

Ing. Marek Loderer

Autoreferát dizertačnej práce

DETEKCIA A ROZPOZNÁVANIE ĽUDSKÝCH TVÁRÍ

na získanie akademickej hodnosti doktor (philosophiae doctor, PhD.)

v doktorandskom študijnom programe: Robotika a kybernetika
v študijnom odbore: 9.2.7 kybernetika

Miesto a dátum: Bratislava, 28. 10. 2016

**SLOVENSKÁ TECHNICKÁ UNIVERZITA V BRATISLAVE
FAKULTA ELEKTROTECHNIKY A INFORMATIKY**

Ing. Marek Loderer

Autoreferát dizertačnej práce

DETEKCIA A ROZPOZNÁVANIE ĽUDSKÝCH TVÁRÍ

na získanie akademickej hodnosti doktor (philosophiae doctor, PhD.)

v doktorandskom študijnom programe:
Robotika a kybernetika

Miesto a dátum: Bratislava, 28. 10. 2016

Dizertačná práca bola vypracovaná v dennej forme doktorandského štúdia

Na Ústave robotiky a kybernetiky
Fakulta elektrotechniky a informatiky, STU
Ilkovičova 3, 812 19 Bratislava

Predkladateľ: Ing. Marek Loderer
Ústava robotiky a kybernetiky
Fakulta elektrotechniky a informatiky, STU
Ilkovičova 3, 812 19 Bratislava

Školiteľ: prof. Ing. Jarmila Pavlovičová, PhD.
Ústava robotiky a kybernetiky
Fakulta elektrotechniky a informatiky, STU
Ilkovičova 3, 812 19 Bratislava

Oponenti: prof. Ing. Jana Tučková, CSc.
Katedra teórie obvodů
Fakulta elektrotechnická, ČVUT
Technická 2, 166 27 Praha 6

Ing. Vanda Benešová, CSc.
Ústav počítačového inžinierstva a aplikovanej informatiky
Fakulta informatiky a informačných technológií, STU
Ilkovičova 2, 842 16 Bratislava

Autoreferát bol rozoslaný:

Obhajoba dizertačnej práce sa koná: o h.

na Fakulte elektrotechniky a informatiky STU,
Ilkovičova 3, 812 19 Bratislava,
v miestnosti

prof. Dr. Ing. Miloš Oravec
dekan fakulty
Slovenská technická univerzita v Bratislave
Fakulta elektrotechniky a informatiky

Abstrakt

Dizertačná práca sa zaoberá aktuálnou problematikou detekcie pokožky, tváří, tvárových črt v obraze, klasifikáciou tváří a emócií.

Úvodná časť sa venuje biometrii ako vede a jednotlivým biometrikám. Nasleduje popis rôznych fáz činnosti biometrických systémov a spôsoby vyhodnotenia ich presnosti, bežne používané databázy tváří, priestory farieb, rôzne metódy extrakcie príznakov, redukcia dimenzie, klasifikácia, zhlukovacie a optimalizačné algoritmy. Publikované riešenia uvedené v prehľade súčasného stavu pokrývajú oblasti ako detekcia ľudskej pokožky, detekcia a rozpoznávanie tváří, rozpoznávanie emócií.

Vlastný výskum sa venuje viacerým oblastiam, ktoré boli analyzované a experimentálne overované. Sú nimi využitie Gáborových filtrov pri lokalizácii tvárových črt, automatizovaný výber reprezentatívnych tréningových vzoriek získaných z kamery, optimalizácia počtu a nastavenia parametrov Gáborových filtrov s cieľom lokalizácie tvárových črt, optimalizácia nastavenia parametrov za účelom výraznej redukcie dimenzie dát, identifikácia významných oblastí tváre pre účely rozpoznávania subjektov a emócií. V závere kapitoly sú uvedené možnosti ďalšieho využitia ako napr.: rozpoznávanie pomocou snímky ucha a detekcia chorobných zmien na sietnici.

V poslednej kapitole sú zhrnuté prínosy tejto práce.

Kľúčové slová: Biometria, biometrický systém, databázy tváří, priestory farieb, detekcia pokožky, detekcia tváří, rozpoznávanie tváří, rozpoznávanie emócií, extrakcia príznakov, redukcia dimenzie, klasifikácia, zhlukovacie algoritmy, optimalizačné algoritmy

Obsah

1	Úvod	1
1.1	Motivácia	1
2	Prehľad publikovaných riešení	2
2.1	Detekcia tvárí	2
2.2	Rozpoznávanie tvárí	2
2.3	Detekcia a rozpoznávanie tvárí	3
2.4	Rozpoznávanie emócií	4
2.5	Významné oblasti tváre	4
2.6	Detekcia chorobných zmien na sietnici	5
3	Ciele dizertačnej práce	6
4	Vlastný výskum	7
4.1	Využitie GF pri lokalizácii tvárových črt	7
4.2	Optimalizácia banky GF pre účely lokalizácie tvárových črt	9
4.3	Automatický výber vzoriek a spôsoby redukcie dimenzie dát	11
4.4	Optimalizácia LBP parametrov	15
4.5	Dôležitosť jednotlivých oblastí tváre	19
4.6	Porovnanie dosiahnutých výsledkov	21
4.7	Možnosti využitia v ďalších oblastiach	23
4.7.1	Rozpoznávanie subjektov pomocou snímky ucha	23
4.7.2	Detekcia chorobných zmien na sietnici	24
5	Prínosy práce	25
	Literatúra	26
	Vlastné publikácie	32
	Účasť na projektoch	34

1 Úvod

Biometria a biometrické systémy sa čoraz viac dostávajú do popredia ľudského záujmu. Na rozdiel od minulosti, kedy sa prevažne využívala na forenzné vyšetovanie kriminálnych činov, sa s ňou čoraz častejšie stretávame aj v bežnom živote. Súvisí to v prevažnej miere s pokrokom, ktorý je spätý s výpočtovou technikou. Sú to napr. notebooky alebo oveľa menšie smartfóny, ktoré vzhľadom na svoju veľkosť dosahujú značný výpočtový výkon.

Táto práca sa venuje problematike detekcie a rozpoznávania ľudských tvárí. Tvár je bezpochyby súčasťou nášho každodenného života. Aj keď človek dokáže nájsť a rozpoznať ľudské tváre, či už na fotografii alebo vo videu bez najmenších problémov, v prípade výpočtovej techniky to už také jednoduché nie je. Zo vzhľadu tváre vieme veľmi ľahko určiť pohlavie, vek, emočný stav či dokonca charakter človeka. Táto schopnosť sa v nás vyvíja od útleho detstva.

Z veľkého množstva známych algoritmov, ktoré je možné využiť, budú v práci popísané niektoré matematické a štatistické metódy, zhlukovacie a optimalizačné algoritmy, umelé neurónové siete a Gáborove filtre (GF).

Aby bolo možné rôzne algoritmy testovať, či dokonca porovnávať, je potrebné využiť rovnaký zdroj dát. Práve na tento účel vznikli a stále vznikajú nové databázy ľudských tvárí. Tieto databázy vznikli s cieľom riešiť určitý problém. Sú to napr. problém detekcie pokožky, rozpoznávanie osôb a emócií, určenie pohlavia, rasy. Okrem toho treba brať do úvahy rôzne okolnosti, ktoré negatívne vplývajú na úspešnosť algoritmov. Ide napr. o svetelné podmienky, natočenie, vzdialenosť od senzoru, pozadie, ale aj maskovanie či dokonca problém len jednej dostupnej vzorky.

Samotný biometrický systém počas svojej existencie prechádza rôznymi fázami. Najprv je potrebné pridať používateľov a systém sa musí tzv. natrénovať. Proces tréningu môže byť výpočtovo a časovo náročný. Existuje mnoho metód extrakcie príznakov, ktoré sa snažia zredukovať priestor vstupných dát a tým prispieť k zníženiu výpočtovej náročnosti.

1.1 Motivácia

Rôzne výskumy a množstvo publikácií ukazuje, že táto oblasť je už do značnej miery preskúmaná. Ale stále je tu priestor pre nové riešenia a úspešnejšie zlepšenia. Biometria ako veda je nástroj, ktorý slúži k spojeniu človeka (biologických procesov) so strojmi. Má obrovskú perspektívu pri využití v bežnom živote. Konkrétnejšie sú to oblasti telekomunikácií, bezpečnosti (bezpečnosť informačných systémov), interakcie človeka s počítačom, robotiky, kybernetiky a mnohé iné. Spomenuté oblasti vyžadujú správne a presné informácie, aby bolo možné dosiahnuť čo najvyššiu úspešnosť. Ďalej je potrebné použiť také postupy a metódy, aby bol daný systém čo najrýchlejší (práca v reálnom čase), v niektorých prípadoch, aby bol schopný pracovať aj na menej výkonných zariadeniach a vysporiadať sa s už spomenutými „komplikáciami“. Preto si daná oblasť stále zasluhuje našu pozornosť.

2 Prehľad publikovaných riešení

Táto kapitola je venovaná prehľadu publikovaných článkov od roku 2005. Obsahuje rôzne spôsoby ako riešiť problém detekcie ľudskej pokožky, detekcie a rozpoznávania tváří. Nasleduje časť zaoberajúca sa problematikou rozpoznávania emócií. Ďalej je to podkapitola venujúca sa detekcii významných oblastí tváre. V závere kapitoly sú uvedené ďalšie oblasti, kde dané prístupy možno využiť (napr. biomedicínske obrazy).

2.1 Detekcia tváří

V prácach [CHU08], [KIT08], [PER11], [BON13] sa autori zaoberali problémom detekcie tváří.

Autori [CHU08] k detekcii tváre využili banku 40 Gáborových filtrov a neurónovú sieť. Sieť sa trénuje na základe masky, ktorá určuje pozíciu tváre v obraze.

Metóda v [KIT08] je tiež založená na príznakoch extrahovaných bankou Gáborových filtrov a ich hodnotení vzhľadom na viaceré testovacie hypotézy (angl. multiple hypothesis testing). Navrhovaná metóda je schopná vysporiadať sa s natočením a rôznou veľkosťou tváre.

Práca [PER11] je venovaná detekcii frontálnych a profilových tváří. Využili štyri rôzne metódy extrakcie príznakov: LBP (angl. Local binary patterns), PCA (angl. Principal components analysis), integrálne histogramy a optimalizované pomery tváre. Posledný typ príznakov predstavuje pomery priemerného jasú medzi dvojicami rôznych oblastí tváre. Pozícia a veľkosť týchto oblastí spolu s parametrami SVM (angl. Support vector machine) klasifikátora boli určené pomocou genetického algoritmu.

V práci [LI12], ktorá sa síce nezaobrá detekciou tváří, je publikovaných niekoľko zaujímavých zmien klasického kaskádového detektora. Sú to nasledujúce úpravy: LBP príznaky slúžia pre výpočet integrálneho obrazu, zmena zastavovacieho kritéria pre zostavenie jednej vrstvy kaskády, podmienka na automatické určenie počtu vrstiev kaskády a spôsob ako určiť skóre preverovanej oblasti.

Práca [BON13] sa okrem detekcie tváre zaoberá aj detekciou ľudských postáv. Využíva LBP, LGP (angl. Local Gradient Patterns) a BHOG (angl. Binary histograms of oriented gradients) príznaky. Príznaky sú klasifikované pomocou kaskádového detektora. Najlepšie výsledky pri detekcii tváre dosiahla kombinácia spomínaných príznakov.

2.2 Rozpoznávanie tváří

Nasledujúce tri práce [MAG11], [RAG13], [ALY10] sa venujú problematike rozpoznávania tváří.

V práci [MAG11] autori použili banku 40 Gáborových filtrov. Výstup po konvolúcii bol podvzorkovaný. Dimenzia dát bola znížená pomocou PCA a LDA (angl. Linear discriminant analysis) metódy. Ako klasifikátor bola použitá trojvrstvová MLP (angl. Multilayer perceptron) sieť.

Práca [RAG13] je doplnením predošlej. Obsahuje porovnanie úspešnosti rozpoznávania na základe dvoch metrík, neurónovej siete a SVM. Posledný spomínaný spôsob klasifikácie

dosiahol najvyššiu úspešnosť 98%.

Podľa autorov [ALY10] vhodnejším nástrojom redukcie dát generovaných Gáborovými filtrami je MSOM sieť (angl. Multiple self-organizing map) skladajúca sa z matice 16*16 SOM sietí. Kombinácia metódy MSOM a navrhovanej metriky dosahuje v priemere o 15% lepšiu úspešnosť ako kombinácia pomocou MSOM s KNN (angl. k-Nearest neighbors algorithm) alebo PCA s KNN.

Nasledujúci prehľadový článok [HUA11] obsahuje prehľad používaných LBP metód v rôznych oblastiach. Zahŕňa prehľad viac než 100 príspevkov. Od pôvodného LBP operátora vzniklo mnoho obmien, pričom každá sa zameriava na určitý problém. Autori poukazujú na problematiku nárastu dimenzie dát. To je možné riešiť pomocou viacerých spôsobov. Ďalšia oblasť nevyhnutná pre extrakciu príznakov je veľkosť, tvar, prekrytie oblastí. Môže byť využitý rôzny počet vzorkovacích bodov alebo priemer. V prehľade sa nachádzajú rôzne databázy tvárí, na ktorých boli vykonané experimenty spolu s dosiahnutou úspešnosťou rozpoznávania.

Nasledujúci článok [SUR12] obsahuje porovnanie 8 rôznych obmien LBP algoritmu. Ako klasifikátor slúžil KNN algoritmus, využívajúci ako metriku tzv. G štatistiku. Prvý experiment bol zameraný na klasifikáciu neutrálnych výrazov z JAFFE databázy, v ďalšom experimente boli obrazy poškodené rôznymi druhmi šumu, tiež boli preverené úspešnosti pre natočené obrazy tvárí z YALE databázy. Ďalej bola preverená invariantnosť voči rôznym svetelným podmienkam. Posledný experiment sledoval úspešnosť klasifikácie emócií.

2.3 Detekcia a rozpoznávanie tvárí

V nasledujúcich príspevkoch [JEM08], [LOP10], [CHI13], [ALA12] sa nachádzajú zdokumentované popisy činnosti automatizovaných biometrických systémov, ktoré riešia problémy spojené s detekciou a rozpoznávaním tvárí.

