

SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

SEMINAR

Detekcija lica u grupnim scenama

T. Antunović, S. Čolaković, E. Smoljan, F. Stamenković, I. Weber

Voditelj: *Marijo Maračić*

Zagreb, prosinac 2014.

SADRŽAJ

1. Uvod i problematika	1
2. Pregled postojećih rješenja	2
2.1. Detekcija lica - Toni	2
2.2. Detekcija lica - Edi	3
2.3. Bostonski maraton	6
3. Predložena implementacija	8
4. Rezultati	9
5. Zaključak	10
6. Literatura	11
7. Sažetak	13

1. Uvod i problematika

2. Pregled postojećih rješenja

2.1. Detekcija lica - Toni

Detekcija lica je prvi korak u sustavima za raspoznavanje lica s ciljem lokalizacije i ekstrakcije lica od pozadine. Ljudsko lice je dinamičan objekt s visokim stupnjem varijabilnosti na slikama, što čini detekciju lica teškim problemom u oblasti računalnog vida.

Sustavi za detekciju lica se mogu podijeliti na sustave temeljene na značajkama (engl. feature-based) i sustave temeljene na slici (engl. image-based) [6]. Oni koji su temeljeni na značajkama mogu vršiti analizu niskog nivoa koja se oslanja na rubove, sive nivoe ili boje, mogu vršiti analizu značajki ili kreirati modele aktivnih oblika (engl. active shape models). Sustavi temeljeni na slici se dijele na tri glavne skupine: neuronske mreže, metode linearnih potprostora, te razni statistički pristupi. U radu koji će poslužiti kao osnova za izradu ovog projekta [3] za detekciju lica se vrši jednostavna analiza niskog nivoa temeljena na segmentaciji boja, slike i višeslojnom filtriranju tako dobivenih regija koristeći različite vrijednosti pragova sličnosti s prosječnim licem.

Sustav za robusnu detekciju lica u realnom vremenu opisan u [16] može poslužiti kao primjer sustava za detekciju lica koji vrši analizu značajki. Temelji se na prikazu slike koji su nazvali “integralna slika”, jednostavnom klasifikatoru izgrađenom koristeći AdaBoost algoritam učenja da izabere najbitnije značajke iz jako velikog skupa potencijalnih značajki, te na kaskadnom kombiniranju klasifikatora koje omogućuju da pozadinske regije budu brzo odbačene i da što je moguće veći dio računanja koncentrira na regije koje imaju veću vjerojatnost da predstavljaju lice.

Modeli aktivnih oblika predstavljaju značajke višeg nivoa od prethodno spomenutih modela. Kada se inicijalizira u blizini značajke model aktivnog oblika će kroz interakciju s lokalnim značajkama poput rubova i osvjetljenosti postepeno zauzeti oblik značajke višeg nivoa. Na taj način se mogu koristiti ne samo za detekciju lica nego i za prepoznavanje lica kroz označavanje bitnih regija poput očiju, obrva, usta i nosa [14].

Pri detekciji lica koristeći neuronske mreže u [12] korišteno je više neuronskih

mreža koje su obavljale različite zadatke. Prva neuronska mreža je vršila procjenu poze potencijalnih regija koje predstavljaju lice. Nakon nje se vršilo pretprocesiranje s ciljem smanjivanja varijacija uzrokovanih osvjetljenjem i razlikama vezanim za kamere. Nakon toga za svaku pozu je korišteno nekoliko neuronskih mreža koje su učile različite stvari iz podataka za treniranje i davale različite rezultate. U posljednjem sloju njihovi izlazi su kombinirani koristeći jednostavnu heuristiku s ciljem povećanja točnih detekcija. Pristup zasnovan na dubokim konvolucijskim neuronskim mrežama [17] na efikasan način izvlači značajke tokom učenja i na Fddb bazi trenutno ostvaruje najbolje rezultate.

Metode linearnih potprostora su metode poput PCA, ICA, LDA i FA. Za ovaj projekt ovakve metode će se koristiti u kontekstu prepoznavanja lica, pa u kontekstu detekcije su spomenuti tek kao jedna od mogućnosti.

