

UNIVERZA V LJUBLJANI
FAKULTETA ZA RAČUNALNIŠTVO IN
INFORMATIKO
LABORATORIJ ZA RAČUNALNIŠKI VID

AVTOMATSKO ISKANJE ČLOVEŠKIH
OBRAZOV NA SLIKAH

Peter Peer

Delo je pripravljeno po razpisu odbora za Prešernove nagrade
študentom za 1998 leto, pod mentorstvom
prof. dr. Franca Soline.

Ljubljana, 1998

Tekst je oblikovan z L^AT_EX urejevalnikom besedil.

Lektor: Nataša Meh, prof. slov.

AVTOMATSKO ISKANJE ČLOVEŠKIH OBRAZOV NA SLIKAH

Peter Peer

Mentor: prof. dr. Franc Solina

POVZETEK

Delo obsega pregled problematike avtomatskega iskanja človeških obrazov na slikah in možnosti izboljšav algoritmov na tem področju. Predlagan je algoritem za avtomatsko iskanja človeških obrazov na slikah, ki poskuša združiti več idej, o katerih so pisali različni avtorji. Algoritem predpostavlja določene olajševalne okoliščine oziroma omejitve. Podane so tudi možne izboljšave, katerih ideje izvirajo iz negativnih rezultatov testa na 43 testnih slikah. Algoritem zaradi omejitev in temu primerne zasnove ni univerzalen. Tako ni primeren kot izhodišče za načrtovanje fleksibilnejšega algoritma. Dovolj učinkovit pa je za aplikacijo, katero omejitve algoritma ne bi ovirale in bi zahtevala hitro delovanje.

KLJUČNE BESEDE

računalniški vid, avtomatsko iskanje, človeški obraz, izločanje značilk obraza, pove-zovanje značilk obraza, iskanje kandidatov za obraz

AUTOMATIC DETECTION OF HUMAN FACES IN IMAGES

Peter Peer

Supervisor: prof. dr. Franc Solina

ABSTRACT

This work contains a proposal for an automatic in-image face detection method, that tries to join several theories, which many authors wrote of before. The method assumes certain circumstances and constraints, respectively. Just because of such constraints it is not applicable universally. Thus, it's not suitable for a starting-point of a more flexible method. On the other hand, however, it is effective enough to be used in an application, where this constraints would not obstruct the employed algorithm and also if the application would require fast execution. Effectiveness of the method was tested on 43 test-images.

KEY WORDS

computer vision, automatic detection, human face, face features, face candidates

Kazalo

Povzetek	III
Abstract	IV
1 Uvod	1
1.1 Primer aplikacije bančnega avtomata	2
1.2 Sistem za iskanje in razpoznavanje obrazov	3
1.2.1 Zaznavanje in sledenje gibanju	4
1.2.2 Iskanje obrazov	4
1.2.3 Izločanje značilnosti obraza	6
1.2.4 Prepoznavanje obraza	6
2 Dela s podobno tematiko	8
2.1 Iskanje obrazov	8
2.1.1 Pristop K.-K. Sunga in T. Poggia - Cambridge, MIT, ZDA, 1998	8
2.1.2 Pristop A. Rowleya in sod. - Pittsburgh, ZDA, 1998	9
2.1.3 Pristop E. Osune in sod. - Cambridge, MIT, ZDA, 1998	10
2.2 Sledenje gibanju obraza	10
2.2.1 Pristop R. Ferauda in sod. - Telekom Francije, 1997	10
2.3 Izločanje značilnosti obraza	11
2.3.1 Pristop K. C. Yowa in R. Cipolle - Cambridge, VB, 1997	11
2.3.2 Pristop S. Jebara in A. Pentlanda - Cambridge, MIT, ZDA, 1997	12
2.3.3 Pristop N. Oliverja in A. Pentlanda - Cambridge, MIT, ZDA, 1997	13
2.3.4 Pristop K. Sobottke in I. Pitasa - Thessaloniki, Grčija, 1998	14
2.4 Prepoznavanje obraza	16
2.4.1 Pristop L. Wiskotta in sod. - Bochum, Nemčija, 1997	17
2.5 Ugotovitev	17
3 Predlagan algoritem za iskanje obrazov	19
3.1 Opis posameznih korakov algoritma	20

3.1.1	Izločanje nepomembnih barv	20
3.1.2	Sivinska slika	22
3.1.3	Mediana filter	22
3.1.4	Algoritem "Rast regij"	23
3.1.5	Izločanje nepomembnih regij	24
3.1.6	Iskanje robov	24
3.1.7	Houghova transformacija	25
3.1.8	Iskanje najboljše krožnice v regiji	27
3.1.9	Iskanje partnerskih krožnic	27
3.1.10	Potrditev kandidata za obraz	27
3.2	Odraž pragov	28
4	Rezultati	34
4.1	Šibke točke algoritma	39
5	Zaključek in nadaljne delo	41
A	Učna množica	43
B	Demo različica sistema A. Rowleya in sod.	45
	Literatura	47

Slike

1.1	Algoritem sistema za iskanje in razpoznavanje obrazov	4
3.1	Osnovni princip delovanja predlaganega algoritma: 1. vhodna slika, 2. izločene nepomembne barve, 3. slika filtrirana z mediano, 4. segmentirane bele regije, 5. izločene nepomembne regije, 6. najdeni robovi, 7. najboljše krožnice znotraj zanimivih regij, 8. izhodna slika	21
3.2	Od leve proti desni: parametra (polmer r in kot θ), ki določata premico v normalni obliki, šop premic skozi robni element in predstavitev šopa premic v parametričnem prostoru	25
3.3	Od leve proti desni: vhodna slika, najdeni robovi, akumulacijsko polje in izbrane premice na vhodni sliki	26
3.4	Zavrnitev kandidata za obraz, ki vsebuje levo oko leve osebe in desno oko desne osebe	30
3.5	Učinkovita omejitev števila belih regij	31
3.6	Zaradi večje oddaljenosti obrazov v ozadju, algoritem slednjima pripiše manjšo verjetnost, saj sta osvetljena veliko slabše kot obraz v ospredju, pravtako pa sta obraza v ozadju v manjši ločljivosti	32
3.7	Lepo razvidna izločitev regij obrvi	33
4.1	Pozitivni rezultati iz baze obrazov PICS	36
4.2	Negativni rezultati iz baze obrazov PICS - na prvi in zadnji sliki pride do napake zaradi barve ozadja, ki ustreza lastnostim bliskavice, na preostalih treh pa zaradi barve las (za konkreten primer glej naslednjo sliko)	37
4.3	Obrazložitev negativnega rezultata iz baze obrazov PICS: iz zaporedja slik je lepo razvidno, da algoritem obdrži regije, katerih najtemnejša točka je presvetla, da bi regija lahko vsebovala oko, razlog za napako pa je neizločeno ozadje, ki ustreza barvi kože (skrajno desna regija)	37
4.4	Pozitivni rezultati iz baze obrazov M2VTS; iz zadnjih dveh slik je razvidno, da algoritem tolerira tudi manjše odklone od frontalnega položaja	37
4.5	Pozitivni rezultati nad realističnimi animacijami obrazov	38
4.6	Pozitivni rezultati nad slikami znanih osebnosti	38

4.7	Negativni rezultati iz baze obrazov M2VTS (sredinska slika) in nad slikami znanih osebnosti - na prvi sliki pride do napake zaradi zlitja regije očesa z regijo ozadja (obraz najde na osnovi kombinacije oko-obrv), na drugi sliki pride do napake zaradi kompleksnejšega ozadja, na tretji sliki zaradi zaliscev, na četrti in peti pa zaradi prevlekega odklona od frontalnega položaja (na peti tudi zaradi napačne izločitve regije očesa); tukaj moramo vzeti v zakup dejstvo, da 2., 4. in 5. slika ne ustrezajo zahtevam predlaganega pristopa navedenim na začetki poglavja 3	38
A.1	Preostala učna množica	43
B.1	Primeri slik, na katerih sta oba pristopa (pristop A. Rowleya in sod. ter v tem delu predlagan pristop) pravilno označila obraze	46
B.2	Primeri slik, ki dokazujejo večjo fleksibilnost sistema A. Rowleya in sod. (glej sliko 4.7)	46

Tabele

4.1	Rezultati testiranja algoritma nad bazo obrazov PICS	35
4.2	Rezultati testiranja algoritma nad bazo obrazov M2VTS	35
4.3	Rezultati testiranja algoritma nad realističnimi animacijami obrazov	35
4.4	Rezultati testiranja algoritma nad slikami znanih osebnosti; tri slike (2., 4. in 5. slika na sliki 4.7) niso ustrezale zahtevam navedenim na začetku poglavja 3 in zato niso vključene v zgornjih rezultatih . . .	39
4.5	Rezultati testiranja algoritma nad celotno testno množico	39

Poglavje 1

Uvod

Raziskave na področju vida od grških filozofov, prek gestaltske šole psihologije na začetku 20. stoletja, do današnjih dni, kažejo na to, da je proces ustvarjanja mentalne predstave o prizoru, ki ga človek opazuje, lahko pojasnjen s spletom različnih mehanizmov. Dvodimenzionalna slika je le eden izmed osnovnih faktorjev mentalne predstavitve prizora, vendar med vsemi faktorji podaja veliko, če ne kar največ, informacij o opazovanem prizoru. Pri določenih problemih računalniškega vida je tako praktično nesmiselno delati s kompleksnejšimi modeli, če nam že enostavni ponujajo dovolj veliko količino informacij.

Disciplina računalniškega vida skuša oponašati sposobnost človeškega vida. Eden njenih ciljev je iskanje objektov na dvodimenzionalnih slikah. Tako kot človek neznan prizor preučuje na način, da ga "razbija" na manjše dele in primerja z znanimi vzorci iz spomina, tudi postopek iskanja objektov segmentira sliko.

Zaradi vse bolj razširjenih in obsežnih slikovnih podatkovnih baz in "živih" video informacij je njihovo inteligentno oziroma avtomatično preiskovanje izredno pomembno. Ljudje, to je človeški obrazi, so eden od najpogostejših in zelo specifičnih objektov, ki jih želimo poiskati na slikah. O pomembnosti problematike pričajo številne raziskave, ki jih na tem področju izvajajo raziskovalne skupine po svetu.

V svetu računalništva nobeno področje ni majhno v smislu, da ga lahko človek povsem obvlada. Na vsako področje namreč vpliva toliko drugih področij (računalniških in drugih), da enostavno ne moremo potegniti razločne meje med posameznimi področji. Ideja z nekega področja je nemalokrat uporabna tudi drugje, slediti tempu, ki ga narekuje takšno življenje, pa je praktično nemogoče in nesmiselno! Zato so strokovnjaki v današnjem svetu največkrat "ozko usmerjeni", kar ni nujno dobro, je pa potrebno za uspešno prilagajanje okolju v katerem živimo!

Tako je seveda tudi pri sistemih za avtomatsko iskanje človeških obrazov na slikah in njihovimi izvedenkami oziroma nadgradnjami! Veliko avtorjev se že več let ukvarja le z določenim podproblemom takšnega sistema, takšnih podproblemov pa je spet mnogo. V novejših člankih so osnovni algoritmi, ki so uporabni na tem področju, dandanes le bežno omenjeni s kakšnim imenom, sistemi pa so dosegli že kompleksnost "velikega" projekta. Nekateri avtorji so si zastavili nalogo sistema

bolj specifično, drugi spet bolj univerzalno. Rezultati, ki jih dosegaajo posamezni algoritmi (sistemi) so primerljivi med seboj; če upoštevamo njihovo kompleksnost in omejitve. Na prvi pogled bi lahko rekli, da vsi algoritmi dosegaajo podobne rezultate, noben od njih pa še vedno ni “popoln”! Pa vendar, kaj pa sploh razumemo pod pojmom podobnost in popolnost?! Večina razvitih algoritmov ima vsaj enega od dveh najpogostejših problemov:

- prevelika računrska (časovna, prostorska) kompleksnost in/ali
- premajhna učinkovitost.

Na osnovi zgornjih opazk je bila začrtana pot katere rezultat je pričujoče delo. Delo obsega pregled problematike avtomatskega iskanja človeških obrazov na slikah in možnosti izboljšav algoritmov na tem področju. Predlagan je algoritem za avtomatsko iskanja človeških obrazov na slikah, ki poskuša združiti več idej, o katerih so pisali različni avtorji (glej literaturo). Algoritem predpostavlja določene olajševalne okoliščine oziroma omejitve. Podane so tudi možne izboljšave, katerih ideje izvirajo iz negativnih rezultatov testa.

Na tem mestu je potrebno izpostaviti dejstvo, da je avtomatsko iskanje obrazov, kot tudi večina drugih avtomatskih iskanj različnih objektov, zahtevna naloga, predvsem zaradi pomembnih variacij vzorcev, ki jih težko parametriziramo analitično.

Namen algoritmov za avtomatsko iskanje človeških obrazov na slikah je očiten: njihov osnovni cilj je segmentirati sliko na regije, ki vsebujejo obraz ali njegove dele in na regije, ki jih lahko v nadaljni obdelavi slike zanemarimo, saj ne predstavljajo obraza ali njegovega dela. Izpostavimo nekaj aplikacij, ki uporabljajo takšne algoritme:

- najpomembnejšo skupino aplikacij predstavljajo aplikacije za nadzor in varnost,
- aplikacije za video konferenčni sistem,
- aplikacije grafične animacije mimike uporabnika,
- aplikacije oddaljenega vodenja kamere,
- ipd.

1.1 Primer aplikacije bančnega avtomata

S stališča računalniške znanosti beseda prihodnost največkrat izpostavlja avtomatiziranost. To pomeni, da težimo k temu, da stroj, recimo računalnik, opravi večino dela sam.

Včasih smo za dvig denarja vedno morali na banko ter tam zaposlenim zaupati svoje želje in namere. Danes lahko marsikatero bančno storitev opravimo drugače,

preko bančnega avtomata. Aplikacija bančnega avtomata spada v skupino aplikacij za nadzor in varnost.

Bančni avtomati so pospešili poslovanje, razbremenili zaposlene, so (praktično) nezmotljivi, dostopni 24 ur na dan, prijazni do uporabnika, itd. Edina negativna lastnost bančnih avtomatov je neoseben odnos, vendar po drugi strani prav ta zagotavlja učinkovitost.

Bančni avtomat je res majhna avtomatizirana banka, vendar ljudje težimo še k večji stopnji avtomatiziranosti. Danes potrebujemo za uporabo bančnega avtomata dve stvari: bančno kartico in osebno številko. Marsikomu povzročata ti dve stvari preglavice, saj moramo imeti bančno kartico vedno pri sebi, osebno številko pa moramo znati na pamet. Največji problem predstavlja prav osebna številka.

Predstavljajte si naslednjo rešitev: uporabnik stopi do bančnega avtomata, ta ga samodejno prepozna ter ga povpraša po storitvah, ki jih želi opraviti. Problemi, ki so danes vsakdanji - bančna kartica, osebna številka, mogoče celo neoseben odnos - izginejo!

Ne, to ni znanstvena fantastika! Veliko raziskovalnih skupin po svetu se ukvarja s takšnimi problemi. Tudi pričujoče delo je prvi korak k zgoraj opisani rešitvi. Namreč, avtomatsko prepoznavanje uporabnika lahko omejimo na prepoznavanje obraza, saj že same oči vsebujejo več informacije kot prstni odtisi, ta informacija pa je tudi zanesljivejša. Če želimo uporabiti to informacijo moramo najprej lokalizirati oči na sliki, saj je bistvena informacija prav v tem delu obraza. V pričujočem delu s predlaganim algoritmom za avtomatsko iskanje človeških obrazov na slikah poskušamo doseči prav to, torej najti obraz in lokalizirati oči!