V [JEM08] autori predstavili automatizovaný systém. Využíva metódy ako predspracovanie, extrakciu rôznych príznakov a klasifikáciu na základe MLP sietí. Vstupom je v tomto prípade farebný obraz. Extrakcia oblastí pokožky v priestore YCbCr slúži na nájdenie polohy tváre. Cannyho detektor nájde hrany okolo tváre a pomocou pomeru výšky a šírky sa všetky obrazy znormálnujú na rovnaký rozmer. Poloha očí a úst závisí od hodnôt oboch chrominančných zložiek. Nosu je priradená poloha v strede medzi očami a ústami. Aplikovaním Sobelovho gradientného operátora dostali polohu 10 významných bodov tváre. Medzi týmito bodmi je možné definovať 7 vzdialeností, ktoré budú predstavovať jeden zo spôsobov extrakcie príznakov. V meraniach využili banku Gáborových filtrov. Testovali aj vplyv rôzneho počtu natočení a taktiež rôzneho počtu filtrov v banke. Pre účely klasifikácie využili dve architektúry MLP siete, ktoré majú svoje výhody a obmedzenia.

Práca [LOP10] sa podrobne zaoberá využitím LBP príznakov pre účely detekcie pokožky, detekciu a rozpoznávanie tvárí. Navrhovaný systém sa skladá z troch modulov. Autor použil rôzne metódy, ktoré sa v jednotlivých moduloch líšia. Na základe meraní a porovnaní je možné vyvodiť nasledujúce závery. V prípade presnosti detekcie pokožky bola zásadná kvalita obrazov. Na účely detekcie tváre postačoval samotný SVM klasifikátor, ktorý prekonal klasifikáciu príznakov redukovaných pomocou PCA. Odolnosť systému voči natočeniu je možné dosiahnuť pridaním natočených vzoriek do tréningového procesu. Použitie χ^2 metriky na porovnávanie LBP histogramov dosiahlo v procese rozpoznávania dostatočnú úspešnosť. Úspešnosť metódy zvýšilo predspracovanie obrazov v podobe ekvalizácie histogramu, maskovanie rohov obrazu a váhovanie jednotlivých blokov obrazu. Redukcia dimenzie LBP príznakov pomocou PCA dosahuje dobré výsledky. Problém identifikácie nového subjektu je možné vyriešiť použitím prahovej hodnoty v procese rozpoznávania.

V práci [CHI13] autori vytvorili automatizovaný dochádzkový systém. Prítomnosť subjektu je detegovaná pomocou Viola-Jones detektora. Získaný obraz tváre sa predspracuje. Príznaky sú extrahované pomocou PCA, LDA, LBP. Euklidovská vzdialenosť, SVM, Baiyesov

klasifikátor (angl. naive Baiyes classifier) budú slúžiť ako klasifikátor. Na základe výsledkov najmenšiu hodnotu chybného prijatia (FAR) dosiahla kombinácia LBP a Euklidovskej vzdialenosti. Pri použití spomenutých metód trvalo tréovanie najkratšie. Využitie SVM alebo Baiyesovho klasifikátora však dosahuje vyššiu úspešnosť rozpoznávania.

Práca [ALA12] sa zaoberá detekciou tváří, extrakciou príznakov a vyhľadávaním tváří v databáze. Rozdiel spočíva v tom, že pri klasifikácii sa systém snaží nájsť rovnaké obrazy, pri vyhľadávaní sa snaží nájsť podobné obrazy. Aj keď ide o mierne odlišnú oblasť, poskytol tento článok zaujímavé prístupy. Na detekciu tváre bol použitý kaskádový detektor. Boli použité tri rôzne priestory farieb. Ako spôsob extrakcie príznakov využili histogram farebných obrazov (obdoba LBP) a PCA. Ukázalo sa, že PCA príznaky založené na obrazoch v HSV a RGB priestore dosiahli dobré výsledky.

2.4 Rozpoznávanie emócií

V nasledujúcich prácach [FAZ09], [LAJ08], [LEE10], [YIN09], [VER13] sa autori venovali problematike rozpoznávania emócií.

Autori [FAZ09] použili databázu Cohn-Kanade, ktorá zachytáva rôzne výrazy tváre. Obrazy manuálne predspracovali a znormalizovali. Využili banku 40 Gáborových filtrov. Výsledok po aplikovaní banky filtrov podvzorkovali a znormalizovali, čím získali vektor príznakov o veľkosti 10240 hodnôt. Pomocou metódy PCA zredukovali ich počet na 138 hodnôt. Tie slúžili ako vstup do LDA metódy, ktorej výstupom je 5 hodnôt. Ako klasifikátor bola použitá PNN (angl. Probabilistic neural network) sieť, ktorá je variantou RBF (angl. Radial basis function) siete. Najlepšie klasifikované emócie boli „prekvapenie“ a „radosť“.

Ďalšia práca, ktorá sa zaoberá rozpoznávaním emócií [LAJ08], však použila banku 40 logaritmických Gáborových filtrov. Použili rovnakú databázu obrazov ako v predošlej práci. Na redukciu dlhého vektora príznakov použili podvzorkovanie a metódu PCA. Jej výstupom bolo 29 hlavných komponentov. Na klasifikáciu využili 6 MLP sietí (pre každú emóciu jedna sieť). Výsledná klasifikácia do triedy bola založená na maximálnej hodnote odpovede niektorej zo sietí. Výsledky však ani zďaleka nedosiahli hodnoty uvedené v [FAZ09]. Najlepšie klasifikovaná bola opäť emócia „prekvapenie“. Zaujímavý je spôsob výberu optimálnych filtrov.

Práca [LEE10] venovaná rozpoznávaniu výrazov tváre používa kombináciu Gáborových filtrov a RBF siete. RBF sieť však obsahuje niekoľko navrhovaných zmien. Zdrojom obrazov bola databáza JAFFE. Výsledkom je zistenie, že dáta redukované pomocou entropie (10%, 20% a 30% veľkosti obrazu spracovaného Gáborovým filtrom) dosahujú najvyššiu úspešnosť.

Autori v práci [YIN09] využili kombináciu metód LBP a SVDA. Základný princíp činnosti SVDA (angl. Support vector discriminant analysis) sa podobá metóde SVM. Zdrojom údajov bola JAFFE databáza. Obrazy boli spracované LBP a rozdelené do blokov tromi rôznymi spôsobmi. Najlepšie výsledky dosiahol LBP histogram pozostávajúci z 15 blokov a SVDA70. Číslo v prípade SVDA zodpovedá dimenzii redukovaného priestoru.

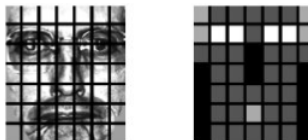
Práca [VER13] sa zaoberá rozpoznávaním piatich emócií. Tvár je v obraze nájdená pomocou Viola-Jones detektora. Príznaky sú extrahované pomocou LBP. Ako klasifikátor boli použité SVM a päť AdaBoost klasifikátorov. Každý obsahoval 100 slabých klasifikátorov. Navrhovaný systém dosiahol v prípade AdaBoost úspešnosť 86,67%. Ako aj v predošlých publikáciách, aj tu boli najlepšie klasifikované emócie „prekvapenie“ a „radosť“.

2.5 Významné oblasti tváre

Človek rozpoznáva známych a celebrity na základe významných alebo výrazných tvárových črt. V prípade, že má byť identita osoby utajená, používajú sa rôzne prostriedky ako prekrytie

alebo rozmazanie oblastí očí. V prípade existencie významných oblastí je možné ignorovať oblasti nízkej priority, a tým pádom znížiť výpočtovú a pamäťovú náročnosť.

Podľa [AHO04] je na nasledujúcom Obr. 2.1 znázornená maska s váhami, ktorú autori určili na základe percentilu. Maska je symetrická a ako uviedli autori, nebola ďalej optimalizovaná a pravdepodobne nemusí byť ani optimálna.



Obr. 2.1: Obrázok tváre rozdelený do 49 blokov. Maska váh - čierne bloky majú váhu 0, tmavo sivé 1, svetlo sivé 2 a biele váhu 4 [AHO04]

Autori Nikisins a Greitans [NIK12] testovali vplyv váh na úrovni jednotlivých príznakov a blokov. Lepšie výsledky dosiahlo pridelovanie váh jednotlivým blokom.

Biglari et al. navrhli v [BIG13], že oči, nos, ústa a obočie hrajú v procese rozpoznávania kľúčovú úlohu. Na základe experimentálnych výsledkov navrhli masku váh. Podľa ich výsledkov bola efektívnejšia, avšak podľa slov autorov taktiež určite nie je optimálna.

Wang et al. [WAN12] navrhli váhy na základe Fisherovho separačného kritéria. Stále ide o symetrickú masku váh.

Happy et al. vo svojom príspevku [HAP12] taktiež preverovali vplyv delenia obrazov do blokov. Na základe toho vyhodnocovali vplyv na výkon a presnosť rozpoznávania nimi navrhovaného biometrického systému. Uviedli, že existuje určitý počet blokov a ďalším zvyšovaním ich počtu sa už úspešnosť rozpoznávania nezvyšuje.

Podobné publikácie existujú aj pre oblasť rozpoznávania emócií. V publikácii Shan et al. [SHA05] hodnoty váh pre jednotlivé oblasti stanovili na základe empirických pozorovaní.

V publikácii [TAI09] autori Taini et al. použili na účely rozpoznávania emócií algoritmus založený na LBP príznakoch. Ich systém pracoval na princípe porovnávania dvojíc emócií.

Ďalšími príkladmi sú dva nasledujúce prístupy uvedené v [LI05] a [ZHA05]. Tie sú však založené na AdaBoost algoritme. Na nasledujúcom Obr. 2.2 sú zobrazené najvýznamnejšie oblasti tváre. Ako je možné vidieť, ide o asymetrické výbery.



Obr. 2.2: Prvé tri najdôležitejšie oblasti tváre identifikované AdaBoost algoritmom [ZHA05]

2.6 Detekcia chorobných zmien na sietnici

Publikácia [ASH14] sa venuje spôsobu detekcie poškodených oblastí na sietnici vplyvom diabetu (tzv. cukrovky). Predspracovanie pozostávalo z využitia zelenej zložky obrazu, rozdelenia obrazu na prekrývajúce sa bloky. Extrakcia príznakov pomocou LBP. Príznačky boli klasifikované SVM s RBF jadrovou funkciou. Na základe priemernej úspešnosti klasifikácie v intervale 86-85% bol definovaný vhodný typ príznakov.

3 Ciele dizertačnej práce

Motiváciou pri tvorbe tejto PhD. práce je navrhnuť a vytvoriť biometrický systém, ktorý bude schopný:

- dosiahnuť vysokú úspešnosť rozpoznávania ľudských tvárí za podmienky nízkej pamäťovej a výpočtovej náročnosti (možnosť využitia v smartfóne, tablete alebo smart televízore - pracovať v reálnom čase)
- umožňovať jednoduché pridanie nového a odobranie existujúceho používateľa

Na základe získaných poznatkov a záverov sme si stanovili nasledujúce ciele:

- Navrhnutie jednotlivých častí systému na detekciu tvárí na základe tvárových črt metódami strojového učenia. Systém bude používať vhodný spôsob:
 - detekcie/lokalizácie významných tvárových črt
 - extrakcie príznakov
 - automatizovaného výberu vzoriek pre proces trénovania
 - klasifikácie podľa typu získaných príznakov

Na základe prehľadu súčasného stavu budú v jednotlivých moduloch použité nasledujúce metódy a postupy:

- lokalizácia tvárových črt a extrakcia príznakov pomocou rôznych modifikácií LBP (lokálne binárne vzory), Gáborových filtrov a AdaBoost algoritmu
- výber vzoriek pomocou zhukovacích algoritmov ako K-means alebo SOM (samoorganizujúca sa mapa)
- klasifikácia príznakov pomocou metód strojového učenia ako neurónové siete, SVM (stroj s podpornými vektormi), metrík atď.
- Návrh rôznych úprav algoritmov alebo nastavení parametrov použitím optimalizačných algoritmov ako genetický algoritmus a simulované žihanie.
- Vyhodnotenie presnosti navrhovaného biometrického systému.
- Preveriť, či je možné navrhovaný biometrický systém využiť pri rozpoznávaní emócií alebo výrazov tváre.

4 Vlastný výskum

Na základe štúdia súčasného stavu problematiky a podľa stanovených cieľov sa budeme venovať oblastiam, ktoré sú popísané v nasledujúcich podkapitolách.

4.1 Využitie GF pri lokalizácii tvárových črt

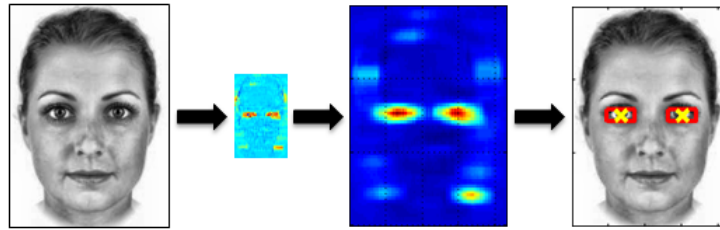
Táto kapitola sa venuje lokalizácii tvárových črt na šedotónovom obraze. V práci [5] sme využili Gáborove filtre. Navrhovaná metóda pozostáva z týchto krokov:

- Extrakcia príznakov
Príznačky sú extrahované pomocou banky dvojrozmerných Gáborových filtrov. Využíva sa len reálna zložka filtra. Nastavenie parametrov filtrov bolo zvolené tak, aby tvarom, pomerom strán a orientáciou zodpovedali tvaru hľadaných črt (oči, ústa). Banka obsahovala 4 filtre, pre každú črtu iné.
- Redukcia dimenzie
Redukciu zabezpečuje výpočet štyroch štatistických veličín. Kde $\tilde{x}_p$ predstavuje percentil ($p = 0, 5$ zodpovedá mediánu), $\bar{x}$ je stredná hodnota, σ^2 je variancia a $\mathcal{E}$ je energia.
- Klasifikácia štatistických hodnôt
Ako klasifikátor bola použitá MLP sieť s dvomi skrytými vrstvami. Konfigurácia siete 16-10-5-1. Trénovacie obrazy sa rozdelili na bloky, ktorých veľkosť bola definovaná veľkosťou hľadanej črt. 40% zo všetkých vzoriek sa použilo na validáciu siete, aby nedošlo k pretrénovaniu a bola dosiahnutá požadovaná úspešnosť klasifikácie.
- Lokalizácia črt
Rekonštrukcia pomocou veľkosti spracovávaného bloku a hľadanie oblastí s maximálnou sumou odpovedí neurónovej siete.