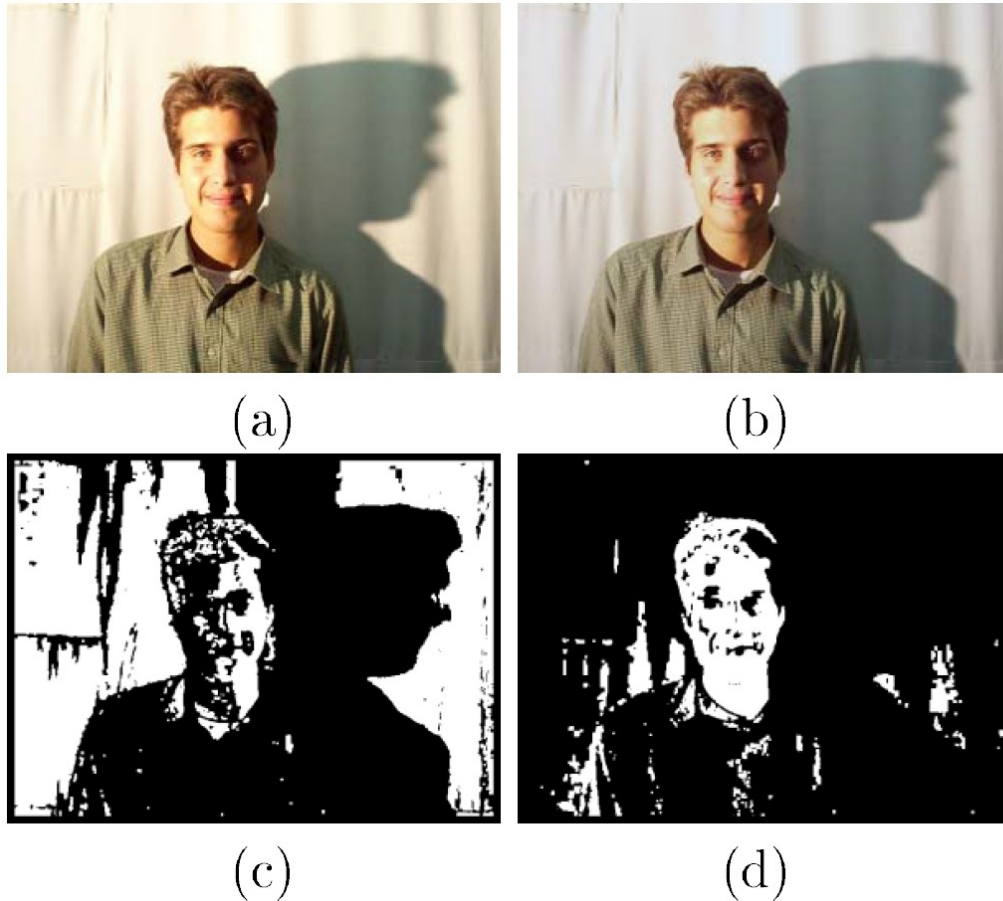
Kao primjer statističkog pristupa u detekciji lica može poslužiti FloatBoost učenje bazirano na AdaBoost algoritmu [9]. FloatBoost nakon svake iteracije AdaBoost učenja koristi povratni mehanizam za direktnu minimizaciju pogreške. Postiže manju pogrešku učenja i generalizacije koristeći manji broj slabih klasifikatora od AdaBoost algoritma.

2.2. Detekcija lica - Edi

Prvi dio zadatka prepoznavanja lica u grupnoj sceni je sama detekcija lica koje treba prepoznati. Ovaj problem je dosta istražen u području računalnog vida i postoje različiti načini kako dobiti zadovoljavajuće rezultate poput prepoznavanja lica po boji ili po kretanju (ili koristeći oboje). U ovom radu se odlučujemo koristiti detekciju lica temeljenu na boji što će se bolje opisati. Postupak detekcije lica temeljene na boji kojeg koriste radovi poput [13] i [3] je sljedeći: detektirati područja na slici koja odgovaraju koži i pronađena područja klasificirati kao lica ili ne-lica.

Kako bi se prvi dio postupka obavio efikasno zaključilo se da je potrebno sliku iz RGB prostora konvertirati u YCrCb ili YIQ prostor i onda izgraditi binarnu sliku (masku) u kojoj je svaki piksel označen ako komponente piksela zadovoljavaju uvjet pripadanja koži. Sam uvjet pripadanja piksela području kože varira kroz radove: u [11] koji za obradu koristi YCrCb prikaz uvjet glasi $90 < Y < 180$, $90 < Cr < 130$, $80 < Cb < 150$, dok je u radu [13] utvrđena i opisana zavisnost između Cr, Cb i Y komponenti te se prvo izvršava nelinearna transformacija Cr i Cb komponenti i nakon toga ispituje uvjet pripadanja. Ove opisane metode su empirijske i moguće je da svaki istraživački tim definira svoje u sklopu svog rada. U radu [13] se prije samog stvaranja binarne

slike početna slika još provlači kroz fazu pretprocesiranja u kojoj se gleda umanjiti utjecaj izvora svjetlosti na boje u slici. Dobivena binarna maska se još dodatno može transformirati operacijama otvaranja, filtriranja, dilatacije, erozije i zatvaranja kako bi se postigle kompaktnije maske koje predstavljaju moguća područja lica. Dobivena područja se iz slike izvlače postupcima segmentacije.