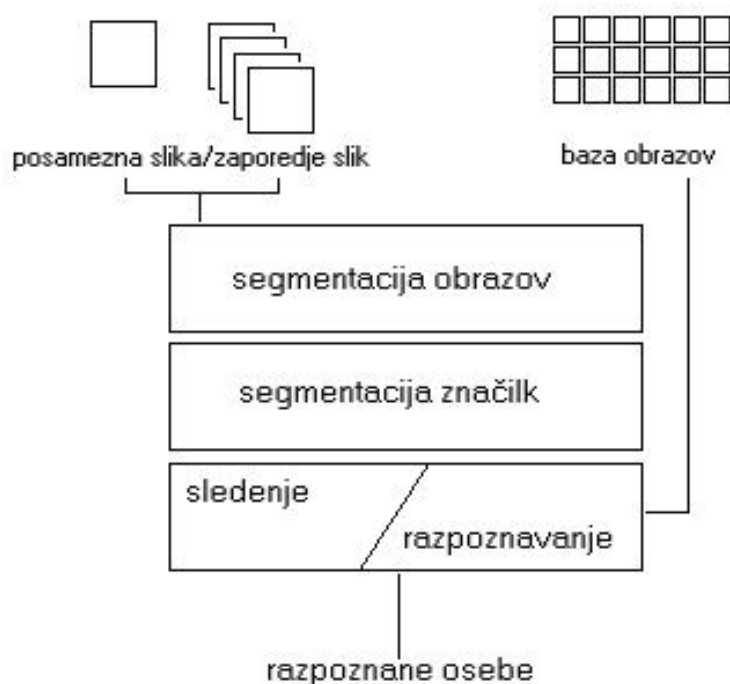
1.2 Sistem za iskanje in razpoznavanje obrazov

V grobem lahko rečemo, da nek sistem za iskanje in razpoznavanje obrazov sestavljajo naslednji deli (slika 1.2):

- zaznavanje gibanja,
- iskanje obraza (obrazov),
- sledenje gibanju obraza,
- izločanje značilnosti obraza in
- prepoznavanje obraza.

Takšen sistem v realnem svetu predstavlja aplikacijo nadzora in varnosti, algoritme takšnega sistema pa lahko uporabljamo še marsikje (npr. pri nadzoru kvalitete izdelkov).

Vsak del je največkrat spet (kompleksna) zbirka algoritmov, ki realizirajo določene zahteve. Velikokrat z algoritmi računalniškega vida v takšnih sistemih sodelujejo tudi različni pristopi k strojnemu učenju (največkrat se uporabljajo nevronske



Slika 1.1: Algoritem sistema za iskanje in razpoznavanje obrazov

mreže). Vendar pa je več kot jasno, da strojno učenje ni vedno na mestu, saj lahko realiziramo določene zahteve bolj učinkovito brez njega!

V naslednjih podpoglavjih so predstavljeni pristopi k zgoraj navedenim podproblemom, poglavje 2 pa nam daje vpogled v kompleksnost posameznega podproblema.

1.2.1 Zaznavanje in sledenje gibanju

Nenapisano pravilo pravi, da moramo pri načrtovanju vsakega postopka izhajati iz dejstva, da so enostavne stvari “zlata vredne” (posnemanje narave in ravnanje po zdravi pameti). Če se tega ne držimo, si že na samem začetku dela na nekem projektu nakopljemo na glavo nepotrebne težave. Šele nato se lotimo problema z orodji, ki nam omogočajo izboljšanje nekega pristopa.

Če upoštevamo to pravilo, potem vemo, da je najenostavnejši pristop k zaznavanju in sledenju gibanja ta, da enostavno naredimo razliko med dvema zaporednima slikama. Z izboljševanjem osnovnega pristopa raste učinkovitost in seveda tudi kompleksnost sistema.

Da realiziramo sledenje gibanju obraza pa moramo seveda najprej rešiti problem iskanja obraza.

1.2.2 Iskanje obrazov

Poznamo kar nekaj očitno različnih pristopov k iskanju obrazov [12, 16]:

- **pristop na osnovi modela (top-down model-based approach)** - Temelji na različnih velikostih modelov obraza. Zaradi učinkovitosti algoritem tega pristopa najprej preišče sliko na podlagi grobega modela, potem pa, ko najde primerno ujemanje modela in vhodne slike, nadaljuje iskanje z manj grobim modelom (Yang, Huang, 1994; Lanitis, 1995). Največkrat pristop predvideva le en model obraza za posamezno velikost (največkrat v frontalnem položaju). Pristop je težko nadgraditi tako, da bi našel tudi obraze v drugačnih (različnih) položajih!
- **pristop na osnovi značilik (bottom-up feature-based approach)** - Pristop izseje iz slike množico bistvenih točk, ki opisujejo obraz in jih združi v kandidate za obraze na podlagi njihovih geometrijskih razmerij (Leung, 1995; Sumi, Ohta, 1995; Yow, Cipolla, 1995)! Ta pristop je enostavno nadgradljiv tako, da bi našel tudi obraze v drugačnih (različnih) položajih, njegova pomankljivost pa je ta, da ne bo delal enako dobro na različno kvalitetnih slikah (večja je ločljivost, večja je kvaliteta), saj se struktura bistvenih točk preveč spreminja, da bi lahko bile uspešno zaznane!
- **pristop na osnovi teksture (texture-based approach)** - Obraze iščemo na podlagi distribucije informacije o sivinah v neki podsliki, pri tem pa uporabljamo Haralickov matrični pristop (Space Gray Level Dependency) [6] (Dai et al., 1994). Tudi ta pristop je težko nadgraditi tako, da bi našel tudi obraze v drugačnih (različnih) položajih!
- **pristop na osnovi nevronske mreže (neural-network approach)** - Celotno sliko razdelimo na regije, te pa predstavimo v neki standardni velikosti oziroma obliki. Tako obdelan del slike pošljemo na vhod nevronske mreže, ki deluje kot filter (Sung, Poggio, 1994; Rowley et al., 1995). Pristop daje (zelo) dobre rezultate nad slikami, ki vsebujejo obraze v frontalnem položaju, učinkovitost pa se poslabša pri ostalih orientacijah obraza. Pristopa (še vedno) ni mogoče nadgraditi tako, da bi našel obraze tudi v profilnem položaju.
- **pristop na osnovi barv (colour-based approach)** - Vsak slikovni element (piksel) označimo na podlagi podobnosti z barvo kože. Na podlagi te transformacije označimo določene regije slike kot obraz, če seveda regija vsebuje primerno oziroma zadostno količino barve kože (Chen et al., 1995; Dai, Nakano, 1996). Zaznamo lahko različne orientacije obraza. Slaba stran pristopa je občutljivost na barvo kože in obliko obraza!
- **pristop na osnovi gibanja (motion-based approach)** - Tu uporabljamo razliko med dvema zaporednima slikama. Tako ločimo dinamično dogajanje od statičnega ozadja. Obraz nato poiščemo na podlagi prisotnih silhuet (Trew et al., 1993) ali barve, ki nam jo da razlika (Schiele, Waibel, 1995). Pomam-

jkljivost tega pristopa je, da ne deluje dobro nad scenami, kjer se premika veliko objektov.

- **pristop na podlagi informacije o globini (depth-based approach)** - V prvem koraku moramo dobiti par stereo posnetkov frontalnega položaja obraza (obrazov). Nato se določi korespondenca med točkami nad širokim intervalom disparitete. Pri tem se uporablja multiresolucijski hierarhični algoritem ujemanja. Nato se na podlagi informacije o globini določijo značilke obraza oziroma obraz.
- Razne kombinacije zgoraj omenjenih pristopov oziroma njihovih delov (tudi z drugimi metodami strojnega učenja [4])!

1.2.3 Izločanje značilnosti obraza

Večina algoritmov za razpoznavanje obrazov oziroma izločanje značilnosti obraza predvideva določene olajševalne okoliščine, ki povzročijo, da iskanje obrazov postane trivialno. Tako so recimo nekateri avtorji v nadaljno obdelavo pošiljali le izrez iz originalne slike, ki je vseboval obraz (Craw et al., 1992; Turk, Pentland, 1991), drugi pa so spet predvideli določene omejitve pri orientaciji obraza ali/in kompleksnosti ozadja [3]. - Problem iskanja obrazov vsekakor ni trivialen, če je ozadje lahko kompleksno, osvetlitev poljubna in orientacija obraza ni vnaprej določena!

Poznamo dva osnovna pristopa k izločanju značilnosti obraza:

- **pristop od spodaj navzgor** - Pri segmentaciji izhajamo iz posameznega slikovnega elementa, le-te pa združujemo v modele, ki opisujejo posamezne značilke obraza.
- **pristop od zgoraj navzdol** - Tukaj si pomagamo z modeli značilk obraza, ki jih primerjamo s posameznimi deli slikovnega prostora.

Opisi značilk so lahko zelo enostavni, naprimer zgolj barvna informacija, lahko pa so tudi zelo kompleksni, naprimer opis s krivuljami.

1.2.4 Prepoznavanje obraza

Poznamo sisteme, ki poskušajo razpoznati obraz na podlagi več slik iste osebe oziroma na podlagi le ene slike. Algoritmi za oba pristopa so lahko zelo različni, ni pa nujno. Na podlagi več slik lahko zgeneriramo 3D model, ki nam nudi veliko več informacij od ene same 2D slike! Takšen sistem pa je najverjetneje s stališča kompleksnosti predrag! Torej trenutno je najprimerneje narediti kompromis, ki zagotavlja kvalitetno 2D informacijo.

Torej kot vhod sistemu podamo sliko ali sekvenco slik. V prvem koraku določimo segmente, ki bi lahko predstavljali obraze, potem pa izločimo značilke obraza. Nato pride na vrsto prepoznavanje, ki se izvrši kot primerjava med izločenimi oziroma

najdenimi značkami in podatkovno bazo sistema. Kot rezultat dobimo identifikacijo oseb(e). Poznani so tudi pristopi, kjer sta prvi in drugi korak zamenjana, tako da se (kandidat za) obraz tvori na podlagi kandidatov za značilke.

Poglavje 2

Dela s podobno tematiko

V tem poglavju je predstavljenih nekaj najzanimivejših (najnovejših) sistemov za avtomatsko iskanje človeških obrazov na slikah ter njihovih izvedenk oziroma nadgradenj. Opis vsakega sistema vsebuje razlago osnovne ideje ter prednosti in slabosti pristopa.

Podpoglavja 2.1, 2.2, 2.3 in 2.4 so urejena tako, da kompleksnost opisanih algoritmov narašča iz podpoglavja v podpoglavje. Prav zaradi tega velikokrat kasnejša podpoglavja vsebujejo tudi opise, ki bi kot zaključene celote morali imeti mesto v podpoglavju z nižjo oznako, vendar zaradi konsistentnosti z literaturo temu ni tako; opis dela enega avtorja oziroma skupine ostaja v enem sklopu.

Pozornost pri opisu je namenjena predvsem iskanju obrazov in značilk obraza.

Med starejšimi sistemi je potrebno omeniti predvsem delo G. Chow in X. Lija [3] ter A. L. Yuilla in sod. [17].

2.1 Iskanje obrazov

2.1.1 Pristop K.-K. Sunga in T. Poggia - Cambridge, MIT, ZDA, 1998

Pristop temelji na strojnem učenju ter omogoča iskanje obrazov v frontalnih položajih (z dovoljeno majhno stopnjo zasuka) na sivinskih slikah s kompleksnim ozadjem [13].

Osnovni postopek je sledeč:

- Nad izsekom oziroma oknom testnega vzorca se najprej izvrši predprocesiranje.
- Na podlagi kanoničnega modela obraza, ki se sestavi na podlagi predhodnega učenja, se naredi primerjava z rezultati prejšnje točke.
- Parametri, ki jih algoritem pridobi na ta način, gredo na vhod nevronske mreže, ki na izhodu pove, ali je v oknu obraz ali ne.

Predprocesiranje se izvrši nad dvema učnima bazama slik (bazo obrazov in skrbno izbrano bazo, v kateri ni obrazov), prav tako pa tudi nad vsakim novim oknom testnega vzorca, ki ga želimo klasificirati.

- Sprememba velikosti. - Vzorec spremeni velikost na 19×19 slikovnih elementov. Algoritem obravnava takšen vzorec kot vektor s 361 komponentami; torej gre za prav toliko razsežen prostor.
- Maskiranje vzorca. - Iz vzorca se odstranijo slikovni elementi ob samih robovih vzorca s pomočjo maske, ki naj bi ohranjala le obraz. S tem je dosežena izločitev nezaželjene napačne klasifikacije zaradi strukture ozadja, zmanjša pa se tudi razsežnost prostora; vektor ima sedaj le 283 komponent.
- Popravek gradienta osvetlitve. - V vzorcu se osvetlijo (zelo) zasenčeni deli, ki so posledica ekstremnih kotov osvetlitve. Operacija odšteje od nemaskiranega okna najustreznejšo ravnino svetlosti.
- Izravnavanje histograma. - Okno postane relativno neodvisno od različnih pogojev osvetlitve, velikokrat pa se izboljša tudi kontrast.

Kanonični model obraza, ali z drugimi besedami, model distribucije vzorcev obraza sestavlja šest grup, ki aproksimirajo distribucijo v 283 razsežnem prostoru. S šestimi grupami je opisan tudi model distribucije vzorcev, na katerih ni obraza. Vsaka grupa je večdimenzionalna Gaussova funkcija s središčem in kovariančno matriko, ki opisuje lokalno distribucijo okoli središča.

Za vsako novo okno testnega vzorca, ki ga želimo klasificirati, se izračuna vektor, ki vsebuje 12 komponent. Vsaka komponenta podaja razdaljo med centrom posamezne grupe kanoničnega modela in centrom posamezne grupe modela, ki ga dobimo iz testnega vzorca. Takšen vektor gre na vhod nevronske mreže - večnivojskega perceptrona.

Avtorja navajata, da se je takšen pristop izkazal za uspešnega tudi pri iskanju značilnk in nalogah razpoznavanja vzorcev v drugih domenah.

Podoben pristop je uporabljen tudi v [9, 11].

2.1.2 Pristop A. Rowleya in sod. - Pittsburgh, ZDA, 1998

Bistvene razlike v pristopu A. Rowleya in sod. [11] glede na pristop opisan v poglavje 2.1.1 so:

- Kot vhod v nevronske mreže se uporabi predprocesiran vhodni vzorec.
- Nevronska mreža je pravzaprav paralelna povezava treh nevronskih mrež, kar naj bi zagotavljalo zaznavanje značilnk obraza. Izhod celotne nevronske mreže poveže ugotovitve vseh treh delov.

- Na osnovi rezultatov več mrež postopek še dodatno izloči napačne klasifikacije. Posamezne mreže so učene na enak način, razlikujejo pa se glede na naključno izbrano začetno inicializacijo uteži, naključno izbrano učno množico slik, na katerih ni obrazov in permutacijo vrstnega reda predstavitve slik.
- Integrirana je heuristika, ki na podlagi ugotovitve, da pristop najde večino obrazov večkrat (na lokacijah, ki so si zelo blizu in na različno velikih vzorcih), še dodatno izloči napačne klasifikacije.

