

**Ing. Marián Beniak**

**Autoreferát dizertačnej práce**

**DETEKCIA A LOKALIZÁCIA ĽUDSKÝCH TVÁRÍ**

**na získanie akademického titulu doktor (philosophiae doctor, PhD.)**

**v doktorandskom študijnom programe: Kybernetika**

**v študijnom odbore 9.2.7 Kybernetika**

**Bratislava, 2016**



**Dizertačná práca bola vypracovaná** v externej forme doktorandského štúdia

**Na** Ústave robotiky a kybernetiky, Fakulta Elektrotechniky a Informatiky

**Predkladateľ:** Ing. Marián Beniak  
Ústav robotiky a kybernetiky, FEI STU Bratislava  
Ilkovičova 3, 812 19 Bratislava

**Školiteľ:** prof. Ing. Jarmila Pavlovičová, PhD.  
Ústav robotiky a kybernetiky, FEI STU Bratislava  
Ilkovičova 3, 812 19 Bratislava

**Oponenti:**

1. Prof. RNDr. Ing. Ivan Bajla, PhD.  
Ústav merania SAV, Dúbravská cesta 9, Bratislava, 841 04
2. RNDr. Elena Šikudová, PhD.  
Katedra aplikovanej informatiky, Fakulta matematiky fyziky  
a informatiky UK, Mlynská dolina, Bratislava, 824 48

**Obhajoba dizertačnej práce sa koná:**

Ústav robotiky a kybernetiky FEI STU v Bratislave, Ilkovičova 3, 812 19  
Bratislava

prof. Dr. Miloš Oravec, PhD.  
dekan FEI STU Bratislava  
Ilkovičova 3, 812 19 Bratislava



## Obsah

<b>1</b>	<b>ÚVOD .....</b>	<b>4</b>
<b>2</b>	<b>PREHĽAD SÚČASNÉHO STAVU PROBLEMATIKY .....</b>	<b>4</b>
2.1	HISTOGRAMOVÉ METÓDY .....	5
2.2	KOMBINÁCIA GAUSSOVYCH PRAVDEPODOBNOTNÝCH FUNKCIÍ .....	5
2.3	BAYESOV KLASIFIKÁTOR .....	5
2.4	MARKOVVE NÁHODNÉ POLIA .....	6
2.5	METÓDY S VYUŽITÍM NEURÓNOVÝCH SIETÍ A SYSTÉMOV STROJOVÉHO UČENIA.....	7
2.6	GABOROVE FILTRE.....	8
<b>3</b>	<b>CIELE PRÁCE.....</b>	<b>10</b>
3.1	FAREBNÁ ANALÝZA .....	10
3.2	GABOROVE PRÍZNAKY .....	10
3.3	VÝKONNOSTNÁ ANALÝZA.....	10
<b>4</b>	<b>VÝSLEDKY DIZERTAČNEJ PRÁCE A VYKONANÉ EXPERIMENTY .....</b>	<b>11</b>
4.1	FAREBNÁ ANALÝZA .....	11
4.2	GABOROVE PRÍZNAKY .....	12
4.3	VÝKONNOSTNÁ ANALÝZA.....	13
<b>5</b>	<b>SUMARIZÁCIA PRÍNOSOV PRÁCE .....</b>	<b>16</b>
5.1	PRÍNOSY V TEORETICKEJ OBLASTI.....	16
5.2	PRÍNOSY V PRAKTICKEJ OBLASTI .....	16
<b>6</b>	<b>ZHODNOTENIE .....</b>	<b>18</b>
<b>7</b>	<b>ÚČASŤ AUTORA V PROJEKTOCH.....</b>	<b>20</b>
<b>8</b>	<b>PUBLIKÁCIE AUTORA .....</b>	<b>21</b>
<b>9</b>	<b>ZOZNAM POUŽITEJ LITERATÚRY .....</b>	<b>22</b>

# 1 Úvod

V našom výskume sme sa rozhodli venovať problematike detekcie a lokalizácie tváří a tvárových črt. Široké uplatnenie nachádza detekcia a lokalizácia tváří a tvárových črt v systémoch počítačového videnia (*computer vision*) a takisto v automatických monitorovacích a zabezpečovacích systémoch. Predmet výskumu je veľmi aktuálny a dáva široké možnosti pre vývoj nových metód a prístupov na riešenie súvisiacich úloh. Oblasť priemyslu, v ktorej detekcia tváří momentálne nachádza veľké uplatnenie, je digitálna fotografia. Digitálne fotoaparáty vyžívajú detekciu tváří za účelom úpravy kontrastu či osvetlenia oblasti tváre v obraze už pri jeho tvorbe. Následne na úlohu detekcie nadväzuje rozpoznávanie výrazu tváre (hlavne úsmevu) či zatvorených očí, aby nasnímaná fotografia bola naozaj dokonalá a nezaznamenávala rôzne nežiaduce javy. Dôležitým medzikrokom k týmto aplikáciám je lokalizácia tvárových črt, ktorá dovedie systém k najvýznamnejším tvárovým črtám, ktoré sa menia v závislosti od výrazu tváre.

Ďalej je tu obrovské množstvo aplikácií, kde detekcia tváre tvorí predstupeň pri rozpoznávaní identity či výrazu tváre. Americký úrad národnej bezpečnosti má k dispozícii systém na detekciu a extrakciu tváří z veľkého množstva tváří, ktoré sa vyskytujú vo veľkých skupinách ľudí v priestoroch ako sú napríklad letiskové haly, verejné zhromaždenia a podobne (napríklad inaugurácie amerických prezidentov). Primárnou oblasťou nášho záujmu je využitie detekcie tváří a lokalizácie tvárových črt v oblasti robotiky a systémov počítačového videnia. V tejto oblasti ide najmä o systémy zabezpečujúce interakciu robota s obsluhou a následné autonómne vyhodnocovanie a plánovanie budúceho správania. Pre tieto systémy predstavuje detekcia tváří a lokalizácia tvárových črt prvý stupeň potrebný pre rozpoznanie užívateľovej identity, jeho emocionálneho stavu či zámeru. Nielen na rozpoznávanie tváří, ale už aj na ich samotnú detekciu sa vo veľkej miere využívajú systémy strojového učenia a celý rad ďalších kybernetických techník.

## 2 Prehľad súčasného stavu problematiky

Prvým krokom pri spracovaní obrazu ľudskej tváre je určenie a vymedzenie polohy tváre na snímke. Táto úloha je však kvôli variabilnosti pozície, orientácie, pózy a osvetlenia tváre veľkou výzvou pre strojové spracovanie obrazu. Môžeme prejsť k definícii pojmu detekcia tváre (*Face Detection*). Pri použití ľubovoľného obrazu je úlohou detekcie tváre rozhodnúť, či sú alebo nie sú na obraze ľudské tváre. Ak sa tam nachádzajú, detekcia musí určiť polohu a rozmer každej z nich.

Faktory vplývajúce na detekciu tváří môžeme rozdeliť nasledovne:

- **Poloha.** Vzhľad tváre na obraze sa mení v závislosti od vzájomnej polohy kamery a tváre, čo môže spôsobiť čiastočné alebo úplné ukrytie či prekrytie jednotlivých črt tváre (oči, pery, nozdry...).
- **Prítomnosť alebo neprítomnosť štruktúrnych komponentov.** Niektoré komponenty tváre, ako sú fúzy, brada či okuliare môžu, ale nemusia byť obsiahnuté v oblasti tváre. To vedie k ďalším variabilitám, ktoré sťažujú detekciu.
- **Výraz tváre.** Vzhľad tváre je do značnej miery ovplyvnený prejavmi nálady, či pocitov.
- **Prekrývanie.** Tváre môžu byť čiastočne zakryté inými objektmi, alebo dokonca prekryté inými tvármi.
- **Orientácia obrazu.** Obraz tváre sa takisto mení v závislosti od rôznej rotácie okolo optickej osi kamery.
- **Vonkajšie podmienky.** Pri snímaní obrazu participujú rôzne vonkajšie činitele ako sú osvetlenie a rôzne charakteristiky kamery, čo do veľkej miery ovplyvňuje vnímanie obrazu a aj jeho spracovanie.

V tejto oblasti je viacero úzko súvisiacich úloh. Vychádzajúc zo Sandeep, Rajagopalan (2002) a Zhang and Zhang (2010) môžeme hovoriť o nasledovných úlohách: Lokalizácia tváre (*Face localization*) zahŕňa určenie pozície jednej tváre (Dasan et. al, 2015). Je to zjednodušená detekcia s predpokladom, že na vstupnom obraze sa nachádza iba jedna ľudská tvár. Detekcia tvárových črt (*Facial features detection*) má za cieľ zistiť prítomnosť a pozíciu jednotlivých črt ako sú oči, nos, nosné dierky, ústa, pery, uši atď. za predpokladu, že na obraze sa nachádza iba jedna ľudská tvár (Zou and de With, 2008), (Hassan and Hossain, 2014). Rozpoznávanie alebo identifikácia tváre (*Face recognition or identification*) väčšinou porovnáva vstupný obraz s databázou a informuje o zhode či nezhode (Vezhnevets and Sazonov, 2003, Oravec, Rozinaj and Beszédeš, 2008). Účelom úlohy nazývanej autentifikácia tváre (*Face authentication*) je overiť identitu jednotlivca vo vstupnom obraze, kým metódy na sledovanie tváre (*Face tracking*) stanovujú polohu a možnú orientáciu tváre v obrazovej sekvencii v reálnom čase (Dey, 2016). Rozpoznávanie výrazu tváre (*Facial expression recognition*) sa zaujíma o identifikáciu citových stavov človeka (šťastie, smútok, znechutenie, atď.) (Pantic and Rothkrantz, 2000).

## 2.1 Histogramové metódy

Histogram v prípade spracovania obrazu má najčastejšie formu tabuľky, ktorá obsahuje v sebe údaje o početnom zastúpení jednotlivých hodnôt vybranej farebnej zložky v rámci množiny všetkých obrazových bodov. Napríklad pri sledovaní početnosti jednej zo zložiek R, G, alebo B vo farebnom obraze s rozmermi  $m \times n$ , kde každá zo zložiek je reprezentovaná hodnotou od 0 do 255, bude histogram obsahovať 256 dvojíc čísel. Prvé číslo vo dvojici bude udávať príslušnú hodnotu farebného komponentu (0-255) a druhý údaj bude reprezentovať jeho početnosť v obraze (od 0 po maximálne  $M \times N$ ). V grafickom vyobrazení má potom histogram formu stĺpcového grafu, v ktorom je viditeľné rozloženie jednotlivých hodnôt príslušného farebného komponentu. Ak hodnoty početností v histograme podelíme celkovým počtom obrazových bodov, dostaneme odhady pravdepodobností, s akou sa príslušná hodnota farebného komponentu vyskytuje v obraze. Jednoduchý histogramovo orientovaný prístup nachádzame v práci Dahal et. al, 2016

Ak vytvoríme histogramový model, ktorý bude sledovať početnosť len v rámci oblasti tváre v obraze, získame tak odhad pravdepodobnosti, s akou daná hodnota farebného komponentu patrí do oblasti tváre. Klasifikátor tak vykonáva triedenie bodov na body s farbou pokožky a na body pozadia a to na základe vopred stanoveného prahu pre daný odhad pravdepodobnosti. Prístup je popísaný v práci Chai and Ngan (1999). Z manuálne extrahovaných oblastí tváre z obrazov v tréningovom súbore vznikol jeden histogram pre každú zo zložiek  $C_r$  a  $C_b$ . Histogram udáva početnosť hodnôt jednotlivých komponentov v celom tréningovom súbore. Podobný prístup využívajúci cloudové aplikačné riešenie nachádzame v Praseetha and Vadivel (2016)

## 2.2 Kombinácia Gaussových pravdepodobnostných funkcií

Veľké množstvo prác používa pre klasifikáciu farby pokožky zmes Gaussových pravdepodobnostných funkcií (MOG – Mixture of Gaussians, GMM – Gaussian Mixture Models) na určenie pravdepodobnosti výskytu farby pokožky na obraze. Základnou výhodou tohto prístupu je schopnosť takéhoto modelu generalizovať pravdepodobnosť výskytu farby pokožky už z malého množstva tréningových dát (Hamouz et. al, 2004, Zhu and Fujimora, 2003).

