

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

Evidenčné číslo: FEI-5384-72758

**KLASIFIKÁCIA TVÁRÍ POMOCOU
KONVOLUČNÝCH SIETÍ PRE ANDROID OS
DIPLOMOVÁ PRÁCA**

2018

Bc. Ján Polaček

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

Evidenčné číslo: FEI-5384-72758

**KLASIFIKÁCIA TVÁRÍ POMOCOU
KONVOLUČNÝCH SIETÍ PRE ANDROID OS
DIPLOMOVÁ PRÁCA**

Študijný program: Aplikovaná informatika
Číslo študijného odboru: 2511
Názov študijného odboru: 9.2.9 Aplikovaná informatika
Školiace pracovisko: Ústav informatiky a matematiky
Vedúci záverečnej práce: Ing. Dominik Sopiak

Bratislava 2018

Bc. Ján Polaček



ZADANIE DIPLOMOVEJ PRÁCE

Autor práce: Bc. Ján Polaček
Študijný program: aplikovaná informatika
Študijný odbor: 9.2.9. aplikovaná informatika
Evidenčné číslo: FEI-5384-72758
ID študenta: 72758

Vedúci práce: Ing. Dominik Sopiak

Miesto vypracovania: Ústav informatiky a matematiky FEI STU

Názov práce: **Klasifikácia tvárí pomocou konvolučných sietí pre Android OS**

Jazyk, v ktorom sa práca vypracuje: slovenský jazyk

Špecifikácia zadania: Cieľom tejto práce je natréňovanie konvolučnej siete a jej aplikácie na problém klasifikácie ľudskej tváre v neriadených podmienkach. Ako prvé je dôležité naštudovať si problematiku konvolučných sietí. Následne je potrebné tieto znalosti aplikovať pri tréňovaní a aplikácie konvolučnej siete na zariadení Android OS.

Úlohy:

1. Naštudovať teóriu konvolučnej siete.
2. Oboznámiť sa s problematikou klasifikácie tvárí.
3. Navrhnuť štruktúru konvolučnej siete.
4. Vytvoriť navrhnutú konvolučnú sieť.
5. Diskusia získaných výsledkov.

Literatúra:

- JAIN, A. *Handbook of Biometrics*. London: Springer-Verlag, 2008. ISBN 978-0-387-71040-2.
- GOODFELLOW, M. -- WILLIAMS, S. -- MORDAESKI, M. *Actinomycetes in Biotechnology*. San Diego : Academic Press, 1988. 501 p.
- Vincent Dumoulin, Francesco Visin: "A guide to convolution arithmetic for deep learning", Cornell University, 23. Mar 2016

Dátum zadania: 18. 09. 2017

Dátum odovzdania: 11. 05. 2018

SÚHRN

SLOVENSKÁ TECHNICKÁ UNIVERZITA V BRATISLAVE
FAKULTA ELEKTROTECHNIKY A INFORMATIKY

Študijný program:	Aplikovaná informatika
Autor:	Bc. Ján Polaček
Diplomová práca:	Klasifikácia tvárí pomocou konvolučných sietí pre Android OS
Vedúci záverečnej práce:	Ing. Dominik Sopiak
Miesto a rok predloženia práce:	Bratislava 2018

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Aenean et est a dui semper facilisis. Pellentesque placerat elit a nunc. Nullam tortor odio, rutrum quis, egestas ut, posuere sed, felis. Vestibulum placerat feugiat nisl. Suspendisse lacinia, odio non feugiat vestibulum, sem erat blandit metus, ac nonummy magna odio pharetra felis. Vivamus vehicula velit non metus faucibus auctor. Nam sed augue. Donec orci. Cras eget diam et dolor dapibus sollicitudin. In lacinia, tellus vitae laoreet ultrices, lectus ligula dictum dui, eget condimentum velit dui vitae ante. Nulla nonummy augue nec pede. Pellentesque ut nulla. Donec at libero. Pellentesque at nisl ac nisi fermentum viverra. Praesent odio. Phasellus tincidunt diam ut ipsum. Donec eget est. A skúška mäččňov a dlžňov.