Slika 2.1: Određivanje područja na slici koja pripadaju koži uz normalni i smanjeni utjecaj svjetlosti.

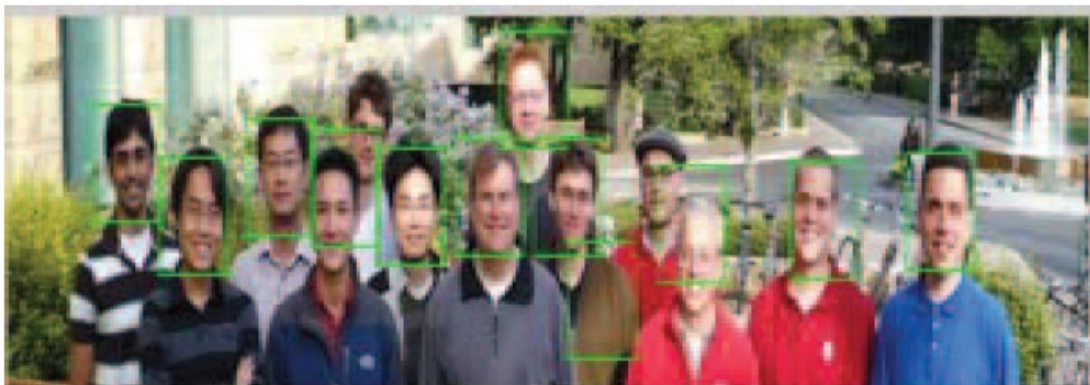
U drugom se dijelu postupka područja slike dobivena segmentacijom moraju klasificirati kao lica odnosno ne-lica zbog toga što se segmentacijom izdvajaju dijelovi slike koji odgovaraju koži što često uključuje i ostale dijelove tijela poput ruku. U radu [13] se tom problemu pristupa tako da se najprije odrede područja u kandidatima za lica koji odgovaraju očima i ustima te se lica prihvaća ako je ocjena pronađenih kandidata bolja od neke granične vrijednosti.

Rad [3] ovom problemu pristupa na malo drugačiji način: na samom početku postoji skup lica koja čine skup za učenje, od tih se lica stvara slika prosječnog lica i za svakog kandidata lica se računa korelacija kandidata sa prosječnim licem. Ako je ta



Slika 2.2: Primjer detekcije lica u radu [13].

korelacija niska kandidat se odbacuje, inače ako je područje kandidata dovoljno veliko on se prihvaća kao područje lica. Test korelacije je sličan načinu na koji se maximal rejection classifier (MRC) koristi za detekciju lica opisanog u radu [5]. Prihvaćanje kandidata na temelju veličine područja se obavlja koristeći prilagođavajuću granicu prihvaćanja kako bi se omogućilo prihvaćanje malih lica na slici, a istodobno odbacivalo manja područja za koje je prolazak testa korelacije moguć (dijelovi tijela poput ruku).



Slika 2.3: Primjer detekcije lica u radu [3].