Razvitih je bilo več različic celotnega sistema, nekaj tudi z namenom pospešitve oziroma želje po delovanju sistema v realnem času, seveda na račun slabše klasifikacijske točnosti.

V svetovnem spletu (internet) je na naslovu

<http://www.ius.cs.cmu.edu/demos/facedemo.html>

na voljo demonstracijska različica razvitega sistema. V dodatku B je podanih nekaj rezultatov tega sistema. Vse slike v dodatku B so služile tudi kot testne slike za predlagan algoritem za avtomatsko iskanje človeških obrazov na slikah, ki je opisan v poglavju 3.

2.1.3 Pristop E. Osune in sod. - Cambridge, MIT, ZDA, 1998

Pristop [9] se glede na pristop opisan v poglavju 2.1.1 razlikuje predvsem po spremenjeni metodi učenja nevronske mreže, t.i. SVM (Support Vector Machine).

Pri večini postopkov učenja je cilj minimizirati napačne klasifikacije (t.i. empirično tveganje), pri SVM pa je cilj minimizirati zgornjo mejo napake posplošitve (t.i. minimizacija strukturnega tveganja). Osnovni problem SVM je, da zahteva veliko spomina. Vektorji podpore (support vectors) so le tiste podatkovne točke, ki so bistvene za rešitev problema; gre za točke, ki ležijo na meji med dvema razredoma. Njihovo število je navadno majhno, dokazano pa je, da je to število proporcionalno napaki posplošitve klasifikatorja (V. Vapnik in sod., 1992).

Tudi za klasifikacijo novega vzorca se uporabi SVM.

2.2 Sledenje gibanju obraza

2.2.1 Pristop R. Ferauda in sod. - Telekom Francije, 1997

Tudi ta pristop [5] uporablja nevronske mreže. Bistvena prednost tega pristopa pa je dobro zaznavanje obrazov, ki so od frontalnega položaja obrnjeni proti profilnemu. Zaznavanje je zadovoljivo do kota 50°. Pomemben je tudi podatek delovanja v realnem času, seveda na zmogljivi delovni postaji, na kateri R. Feraud in sod. razvijajo aplikacijo videofona z možnostjo sledenja osebi (sistem LISTEN).

Vhod v nevronske mreže je predprocesirano okno vzorca oziroma slike velikosti 15×20 . Predprocesiranje vsebuje ravnanje histograma (okno postane relativno neodvisno od različnih pogojev osvetlitve), glajenje (odpravljanje šuma) in normalizacija (operacija odštevanja) z vzorcem povprečnega obraza, dobljenega nad učno množico.

Tako dobljeni vhod nato vstopa v 3 neodvisne nevronske mreže (večnivojske perceptrone):

- Dve izmed njih dajo na izhodu verjetnost, da je na vhodu obraz, pri čemer je bila ena naučena z učno domeno obrazov v frontalnem položaju, druga pa z učno domeno obrnjenih obrazov. Podobna ideja je uporabljena tudi v [11] (poglavje 2.1.2).
- Tretja nevronska mreža da na izhodu verjetnost posamezne domene; torej ima dva izhoda. Mreža se nauči iz že naučenih mrež prejšnje točke ter enovite baze vzorcev obrazov (obrazi v frontalnem položaju in obrnjeni obrazi) in vzorcev, na katerih ni obrazov.
- Na podlagi zgornjih verjetnosti se izračuna verjetnost, da je na vhodu obraz.

Večkratna zaznava istega obraza služi kot ocena dejanske prisotnosti obraza na določeni lokaciji.

Takšen pristop k iskanju obrazov je integriran v sistem LISTEN (Locating Individual Speaker and Tracking ENVIRONMENT), ki ima sposobnost lociranja in sledenja človeškemu obrazu ob neznanem ozadju v realnem času. Sistem se lahko nahaja v dveh načinih: načinu iskanja ali sledenja. V načinu iskanja je kamera fiksirana in deluje modul iskanja. Ko je (najbližji) obraz najden, se sistem preklopi v način sledenja in kamera sledi najdenemu obrazu. Pri tem se modul iskanja izključi, modula zaznavanja gibanja in barve kože pa ostajata vključena. Ta dva modula v času iskanja iz slike izločata regije (okna), ki gredo na vhod modula iskanja.

2.3 Izločanje značilnosti obraza

2.3.1 Pristop K. C. Yowa in R. Cipolle - Cambridge, VB, 1997

Da dobimo vpogled v kompleksnost problema za izločanje značilnosti obraza, si oglejmo primer rešitve problema [16]:

- določi zanimive točke (filtriranje),
- iskanje robov,
- ohranitev zanimivih (paralelnih) robov,

- grupiranje točk v kandidate za obraze (obris obraza, obrvi, oči, nosa in ust) in
- izbira resničnih obrazov glede na verjetnost obraza (s pomočjo verjetnostne mreže).

Algoritem temelji na veliki količini geometrijskih in barvnih (sivinskih) informacijah.

Postopek se dobro obnese nad slikami z razgibanim ozadjem in različnimi orientacijami obraza ($\pm 30^\circ$). Slabosti takšnega pristopa pa so problemi kot npr. detekcija majhnih obrazov, včasih se zgodi, da ima večjo verjetnost kandidat, ki ni obraz, šumne slike in premalo podanih značilnosti obraza (narisani obrazi).

Ta pristop k iskanju obraza (in izločanju njegovih značilnosti) je izboljšava t.i. pristopa na osnovi značilk (glej poglavje 1.2.2). Ta pristop je eden najbolj fleksibilnih, saj je "enostavno" nadgradljiv glede na različne velikosti in orientacije obraza.

2.3.2 Pristop S. Jebara in A. Pentlanda - Cambridge, MIT, ZDA, 1997

Sistem [7] v realnem času omogoča avtomatsko iskanje, modeliranje in sledenje značilkam obraza s pomočjo 3D informacije. 3D informacija omogoča obdelavo obrnjenih obrazov (proti profilnemu položaju).

Glavne komponente algoritma za iskanje značilk so:

- Klasifikacija kože. - Nad testno množico obrazov se izvede grupiranje distribucije slikovnih elementov v barvnem prostoru RGB. Cilj tega je najti model verjetnostne porazdelitve barve kože. Testna množica ni vsebovala temnopoltih obrazov.

Ko dobi sistem na vhod novo sliko, se vsak slikovni element slike na podlagi zgornjega modela oceni kot slikovni element kože oziroma kot slikovni element nečesa drugega. Nato se izvrši še grupiranje slikovnih elementov kože.

- Transformacija simetrije. - Sedaj se na osnovi konture obraza določi okno, v katerem lahko pričakujemo oči. S t.i. algoritmom "Transformacije simetrije" se poiščejo oči. Ta algoritem zazna pomembne robne konfiguracije, ki obkrožajo objekt. Kot kandidat za oko se uporabi najtemnejši slikovni element simetričnega objekta.

Usta se poiščejo na osnovi izločitve temnega horizontalnega dela, na osnovi iskanja največjega vertikalnega gradienta med očmi in usti pa se poišče tudi vertikalna lokacija nosu. Problem brkov ni izpostavljen.

- 3D normalizacija. - Ker vsi obrazi niso v frontalnem položaju, se najprej izvrši izkrivljanje obraza v kanoničen frontalen položaj s pomočjo 3D modela. 3D

model je bil zgrajen predhodno iz baze. Sedaj se izvrši poravnavanje značilk modela in najdenih značilk na vhodni sliki.

Ko je najdeno najboljše ujemanje vseh štirih značilk, se tekstura obraza iz vhodne slike prenese na 3D model. Model se nato obrne v frontalen položaj. Sedaj se iz modela naredi 2D slika obraza v frontalnem položaju (t.i. popačen obraz). S pomočjo operacije ujemanja histogramске informacije leve in desne strani tako dobljenega obraza z željenim histogramom, se izvrši še normalizacija osvetlitve.

- Analiza v prostoru obrazov. - Prostor obrazov je vnaprej podana baza obrazov z ročno določenimi značilkami. S projekcijo popačenega obraza v ta prostor, se aproksimira njegova razdalja od učne množice obrazov.

Do sedaj algoritem pozna lokacijo oči, ust ter vertikalo na kateri je lokacija nosu. Do točne horizontalne lokacije nosu pride s pomočjo 12. različnih normalizacij in projekcij v prostor obrazov, narejenih vzdolž horizontalne linije spodnjega dela nosu. Za lokacijo nosu se sedaj izbere tista lokacija nosu iz množice 12. popačenih obrazov, ki ima minimalno razdaljo od učne množice obrazov.

Ko so lokacije značilk znane, se prične sledenje na osnovi 2D SSD korelacijskih krp. Pristop sledenju značilk na osnovi korelacije se začne s sledenjem v smislu najbližje soseščine, kjer se lokalno iščejo značilke. Takšno sledenje samo po sebi ni zmožno zaznavati nelinearnih sprememb (npr. iz slike izgine del obraza, obračanje obraza proti profilu,...), zato si algoritem za sledenje pomaga s 3D informacijo. Pomaga si tudi s parametrično opisanim prostorom 3D (lasersko skeniranih) glav. Tako preprečuje ocenjevanje oziroma obdelavo napačnih 3D oblik. Končna filtrirana 3D struktura obraza in 3D ocena poze obraza služita kot kontrola 2D sledenja v naslednji iteraciji sistema, s tem pa je rešenih nekaj omejitev 2D predstavitve.

2.3.3 Pristop N. Oliverja in A. Pentlanda - Cambridge, MIT, ZDA, 1997

V [8] je predstavljen sistem, ki v realnem času sledi, opisuje oblike in klasificira izraze človeških obrazov in ust. Uporablja se informacija o 2D značilnostih regij. Regije so grupe slikovnih elementov, ki si delijo lastnosti slike na nizkem nivoju. Prav pristop z regijam omogoča delovanje v realnem času.

Bistveni deli sistema, poimenovanega LAFTER (Lips And Face TrackER), so:

- Inkrementalna metoda Maksimizacije pričakovanja. - Metoda zagotavlja adaptivno in nenehno izboljševanje opisa regije (modela). S tem postaja opis odpornejši proti spremembam v osvetlitvi in različnim barvam polti (med barvami polti ni omenjene temne polti). Začetni model verjetnostne porazdelitve barve kože se dobi po isti metodi pred samim začetkom delovanja sistema.

- Nov pristop opisa regij. - Temelji na ideji najkrajšega opisa (MDL - Minimum Description Length).
- Metoda klasifikacije podatkov o obliki oziroma izrazu ust s pomočjo HMM (Hidden Markov Method).

Pristop k problemu iskanja obraza uporablja grobo informacijo o barvi, velikosti in obliki regije. Za segmentacijo se uporabi algoritem "Rast regij", ki je optimiziran tako, da dosega veliko večjo hitrost kot standarden algoritem. Do končne odločitve o prisotnosti obraza pride s pomočjo informacije o velikosti in obliki regije, ki se primerja z informacijo o kanonični velikosti in obliki obraza. Kot informacija o obliki služi štirikotnik okoli regije. S pomočjo Kalmanovega filtra sistem vzdržuje oceno središč in oblik vseh regij, kar izkorišča za to, da kamera sledi gibanju tako, da je obraz vedno v sredini in enake velikosti.

Ko je znana informacija o lokaciji obraza (središče, širina, višina, kot zasuka) se s pomočjo statistike določi štirikotnik, znotraj katerega mora sistem poiskati usta. Tudi tukaj se uporablja isti princip segmentacije. Ker pa je ta informacija na tem mestu nezadostna, se uporabi tudi informacija o regiji, ki obkroža usta. Ob vsakem iterativnem koraku se zaradi učinkovitosti izračuna središče ust. Oblika ust je karakterizirana s širino in višino ter štirikotnikom okoli njih. Takšen vektor parametrov se uporablja pri klasifikaciji izraza na človeškem obrazu. Neobčutljivost glede na kot zasuka obraza (nagnjenost glave) je dosežena z izračunom kota nagnjenosti ter nato zasuka regije. Usta so torej vedno bolj ali manj v horizontalni legi.

Z uporabo opisanega vektorja parametrov, ki opisuje izraz ust, je bila naučena HMM. Naučila se je razlikovati med petimi osnovnimi izrazi: normalen, smejoč, žalosten, odprta usta, obširen smeh (smeš in odprta usta.)

Avtorja izpostavljata dejstvo, da je s tem pristopom možno (prvič) v realnem času brati z ustnic in razpoznavati izraze v neomejenem okolju pisarne.

Sistem je integriran v več aplikacij: aplikacijo za video konferenčni sistem, aplikacijo grafične animacije mimike uporabnika, aplikacijo oddaljenega vodenja kamere,...

2.3.4 Pristop K. Sobottke in I. Pitasa - Thessaloniki, Grčija, 1998

Avtorja [12] predstavita rešitvi prvih dveh korakov problema prepoznavanja obrazov, torej iskanja obrazov in izločanja značilk, predlagata pa tudi rešitev za sledenje obrazu in značilkam.

Pristop izkorišča dejstvo, da so obrazi zelo dobro opisljivi z barvo kože in ovalno obliko. Pri izločanju značilk se uporabi informacija o sivinskih nivojih znotraj regij, ki naj bi predstavljale obraz(e).

Postopek je sledeč; na vhodu imamo sivinsko sliko:

- Najprej se izvrši segmentacija na osnovi barv. Kot osnovni barvni prostor je vzet HSV (Hue-Saturation-Value), saj je kompatibilen s človeškim zaznavanjem barv. Dobimo dvobitno sliko.
- Najboljša aproksimacija ovalne oblike obraza je elipsa. Segmentacijo elips lahko naredimo na podlagi robov ali regij. Ta postopek uporablja informacijo o regijah, saj naj bi le-ta zagotavljala večjo učinkovitost pri boju zoper šum in spremembam osvetlitve.
 - Postopek najde najprej povezane komponente. Uporabi se algoritem “Rast regij”.
 - Nato preveri ali je vsaka povezana komponenta približno enaka elipsi. Najbolje se prilegajočo elipso najde na podlagi momentov. Za opis elipse potrebujemo pet parametrov: središče elipse je težišče povezane komponente, do orientacije elipse pride preko najmanjšega vztrajnostnega momenta, podobno pa preko vztrajnostnih momentov dobi tudi vrednosti obeh polosi.
 - Na podlagi znanih parametrov se izvrši redukcija kandidatov za obraze. Upošteva se orientacija in razmerje med polosema.
 - Kandidate za obraze, ki so ostali, oceni na podlagi ujemanja povezane komponente in najboljše prilegajoče se elipse. Uporabi se prag, ki število kandidatov za obraz še zmanjša.
- Kandidate za obraz preveri z iskanjem značilnk znotraj povezanih komponent; nad vhodno sliko. Pri tem se naslanja na ugotovitev, da se značilke obraza (obrvi, oči, nosnici, usta in brada) ločijo od ostalega obraza po (izrazito) temnejši barvi. - Le kako bi se takšen pristop obnesel pri obdelavi obrazov temne polti!?!?
 - Na podlagi zgornje ugotovitve se temnejši deli še poudarijo s pomočjo dveh operacij morfiranja: sivinske erozije in izboljšanja kontrasta.
 - Obraz opiše kot ozvezdje minimumov in maksimumov sivinskega reliefa (uporaba histogramске informacije). Ker so vse značilke obraza orientirane horizontalno, se najprej izvede normalizacija orientacije povezane komponente, če je ta seveda potrebna.
 - * Izračuna se y -projekcija (histogram) sivinskega reliefa - za vsako vrstico povezane komponente se določi povprečna vrednost sivine v vrstici. Tako dobljeni rezultat zgladi s filtrom povprečja velikosti 3×3 . Pomembni minimumi so določeni tako, da se preveri gradient vsakega minimuma z najbližjim maksimumom. Minimum naj bi predstavljal značilko.
 - * Za tako dobljene minimume se izračunajo x -projekcije (histogrami) sivinskega reliefa. Povprečijo se vrednosti sivin treh vrstic za vsako

kolono. Tako dobljeni rezultat se spet zgladi s filtrom povprečja velikosti 3×3 , določijo pa se tudi pomembni ekstremi, ki naj bi predstavljali posamezno značilko.