V Jones a Regh (2002) uvádzajú autori porovnanie výkonnosti klasifikátora pracujúceho na základe histogramu a na základe kombinácie Gaussových rozdelení pravdepodobnosti. Funkcia hustoty zmesi Gaussových rozdelení je vyjadrená ako suma jadier Gaussových pravdepodobnostných funkcií. Modely pre farbu pokožky a farbu pozadia sú tréňované oddelene. V každom modeli bolo použitých 16 Gaussových pravdepodobnostných funkcií a boli dosiahnuté pomocou paralelnej implementácie štandardného EM (Expectation Maximization) algoritmu (Dempster, Laird and Rubin (1977).

Porovnanie rôznych prístupov z hľadiska výpočtovej náročnosti a potrebného dátového priestoru ukazuje jednu z nevýhod týchto metód. Získanie lineárnej kombinácie Gaussovských modelov trvá rádovo dlhšie. Naproti tomu získanie histogramového modelu sa z časového hľadiska omnoho rýchlejšie. Podobne značné rozdiely badať aj pri testovaní oboch modelov, teda pri klasifikácii obrazových bodov. Pri výpočte pravdepodobnosti výskytu hodnoty určitej farebnej zložky musíme pri zmesovom modeli počítať so všetkými použitými Gausiánmi, čo predstavuje pomerne veľkú výpočtovú náročnosť.

## 2.3 Bayesov klasifikátor

Vezhnevets et. al (2003) uvádzajú základné princípy tohto prístupu. Využívajú sa tu informácie získané z histogramového farebného modelu. Hodnota  $P_{skin}(c)$  vypočítaná podľa (1) je podmienená pravdepodobnosť  $P(c/skin)$  – pravdepodobnosť, že pozorovaná farba  $c$  je bod, ktorý patrí do oblasti tváre. Parameter  $skin[c]$  označuje početnosť bodov s farbou  $c$  a  $Norm$  je označenie pre faktor, ktorý použijeme pri normovaní, čím získame výslednú pravdepodobnosť výskytu určitej farby. Najčastejšie je parametrom  $Norm$  maximálna hodnota v histograme.

$$P_{skin}(c) = \frac{skin[c]}{Norm} \quad (1)$$

Prilievavejšie meradlo, ktoré je omnoho výhodnejšie použiť pri detekcii tváre, je pravdepodobnosť  $P(\text{skin}|c)$  – pravdepodobnosť javu, že určitý bod je bodom pokožky, ak je jeho farba  $c$ . A tu sa dostávame ku kľúčovému pravidlu, ktoré dalo názov celému prístupu a je ním Bayessovo pravidlo. V našom prípade bude mať tvar (2).

$$P(\text{skin}|c) = \frac{P(c|\text{skin}) \cdot P(\text{skin})}{P(c|\text{skin}) \cdot P(\text{skin}) + P(c|\neg\text{skin}) \cdot P(\neg\text{skin})} \quad (2)$$

$P(c|\text{skin})$  a  $P(c|\neg\text{skin})$  sú priamo vypočítané z histogramov pre farbu pokožky a pozadia. Apriórne pravdepodobnosti  $P(\text{skin})$  a  $P(\neg\text{skin})$  môžu byť takisto odhadnuté z celého množstva vzoriek s farbou pokožky a pozadia v tréningovej množine. Nerovnosť  $P(\text{skin}|c) \geq \Theta$ , kde  $\Theta$  je prahová hodnota, môže byť použitá ako pravidlo pri detekcii farby pokožky. Opäť sa dostávame ku ROC krivkám, ktoré ukazujú vzťah medzi správnymi a nesprávnymi detekciami pre dané klasifikačné pravidlo ako funkciu detekčného prahu. ROC krivka pre detekčné pravidlo  $P(\text{skin}|c) \geq \Theta$  je invariantná vzhľadom na voľbu apriórnych pravdepodobností a to kvôli povahe Bayessoého modelu. To znamená, že hodnotu  $P(\text{skin})$  ovplyvňuje iba voľba prahu. Samotný výpočet (4) nie je explicitne potrebný, ale zaujíma nás iba vzájomný vzťah medzi  $P(c|\text{skin})$  a  $P(c|\neg\text{skin})$  a nie ich presné hodnoty

Porovnaním (2) a rozhodovacieho pravidla  $P(\text{skin}|c) \geq \Theta$  dostaneme nové rozhodovacie pravidlo. Po úpravách dostaneme ako výsledok vzťah (3).

$$\frac{P(c|\text{skin})}{P(c|\neg\text{skin})} > K \cdot \frac{1 - P(\text{skin})}{P(\text{skin})} \quad (3)$$

Z uvedeného vyplýva, že voľba apriórnych pravdepodobností neovplyvňuje správanie detektora – pre akúkoľvek apriórnu pravdepodobnosť  $P(\text{skin})$  je možné vybrať vhodnú hodnotu  $K$ , ktorá nám dá rovnaký detekčný prah.

Ak sa pozrieme na pravdepodobnostné metódy detekcie tváří a chceme zhrnúť ich vlastnosti a výhody ich použitia, musíme ako prvú spomenúť schopnosť generalizácie. Pravdepodobnostné modely majú schopnosť získať z malého počtu vzoriek v tréningovom súbore pomerne veľké množstvo informácií o distribúcii farby pokožky v určitom farebnom modeli. Nevýhodou týchto metód je často ich výpočtová náročnosť, kde na výpočet výsledných pravdepodobností musíme počas samotnej detekcie vynaložiť veľa hardvérových ale aj časových prostriedkov.

## 2.4 Markovove náhodné polia

Vybrali sme prácu Dass a Jain (2002). Algoritmus je navrhnutý pre šedoúrovňové obrazy a dané MRF je navrhnuté pre veľkosť vstupného okna  $20 \times 15$  obrazových bodov. Kvôli jednoduchosti je použitých iba 16 úrovní šedej. Ako tréningový súbor bolo použitých 2000 obrazov tváří aj obrazov, ktoré tváre neobsahovali. Obrazové okno je reprezentované ako množina bodov  $s$ , ktoré nadobúdajú hodnoty šedej  $x_s$ . Na odhad pravdepodobnostných pomerov v 8-susedstve obrazového bodu je použitý vzťah (4)

$$p(x_s|x_{-s}) = \frac{\exp\{H(x_s|x_{-s})\}}{\sum_{x_s=0}^{L-1} \exp\{H(x_s|x_{-s})\}} \quad (4)$$

pričom

$$H(x_s|x_{-s}) = \alpha \cdot x_s + \sum_{t \in N_s} \beta_{st} \cdot x_s \cdot s_t$$



kde  $x_s$  je úroveň šedej aktuálne posudzovaného bodu,  $x_s$  sú úrovne šedej všetkých bodov v obraze okrem daného bodu a  $x_t$  je bod v susedstve daného bodu. Parametre  $\alpha$  a  $\beta$  sa získavajú z tréningovej množiny. Spomínané LPL pravidlo, ktoré sa uplatňuje pri tréningu modelu, má tvar (5).

$$LPL(\alpha, \beta) = \sum_{s=1}^N \sum_{s=1}^{\#S} \left( \alpha_s \cdot x_s^{(k)} + \sum_{t \in N_s} \beta_{st} \cdot x_s^{(k)} \cdot x_t^{(k)} - \log \left( \sum_{x_s=0}^{L-1} \exp \left[ a_s \cdot x_s + \sum_{t \in N_s} \beta_{st} \cdot x_s \cdot x_t^{(k)} \right] \right) \right) \quad (5)$$

kde  $\alpha = \{\alpha_s, s \in S\}$ ,  $\beta = \{\beta_{st}, s \in S, t \in N_s\}$ ,  $x^{(k)} = (x_1^{(k)}, x_2^{(k)}, \dots, x_M^{(k)})$  sú pozorované odtiene šedej  $k$ -teho tréningového obrazu, pričom  $k = 1, 2, \dots, N$ . Množina  $S$  je množina všetkých obrazových bodov a jej prvky sú označované ako  $s$ . Množina  $N_s$  je množina bodov susedstva a  $t$  sú jej prvky. Premenná  $N$  je počet tréningových obrazov. Konštanta  $L$  určuje počet úrovní šedej, ktoré sú použité v obraze. Číslo  $\#S$  reprezentuje celkový počet obrazových bodov jedného obrazu.

Výpočtom upravujeme parametre modelu  $\alpha$  a  $\beta$ . Následne po tréningu prebieha klasifikácia, ktorá sa riadi odlišným pravidlom LPR, podľa ktorého je zostrojená klasifikačná podmienka. Ak je podmienka splnená, obraz je klasifikovaný ako obraz tváre, v opačnom prípade je klasifikovaný ako obraz, ktorý neobsahuje tvár.

## 2.5 Metódy s využitím neurónových sietí a systémov strojového učenia

Jeden z najznámejších je Rowleyho systém (Rowley et. al (1998)), ktorý je použitý na detekciu čelných priamych zobrazení tvárí v šedo-tónových obrazoch. Algoritmus pracuje na princípe jednej alebo viacerých neurónových sietí aplikovaných priamo na časti vstupného obrazu a následnom použití rozhodovacieho bloku, ktorý na základe výstupov NN rozhoduje, či sa v obraze tvár nachádza, alebo nie. Jeden z novších prístupov nachádzame v Kumar, Kar a Cahndra (2014). Tento prístup využíva podporné vektorové stroje (Support Vector Machine) na detekciu tvárí v zašumených obrazoch.

Pri nasadení v úlohách detekcie tváre našli veľké uplatnenie aj Kohonenove samoorganizujúce sa mapy (Self-Organizing Map - SOM). Systém navrhnutý Kohonenom v osemdesiatych rokoch je jeden z najpopulárnejších prístupov používaných pre umelé neurónové siete. V Brown et al. (2000) prinášajú autori návrh detektora farby pokožky založený práve na princípe SOM. Metóda využíva dve SOM, pričom jedna z nich je natrénovaná na množine obsahujúcej iba tváre a druhá na obrazoch, ktoré tváre vôbec neobsahujú. Trénovanie prebehlo na súbore takmer 500 manuálne označených obrazov. Výkonnosť detektora bola testovaná na testovacom súbore, ktorý vytvorili autori. Ten istý súbor bol použitý aj na testovanie a pre lepšiu validáciu výsledkov bol na testovanie použitý aj súbor Compaq Skin Database (Jones a Regh (2002)). Na tejto metóde bolo na testovanie použitých niekoľko farebných priestorov a výsledkom bolo, že detektor nevykazoval viditeľné zmeny výkonnosti pre rôzne farebné priestory. Výkonnosť na súbore dát, ktorý vytvorili autori bola o čosi lepšia, ako pre zmesový model Gaussiánov, zatiaľ čo na Compaq databáze bola výkonnosť SOM horšia oproti RGB histogramovej metóde.