Kľúčové slová: Android

ABSTRACT

SLOVAK UNIVERSITY OF TECHNOLOGY IN BRATISLAVA
FACULTY OF ELECTRICAL ENGINEERING AND INFORMATION TECHNOLOGY

Study Programme:	Applied Informatics
Author:	Bc. Ján Polaček
Master's thesis:	Face classification using CNN for Android OS
Supervisor:	Ing. Dominik Sopiak
Place and year of submission:	Bratislava 2018

On the other hand, we denounce with righteous indignation and dislike men who are so beguiled and demoralized by the charms of pleasure of the moment, so blinded by desire, that they cannot foresee the pain and trouble that are bound to ensue; and equal blame belongs to those who fail in their duty through weakness of will, which is the same as saying through shrinking from toil and pain. These cases are perfectly simple and easy to distinguish. In a free hour, when our power of choice is untrammelled and when nothing prevents our being able to do what we like best, every pleasure is to be welcomed and every pain avoided. But in certain circumstances and owing to the claims of duty or the obligations of business it will frequently occur that pleasures have to be repudiated and annoyances accepted. The wise man therefore always holds in these matters to this principle of selection: he rejects pleasures to secure other greater pleasures, or else he endures pains to avoid worse pains.

Keywords: Android

Pod'akovanie

I would like to express a gratitude to my thesis supervisor.

Contents

Úvod	1
1 Rozpoznávanie tváre	2
1.1 Rozdelenie	3
1.2 Proces identifikácie	4
1.3 Techniky rozpoznávania tvári	5
1.3.1 Eigenfaces (PCA)	6
1.3.2 LDA, Fisherfaces	6
1.4 Detekcia tváre	7
1.5 Hľadania bodov tváre	7
1.6 Meranie úspešnosti	7
1.7 Súčasný stav	7
Záver	8

Zoznam algoritmov

Zoznam výpisov

Úvod

Tu bude krásny úvod s diakritikou atď.

A možno aj viac riadkový úvod.

Rozpoznávanie tváre

Rozpoznávanie tváre je jedným z úkonov, ktoré človek robí pravidelne a bez námahy v každodennom živote. Široká dostupnosť silných počítačov za nízku cenu vytvára obrovský záujem o automatické spracovanie digitálneho obrazu v širokom spektre aplikácií, ako sú napríklad biometrická autentifikácia, monitorovanie osôb, interakcia s počítačom či spravovanie multimédií. Výskum a vývoj v oblasti rozpoznávania tvárí nasleduje tento trend automaticky.

Hlavnými výhodami využívania rozpoznávania tvárí voči iným biometrickým metódam ako napríklad odtlačky prstov či dúhovka, je ich prirodzené a nerušivé používanie, ale najmä možnosť použitia na väčšiu vzdialenosť. Zo šiestich biometrických metód (tvár, odtlačky prstov, dlaň, hlas, dúhovka, podpis) je podľa The International Civil Aviation Organization(ICA0)[[icao](#)] rozpoznávanie tváre primárnou metódou pri kontrole identity na letisku.

Prvým automatickým systémom na rozpoznávanie tvárí bol podľa[[handbookface](#)], systém navrhnutý Takeo Kanadem v jeho práci[[kanade1974](#)] z roku 1974. Za ním nasledovalo obdobie bez výraznejšieho pokroku v automatickom rozpoznávaní tváre, až do roku 1990, kedy Sirovich a Kirby zverejnili článok[[kirby1990application](#)], v ktorom popisujú využitie nízko dimenzionálnych reprezentácií tváre odvodených z Karhunen-Loevovej transformácie alebo Principal Component Analysis(PCA). Podľa Jaina v[[handbookface](#)] bol ďalším veľkým míľnikom práca[[turk1991eigenfaces](#)] od Turka a Pentlanda na vlastných vektoroch tváre (eigenface), ktorá znovu naštartovala výskum v oblasti rozpoznávania tváre. Medzi ďalšie míľniky tiež Jain zaraďuje prácu na Fisherovej metóde[[belhumeur1997eigenface](#)] ktorá aplikuje Linear Discriminant Analysis(LDA) po aplikovaní PCA k dosiahnutiu vyššej presnosti, výskum na lokálnych Gaborových filtroch[[wiskott1997face](#)] k dosiahnutiu efektívnejších príznakov tváre a prístup AdaBoost učenia, založeného na architektúre kaskádneho klasifikátora pre detekciu v reálnom čase[[viola2001rapid](#)].

Od obdobia kedy bola narhnutá Eigenface metóda nastal veľký pokrok v oblasti rozpoznávania tváre. V kontrolovaných podmienkach, kde je možné ovládať svetelnosť, postoj osoby či výraz tváre, prekonáva automatické rozpoznávanie tváre ľudí, a to najmä pokiaľ databáza obsahuje veľké množstvo snímok tváre. Napriek tomu rozpoznávanie tváre stále čelí mnohým výzvam, najmä problematike rozpoznávania tváre v neriadených podmienkach.