Zdrojom dát boli dve databázy: Face Research [FARE] a PAML [FADA]. Prvá databáza obsahuje 100 subjektov (neutrálny výraz, frontálne tváre, kontrolované pozadie) viď Obr. 4.1. Veľkosť obrazov je 400*300 pixelov. Webová stránka obsahuje možnosť vypočítať priemerné tváre (angl. *averager*), ktoré slúžili ako trénovacie obrazy. Druhá databáza obsahuje 180 subjektov (neutrálny výraz, frontálne tváre, kontrolované pozadie), ktoré boli použité ako testovacie obrazy (Obr. 4.2). Veľkosť obrazov je 450*300 pixelov. Obrazy oboch databáz majú podobné vlastnosti a kvalitu. Všetky boli normalizované na veľkosť 300*215 pixelov a ekvalizované pomocou adaptívnej metódy (CLAHE).

Metóda detekcie využíva posúvajúce sa okno (angl. *sliding window*), ktorého veľkosť závisí od hľadanej črt. Posun okna bol empiricky zvolený na 5 pixelov v horizontálnom aj vertikálnom smere.

Použitý posun spôsobil, že výstup MLP siete bol päťnásobne menší ako originálny obraz. Pre lokalizáciu bolo potrebné tento výstup zväčšiť. To je možné vyriešiť pomocou interpolácie alebo nahradením jednej hodnoty celým spracovaným blokom a pridávané hodnoty postupne sčítovať. Interpolácia sa však neosvedčila. Príklad druhého spôsobu je znázornený na Obr. 4.1.



Obr. 4.1: Originálny obraz (vľavo), výstup MLP siete, rekonštrukcia obrazu a lokalizácia črt v originálnom obraze (vpravo) [5]

V prípade očí je možné vypočítať uhol, ktorý zvierajú s horizontálnou osou. Prijatá bola len taká poloha očí, kde bol uhol v intervale $\langle -15^\circ, 15^\circ \rangle$. Inak bola nájdená poloha zamietnutá, čím bol obmedzený vplyv chybnjej detekcie.



Obr. 4.2: Testovacie obrazy z [FADA] po aplikovaní CLAHE. Príklady úspešnej lokalizácie očí a úst [5].

Maximálna úspešnosť detekcie (ACC) navrhovaného systému je uvedená v Tab. 4.1.

Tabuľka 4.1: Úspešnosť lokalizácie [5]

Črta	Správne lokaliz. (%)	Nesprávne lokaliz. (%)	Správne odmiet. zlá poloha (%)	ACC [%]
Oči	108 (60,00)	23 (12,78)	49 (27,22)	60,00
Ústa	110 (61,11)	70 (38,89)	0 (0,00)	61,11

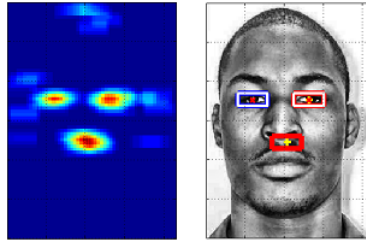
Na základe uskutočnených meraní bolo možné vyvodiť nasledujúce závery:

- Použitie štatistických hodnôt výrazným spôsobom redukuje dimenziu dát
- Trénovanie MLP siete bolo zastavené na základe chyby vypočítanej pomocou validačnej množiny vzoriek
- Ukázalo sa, že dve skryté vrstvy majú schopnosť dosiahnuť lepšie výsledky ako jedna skrytá vrstva a proces tréningu bol zároveň výrazne rýchlejší
- MLP sieť nie je citlivá na nastavenie parametrov Gáborových filtrov
- Popísaný spôsob rekonštrukcie výstupu MLP siete mal lepšie výsledky ako samotná interpolácia alebo interpolácia následne upravená pomocou prahu

Druhá práca [13] sa zaoberala porovnaním úspešnosti detekcie tvárových črt pomocou klasických, logaritmických Gáborových filtrov a ich kombináciou. Na rozdiel od predošlej práce boli namiesto mediánu použité 25% a 75% percentily. Hodnoty štatistických veličín boli vypočítané z magnítudy odpovedí komplexných filtrov.

Pretože testovacia databáza obsahovala subjekty ázijského a afroamerického vzhľadu, boli pridané zodpovedajúce tréningové subjekty. Ďalej boli pridané aj subjekty staršieho veku, nakoľko sú tiež početne zastúpené v testovacej množine.

V tomto prípade bolo použitých 6 filtrov a 5 štatistík. Vstupný vektor do MLP siete obsahoval vektor 30 hodnôt. Podrobnou analýzou výsledkov predošlej práce sme zistili, že vo veľkom počte prípadov boli v rekonštruovanom obraze tri oblasti s maximálnou sumou odpovedí neurónovej siete. Tento systém bral do úvahy tri oblasti a počítal tri uhly, ktoré navzájom zvierali. Ten uhol, ktorý spĺňal interval natočenia bol považovaný za správne riešenie. Ak ani jeden uhol nespĺňal podmienku $\langle -15^\circ, 15^\circ \rangle$, nájdené polohy boli zamietnuté.



Obr. 4.3: Nájdenie správnej polohy očí z troch kandidátov

Maximálna úspešnosť navrhovaného systému je uvedená v Tab. 4.2.

Tabuľka 4.2: Úspešnosť lokalizácie [13]

Črta	Banka filtrov	Správne lokaliz. (%)	Nesprávne lokaliz. (%)	Správne odmiet. zlá poloha (%)	ACC [%]
Oči	GF	74,81	9,26	15,93	74,81
	LGF	51,85	23,33	24,82	51,85
	Oba typy	78,52	9,81	11,67	78,52
Ústa	GF	57,41	42,59	0,00	57,41
	LGF	68,15	31,85	0,00	68,15
	Oba typy	71,30	28,70	0,00	71,30

Prínosom danej metódy je vyššia úspešnosť a menší počet odmietnutých prípadov. To bolo dosiahnuté väčším počtom filtrov, kde kombinácia filtrov umožňuje vytvoriť vektor príznačiek dĺžky 60 hodnôt. Pomer odmietnutých vzoriek je znížený pomocou vyhodnocovania troch uhlov.

4.2 Optimalizácia banky GF pre účely lokalizácie tvárových črt

V práci [8] podobne ako v predošlých publikáciách [5, 13] sme využili banku GF pre lokalizáciu tvárových črt. Tu sme sa však venovali aj návrhu optimalizovanej banky Gáborových filtrov. Pri hľadaní optimálnej banky filtrov sme vytvorili pomerne veľkú banku obsahujúcu 72 filtrov. Filtre mali 8 rôznych natočení a 9 rôznych hodnôt pre šírku pásma bw . Podobne ako v spomínaných publikáciách bola použitá metóda plávajúceho okna. Pre každý blok sa vypočítal rozptyl, energia, medián, horný a dolný kvantil. Vektor 360 hodnôt slúžil ako vstup do MLP siete. Na základe pozorovania váh neurónovej siete sme prišli k záveru, že významné informácie sú práve na váhach zodpovedajúcich rozptylu a energii. Hľadali sme banky pozostávajúce z 8 filtrov. Vzhľadom na túto skutočnosť je možné vstupný vektor redukovať len na 16 hodnôt. Zlepšenie úspešnosti detekcie bolo o 3-9%, pretože došlo k zníženiu počtu

zle lokalizovaných črt. Daný rozdiel bol zistený pri porovnaní s inou bankou publikovanou v [LI10].

Táto práca [4] je pokračovaním predošlej [8]. Hlbšia analýza spočívala v sledovaní odpovedí jednotlivých filtrov s rôznym nastavením vzhľadom na hľadané tvárové črty. Pre detailnú analýzu boli použité rovnaké ukazovatele (energia ε a rozptyl σ).

Zaviedli sme nový parameter $G_i(bw, \theta)$ (informatívny parameter potencionalneho rastu), ktorý zodpovedá rozdielu medzi $L_2(bw, \theta)$ a minimálnou $L_{2(bw, n)}(bw, \theta)$ (minimálna hodnota $L_2(bw, \theta)$ vzhľadom na konštantnú hodnotu bw). Predpokladáme, že vyššia hodnota tohto parametra $G_i(bw, \theta)$ poskytuje lepšiu informáciu ohľadom filtrovanej oblasti vzhľadom na iné filtre. Na základe tohto parametra je možné vidieť, že niektoré filtre majú výrazne vyššiu odpoveď na danú črtu tváre. Podobným spôsobom boli vybrané filtre pre ústa a nos.

Na vyhodnotenie úspešnosti bola podobne ako v predošlom prípade použitá neurónová sieť. 100 vzoriek z PAML databázy tvorilo testovaciu množinu. Príklad úspešnej lokalizácie je na Obr. 4.4.



Obr. 4.4: Príklad správnej detekcie tvárových črt

Tabuľka 4.3: Porovnanie úspešnosti lokalizácie navrhovaných bánk Gáborových filtrov obsahujúce 2, 4, 6, 8 a 16 filtrov

Tvárová črta	Počet filtrov v banke				
	16 Správne lok.	8 Správne lok.	6 Správne lok.	4 Správne lok.	2 Správne lok.
Oči	86	86	89	83	20
Ústa	89	89	85	92	41
Nos	92	95	89	90	21

Na základe tejto publikácie sme dospeli k nasledujúcim záverom:

- špecifické textúry (črty tváre) zodpovedajú špecifickému nastaveniu parametrov GF (v tomto prípade bw a θ)
- obe veličiny (rozptyl a energia) sú vhodné na lokalizáciu tvárových črt, pretože majú dostatočnú výpočtovú hodnotu
- znižovanie počtu filtrov v banke nemá zásadný dopad na úspešnosť navrhovaného systému
- nižší počet filtrov ako 6 už nie je vhodný pre úspešnú lokalizáciu
- 16 filtrov má porovnateľné výsledky ako 8 filtrov, ale predstavuje vyššiu výpočtovú náročnosť

4.3 Automatický výber vzoriek a spôsoby redukcie dimenzie dát

Táto kapitola je venovaná návrhu systému na inteligentný prístup a ovládanie smart TV v rámci 7. rámcového projektu HBB-NEXT. Jedna z častí smart TV obsahuje systém na detekciu, lokalizáciu, sledovanie tvárí a rozpoznávanie užívateľov smart TV v reálnom čase. Vstupným sensorom je kamera Kinect, ktorá dokáže zaznamenať 25 až 30 snímok za sekundu. Bolo potrebné navrhnuť spôsob výberu vzoriek z nazhromaždenej množiny a stanoviť ich počet.

Odpovede na túto otázku obsahuje práca [6]. Zdrojom obrazov sa stala CMU PIE (Pose, Illumination and Expression) databáza [PIED]. Bola vytvorená na Carnegie Mellon University. Vzhľadom na rozličné natočenia a svetelné podmienky, ktoré databáza poskytuje sme sa rozhodli použiť práve tie, ktoré najviac zodpovedajú predpokladanému správaniu používateľa smart televízora. Konkrétne tieto tri pózy: *C05* (24 obr./subj.), *C27* (49 obr./subj.) a *C29* (24 obr./subj.). Príklady sú uvedené na nasledujúcom Obr. 4.5. Databáza obsahuje 68 subjektov. Celkový počet obrazov, ktorý sme použili bol 6596. Veľkosť predspracovaných a orezaných vzoriek je 64*64 pixelov.



Obr. 4.5: CMU PIE: Póza C05, C27 a C29 [6]

Na základe stanoveného postupu a metód ako K-means a SVM sme použili 6 vzoriek na subjekt. Táto hodnota je vhodná aj pri väčšom počte subjektov. Tiež je dostatočná pre zastúpenie rôznych póz, ktoré môžu byť zosnímané v procese registrácie subjektu. Posledným dôvodom pri výbere počtu vzoriek bola topológia dvojrozmernej SOM siete. Párny počet neurónov je vhodnejší ako nepárny v prípade, ak chceme dosiahnuť, aby topológia siete mala pravouhlý (približne štvorcový) tvar.

Odpoveď na prvú časť problému sme zistili pomocou rôzneho predspracovania a následného výberu tréningových vzoriek.

- Predspracovanie
žiadne (originálne obrazy), globálna ekvalizácia histogramu, adaptívna ekvalizácia histogramu (CLAHE)
- Výber 6 vzoriek pomocou rôznych algoritmov
 - K-means bol nastavený pomocou vstupného parametra $k = 6$. Použitá bola metrika L_1 a 10 opakovaní pre výpočet zhľukov.
 - Topológia dvojrozmernej SOM siete bola 2*3, čo zodpovedá 6 neurónom.
 - DBSCAN však nie je možné takto nastaviť. Hodnoty parametrov $MinPts = 3$ a $\epsilon_{opt} = 0,65\epsilon$ vypočítané podľa definovaného vzťahu približne v priemere zodpovedalo 6 vzorkám na subjekt.
- Výpočet stredov zhľukov
Tento krok je potrebný len v prípade DBSCAN-u, pretože jeho výstupom sú identifikované zhľuky. Stred zhľuku bol vypočítaný ako priemer všetkých *základných* (angl. core) bodov.
- Rekonštrukcia (výsledný výber)
Každý stred zhľuku sa porovnával s použitou podmnožinou CMU PIE databázy. Na

základe najmenšej hodnoty Euklidovskej vzdialenosti (L_2) boli identifikované výsledné trénovacie vzorky.

Tabuľka 4.4: Výsledky výberu trénovacích vzoriek tromi algoritmi (priemerný počet obrazov) [6]

Druh použitej ekvaliz.	Zhlukovací algoritmus									
	K-means			SOM			DBSCAN			
	C05	C27	C29	C05	C27	C29	C05	C27	C29	Priemer
Bez ekval.	0,588	4,941	0,471	2,603	3,074	0,323	1,397	3,427	1,029	5,853
Globálna	1,382	3,529	1,089	2,265	2,632	1,103	1,412	3,000	1,735	6,147
Adaptívna	1,500	3,206	1,294	2,147	2,000	1,853	1,250	3,015	1,617	5,882

Na základe analýzy výstupov meraní je možné urobiť nasledujúce závery:

- 6 trénovacích vzoriek na subjekt je dostatočne veľa na pokrytie rôznych póz pri snímaní
- ekvalizácia histogramu zabezpečila, že zhlukovacie algoritmy sa nezamerali na výber vzoriek podľa rôznej úrovne osvetlenia
- SOM a K-means sú vhodné metódy na výber vzoriek, kde je ich počet vopred známy
- Nastavenie hodnoty parametra ϵ pre DBSCAN algoritmus môže byť niekedy dosť komplikované
- DBSCAN je menej citlivý na zmenu osvetlenia pri výbere vzoriek
- Vybraný priemerný počet vzoriek frontálnej pózy (C27 - obsahuje dvojnásobne vyšší počet vzoriek v porovnaní s ostatnými pózami) je vyšší ako priemerný počet zvyšných póz, čo demonštruje Obr. 4.6.