Uvádzame aj viacero novších prístupov. AdaBoost algoritmus (Adaptive Boosting) je pomerne nový prístup na vytváranie klasifikátorov, ktoré sú schopné klasifikovať aj nelineárne separovateľné množiny. Na jednej strane je AdaBoost algoritmus citlivý na tzv. šum vo vstupných dátach, no na strane druhej je v porovnaní s inými algoritmami nezvyčajne odolný voči pretrénovaniu. Príklad použitia pri detekcii tvárí nájdeme v (Shuang, 2014 a Liuliu and Mianyang, 2016) kde sa Adaboost využíva v spolupráci s PCA. Okrem toho sa na tvorbu klasifikátorov v súčasnosti používa PAWR algoritmus (pocket algorithm with ratchet) uvedený v Gallant (1990). Tento algoritmus vychádza z pôvodného algoritmu tréningu pre doprednú neurónovú sieť – perceptrón. Pocket algorithm with ratchet (označovaný ako "Pocket" perceptrón algoritmus) vychádza z pôvodného algoritmu tréningu pre doprednú neurónovú sieť typu perceptrón. "IG" algoritmus (Information Gain) – názov by sa dal preložiť aj ako algoritmus informačného zisku – sa používa pri vytváraní klasifikátorov nazvaných rozhodovacie stromy.

Ak chceme porovnať vlastnosti NN v aplikácii na detekciu tvárí v obraze, musíme v prvom rade spomenúť ich výkonnosť a presnosť. Kvôli pomerne vysokej náročnosti implementácie sa používajú hlavne

v oblasti zložitejších systémov, ale aj tu existujú jednoduchšie alternatívy, ktoré sa dajú nasadiť v akýchkoľvek systémoch.

## 2.6 Gaborove filtre

Štruktúra a tvar týchto filtrov sú odvodené od činnosti ľudského vnímania tak, ako ho vykonávajú bunky vo vizuálnom kortexe ľudského mozgu. Mozog venuje špeciálnu pozornosť práve orientácii a priestorovej frekvencii (Huang, Shimizu a Bobatake, 2005).

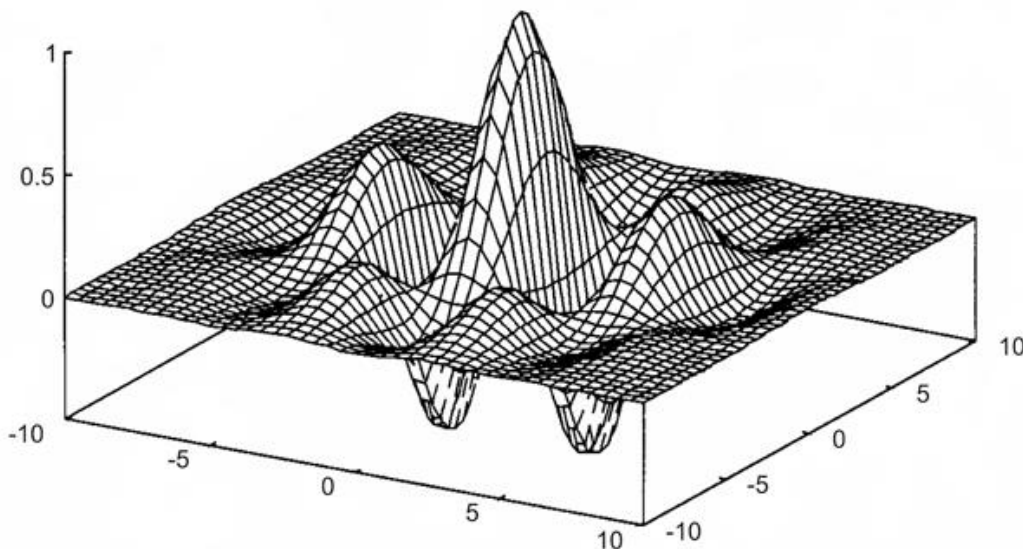
Hoci teoretický základ týchto priestorových filtrov siaha až do päťdesiatych rokov minulého storočia, je stále k dispozícii viacero nových prístupov, ktoré využívajú Gaborove filtre ako nástroj na extrakciu informácií o priestorovom a frekvenčnom rozložení obrazového signálu. Vo väčšine aplikácií sú Gaborove filtre použité v kombinácii s neurónovými sieťami. Všeobecný tvar Gaborovej funkcie  $g(x,y)$  je uvedený v (6).

$$g(x, y) = s(x, y) \cdot w_r(x, y) \quad (6)$$

Funkcia  $s(x,y)$  sa nazýva nosná a má tvar komplexnej sinusoidálnej funkcie podľa vzťahu (7). Druhá funkcia  $w_r(x,y)$  sa nazýva obálka a má tvar Gaussovej funkcie podľa vzťahu (8). Všeobecná Gaborova funkcia je vyobrazená na Obr. 1.

$$s(x, y) = \exp(j(2 \cdot \pi \cdot (u_0 \cdot x + v_0 \cdot y) + P)) \quad (7)$$

$$w_r(x, y) = K \cdot \exp(-\pi \cdot (a^2 \cdot (x - x_0)_r^2 + b^2 \cdot (y - y_0)_r^2)) \quad (8)$$



Obr. 1 Príklad Gaborovej funkcie.

Pre našu oblasť záujmu sú Gaborove filtre používané na detekciu tvárí a ich rozpoznávanie. Nasadenie Gaborových filtrov sa zvyčajne kombinuje s viacvrstvovým perceptrónom a algoritmi ako je AdaBoost a pod. Ďalej sa bližšie zameriame na detekciu tvárí. Pre lepšiu ilustráciu popíšeme konkrétny algoritmus. Huang, Shimizu a Kobatake (2005) použili banku Gaborových filtrov s polynomiálnou neurónovou sieťou. Ako metóda bola zvolená analýza hlavných komponentov (PCA – principal component analysis).

V spomínanom prístupe sa na detekciu príznakov používa metóda tzv. plávajúceho okna. Okno postupne prechádza cez celý obraz a vstupuje do konvolúcie s Gaborovými filtermi. Tým dostávame tzv. Gaborovu reprezentáciu obrazu. Z týchto odpovedí sú extrahované tri typy vektorov príznakov: reálna zložka, magnitúda a magnitúda+fáza. Tieto vektory tvoria vstup do polynomiálnej neurónovej siete (PNN), ktorá na

výstupe označí podobnosť aktuálneho plávajúceho okna s tvárovou, resp. netvárovou oblasťou. Spomínaná sieť je jednovrstvová, takže potrebuje oveľa menší čas na učenie ako viacvrstvový perceptrón.

Keďže vektory príznakov, ktoré sme opísali, majú príliš veľké rozmery, čo vedie k značnej veľkosti samotnej PNN siete, musíme dané vektory rozmerovo zredukovať. Využíva sa na to analýza základných komponentov, ktorej výsledkom sú tzv. vlastné vektory (eigenvectors).

Použitá PNN sieť je trénovaná na tvárových a netvárových vzorkách. Inicializácia váh je náhodná, ale keďže ide o jednovrstvovú sieť, nemá inicializácia váh vplyv na rýchlosť učenia. Sieť bola trénovaná na 2987 obrazoch tváří. Z každého obrazu je vybratá vzorka s veľkosťou 20x20 obrazových bodov. Natiahnutím a zmenšením rozmerov sú získané ďalšie 4 vzorky. Z výsledných piatich vzoriek sú ďalej urobené zrkadlové obrazy, čo umožňuje pokryť variácie, ktoré sa vyskytujú v reálnom svete. Experimentálne výsledky opísaného algoritmu boli vyhodnotené na viacerých testovacích súboroch, kde boli dosiahnuté očakávané dobré výsledky.

### **3 Ciele práce**

Na základe štúdia odbornej a vedeckej literatúry a poznatkov získaných v problematike strojového rozpoznávania ľudských tvárí sme si stanovili nasledovné ciele dizertačnej práce:

#### **3.1 Farebná analýza**

Analýza farby pokožky v priestore farieb, ktorý je založený na oddelení luminančnej zložky od chrominančných zložiek a nájdenie vhodného modelu, ktorý bude využívať chrominančné zložky. Vychádzame zo známeho predpokladu, že jasová zložka neovplyvňuje významným spôsobom farbu. Na základe analýzy histogramov chrominančných zložiek navrhne 2D model farebného histogramu farby pokožky. Na spracovanie a kódovanie digitálneho videa je podľa odporúčania štandardu CCIR-601 pre vysielanie v Európe k dispozícii farebný priestor YCbCr, ktorý je veľmi blízky farebnému priestoru v systéme PAL (YUV).

#### **3.2 Gaborove príznaky**

Analýza a vyhodnotenie účinnosti príznakov získaných rôznymi Gaborovými filtermi pri lokalizácii jednotlivých tvárových črt a určenie vhodnej reprezentácie použiteľnej pre strojové spracovanie.

Vyhodnocovanie príznakov získaných použitím Gaborových filtrov môžeme nájsť vo viacerých prácach. My si kladieme otázku, ako vybrať optimálnu bázu Gaborových filtrov, to znamená ich počet a hodnoty ich parametrov. Štandardne sa stretávame s bankou 40 filtrov (5 frekvencií /6 orientácií), prípadne aj s inými stratégiami (4/6, 3/6) bez bližšieho zdôvodnenia výberu (Shen and Bai 2006). Naším cieľom je dosiahnuť redukciu počtu príznakov použitím minimalizovanej banky filtrov, pričom budeme aplikovať len filtre, ktoré budú z hľadiska nášho vyhodnotenia vhodné na detekciu konkrétnych črt tváre. Vytvorenie novej metódy detekcie ľudskej tváre vychádzajúcej z existujúcich metód, založenej na detekcii farby pokožky použitím vlastného farebného modelu a lokalizácii tvárových črt použitím optimálnej banky Gaborových filtrov vytvorenej na základe analýzy textúr.

#### **3.3 Výkonnostná analýza**

Vyhodnotenie navrhnutých postupov z hľadiska výkonnosti.