Rozdelenie

Ako aj ostatné biometrické systémy, aj rozpoznávanie tváre funguje v jednom alebo oboch z režimov:

- Verifikácia (autentifikácia)
- Identifikácia tváre

Pri verifikácii tváre ide o porovnanie jedna k jednej, čo znamená, že jeden snímok tváre sa porovnáva len jedným záznamom identity v databáze, za ktorú sa prehlasuje. Typickým využitím tohoto režimu je samoobslužná kontrola identity prostredníctvom elektronického pasu[**handbookface**].

Identifikácia tváre zahŕňa porovnanie jedna k mnohým, čo znamená, že jeden snímok tváre sa porovnáva s viacerými záznamami identít v databáze a vyberie jednu[**handbookface**]. V niektorých prípadoch využitia je postačujúce nájsť len najpodobnejšiu identitu. V iných prípadoch, ako napríklad sledovanie podozrivých osôb, je okrem nájdania najpodobnejšej tváre potrebné zaviesť aj prah spoľahlivosti, a tie tváre ktoré dosiahli mieru podobnosti väčšiu ako je prah, sú zaznamenané.

Úspešnosť systému na rozpoznávanie tváre závisí vo veľkej miere na množstve variabilných faktorov, ako je osvetlenie, poloha tváre, mimika, vek, make-up, účes, brada či pohyb tváre. Na základe týchto faktorov Jain rozdeľuje[**hanbookface**] na 2 kategórie vzhľadom na miery spolupráce užívateľov:

- Scenár so spolupracujúcim užívateľom
- Scenár s nespupracujúcim užívateľom

Prípad spolupracujúceho užívateľa je využívaný napríklad pri prihlasovaní do počítača, riadenie fyzického prístupu, elektronické pasy (e-passport), teda prípady kedy má užívateľ záujem spolupracovať na správnom zoznámaní tváre (napr. pod správnym zoznámaním tváre môžeme rozumieť napríklad snímok tváre z predu s neutrálnym výrazom a otvorenými očami) kvôli prístupu alebo povoleniu.

V prípade nespupracujúceho užívateľa, ktorý je typické pre už spomínané sledovanie podozrivých, si osoba nie je vedomá toho, že podlieha identifikácii. Čo sa týka vzdialenosti medzi tvárou osoby a kamerou, sa v spolupracujúcom scenári využíva krátka vzdialenosť, typicky o jedného metra, a ide teda o oveľa jednoduchšiu úlohu v porovnaní s identifikáciou nespupracujúcej osoby často na väčšiu vzdialenosť.

Proces identifikácie

Jain popisuje rozpoznávanie tváre takto: “Rozpoznávanie tváre sa zaraďuje medzi problémami rozpoznávania vzorov, kde tvár, ktorá je reprezentovaná ako trojzormerný objekt, podlieha odlišnostiam vo svetle, postoji, výraze a iných faktoroch, potrebuje byť identifikovaná na základe zozbieraných obrázkov“[**handbookface**]. Zatiaľ čo roznávanie tváre z dvojrozmerného snímku je dnes bežne používané vo väčšine prípadov, v niektorých, najmä tých ktoré vyžadujú vyššiu bezpečnosť, sa využíva trojrozmerný snímok tváre, prípadne snímky tváre mimo bežne viditeľného spektra - napr. termografický snímok tváre. Systém na rozpoznávanie tváre sa podľa Jaina[**handbookface**] vo všeobecnosti skladá zo štyroch základných častí, ako je ukázané na obrázku1:

- Detekcia tváre a lokalizácia bodov tváre
- Normalizácia tváre
- Extrakcia príznakov
- Hľadanie zhody tváre