2.3. Bostonski maraton

Zanimljiv primjer korištenja računalnog vida u svrhu detekcije i raspoznavanja lica je bombaški napad na Bostonski maraton koji se dogodio 2013. godine. Detekcija na slikama maratona se raspravlja u radu [1], gdje se predlaže algoritam baziran na Viola-Jones algoritmu [15]. Koristeći integralnu sliku i pomični prozor, algoritam prolazi kroz cijelu sliku te provjerava za svaki njen dio da li se na njemu nalazi neko lice. U sklopu ovog rada primljeno je na znanje da na značajke lica uvelike utječe njihova nakrenutost pa je algoritam dodatno opremljen detekcijom lica koja su okrenuta prema naprijed, ulijevo ili udesno. Nakon što je izračunao pripadnost svakom od te tri grupe, računa konačnu vrijednost iz te ta tri faktora pripadnosti te provjerava da li konačna pripadnost prelazi postavljeni prag. Ukoliko prelazi, algoritam zaključuje da se na prozoru nalazi lice. Algoritam je, osim na slikama maratona, testiran i na FDDB ispitnom skupu te je dao bolje rezultate od ostalih sličnih algoritama, uz više ispravnih i manje pogrešnih detekcija. Nažalost, algoritam se pokazao osjetljivim na utjecaj vizualnog šuma i prekrivanja lica, a problem pozicije lica nije u potpunosti riješen. Raspoznavanje lica je također provedeno nad slikama Bostonskog maratona, a time se bavi rad [8], koji testira dva komercijalna sustava za raspoznavanje lica, PittPatt 5.2.2 i NeoFace 3.1 nad slikama nekooperativnih osumnjičenika, zbog čega su same slike slabije kvalitete. Valja napomenuti da je istraga vezana uz maraton mogla biti puno kraća uz valjane algoritme raspoznavanja, budući da su se slike lica braće Tsarnaev, glavnih osumnjičenika, nalazile u službenoj bazi podataka. Iako ne nudi uvid u rad samih sustava, rad prikazuje varijacije u dobivenim rezultatima ovisno o načinu izbora skupa sa kojim se lica uspoređuju. Sustavi su prvo testirani nad cijelom bazom podataka koja se sastoji od milijun profila te su za svaku danu sliku dobivene tri osobe koje najviše odgovaraju traženom licu, pri čemu je samo NeoFace uspio ispravno prepoznati jednu sliku mlađeg brata. Nakon toga su za svakog brata uneseni dodatni podaci poput spola, rase i dobi što je smanjilo broj mogućih osoba za otprilike šest puta u oba slučaja, a povećanje kvalitete je bilo proporcionalno tom smanjenju. Na kraju se isprobala varijanta sa spajanjem rezultata za svaku osobu koju se pokušalo prepoznati u bazi, čime se dobivaju bolji rezultati ako su pojedinačne slike davale slične rezultate, ali gori ukoliko nisu. Sve u svemu, rad dobro pokazuje utjecaj pristupa testnom skupu. Rad [2] predstavlja još jedan algoritam raspoznavanja lica u kojem autori nastoje omogućiti precizno raspoznavanje bez obzira na gore navedene smetnje, poput prekrivanja, položaja i vizualnog šuma. Koriste naučeni koder kako bi dobili opisnik koji se pred kraj obrađuje PCA (principle component analysis) metodom kako bi se smanjile nje-

gove dimenzije i poboljšale performanse algoritma. Detaljnije, nakon detektiranja lica pronalaze se komponente lica poput nosa, očiju i obraza te se nakon filtriranja šuma dobivaju digitalne reprezentacije njihovih značajki u obliku vektora. Vektori se normaliziraju te se pomoću njih utvrđuje da li se radi o traženom licu. Pretpostavlja se, naravno, da je koder naučen na značajkama lica koja su korištena u treniranju. Algoritam je također u stanju paralelno računanju značajka lica odrediti položaj lica te tu informaciju iskoristiti u konačnoj odluci, čime se poboljšava preciznost. Možda najzanimljiviji dio algoritma je mogućnost kombiniranja više opisnika koristeći poptorni vektor, čime algoritam u konačnosti postiže preciznost od visokih 84.45 % na LFW [7] bazi slika. Većina algoritama za prepoznavanje lica koriste neki oblik preprocesiranja slike kako bi se dobili podaci primjereni za korištenje odgovarajućih algoritama za raspoznavanje. U radu koji nam je dan kao glavna osnova za izradu projekta se koristi i metoda PCA (principle component analysis) i metoda ICA (independent component analysis), što čini rad [4] dosta zanimljivim budući da se fokusira upravo na usporedbu performansi te dvije metode. U literaturi ne postoji koncepcija o tome koja je metoda bolja u kojim situacijama, ali rad dolazi do zaključka da implementacije ICA metode u prosjeku daju bolje performanse i višu preciznost. Za kraj, rad [10] daje pregled smjera u kojem se trenutno kreću sustavi za raspoznavanje lica i algoritmi koji se pri tome koriste. Opisuje nagli porast interesa za područje u proteklom godinama zbog napretka na području strojnog učenja i računalne grafike te potražnje za sustavima koji koriste tehnologije kojima se bavi računalni put, a neka od primjena su prepoznavanje profilnih slika, automatsko praćenje i promatranje većeg broja ljudi, digitalna rekonstrukcija lica itd. Autor dijeli algoritme za prepoznavanje lica u dvije velike skupine. Prva skupina su algoritmi koji se osnivaju na izgledu, odnosno slikama pojedinaca te se obično koriste u paru sa vektorskim prikazom slika koje obrađuju, što znači da u ovu skupinu pripadaju prije navedeni radovi, a na internetu se nude razne baze podataka slika poput *AT&T Cambridge* ¹. Sa druge strane su algoritmi bazirani na modelu, koji nastoje promatrati lica kao 3D modele koji vjerodostojno prikazuju lice sa svim njegovim značajkama te pomoću toga vrše prepoznavanje. Autor navodi neke od algoritama obje skupine te opisuje prednosti i mane svake skupine.