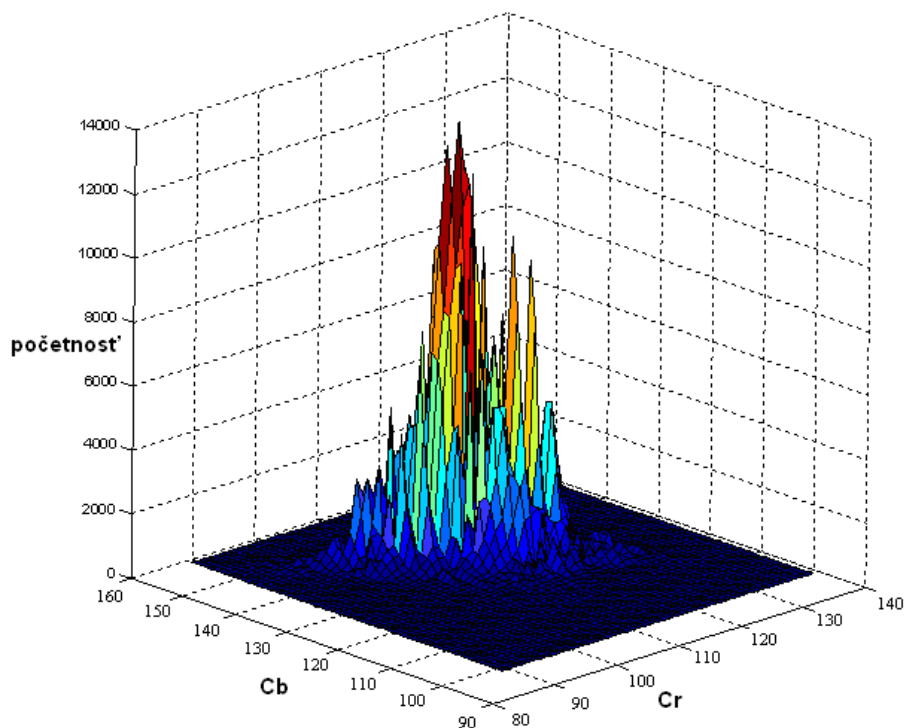
## 4 Výsledky dizertačnej práce a vykonané experimenty

### 4.1 Farebná analýza

Farebnú analýzu farby ľudskej pokožky sme vykonali vo farebnom priestore  $YCbCr$ . Tento farebný priestor poskytuje oddelenie chrominančných zložiek od jasovej. Pri analýze histogramových modelov farby ľudskej pokožky sme sa zamerali na získanie invariantných charakteristík farby pokožky, ktoré by sme mohli použiť naprieč rôznym typom farby pleti ako aj meniacich sa vonkajších podmienok, ako je napríklad osvetlenie.

Analýzou histogramových modelov pre chrominančné zložky  $Cb$  a  $Cr$  pre rôzne typy farby pleti sme spozorovali vysokú mieru podobnosti histogramových modelov pre farbu pokožky ázijského, európskeho a afrického typu pokožky. Tieto tri modely nám umožnili tvorbu zjednoteného histogramového modelu farby pokožky, ktorý poskytuje vyhovujúce diskriminačné vlastnosti pri detekcii oblastí s farbou ľudskej pokožky.

V ďalšej analýze sme vytvorili kombinovaný histogram oboch chrominančných zložiek  $Cr$  a  $Cb$ . Namiesto sledovania početnosti každej zložky osobitne, navrhli sme sledovanie početnosti kombinácie hodnôt jednotlivých zložiek. Takto navrhnutý farebný model poskytuje zlepšenie diskriminačných schopností získaného farebného modelu pri detekcii tváří. Dosiahnuté výsledky boli publikované v práci Beniak, Pavlovičová a Oravec (2008). Výsledkom je 2D histogram (Obr. 2), na základe ktorého potom prebieha klasifikácia jednotlivých obrazových bodov. Výsledky porovnania diskriminačných schopností s konkurenčným modelom uvádzame v Tab. 1.



**Obr. 2 3D reprezentácia kombinovaného histogramu reprezentujúca distribúciu farebných zložiek  $Cb$  a  $Cr$  v oblasti tváre.**

Použitie takéhoto modelu pre farbu ľudskej pokožky vedie ku zlepšeniu účinnosti celého systému. Na testovacím súbore 100 obrazov z databázy FERET dosiahol algoritmus vyššiu úspešnosť v zmysle menšieho počtu falošných detekcií. Porovnanie tohto parametra pre jednotlivé typy tváří uvádzame v Tab. 1.

Tab. 1 Porovnanie výkonnosti algoritmov – počet falošných detekcií v troch podmnožinách testovacieho súboru.

		2D model farby pokožky (Beniak et. al 2011)		Oddelené modely pre Cb a Cr (Oravec et. al, 2008)	
Podmnožina	Veľkosť súboru	FP (počet)	FP (%)	FP (počet)	FP (%)
African	14	3	17,65	9	64,3
Asian	7	0	0	2	28,57
European	79	10	12,66	12	15,19
All	100	13	13	6	23

## 4.2 Gaborove príznaky

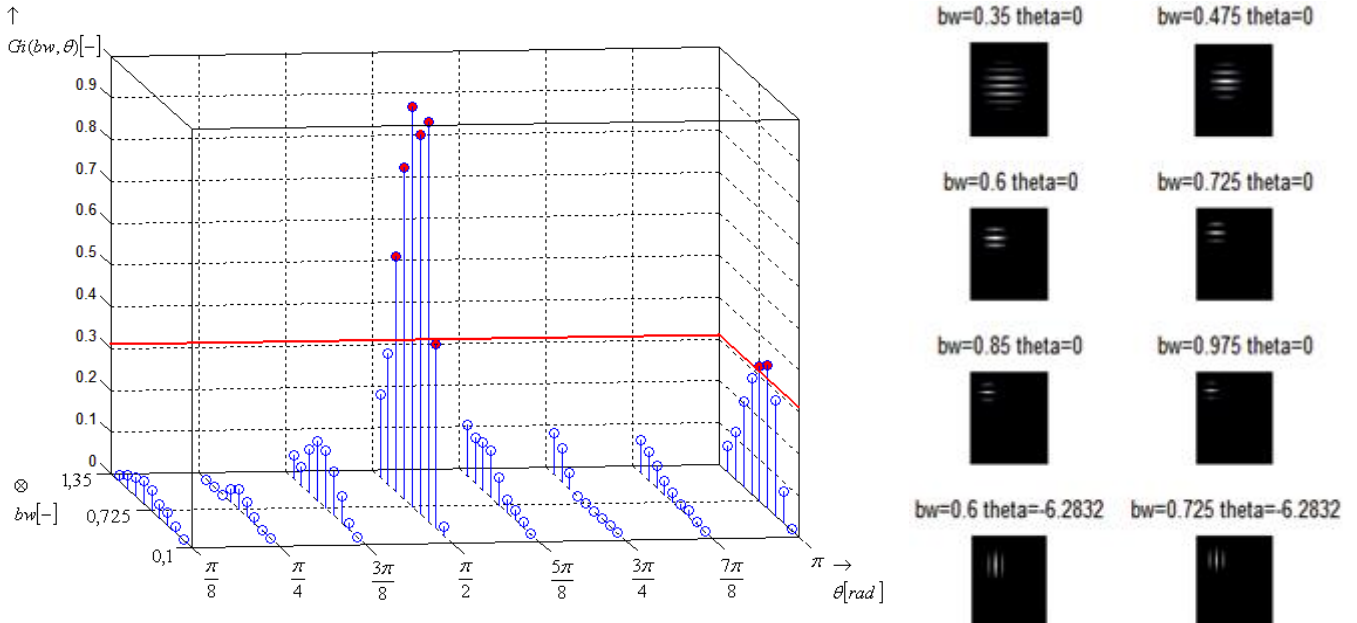
V analýze Gaborových príznakov a ich účinnosti sme sa sústredili na problematiku výberu optimálnej banky filtrov na lokalizáciu tvárových črt. Do optimálnej banky filtrov boli vybraté iba také filtre, ktoré poskytnú relevantné dáta pre dané použitie a veľkosť banky filtrov. Analýzu sme vykonali na základe štatistických veličín odpovedí jednotlivých filtrov (smerodajná odchýlka a energia). Z nich odvodený parameter nárastu informačného potenciálu  $G_i(bw, \theta)$  (Informational Potential Growth Parameter) sme použili ako diskriminačné kritérium pri výbere filtra. Parameter je definovaný podľa nasledujúceho vzťahu (9):

$$G_i(bw, \theta) = D_n(bw, \theta) - D_{n\_min}(bw, \theta) \quad (9),$$

$$\text{kde } D_n(bw, \theta) = \sqrt{\sigma_n^2(bw, \theta) + e_n^2(bw, \theta)}.$$

Veličina  $\sigma_n^2(bw, \theta)$  je rozptyl a  $e(bw, \theta)$  je energia signálu, ktorý reprezentuje odpoveď Gaborovho filtra so šírkou pásma  $bw$  a priestorovou orientáciou  $\theta$ .

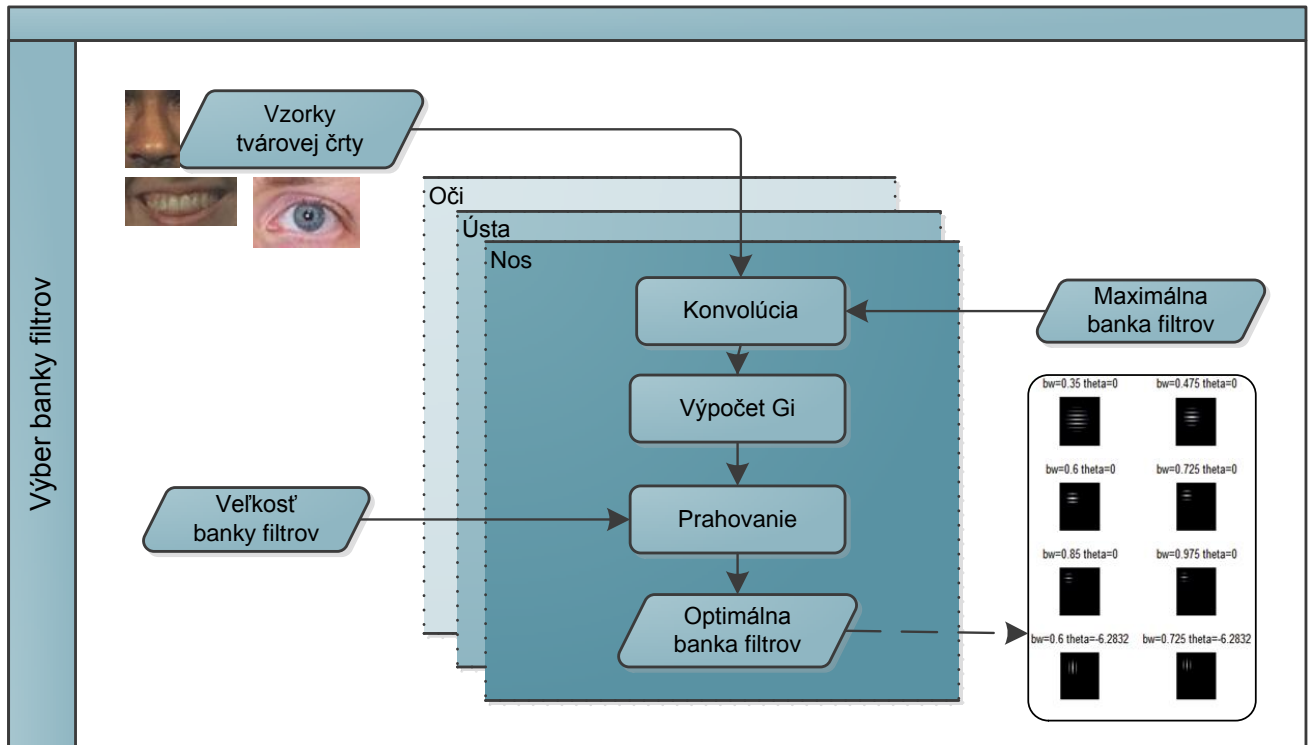
Do banky filtrov sú vyberané najlepšie filtre podľa parametra  $G_i(bw, \theta)$ . Výberový prah je určený na základe požadovanej veľkosti banky filtrov. Ilustráciu tohto postupu uvádzame na Obr. 3.



