia prahu, ktoré je v nasledujúcej kapitole používané pre vykreslenie ROC krivky.

Parametre, ktoré sú podporované pri spúšťaní tejto aplikácie, sú popísané v prílohe [A](#).

Kapitola 4

Experimenty

Táto kapitola popisuje experimenty vykonané na nami implementovaných metódach, ktoré boli prevedené za účelom nielen vyhodnotenia ich úspešnosti pri rozpoznávaní na rôznych sadách fotografií, ale aj s cieľom pochopiť, akým spôsobom ich výsledky ovplyvňuje zmena použitej konfigurácie. Na začiatku kapitoly sú preto predstavené zvolené dátové sady, nad ktorými boli jednotlivé experimenty spúšťané, a tiež popísaná metodika experimentovania. Ďalej nasleduje prezentovanie výsledkov meraní pomocou CMC a ROC kriviek na identifikačnej a verifikačnej úlohe, avšak samotné sledovanie závislosti zmeny parametrov na úspešnosti rozpoznávania bolo vykonané len na identifikačných experimentoch.

V závere kapitoly je analyzovaná rýchlosť a pamäťová náročnosť implementácie, stručne zhodnotené výsledky experimentov a riešenia ako celku spolu s navrhnutými možnosťami ďalšieho napredovania v oblasti, ktorou sa zaoberá táto práca.

4.1 Použité dátové sady

Experimentovanie na uzavrenej množine osôb bolo vykonané na troch rozličných voľne dostupných dátových sadách, ktoré boli vybrané tak, aby reprezentovali špecifické pozorované vlastnosti metód.

Ako prvá dátová sada bola vybraná dátová sada vytvorená Markusom Weberom na Kalifornskom technologickom inštitúte (Caltech) obsahujúca 450 farebných fotografií priameho pohľadu do objektívu približne 27 osôb [36]. Tieto fotografie boli nafotené v bežných situáciách, čo spôsobilo, že niektoré fotografie obsahujú príliš tmavé, preexponované alebo iným spôsobom nevhodné fotografie pre naše experimenty. Taktiež obsahuje výrazné rozdiely v počtoch snímok jednej osoby. Preto bola táto dátová sada manuálne prečistená a vybralo sa z nej 17 subjektov, ktoré obsahovali dostatočný počet snímok – konkrétne 16. Na normalizáciu boli využité koordináty pozície tváre z Matlab súboru priloženého k dátovej sade, a ďalej boli pri vyhodnocovaní v rámci predspracovania orezané jednotlivé snímky na štvorcový tvar a skonvertované do odtieňov šedej farby. Ukážka fotografií z tejto dátovej sady je k dispozícii na predchádzajúcej vizualizácii 3.4.

Druhou množinou bola *ORL databáza tvári* [4] zachytávajúca 40 rozličných osôb z prostredia Laboratórií AT&T Univerzity v Cambridgi. Každá osoba je zachytená vždy na 10-tich čiernobielych fotografiách v nemenných svetelných podmienkach s prevažne rozličným natočením pozície hlavy a s občasou malou variabilitou výrazu tváre. Ukážka fotografií tejto dátovej sady je na obrázku 4.1.



Obr. 4.1: Ukážka snímok jednej osoby z dátovej sady ORL databáza tvárí [4].

Poslednou vybranou množinou bola *Databáza tvárí Yalovej univerzity* [33, 5], ktorá obsahuje 165 čiernobielych fotografií 15-tich rôznych osôb. Každá osoba je zachytená na 11-tich snímkoch v kontrolovaných variabilných svetelných podmienkach, s rozličnými výrazmi tváre a čiastočným prekrytím tváre okuliarmi. Ako je možné vidieť na obrázku 4.2, konkrétne ide o neutrálny výraz tváre nasvietený priamo, z ľavej a pravej strany a ďalej fotografie zachytávajúce žmurkanie, ospalosť, smútok alebo prekvapenie. Táto dátová sada bola normalizovaná použitím detekcie tváre implementovanej v našej aplikácii a upravená na štvorcový tvar orezaním časti čela a brady z fotografie.



Obr. 4.2: Ukážka normalizovaných snímok jednej osoby z dátovej sady Databáza tvárí Yalovej univerzity [33].

4.2 Metodika experimentovania

Na prevedenie experimentov bolo po zvolení vhodných dátových sád tiež potrebné vybrať východzie nastavenia, ktoré boli používané a postupne modifikované v jednotlivých meraniach. Táto sekcia sa preto zaoberá ich voľbou, popisuje všeobecný proces merania experimentov a následne aj spôsob ich vyhodnocovania.

4.2.1 Nastavenie východných parametrov pri experimentovaní

Hodnoty jednotlivých parametrov pri experimentoch boli zvolené s cieľom zachovať čo najvyššiu úspešnosť rozpoznávania avšak s ohľadom na rýchlosť vyhodnocovania. Preto v experimentoch tejto sekcie bolo použité rozlíšenie snímok tvárí 128 na 128 pixelov. Nastavenie parametrov, ktoré neboli predmetom pozorovania daného experimentu, boli zvolené odhadmi získanými v priebehu implementácie a testovania našej aplikácie na hodnoty, ktoré prinášali subjektívne najúspešnejšie výsledky:

- analýza hlavných komponent (ďalej označovaná aj ako *Eigenfaces*):
 - maximálny počet použitých vlastných tvárí: 40,
- lokálne binárne vzory (*LBP*):
 - horizontálny počet regiónov: 6,
 - vertikálny počet regiónov: 6,
 - počet bodov operátora: 6,

- polomer operátora: 6,
- histogramy orientovaných gradientov (*HOG*):
 - horizontálny počet regiónov: 6,
 - vertikálny počet regiónov: 6,
 - počet košov: 70,
 - samostatné koše pre celý podporovaný úhlový rozsah.

Keďže sa jednalo o rozpoznávanie na uzavrenej množine osôb, nastavenie prahu analýzy hlavných komponent určujúcej či ide o ľudskú tvár bolo vždy zvolené na maximálnu prípustnú hodnotu. Váhy regiónov v metódach LBP a HOG boli vo všetkých experimentoch nastavené na hodnotu 1,0.

4.2.2 Proces experimentovania a vyhodnocovania

Na získavanie výsledkov experimentov bola použitá konzolová verzia implementovanej aplikácie parametrizovane spúšťaná prostredníctvom pripravených shellových skriptov v jazyku *Bash*. Namerané hodnoty boli následne vyhodnotené pomocou pomocných skriptov v jazyku *Python* za použitia knižnice *Scikit-learn* [28] na vypočítanie metriky AUC lichobežníkovou metódou a zvizualizované vo forme grafov pomocou nástroja *Gnuplot* [12].

Každý experiment bol nameraný a vyhodnotený na každej dátovej sade, a to v niekoľkých variantách podľa dostupného počtu snímok na jednu osobu. Dostupný počet N snímok danej osoby bol vždy rozdelený na „trénovaciú“ množinu obsahujúcu prvých T fotografií, ktorá bola použitá pri extrakcii charakteristických rysov, a zvyšný počet $R = N - T$ bolo použitých na vyhodnocovanie experimentu. Experimentovanie pre dané sady s nasledujúcimi počtami snímok v trénovacej množine:

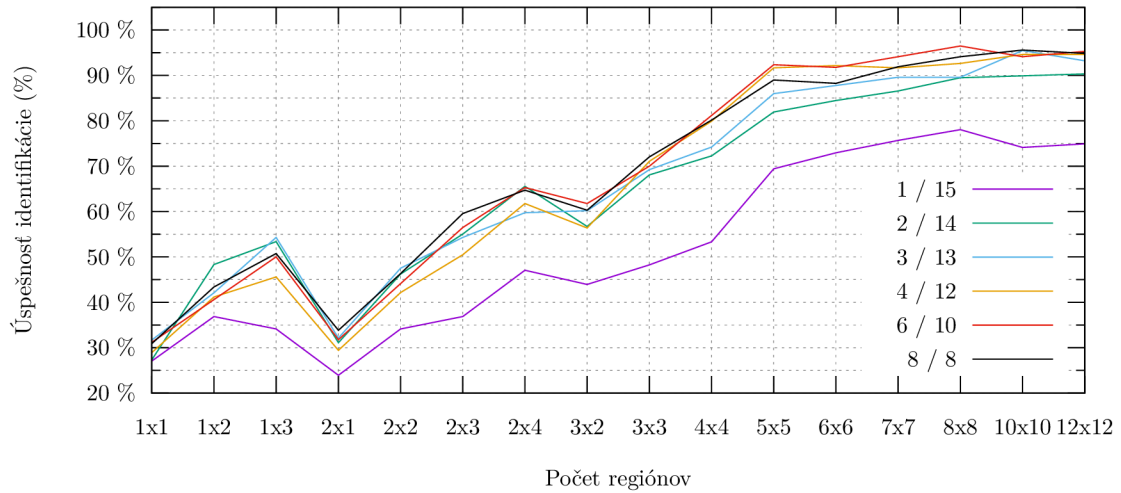
- pre 17 osôb dátovej sady „Caltech” s $N = 16$: $T = [1, 2, 3, 4, 6, 8]$
- pre 40 osôb dátovej sady „ORL” s $N = 10$: $T = [1, 2, 3, 4]$
- pre 15 osôb dátovej sady „YALE” s $N = 15$: $T = [1, 2, 3, 4, 5]$

Mierou úspešnosti identifikácie v identifikačných experimentoch bola zvolená percentuálna hodnota pravdepodobnosti, ktorú nadobúda CMC krivka na prvej pozícii, teda počet správne označených identít našim systémom z celkového počtu snímok danej dátovej sady použitých na vyhodnocovanie.