- Do ocene, kako dobro par značilk (minimumov ali maksimumov) ustreza zahtevam, pride na osnovi teorije mehke logike. Za vsako zahtevo je podana t.i. funkcija članstva, ki z določeno stopnjo verjetnosti oceni, kako dobro je posamezni zahtevi ugodeno. Na podlagi praga se opravi redukcija kandidatov za značilke.
- Izvrši se združevanje v grupe glede na koordinate levega in desnega minimuma oziroma maksimuma. Pri tem izhajamo iz x -projekcij. Smisel tega je grupiranje podobnih kandidatov. Do grup pridemo s pomočjo Nenadzorovanega min-max algoritma.
- Izbere se najboljše ozvezdje značilk, ki opisuje obraz. Vsaka od možnosti se oceni na podlagi vertikalne simetrije ozvezdja, razdalj med posameznimi značilkami in prej izračunane ocene posameznega kandidata za obraz.

Parametri barvnega prostora HSV so bili nastavljeni tako, da je sistem zaznal le obraze svetle (bele, rumene) polti, kako dobra je segmentacija obrazov druge polti pri teh parametrih, pa še ni znano.

Pristop najde tudi obraze v profilu; slikovnega dokaza v delu ni. Večjo težo posveča iskanju obrazov z vsemi značilkami. To je doseženo tako, da se število najdenih značilk pomnoži z utežjo.

Avtorja pravita, da so rezultati (zelo) zadovoljivi. - Do problemov je prišlo pri naslednjih primerih:

- če so najdene le značilke oči ali obrvi,
- če so kandidati za značilke raztreseni široko okoli prave pozicije, se lahko izbere napačno središče grupe,
- oblika pričeske lahko vpliva na rezultat in
- v primeru (popolne) poraščenosti obraza.

Ko je obraz segmentiran in so poiskane tudi značilke, lahko sledimo gibanju. Sledenje obrisu obraza je rešeno s pomočjo aktivnih (deformirajočih se) krivulj (t.i. kač). Krivulja se spreminja na podlagi barvnih karakteristik. Sledenje značilkam pa je rešeno s pomočjo ujemanja blokov.

2.4 Prepoznavanje obraza

Potencialni uporabniki nekega sistema so vedno vsaj malo skeptični glede podanih rezultatov posameznih avtorjev, zato si želijo, da bi sisteme testirala neka neodvisna

skupina, ki naj ima zelo dodelan pristop k ocenjevanju. Iz ocene naj bo razvidna tako primerjava kot tudi smernice za izboljšanje postopkov.

FERET (FAce REcognition Technology) [10] je program, ki oceni različne pristope k prepoznavanju obrazov. Zaupanje v oceno je doseženo z ogromno, neodvisno podatkovno bazo in dodelano metodo vrednotenja učinkovitosti posameznega sistema. Avtorji FERETA testirajo pristope približno na leto in pol, testirajo pa pristope avtorjev, ki se samostojno odločijo za sodelovanje. Ob gradnji sistema za razpoznavanje je to vsekakor zelo zanimiva referenca.

2.4.1 Pristop L. Wiskotta in sod. - Bochum, Nemčija, 1997

Zelo zanimiv in uspešen pristop temelji na teoriji grafov [14]. Sistem uporablja v osnovi naslednji postopek:

- iz slike izloči obraz in ga pretvori v neko standardno obliko (množica bistvenih oziroma opisnih točk),
- obraz opiše z (neusmerjenim) grafom, kjer so vozlišča prej določene bistvene točke,
- vsako vozlišče hrani nekakšen sklad možnih objektov (npr. nosov),
- algoritem ob primerjavi oziroma razpoznavi izbere najboljše možnosti iz posameznih skladov in na podlagi ujemanja razpozna osebo.

Sistem, potem ko zgradi graf nad slikovno bazo vseh oseb in sliko osebe, ki jo mora prepoznati, potrebuje malo računske zmogljivosti, saj prepoznavanje poteka kar na osnovi primerjave med zgrajenima grafoma oziroma podobnostih med posameznimi vozlišči grafa.

Takšen sistem je zelo fleksibilen, namenjen pa je prepoznavanju oseb znotraj znanega, torej vnaprej podanega, razreda. V primerjavi s sistemi, ki uporabljajo nevronske mreže, ta sistem ne potrebuje nobenega dodatnega učenja, če slikovno bazo oseb razširimo!

S pomočjo iste teorije in pristopa lahko sistem ugotavlja tudi spol osebe in prisotnost očal in/ali brade [15].

Zanimivi so tudi pristopi, ki temeljijo na 3D informaciji [1].

2.5 Ugotovitev

Glede na vrstni red zgornjih podpoglavij je zanimiva ugotovitev, da se pristop z nevronskimi mrežami uporablja le v sistemih za golo iskanje obrazov, nobeno delo pa ne raziskuje možnosti nadgradnje, čeprav je nadgradnja mogoča; v [13] je omenjeno, da se je takšen pristop izkazal za uspešnega tudi pri iskanju značilnih in nalogah razpoznavanja vzorcev v drugih domenah.

Značilke na podlagi predprocesiranega vhodnega okna in izhoda nevronske mreže, ki pove, ali okno vsebuje obraz ali ne, ni težko najti. Ena od idej je ta, da si pomagamo s t.i. povprečnim obrazom. Med njim in vhodnim oknom se nato naredi primerjava, katere rezultat so posamezne značilke obraza. Na ta način so značilke označene, vendar niso opisane tako, da bi ta opis lahko služil kot pomemben parameter za razpoznavanje obraza. Za rešitev tega problema je torej nujen drugačen, mogoče dodaten, pristop.

Druga zanimiva ugotovitev pa je ta, da noben sistem, razen sistemi, ki temeljijo na nevronskih mrežah¹, ne išče oziroma ne najde obrazov temne polti. V slikovnem materialu del so le obrazi svetle polti.

¹Pri nevronskih mrežah se problemu temne polti lahko izognemo z ustreznim predprocesiranjem (glej poglavje 2.1.1).

Poglavje 3

Predlagan algoritem za iskanje obrazov

Algoritem združuje ideje pristopa na osnovi značilik in pristopa na osnovi barv (glej poglavje 1.2.2). Osnovni omejitvi predlaganega algoritma tako izvirata že iz omejitev omenjenih pristopov:

- vhodna slika mora biti v dovolj veliki ločljivosti; obraz mora biti dovolj velik in
- občutljivost na barvo kože.

Osnovna ideja algoritma: na sliki najdi takšne regije, ki vsebujejo kandidate za oko, nato na podlagi geometrijskih značilnosti obraza poskusi združiti kandidate v par oči in končno kandidata za obraz potrdi ali zavrne na podlagi barvne informacije obraza.

Zasnovan je nad množico dokaj različnih slik, t.i. učno množico (pri tem ni uporabljeno strojno učenje). Cilj algoritma je bil doseči maksimalno klasifikacijsko točnost nad slikami, ki ustrezajo naslednjim zahtevam oziroma omejitvam (poleg zgoraj omenjenima):

- delovanje v realnem času na danes standardnem osebem računalniku,
- enostavno ozadje,
- enakomerna osvetlitev slike (ateljejska osvetlitev),
- obrazi svetle polti, ki morajo biti v celoti na sliki (frontalen položaj) in
- obrazi obrnjeni največ do 30° .

Učinkovitost algoritma je bila testirana nad neodvisno množico slik, t.i. testno množico.

Algoritem je nastajal korak za korakom, kjer je posamezen korak predstavljal neko zaključeno celoto. Vsak korak algoritma svoj rezultat zapiše v sliko. Osnovni princip delovanja algoritma je prikazan na sliki 3.1, ki pripada učni množici.

Pragovi integrirani v algoritem igrajo pomembno vlogo pri ugotavljanju njegove učinkovitosti. Zastavljeni so dokaj ohlapno (tolerantno), učinkoviti pa postanejo kot zaporedje. Vsi pragovi so bili določeni eksperimentalno nad učno množico.

Predlagan algoritem za iskanje obrazov

Vhod: slika v formatu BMP

Izhod: označeni obrazi na vhodni sliki

Osnovni koraki algoritma so:

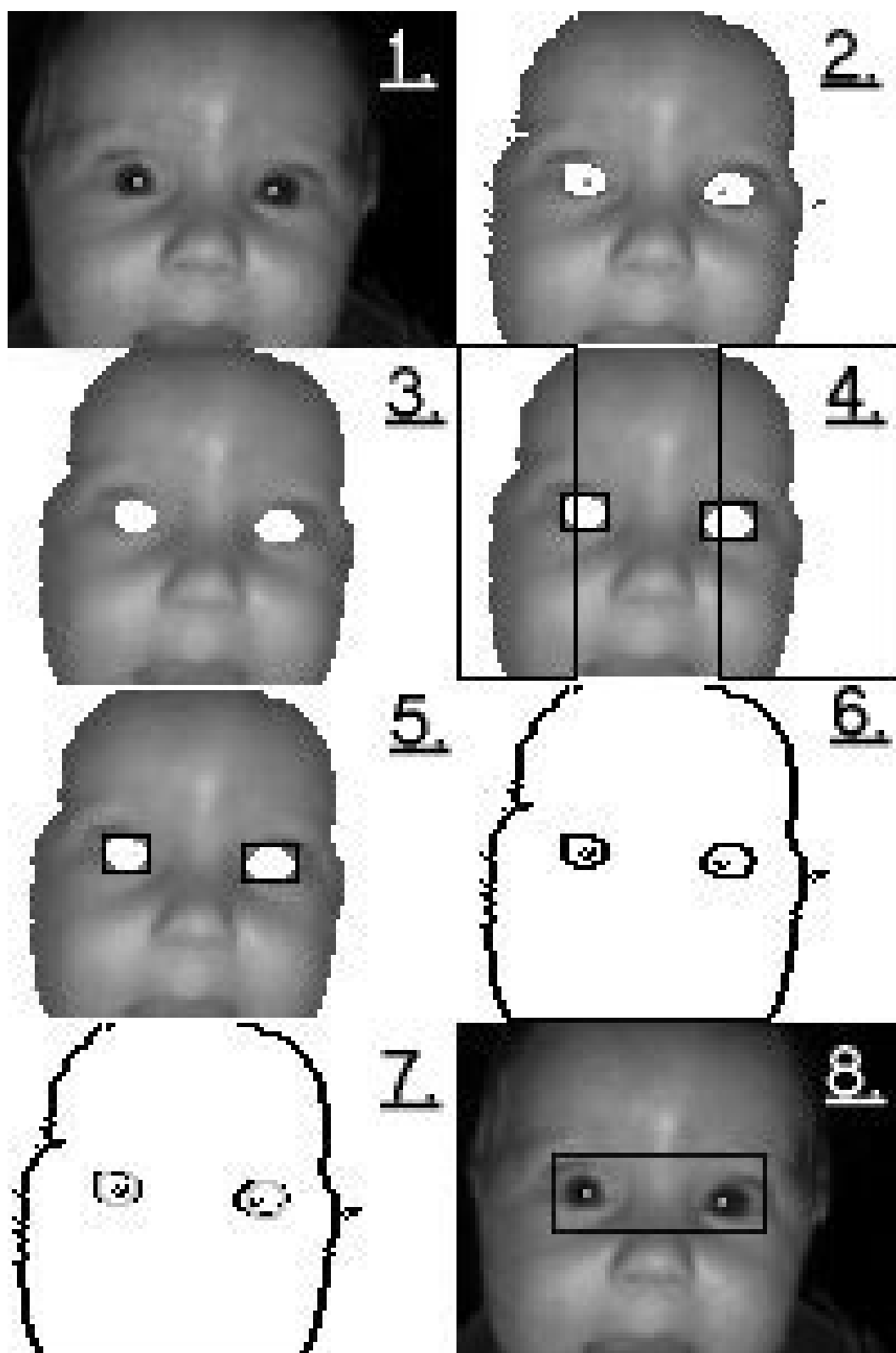
1. Iz barvne slike algoritem izloči popolnoma nepomembne barve (tiste, ki nikakor ne morejo predstavljati obraza). Nepomembne barve zamenja z belo barvo.
2. Tako obdelano sliko spremeni v sivinsko sliko (format PGM).
3. Sliko filtrira z mediano.
4. S pomočjo algoritma "Rast regij" izvrši segmentacijo belih regij.
5. Izloči regije, ki ne morejo vsebovati očesa.
6. Poišče robove na sliki s pomembnimi sivinami.
7. Znotraj ohranjenih regij izvrši iskanje krožnic (kandidatov za oko) nad najdenimi robovi s pomočjo Houghove transformacije.
8. Za vsako regijo najde najboljšo krožnico.
9. Na podlagi geometrijskih značilnosti obraza išče ustrezne partnerske krožnice.
10. Za potrditev kandidata za obraz se uporabi barvna informacija obraza. Na podlagi te informacije (grobno) oceni tudi verjetnost kandidata.

3.1 Opis posameznih korakov algoritma

3.1.1 Izločanje nepomembnih barv

Popolnoma nepomembne barve so tiste, ki nikakor ne morejo predstavljati obraza. Nepomembne barve algoritem zamenja z belo barvo.

Zaradi zgornje definicije nepomembnih barv se učinkovitost algoritma močno zmanjša, saj se omeji le na iskanje obrazov svetle polti. (Podoben pristop bi se sicer lahko uporabil tudi pri iskanju obrazov temne polti.)



Slika 3.1: Osnovni princip delovanja predlaganega algoritma: 1. vhodna slika, 2. izločene nepomembne barve, 3. slika filtrirana z mediano, 4. segmentirane bele regije, 5. izločene nepomembne regije, 6. najdeni robovi, 7. najboljše krožnice znotraj zanimivih regij, 8. izhodna slika

Upragovljanje barvne informacije posameznih slikovnih elementov se je izkazal za vse prej kot stabilen postopek, zaradi enostavnosti pa je vsekakor uporaben pri obdelavi slik, ki ustrezajo določenim zahtevam. Za večjo univerzalnost postopka bi morali uporabiti kompleksnejši oziroma stabilnejši postopek (npr. model verjetnostne porazdelitve barve kože [7]).

Barva kože je določena z naslednjim izrazom (oznake R , G in B so komponente barvnega prostora RGB):

```
% Koža ob enakomerni osvetlitvi =>
R > 95 IN G > 40 IN B > 20 IN
max{R, G, B} - min{R, G, B} > 15 IN
    % komponente RGB ne smejo biti blizu skupaj - izločevanje sivin
|R - G| > 15 IN
    % tudi R in G komponenti ne smeta biti blizu skupaj -
    % ne gre za svetlo polt
R > G IN R > B    % R je največja komponenta
```

ALI

```
% Koža ob uporabi bliskavice ali (rahle) osvetlitve s strani =>
R > 220 IN G > 210 IN B > 170 IN
|R - G| ≤ 15 IN    % R in G komponenti sta blizu skupaj
R > B IN G > B    % B je najmanjša komponenta
```

3.1.2 Sivinska slika

Sliko, ki vsebuje le pomembne barve, algoritem spremeni v sivinsko sliko (format PGM). To ne pomeni, da barvne informacije ne bomo več potrebovali - sprememba se naredi le zaradi poenostavitve naslednjih korakov algoritma.