Obr. 3 Výber banky filtrov s veľkosťou 8 (Filter Bank 8) a jemu prislúchajúca prahová hodnota (vyznačená červenou). Vybraná banka filtrov je uvedená vpravo. Banka filtrov poslúži na lokalizáciu očí.

Banku filtrov, ktorú sme takýmto spôsobom vybrali, použijeme v algoritme na detekciu tváří, na ktorej sú lokalizované tvárové črty. Algoritmus využíva Gaborove príznaky, z ktorých získa štatistickou analýzou veľké množstvo príznakov. Tieto príznaky sú ďalej spracované viacvrstvom perceptrónom, ktorý na základe naučených poznatkov dokáže vyhodnotiť, či daná vzorka obsahuje tvárovú črtu, alebo nie. Takýmto spôsobom je skenovaná kandidátska oblasť tváre a vyhodnocuje sa prítomnosť hlavných tvárových črt (oči, ústa, nos). Ak kandidátska oblasť tváre obsahuje minimálne dve tvárové črty, je táto oblasť validovaná ako oblasť tváre. Takýmto spôsobom dokážeme zlepšiť výkonnosť detektora využívajúceho farebnú analýzu.

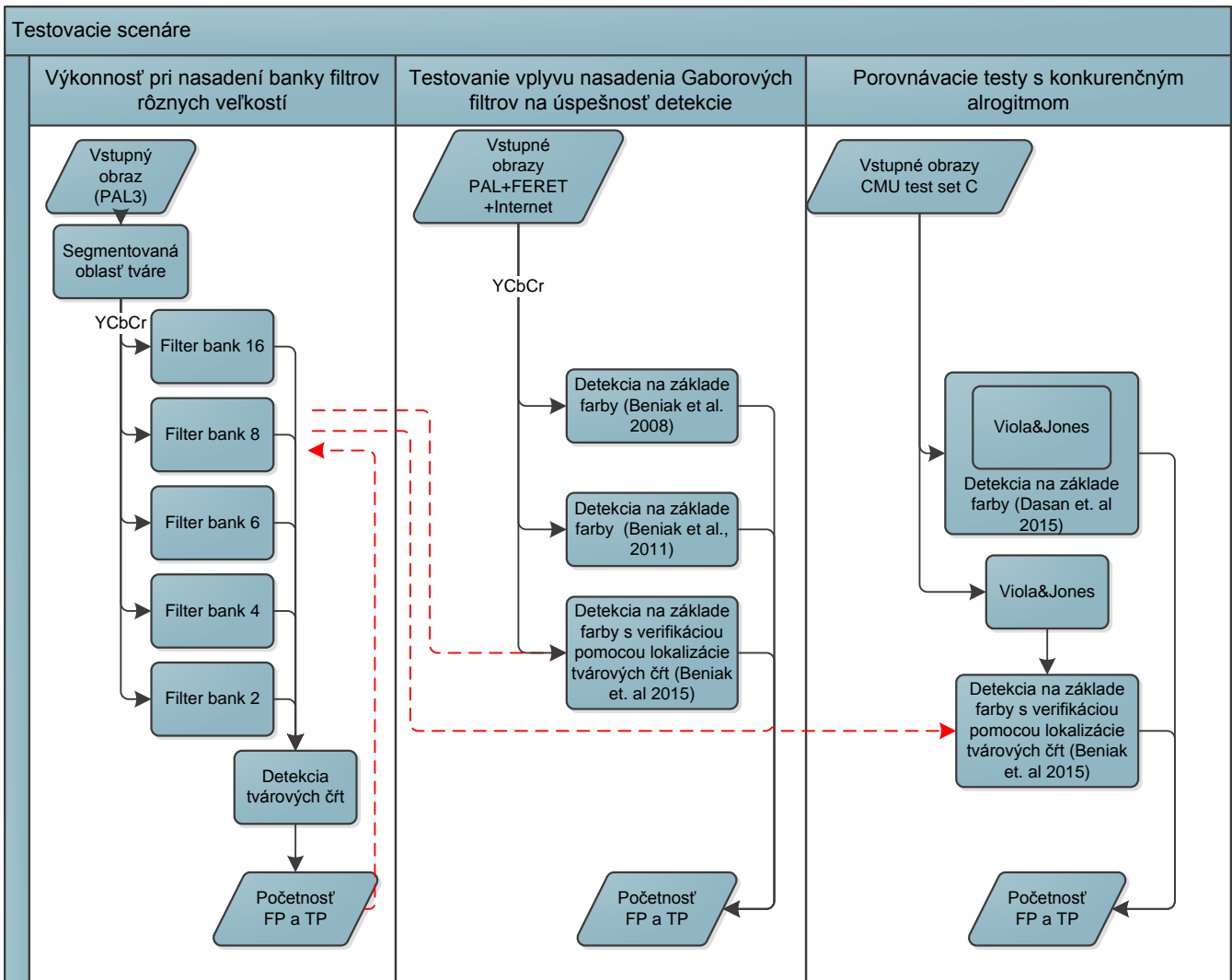
Hore uvedený postup validácie kandidátskej oblasti uvádzame v grafickej podobe na Obr. 4.



Obr. 4 Tok dát pre proces výberu optimálnej banky Gaborových filtrov

### 4.3 Výkonnosťná analýza

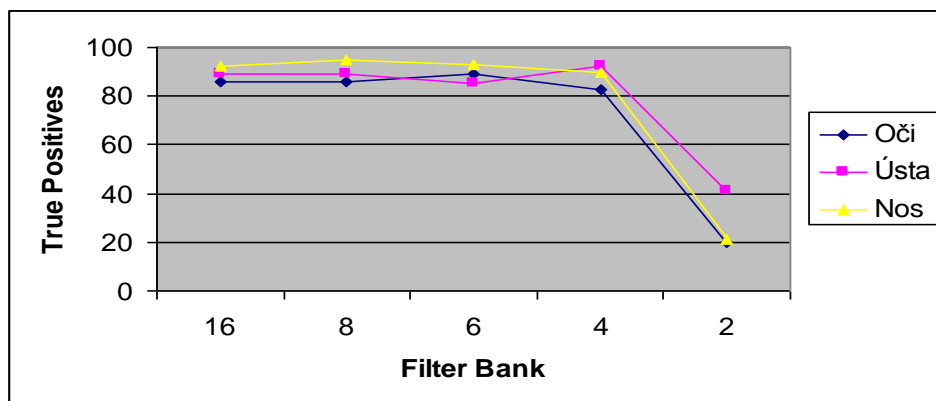
Naše testovanie malo tri časti. Najprv sme podrobili testovaniu výkonnosti navrhovaný algoritmus s použitím banky filtrov rôznych veľkostí. Banku filtrov s najlepšimi výsledkami sme potom nasadili v testoch výkonnosti spolu s inými algoritmami nevyužívajúcimi Gaborove filtre. Tretí stupeň testovania spočíval v porovnávacích testoch výkonnosti s algoritmom využívajúcim Gaborove filtre a ktorého architektúra je porovnateľná s našim návrhom. Prehľad navrhnutých testovacích scenárov uvádzame na Obr. 5.



Obr. 5 Prehľad testovacích scenárov

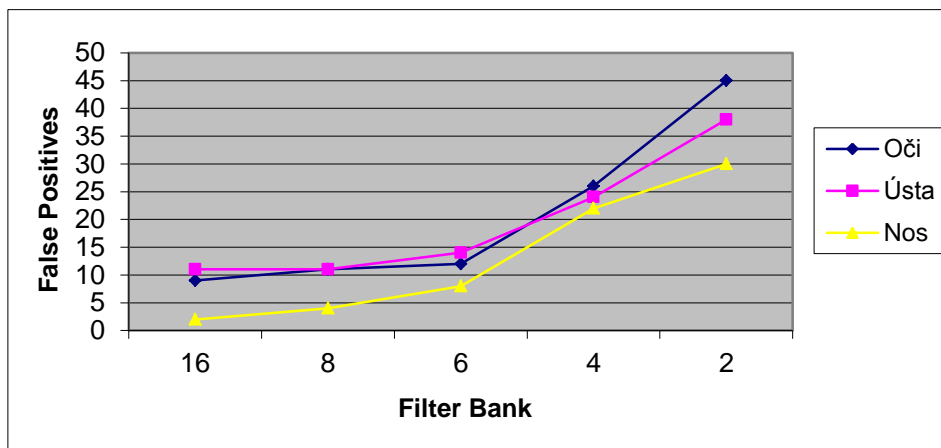
Výsledky testovania nasadenia banky filtrov rôznych veľkostí vyjadríme graficky osobitne pre True Positives (Obr. 6) aj False Positives (Obr. 7).

Z grafov vidíme, že banka s veľkosťou 6 filtrov vykazovala nepochybne dobré diskriminačné schopnosti a v porovnaní s bankou veľkosti 8 sledujeme iba mierny nárast False Positives (maximálne 4% v prípade falošných detekcií nosa). Pre ďalšie testovanie sme vybrali banku filtrov s veľkosťou 8 (Filter Bank 8).



Obr. 6 True Positives pre banky filtrov s veľkosťou 16, 8, 6, 4 a 2 filtre.





Obr. 7 False Positives pre banky filtrov s veľkosťou 16, 8, 6, 4 a 2 filtre.

Testovanie výkonnosti nasadenia Gaborových filtrov na verifikácia kandidátskych oblastí prebiehalo na kombinovanej databáze zloženej z databázy PAL, FERET a z Internetového vyhľadávania. Z tabuľky sumarizujúcej výsledky vyplýva, že nami navrhovaný algoritmus ponúka výrazné zmenšenie počtu falošných detekcií (Tab. 2).

Tab. 2 Porovnanie výkonnosti testovaných metód.

Parameter	Metóda		
	Beniak et al. (2008)	Beniak et al. (2011)	Mod. alg. verifikovaný detekciou tvárových črt
Test set	100	100	100
True Positives	82	98	98
False Positives	11	23	6

Testovanie výkonnosti s algoritmom (Dasan et al. (2015), ktorý má porovnateľnú štruktúru, budeme vykonávať na databáze obrazov CMU. Keďže algoritmy používajú na vstupe odlišný model farebnej reprezentácie, ich vzájomné porovnanie bude možné iba po eliminácii vplyvu spracovania na celkové výsledky. Za týmto účelom vyberieme kandidátske regióny z databázy CMU pomocou algoritmu Viola&Jones (Viola and Jones, 2004). Takto vybrané kandidátske oblasti podrobíme validácii pomocou detekcie tvárových črt ako sme to navrhli vyššie.

Porovnanie True Positives and False Positives parametrov uvádzame v tabuľke Tab. 3.

Tab. 3 Porovnanie výkonnosti testovaných metód.

Parameter	Metóda	
	Dasan et al. (2015)	Mod. alg. verifikovaný detekciou tvárových črt
Detection Rate %	90,3	89,8
True Positives #	180	180
False Positives #	3	4

Z uvedeného porovnania vidíme, že nami navrhovaný postup dosahuje veľmi dobré výsledky v rámci parametra Detection Rate a vyznačuje sa iba mierne vyšším počtom falošných detekcií (FP). Nami navrhovaný algoritmus teda dosahuje výkonnostné parametre porovnateľné s algoritmom publikovaným v Dasan et al. (2015). Takýto výsledok je pozoruhodný, nakoľko algoritmus Dasan et. al (2015) používa banku filtrov veľkosti 40 filtrov. Nami navrhovaný postup teda priniesol porovnateľné výkonnostné parametre použitím banky filtrov veľkosti 8. Z výpočtového hľadiska sme teda zredukovali počet spracovávaných príznakov o 80% pri porovnateľných výkonnostných parametroch. Takáto optimalizácia algoritmu predstavuje nepochybný prínos.