Obr. 4.6: Obrazy vybrané K-means algoritmom po aplikovaní: adaptívnej (prvý riadok), globálnej ekvalizácii histogramu (druhý riadok) a bez ekvalizácie (tretí riadok) [6]

Práca [7] je pokračovaním a doplnením predošlej [6], pretože sa zaoberá výberom vzoriek a následným spracovaním príznakov. Metóda pozostávala z týchto krokov:

- Predspracovanie
žiadne (originálne obrazy), globálna ekvalizácia histogramu, adaptívna ekvalizácia histogramu (CLAHE)

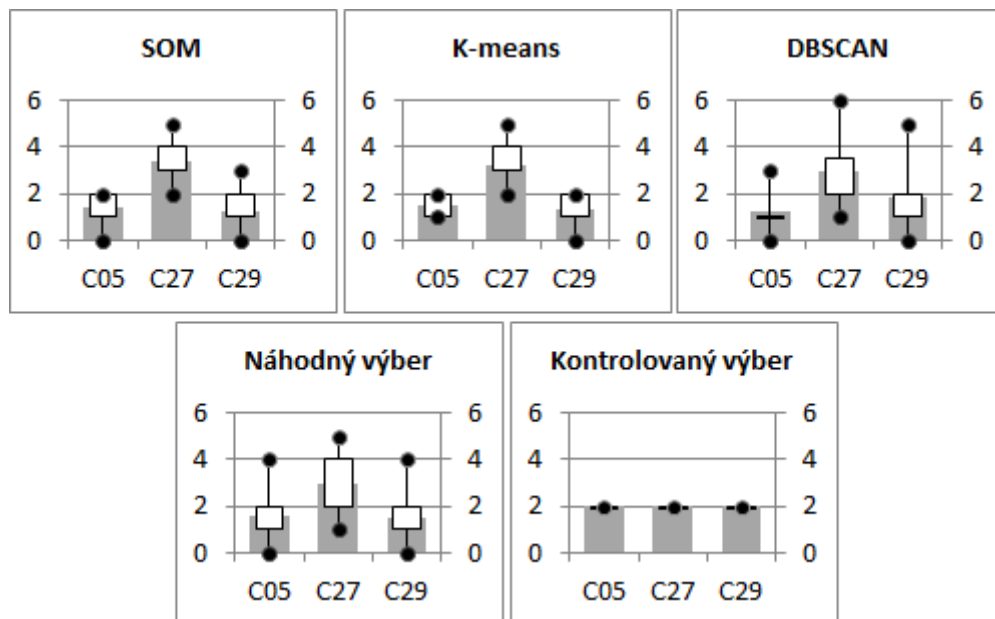
- Metódy extrakcie príznakov a redukcie dimenzie
Porovnanie štatistických metód ako PCA a LDA. Metóda na spracovanie textúr $LBP_{8,2}^{u2}$. Kombinácia predošlých $LBP_{8,2}^{u2}$ a PCA.
- Výber vzoriek pomocou zhlukovacích algoritmov
Využitie algoritmov SOM, K-means a DBSCAN. V prípade DBSCAN-u bolo potrebné na základe vstupných hodnôt hľadať vhodnú hodnotu parametra ϵ .

Pridanie *náhodného* (6 vzoriek je náhodne vybratých zo všetkých 97 vzoriek pre daný subjekt) a *čiastočne kontrolovaného* (vzhľadom na tri rôzne pózy sa náhodne vyberú po dve vzorky) výberu.

- Vyhodnotenie výsledkov
Použitie nelineárneho SVM klasifikátora.

Pretože histogram vytvorený z LBP príznakov sa skladá z histogramov jednotlivých blokov zvolili sme 354 príznakov. Výsledný histogram vznikol zo 6 riadkov po 59 binov. Rovnaký počet hlavných komponentov bol použitý z PCA metódy. Výstupom LDA metódy bolo 67 príznakov. V prípade kombinácie (LBP a PCA) sme pri LBP využili 1180 hodnôt. Histogram bol v tomto prípade tvorený z 20 riadkov po 59 binov. Následnou aplikáciou PCA sme tiež dostali spomínaných 354 príznakov.

Podobne ako v práci [6], aj tu sa potvrdil fakt, že zhlukovacie algoritmy vyberajú častejšie frontálne obrazy ako tie s rôznym natočením. Demonstrujú to grafy na Obr. 4.7. Pre porovnanie úspešnosti klasifikácie jednotlivých spôsobov výberu tréningových vzoriek sme pridali náhodný výber, ktorý by mal zodpovedať najhoršiemu scenáru a čiastočne kontrolovaný výber, ktorý by mal hypoteticky simulovať najlepší scenár.



Obr. 4.7: Počet vybraných vzoriek pomocou rôznych algoritmov po aplikovaní CLAHE. Sivé stĺpce zodpovedajú priemeru, biele stĺpce zodpovedajú 25% a 75% percentilu a čierne kolieska zodpovedajú minimu a maximu [7].

Pre účely klasifikácie bolo použitých 408 tréningových vzoriek (6*68 subjektov) a 6180 testovacích vzoriek (zvyšné nepoužité vzorky). Určenie výslednej úspešnosti (Tab. 4.5) vyžadovalo nájsť vhodné nastavenie SVM parametrov pre každý typ vstupných dát. Zvýraznené hodnoty predstavujú maximálnu úspešnosť klasifikácie vzhľadom na spôsob výberu tréningových vzoriek.

Tabuľka 4.5: Úspešnosť klasifikácie vzhľadom na rôzne spôsoby predspracovania, extrakcie príznakov a výberu tréningových vzoriek [7]

Extrakcia Redukcia	Ekvaliz. histog.	Výber				
		SOM	K-means	DBSCAN	Náhod.	Kontrol.
Žiadna	CLAHE	92,7117	93,7621	87,9948	81,4480	86,6516
	Hist.	90,5139	90,9179	91,7193	81,3025	86,0698
	Žiadne	78,4745	79,9127	81,8049	69,6671	73,4971
PCA	CLAHE	92,8410	91,6936	86,4330	81,3995	86,5869
	Hist.	90,2069	91,0149	91,3732	81,2379	86,0052
	Žiadne	77,8927	80,1551	79,1216	69,3277	73,2870
LDA	CLAHE	77,3917	77,0200	73,6749	61,4738	68,6328
	Hist.	72,3497	76,2120	71,0543	61,8778	69,1661
	Žiadne	68,9722	68,5036	65,9877	54,9935	60,0194
LBP	CLAHE	88,5603	88,3323	83,0454	75,1616	80,9955
	Hist.	89,4796	89,8352	83,3818	77,8442	81,1732
	Žiadne	89,7867	89,8998	84,4786	77,6341	81,1732
LBP+PCA	CLAHE	93,6985	93,3258	88,4227	84,6962	90,6593
	Hist.	93,6490	93,5036	87,7564	85,2295	89,2372
	Žiadne	94,4732	93,1965	87,6737	85,3103	89,3665

Na základe dosiahnutých výsledkov v Tab. 4.5 je možné vyvodiť nasledujúce závery:

- Predspracovanie zlepšuje činnosť algoritmov PCA a LDA.
- Metóda LBP je invariantná na zmenu osvetlenia obrazov. Predspracovanie pomocou CLAHE metódy na blokoch 8*8 vrátane použitia bilineárnej interpolácie medzi blokmi má dokonca negatívny vplyv na klasifikáciu.
- Hypotéza úspešnej klasifikácie pri rovnakom počte vzoriek (viď. náhodný kontrolovaný výber) pre každú pózu sa nepotvrdila. Z toho vyplýva, že zmena osvetlenia, prekrytie tváre (okuliare), prípadne zmena výrazu tváre môžu mať z hľadiska extrakcie príznakov väčšiu váhu pri výbere vzoriek ako samotné natočenie tváre.
- DBSCAN nie je vhodným algoritmom pre výber vzoriek (obťažne nastavenie parametra ϵ).
- V prípade pridania, resp. odstránenia používateľa musia mať metódy PCA a LDA uložené tréningové vzorky ostatných používateľov a projekcie musia byť znovu prepočítané.
- Metóda PCA s 354 príznakmi dosiahla približne rovnakú úspešnosť ako použitie originálnych obrazov o veľkosti 4096 pixelov.
- Metóda LBP dosiahla o niečo nižšiu úspešnosť ako PCA, ale je značne výpočtovo jednoduchšia. V prípade pridania, resp. odstránenia používateľa nie je potrebné historizovať pôvodné snímky všetkých subjektov.
- Kombinácia metód LBP a PCA s použitím SOM siete alebo K-means algoritmu dosiahla najlepšie výsledky.
- Kombinácia metód LBP a PCA umožňuje relatívne jednoduché pridanie, resp. odstránenie používateľa. Pri prepočítaní PCA projekcie (napr. pridanie používateľa) je potrebné uchovať 1180 LBP príznakov pre každý subjekt. Ide tu však o značnú úsporu pamäte voči veľkosti originálnych obrazov.

Motiváciou pre automatizovaný výber vzoriek je fakt, že v publikáciach možno nájsť len veľmi málo zmienok o výbere tréningových vzoriek. Väčšinou je výber náhodný [SUR12], [LIA07], použije sa empirický počet [SHA06],[TAN10], polovica vzoriek z databázy [LIA07] alebo sa vykoná vzájomná validácia so zmenou tréningových a testovacích vzoriek [ZAV11], [SUR12]. Ako je však vidieť v Tab. 4.5 náhodný výber vzoriek určite nie je optimálnym riešením.

4.4 Optimalizácia LBP parametrov

V práci [9] sme sa venovali optimálnemu nastaveniu rôznych parametrov týkajúcich sa LBP príznakov v rámci klasifikácie ľudských tvárí. Optimalizácia pomocou genetického algoritmu zahŕňala: typ príznakov (LBP, LGP, NRLBP), druh mapovania (žiadne, U2, RI, RIU2), typ metriky ($L1$, $L2$, χ^2 , EMD), redukciu dimenzie vektoru príznakov (optimálne delenie obrazu príznakov do blokov) a maximalizáciu úspešnosti klasifikácie.

Na testovacie účely a porovnanie výsledkov boli použité tri rôzne databázy:

- CMU PIE [PIED]
Databáza už bola popísaná v kapitole 4.3. Bol použitý rovnaký počet subjektov, a aj počet vzoriek na subjekt.
- Extended YALE Database B
Databáza bola vytvorená na University of California, San Diego. Obsahuje 38 subjektov. Subjekty sú nasnímané v rôznych svetelných podmienkach (zdroj svetla pod rôznymi uhlami). Každý subjekt má 68 obrazov. Databázu tvorí 2584 obrazov. Veľkosť obrazov je 192×168 pixelov.



Obr. 4.8: Príklady z Ext. YALE (tri rôzne svetelné podmienky)

- AT&T (ATT/ORL) [ATTD]
Databáza vytvorená na Cambridge University Computer Laboratory. Obsahuje 40 subjektov. Subjekty sú nasnímané v rôznych pózach. Každý subjekt má 10 obrazov. Databázu tvorí 400 obrazov. Veľkosť obrazov je 112×92 pixelov.

Tréningové vzorky boli identifikované na základe výsledkov zhukovacieho algoritmu K-means. Bolo vybraných 6 tréningových vzoriek na subjekt. Vzorky boli vyberané z obrazov spracovaných pomocou $LBP_{8,2}^{u2}$. Každý stred nájdený algoritmom K-means bol nahradený obrazom príznakov s najmenšou hodnotou L_2 . Na základe nájdeného obrazu príznakov bol definovaný originálny obraz z databázy.

Ako nástroj optimalizácie slúžil genetický algoritmus. Nastavenie parametrov bolo nasledovné. Dĺžka chromozómu $c_l = 5$ (typ LBP, typ mapovania, rozmery bloku, typ metriky), počet jedincov v generácii $p_g = 20$ (menší počet jedincov spôsobuje rýchlejší vývoj populácie), počet generácií $g_n = 40$, pravdepodobnosť rekombinácie $p_r = 0,5$ a pravdepodobnosť mutácie $p_m = 0,8$ (na úrovni chromozómu 0,04). Kriteiálna funkcia mala tvar:

$$f(d, x, y, acc) = 0,25 \frac{d}{xy} + (1 - acc) \quad (4.1)$$

kde d je dĺžka vektoru príznakov, x, y sú rozmery originálneho obrazu, acc zodpovedá hodnote ACC a konštanta 0,25 definuje váhu pre dôležitosť redukcie dimenzie dát. V prípade, že $d > xy/3$ (veľkosť vektora príznakov je väčšia ako 1/3 obrazu) bude daný chromozóm ignorovaný.

Najlepšie výsledky sú pre jednotlivé databázy uvedené v nasledujúcej tabuľke.

Tabuľka 4.6: Optimalizácie vykonané na troch databázach

Data-báza	Veľkosť obrazu	Typ príznakov	Veľkosť blokov	Počet blokov	Počet príznakov	Metrika	ACC (%)
CMU PIE	48*48	LBP RIU2	1*44	44	440	L1	85,326
Ext. YALE	48*42	LBP RIU2	1*38	44	440	χ^2	94,056
ATT	56*46	LBP U2	26*42	2	118	χ^2	100,000
	84*69	NRLBP RIU2	20*13	20	180	L1, χ^2	100,000
	112*92	NRLBP RIU2	12*88	9	81	χ^2	100,000

Na základe predošlej tabuľky sme dospeli k nasledujúcim záverom:

- CMU PIE
Najlepší výsledok klasifikácie 85,326% bol dosiahnutý na obrazoch veľkosti 48*48 pixelov. Odporúčané delenie obrazu príznakov na bloky je 1*44 pixelov.
- Ext. YALE
Úspešnosť rozpoznávania bola 94,056%. V tomto prípade sa výsledok podobá databáze CMU PIE, pretože veľkosť obrazu bola 48*42 pixelov a odporúčaná veľkosť bloku je 1*38 pixelov.
- ATT
Bezchybná klasifikácia subjektov bola dosiahnutá trikrát. Vzhľadom na výpočtovú náročnosť je možné použiť obrazy najmenšej veľkosti (56*46 pixelov).
- Odporúčaný typ príznakov je LBP s mapovaním RIU2 a U2
- Odporúčanou metrikou je χ^2
- Veľkosť vstupného obrazu by mala byť aspoň približne 50*50 pixelov
- Metrika EMD nie je vhodná pre daný typ klasifikácie
- LBP príznaky nemajú dostatočnú diskriminačnú schopnosť pre daný problém

Nasledujúca publikácia [15] obsahuje rozšírené porovnanie viacerých modifikácií LBP, kde sme sa podobne ako v predošlej práci [9] zamerali na optimalizáciu podobných parametrov. Rozdiel spočíva v pridaní nasledujúcich doplnení:

- pridaný ďalší typ príznakov - LTP (angl. Local ternary patterns)
- vytvorenie tzv. *pyramídy obrazov* rôznych škál, z ktorých boli extrahované príznaky
- preskúmanie vplyvu rôznej veľkosti polomeru R

Na detailnú analýzu viacerých modifikácií LBP sme použili štyri rôzne databázy tvári. Aby bolo možné dosiahnuté výsledky porovnať, zjednotili sme veľkosti obrazov všetkých databáz. Boli použité všetky nasledujúce databázy: CMU PIE, Ext. YALE, ATT a FERET (podmnožiny pomenované „BIG4“ a „BIG6“).