Všetky výsledky experimentov sú dostupné na médiu priloženom k tejto práci. Ukážka grafu vyhodnocujúceho jeden z experimentov je na obrázku 4.3. Výsledok tohoto aj ďalších experimentov vo všeobecnosti ukazuje o zhruba tretinu až polovicu horšie výsledky, ak v trénovacej množine je len jedna snímka osoby, ako v prípade ostatných variánt. V tých sú výsledky veľmi podobné a líšia sa len na úrovni pár percent, pričom najlepšie výsledky dosahujú krivky 4-6 snímok použitých v trénovacej množine.

V nasledujúcich sekciách tejto práce sú porovnané krivky experimentov vo variante $T = 4$.

Ex 3: HOG (Caltech, 17 identít)



Obr. 4.3: Ukážka výsledkov experimentu so sledovaním závislosti počtu regiónov na úspešnosti identifikácie vykonaná v rôznych variantách veľkosti tréningovej množiny na dátovej sade Caltech.

4.3 Vyhodnotenie rozpoznávania pri odhadnutých parametroch

Ako už bolo spomínané v podsekcii 4.2.2, v tejto sekcii sú porovnávané výsledky experimentov vykonané na jednotlivých dátových sádach pri použití prvých 4 snímkov každej osoby na vytvorenie šablón v databáze, ktoré sú využité na rozpoznávanie 12 (Caltech), 6 (ORL) resp. 11 (YALE) osôb vo fáze vyhodnocovania.

V experimente tejto sekcie je sledovaná výkonnosť jednotlivých metód pri verifikácii a rozpoznávaní vo východzej konfigurácii uvedenej v podsekcii 4.2.1.

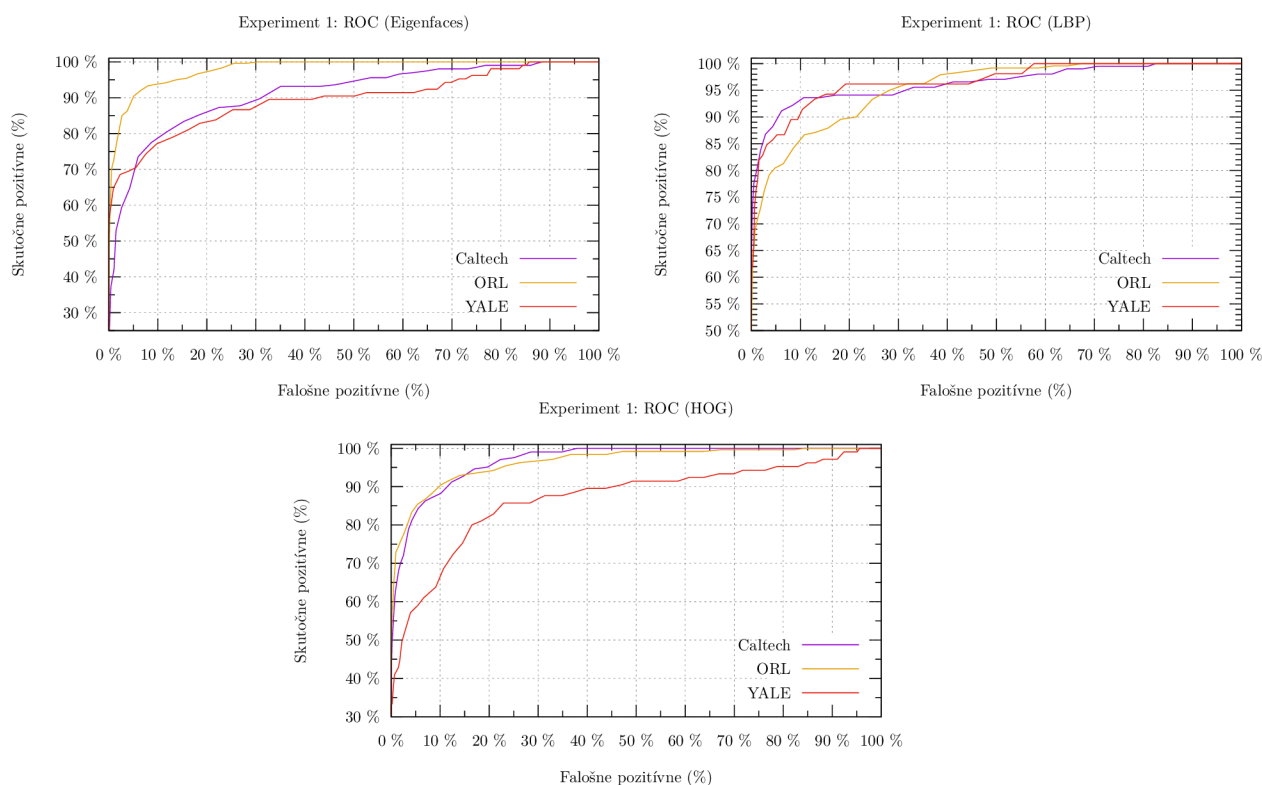
Grafy na obrázku 4.4 porovnávajú ROC krivky jednotlivých metód, pričom na získanie dát bolo vykonaných 100 meraní s postupne posúvaným prahom verifikácie. Hodnoty AUC vyšli pre metódy uvedené v poradí Eigenfaces, LBP a HOG nasledovne:

- Caltech: 0,912; 0,964 a 0,968;
- ORL: 0,980; 0,954 a 0,964;
- YALE: 0,895; 0,965 a 0,868.

Vidíme, že metóda Eigenfaces je najefektívnejšia pri dátovej sade ORL a HOG má najhoršie výsledky na sade YALE. Celkovo najúspešnejší priemerný AUC 0,961 verifikácie prináša v tejto konfigurácii metóda LBP.

Vyhodnotenie úspešnosti identifikácie jednotlivých metód na rovnakej konfigurácii je znázornené na grafe 4.5. Výsledky AUC uvedené v rovnakom poradí ako v prípade verifikačného experimentu vyšli nasledujúce:

- Caltech: 0,937; 0,992 a 0,996;
- ORL: 0,986; 0,982 a 0,990;



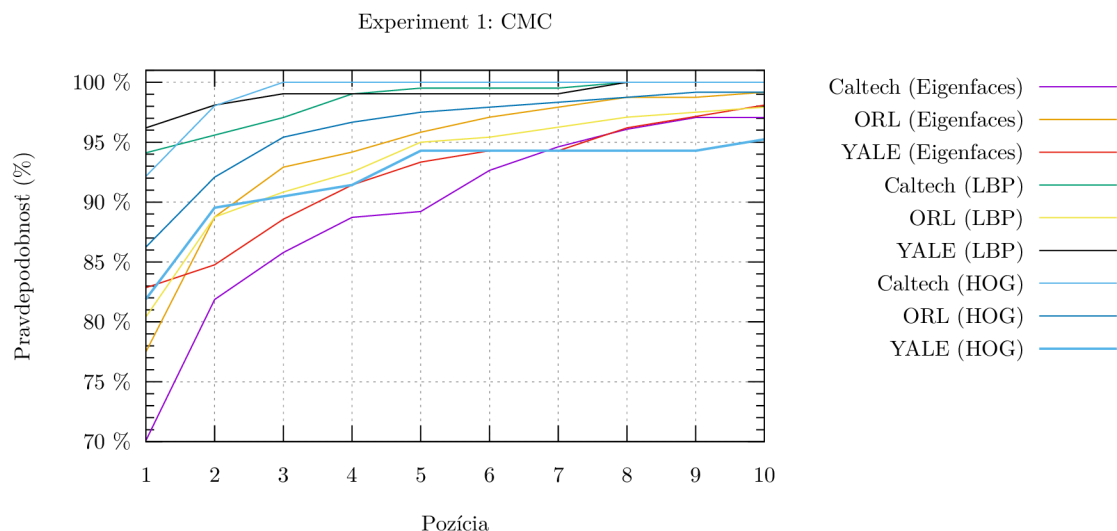
Obr. 4.4: Porovnanie ROC kriviek jednotlivých implementovaných metód pri verifikácii na troch rôznych dátových sadách.

- YALE: 0,950; 0,994 a 0,945.

Môžeme vidieť, že z pohľadu metriky AUC metóda HOG prináša najhoršie výsledky s priemerným AUC 0,963 a naopak najlepšie má opäť metóda LBP s priemerným AUC 0,986, pred Eigenfaces s výsledkom 0,975. Avšak pri hodnotení podľa priemernej úspešnosti identifikácie len na základe hodnoty prvej pozície na CMC krivke dosahuje najlepšie výsledky LBP s hodnotou 90,242% a najhoršie výsledky dosahuje metóda Eigenfaces s hodnotou 86,771%.