## 5 Sumarizácia prínosov práce

### 5.1 Prínosy v teoretickej oblasti

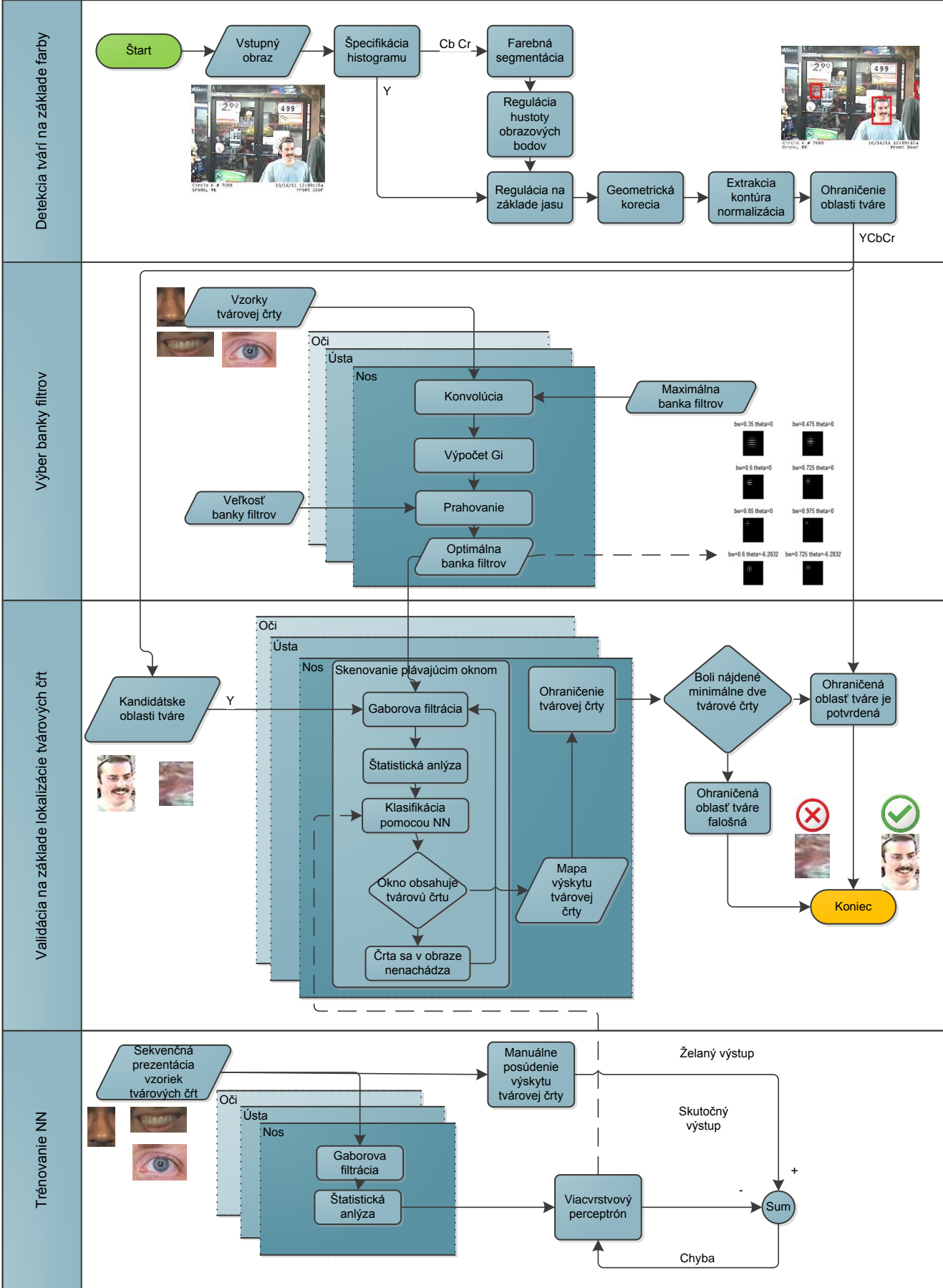
- Teoretické poznatky zhromaždené v tejto práci ponúkajú ucelený pohľad na problematiku spracovania obrazov ľudských tvári za účelom detekcie tvári. Nielen prehľad používaných metód, ale hlavne predstavenie základných nástrojov na spracovanie obrazov tvári poskytuje solídnu základňu pre vedecké bádanie v tejto oblasti.
- Pri detekcii farby pokožky sme navrhli spôsob kombinovaného vyhodnocovania jednotlivých chrominančných zložiek pomocou 2D histogramu. Analýza poskytuje nástroj pre klasifikáciu jednotlivých obrazových bodov na základe kombinácie hodnôt chrominančných zložiek.
- V teoretickej rovine sme navrhli postup ako efektívne vybrať optimálnu banku Gaborových filtrov pre detekciu či lokalizáciu rôznych typov textúr. Navrhnutý postup bol navrhnutý pre potreby lokalizácie tvárových črt, ale je možné ho nasadiť na detekciu ľubovoľných textúr.
- Potvrdili sme, že na opis textúr sa dajú veľmi efektívne využiť osobitne dve štatistické charakteristiky – smerodajná odchýlka a energia, ako parametre 2D signálu, ktorý reprezentuje jasovú zložku danej textúry. Navrhnutá metóda analýzy Gaborových filtrov z hľadiska ich informačného potenciálu pri detekcii konkrétnych typov textúr predstavuje originálny prínos v danej oblasti.
- Definovaním parametra nárastu informačného potenciálu (Informational Growth Parameter -  $Gi(bw, \theta)$ ) sme navrhli metriku, ako kritérium na výber filtrov do banky filtrov pre želané použitie. Metóda výberu filtrov do banky je možné aplikovať pri detekcii ľubovoľných textúr.
- Navrhnutý algoritmus pre výber optimálnej banky filtrov danej veľkosti je dizajnovaný ako úplne automatický, pričom sú potrebné tri vstupy. Prvým je maximálna banka filtrov, z ktorej budeme vyberať optimálnu banku. Druhým sú vzorky textúry na detekciu ktorej budeme optimálnu banku filtrov navrhovať. Posledným parametrom je želaná veľkosť optimálnej banky filtrov. Algoritmus na základe týchto vstupov dokáže nájsť optimálnu banku filtrov definovanú parametrami jednotlivých filtrov.

### 5.2 Prínosy v praktickej oblasti

- V praktickom pohľade ponúka práca komplexný algoritmus na detekciu tvári a lokalizáciu tvárových črt, ktorý vďaka svojej modularite dovoľuje jeho nasadenie pri širokej škále aplikácií od zarovňavania tvári do databázy až po detekciu úsmevu, hnevu či iných emocionálnych stavov.
- Aplikácia implementovaná v prostredí Matlab môže byť ďalej modifikovaná a jej jednotlivé časti môžu slúžiť pre ďalší vývoj.
- Navrhnutý postup detekcie farby pokožky pomocou 2D histogramu poskytuje zlepšenie výkonnostných parametrov detektora oproti použitiu nezávislých histogramov pre jednotlivé chrominančné zložky.
- V navrhovanej podobe je proces výberu optimálnej banky filtrov možné realizovať ako úplne automatický. Jedinou interakciu obsluhy je poskytnutie tréningových vzoriek danej textúry.
- Navrhnutým konceptom detekcie tvári a jej verifikácie na základe detekcie tvárových črt sme zlepšili výkonnostné parametre existujúceho algoritmu (Beniak et al. 2011).

Zhrnutie publikovaných prínosov a ich vzájomné prepojenie sumarizuje Obr. 8.

Komplexný systém na detekciu tváří v obrazoch



Obr. 8 Prehľad navrhnutých riešení a ich vzájomných vzťahov

## 6 Zhodnotenie

Cieľom našej práce bolo poskytnúť prehľad súčasného stavu problematiky detekcie tváří a s tým súvisiacej problematiky detekcie tvárových črt. Z veľkého množstva metód a prístupov sme sa ďalej zamerali na detekciu tváre na základe farby pokožky a následné využitie Gaborových filtrov na detekciu tvárových črt. Gaborove filtre sa dajú použiť na detekciu všeobecných textúr, ale rovnako dobre poslúžia na detekciu tváří a špecificky tvárových črt, ktoré sa vyznačujú osobitou textúrou.

Kombináciou existujúcich prístupov, ich prispôbením a vylepšením sme získali komplexný systém na detekciu tváří a tvárových črt. Na dosiahnutie tohto cieľa sme využili širokú škálu analytických a experimentálnych nástrojov, ktoré nám pomohli navrhnuť optimálny systém z hľadiska výkonnostných parametrov. Spomedzi použitých nástrojov pokladáme za najdôležitejšiu štatistickú analýzu. Okrem toho sme využili systémy strojového učenia, ktoré tvoria jeden zo základných nástrojov kybernetiky.

Štatistickým vyhodnotením vlastností textúr jednotlivých tvárových črt sme vytvorili metódu na výber vhodnej banky filtrov pre konkrétnu aplikáciu. Jednotlivé štatistické parametre sme nevyhodnocovali osobitne, ale navrhli sme spôsob, ktorým sa dajú jednotlivé veličiny popisujúce vlastnosti textúr, skombinovať. Uvedená metrika v sebe kombinuje smerodajnú odchýlku a energiu signálu. Vytvorením parametra, ktorý sme nazvali parameter nárastu informačného potenciálu (Informational Potential Growth Parameter -  $Gi(bw, \theta)$ ) sme dostali metriku, na základe ktorej sa dajú veľmi jednoducho vyberať filtre do banky filtrov pre želané použitie a veľkosť banky. Metóda výberu filtrov do banky sa dá použiť nielen pre detekciu tvárových črt, ale aj pre detekciu ľubovoľných textúr.

Na základe tohto parametra  $Gi$  sme schopní vybrať do banky filtre, ktoré nám pri následnom spracovaní poskytnú čo najviac informácií o uvažovanej textúre. Takýmto spôsobom vieme výrazne znížiť počet filtrov v banke potrebných na získanie potrebného množstva informácií pre ďalšie spracovanie. Vykonané testovanie to iba potvrdzuje. Dosiahli sme uspokojivé výsledky s použitím banky filtrov veľkosti 8. V porovnaní s metódami štandardne používajúcimi banky filtrov veľkosti 40 sme tak výrazne zredukovali výpočtovú náročnosť celého procesu.

Navrhovaný postup detekcie a lokalizácie tvárových črt sme otestovali a vyhodnotili jeho výkonnosť v zmysle počtu správnych a falošných detekcií pri použití bank filtrov rôznych veľkostí.

Aby sme však dosiahli kvalitatívne zlepšenie nami navrhnutého spôsobu detekcie tváří, vytvorili sme kombinovaný prístup, ktorý ponúka vysoké hodnoty korektných detekcií ako aj veľmi malé hodnoty počtu falošných detekcií. Toto zlepšenie sme dosiahli synergickým efektom dvoch prístupov: detektor tváří na základe farby pokožky vyberie väčšinu kandidátskych regiónov, zatiaľ čo detektor tvárových črt ponúkne validáciu kandidátskych regiónov na základe prítomnosti tvárových črt.