Na testovacie účely bola vytvorená databáza *Big Faces*, ktorá je podmnožinou databázy FERET. Databáza FERET bola vytvorená na univerzite George Mason a v spolupráci s US Army Research Laboratory. Big Faces tvoria iba frontálne šedotónové obrazy z množín „fa“ a

„fb“. Obrazy boli zarovnané vzhľadom na koordináty očí, upravený histogram a nepodstatné časti obrazu boli zamaskovné. Vytvorili sme dve množiny „BIG4“ a „BIG6“. Rozdiel v názve uvádza počet vzoriek na subjekt. BIG4 obsahuje aspoň 4 snímky a teda obsahuje 246 subjektov a 1223 obrazov. BIG6 obsahuje aspoň 6 snímok a tvorí ju 73 subjektov a 531 obrazov. Veľkosť obrazu je 154×134 pixelov.



Obr. 4.9: Pôvodný obraz z databázy FERET a obraz z databázy Big Face

Aj počet trénovacích vzoriek bol zvolený tak, aby bolo možné dané výsledky porovnať. V predošlej publikácii to bolo práve 6 vzoriek. Ich počet sa teda znížil na 5 okrem prípadu BIG4, kde to boli len 3 vzorky. Taktiež spôsob výberu trénovacích vzoriek bol pozmenený podľa vykonaných experimentov. Predtým to bolo na základe obrazov príznakov $LBP_{8,2}^{u2}$ pre všetky databázy.

1. Ext. YALE, ATT, FERET (Big Faces)

Najlepšie výsledky klasifikácie boli dosiahnuté v prípade, že trénovacie vzorky boli vybrané z originálnych obrazov, na ktorých sme zvýšili kontrast (adaptívnou histogramovou ekvalizáciou - CLAHE). Pomocou zhlučovacieho algoritmu K-means a najmenšej hodnoty miery L_2 vzhľadom k upraveným obrazom boli identifikované trénovacie vzorky. Daný prístup zdôvodňuje fakt, že databázy obsahujú len frontálne vzorky alebo malý počet natočení. Rôzne svetelné podmienky je schopná korigovať úprava histogramu.

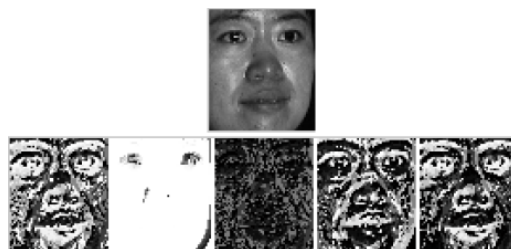
2. CMU PIE

Predošlý prístup nebol vhodný pre túto databázu. Vysvetlením je zrejme fakt, že subjekt obsahuje okrem frontálnej pózy aj dve ďalšie natočenia (C05 a C29) vo väčšom počte vzoriek a tiež subjektov.

Lepšie výsledky klasifikácie sme dosiahli, keď sme ako trénovacie vzorky použili tie, ktoré sme získali na základe výberu z obrazov príznakov. Pre každú modifikáciu LBP a spôsob mapovania príznakov bola vytvorená množina trénovacích vzoriek. Daných 16 druhov obrazov príznakov separátne tvorilo vstupy do zhlučovacieho algoritmu. Na základe najmenšej hodnoty L_2 boli vzhľadom na obrazy príznakov identifikované trénovacie vzorky v pôvodnej databáze. V procese klasifikácie a optimalizácie boli použité príslušné trénovacie vzorky.

Takto zvolené trénovacie vzorky spôsobili nárast úspešnosti klasifikácie v intervale 5-10%. Obrazy príznakov sú lepšie schopné zachytiť rôzne typy natočenia.

Na Obr. 4.10 sú pre porovnanie zobrazené obrazy príznakov všetkých modifikácií LBP.



Obr. 4.10: Originálny obrázok z CMU PIE (prvý riadok). Obrazy príznakov: LBP, LGP, NRLBP, negatívna a pozitívna zložka LTP.

Bola vytvorená pyramída normalizovaných obrazov v troch rôznych mierkach. Výhodou daného prístupu je možnosť využiť príznaky mikro a makro štruktúr tváre daných subjektov. Nevýhodou je však vyššia výpočtová náročnosť.

Nastavenie parametrov genetického algoritmu: dĺžka chromozómu $c_l = 5$, počet jedincov v populácii $p_g = 30$, počet generácií $g_n = 100$, pravdepodobnosť rekombinácie $p_r = 0,5$ a pravdepodobnosť mutácie $p_m = 0,5$ (na úrovni chromozómu 0,017). Kriteálna funkcia bola:

$$f(d, x, y, acc) = 0,01 \frac{d}{xy} + (1 - acc) \quad (4.2)$$

kde d je dĺžka vektoru príznakov, x, y sú rozmery originálneho obrazu, acc zodpovedá hodnote ACC a konštanta 0,01 definuje váhu pre dôležitosť redukcie dimenzie dát. V prípade, že $d > xy/3$ (veľkosť vektora príznakov je väčšia ako 1/3 obrazu) bude daný chromozóm ignorovaný.

V nasledujúcich Tab. 4.7 a Tab. 4.8 sú uvedené najlepšie priemerné výsledky pre jednotlivé databázy. Priemerné výsledky sú vypočítané na základe šestnásobnej vzájomnej validácie.

Tabuľka 4.7: Optimálne nastavenie parametrov (škálované obrazy)

Použitá databáza	Veľkosť obrazu	R	Typ príznakov	Veľkosť bloku	Počet príznakov	Metrika	ACC [%]
CMU PIE	32*28 48*42 64*56	3	LBP RIU2	26*22 1*36 2*25	1010	χ^2	87,675
Ext. YALE	32*28 48*42 64*56	3	LBP RIU2	26*22 1*36 2*50	720	χ^2	93,949
ATT	32*28 48*42 64*56	3	LTP U2	26*22 42*36 29*50	236	χ^2	99,500
BIG4	32*28 48*42 64*56	2	LTP RIU2	28*6 4*38 10*4	930	χ^2	98,660
BIG6	32*28 48*42 64*56	3	LTP RIU2	2*22 7*6 2*50	1130	L1	99,498

Tabuľka 4.8: Optimálne nastavenie parametrov

Použitá databáza	Veľkosť obrazu	R	Typ príznakov	Veľkosť bloku	Počet príznakov	Metrika	ACC [%]
CMU PIE	64*56	3	LBP RIU2	1*25	1160	χ^2	87,274
Ext.YALE	64*56	2	LBP RIU2	2*13	1200	L1	95,079
ATT	48*42	2	LTP U2	11*38	236	L1	99,500
	64*56		LBP U2	15*52	236	L1	99,500
BIG4	64*56	2	NRLBP RIU2	6*4	1170	L1	98,316
BIG6	64*56	1	LTP RIU2	2*18	930	L1	99,096

Na základe všetkých Tab. 4.7 a Tab. 4.8 sme dospeli k nasledujúcim záverom:

- najlepšie výsledky dosiahli LBP príznaky, ktoré sú bežne používané v literatúre a LTP

príznamy, ktoré majú schopnosť redukovať šum (záporom operátora je korelácia dát medzi pozitívnou a negatívnou zložkou)

- aj parameter R má určitý vplyv na výslednú úroveň klasifikácie; optimálna je bežne používaná hodnota $R = 2$, ale aj $R = 3$ môže byť využité pri klasifikácii
- typ mapovania RIU2 príznamov poskytuje významnú redukciu dimenzie a invariantnosť voči natočeniu; taktiež mapovanie U2 môže byť využité pri klasifikácii
- ako vhodné metriky sa osvedčili najmä χ^2 a L_1 , ktoré sú výpočtovo nenáročné a dosahujú dobrú úroveň klasifikácie
- použitie škálovaných obrazov v danej podobe v minimálnej miere ovplyvnilo úspešnosť klasifikácie (PIE a BIG6 = +0,40%, BIG4 = +0,34%, ATT = 0,00 a Ext. YALE = -1,13% - čo negatívne ovplyvnila výrazná redukcia dimenzie)
- databáze s malým počtom subjektov a obrazov akou je ATT postačuje malý vektor príznamov (236), zatiaľ čo väčšie databázy potrebujú aspoň 930 príznamov na dosiahnutie požadovanej úspešnosti (zapríčiňuje to počet subjektov, počet obrazov, rôzne pózy - čo má v konečnom dôsledku vplyv aj na tvar a veľkosť blokov
- tvar a veľkosť blokov sa mení - *tenké riadky* (PIE - 1*25, ATT - 11*38 pixelov), *malé obdĺžnikové oblasti* (Ext. YALE - 2*13, BIG6 - 2*18 pixelov) a *takmer štvorcové oblasti* (BIG4 - 6*4 pixelov)
- tvar a veľkosť blokov sa mení aj pri škálovaných obrazoch - *tenké riadky* (PIE - 1*36, Ext. YALE - 1*36, 2*50, BIG4 - 4*38, BIG6 - 2*22, 2*50 pixelov), *malé obdĺžnikové oblasti* (PIE - 2*25, BIG4 - 28*6 pixelov), *takmer štvorcové oblasti* (BIG4 - 10*4, BIG6 - 7*6 pixelov) a *celé obrázky* (PIE, Ext. YALE, ATT - 26*22 pixelov)
- v predošlej práci [9] (Tab. 4.6) sme dosiahli porovnateľnú úspešnosť klasifikácie, aj napriek nižšiemu počtu tréningových vzoriek (5 voči 6 vzorkám na subjekt)
- v článku [15] sú uvedené výsledky získané pomocou genetického algoritmu; taktiež sme využili simulované žihanie avšak bez významnej zmeny vo výsledkoch
- najlepšie výsledky dosiahli nasledujúce typy príznamov: LBP RIU2, LTP RIU2 a LTP U2
- riadky v Tab. 4.7 ukazujú, že pyramída z obrazov rôznych mierok dosahuje dobrú úroveň rozpoznávania, má vplyv na počet príznamov a nastavenie parametrov: mapovanie RIU2, $P = 8$, $R = 3$, χ^2 metrika
- riadky v Tab. 4.8 ukazujú, že najlepšia veľkosť obrazu pri použití jednej mierky je približne 64*56 pixelov; menšie rozmery vstupného obrazu už nie sú vhodné; nastavenie parametrov: mapovanie RIU2, $P = 8$, $R = 2$, L_1 metrika
- metriky EMD, L_2 sa vyskytli v procese optimalizácie v minimálnej miere
- LBP príznamy opakovane nemajú dostatočnú diskriminačnú schopnosť

4.5 Dôležitosť jednotlivých oblastí tváre

Na základe optimalizácie parametrov podľa rovnakej metodiky ako v [15] sme identifikovali úspešnosť klasifikácie emócií použitej JAFFE databázy.

Nami navrhovaná metóda detekcie významných oblastí tváre využívala nasledujúce:

- rozmery obrazov 64*56 pixelov (podobne ako v našich predošlých publikáciách)
- výber trénovacích vzoriek pomocou K-means zhlukovacieho algoritmu
- BIG4 - 3 trénovacie vzorky na subjekt
- BIG6 - 5 trénovacích vzoriek na subjekt
- JAFFE - 12 trénovacích vzoriek na emóciu (1,2 vzorky na subjekt)
- typ príznakov (BIG4 - NRLBP RIU2, BIG6 - LBP U2, JAFFE - NRLBP U2), $R = 2$, $P = 8$ (viď Tab. 4.8)
- rozmery blokov (viď Tab. 4.8)
- klasifikácia subjektov pomocou váženej χ^2 metriky

$$\chi^2(\mathbf{x}, \mathbf{y}, \boldsymbol{\omega}) = \sum_{i,j} \omega_j \frac{(\mathbf{x}_{i,j} - \mathbf{y}_{i,j})^2}{(\mathbf{x}_{i,j} + \mathbf{y}_{i,j})} \quad (4.3)$$

kde $\mathbf{x}, \mathbf{y}$ sú riadkové vektory, i zodpovedá i-temu binu v histograme a $\boldsymbol{\omega}_j$ je vektor váhových koeficientov oblasti j

- identifikácia významných oblastí tváre pomocou simulovaného žihania (váhy: 1 - dôležitá oblasť, 0 - ignorovaná oblasť)
- definícia kritériálnej funkcie - maximalizácia úspešnosti klasifikácie subjektov na základe definovanej metriky (4.3)
- vykonaná 6-násobná vzájomná validácia

V rámci hľadania dôležitých oblastí tváre sme fixne definovali maximálny počet použitých blokov. Hodnoty boli pre jednotlivé databázy zvolené tak, aby mal vektor príznakov maximálne polovičný počet príznakov v porovnaní bez použitia masky váh (Tab. 4.9).

Tabuľka 4.9: Výsledky dosiahnuté pomocou simulovaného žihania [10]

Databáza	BIG4	BIG6	JAFFE
Typ príznakov	<i>NRLBP_{8,2} RIU2</i>	<i>LBP_{8,2} U2</i>	<i>NRLBP_{8,2} U2</i>
Veľkosť blokov (pixely)	6*4	5*52	3*52
Počet príznakov bez váh	1170	708	600
Počet príznakov s váhou 1	450	345	300
ACC bez váh (%)	97,801	99,096	97,933
ACC s váhou 1 (%)	97,457	98,996	97,674

Ako je možné vidieť na Obr. 4.11 metóda je vhodná pre klasifikáciu subjektov a aj rozpoznávanie emócií.



Obr. 4.11: Príklad z BIG4, BIG6 a JAFFE databázy a optimalizované masky váh (tmavý obdĺžnik znamená 0, biely 1).