4.4 Závislosť konfigurácie jednotlivých metód na úspešnosti rozpoznávania

Cieľom experimentov uvedených v tejto sekcii je sledovanie, ako sa vyvíja úspešnosť rozpoznávania pri zmenách nastavenia konkrétneho parametra v jednotlivých metódach, čo zároveň ovplyvňuje proces vytvárania vektora charakteristických rysov pri analyzovaní snímky. Východzia konfigurácia bola popísaná a nameraná v predchádzajúcej sekcii 4.2.1. Tieto experimenty boli merané na identifikačnej úlohe procesom popísaným v podsekcii 4.2.2 a ako už bolo spomínané, za metriku úspešnosti je preto považovaná hodnota 1. pozície na CMC krivke.



Obr. 4.5: Porovnanie prvých 10 pozícií CMC kriviek jednotlivých implementovaných metód pri identifikácii na uzavrenej množine osôb na troch dátových sadách.

4.4.1 Počet vlastných tvárí

Cieľom tohoto experimentu bolo zistiť, ako ovplyvňuje použitý počet vlastných tvárí úspešnosť identifikácie subjektov. Pri experimentovaní boli vlastné tváre získané vo fáze extrakcie charakteristických rysov analýzy hlavných komponent, ktorá bola znázornená na obrázku 2.1, postupne limitované od použitia 1 – tej s najvyšším vlastným číslom, až po ich maximálny dostupný počet, ktorý bol obmedzený počtom snímok použitej trénovacej množiny.

Z výsledkov merania zaznačených do grafu na obrázku 4.6 je zrejmé, že na použitých dátových sadách je dostatočné používať najviac 40 vlastných tvárí, a akýkoľvek vyšší počet použitý pri identifikácii prináša len zanedbateľné zlepšenie za cenu za cenu dlhšieho výpočtu. V našom prípade bola priemerná úspešnosť identifikácie pri použití maximálne 40 vlastných tvárí na úrovni 76,818%, pričom priemerné zlepšenie pri použití maximálneho detegovaného počtu vlastných tvárí bolo len o 0,58%.

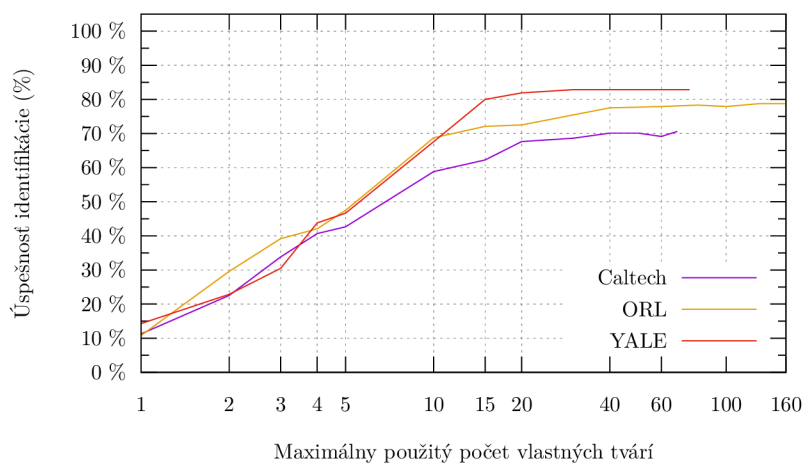
4.4.2 Množstvo regiónov u metód LBP a HOG

V ďalšom experimente bolo pozorované ako v metódach LBP a HOG vplyva spôsob rozdelenia spracovávaného obrázka na úspešnosť identifikácie.

Na grafoch 4.7 vidíme, že horizontálne rozdelenie obrázka na dve polovice prináša približne rovnako slabé výsledky ako použitie len jedného lokálneho regiónu pokrývajúceho celý obrázok. Na porovnaní regiónov 2×3 a 3×2 , resp. 1×2 a 2×1 je zasa možné vidieť, že dobrá úspešnosť identifikácie viac závisí na vertikálnom delení obrázka.

Avšak závislosť na vertikálnom delení nemaximalizuje možnú dosiahnuteľnú úspešnosť. Pri porovnaní hodnôt rozdelenia 1×2 a 2×2 alebo 1×3 , 2×3 a 3×3 alebo 2×4 a 4×4 môžeme vidieť, že pre rozdelenie $w \times h$ má priemerná úspešnosť stúpajúci trend pre $w \leq h$, a teda je aspoň na nami skúmaných dátových sadách najlepšia pri rozdelení $w = h$.

Experiment 2: Eigenfaces



Obr. 4.6: Závislosť použitého počtu vlastných tvárí vzhľadom na úspešnosť identifikácie analýzou hlavných komponent na uzavrenej množine osôb troch rôznych dátových sád.

Výraznejšie stúpajúca tendencia úspešnosti identifikácie platí približne do rozdelenia regiónov na 5×5 v prípade HOG, resp. 6×6 pre LBP, za hranicou 8×8 , resp. 10×10 prichádza naopak už ku klesaniu.

4.4.3 Počet bodov a polomer LBP operátora

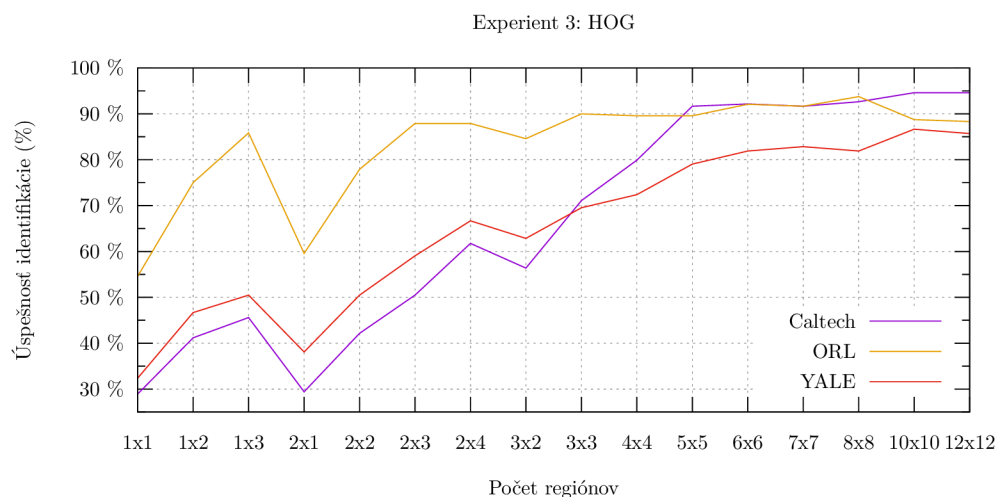
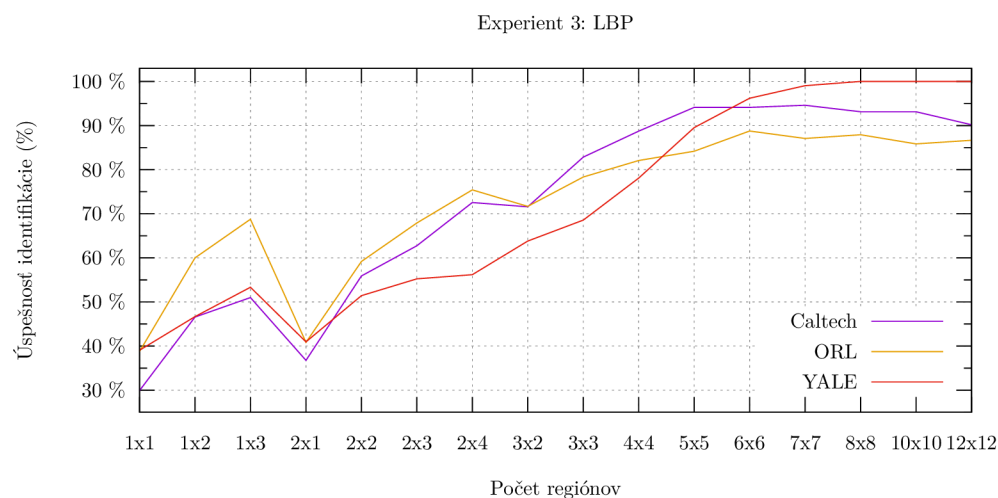
Pri nezávislom skúmaní vplyvu veľkosti a polomeru LBP operátora, ktorý sa využíva na vytváranie deskriptorov pixelov analyzovaného obrázka v rovnomennej metóde, na úspešnosť identifikácie sme namerali hodnoty zaznačené v grafoch na obrázku 4.8.

S počtom bodov operátora vzrastá počet košov lokálnych histogramov tejto metódy. Z prvého grafu je zrejmé, že výrazné zlepšenie nastáva, ak je počet použitých bodov aspoň 3. V prípade použitia operátora v tvare pravidelného pentagonu dochádza v dvoch z troch dátových sád k niekoľkopercentnému poklesu, avšak pri pridávaní bodov dochádza k ďalšiemu stúpaniu priemernej hodnoty úspešnosti približne až do počtu 10. Celkovo je zlepšenie priemernej úspešnosti medzi použitím 3 a 10 bodov len o 3,193% z hodnoty 88,715%, pričom počet výpočtov na získanie hodnoty deskriptora jedného pixelu vzrastie vyše 3-násobne. Na interpolovanie intenzity jedného bodu je pritom potrebných podľa rovníc bilineárnej interpolácie 2.2 až 25 matematických operácií.