Nárast hodnoty počtu korektných detekcií z 86% na 98% a k tomu pokles počtu falošných detekcií až na hodnotu 6% ponúka veľmi účinný nástroj na detekciu tváří v obrazoch.

Testovanie navrhnutého algoritmu a porovnanie výkonnosti s podobným prístupom (Dasan et al., 2015) prinieslo takisto sľubné výsledky. Porovnateľné hodnoty parametra Detection rate, ako aj nízke hodnoty falošných detekcií ukazujú, že nami navrhnutý algoritmus dokáže konkurovať porovnateľným prístupom. Zásadný prínos predstavuje dosiahnutie výkonnostných parametrov porovnateľných s konkurenčným algoritmom pri použití banky filtrov s veľkosťou iba 8 filtrov. To predstavuje pokles počtu spracovávaných príznakov až o 80%.

Ako vidíme z dosiahnutých výsledkov, kombinácia viacerých charakteristík obrazu tváre a synergický efekt viacerých detektorov, ktoré postupne validujú kandidátske oblasti v obraze, sa ukazuje ako cesta pre dosiahnutie lepších výsledkov detekcie. Navrhovaný postup je z hľadiska výkonnosti vyhovujúci pre daný typ obrazov. Jeho nasadenie pre ľubovoľný obrazový set si vyžaduje úvodnú normalizáciu obrazov tváre. Potrebná je normalizácia veľkosti a natočenia tváří.

Nakoľko navrhovaný detektor na základe farby pokožky je do veľkej miery robustný vzhľadom na zmenu svetelných podmienok, nie je potrebná farebná normalizácia. Obmedzenie má teda iba detektor tvárových črt, ktorý vyžaduje správne natočenie tváre v obraze (priamy vzpriamený pohľad).

Samozrejme, že navrhnutý prístup má viacero možností na zlepšenie jeho výkonnosti a všestrannosti. Prvým z nich je už spomenutá normalizácia vstupných obrazov, ktorá umožní nasadenie prístupu pre ľubovoľné množiny obrazov – obsahujúce nielen vzpriamené obrazy tváří.

Isté zlepšenie výkonnosti môže priniesť širšia množina vzoriek tvárových črt použitá pre učenie neurónovej siete. Ba čo viac, aj samotný druh použitej neurónovej siete by podľa môjho názoru dokázal zlepšiť úspešnosť lokalizácie tvárových črt. Použitý viacvrstvovým perceptrón je jednoduchý z hľadiska implementácie, avšak neponúka možnosť paralelizácie.

Na záver môžeme konštatovať, že ciele, ktoré sme si stanovili na začiatku nášho výskumu, sme splnili v plnom rozsahu. Ponúkli sme ucelený obraz o problematike detekcie tváří a tvárových črt, poskytli sme prehľad metód a prehľadne sme spracovali metodiku nášho skúmania. Testovanie výkonnosti poskytlo uspokojivé výsledky. Naše skúmanie a jeho výstupy boli ponúknuté vedeckej komunite v rámci publikácie Beniak et al. (2015). Ponúknuté spôsoby zlepšenia výkonnosti či robustnosti algoritmu dávajú priestor pre ďalší výskum v tejto oblasti.

## 7 Účast' autora v projektoch

- **HBB-NEXT Next-Generation Hybrid Broadcast Broadband**  
Projekt 7. rámcového programu FP7-ICT-2011-7-287848, 2011-2014
- **Pokročilé algoritmy spracovania obrazov na efektívne vyhľadávanie a kódovanie ľudských tvárí**  
VEGA 1/0961/11, 2011-2013
- **Inovácia vzdelávania v číslicovom spracovaní obrazu, v biometrii a v strojovom učení a neurónových sieťach**  
KEGA 010 – 4/2014STU
- **Návrh pokročilých metód biometrického rozpoznávania na základe obrazov tváre a dúhovky**  
VEGA 1/0529/13, 2013-2016
- **Návrh metód analýzy a klasifikácie pre biometrické rozpoznávanie obrazov tvárí a prevádzku komunikačných sietí**  
VEGA 1/0214/10, 2010-2011
- **Rozpoznávanie obrazov ľudských tvárí ako súčasť biometrických metód pre zvyšovanie bezpečnosti otvorenej spoločnosti**  
VEGA 1/3117/06, 2006-2008, (Vedecká grantová agentúra MŠ SR a SAV na základe hodnotenia vydala Certifikát o úspešnom ukončení projektu a dosiahnutí vynikajúcich výsledkov, apríl 2009)

## 8 Publikácie autora

**Beniak, M., Pavlovičová, J., Oravec, M. (2011)** '3D Chrominance Histogram Based Face Localization', Int. Journal of Signal and Imaging Systems Engineering (IJSISE), 2011 - Vol. 4, No.1 pp. 3 - 12, ISSN (Online): 1748-0701, www.inderscience.com/ijsise, /Scopus /Začlenenie ADM

Citácie:

1. Heribanová, P., Polec, J., Martinovič, M. (2013) Video Quality Control Using a ROI and Two-Component Weighted Metrics, World Academy of Science, Engineering and Technology, International Journal of Electrical, Robotics, Electronics and Communications Engineering Vol.7 No.7, 2013, ISSN 1307-6892, /GoogleScholar/
2. Polec, J., Heribanová, P., Hirner, T. (2013) Key Frames Extraction for Sign Language Video Analysis and Recognition, World Academy of Science, Engineering and Technology, Vol.7, No. 6 2013-06-24 pp.444-448, ISSN 1307-6892, /GoogleScholar/

**Beniak, M., Pavlovičová, J., Loderer, M. (2015)** 'Minimal Gabor Filter Bank for Face Features detection', International Review on Computers and Software, IRECOS, ISSN: 1828-6003 Vol. 9, No. 12 (2014), pp. 1977-1983. /Scopus /Začlenenie ADM

**Beniak, M., Pavlovičová, J. (2014)** 'Optimal Gabor Filter Bank Selection for Face Features Localization', Proc. of Redžúr 2014, International Workshop on Speech and Signal processing, ISBN 978-80-227-4162-0, 13 May 2014, Dubrovnik, Croatia, pp. 57-60. začlenenie AFC

**Beniak, M., Pavlovičová, J. (2011)** Luminance Component in Skin Color Detection, In: Proc. of the 18th International Conference on Systems, Signals and Image Processing IWSSIP 2011, June 16 – June 18, 2011, Sarajevo, Bosna and Herzegovina, ISBN 978-1-4577-0074-3, pp. 1-4. /Scopus /, začlenenie AFC

Citácie:

3. Jang, S.-W., Jung, M (2013) Skin region extraction combining color and depth features, Information (Japan), Vol. 16, No. 4, April 2013, pp. 2575-2584, ISSN: 13434500, /Scopus /

**Pintér, A., Pavlovičová, J. and Beniak, M. (2009)** Binary and Grayscale Morphology in Image Enhancement, Proc. of Redžúr 2009, 3rd Int. Workshop on Speech and Signal Processing, Sept. 24, 2009, Bratislava, Slovak Republic, zborník CD, ISBN 978-80-227-3137-9, pp. 16-19, začlenenie AFD

**Beniak, M., Pavlovičová, J., Oravec, M. (2008)** 'Automatic Face Detection Based on Chrominance Components Analysis', Proc. of the 15th International Conference on Systems, Signals and Image Processing IWSSIP 2008, 25.-28. jún 2008, Bratislava, Slovenská republika, ISBN 978-80-227-2880-5, pp. 475-478. /WebOfScience /, začlenenie AFD

Citácie:

4. A Cheddad, J Condell, K Curran, P Mc Kevitt (2009) A skin tone detection algorithm for an adaptive approach to steganography, Signal Processing, Elsevier, ISSN: 0165-1684, Vol. 89, No. 12, December 2009, pp. 2465–2478, 2009; CC (Current Contents), /WebOfScience/
5. Hung-Ming Sun (2010) Combining Multiple Skin Models for Adaptable Skin Detection, Journal of Computer Science and Application, Vol.6, No.1, June 2010, pp. 141-154, ISSN 1817-0013, Published by National Pingtung University of Education, /GoogleScholar/

**Beniak, M., Pavlovičová, J. (2011)** 'Image Enhancement for Skin Color Detection in YCbCr Color Space', Proceedings of 13th Conference of Doctoral Students ELITECH '11, ISBN 978-80-227-3500- začlenenie ADF

**Beniak, M., Pavlovičová, J. (2006)** 'Lokalizácia tvári vo videotelefónnych aplikáciách', EE časopis pre elektrotechniku a energetiku. Roč. 12, špeciálne vydanie, pp. 127-129., začlenenie ADF

**Beniak, M., Pavlovičová, J. (2007)** 'Segmentácia ľudských tvári - farebná analýza', EE časopis pre elektrotechniku a energetiku, roč. 13, špeciálne vydanie, ISSN 1335-2547, pp. 156-159. začlenenie ADF

## 9 Zoznam použitej literatúry

- Albiol, A., Torres, L., Delp, E., J. (2001)** 'Optimum Color Space for Skin Detection', s.l. : Politechnic University of Valencia, Spain, Politechnic University of Catalonia, Spain, Purdue University, USA, 2001.
- Baskan, S., Bulut, M. M., Atalay, V. (2002)** 'Projection based method for segmentation of human face and its evaluation Source', Pattern Recognition Letters, ISSN:0167-8655, Vol. 23, No. 14, pp. 1623-1629.
- Bandzi, P., Oravec, M., Pavlovičová, J. (2007)** 'New Statistics for Texture Classification Based on Gabor Filters', Radio Engineering, Vol. 16, No. 3, pp. 133-137, 2007.
- Brown, D., Craw, I., Lewthwaite, J., (2000)** 'A SOM-based approach to skin detection with application in real-time systems' Proceedings IEEE Region Ten Conference (TENCON'2000), Vol. 2, pp. 421-424.
- Canny, J. (1986)** 'A computational approach to edge detection', IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 8, No. 6., pp. 679-698.
- Dahal, B., Alsadoon, A., Prasad, P. W. C., Elchouemi, A. (2016)**. 'Incorporating Skin Color for Improved Face Detection and Tracking System', Helwet Packard Enterprise, 2016.
- Dass, S. C., Jain, A. K. a Lu, X. (2002)** 'Face Detection And Synthesis Using Markov Random Field Models', Proc. International Conference on Pattern Recognition, Quebec City, August 11-15, pp. 680-687.
- Dasan, M., Alqudah, A., Debeir, O. (2015)** 'Face Detection using Viola and Jones Method and Neural Networks', Proc. International Conference on Information, and Communication Technology Research (ICTRC 2015), pp. 40-43.
- Dempster, A.P., Laird, N.M., Rubin, D.B. (1977)**. 'Maximum Likelihood from Incomplete Data via the EM Algorithm', Journal of the Royal Statistical Society. Series B (Methodological), No. 39., pp. 1-38.
- Dey, A. (2016)**. 'A Contour based Procedure for Face Detection and Tracking from Video', Third International Conference on Recent Advances in Information Technology, 2016.
- Chai, K. D. and Ngan., N. (1999)** 'Face Segmentation Using Skin-Color Map in Videophone Applications', IEEE Tans. on Circuits and Systems for Video Technology, Vol. 9, No. 4, pp. 551-564.
- Chai, D. and Bouzerdoum, A. (2000)** 'A Bayesian approach to skin color classification in YCbCr color space', Proceedings of TENCON 2000, Vol. 2, pp. 421-424.
- Chen, J., Taskarin, C. M., Albiol, A., Bouman, C. A., Depl, E. J. (1999)** 'Vibe: A video indexing and browsing environment', IEEE-SPIE Conference on Multimedia Storage and Archiving Szstems IV, Boston, USA, september 1999.
- Fritsch, J., Lang, S., Kleinhängenbrock, M., Fink, G. A., Sagerer G. (2002)** 'Improving Adaptive Skin Color Segmentation by Incorporating Results from Face Detection', Technical Faculty, Bielefeld University, GERMANY 2002.
- Gabor, D. (1946)** 'Theory of communications', Inst. Elect. Eng., Vol. 93, pp. 429-457.
- Gallant, S. I. (1990)** 'Perceptron-based learning algorithms', IEEE Transactions on Neural Networks, Vol. 1, No. 2, pp. 179-191.
- Garcia, Ch. (2004)** 'A Survey of Face Detection and Recognition Techniques', Séminaire CNRT TIM – Vision par ordinateur pour les Télécommunications, 2004.
- Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I. H. (2009)** 'The WEKA Data Mining Software: An Update', SIGKDD Explorations, Vol. 11, No. 1, 2009.
- Hasan, M, Hossain, F. (2014)** 'Facial Features Detection in color Images Based on Skin Color Segmentation', International conference on Informatics, Electronics and Vision, 2014, pp. 1-5.
- Haykin, S. (1999)** 'Neural Networks: A Comprehensive Foundation', 2nd ed., Uppersaddle, N.J.: Prentice-Hall, 1999.