Výsledkami práce sme dospeli k nasledujúcim záverom:

- navrhovaná metóda umožňuje významnú redukciu dimenzie dát so zachovaním úspešnosti klasifikácie
- na základe Obr. 4.11 (vľavo) sú pre rozpoznávanie subjektov dôležitými miestami tváre: oko a jeho okolie, koreň nosa a oblasť nosových dierok, čelo, časť líca a brady, pričom sú dané oblasti situované na pravej strane tváří
- na základe Obr. 4.11 (v strede) zodpovedá dôležitým oblastiam tváre: horná (čelo a obočie) a dolná časť (brada, ústa a okolie) obrazu
- na základe Obr. 4.11 (vpravo) sú pre rozpoznávanie emócií dôležité oblasti tváre: horná (čelo) a najmä stredná časť hlavy ako oči, líca a horná pera
- vymenované oblasti často obsahujú textúry v podobe vlasov, chlupov, fúzov, brady a v prípade očí objekty ako napr. okuliare, pričom tieto informácie môžu byť použité v procese klasifikácie
- v prípade množiny BIG4 boli použité bloky takmer štvorcového tvaru, čo umožnilo asymetrický výber dôležitých oblastí (zodpovedá to výberu v publikáciách [TAI09], [LI05], [ZHA05])
- v prípade BIG6 a JAFFE boli použité obdĺžnikové bloky (po riadkoch), čo malo za dôsledok symetrický výber oblastí (zodpovedá to výberu v publikáciách [NIK12], [SHA05])
- výber oblastí do určitej miery závisí od obsahu databázy, počtu vzoriek, type použitých príznakov a veľkosti blokov
- výhodou navrhovaného prístupu je zníženie pamäťovej a výpočtovej náročnosti, možnosti nahradiť LBP príznaky iným druhom príznakov, simulované žihanie iným optimalizačným algoritmom, možnosť optimalizovať väčší počet parametrov atď.
- obmedzenie spočíva v opakovaných výpočtoch a najmä v prípade využitia náročnejších algoritmov ako SVM, ANN, hlbokých (angl. deep) sietí (konvolučných sietí)

4.6 Porovnanie dosiahnutých výsledkov

Na základe údajov uvedených v predošlých kapitolách 4.3, 4.4 a 4.5 sme vytvorili nasledujúcu Tab. 4.10, v ktorej sú uvedené publikované a nami dosiahnuté výsledky. Pre prehľadnosť sme experimenty usporiadali do stĺpcov a označili nasledovne:

- Exp. č. 1 - Zodpovedá problému automatického výberu tréningových vzoriek a článku [7], ktorý sa nachádza na str. 12 a maximálna hodnota je uvedená v Tab. 4.5.
- Exp. č. 2 - Zodpovedá problému optimalizácie LBP parametrov a článku [9], ktorý sa nachádza na str. 15 a hodnoty sú uvedené v Tab. 4.6.
- Exp. č. 3 - Zodpovedá tiež problému optimalizácie LBP parametrov a článku [15], ktorý sa nachádza na str. 16. Hodnoty pre jednu škálu obrazov sú uvedené v Tab. 4.8, hodnoty pre tri škály sa nachádzajú v Tab. 4.7.
- Exp. č. 4 - Zodpovedá problému identifikácie významných oblastí tváre a článku [10], ktorý sa nachádza na str. 19 a hodnoty pochádzajú z Tab. 4.9.

Tabuľka 4.10: Porovnanie publikovaných a nami dosiahnutých výsledkov

Databáza a referen.	Publik. ACC [%]	Dosiahnutá ACC [%]				Roz- diel [%]
		č. 1	č. 2	č. 3		č. 4
				1 škála	3 škály	
FERET(fb) [SHA06]	99,60					-0,940
[NGU12]	99,60					-0,940
FERET(fc) [SHA06]	99,00			98,316	98,660	-0,340
[NGU12]	99,50					-0,840
FERET(-) [NIK14]	99,00					-0,340
FERET [LIA07]	93,16			99,096	99,498	+6,338
ATT [LIA07]	98,50					+1,500
[WAN13]	≈98,7		100,000	99,500	99,500	+1,300
Ext. YALE [TAN10]	99,28		94,056	95,079	93,949	-4,201
[FAN14]	≈96,0					≈ -0,9
CMU PIE [REN13]	99,85					-5,377
[FAN14]	≈91,4	94,473	85,326	87,274	87,675	≈+3,1
JAFFE [ZAV11]	96,20					+1,474
						+1,733
[GIR14]	98,57					-0,896
						-0,637

Na základe porovnania úspešností a uvedených rozdielov je potrebné uviesť aj slovné zhodnotenie. Úplné informácie sa nachádzajú v dizertačnej práci. Uvedené detaily sa týkajú počtu tréovacích a testovacích obrazov/subjektov, použitých metód a najmä počtu klasifikovaných príznakov.

FERET (fb, fc, -)

Nami dosiahnuté výsledky dosahujú maximálny rozdiel v intervale $-0,94$ a $-0,34\%$. Avšak vzhľadom na uvedené detaily dosahujú úspešnosti klasifikácie porovnateľné so sofistikovanejšími súčasnými metódami.

FERET

Výsledok dosahuje maximálny rozdiel $+6,3\%$, čo je do určitej miery spôsobené vyšším počtom tréovacích vzoriek v našom prípade. Ale aj v prípade použitia množiny BIG4 dosiahnutú úspešnosť možno považovať za zlepšenie klasifikácie.

ATT

V oboch prípadoch bolo dosiahnuté zlepšenie minimálne o $+1,3\%$. Úplne porovnateľné výsledky sú vo stĺpci Exp. č. 3, čo potom zodpovedá zlepšeniu o minimálne $+0,8\%$.

Ext. YALE

Zásadný rozdiel v oboch prípadoch spočíva v rádovo vyššom počte príznakov. Ak považujeme daný fakt za negatívny jav, potom sme dosiahli porovnateľné výsledky so súčasným stavom.

CMU PIE

Aj tu sa podobne ako v prípade použitia Ext. YALE databázy môžeme odvolať na niekoľkonásobne dlhší vektor príznakov. Za prínos považujeme robustnosť nami navrhovanej metódy, ktorú je možné aplikovať nielen na malé, ale aj na veľké databázy.

JAFFE

Aj napriek spomenutému rozdielu v spôsobe rozdelenia klasifikovaných emócií aj tu možno argumentovať vyšším počtom klasifikovaných príznakov. Ako je možné vidieť, nami navrhovaná metóda má využitie aj v oblasti rozpoznávania emócií.

4.7 Možnosti využitia v ďalších oblastiach

Získané vedomosti z literatúry a našich dosiahnutých výsledkov je možné použiť aj v iných oblastiach. Metódy extrakcie príznakov, zhukovacie a klasifikačné algoritmy sme využili pri návrhu modulu na rozpoznávanie subjektov na základe snímky ucha. Ďalej sme analyzovali možnosti strojového učenia pre detekciu chorobných zmien na sietnici.

4.7.1 Rozpoznávanie subjektov pomocou snímky ucha

Rozvoj mobilných zariadení a ich masívne používanie znamená nové príležitosti pre biometriu. Nakoľko však tvár nie je jediný spôsob ako rozpoznávať ľudí, vytvorili sme mobilnú aplikáciu, ktorá rozpoznáva subjekty na základe tvaru a štruktúry ucha [12]. Navrhnutý systém môže byť použitý pri tvorbe multimodálneho biometrického systému. Boli vytvorené dve verzie aplikácie pre mobilné platformy Android OS a Windows Phone. Aplikácia pracuje v reálnom čase a v nasledujúcich dvoch módoch: registrácia používateľa (získanie biometrických dát, extrakcia príznakov a ich uloženie do databázy na vzdialenom serveri) a rozpoznávanie (získanie biometrických dát, extrakcia príznakov, spracovanie a klasifikácia na serveri, zobrazenie výsledku).

Pre účely tréovania a testovania boli použité dve rôzne databázy uší: nami vytvorená databáza (pozostávala z 20 obrazov pravého ucha získaných pomocou prehliadača Google; obrazy mali rôznu veľkosť a rozlíšenie; okrem jedného obrazu boli všetky farebné; uši neboli prekryté) a WPUTE databáza (štandardná databáza; veľkosť obrazov bola 500*380 pixelov; každý subjekt obsahoval 4 vzorky; obrazy boli farebné a niekedy sú uši prekryté).

Boli použité nasledujúce fázy a metódy:

- **predspracovanie**
podvzorkovanie obrazov získaných kamerou zabudovanou v mobilnom zariadení, transformácia na šedotónový obraz, ekvalizácia histogramu
- **detekcia**
získanie polohy ucha pomocou jedného z nasledujúcich detektorov: kaskádový detektor využívajúci Haarove príznaky (Viola-Jones algoritmus) alebo nami vytvorený a natré-
novaný detektor využívajúci LBP príznaky a orezanie obrazu na 160*80 pixelov
- **extrakcia príznakov a redukcia dimenzie dát** (PCA alebo $LBP_{8,1}^{u2}$)
- **klasifikácia dát** (Euklidova vzdialenosť (PCA) alebo χ^2 vzdialenosť (LBP))

Nasledujúca Tab. 4.11 obsahuje dosiahnuté výsledky.

Tabuľka 4.11: Porovnanie výsledkov jednotlivých databáz [12]

Databáza	ACC [%]	
	LBP	PCA
Google (bez natočenia)	90	100
Google (natočenie max. $\pm 10^\circ$)	65	55
WPUTE	60	40

Na základe výsledkov je možné stanoviť nasledujúce závery. Navrhnutý biometrický systém bol testovaný a tréovaný na malom počte vzoriek s nízkou kvalitou. To sa negatívne odrazilo na úspešnosti klasifikácie. Čo sa týka použitého druhu príznakov, LBP v podstate dosahuje o niečo lepšie výsledky ako metóda PCA. Aby navrhnutý systém dosahoval akceptovateľnú úroveň úspešnosti, vyžaduje si proces detekcie a extrakcie príznakov vykonať hlbšiu analýzu.

4.7.2 Detekcia chorobných zmien na sietnici

Ďalšou oblasťou, kde je možné spomínané metódy využiť, sú medicínske obrazy. Vplyvom životného štýlu, nesprávnou životosprávou alebo vplyvom genetiky môže ľudský organizmus ochorieť na diabetes (tzv. cukrovka). Vplyvom zvýšenej hladiny cukru dochádza k zmenám na cievach a tie sa môžu prejaviť na sietnici oka. Diabetická retinopatia (DR) je pojem označujúci patologické zmeny sietnice a ciev.

V našom nasledujúcom výskume sme sa snažili zlepšiť, resp. doplniť nasledovné:

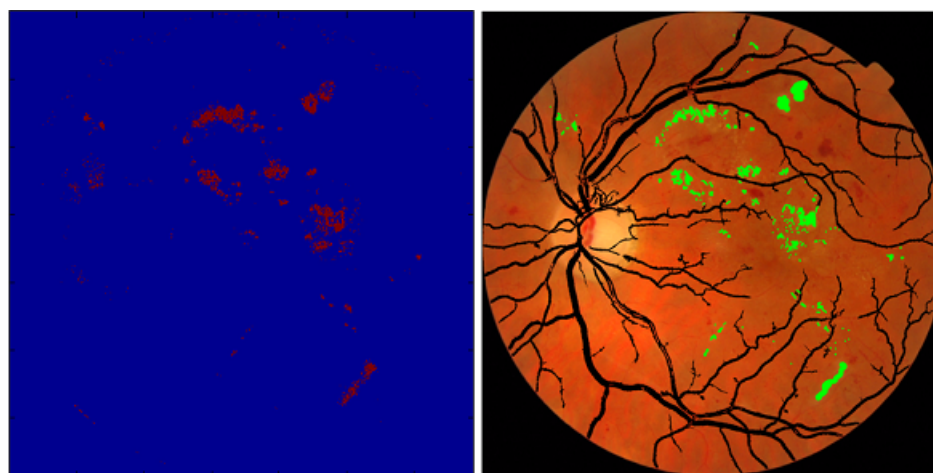
1. detekcia chorobných zmien na úrovni blokov
2. detekcia chorobných zmien aj na úrovni pixelov
3. porovnanie výsledkov na databáze Messidor
4. porovnanie výsledkov pri použití klasifikátorov ako MLP alebo RBF siete
5. predspracovanie obrazov pomocou Foracchia algoritmu [FOR05]
6. využitie nielen G zložky, ale farebných obrazov (napr. chominančných zložiek z YCgCr alebo YCbCr modelu)

Systém bude obsahovať nasledujúcu postupnosť krokov:

1. predspracovanie
pozostáva z aplikovania Foracchia algoritmu [FOR05] na zelenú zložku a rozdelenia obrazov na bloky
2. detekcia exudátov na úrovni blokov
pomocou LBP príznakov, využitie napr.: RBF, MLP, SVM klasifikátora na detekciu
3. detekcia exudátov na úrovni pixelov
použitie YCgCr/YCbCr, využitie napr.: RBF, MLP, SVM klasifikátora na detekciu

Na nasledujúcom Obr. 4.12 je uvedený príklad detekcie chorobných zmien nami navrhovanou metódou.

Nakolko v procese písania, resp. tlačenia tejto práce neboli vykonané potrebné experimenty, nie je možné s istotou stanoviť úspešnosť navrhovaného systému, a ani vyvodiť žiadne závery. Predpokladáme, že navrhovaný systém má potenciál dosiahnuť dobré výsledky.



Obr. 4.12: Poloha detegovaných chorobných zmien (tvrdé exudáty) a snímka sietnice bez ciev a označená poloha exudátov

5 Prínosy práce

Vzhľadom na stanovené ciele, vykonané experimenty a získané poznatky považujeme za prínosy tejto dizertačnej práce nasledovné:

- Návrh metódy lokalizácie tvárových črt, pomocou malej banky GF, kde redukcia dimenzie dát bola uskutočnená pomocou výpočtu štatistických veličín klasifikovaných MLP sieťou (kapitola 4.1).
- Návrh automatizovaného výberu trénovacích vzoriek (obrazov tvári používateľov) v rámci 7. rámcového projektu HBB-NEXT. Jedným z cieľom projektu HBB-NEXT bol prístup k užívateľskému kontu a nastaveniam v rámci smart TV pomocou snímacej kamery Kinect. Zozbierané obrazy tváre boli vyberané pomocou zhlukovacieho algoritmu (K-means). Extrakcia príznakov a redukcia dimenzie dát bola zabezpečená pomocou kombinácie metód LBP a PCA. Klasifikácia získaných príznakov metódou SVM (kapitola 4.3).
- Nájdenie optimálneho rozdelenia obrazov príznakov do blokov vzhľadom na použité databázy tvári a nastavenie ďalších parametrov pomocou genetického algoritmu. Optimalizačný proces zahŕňal porovnanie viacerých modifikácií metódy LBP (kapitola 4.4), za účelom dosiahnutia najvyššej úspešnosti rozpoznávania subjektov pri značne redukovanej dimenzii vektora príznakov.
- Vyhodnotenie úspešnosti navrhovaných a optimalizovaných systémov s existujúcimi prístupmi (kapitola 4.6).
- Návrh metodiky na identifikáciu významných oblastí tváre pomocou optimalizačného algoritmu (simulované žihanie) (kapitola 4.5).
- Aplikácia navrhovaných systémov v oblasti rozpoznávania emócií. Nájdenie optimálneho rozdelenia príznakov do blokov a identifikácia významných oblastí tváre dôležitých pre klasifikáciu emócií a výraznú redukciu dimenzie dát (kapitola 4.5).
- Spolupráca pri návrhu mobilnej aplikácie na rozpoznávanie subjektov pomocou snímky ucha využívajúca LBP príznaky a Haarov detektor založený na LBP príznakoch (kapitola 4.7.1). Návrh použitého typu príznakov, rozdelenie do blokov, výber vzoriek a spôsobov klasifikácie príznakov.
- Návrh metodiky na detekciu chorobných zmien (napr. tvrdých exudátov) na sietnici na úrovni blokov a pixelov pomocou LBP príznakov, YCbCr/YCgCr farebného modelu a Foracchia algoritmu. Príznaky sú klasifikované SVM algoritmom (kapitola 4.7.2).