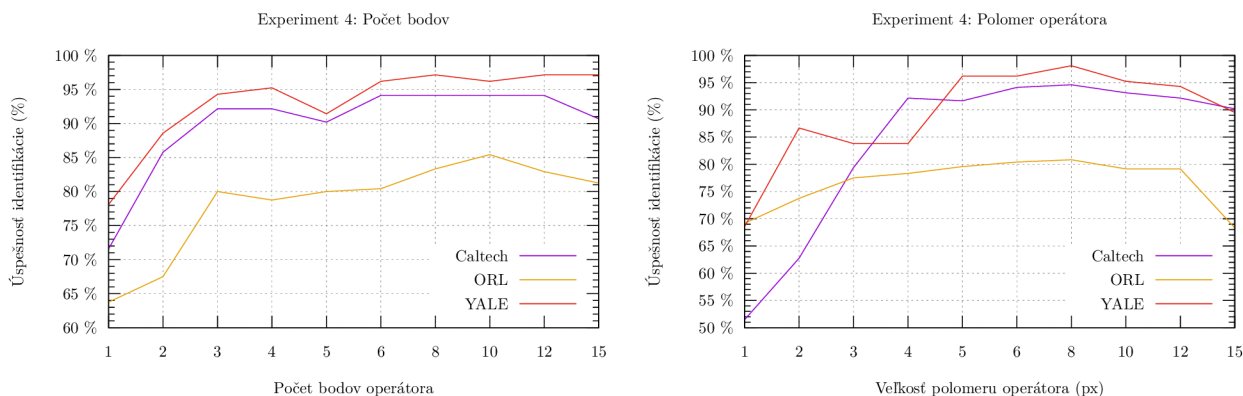
Aj keď veľkosť polomeru LBP operátora nemá vplyv na časovú náročnosť metódy, môžeme vidieť, že najlepšie výsledky dosahuje identifikácia na všetkých dátových sádach pri použití polomeru o veľkosti 8 pixelov. Pri použití polomeru pod 5 alebo nad 12 pixelov dochádza k výraznému zhoršeniu úspešnosti identifikácie.

4.4.4 Počet a rozsah košov HOG histogramu

Štvrtý experiment bol zameraný na zistenie, do akej miery vplýva počet a rozsah košov použitých v metóde HOG na vytvorenie lokálneho histogramu každého regiónu. Experiment bol nameraný v dvoch variantách – pre koše reprezentujúce celý rozsah $\langle -\pi, \pi \rangle$ alebo



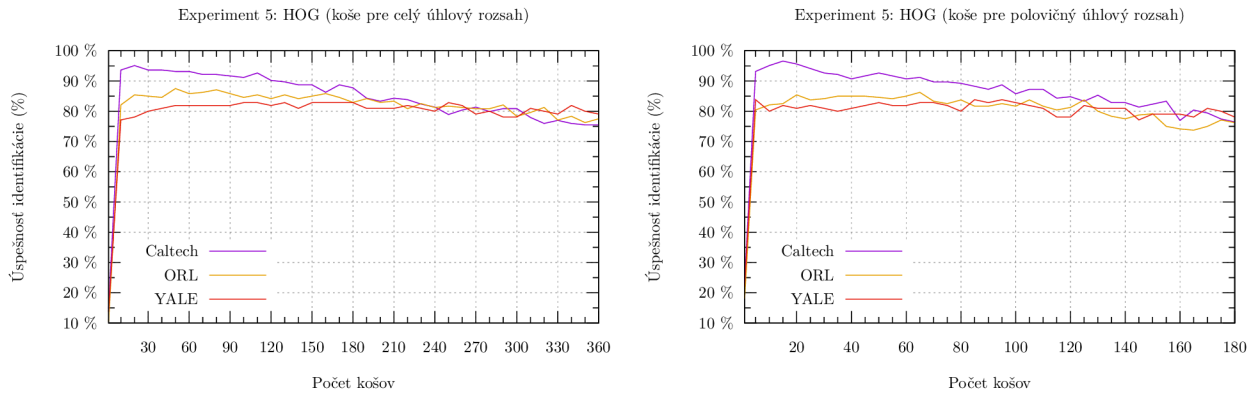
Obr. 4.7: Závislosť počtu regiónov vzhľadom na úspešnosť identifikácie metódami LBP a HOG na uzavrenej množine osôb troch rôznych dátových sád.



Obr. 4.8: Závislosť veľkosti a polomeru LBP operátora vzhľadom na úspešnosť identifikácie na uzavrenej množine osôb troch rôznych dátových sád.

polovičný rozsah $\langle 0, \pi \rangle$, pričom v druhej variante boli záporné hodnoty invertované prostredníctvom funkcie absolútnej hodnoty na kladné.

Vidíme, že v oboch prípadoch je na obrázku 4.9 vývoj približne podobný a priemerná úspešnosť sa v jednotlivých miestach grafu líši približne len o 1% v prospech varianty reprezentujúcej celý rozsah $\langle -\pi, \pi \rangle$. Tá je tiež v jej najlepšom mieste 50 košov s priemernou úspešnosťou identifikácie 87,514% lepšia len o 0,483% v porovnaní s najlepším miestom druhej varianty na 40 košoch.



Obr. 4.9: Závislosť počtu košov metódy HOG vzhľadom na úspešnosť identifikácie na uzavrenej množine osôb troch rôznych dátových sád.

4.4.5 Rozlíšenie analyzovaných snímok

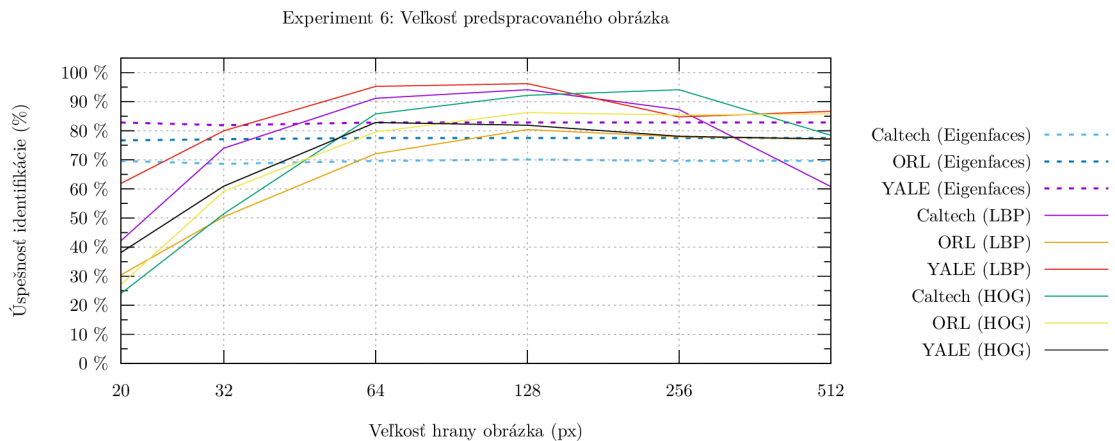
Posledný experiment zameraný na hľadanie optimálneho nastavenia parametru sa týkal veľkosti snímky, ktorá vstupuje do analýzy metódou. V rámci predspracovania sa tvár na snímke oreže na štvorcový tvar (kvôli možnosti použitia s metódou Eigenfaces, keďže zvyšné metódy nevyžadujú takúto podmienku) a zväčší, resp. zmenší, sa na požadovanú veľkosť hrany. Na obrázku 4.10 sú zaznačené hrany, pre ktoré bola úspešnosť na jednotlivých dátových sádach meraná. Najmenšia nameraná hodnota hrany 20 pixelov, je minimálna prípustná hodnota metódy LBP pri nastavení popísanom v podsekcii 4.2.1.

Z výsledkov vyplýva, že metódy LBP a HOG závisia na veľkosti použitej snímky, pričom v závislosti od dátovej sady môžeme vidieť najlepšie výsledky v oblasti merania 96 až 192 pixelov. Úspešnosť metód výrazne klesá pri použití rozlíšenia 64 pixelov a nižšom, a v prípade dátovej sady Caltech aj nad 128 pixelov. Pri zvyšných dvoch dátových sádach k výraznému poklesu do 256 pixelov nedochádza.

Zaujímavá je tiež vypozerovaná vlastnosť metódy analýza hlavných komponent, a tou je (až na drobné výkyvy) invariantnosť úspešnosti voči použitému rozlíšeniu vstupného obrázka.

4.5 Vyhodnotenie rozpoznávania pri použití optimálnych hodnôt parametrov

V poslednom experimente tejto sekcie sme použili sa inšpirovali hodnotami nameranými v predchádzajúcich podsekciiach, ktoré prinášali v jednotlivých metódach najlepšie výsledky. V porovnaní s konfiguráciou popísanou v podsekcii 4.2.1 to znamená nasledujúce zmeny:



Obr. 4.10: Závislosť použitej veľkosti strany štvorcovej snímky na úspešnosti identifikácie jednotlivých metód na uzavrenej množine osôb troch rôznych dátových sád.