- Hamouz, M., Kittler, J., Kamarainen, J.-K., Paalanen, P., Kalviainen, H. (2004)** 'Affine-invariant face detection and localization using GMM-based feature detector and enhanced appearance model', Proceedings of Sixth IEEE International Conference on Automatic Face and Gesture Recognition, ISBN: 0-7695-2122-3, pp. 67-72.
- Huang, L.-L., Shimizu, A., Kobatake, H. (2005)** 'Robust face detection using Gabor filter features', Pattern Recognition Letters, Vol. 26, pp. 1641-1649.
- Jones, M.J. and Rehg, M. (2002)** 'Statistical color models with application to skin detection,' *International Journal of Computer Vision*, Vol. 46, No. 1, pp. 81-96.
- Kakumanu, P., Makrogiannis, S. and Bourbakis, N. (2007)** 'A survey of skin-color modeling and detection methods', Pattern Recognition, No. 40, pp. 1106 - 1122.
- Kamarainen J.-K., Kyrki,V., Kälviäinen, H. (2006)** 'Invariance Properties of Gabor Filter-Based Features—Overview and Applications'. IEEE Trans. on Image processing, Vol. 15, No.5, pp. 1088-1099.
- Kovač, J. , Peer, P. , Solina, F. (2003)** ,*Human Skin Colour Clustering for Face Detection*', Faculty of Computer and Information Science, University of Ljubljana, Slovenia, 2003.
- Kuchi, P., Gabbur, P., Bhat, P. S., Davinds, S. (2002)** ,*Human Face Detection and Tracking using Skin Color Modeling and Connected Component Operators*', Department.of Ellectrical nad Computer Engineering, Karnataka Regional Engineering College, Surathkal, Karnataka, India, 2002.
- Kumar, S., Kar, A., Cahndra, M. (2014)** ,*SVM Based Adaptive Median Filter Design for Face Detection in Noisy Images*', International Conference on Signal Processing and Integrated Networks (SPIN), 2014, pp. 695-698.
- Li, J., Wang, Y., Tan, T., Jain, A. K. (2004)** ,*Live Face Detection Based on the Analysis of Fourier Spectra*', Chinese Academy of Science, Beujing, China, 2004.
- Lichtenauer, J. F., Reinders, M. J. T., Hendriks, E. A. (2007)** ,*A Self-Calibrating Chrominance Model Applied to Skin Color Detection. Information and Communication Theory Group*',*Proceedings of the 2nd International Conference on Computer Vision Theory and Applications (VISAPP)*, March 2007.
- Lim, S. J. (1990)** ,*Two-dimensional Signal and Image Processing*, Prentice-Hall International Inc., 1990.
- Liuliu, W., Mingyang, L. (2016)** .*Multi-pose Face Detection Research based on Adaboost*', Eight International Conference on Measuring Technology and Mechatronics Automation, 2016, pp. 409-412.
- Mao, L., W., Zhang, K., Chai, T. (2010)** 'Selection of Gabor Filters for Improved Texture Feature Extraction', Proceedings of 2010 IEEE International Conference on Image Processing, September 2010, Hong Kong, pp. 361 - 364.
- Mohan, A.R. and Sudha, N. (2009)** 'Fast face detection using boosted eigenfaces', IEEE Symposium on Industrial Electronics & Applications, Vol. 2, pp. 1002-1006.
- Minear, M, Park, D.,C. (2004)** 'A lifespan database of adult facial stimuli', Behavior Research Methods, Instruments and Computers, Vol. 36, No. 4, 2004, pp. 630-633.
- Ming-Chieh Chi, Mei-Juan Chen, Chia-Hung Yeh and Jyong-An Jhu (2008)** 'Region-of-interest video coding based on rate and distortion variations for H.263+', *Signal Processing: Image Communication*, Vol. 23, pp. 194-211.
- Mohamed, A., Weng, Y., Jiang, J., Ipson, S. (2008)** ,*Face detection based on Neural Network using robust Skin Color Segmentation*', Proceedings of the International Multi-Conference on Systems, Signals and Devices, 2008.
- Oravec,M., Rozinaj,G., Beszédeš, M. (2008)** 'Detection and Recognition of Human Faces and Facial Features'. *Speech, Audio, Image and Biomedical Signal Processing Using Neural Networks*. Editors: Bhanu Prasad and S. R. Mahadeva Prasanna, Springer-Verlag, Berlin Heidelberg, ISBN: 978-3-540-75397-1, 2008, pp. 283-301
- Pantic, M., Rothkrantz, L. J. M. (2000)** 'Automatic Analysis of Facial Expressions: The State of Art', *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 22, No.12, pp. 1424-1445.
- Phillips, Wechsler, Huang and Rauss (1998)** 'The Facial Recognition Technology (FERET) Database. [http://www.itl.nist.gov/iad/humanid/feret/feret\\_master.html](http://www.itl.nist.gov/iad/humanid/feret/feret_master.html), [prístup 15/01/2010].

- Popovici, V. , Rodriguez, Y. , Marcel, S., Thiran, J.-Ph. (2004)** ,*On Performance Evaluation of Face Detection and Localization Algorithm*’, Proc. of 17th Intl. Conference on Pattern Recognition, pp.313-317, UK, 2004.
- Rowley, A. H., Bahuja, S., Kanade, T. (1997)** CMU Image Data Base. [http://vasc.ri.cmu.edu/idb/html/face/frontal\\_images/index.html](http://vasc.ri.cmu.edu/idb/html/face/frontal_images/index.html)
- Rowley, A. H., Bahuja, S., Kanade, T. (1998)** ,*Neural Network-Based Face Detection*’, IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, Vol. 20, p. 1, 1998.
- Praseetha, V. M., Vadivel, S. (2016)** ‘*Face Extraction using Skin Color and PCA Face Recognition in a Mobile Cloudlet Environment*’, Fourth IEEE International Conference on Mobile Cloud Computing, Services, and Engineering, 2016, pp. 41-45.
- Sahbi, H., Boujemaa, N. (2002)** ‘*Coarse to Fine Face Detection Based on Skin Color Adaption*’, INRIA Rocquencourt, Le Chesnay, FRANCE 2002.
- Samaria, F. a Harter, A. (1994)** ‘Parameterisation of a Stochastic Model for Human Face Identification’, Proceedings of 2nd IEEE Workshop on Applications of Computer Vision, Sarasota FL, December 1994.
- Sandeep, K., Rajagopalan, A. N. (2002)** ‘*Human Face detection in Cluttered Color Images Using Skin Color and Edge Information*’, Department of Electrical Engineering, Indian Institute of Technology – Madras Chennai, INDIA,2002.
- Shen, L., Bai, L. (2006)** ‘A review on Gabor wavelets for face recognition’, *Pattern Anal applic*, Vol. 9., pp. 273-292.
- Singh, S. K., Chauhan, D. S., Vatsa, M., Singh, R. (2003)** ,*A Robust Skin Color Based Face Detection Algorithm*’, Department of Computer Science and Engineering, Institute of Engineering and Technology, Jaunpur, INDIA 2003.
- Sobottka, K. a Pitas, I. (1998)** ‘*A Novel Method for Automatic Face Segmentation, Facial Feature Extraction and Tracking*’, Signal Processing: Image Comm., Vol. 12, No. 3, pp. 263-281, 1998.
- Shuliang Wang, Xiao Feng, Hehua Chi, Xiuling Wang (2009)** ‘*Localization and extraction on the eyes, nose and mouth of human face*’, IEEE International Conference on Granular Computing, 2009, ISBN: 978-1-4244-4830-2, pp. 561-564.
- Shuang, L. (2014)** ‘*A novel face detection algorithm based on PCA and Adaboost*’, Sixth International Conference on Measuring Technology and Mechatronics Automation, 2014, pp. 38-41.
- Sirohey S., Rosenfeld A. (2001)** ‘*Eye detection in a face image using linear and nonlinear filters*’, Pattern Recognition , No. 34, pp. 1367-1391.
- Serrano, A., de Diego, I., M., Conde, C., Cabello, E. (2009)** ‘*Recent advances in face biometrics with Gabor wavelets: A review*’, Pattern Recognition Letters, Vol. 31, pp. 372–381.
- Taskiran, C., Chen, J.-Y., Albiol, A., Torres, L., Bouman, Ch.A., Delp, E.J (2004)** ‘ViBE: A compressed video database structured for active browsing and search’, *IEEE transactions on multimedia*, Vol 6, No. 1, pp. 103-118.
- Verma et al. (2014)** ‘*Face detection using Skin color Modeling and Geometric Feature*’, In ternational Conference on Informatics, electronics and Vision, 2014.
- Vezhnevets, V., Sazonov, V., Andreeva, A. (2003)** ,*A Survey on Pixel-Based Skin Color Detection Techniques*’, Graphics and Media Laboratory, Faculty of Computational Mathematics and Cybernetics, Moscow State University, Moscow, RUSSIA, 2003.
- Viola, P, Jones, M. (2004)** ,*Robust Real-time Face Detection*’, International Journal of Computer Vision, 2001, Vol. 2, pp. 137-154.
- Yang, M.-H., Kriegman, D.J. and Ahuja, N. (2002)** ‘Detecting Faces in Images: A Survey’, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 24, No.1, pp. 34-57.
- Zhang, Z. (2006)** ,*Skin/Hair Region Segmentation using Texture Analysis*’, CUBS, 2006.

**Zhang W. et al. (2007)** '*Precise eye localization with adaboost and fast radial symmetry*', Lecture Notes in Computer Science, 4456, pp. 1068-1077.

**Zhang Ch., Zhang, Z. (2010)** '*A Survey of Recent Advances in Face Detection*', Technical Report MSR-TR-2010-66.

**Zuo, F. a de With, P., H., N. (2008)** '*Facial feature extraction by cascade of model-based algorithms*', *Signal Processing: Image Communication*, No. 23, pp. 194-211.