¹<http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>

3. Predložena implementacija

4. Rezultati

5. Zaključak

Zaključak.

6. Literatura

- [1] Jeremiah R Barr, Kevin W Bowyer, i Patrick J Flynn. The effectiveness of face detection algorithms in unconstrained crowd scenes. U *Applications of Computer Vision (WACV), 2014 IEEE Winter Conference on*, stranice 1020–1027.
- [2] Zhimin Cao, Qi Yin, Xiaoou Tang, i Jian Sun. Face recognition with learning-based descriptor. U *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, stranice 2707–2714, June 2010. doi: 10.1109/CVPR.2010.5539992.
- [3] D. N. Chandrappa i M. Ravishankar. Automatic face recognition in a crowded scene using multi layered clutter filtering and independent component analysis. U Ajith Abraham, Albert Y. Zomaya, Sebastián Ventura, Ronald Yager, Václav Snásel, Azah Kamilah Muda, i Philip Samuel, urednici, *ISDA*, stranice 552–556. IEEE, 2012. ISBN 978-1-4673-5117-1. URL <http://dblp.uni-trier.de/db/conf/isda/isda2012.html#ChandrappaR12>.
- [4] Bruce A. Draper, Kyungim Baek, Marian Stewart Bartlett, i J. Ross Beveridge. Recognizing faces with pca and ica. *Comput. Vis. Image Underst.*, 91(1-2):115–137, Srpanj 2003. ISSN 1077-3142. doi: 10.1016/S1077-3142(03)00077-8. URL [http://dx.doi.org/10.1016/S1077-3142\(03\)00077-8](http://dx.doi.org/10.1016/S1077-3142(03)00077-8).
- [5] Michael Elad, Yacov Hel-or, i Renato Keshet. Pattern detection using maximal rejection classifier. U *In Int. Workshop on Visual Form*, stranice 28–30, 2000.
- [6] Erik Hjelmås i Boon Kee Low. Face detection: A survey. *Computer Vision and Image Understanding*, 83(3):236–274, Rujan 2001. doi: 10.1006/cviu.2001.0921. URL <http://dx.doi.org/10.1006/cviu.2001.0921>.
- [7] Gary B. Huang, Manu Ramesh, Tamara Berg, i Erik Learned-miller. Labeled faces in the wild: A database for studying face recognition in unconstrained environments.

- [8] Joshua C. Klontz i Anil K. Jain. A case study on unconstrained facial recognition using the boston marathon bombings suspects, 2013.
- [9] Stan Z. Li i ZhenQiu Zhang. Floatboost learning and statistical face detection. *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, 26(9):2004, 2004.
- [10] Xiaoguang Lu. Image analysis for face recognition.
- [11] Md. Abul Bashar Md. Hafizur Rahman, Suman Chowdhury. An automatic face detection and gender classification from color images using support vector machine.
- [12] Henry A. Rowley, Student Member, Shumeet Baluja, i Takeo Kanade. Neural network-based face detection. *IEEE Transactions On Pattern Analysis and Machine intelligence*, 20:23–38, 1998.
- [13] Andrew Senior, Rein-Lien Hsu, Mohamed Abdel Mottaleb, i Anil K. Jain. Face detection in color images. *IEEE Trans. Pattern Anal. Mach. Intell.*, 24(5):696–706, Svibanj 2002. ISSN 0162-8828. doi: 10.1109/34.1000242. URL <http://dx.doi.org/10.1109/34.1000242>.
- [14] Prabhu Utsav i Keshav Seshadri. Facial recognition using active shape models, local patches and support vector machines.
- [15] Paul Viola i Michael Jones. Rapid object detection using a boosted cascade of simple features. stranice 511–518, 2001.
- [16] Paul Viola i Michael Jones. Robust real-time face detection. *International Journal of Computer Vision*, 57:137–154, 2004.
- [17] Cha Zhang i Zhengyou Zhang. Improving multiview face detection with multi-task deep convolutional neural networks. U *Applications of Computer Vision (WACV), 2014 IEEE Winter Conference on*, stranice 1036–1041. IEEE, 2014.

7. Sažetak

Sažetak.