Sivinsko sliko dobimo po formuli (oznake R , G in B so komponente barvnega prostora RGB):

$$sivina = \sqrt{(R^2 + G^2 + B^2)}/3.$$

Formulo uporabimo nad posameznimi slikovnimi elementi slike.

3.1.3 Mediana filter

Mediana filter je nelinearen filter. To pomeni, da velja:

$$mediana(A(x) + B(x)) \neq mediana(A(x)) + mediana(B(x)).$$

Filter ohranja detajle (ne gladi robov), popravi t.i. sol in poper šum, slabo pa odpravi Gaussov šum.

Slika se filtrira s filtrom oziroma oknom velikosti 3×3 . Znotraj okna je delovanje sledeče:

- slikovni elementi se uredijo po velikosti glede na vrednost sivine,
- sredinski slikovni element okna dobi vrednost sivine, ki je v sredini urejenega seznama.

Takšen postopek se izvrši nad celotno sliko (okno premikamo preko celotne slike), zanemarijo se le robni slikovni elementi (ostanejo nespremenjeni).

Algoritem s tem doseže zabrisanje nepomembnih belih segmentov, večjo kompaktnost zanimivih belih segmentov, izloči pa tudi slikovne elemente oziroma manjše segmente barve kože, ki so obkroženi z belimi slikovnimi elementi.

3.1.4 Algoritem “Rast regij”

Algoritem uporabimo za segmentacijo belih regij. Bele regije predstavljajo regije izločenih, nepomembnih barv, med njimi tudi regij, ki naj bi vsebovale oči. Posamezne bele regije loči med seboj na podlagi t.i. barvanja, saj vsaki regiji določi številko, ki predstavlja njeno barvo.

Psevdo koda osnovnega algoritma:

```
barva=1
za vsak slikovni element  $T = (x, y)$  slike (razen za robne elemente) delaj
  če element  $T$  bele barve, potem
    če element nad elementom  $T$  bele barve in element levo od elementa  $T$  druge barve, potem
      element  $T$  pobarvaj z isto barvo kot je pobarvan zgornji element
    če element levo od elementa  $T$  bele barve in element nad elementom  $T$  druge barve, potem
      element  $T$  pobarvaj z isto barvo kot je pobarvan levi element
    če sta element levo od elementa  $T$  in element nad elementom  $T$  bele barve, potem
      element  $T$  pobarvaj z isto barvo kot je pobarvan levi element
    zagotovimo ekvivalentnost barv leve in zgornje regije
    če sta element levo od elementa  $T$  in element nad elementom  $T$  ne bele barve, potem
      element  $T$  pobarvaj z barvo  $barva$ 
   $barva = barva + 1$ 
```

Na izhodu dobimo pobarvane regije, ki hranijo poleg barve (torej številke barve) še naslednje podatke:

- število belih elementov v regiji in
- koordinati, ki opisujeta okvir regije.

Število možnih barv oziroma različnih regij na sliki je v implementaciji omejeno navzgor s številom 256, v splošnem pa bi lahko vzeli poljubno število. Iz rezultatov je razvidno, da je po izvršitvi algoritma “Rast regij” iz slike vedno segmentiranih manj regij od 10% omenjenega števila. Omenjeno število pa mora biti veliko večje od števila segmentiranih regij zato, ker algoritem vnaprej ne ve, ali se bosta dve regiji zlili v eno ali ne. Tako lahko na primer algoritem na sredini opravljenega dela uporablja dosti večje število barv kot na koncu.

3.1.5 Izločanje nepomembnih regij

Nepomembne regije so definirane kot tiste, ki ne morejo vsebovati očesa. Izločijo se regije, ki ustrezajo vsaj enemu od spodnjih pogojev:

- če je razmerje med številom belih elementov znotraj regije in površino okvirja regije manjše od 45%,
- če okvir regije zavzema več kot 40% celotne slike,
- če je širina okvirja regije manjša od višine, velja pa še vsaj eden od pogojev:
 - če je razmerje med številom belih elementov znotraj regije in površino okvirja regije manjše od 66%,
 - če je razmerje med višino in širino okvirja regije večje od 2,
- če je razmerje med širino in višino okvirja regije večje od 2,5 in velja, da je površina okvirja regije večja od 4,
- če sta širina in višina okvirja regije večja od 10 in velja, da je vsaj eden izmed kotnih slikovnih elementov okvirja regije bel (namreč, če je regija dovolj velika in vsebuje oko, potem v nobenem kotu okvirja regije ni belega slikovnega elementa) in
- če je razdalja med središčema dveh regij manjša kot 10, ne obdrži zgornje regije oziroma raje kot izločitev zgornje, ne obdrži tiste z manjšo površino okvirja regije, če je razlika v višini manjša ali enaka 3 (s tem izločimo regije obrvi ter manjše regije, ki lahko predstavljajo šum ob regiji očesa).

Hevristična pravila so bila sestavljena na podlagi značilnosti učne množice. Najprej se je avtomatsko, torej s programom, obdelala željena informacija na sliki, nato pa se je ta informacija uporabila pri ročni izpeljavi pogojev, pri tem pa so se upoštevale tudi značilnosti obraza. Pogoji so dodatno omiljeni tako, da se od konkretno dobljenih števil dopušča odstopanje do 10%.

3.1.6 Iskanje robov

Ker so iz slike izločene nepomembne barve, je algoritem iskanja robov zelo enostaven in vendar učinkovit. Bistveni robovi za algoritem so tako ali tako le tisti, ki naj bi opisovali oči.

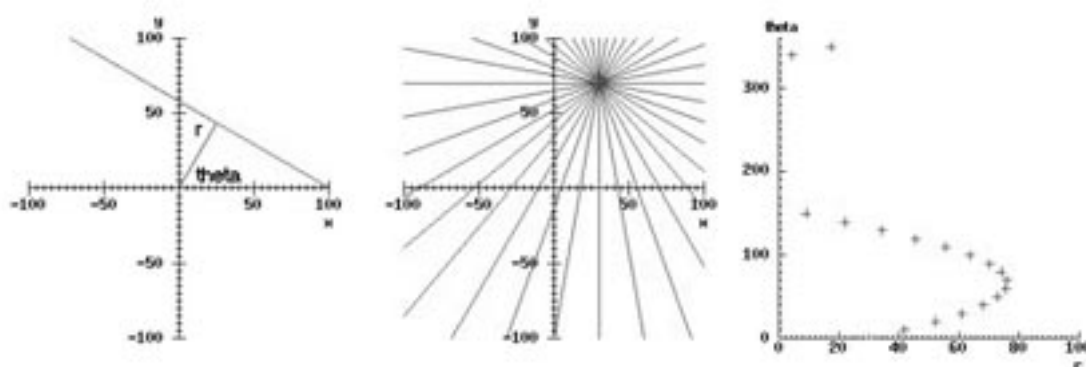
Postopek iskanja robov:

- Najprej sivinsko sliko z obdržanimi pomembnimi barvami (sivinami) premakne za en slikovni element v smeri stranske diagonale, nato pa najde rob na osnovi primerjave istoležnih slikovnih elementov v originalni in premaknjeni sliki. Kot rob označi tisti slikovni element, katerega absolutna vrednost razlike istoležnih slikovnih elementov je večja od 90.

- Enak postopek uporabi tudi nad premaknjeno sliko v smeri glavne diagonale.
- Rezultata obeh zgornjih korakov združi v eno sliko (odebeljitev robov).

3.1.7 Houghova transformacija

Houghova transformacija je uporabna pri iskanju vseh krivulj, ki se dajo zapisati kot parametrična funkcija. Zaradi lažjega razumevanja si oglejmo, kako se s pomočjo te transformacije iz slike izločijo linearni segmenti.



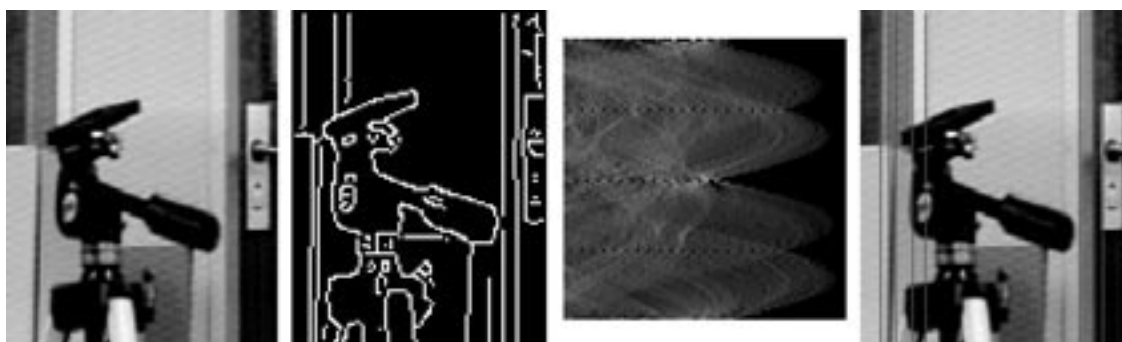
Slika 3.2: Od leve proti desni: parametra (polmer r in kot θ), ki določata premico v normalni obliki, šop premic skozi robni element in predstavitev šopa premic v parametričnem prostoru

Premico zapišemo v normalni obliki:

$$r = x \cos \theta + y \sin \theta.$$

Imamo dva parametra (polmer r in kot θ), ki določata parametrični prostor. Točka (r, θ) v parametričnem prostoru določa premico. Da lahko poiščemo linearne segmente, moramo poiskati vsem možnim premicam skozi robne elemente (točke, ki sestavljajo najdene robove na sliki) točke v parametričnem prostoru. Pri reševanju uporabimo t.i. akumulacijsko polje. Element akumulacijskega polja ustreza majhnemu kosu parametričnega prostora. Ob inicializaciji ima vsak element akumulacijskega polja vrednost 0. Vsako točko, ki pripada določeni premici, zabeležimo v parametričnem prostoru tako, da povečamo vrednost ustreznega elementa akumulacijskega polja. Razdelitev parametričnega prostora v elemente akumulacijskega polja imenujemo kvantizacija. - Slika 3.2.

Element akumulacijskega polja, ki na koncu vsebuje veliko vrednost, ustreza premici na sliki, na kateri leži linearni segment. Glede na uporabljen kvantizacijo, so tudi dobljeni parametri premice ustrezni približki resničnih parametrov. Ker je v primeru iskanja linearnih segmentov akumulacijsko polje dvorazsežno, ga lahko prikažemo kot sivinsko sliko, s katere so lepo razvidni vrhovi akumulacijskega polja - zanimive premice. - Slika 3.3.



Slika 3.3: Od leve proti desni: vhodna slika, najdeni robovi, akumulacijsko polje in izbrane premice na vhodni sliki

Psevdo koda iskanja linearnih segmentov:

```

za vsak robni element  $T = (x, y)$  slike delaj
  za  $\theta = 0$  do 360 delaj
     $r = x \cos \theta + y \sin \theta$ 
    zaokroži  $r$ 
    povečaj vrednost elementa  $(r, \theta)$  v akumulacijskem polju
  
```

Bistveni problemi Houghove transformacije so:

- kvantizacija - Večje je akumulacijsko polje, boljši je približek krivulje, večja pa je časovna kompleksnost postopka, in obratno, manjše je akumulacijsko polje, slabši je približek krivulje in manjša je časovna kompleksnost.
- prag - Kje postaviti mejo, ki določa, katere krivulje vsebujejo dovolj elementov?
- večparametrične krivulje - Akumulacijsko polje ima več dimenzij, kar občutno poveča časovno kompleksnost postopka.

V algoritmu je uporabljena Houghova transformacija za iskanje krožnic. Čeprav je akumulacijsko polje trirazsežno, je hitro izvajanje zagotovljeno tako, da se izvrši transformacija le nad zanimivimi regijami slike.

Psevdo koda iskanja krožnih segmentov v algoritmu:

```

za vsak robni element  $T = (x, y)$  regije delaj
  za  $x_0 = 1$  do širine regije delaj
    za  $y_0 = 1$  do višine regije delaj
       $r = \sqrt{(x - x_0)^2 + (y - y_0)^2}$ 
      zaokroži  $r$ 
      če  $r \leq \min\{\frac{\text{širine regije}}{2}, \frac{\text{višine regije}}{2}\}$ , potem
        povečaj vrednost elementa  $(r, x_0, y_0)$  v akumulacijskem polju
    
```

Predzadnja vrstica pove, da iščemo le krožnice, katerih polmer je največ polovica širine oziroma višine regije. S tem je zagotovljeno, da je krožnica v večini vsebovana v regiji, hkrati pa omejimo velikost akumulacijskega polja. Zaradi majhnega prostora problem kvantizacije ne pride do izraza. Posamezna koordinata v akumulacijskem polju tako ustreza velikosti enega slikovnega elementa.

3.1.8 Iskanje najboljše krožnice v regiji

Tudi problemu praga pri Houghovi transformaciji se poskušamo izogniti tako, da omejimo polmer krožnice navzgor.

Po opravljeni transformaciji, algoritem poišče krožnico z največ elementi, nato pa poišče povprečno središče krožnic, ki imajo takšno število elementov in seveda enak polmer. Tako dobljena krožnica se vzame kot najboljši kandidat za segment očesa v regiji.

3.1.9 Iskanje partnerskih krožnic

Krožnici ne moreta biti partnerki (kandidat za par oči), če velja vsaj eden izmed pogojev:

- če je obraz obrnjen za več kot 30° ,
- če sta polmera najboljše krožnice enaka 1, potem razdalja med središčema krožnic ne sme biti iz intervala $[13,27]$,
- če ima maksimalen polmer vrednost 2 ali 3, potem razdalja med središčema krožnic ne sme biti iz intervala $[13,50]$, prav tako pa ne moreta biti partnerki krožnici s polmeroma 1 in 3,
- če sta polmera večja od 3, potem razdalja med središčema krožnic ne sme biti manjša od 16 ali pa med krožnicama ne sme biti prostora za manjši polmer in
- če je razmerje med površinama okvirjev regij manjše od 10%.

Hevristična pravila so bila sestavljena na podlagi značilnosti učne množice. Najprej se je avtomatsko, torej s programom, obdelala željena informacija na sliki, nato pa se je ta informacija uporabila pri ročni izpeljavi pogojev, pri tem pa so se upoštevale tudi značilnosti obraza. Pogoji so dodatno omiljeni tako, da se od konkretno dobljenih števil dopušča odstopanje do 10%.