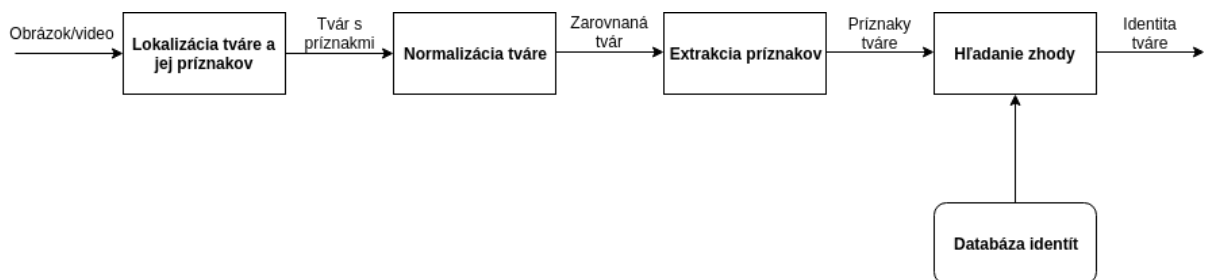


Figure 1: Proces rozpoznávania tváre

Pri detekcii tváre ide primárne o oddelenie oblasti tváre od pozadia snímku. V prípade videa je potrebné sledovať (track) detekovanú tvár naprieč niekoľkými snímkami videa pomocou komponentu na sledovanie pohybu tváre. Zatiaľ čo detekcia tváre poskytuje len hrubý odhad polohy a veľkosti tváre, lokalizácia bodov tváre najde už konkrétne časti tváre, ako sú napríklad oči, nos, obrys tváre a podobne. Lokalizácia bodov tváre je zväčša vykonávaná osobitným komponentom na lokalizáciu bodov alebo komponentom na vyrovnanie (alignment) tváre.[**handbookface**]

Normalizácia tváre ide o normalizáciu tváre v geometrickom a fotometrickom zmysle. Tento krok je nevyhnutný, pretože sa od najnovších a najlepších (state-of-the-art) rozpoznávacích metód očakáva, že dokážu rozpoznať tvár v rôznych polohách a rôznom svetle. Geometrická normalizácia vykonáva transformáciu tváre do štandardného formátu snímku

pomocou orezania (crop) tváre. Snímok je následne zakrivený (warp) a upravený (morph) kvôli ešte lepšej a presnejšej normalizácii tváre. Úlohou fotometrickej normalizácie je spracovanie snímku na základe osvetlenia či farebnej škály.[**handbookface**]

Extrakcia príznakov je vykonávaná na normalizovanom snímku tváre, s cieľom vybrať charakteristické informácie, ktoré sú užitočné pri rozlišovaní tvárí rozdielnych osôb, pričom odolný voči odchýlkam v geometrickej a fotometrickej normalizácii. Extrahované príznaky sú následne použité pri hľadaní zhody s identitou.[**handbookface**]

Pri hľadaní zhody tváre sa porovnávajú extrahované príznaky zo vstupnej tváre s jednou alebo viacerými tvármi ktoré už sú zapísané v databáze. Výsledkom porovnania s jednou tvárou je výsledkom odpoveď áno alebo nie (verifikácia). V prípade porovnávania s viacerými tvármi je výsledkom identita vstupnej tváre, za predpokladu, že identita ktorá je nájdená presiahne prah spoľahlivosti, inak vráti informáciu, že tvár je neznáma. V súčasnosti je v tejto oblasti najväčšou výzvou nájsť spoľahlivé meranie vhodné na určenie podobnosti príznakov tváre.[**handbookface**]

Presnosť systému na rozoznávanie tváre vo veľkom závisí na správnosti extrahovaných príznakov tváre, ktoré naopak, závisia od správnej lokalizácii a normalizácii tváre.

Techniky rozpoznávania tvári

Zhao rozdeľuje[**zhao2003face**] algoritmy rozpoznávania do dvoch základných kategórií, vzhľadom na spôsob extrakcie príznakov:

- Metódy založené na príznakoch (feature-based)
- Metódy založené na vzhľade (appearance-based)

Metódy založené na príznakoch využívajú rôzne vlastnosti a geometrické atribúty na popis tváre, ako sú napríklad vzdialenosti či uhly medzi bodmi tváre. Na druhej strane, metódy založené na vhlade využívajú globálne vlastnosti tvárového vzoru. Typickou črtou algoritmov založených na vhlade je výpočet bazových vektorov, na ktoré je následne tvár premietnutá. Koeficienty takejto projekcie sú potom použité k efektívnej reprezentácii údajov tváre[**handbookbio**]. Oblúbené algoritmy ako sú Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), Independent Component Analysis (ICA), Local Feature Analysis (LFA), Manifolds, Correlation Filters alebo Tensorfaces sú založené práve na vhlade tváre. Holistický prístup k rozpoznávaniu tváre má však často problémy pri rôznych polohách tváre[**handbookbio**].