Aj keď jednotlivé moduly fungujú samostatne, je možné ich spojiť a použiť pri tvorbe multimodálneho biometrického systému. Štúdium bolo zamerané na oblasť biometrie, konkrétne biometrické rozpoznávanie vzorov pomocou metód strojového učenia.

Literatúra

- [AHO04] AHONEN, T. - HADID, A. - PIETIKÄINEN, M. *Face Recognition with Local Binary Patterns*. In *European Conference on Computer Vision (ECCV)*. ISSN 0302-9743. Springer. 2004. p. 469–481.
- [ALA12] ALATTAB, A. - KAREEM, S. *Efficient Method of Visual Feature Extraction for Facial Image Detection and Retrieval*. In *4th International Conference on Computational Intelligence, Modelling and Simulation*. ISSN 2166-8531, 2012, p. 220-225.
- [ALS98] ALSABTI, K. - RANKA, S. - SINGH, V. *An Efficient K-Means Clustering Algorithm*. In *Proceedings of IPPS/SPDP Workshop on High Performance Data Mining*. 1998, p. 6.
- [ALY10] ALY, S. et al. *Robust Face Recognition Using Multiple Self-Organized Gabor Features and Local Similarity Matching*. In *20th International Conference on Pattern Recognition*. ISSN 1051-4651, 2010, p. 2909-2912.
- [ANK99] ANKERST, M. et al. *OPTICS: Ordering Points To Identify the Clustering Structure*. In *ACM SIGMOD international conference on Management of data*. ISBN: 1-58113-084-8, 1999, vol. 28, no. 2, p. 49-60.
- [ASH14] ASHRAF, M. - HABIB, Z. - JUSSAIN, M. *Texture Feature Analysis of Digital Fundus Images for Early Detection of Diabetic Retinopathy*. In *Computer Graphics, Imaging and Visualization (CGIV)*. ISBN 978-1-4799-5720-0 , 2014, p. 57-62.
- [BEL97] BELHUMEUR, P. - HESPANHA, J. - KRIEGMAN, D. *Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection*. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*. ISSN 0162-8828, 1997, vol. 19, no. 7, p. 711-720.
- [BIG13] BIGLARI, M. - MIRZAEI, F. - EBRAHIMPOUR-KOMEH, H. *Illumination invariant face recognition using SQI and weighted LBP histogram*. In *First Iranian Conference on Pattern Recognition and Image Analysis (PRIA)*. 2013. p. 1-7.
- [BON13] BONGJIN, J. - INHO, C. - DAIJIN, K. *Hybridization for Accurate Face and Human Detection*. In *Pattern Analysis and Machine Intelligence*. ISSN 0162-8828, 2013, vol. 35, no. 6, p. 1423-1436.
- [BUR98] BURGESS, Ch. *A tutorial on support vector machines for pattern recognition*. In *Data Mining and Knowledge Discovery*. ISSN 1384-5810, 1998, vol. 2, no. 2, p. 121-167.
- [CAR06] CARANDINI, M. *What simple and complex cells compute*. In *Physiological Society*. 2006, p. 463-466.
- [CHA07] CHAN, C. - KITTLER, J. - MESSER, K. *Multi-scale local binary pattern histograms for face recognition*. In *International conference on Advances in Biometrics*. ISSN 0302-9743. 2007. p. 809-818.

- [DAU85] DAUGMAN, J. *Uncertainty Relation for Resolution in Space, Spatial Frequency, and Orientation Optimized by Two-Dimensional Visual Cortical Filters*. In *Journal of Optical Society America*. 1985, vol. 2, no. 7, p. 1160-1169.
- [EKM93] EKMAN, P. *Facial Expression and Emotion*. In *American Psychologist*. 1993. p. 384-392.
- [EST96] ESTER, M. et al. *A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise*. In *Data Mining and Knowledge Discovery*. ISSN 1384-5810, 1996, vol. 2, no. 2, p. 169-194.
- [FAN14] FAN, K. - HUNG, T. *A Novel Local Pattern Descriptor—Local Vector Pattern in High-Order Derivative Space for Face Recognition*. In *IEEE Transactions on Image Processing*. ISSN 1941-0042, 2014, vol. 23, no. 7, p. 2877-2891.
- [FAW03] FAWCETT, T. *Roc graphs: Notes and practical considerations for data mining researchers*. In *ReCALL (Technical Report Hewlett Packard HPL-2003-4)*. 2004, vol. 31, p. 1-38.
- [FAZ09] FAZLI, S. - AFROUZIAN, R. - SEYEDARABI, H. *High-Performance Facial Expression Recognition Using Gabor Filter and Probabilistic Neural Network*. In *IEEE International Conference on Intelligent Computing and Intelligent Systems*. ISBN 978-1-4244-4754-1, 2009, vol. 4, s 93-96.
- [FIE87] FIELD, D. *Relations between the statistics of natural images and the response properties of cortical cells*. In *Journal of Optical Society America*. 1987, vol. 4, no. 12, p. 2379-2394.
- [FOR05] FORACCHIA, M. - GRISAN, E. - RUGGERI, A. *Luminosity and contrast normalization in retinal images*. In *Medical Image Analysis*. 2005, vol. 9, no. 5, s 179-190.
- [GAB46] GABOR, D. *Theory of Communication*. In *JIEEE*. 1946, vol. 93, no. 26, p. 429-457.
- [GIR14] GIRISH, G. - SHRINIVASA, N. - PRADIP, D. *Face Recognition using MB-LBP and PCA: A Comparative Study*. In *International Conference on Computer Communication and Informatics*. 2014. p. 6. ISBN 978-1-4799-2352-6
- [GRE08] GREGORY, P. – SIMON, M. *Biometrics for dummies*. NJ : Wiley Publishing, 2008, 308 p. ISBN 978-0-470-29288-4.
- [HAP12] HAPPY, S. - GEORGE, A. - ROUTRAY, A. *A Real Time Facial Expression Classification System Using Local Binary Patterns*. In *International Conference on Intelligent Human Computer Interaction (IHCI)*. 2012. p. 1–5.
- [HAY98] HAYKIN, S. *Neural Networks: Comprehensive Foundation*. 2. vyd. NJ : Prentice Hall, 1998. 842 p. ISBN 0132733501.
- [HUA11] HUANG, D. *Local Binary Patterns and Its Application to Facial Image Analysis: A Survey*. In *IEEE Transactions on Systems, Man and Cybernetics*. ISSN 1558-2442, 2011, vol. 41, no. 6, p. 765-781.
- [CHA11] CHANDRAPPA, D. - RAVISHANKAR, M. - RAMESHBABU, D. *Face Detection in Color Images using Skin Color Model Algorithm based on Skin Color Information*. In *3rd International Congress on Electronics Computer Technology*. ISBN 978-1-4244-8679-3, 2011, vol. 1, p. 254-258.

- [CHI13] CHINTALAPTI, S. - RAGHUNADH, M. *Automated attendance management system based on face recognition algorithms*. In *Computational Intelligence and Computing Research*. ISBN 978-1-4799-1594-1, 2013, p. 1-5.
- [CHU08] CHUAN-XU, W. - XUE, L. *Face detection using BP network combined with Gabor wavelet transform*. In *9th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing*. ISBN 978-0-7695-3263-9, 2008, p. 526-529.
- [JEM08] JEMAA, Y. - KHANFIR, S. *Automatic Gabor Features Extraction for Face Recognition using Neural Networks*. In *1st Workshops on Image Processing Theory, Tools and Applications*. ISBN 978-1-4244-3321-6, 2008, p. 1-6.
- [JIA13] JIA, X. et al. *Multi-scale Block Local Ternary Patterns for Fingerprints Vitality Detection*. In *International Conference on Biometrics (ICB)*. ISSN 2376-4201. 2013. p. 1-6.
- [KAN00] KANADE, T. - COHN, J. - TIAN, Y. *Comprehensive database for facial expression analysis*. In *IEEE International Conference on Automatic Face and Gesture Recognition*. 2000. p. 46-53.
- [KIT08] KITTLER, J et al. *Image Feature Localization by Multiple Hypothesis Testing of Gabor Features*. In *IEEE Transactions on Image Processing*. 2008, vol. 17, no. 3, p. 311-325.
- [KOH82] KOHONEN, T. *Self-organizing formation of topologically correct feature maps*. In *Biological Cybernetics*. 1982, vol. 43, no. 1, p. 59-69.
- [KOH95] KOHAVI, R. *A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection*. In *14th International Joint Conference on Artificial Intelligence*. ISBN 1-55860-363-8, 1995, vol. 2, p. 1137-1143.
- [KRU99] KRUIZINGA, P. - PETKOV N. *Non-linear operator for oriented texture*. In *IEEE Trans. on Image Processing*. ISSN 1057-7149, 1999, vol. 8, no. 10, p. 1395-1407.
- [KVA00] KVASNIČKA, V. - POSPÍCHAL J. - TIŇO P. *Evolučné algoritmy*. 1. vyd. Bratislava : Vydavateľstvo STU BA, 2000. 215 s. ISBN 80-227-1337-5.
- [LAJ08] LAJEVARDI, S. - LECH, M. *Facial Expression Recognition Using Neural Networks and Log-Gabor Filters*. In *Digital Image Computing: Techniques and Applications*. ISBN 978-0-7695-3456-5, 2008, p. 77-83.
- [LEE10] LEE, Ch. - SHIH, Ch. *Gabor Feature Selection and Improved Radial Basis Function Networks for Facial Expression Recognition*. In *International Conference on Information Science and Applications*. ISBN 978-1-4244-5941-4, 2010, p. 1-8.
- [LI12] LI, Q. - NIAZ, U. - MIERALDO, B. *An Improved Algorithm on Viola-Jones Object Detector*. In *10th International Workshop on Content-Based Multimedia Indexing*. ISSN 1949-3983, 2012, p. 1-6.
- [LI05] LI, S. et al. *Learning to fuse 3D+ 2D based face recognition at both feature and decision levels*. In *Second International Workshop (AMFG)*. ISSN 0302-9743. Springer. 2005. p. 44-54.
- [LI10] LI, W. - MAO, K. - ZHANG, H. - CHAI, T. *Selection of Gabor Filters for Improved Texture Feature Extraction*. In *17th International Conference on Image Processing*. ISSN 1522-4880. 2010, p. 361-364.

- [LIA07] LIAO, S. - CHUNG, A. *Face Recognition by Using Elongated Local Binary Patterns with Average Maximum Distance Gradient Magnitude*. In *Conference on Computer Vision*. ISSN 0302-9743 2007, p. 672-679
- [LOP10] LÓPEZ, L. *Local Binary Patterns applied to Face Detection and Recognition*. Universitat Politècnica de Catalunya, 2010, p. 146.
- [LYO98] LYONS, M. et al. *Coding Facial Expressions with Gabor Wavelets*. In *International Conference on Automatic Face and Gesture Recognition*. 1998. p. 200-205.
- [MAG11] MAGESHKUMAR, C. et al. *Gabor features and LDA based Face Recognition with ANN classifier*. In *International Conference on Emerging Trends in Electrical and Computer Technology*. ISBN 978-1-4244-7923-8, 2011, p. 831-836.
- [MON13] MONTENEGRO, J. - GOMEZ, W. - SANCHEZ-ORELLANA, P. *A comparative study of color spaces in skin-based face segmentation*. In *Electrical Engineering, Computing Science and Automatic Control*. ISBN 978-1-4799-1460-9, 2013, p. 313-317.
- [NGU12] NGUYEN, T. - VU, N. - CAPLIER, A. *Face Recognition using Multi-modal Binary Patterns*. In *International Conference on Pattern Recognition*. ISSN 1051-4651. 2012, p. 2343-2346
- [NIK12] NIKISINS, O. - GREITANS, M. *A mini-batch discriminative feature weighting algorithm for LBP - Based face recognition*. In *International Conference on Imaging Systems and Techniques (IST)*. ISSN 1558-2809. 2012. p. 170–175.
- [NIK14] NIKAN, S. - AHMADI, M. *Effectiveness of Various Classification techniques on Human Face Recognition*. In *International Conference on High Performance Computing & Simulation*. ISBN 978-1-4799-5313-4. 2014. p. 651-655.
- [OJA02] OJALA, T - PIETIKÄINEN, M - MÄENPÄÄ, T. *Multiresolution Gray Scale and Rotation Invariant Texture Classification with Local Binary Pattern*, In *IEEE Transactions on Pattern Analysis and Machine Intelligence*. ISSN: 0162-8828, 2002, vol. 24, no. 7, p. 971–987.
- [ORA98] ORAVEC, M. a kol. *Neurónové siete pre číslicové spracovanie signálov*. 1. vyd. Bratislava : FABER, 1998. 197 s. ISBN 80-967503-9-9.
- [PER11] PEREIRA, E. et al. *Investigation of Local and Global Features for Face Detection*. In *IEEE Symposium on Computational Intelligence for Multimedia, Signal and Vision Processing*. ISBN 978-1-4244-9913-7, 2011, p. 114-121.
- [PHI00] PHILLIPS, P. et al. *The FERET evaluation methodology for face recognition algorithms*. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*. ISSN 0162-8828. 2000. vol. 22, no. 10. p. 1090-1104.
- [RAG13] RAGUL, G. et al. *Comparative study of statistical models and classifiers in face recognition*. In *Information Communication and Embedded Systems*. ISBN 978-1-4673-5786-9, 2013, p. 623-628.
- [REN13] REN, J. - JIANG, X. - YUAN, J. *Relaxed local ternary pattern for face recognition*. In *International Conference on Image Processing*. ISSN 2381-8549, 2013, p. 3680-3684.
- [RUS05] RUST, N. - MOVSHON, A. *In praise of artifice*. In *Nature Neuroscience*. 2005, vol. 8, no. 12, p. 1647-1650.