3.1.10 Potrditev kandidata za obraz

Za potrditev kandidata za obraz se ponovno uporabi barvna informacija. Najprej se izračuna točka na sredini med središčema krožnic, nato pa se pravokotno (navzdol) na daljico, ki predstavlja razdaljo med središčema izračuna točka, ki je od

sredine oddaljena za polovico omenjene razdalje. Točka naj bi bil približek sredine nosu. Sedaj pogleda, če je krožnica s središčem v tej točki in polmerom velikosti razdalje med izračunanima točkama v celoti na sliki. Krožnica naj bi opisovala del obraza okoli nosu. Če krožnica ni v celoti na sliki, potem kandidata za obraz zavrne, sicer pa pogleda v notranjost krožnice. Znotraj krožnice se izračuna odstotek slikovnih elementov, ki so barve kože ter najsvetlejša in najtemnejša vrednost (sivina) slikovnih elementov barve kože. Če je odstotek slikovnih elementov barve kože manjši od 85% ali pa, če je razlika med najsvetlejším in najtemnejšim slikovnim elementom manjša od 60¹, potem kandidata za obraz zavrne, sicer pa ga označi na osnovi tako dobljenih vrednosti (oceni verjetnost kandidata):

- najmanjšo verjetnost pripiše kandidatu, ki ima znotraj krožnice vsaj 85% slikovnih elementov barve kože in razliko med najsvetlejším in najtemnejšim slikovnim elementom vsaj 60 - takšnega kandidata označi na izhodni sliki s sivino 200,
- srednjo verjetnost pripiše kandidatu, ki ima znotraj krožnice vsaj 90% slikovnih elementov barve kože in razliko med najsvetlejším in najtemnejšim slikovnim elementom vsaj 80 - takšnega kandidata označi na izhodni sliki s sivino 100 in
- največjo verjetnost pripiše kandidatu, ki izpolnjuje vsaj enega izmed naslednjih pogojev (takšnega kandidata označi na izhodni sliki s črno barvo):
 - znotraj krožnice ima vsaj 90% slikovnih elementov barve kože in razliko med najsvetlejším in najtemnejšim slikovnim elementom vsaj 120,
 - znotraj krožnice ima vsaj 95% slikovnih elementov barve kože in razliko med najsvetlejším in najtemnejšim slikovnim elementom vsaj 100,
 - znotraj krožnice ima 100% slikovnih elementov barve kože.

Nos kot središče kroga je izbran zato, ker je ta del obraza pri večini ljudi podoben, hkrati pa predstavlja najbolj neodvisen del obraza od raznih posebnih značilnosti posameznika (npr. nošenje očal, prisotnost brade, pričeska preko čela,...).

3.2 Odraz pragov

Kako pragovi, opisani v poglavju 3.1, vplivajo na učinkovitost algoritma, si oglejmo na slikah 3.4, 3.5, 3.6 in 3.7. Vse vhodne slike pripadajo učni množici.

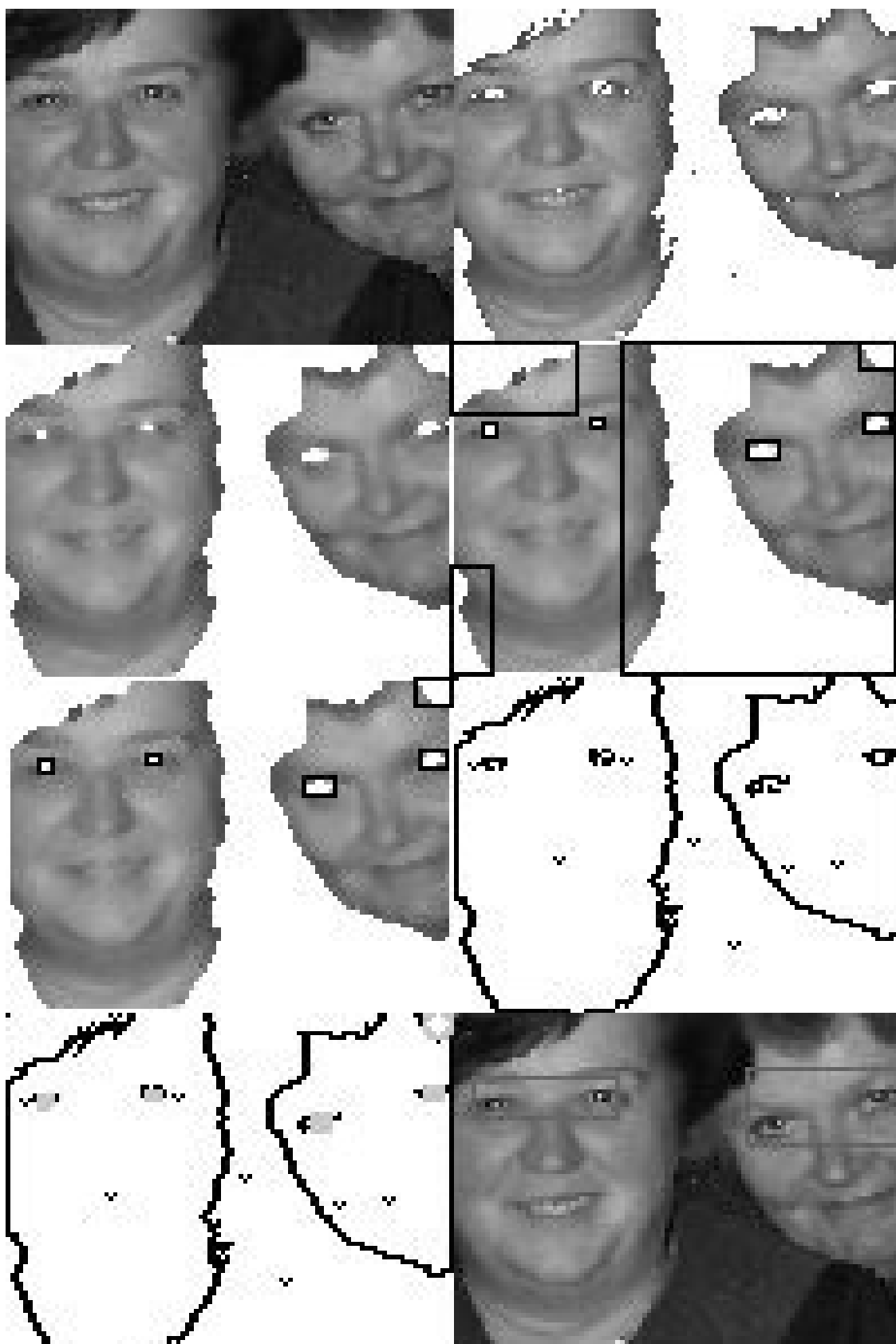
¹Razlika med najsvetlejším in najtemnejšim slikovnim elementom predstavlja eno od značilnosti obraza. Namreč nos zaradi svojih geometrijskih značilnosti povzroči razliko v svetlosti barve kože v svoji okolici. Na ta način na primer izločimo iz nadaljnje obdelave del stene, ki ima barvo kože, ne izpolnjuje pa pogoja o zahtevani minimalni razliki med najsvetlejším in najtemnejšim slikovnim elementom.

Iz slike 3.4 je lepo razvidna zavrnitev kandidata za obraz, ki vsebuje levo oko leve osebe in desno oko desne osebe, saj je odstotek slikovnih elementov barve kože znotraj krožnice, ki naj bi vsebovala nos, manjši od 85%.

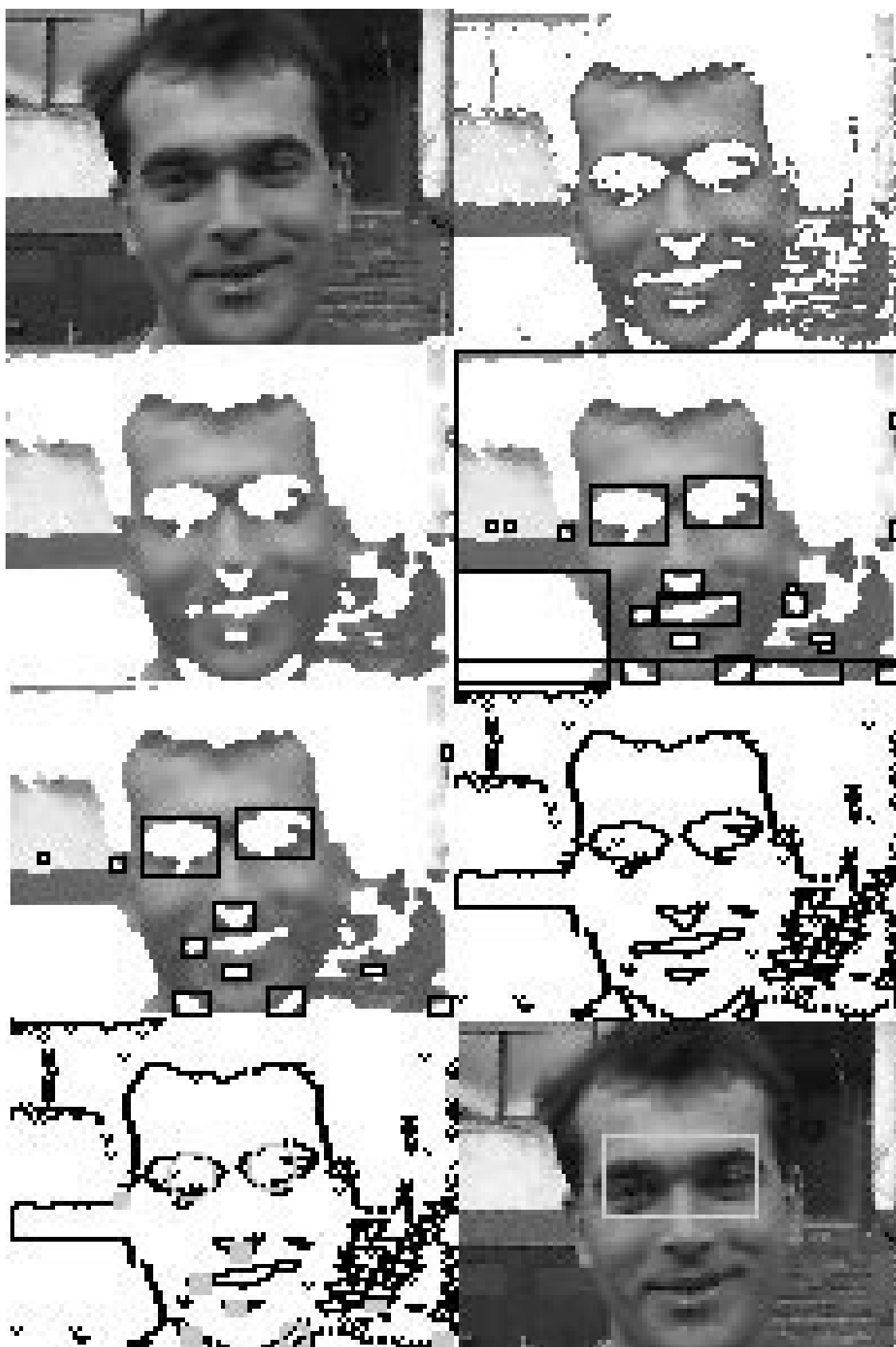
Bistveni značilnosti slike 3.5 sta kompleksno ozadje, ki se delno ujema z barvo kože in slaba osvetlitev osebe. To botruje velikemu številu belih regij, ki pa jih algoritem učinkovito omeji.

Zaradi ustrezne osvetlitve algoritem najde na sliki 3.6 vse tri obraze, zaradi večje oddaljenosti obrazov v ozadju, pa slednjima pripiše manjšo verjetnost, saj sta osvetljena veliko slabše kot obraz v ospredju. Razlika med najsvetlejším in najtemnejšim slikovnim elementom znotraj krožnice, ki naj bi vsebovala nos, ne ustreza pragu za višjo verjetnost. Seveda je rezultat tudi posledica dejstva, da je obraz v ospredju v večji ločljivosti kot obraza v ozadju.

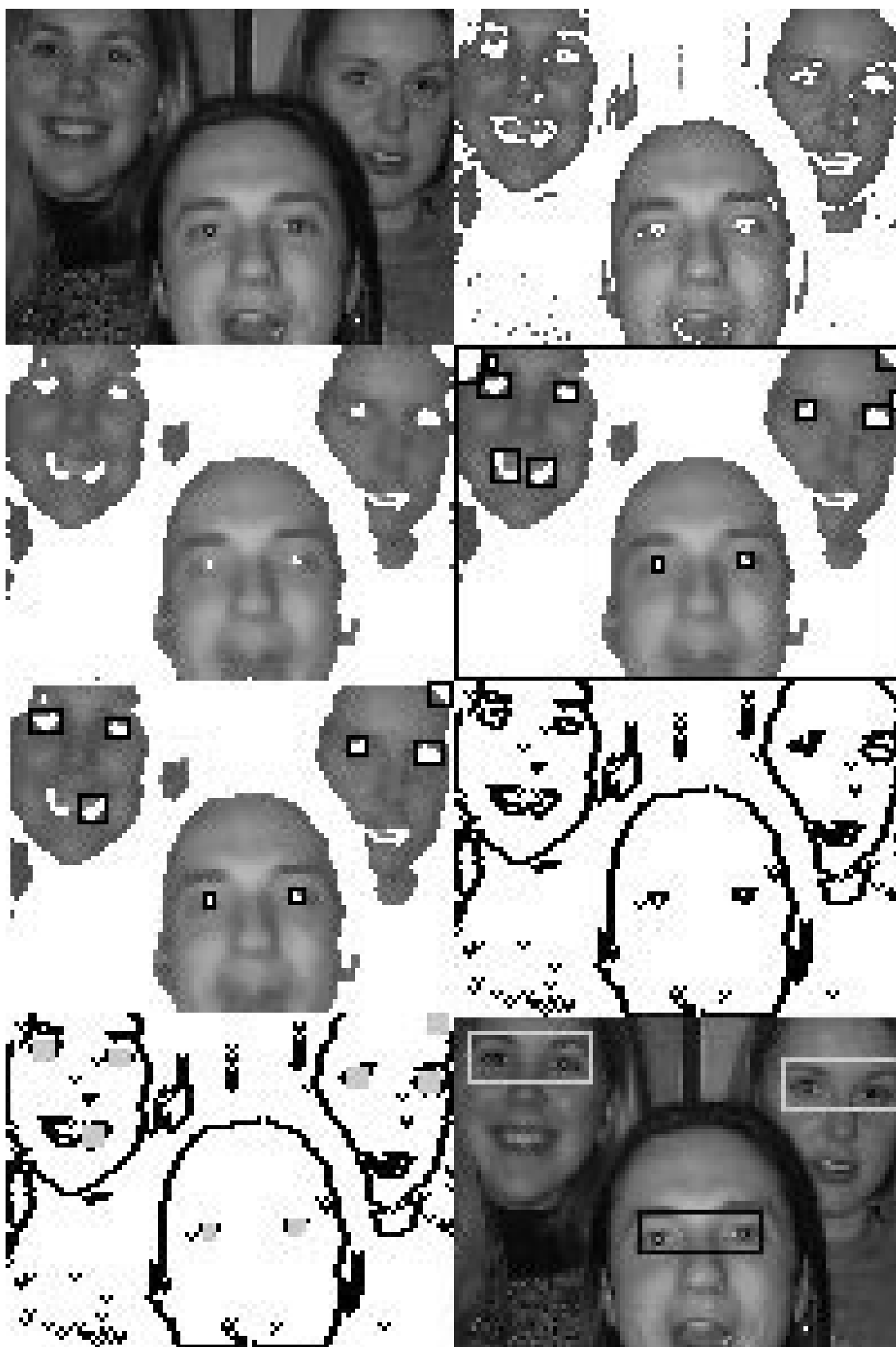
Slika 3.7 predstavlja tipičen element množice, ki ustreza vsem zahtevam navedenim na začetku poglavja 3. Iz slike je lepo razvidna izločitev regij obrvi.