- v metóde analýza hlavných komponent to znamená použitie maximálneho dostupného počtu vlastných tvárí,
- v prípade metódy LBP použitie 7×7 regiónov, 8 bodov operátora s polomerom 8 pixelov (empiricky bolo zistené, že v kombinácii s ostatnými nastaveniami prináša 8 bodov lepšie výsledky ako 10),
- v metóde HOG použitie 8×8 regiónov a 50 košov reprezentujúcich rozsah $\langle -\pi, \pi \rangle$.

Krivky jednotlivých dátových sád v grafoch na obrázkoch 4.4 a 4.5 sme spriemerovali tak, aby každú metódu reprezentovala jediná krivka vyjadrujúca jej priemernú úspešnosť vypočítanú zo všetkých použitých sád. Rovnakým postupom sme vypočítali krivky pri použití nových nastavení, a tieto krivky sme porovnali.

V grafe na obrázku 4.13 môžeme vidieť zmeny v úspešnosti identifikácie vyjadrenej prostredníctvom CMC kriviek. Na prvý pohľad je zrejmé, že vo všetkých troch prípadoch došlo k zlepšeniu úspešnosti minimálne na prvých troch pozíciách. V prípade metódy analýza hlavných komponent bolo pôvodné obmedzenie na použitie maximálne 40 vlastných tvárí posunuté na dostatočne vysoké číslo, aby mohli byť použité všetky, ktoré metóda detegovala. Ako už bolo konštatované v podsekcii 4.4.1, zanedbateľné zlepšenie o 0,58% na prvej pozícii na priemernú úspešnosť 77,40%, ku ktorému došlo, nie je vo väčšine aplikácií vhodné za cenu niekoľkonásobného zvýšenia počtu potrebných výpočtov na jeho dosiahnutie. Zlepšenie metód LBP a HOG je na prvých pozíciách už výraznejšie – o 1,15% na 91,39% a o 2,17% na 88,94%. Aktuálne hodnoty AUC pre tento experiment sú 0,949, 0,977 a 0,977 vymenované postupne pre metódy Eigenfaces, LBP a HOG.

Na obrázku 4.14 sú zasa zaznačené predchádzajúce priemerné a aktuálne ROC krivky. V prípade metód Eigenfaces a HOG ide o prakticky nepatrné zmeny, pričom v prípade prvej menovanej metódy prináša zmena konfigurácie horšie výsledky verifikácie ako v pôvodnej konfigurácii o 0,003 na 0,924. Metóda HOG si zmenou konfigurácie polepšila v hodnote AUC o 0,008 na 0,931. Najvýraznejšiu zmenu zaznamenala metóda LBP, ktorá pri použití nových parametrov zvýšila hodnotu oblasti pod ROC krivkou o 0,0107 na 0,934.

4.6 Analýza vzdialenosti šablón a rozpoznávanej snímky

Grafy na obrázkoch 4.11 a 4.12 zachytávajú namerané hodnoty vzdialeností rozpoznávanej tváre od šablón v databáze počas procesu identifikácie. Na grafe sú zaznačené rozpätia, ktoré sa v priebehu rozpoznávania danej dátovej sady získali, pričom je vždy zaznačená minimálna a maximálna vzdialenosť, ktorá bola v danej dátovej sade pri vyhodnocovaní nameraná a priemerná vzdialenosť zo všetkých nameraných vzdialeností. Dátová sada „*Neznáme*” bola zameraná na získanie hodnôt pri porovnávaní neznámych tvárí so šablónami v databáze. Ako tréningová množina bola použitá tréningová množina dátovej sady Caltech, a ako vyhodnocovacia množina boli použité snímky vyhodnocovacej množiny dátovej sady ORL. Jednotlivé zhluky zaznačení boli namerané s totožnou konfiguráciou, avšak kvôli prehľadnosti boli vykreslené s drobnými horizontálnymi odchýlkami.

Môžeme vidieť, že merania v oblasti rovnakej konfigurácie vykazujú podobné vlastnosti. Najdôležitejšou pre identifikáciu je minimálna vzdialenosť, ktorá sa výrazne vychýľuje z priemernej nameranej vzdialenosti a umožňuje dostatočne odlíšiť správnu šablónu v databáze. Z obrázku tiež vyplýva, že minimálna vzdialenosť neznámych tvárí je dostatočne vzdialená tej pre známe tváre, avšak maximálna vzdialenosť je typicky na približne rovnakej úrovni ako v prípade známych tvárí.

Taktiež si môžeme všimnúť zaujímavú závislosť. V prípade metódy analýza hlavných komponent sa so vzrastajúcim rozlíšením analyzovaných snímok mimoriadne predvídateľne posúva aj rozpätie hodnôt vzdialeností, s ktorými metóda pracuje. Vidíme, že s každým zdvojnásobením veľkosti snímky sa posunie rozsah o približne pol rádu nahor. To vysvetľuje vypozerovanú invariantnosť voči veľkosti obrázka popísanú v podsekcii 4.4.5. V prípade zvyšných dvoch metód už takáto pravidelnosť zrejme nie je. Môžeme si však všimnúť, že so vzrastajúcou veľkosťou snímky dochádza k rozostupu medzi minimálnou a maximálnou hodnotou vzdialenosti nameranej na danej dátovej sade, avšak k predvídateľnému posunu rozpätia nedochádza. Je tiež zaujímavé, že na rozdiel od Euklidovskej vzdialenosti, ktorá sa postupne zvyšuje, v prípade χ^2 -vzdialenosti dochádza so zväčšujúcim sa rozlíšením snímky k znižovaniu hodnôt vzdialeností.

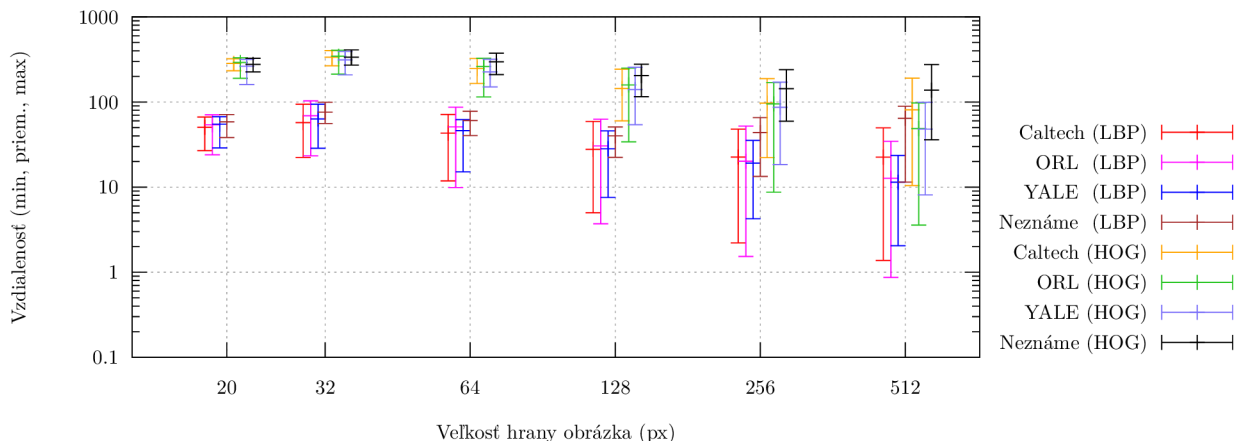
4.7 Rýchlosť a pamäťové nároky implementácie

Výsledky prezentované v nasledujúcej sekcii boli merané notebooku Asus[®] X555LN s 2-jadrovým procesorom Intel[®] Core[™] i5 1,6GHz na operačnom systéme Ubuntu 15.04 v konfigurácii s vypnutými ladiacimi symbolmi a kontrolami funkcie `assert()`.

Samotné grafické rozhranie aplikácie zaberalo po spustení v pamäti RAM 83 MB, 100,3 MB po načítaní 272 obrázkov dátovej sady Caltech, 101,6 MB pri 400 obrázkoch sady ORL a 91,9 MB pri 165 obrázkoch sady YALE. To činí približne 50 kB pamäti na každý načítaný obrázok veľkosti 128×128 pixelov. Detegovanie 40 vlastných tvárí a extrakcia charakteristických rysov metódou analýzy hlavných komponent si vyžiadala približne ďalších 10 MB, extrakcia metódou LBP približne 2,5 MB a extrakcia metódou HOG zhruba 3MB. Tieto čísla sú vyššie ako nutné minimum potrebné pre výpočet z dôvodu alokovania miesta aj pre komponenty vizualizujúce vlastné tváre a šablóny v databáze.

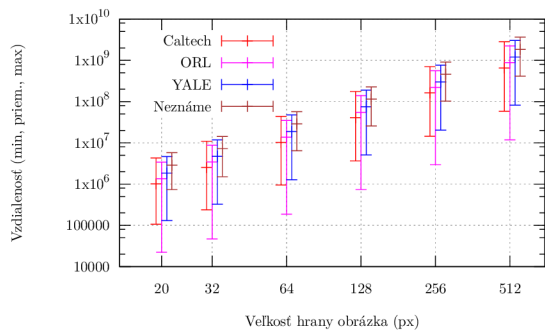
Rýchlosť spracovania jednotlivých metód je znázornená v grafoch na obrázku 4.15. Sú na nej porovnané priemerné časy potrebné na spracovanie jednej snímky vo fáze extrakcie a rozpoznávania pre jednotlivé metódy v konfigurácii popísanej v podsekcii 4.2.1 pri použití štvorcovej snímky o rozlíšení 128×128 a 512×512 pixelov.