- [SAT13] SATPATHY, A. - JIANG, X. - ENG, H. *Human Detection Using Discriminative And Robust Local Binary Pattern*. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. ISSN 1520-6149. 2013. p. 2376-2380.
- [SEE13] SEE, Y. - NOOR, N. - LAI, A. *Hybrid face detection with skin segmentation and edge detection*. In *Signal and Image Processing Applications*. ISBN 978-1-4799-0267-5 , 2013, p. 406-411.
- [SHA05] SHAN, C. - GONG, S. - MCOWAN, P. *Robust facial expression recognition using local binary patterns*. In *IEEE International Conference on Image Processing (ICIP)*. ISSN 1522-4880. 2005. vol. 2, p. 370-373.
- [SHA06] SHAN, S. et al. *Ensemble of Piecewise FDA Based on Spatial Histograms of Local (Gabor) Binary Patterns for Face Recognition*. In *International Conference on Pattern Recognition*. ISSN 1051-4651. 2006, vol. 4. p. 606-609.
- [SHI08] SHIRDHONKAR, S. - JACOBS, D. *Approximate earth mover's distance in linear time*. In *IEEE Computer Vision and Pattern Recognition*. 2008. p. 1-8.
- [SIM02] SIM, T. - BAKER, S. - BSAT, M. *CMU Pose, Illumination, and Expression (PIE) database*. In *Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition*. 2002. p. 46-51.
- [SUR12] SURULIANDI, A. - MEENA, K. - ROSE, R. *Local binary pattern and its derivatives for face recognition*. In *IET Computer Vision*. ISSN 1751-9640. 2012. vol. 6, no. 5, p. 480-488.
- [TAI09] TAINI, M. - ZHAO, G. - PIETIKÄINEN, M. *Weight-based facial expression recognition from nearinfrared video sequences*. In *Scandinavian Conference (SCIA)*. ISSN 0302-9743. Springer. 2009. p. 239-248.
- [TAN10] TAN, X. - TRIGGS, B. *Enhanced local texture feature sets for face recognition under difficult lighting conditions*. In *IEEE Transactions on Image Processing*. ISSN 1057-7149. 2010. vol. 19, no. 6. p. 1635-1650.
- [TUR91] TURK, M. - PENTLAND, A. *Eigenfaces for recognition*. In *Journal of Cognitive Neuroscience*. ISSN 0898-929X, 1991, vol. 3, no. 1, p. 71-86.
- [VAP95] CORTES, C. - VAPNIK, V. *Support-Vector Networks*. In *Machine Learning*. ISSN 0885-6125, 1995, vol. 20, no. 3, p. 273-297.
- [VER13] VERMA, R. - DABBAGH, M. *Fast facial expression recognition based on local binary patterns*. In *Electrical and Computer Engineering*. ISSN 0840-7789, 2013, p. 1-4.
- [WAN12] WANG, B. - LI, W. - LIAO, Q. *Face recognition based on nonsubsampling contourlet transform and block-based kernel Fisher linear discriminant*. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. ISSN 1520-6149. 2012. p. 1533-1536.
- [WAN13] WANG, B. et al. *Face Recognition Based on the Feature Fusion of 2DLDA and LBP*. In *International Conference on Information, Intelligence, Systems and Applications*. ISBN 978-1-4799-0771-7. 2013. p. 1-6.
- [VIO01] VIOLA, P. - JONES, M. *Rapid object detection using a boosted cascade of simple features*. In *IEEE Conference on Computer Vision and Pattern Recognition*. ISSN 1063-6919, 2001, vol. 1, p. 511-518.

- [YIN09] YING, Z. CAI, L. *Support Vector Discriminant Analysis on Local Binary Patterns for Facial Expression Recognition*. In *2nd International Congress on Image and Signal Processing*, ISBN 978-1-4244-4129-7, 2009, vol. 1, p. 1-4.
- [ZAV11] ZAVASCHI, T. - KOERICH, A. - OLIVEIRA, L. *Facial expression recognition using ensemble of classifiers*. In *International Conference on Acoustics, Speech and Signal Processing*. ISSN 2379-190X, 2011, p. 1489-1492.
- [ZHA03] ZHANG, D. - LU, G. *Evaluation of similarity measurement for image retrieval*. In *Neural Networks and Signal Processing*, 2003, vol. 2, p. 928-931.
- [ZHA05] ZHANG, G. et al. *Boosting local binary pattern (LBP)-based face recognition*. In *Chinese Conference on Biometric Recognition (SINOBIOMETRICS)*. ISSN 0302-9743. Springer. 2005. p. 179-186.
- [ZHA10] ZHAO, M. - LI, P. - WANG, L. *A Novel Complete Face Detection Method in Color Images*. In *3rd International Congress on Image and Signal Processing*. ISBN 978-1-4244-6513-2, 2010, vol. 4, p. 1763-1766.
- [ZHO10] ZHONG, Z. *A New Adaptive Color Space YCH for Face Detection*. In *International Conference on Intelligent System Design and Engineering Application*. ISBN 978-1-4244-8333-4, 2010, vol. 1, p. 84-89.
- [ATTD] *The Database of Faces*. [online]. Dostupné na internete: <<http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>>.
- [CVTI] *Computer Vision Test Images*. [online]. Dostupné na internete: <<http://www.cs.cmu.edu/~cil/v-images.html>>.
- [FADA] *Face Database: Park Aging Mind Lab*. [online]. Dostupné na internete: <<http://agingmind.utdallas.edu/facedb>>.
- [FARE] DEBRUINE, L. - JONES, B. *Face Research: Experiments about face and voice preferences*. [online]. Dostupné na internete: <<http://faceresearch.org/>>.
- [FRHD] *Face Recognition Homepage: Databases*. [online]. Dostupné na internete: <<http://www.face-rec.org/databases/>>.
- [JAFD] *Facial Expression Database: Japanese Female Facial Expression (JAFFE) Database*. [online]. Dostupné na internete: <<http://www.kasrl.org/jaffe.html>>.
- [MESD] *ADCIS Download Third Party: Messidor Database*. [online]. Dostupné na internete: <<http://www.adcis.net/en/Download-Third-Party/Messidor.html>>.
- [PIED] *Robotics Institute: PIE Database*. [online]. Dostupné na internete: <http://www.rti.cmu.edu/research_project_detail.html?project_id=418&menu_id=261>.
- [YALD] *The Extended Yale Face Database B*. [online]. Dostupné na internete: <<http://vision.ucsd.edu/~leekc/ExtYaleDatabase/ExtYaleB.html>>.

Vlastné publikácie

- [1] LODERER, M. *Systémy v ochrane objektov* : Bakalárska práca. Bratislava : FEI STU, 2010. 64 s.
- [2] LODERER, M. *Kryptoanalýza funkcie JH* : Diplomová práca. Bratislava: FEI STU, 2012. 58 s.

ADF – vedecké práce v domácich nekarentovaných časopisoch

- [3] ADAMČEK, P. - **LODERER, M.** (30%) - ZAJAC, P. *A Comparison on Local Reduction and SAT-Solver Based Algebraic Cryptanalysis of JH and Keccak*. In *Tatra Mountains Mathematical Publications*. ISSN 1210-3195, 2012, vol. 53, p. 1-20.

ADM - vedecké práce v zahraničných časopisoch registrovaných v databázach Web of Science alebo SCOPUS

- [4] BENIAK, M - PAVLOVIČOVÁ, J. - **LODERER, M.** (20%) *Minimal Gabor Filter Bank for Face Features Localization*. In *International Review on Computers and Software - IRECOs*. 2014. p. 1977-1983.

AFC – publikované príspevky na zahraničných vedeckých konferenciách

- [5] **LODERER, M.** (50%) - PAVLOVIČOVÁ, J. - ORAVEC, M. *Selection of Gabor Features for Face Features Localization*. In *Mezinárodní Masarykova konference pro doktorské a mladé vědecké pracovníky: Sborník příspěvků z mezinárodní vědecké konference*. Hradec Králové: Magnanimitas, 2012. ISBN 978-80-905243-3-0. p. 3247-3255.
- [6] BAN, J. - FÉDER, M. - JIRKA, V. - **LODERER, M.** (20%) - OMELINA, Ľ. - ORAVEC, M. - PAVLOVIČOVÁ, J. *An Automatic Training Process Using Clustering Algorithms for Face Recognition System*. In *Proceedings ELMAR: 55th International Symposium*. ISSN 1334-2630. 2013. p. 15-18.

Citácie:

HONG, H. et al. *Fuzzy System-Based Face Detection Robust to In-Plane Rotation Based on Symmetrical Characteristics of a Face*. In *Symmetry journal*. ISSN 2073-8994. 2016, vol. 8, no. 8, p. 75.

- [7] **LODERER, M.** (50%) - PAVLOVIČOVÁ, J. - FÉDER, M. - ORAVEC, M. *Data Dimension Reduction in Training Strategy for Face Recognition System*. In *IWSSIP: 21st International Conference on Systems, Signals & Image Processing*. ISSN 2157-8672. 2014. p. 263-266.

Citácie:

PUJOL, F. - MORA, H. - GIRONA-SELVA, J. *A connectionist computational method*

for face recognition Methodologies. In *International Journal of Applied Mathematics and Computer Science*. 2016, vol. 26, no. 2. p. 451–465.

- [8] BENIAK, M - **LODERER, M.** (20%) - PAVLOVIČOVÁ, J. *Optimal Gabor Filter Bank Selection for Face Features Localization*. In *Redžúr: 8th International Workshop on Multimedia and Signal Processing*. ISBN 978-80-227-4162-0. 2014. p. 57-60.
- [9] **LODERER, M.** (60%). - PAVLOVIČOVÁ, J. *Optimization of LBP Parameters*. In *ELMAR: 56th International Symposium*. ISSN 1334-2630. 2014. p. 111-114.

Citácie:

ALSALIBI, B. et al. *The Impact of Bio-Inspired Approaches Toward the Advancement of Face Recognition*. In *Journal ACM Computing Surveys (CSUR) Surveys Homepage archive*. 2015, vol. 48, no. 1. p. 33.

ORAVEC, M. *Feature Extraction and Classification by Machine Learning Methods for Biometric Recognition of Face and Iris*. In *International Symposium ELMAR*. 2014. p 1-4.

- [10] **LODERER, M.** (50%) - PAVLOVIČOVÁ, J. - ORAVEC, M. - MAZANEC, J. *Face Parts Importance in Face and Expression Recognition*. In *IWSSIP: 22nd International conference on systems, signals and image Processing*. ISSN 2157-8672. 2015. p. 188-191.

AFD – publikované príspevky na domácich vedeckých konferenciách

- [11] **LODERER, M.** (60%) - REPKA, M. *Časový útok využívajúci cache kolízie na posledné kolo šifry AES*. In *Študentská vedecká a odborná činnosť. Zborník vybraných prác*. ISBN 978-80-227-3508-7, 2011, s. 20-24. Dostupné na internete: <http://www.posterus.sk/?p=12812>.
- [12] ORAVEC, M. - PAVLOVIČOVÁ, J. - SOPIAK, D. - JIRKA, V. - **LODERER, M.** (10%) - LEHOTA, Ľ. - VODIČKA, M. - FAČKOVEC, M. - MIHALÍK, M. - TOMÍK, M. - GERÁT, J. *Mobile Ear Recognition Application*, In *IWSSIP: 23rd International conference on systems, signals and image Processing*. 2016. p. 1-4.
- [13] **LODERER, M.** (60%) - PAVLOVIČOVÁ, J. *Comparison of Gabor Filters for Facial Features Localization*. In *Redžúr: 7th International Workshop on Multimedia and Signal Processing*. Bratislava: Nakladateľstvo STU, 2013. ISBN 978-80-227-3921-4. p. 61-64.

AFH – abstrakty vedeckých príspevkov z domácich konferencií

- [14] ADAMČEK, P. - **LODERER, M.** (20%) - ZAJAC, P. *Using Local Reduction to Find Preimages for Round-Reduced JH and Keccak*. In *Tatracrypt: 12th Central European Conference on Cryptology*. Bratislava : Slovak Academy of Sciences, 2012. p. 10-11.

Nepublikované

- [15] **LODERER, M.** (60%) - PAVLOVIČOVÁ, J. - ORAVEC, M. *Local Binary Patterns and Local Binary Patterns Variations in Face Recognition*. Podané na World Congress of the International Federation of Automatic Control (IFAC 2017).

Účasť na projektoch

1. HBB-NEXT project, Next-Generation Hybrid Broadcast Broadband, project of 7th Frame Program FP7-ICT-2011-7-287848, coordinator Rundfunk Berlin Brandenburg, Germany.
2. Pokročilé algoritmy spracovania obrazov na efektívne vyhľadávanie a kódovanie ľudských tvárí (Advanced Image Processing Algorithms for Effective Localization and Coding of Human Faces), grant of the Slovak scientific grant agency VEGA 1/0961/11 (principal investigator J. Pavlovičová, FEI STU), 2011-2013.
3. Návrh pokročilých metód biometrického rozpoznávania na základe obrazov tváre a dúhovky (Proposal of Advanced Methods of Biometric Recognition Based on Face and Iris Images), grant of the Slovak scientific grant agency VEGA 1/0529/13 (principal investigator M. Oravec, FEI STU), 2013-2016.
4. BioDaT, Biometrické rozpoznávanie na základe obrazov Dúhovky a Tváre (Biometric Recognition Based on Face and Iris Images), research grant from Nadácia Tatrabanky (Foundation of Tatrabanka), (principal investigator M. Oravec), 2013.
5. Inovácia vzdelávania v číslicovom spracovaní obrazu, v biometrii a v strojovom učení a neurónových sieťach (Education Innovation in Digital Image Processing, Biometrics, Machine Learning and Neural Networks), grant of the Slovak cultural and educational grant agency KEGA 010STU-4-2014 (principal investigator J. Pavlovičová, FEI STU), 2014-2016.
6. Pokročilé metódy spracovania obrazu z vizuálnych systémov a ich implementácia do vzdelávacieho procesu, grant of the Slovak cultural and educational grant agency KEGA 003STU-4/2014 (principal investigator F. Duchoň, FEI STU), 2014-2016.