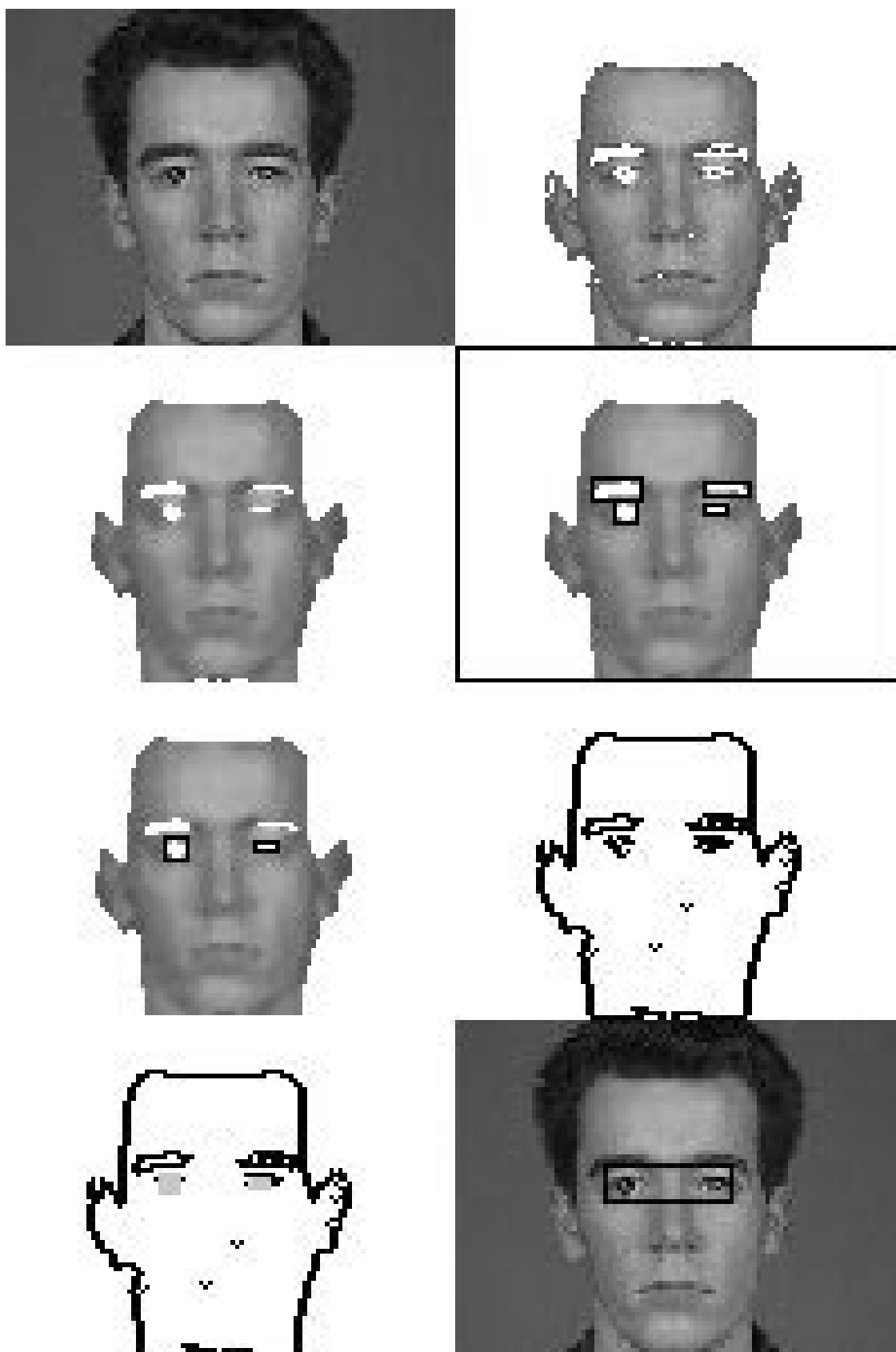
Slika 3.4: Zavrnitev kandidata za obraz, ki vsebuje levo oko leve osebe in desno oko desne osebe



Slika 3.5: Učinkovita omejitev števila belih regij



Slika 3.6: Zaradi večje oddaljenosti obrazov v ozadju, algoritem slednjima pripiše manjšo verjetnost, saj sta osvetljena veliko slabše kot obraz v ospredju, pravtako pa sta obraza v ozadju v manjši ločljivosti



Slika 3.7: Lepo razvidna izločitev regij obrvi

Poglavje 4

Rezultati

Iz priloge A so razvidni še rezultati nad preostalo učno množico. Algoritem nad učno množico dosega 100 odstotno klasifikacijsko točnost, saj je služila kot osnova pri snovanju algoritma. Tako za testiranje učinkovitosti algoritma ta množica ni verodostojna.

Da dobimo verodostojne rezultate o učinkovitosti algoritma, je bil algoritem testiran nad neodvisno testno množico. To množico sestavljajo slike iz dveh javnih baz obrazov, nekaj realističnih animacij obrazov ter nekaj slik znanih osebnosti:

- javna baza obrazov PICS (Psychological Image Collection at Stirling) - 31 slik s po enim obrazom na sliko, z enostavnim ozadjem, osvetlitev se (rahlo) spreminja, do izraza pride tudi uporaba bliskavice,
- javna baza obrazov M2VTS (Multi Modal Verification for Teleservices and Security applications database) - šest slik s po enim obrazom na sliko, z enostavnim ozadjem, pri enakomerni osvetlitvi; na dveh slikah je obraz rahlo obrnjen proti profilnemu položaju,
- realistične animacije obrazov - trije obrazi na dveh slikah, z enostavnim ozadjem,
- slike znanih osebnosti - sedem slik s po enim obrazom na sliko, na dveh slikah je obraz obrnjen proti profilnemu položaju, ena slika pa ima kompleksnejše ozadje.

Skupna značilnost vseh slik je, da so bolj ali manj primerne za identifikacijske dokumente. Slike znanih osebnosti so vključene v test zato, ker so te slike posneli profesionalni fotografi; tudi pretvorba v digitalno obliko je bila narejena z njihove strani.

Vse slike v testni množici ne ustrezajo zahtevam navedenim na začetku poglavja 3, vendar prav te dobro ilustrirajo pomankljivosti predlaganega algoritma. Slike, ki tako že vnaprej niso imele dobrega izhodišča so:

- slike, na katerih je obraz obrnjen proti profilnemu položaju in

- slike s kompleksnejšim ozadjem.

Takšnih slik je skupaj v testni množici pet, od tega je na dveh algoritem uspešno našel obraz. Dejstvo, da sta na obeh slikah obraza obrnjena proti profilnemu položaju, pove, da algoritem tolerira tudi manjše odklone od frontalnega položaja.

Dobrega izhodišča nimajo tudi tiste slike, ki imajo v ozadju barve, ki ustrezajo lastnostim bliskavice. Zaradi padanja sence ob uporabi bliskavice je tako do problemov prihajalo tudi pri slikah iz baze PICS, vendar je algoritem ta problem v večini primerov dobro rešil.

Tabele 4.1, 4.2, 4.3 in 4.4 podajajo rezultate testiranja algoritma nad posamezno podmnožico testne množice, tabela 4.5 pa nad celotno testno množico.

Vsi rezultati so razvidni iz slik 4.1, 4.2, 4.3, 4.4, 4.5, 4.6 in 4.7. Pri negativnih rezultatih je podana tudi razlaga, kaj je povzročilo napako.

baza obrazov PICS - 31 slik - 31 obrazov	
število najdenih obrazov	število zadetkov
37	31

Tabela 4.1: Rezultati testiranja algoritma nad bazo obrazov PICS

baza obrazov M2VTS - 6 slik - 6 obrazov	
število najdenih obrazov	število zadetkov
7	6

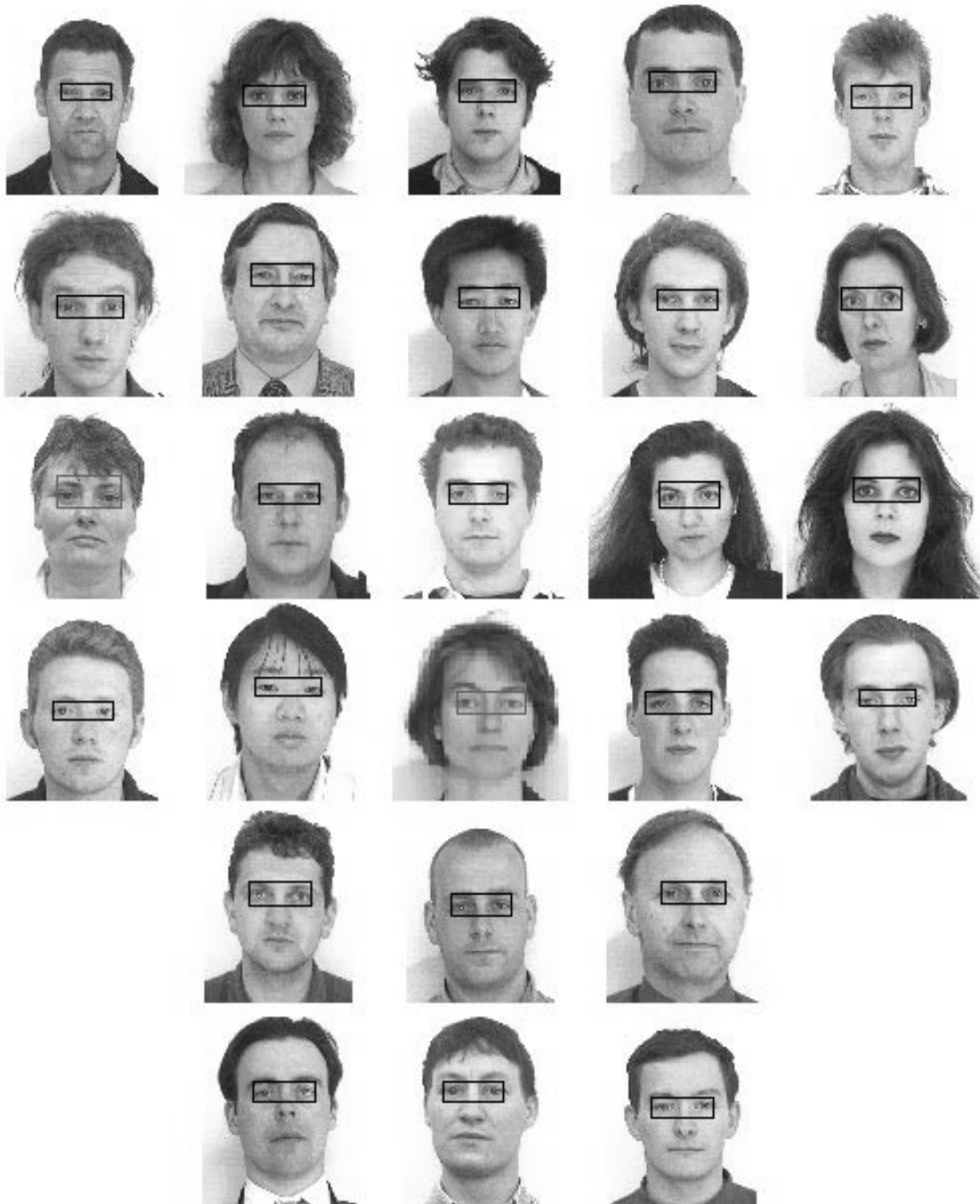
Tabela 4.2: Rezultati testiranja algoritma nad bazo obrazov M2VTS

realistične animacije obrazov - 2 sliki - 3 obrazi	
število najdenih obrazov	število zadetkov
3	3

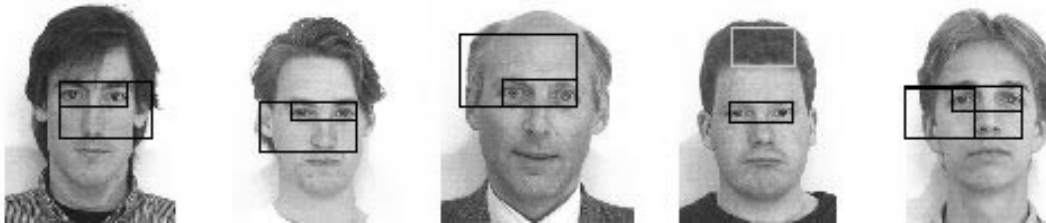
Tabela 4.3: Rezultati testiranja algoritma nad realističnimi animacijami obrazov

Iz rezultatov lahko sklepamo, da ob upoštevanju vseh omejitev algoritma, le-ta lahko doseže skoraj 100 odstotno klasifikacijsko točnost, vendar to pomeni, da se omejimo na določeno domeno slik, ki ustrezajo točno določenim zahtevam in se oddaljimo od željene univerzalnosti! To pa ni nujno slabo, saj je algoritem dovolj učinkovit, da bi lahko bil integriran v aplikacijo, katero omejitve algoritma ne bi ovirale, sama pa bi zahtevala hitro delovanje.

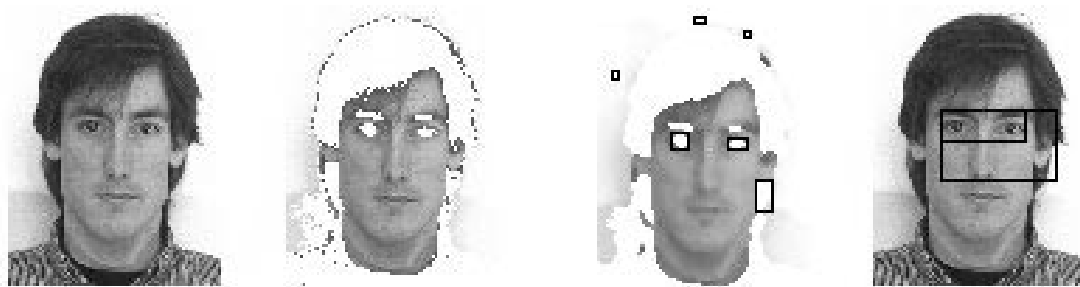
Algoritem bi naprimer lahko služil kot uvodni del sistema za razpoznavanje obrazov, saj je iz rezultatov razvidno, da nad testno množico v večini primerov pravilno označi obraz (točen podatek: 43 zadetkov/44 obrazov=97.7%), pogoške pa bi sistem tako izločil na osnovi ujemanja z bazo obrazov, ki jih že pozna.