Experiment 7: Vzdialenosť rozpoznávanej tváre od šablón (LBP a HOG)



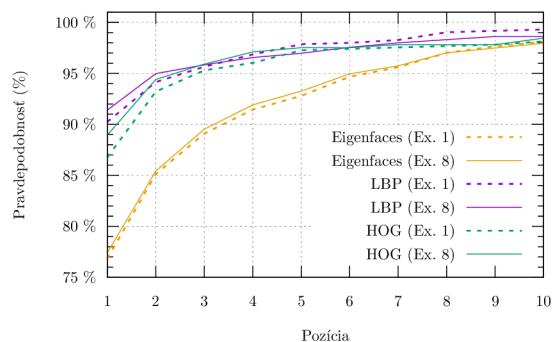
Obr. 4.11: Závislosť vzdialenosti medzi rozpoznávanou snímkou a šablónou z databázy na použitom rozlíšení snímky pri metódach LBP a HOG.

Experiment 7: Vzdialenosť rozpoznávanej tváre od šablón (Eigenfaces)



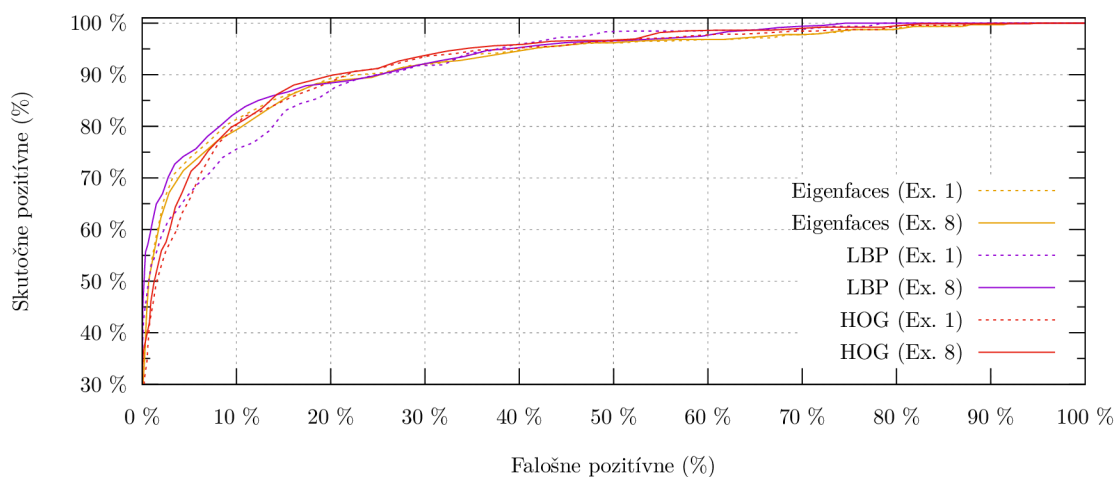
Obr. 4.12: Závislosť vzdialenosti medzi rozpoznávanou snímkou a šablónou z databázy na použitom rozlíšení snímky pri metóde Eigenfaces.

Experiment 8: CMC



Obr. 4.13: Porovnanie priemernej CMC krivky z experimentu č.1 a experimentu č.8 pre každú metódu.

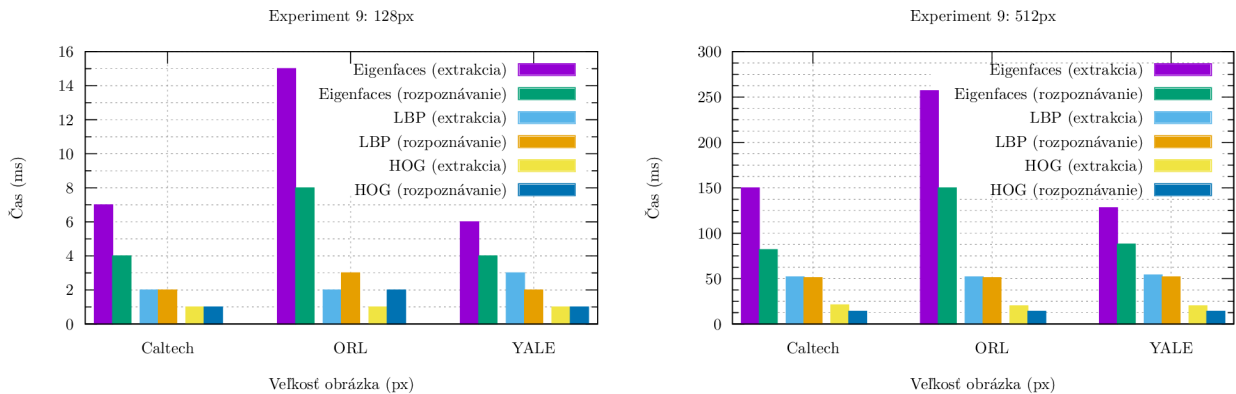
Experiment 8: ROC



Obr. 4.14: Porovnanie priemernej ROC krivky z experimentu č.1 a experimentu č.8 pre každú metódu.

Môžeme vidieť, že časy spracovania sú naprieč rôznymi rozlíšeniami vo svojich pomerných veľkostiach približne rovnaké, aj keď v prípade väčšieho rozlíšenia ide o zhruba 20-násobný nárast času potrebného na spracovanie každej snímky. Pripomeňme, že dátová sada Caltech obsahuje spolu 272 snímok 17-tich osôb, dátová sada ORL 400 snímok, na ktorých je 40 rôznych osôb, a dátová sada YALE 165 snímok s 15-timi osobami, pričom vo fáze extrakcie boli použité prvé 4 snímky každej osoby.

Je zrejmé, že suverénne najrýchlejšie spracovanie vykazuje metóda HOG a za ňou nasleduje LBP, ktoré si držia svoju rýchlosť približne rovnakú na všetkých dátových sadách. Analýza hlavných komponent je nielen najpomalšia, ale jej čas spracovania vzhľadom na veľkosť trénovacej množiny narastá do nepriateľných hodnôt pri práci s veľkými dátovými sadami.



Obr. 4.15: Porovnanie priemerného času spracovania jednej snímky pomocou jednotlivých metód pri používaní štvorcových obrázkov o veľkosti 128 a 512px.

4.8 Vyhodnotenie výsledkov experimentovania

Jednotlivé metódy vykazujú rozdielne kvalitatívne vlastnosti pri rozpoznávaní vzhľadom na rozdielne charakteristiky použitých dátových sád ale napríklad aj rôznu rýchlosť vyhodnocovania, a preto je ťažké označiť niektorú metódu za najlepšiu.

Vo všeobecnosti priniesla najlepšiu priemernú úspešnosť metóda lokálnych binárnych vzorov, kde v experimente č. 8 dosiahla úspešnosť identifikácie 91,39%, za ktorou nasledovala metóda histogramov orientovaných gradientov s o 2,45% horším výsledkom a až o ďalších vyše 10% nižšie sa umiestnila metóda analýzy hlavných komponent. Metóda LBP tiež dokázala ako jediná identifikovať dátovú sadu YALE (ktorej špecifickými komplikáciami boli zmeny osvetlenia, expresivita alebo prekrytie tváre okuliarmi) vo vhodnej konfigurácii až so 100%-nou úspešnosťou. Najväčší vplyv na úspešnosť identifikácie metódy LBP aj HOG má vhodne zvolený počet regiónov, pričom v prípade LBP prináša najoptimálnejšie výsledky rozdelenie 7×7 . Taktiež príliš nízky či vysoký počet bodov operátora LBP alebo jeho polomer môže výrazne znižovať úspešnosť vyhodnocovania, a preto je najlepšie použiť počet bodov približne 8 až 10 o polomere 8 pixelov.

Najmenej úspešná analýza hlavných komponent však oplýva vlastnosťou, ktorá ju v istých situáciách môže významne zvýhodniť – jej úspešnosť je invariantná voči veľkosti použitej snímky na rozpoznávanie. Takže, na rozdiel od zvyšných dvoch metód, si dokáže

spoľahlivo poradiť aj so vstupom o veľkosti len 20×20 pixelov. Najvýznamnejší vplyv na výsledky tejto metódy má vhodné nastavenie počtu vlastných tvárí použitých na rozpoznávanie, pričom aj keď väčší použitý počet prináša vyššiu úspešnosť, nad limitom počtu približne 40 vlastných tvárí ide o zanedbateľné zlepšenie.

Metóda HOG je zasa dostatočne obstojný konkurent metódy LBP, pričom jej bezsporne najväčšou výhodou je jej rýchlosť. Taktiež jej úspešnosť na dátovej sade ORL simulujúcej rozličné natočenie hlavy dosahovala oveľa lepšie výsledky ako v prípade LBP. Okrem nastavenia počtu regiónov, ktoré je podľa výsledkov našich meraní najvhodnejšie nastaviť na rozdelenie 8×8 , je tiež výhodné použitie nižšieho počtu košov lokálnych histogramov na úrovni obmedzenej približne na 50.