Eigenfaces (PCA)

Podstatou metódy Eigenfaces navrhnutej Turkom a Pentlandom[**turk1991eigenfaces**], tiež známej ako PCA, je hľadanie najmenej kvadratickej chyby lineárneho podpriestoru, ktorý mapuje dáta z originálneho N-rozmerného priestoru na M-rozmerný priestor príznačov, kde $M \ll N$. Týmto dosiahneme redukcii dát do M rozmeného priestoru využitím M vlastných vektorov matice kovariance, ktoré korešpondujú s najväčšími hodnotami vlastných čísel[**handbookbio**]. Konečné bázové vektory ktoré sú najvhodnejšie na popis dát, sú nájdené pri procese optimalizácie, ktorej podstatou je maximalizácia variance premietnutých dát. Jain popisuje[**handbookbio**] proces výberu bázových PCA vektorov W optimalizačnou funkciou(1), kde S_T označuje úplne maticu rozptylu, ktorá obsahuje kovariancie dát tváre.

$$W_{PCA} = \operatorname{argmax} |W^T S_T W| = [w_1, w_2, \dots, w_m] \quad (1)$$



Figure 2: Prvých 6 bázových vektorov Eigenfaces, prebraté z[**handbookbio**]

Na obrázku2 vidíme príklad vlastných vektorov vybraných metódou Eigenfaces, na obrázkoch po procese normalizácie. Metóda PCA je vhodná na reprezentáciu dát, čo neznamená, že je vhodná aj na rozdeľovanie do tried.

LDA, Fisherfaces

Metóda LDA[**duda2012pattern**] je oveľa vhodnejšia k hľadaniu projekcií, ktoré dobre oddeľujú rozdielne triedy. Táto metóda je založená na hľadaní optimálnych projekčných vektorov, ktoré optimalizujú pomer medzitriednych (within class) a vnútrotriednych (between class) vzdialeností - maximalizuje oddelenie tried v premietnutom priestore. Optimálne bázové vektory LDA, môžu byť podľa[**handbookbio**] definované ako rovnica(3), kde S_B značí medzitriednu maticu rozptylu a kde S_W vnútrotriednu maticu rozptylu.

$$W_{LDA} = \operatorname{argmax} \frac{|W^T S_B W|}{|W^T S_W W|} \quad (2)$$

Zvyčajne pri riešení problematiky rozpoznávania tváre (a väčšine ostatných problémov rozpoznávania vzorov obrázkov) je množstvo tréningových obrázkov menší ako počet

pixelov (dimenzionalita dát), čo znamená, že vnútrotriedová matica rozptylu S_W je singulárna, čo je problémom pre LDA. Na vyriešenie tohoto problému singulárnej matice sa na najprv vykonáva PCA na redukovanie dimenzionality dát a až následne sa aplikuje LDA v menej rozmernom podpriestore PCA. Na základe týchto zmien bolo dosiahnuté zlepšenie výsledkov v porovnaní s tradičnou PCA metódou. Projekčné vektory, ktorých príklad môžeme vidieť na obrázku 3 Fisherfaces sú tie, ktorú spĺňajú maximalizujú výsledok optimalizačnej funkcie(??).

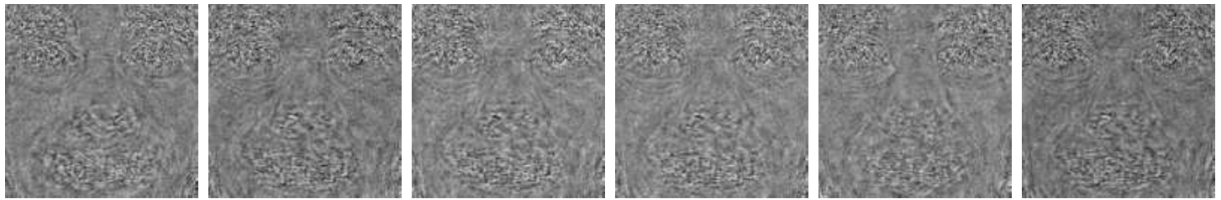


Figure 3: Prvých 6 bazových vektorov Fisherfaces, prebraté z[**handbookbio**]

$$W_{LDA} = \underset{W}{\operatorname{argmax}} \frac{|W^T W_{PCA}^T S_B W_{PCA} W|}{|W^T W_{PCA}^T S_W W_{PCA} W|} \quad (3)$$

Detekcia tváre

Hľadania bodov tváre

Meranie úspešnosti

Súčasný stav

Záver

Conclusion is going to be where?

Here.