V oboch metódach LBP aj HOG, v ktorých je región popísaný histogramom a globálny deskriptor sa skladá z týchto lokálnych histogramov, ovplyvňuje použitý počet regiónov veľkosť vektora príznakov, a tým rýchlosť aj úspešnosť rozpoznávania. Použitím nízkeho počtu regiónov sa síce skráti dĺžka globálneho deskriptora a okrem nižšej spotreby pamäte sa zvýši aj rýchlosť rozpoznávania, daň za tieto výhody je však v podobe straty detailnejšej informácie o štruktúre tváre na analyzovanom obrázku. Naopak, pri použití príliš vysokého množstva regiónov pôsobí veľmi detailná informácia o štruktúre kontraproduktívne, takže okrem zvyšovania doby spracovania a spotreby pamäte kvôli extrémne dlhému globálnemu deskriptoru tiež začne klesať úspešnosť rozpoznávania.

4.9 Zhodnotenie riešenia a možnosti ďalšieho napredovania

Výsledky experimentov prevedených v tejto kapitole ukázali, že naša implementácia prináša riešenie, ktoré je po vhodnom nakonfigurovaní možné použiť v reálnych aplikáciách vyžadujúcich rozpoznávanie ľudskej tváre za účelom nielen identifikácie ale aj verifikácie osôb. Experimentami sme zistili, že každá metóda vykazuje iné vlastnosti, a vďaka tomu môže nájsť svoje špecifické uplatnenie. Taktiež je prijateľná aj rýchlosť rozpoznávania, ktorá konkrétne v prípade metódy histogramov orientovaných gradientov môže spomedzi všetkých troch analyzovaných metód najľahšie dosahovať na obmedzenej množine osôb a vhodnom zariadení vyhodnocovanie v reálnom čase.

Keďže naša implementácia bola naprogramovaná tak, aby bola skompilovateľná ako zdieľaná knižnica so závislosťou len na funkciách multiplatformnej knižnice *OpenCV*, je jednoduché využiť jej funkčnosť v aplikáciách cielených na počítače, mobilné zariadenia alebo napríklad aj herné konzoly.

Jedným z možných napredovaní vývoja tejto implementácie je napríklad úplne odstránenie závislosti na spomínanej knižnici *OpenCV* vlastnou implementáciou základných vektorových operácií nad obrázkovými reprezentáciami. V prípade záujmu o napredovanie vo výskumnom smere sa zasa ponúka príležitosť analyzovať chovanie implementácie a samotných metód aj na dátových sadách otvorenej množiny osôb, teda zahrnúť vo vyhodnocovaní aj tváre neznámych osôb. Existuje tiež veľké množstvo ďalších metód od využívania 3-rozmernej reprezentácie tváre na jej rozpoznávanie až po najnovšie poznatky v aplikovaní konvulčných neurónových sietí, ktorých vlastnosti by mohli byť podrobnejšie preskúmané v rôznych podmienkach použitia.

Záver

V tejto bakalárskej práci som sa zaoberal popisom, návrhom a implementáciou troch metód zameraných na extrakciu charakteristických rysov tváre v úlohe rozpoznávania osôb z fotografií a aplikáciou, ktorá umožňuje jednoduché experimentovanie s možnosťami konfigurácie, ktorú ponúkajú. Na jej vypracovanie sa bolo potrebné oboznámiť s informáciami problematiky biometrie tváre a postupne vyriešiť rôzne problémy, ktoré sa pri návrhu a samotnej realizácii objavili.

Medzi prvé potrebné kroky patrilo získanie prehľadu o aktuálnej situácii v tejto doméne, jej vývoji a princípoch, na ktorých je založená. Po identifikácii vhodných reprezentantov, ktorí sa stali predmetom tejto práce, bolo potrebné dohľadať a naštudovať vhodné publikácie, z ktorých bolo možné pochopiť fungovanie jednotlivých metód, matematické základy, na ktorých stoja, a čerpať informácie pre vytvorenie uceleného a pre čitateľa tejto práce pochopiteľného prehľadu ich princípu fungovania.

Následne som sa mohol začať zamýšľať nad možnosťami realizácie jednotlivých metód a navrhnúť algoritmy ich fungovania. Pri vytváraní týchto návrhov sa objavili ďalšie problémy komplikujúce realizáciu, pre ktoré bolo potrebné hľadať riešenia a nakoniec aj navrhnúť efektívne spôsoby zakomponovania týchto riešení do výsledného produktu.

Výsledkom tejto práce je vlastná implementácia multiplatformnej knižnice obsahujúcej 3 rôzne metódy určené na rozpoznávanie ľudskej tváre a tiež grafické rozhranie umožňujúce konfigurovanie jednotlivých špecifických parametrov týchto metód a testovanie úspešnosti identifikácie na uzavrenej množine osôb na zvolených dátových sadách. Dodatočne bolo naprogramované aj konzolové rozhranie umožňujúce parametrizovateľné spúšťanie a vyhodnocovanie nielen identifikačných ale aj verifikačných úloh pre účely experimentov.

Na naimplementovaných metódach boli prevedené experimenty, ktorých cieľom bolo vyhodnotiť ich úspešnosť na identifikačných aj verifikačných úlohách pri rozličných vlastnostiach použitých snímok tváre. Na tento účel boli zvolené tri voľne dostupné dátové sady nafotených snímok tvárí, z ktorých sa každá vyznačuje inými charakteristickými znakmi. V experimentoch bola ďalej meraná rýchlosť rozpoznávania a pamäťové nároky našej implementácie a tiež sledované, ako ovplyvňuje zmena rôznych parametrov jednotlivých metód úspešnosť identifikácie na uzavrenej množine osôb. Nakoniec boli nájdené parametre prinášajúce vo všeobecnosti najlepšie výsledky na nami používaných dátových sadách.

Z experimentovania vyplynulo, že najlepšie výsledky identifikácie aj verifikácie vykazuje metóda lokálnych binárnych vzorov, ktorej priemerná úspešnosť identifikácie dosiahla až 91,39%. Ide o metódu prinášajúcu prijateľný kompromis medzi úspešnosťou a rýchlosťou rozpoznávania. Metóda histogramov orientovaných gradientov, ktorá sa umiestnila s o 2,45% nižšou priemernou úspešnosťou na druhej pozícii, je zasa najrýchlejšou spomedzi všetkých metód a je najvhodnejšou na zakomponovanie do aplikácií vyžadujúcich rozpoznávanie v reálnom čase. V experimentoch najhoršie obstála jedna z historicky prvých metód určených na rozpoznávanie tvárí analýza hlavných komponent s priemernou úspešnosťou identifikácie

77,40%. Aj keď je tiež suverénne najpomalšia pri rozpoznávaní, jej výhodou v porovnaní s ostatnými metódami je jej nezávislosť na použitej veľkosti snímky tváre, a tak už od rozmerov 20×20 pixelov dokázala udržať úspešnosť identifikácie na prakticky nezmenenej hodnote.

Pri hľadaní odpovede na najlepšie nastavenie parametrov sme zistili, že metódu lokálnych binárnych vzorov, rovnako ako aj metódu histogramov orientovaných gradientov, výrazne ovplyvňuje počet použitých regiónov, z ktorých sú vytvárané lokálne histogramy. Toto nastavenie totiž vplýva na dĺžku vyextrahovaného vektora charakteristických rysov, ktorým je globálny deskriptor zložený z lokálnych histogramov, a tým pádom aj na detailnosť uchovania informácie o štruktúre tváre. Najlepšie výsledky prinieslo pravidelné rozdelenie o počte 7×7 , resp. 8×8 regiónov. V prípade prvej menovanej metódy je tiež dôležité správne nastavenie počtu použitých bodov a polomer LBP operátora, a to na približne 8 až 10 bodov a polomer 8 pixelov. Zvyšujúci sa počet použitých košov v druhej metóde zasa nepriamo úmerne vplýva na jej úspešnosť a najlepšie je ich použiť približne 50. V prípade analýzy hlavných komponent je jediným, a zároveň parametrom významne ovplyvňujúcim výsledky, počet vlastných tvárí použitých na rozpoznávanie, pričom najlepší kompromis medzi rýchlosťou a úspešnosťou prináša obmedzenie tohoto počtu na 40.

Literatúra

- [1] Ahonen, T.; Hadid, A.; Pietikäinen, M.: Face recognition with local binary patterns. In *European conference on computer vision*, Springer, 2004, s. 469–481.
- [2] Ahonen, T.; Hadid, A.; Pietikäinen, M.: Face description with local binary patterns: Application to face recognition. *IEEE transactions on pattern analysis and machine intelligence*, ročník 28, č. 12, 2006: s. 2037–2041.
- [3] Atlassian: *BitBucket*. [Online; navštívené 18.4.2017].
URL <https://bitbucket.org/>
- [4] AT&T Laboratories Cambridge: *The Database of Faces*. [Online; navštívené 12.3.2017].
URL http://www.cl.cam.ac.uk/Research/DTG/attarchive/pub/data/att_faces.tar.Z
- [5] Belhumeur, P. N.; Hespanha, J. P.; Kriegman, D. J.: Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Transactions on pattern analysis and machine intelligence*, ročník 19, č. 7, 1997: s. 711–720.
- [6] Bityukov, S.; Maksimushkina, A.; Smirnova, V.: Comparison of histograms in physical research. *Nuclear Energy and Technology*, ročník 2, č. 2, 2016: s. 108–113.
- [7] Chen, Y.: *Student Work: Sobel Filter*. [Online; navštívené 18.4.2017].
URL <http://www.seas.upenn.edu/~cis565/StudentWork.htm>
- [8] Dalal, N.; Triggs, B.: Histograms of oriented gradients for human detection. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, ročník 1, IEEE, 2005, s. 886–893.
- [9] Déniz, O.; Bueno, G.; Salido, J.; aj.: Face recognition using histograms of oriented gradients. *Pattern Recognition Letters*, ročník 32, č. 12, 2011: s. 1598–1603.
- [10] Diniz, F. A.; Neto, F. M. M.; Júnior, F. d. C. L.; aj.: A Facial Recognition System Based on Techniques of Principal Component Analysis and Autofaces with K-NN, K-Star and Random Forest Classifiers. *Research Notes in Information Science*, ročník 12, 2013: str. 7.
- [11] Fisher, R. A.: The use of multiple measurements in taxonomic problems. *Annals of eugenics*, ročník 7, č. 2, 1936: s. 179–188.
- [12] Gnuplot: *A portable command-line driven multiplatform graphing utility*. [Online; navštívené 2.5.2017].
URL <http://www.gnuplot.info/>

- [13] Guyon, I.; Elisseeff, A.: An introduction to feature extraction. In *Feature extraction*, Springer, 2006, s. 1–25.
- [14] Güneş, F.: *Principal Component Analysis for Dimensionality Reduction*. [Online; navštívené 15.12.2016].
URL <http://blogs.sas.com/content/subconsciousmusings/2015/10/26/principal-component-analysis-for-dimensionality-reduction/>
- [15] Hadid, A.: The local binary pattern approach and its applications to face analysis. In *2008 First Workshops on Image Processing Theory, Tools and Applications*, IEEE, 2008, s. 1–9.
- [16] Jafri, R.; Arabnia, H. R.: A survey of face recognition techniques. 2009.
- [17] Lata, Y. V.; Chandra, K.; Tungathurthi, B.; aj.: Facial recognition using eigenfaces by PCA. 2009.
- [18] Martínez, A. M.: Recognizing expression variant faces from a single sample image per class. In *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*, ročník 1, IEEE, 2003, s. I–353.
- [19] Martínez, A. M.; Kak, A. C.: Pca versus lda. *IEEE transactions on pattern analysis and machine intelligence*, ročník 23, č. 2, 2001: s. 228–233.
- [20] Microsoft News: *Microsoft Xbox One With Kinect Can Map Your Face To In-game Models*. [Online; navštívené 28.12.2016].
URL <https://mspoweruser.com/microsoft-xbox-one-with-kinect-can-map-your-face-to-in-game-models/>
- [21] Ojala, T.; Pietikäinen, M.; Harwood, D.: A comparative study of texture measures with classification based on featured distributions. *Pattern recognition*, ročník 29, č. 1, 1996: s. 51–59.
- [22] Ojala, T.; Pietikainen, M.; Maenpaa, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on pattern analysis and machine intelligence*, ročník 24, č. 7, 2002: s. 971–987.
- [23] Poh, N.; Chan, C.; Kittler, J.; aj.: D3. 3: Description of Metrics For the Evaluation of Biometric Performance. *Evaluation*, ročník 1, 2011.
- [24] Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; aj.: Numerical Recipes 3rd Edition: The Art of Scientific Computing. 2007.
- [25] Rahim, M. A.; Azam, M. S.; Hossain, N.; aj.: Face Recognition using Local Binary Patterns (LBP). *Global Journal of Computer Science and Technology*, ročník 13, č. 4, 2013.
- [26] Reid, P.: *Biometrics and Network Security*. Prentice Hall PTR, 2003, ISBN 0131015494.
- [27] Salhi, A. I.; Kardouchi, M.; Belacel, N.: Fast and efficient face recognition system using random forest and histograms of oriented gradients. In *Biometrics Special Interest Group (BIOSIG), 2012 BIOSIG-Proceedings of the International Conference of the*, IEEE, 2012, s. 1–11.

- [28] Scikit-learn: *Machine Learning in Python*. [Online; navštívené 2.5.2017].
URL <http://scikit-learn.org/stable/>
- [29] Sirovich, L.; Kirby, M.: Low-dimensional procedure for the characterization of human faces. *Josa a*, ročník 4, č. 3, 1987: s. 519–524.
- [30] Sobel, I.; Feldman, G.: *An Isotropic 3x3 Image Gradient Operator*. [Online; navštívené 17.4.2017].
URL https://www.researchgate.net/publication/239398674_An_Isotropic_3_3_Image_Gradient_Operator
- [31] Tan, X.; Chen, S.; Zhou, Z.-H.; aj.: Face recognition from a single image per person: A survey. *Pattern recognition*, ročník 39, č. 9, 2006: s. 1725–1745.
- [32] Turk, M.; Pentland, A.: Eigenfaces for recognition. *Journal of cognitive neuroscience*, ročník 3, č. 1, 1991: s. 71–86.
- [33] UCSD Computer Vision: *Yale Face Database*. [Online; navštívené 17.3.2017].
URL http://vision.ucsd.edu/datasets/yale_face_dataset_original/yalefaces.zip
- [34] Vincent, O.; Folorunso, O.: A descriptive algorithm for sobel image edge detection. In *Proceedings of Informing Science & IT Education Conference (InSITE) 2009*, InSITE, 2009.
- [35] Viola, P.; Jones, M. J.: Robust real-time face detection. *International journal of computer vision*, ročník 57, č. 2, 2004: s. 137–154.
- [36] Weber, M.: *Frontal face dataset*. [Online; navštívené 11.3.2017].
URL http://www.vision.caltech.edu/Image_Datasets/faces/faces.tar
- [37] Wegner, P.: A technique for counting ones in a binary computer. *Communications of the ACM*, ročník 3, č. 5, 1960: str. 322.
- [38] Woollaston, V.: *Facebook can tag you in photos AUTOMATICALLY: Social network starts rolling out DeepFace recognition feature*. [Online; navštívené 28.12.2016].
URL <http://www.dailymail.co.uk/sciencetech/article-2946186/Facebook-soon-tag-photos-AUTOMATICALLY-Social-network-starts-rolling-DeepFace-feature.html>
- [39] Yin, R. K.: Looking at upside-down faces. *Journal of experimental psychology*, ročník 81, č. 1, 1969: str. 141.

Zoznam príloh

Príloha A: Popis použitia konzolovej aplikácie

Obsah CD

- Zdrojové kódy implementácie tejto práce v adresári `/src/`
- Zdrojové kódy experimentov, použité dátové sady a výsledky experimentov v adresári `/experiments/`
- Táto práca vo formáte PDF v adresári `/thesis/`
- Zdrojové kódy tejto práce vo formáte systému \LaTeX v adresári `/thesis/latex/`

Príloha A: Popis použitia konzolovej aplikácie

- `-h` , `--help` – zobrazenie nápovedy
- `-v` , `--version` – zobrazenie verzie programu
- `--eigenfaces` – použije pri vyhodnocovaní metódu analýzy hlavných komponent
- `--lbp` – použije pri vyhodnocovaní metódu lokálnych binárnych vzorov
- `--hog` – použije pri vyhodnocovaní metódu histogramov orientovaných gradientov
- `-t <cesta>`, `--training-directory <cesta>` – cesta k priečinku s „trénovacou“ množinou snímok
- `-e <cesta>`, `--evaluation-directory <cesta>` – cesta k priečinku s „vyhodnocovacou“ množinou snímok
- `-o <súbor>`, `--output-file <súbor>` – CSV súbor s výsledkami vyhodnocovania
- `--img-preprocess` – povolí predspracovanie obrázkov pri ich načítavaní
- `--img-detectface` – zapne orezanie na detekovanú tvár v obrázku
- `--img-dimension <int>` – zmení veľkosť obrázku na štvorcový tvar o veľkosti pixelov špecifikovanej nastavením
- `--img-crop` – zapne orezanie na štvorcový tvar namiesto východzieho rozťahnutia obrázku
- `--ef-count <int>` – obmedzí použitie vlastných tvárí na maximálne počet špecifikovaný nastavením
- `--ef-isfacethreshold <float>` – nastavenie prahu pre určenie či ide o tvár v metóde analýzy vlastných komponent
- `--XXX-wregions <int>` – špecifikácia počtu regiónov použitých na šírku v metóde `lbp` alebo `hog`
- `--XXX-hregions <int>` – špecifikácia počtu regiónov použitých na výšku v metóde `lbp` alebo `hog`
- `--XXX-weights <string>` – pole váh jednotlivých regiónov v metóde `lbp` alebo `hog` oddelených čiarkou

- `--lbp-points <int>` – špecifikácia použitého počtu bodov LBP operátora
- `--lbp-radius <int>` – špecifikácia použitého polomeru LBP operátora v pixeloch
- `--hog-bins <int>` – špecifikácia použitého počtu košov v metóde HOG
- `--hog-twoangled` – použitie polovičného rozsahu a mapovanie záporných úhlov do kladného rozsahu v metóde HOG
- `--verification-steps <int>` – použitie verifikácie vyhodnotenej so špecifikovaným počtom posunutí práhu namiesto identifikácie