Slika 4.1: Pozitivni rezultati iz baze obrazov PICS



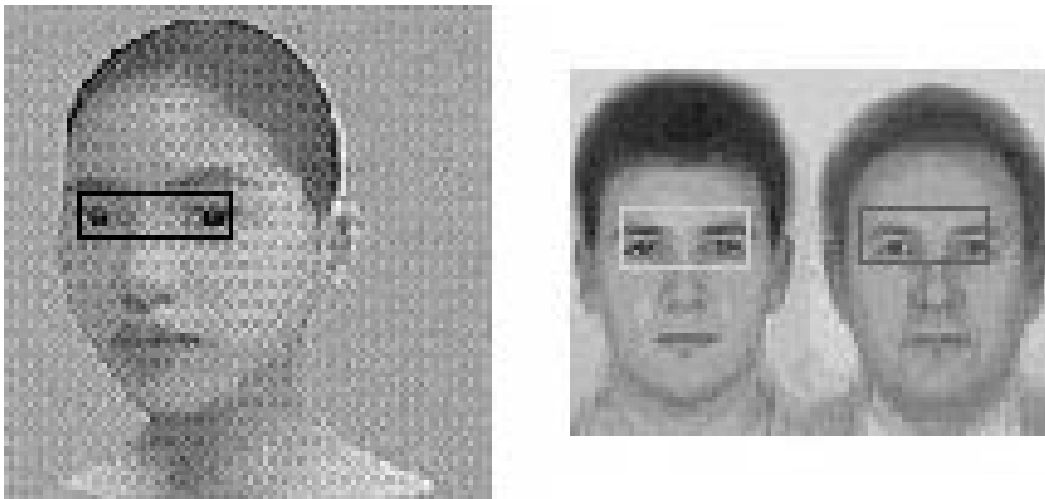
Slika 4.2: Negativni rezultati iz baze obrazov PICS - na prvi in zadnji sliki pride do napake zaradi barve ozadja, ki ustreza lastnostim bliskavice, na preostalih treh pa zaradi barve las (za konkreten primer glej naslednjo sliko)



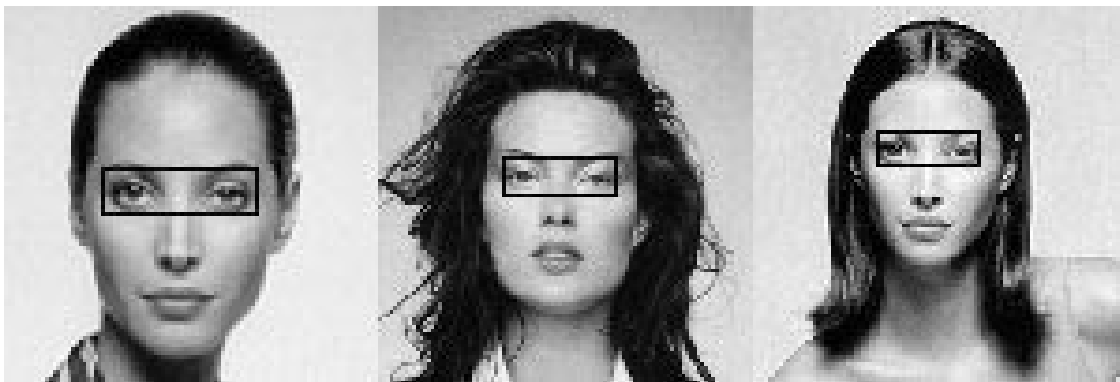
Slika 4.3: Obrazložitev negativnega rezultata iz baze obrazov PICS: iz zaporedja slik je lepo razvidno, da algoritem obdrži regije, katerih najtemnejša točka je presvetla, da bi regija lahko vsebovala oko, razlog za napako pa je neizločeno ozadje, ki ustreza barvi kože (skrajno desna regija)



Slika 4.4: Pozitivni rezultati iz baze obrazov M2VTS; iz zadnjih dveh slik je razvidno, da algoritem tolerira tudi manjše odklone od frontalnega položaja



Slika 4.5: Pozitivni rezultati nad realističnimi animacijami obrazov



Slika 4.6: Pozitivni rezultati nad slikami znanih osebnosti



Slika 4.7: Negativni rezultati iz baze obrazov M2VTS (sredinska slika) in nad slikami znanih osebnosti - na prvi sliki pride do napake zaradi zlitja regije očesa z regijo ozadja (obraz najde na osnovi kombinacije oko-obrv), na drugi sliki pride do napake zaradi kompleksnejšega ozadja, na tretji sliki zaradi zaliscev, na četrti in peti pa zaradi prevelikega odklona od frontalnega položaja (na peti tudi zaradi napačne izločitve regije očesa); tukaj moramo vzeti v zakup dejstvo, da 2., 4. in 5. slika ne ustrezajo zahtevam predlaganega pristopa navedenim na začetki poglavja 3

znane osebnosti - 4 ustrezne slike - 4 obrazi	
število najdenih obrazov	število zadetkov
4	3

Tabela 4.4: Rezultati testiranja algoritma nad slikami znanih osebnosti; tri slike (2., 4. in 5. slika na sliki 4.7) niso ustrezale zahtevam navedenim na začetku poglavja 3 in zato niso vključene v zgornjih rezultatih

celotna testna množica - 43 slik - 44 obrazov	
število najdenih obrazov	število zadetkov
51	43

Tabela 4.5: Rezultati testiranja algoritma nad celotno testno množico

4.1 Šibke točke algoritma

Rezultati testiranja algoritma so primarno ocenili učinkovitost algoritma, sekundarno pa izpostavili njegove šibke točke:

Sliki 4.2 in 4.3 izpostavljata predvsem dejstvo, da pride do napak pri klasifikaciji zaradi neizločenega ozadja in barve las. Obdržijo se tudi regije, katerih najtemnejša točka je presvetla, da bi regija lahko vsebovala oko. Takšne regije lahko na podlagi te informacije zagotovo izločimo. Lasje predstavljajo problem le v toliko, če so neenakomerno osvetljeni (oziroma če so v skrajnosti pobarvani z več različnimi barvami), sicer jih algoritem tako ali tako večji del ohrani oziroma odstrani. Najbolj problematična barva las je svetlo rjava, saj ta barva ustreza spodnji meji barve kože v barvnem prostoru RGB.

Ostali negativni primeri (slika 4.7) večinoma ne ustrezajo omejitvam algoritma (2., 4. in 5. slika na sliki 4.7 ne ustrezajo zahtevam navedenim na začetku poglavja 3). Najbolj pereč problem predstavlja negativen primer, ki nastopi zaradi zalisev (3. slika na sliki 4.7). Rešitev tega problema bi lahko iskali v bolj izpopolnjenem modelu barvne informacije znotraj krožnice, ki naj bi vsebovala nos. Prva slika na sliki 4.7 potrjuje že izpostavljeno dejstvo (poglavje 3.1.1), da je upravljanje barvne informacije posameznih slikovnih elementov vse prej kot stabilen postopek.

Skupno rešitev vseh treh problemov, torej problema neizločenega ozadja, barve las in zalisev, bi lahko iskali tudi v dodatni hevristici, ki bi izkoriščala ugotovitev, da napačno klasificirani obrazi največkrat vsebujejo pravilno označen obraz ali pa vsaj eno značilko, torej oko. Po drugi strani velja tudi, če primerjamo označitve, ki se prekrivajo, da je pravilno označen obraz po površini vedno najmanjši. S pomočjo takšne hevristike bi se število najdenih obrazov v tabeli 4.5 zmanjšalo za 6, kar bi pomenilo, da bi bila napačno klasificirana le dva primera in sicer predzadnja slika na sliki 4.2 in prva slika na sliki 4.7.

Izpostavljene šibke točke in možne rešitve problemov oziroma razširitve algoritma predstavljajo smernice za nadaljne delo.

Poglavje 5

Zaključek in nadaljne delo

Avtomatsko iskanje obrazov je, kot tudi večina drugih avtomatskih iskanj različnih objektov, zahtevna naloga, predvsem zaradi pomembnih variacij vzorcev, ki jih težko parametriziramo analitično.

Zaradi vse bolj razširjenih in obsežnih slikovnih podatkovnih baz in "živih" video informacij je njihovo inteligentno oziroma avtomatično preiskovanje izredno pomembno. Ljudje, to je človeški obrazi, so eden od najpogostejših in zelo specifičnih objektov, ki jih želimo poiskati na slikah. O pomembnosti problematike pričajo številne raziskave, ki jih na tem področju izvajajo raziskovalne skupine po svetu.

Večina razvitih algoritmov ima vsaj enega od dveh najpogostejših problemov:

- prevelika računska (časovna, prostorska) kompleksnost in/ali
- premajhna učinkovitost.

Pri določenih problemih računalniškega vida je praktično nesmiselno delati s kompleksnejšimi modeli, če nam že enostavni ponujajo dovolj veliko količino informacij. Ta stavek je predstavljal vodilo pri snovanju predlaganega algoritma.

Algoritem je zasnovan nad množico dokaj različnih slik, t.i. učno množico (pri tem ni uporabljeno strojno učenje). Cilj algoritma je bil doseči maksimalno klasi-fikacijsko točnost nad slikami, ki ustrezajo naslednjim zahtevam oziroma omejitvam:

- vhodna slika mora biti v dovolj veliki ločljivosti; obraz mora biti dovolj velik¹,
- občutljivost na barvo kože (omejimo se na svetlo polt),
- delovanje v realnem času na dandanes standardnem osebнем računalniku,
- enostavno ozadje,
- enakomerna osvetlitev slike (ateljejska osvetlitev),

¹Slike v učni množici so velike 100×75 slikovnih elementov, slike v testni množici pa 100×100 slikovnih elementov. Razmerje med stranicama formata 100×75 ustreza razmerju med stranicama standardne fotografije.

- obrazi svetle polti, ki morajo biti v celoti na sliki (frontalen položaj) in
- obrazi obrnjeni največ do 30° .

Algoritem združuje ideje pristopa na osnovi značilnk in pristopa na osnovi barv (glej poglavje 1.2.2). Prvi dve omejitvi predlaganega algoritma tako izvirata že iz omejitev omenjenih pristopov.

Učinkovitost algoritma je bila testirana nad neodvisno množico slik, t.i. testno množico. Iz rezultatov lahko sklepamo, da ob upoštevanju vseh omejitev algoritma, le-ta lahko doseže skoraj 100 odstotno klasifikacijsko točnost, vendar to pomeni, da se omejimo na določeno domeno slik, ki ustrezajo točno določenim zahtevam in se oddaljimo od željene univerzalnosti! To trditev lahko podkrepimo z dejstvom, da nad testno množico algoritem v večini primerov pravilno označi obraz (točen podatek: 43 zadetkov/44 obrazov=97.7%).

Univerzalnosti pa algoritem zaradi omejitev in temu primerne zasnove ne more doseči. Tako ni primeren kot izhodišče za načrtovanje fleksibilnejšega algoritma, dovolj učinkovit pa je, da bi lahko bil integriran v aplikacijo, katero omejitve algoritma ne bi ovirale, sama pa bi zahtevala hitro delovanje.

Smernice za nadaljne delo lahko torej iščemo na več koncih:

- nadgraditev algoritma tako, da bi dosegal 100 odstotno klasifikacijsko točnost nad določeno domeno slik, ki ustrezajo točno določenim zahtevam,
- zasnova fleksibilnejšega algoritma,
- zasnova aplikacije, ki bi uspešno uporabljala predlagan algoritem,
- ...

Dodatek A

Učna množica



Slika A.1: Preostala učna množica

Rezultati nad preostalo učno množico so prikazane na sliki A.1; v učno množico spadajo tudi slike 3.1, 3.4, 3.5, 3.6 in 3.7. Seveda algoritem nad učno množico dosega 100 odstotno klasifikacijsko točnost, saj je služila kot osnova pri snovanju algoritma. Tako za testiranje učinkovitosti algoritma ta množica ni verodostojna. Celotna učna množica vsebuje 24 slik.

Dodatek B

Demo različica sistema A. Rowleya in sod.

Pristop A. Rowleya in sod. [11] je opisan v poglavju 2.1.2¹. Izpostavimo le nekaj bistvenih lastnosti pristopa:

- temelji na strojnem učenju (nevronske mreže),
- omogoča iskanje obrazov v frontalnem položaju in
- na vhodu dobi sivinsko sliko.

Že iz naštetih lastnosti je razvidna ogromna razlika med tem pristopom in v pričujočem delu predlaganim pristopom za avtomatsko iskanje človeških obrazov na slikah (glej poglavje 3).

Na sliki B.1 je združenih več slik iz testne množice (glej slike 4.1, 4.4, 4.5 in 4.6). Predlagan algoritem je na vseh slikah pravilno označil obraz. Enak zaključek lahko podamo tudi za sistem A. Rowleya in sod..

Slika B.2 je ekvivalentna sliki 4.7. Iz nje lahko sklepamo, da je sistem A. Rowleya in sod. fleksibilnejši od v tem delu predlaganega pristopa, saj dovoli večje odmike obraza od frontalnega položaja in tolerira večjo stopnjo šuma na sliki. Tukaj moramo vzeti v zakup dejstvo, da 2., 4. in 5. slika na sliki 4.7 ne ustrezajo zahtevam predlaganega pristopa navedenim na začetku poglavja 3, pravtako pa velja, da na 2. in 3. sliki na sliki 4.7 ne bi prišlo do napačne klasifikacije, če bi v predlagan postopek dodali hevrstiko opisano v poglavju 4.1.

Torej A. Rowley in sod. so razvili učinkovit sistem, katerega šibki točki sta kvečjemu časovna kompleksnost in zahteva po računski moči.

¹Demonstracijska različica sistema A. Rowleya in sod. je na voljo na spletnem naslovu <http://www.ius.cs.cmu.edu/demos/facedemo.html>.



Slika B.1: Primeri slik, na katerih sta oba pristopa (pristop A. Rowleya in sod. ter v tem delu predlagan pristop) pravilno označila obraze



Slika B.2: Primeri slik, ki dokazujejo večjo fleksibilnost sistema A. Rowleya in sod. (glej sliko 4.7)

Literatura

- [1] J. Ben-Arie, D. Nandy, Representation Of Objects In A Volumetric Frequency Domain With Application To Face Recognition, *IEEE CVPR Proceedings*, pp. 615-620, 1997.
- [2] J. Canny, A Computational Approach to Edge Detection, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-8, no. 6, pp. 679-698, 1986.
- [3] G. Chow, X. Li, Towards a System for Automatic Facial Feature Detection, *Pattern Recognition*, vol. 26, no. 12, pp. 1739-1755, 1993.
- [4] J. Demšar, F. Solina, Searching for Faces in Image Date Bases Using Machine Learning, *Speech and Image Understanding, Proceedings of 3rd Slovenian-German and 2nd SDRV Workshop*, Ljubljana, Slovenia, 1996, pp. 199-207.
- [5] R. Feraud, O. Bernier, J.-E. Viallet, M. Collobert, D. Collobert, A Conditional Mixture of Neural Networks for Face Detection, Applied to Locating and Tracking an Individual Speaker, *CAIL'97*, Kiel, Germany, pp. 464-471, 1997.
- [6] R. M. Haralick, K. Shanmugam, I. Dinstein, Textural Features for Image Classification, *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-3, no. 6, pp. 610-621, 1973.
- [7] T. S. Jebara, A. Pentland, Parametrized Structure from Motion for 3D Adaptive Feedback Tracking of Faces, *IEEE CVPR Proceedings*, pp. 144-150, 1997.
- [8] N. Oliver, A. Pentland, LAFTER: Lips and Face Real Time Tracker, *IEEE CVPR Proceedings*, pp. 123-129, 1997.
- [9] E. Osuna, R. Freund, F. Girosi, Training Support Vector Machines: an Application to Face Detection, *IEEE CVPR Proceedings*, pp. 130-136, 1997.
- [10] P. J. Phillips, H. Moon, P. Rauss, S. A. Rizvi, The FERET Evaluation Methodology for Face-Recognition Algorithms, *IEEE CVPR Proceedings*, pp. 137-143, 1997.

- [11] H. A. Rowley, S. Baluja, T. Kanade, Neural Network-Based Face Detection, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 1, pp. 23-38, 1998.
- [12] K. Sobottka, I. Pitas, A novel method for automatic face segmentation, facial feature extraction and tracking, *Signal processing: Image communication*, 12, pp. 263-281, 1998.
- [13] K.-K. Sung, T. Poggio, Example-Based Learning for View-Based Human Face Detection, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 1, pp. 39-51, 1998.
- [14] L. Wiskott, J.-M. Fellous, N. Krüger, C. von der Malsburg, Face Recognition by Elastic Bunch Graph Matching, *CAIL'97*, Kiel, Germany, pp. 456-463, 1997.
- [15] L. Wiskott, Phantom Faces for Face Analysis, *CAIL'97*, Kiel, Germany, pp. 480-487, 1997.
- [16] K. C. Yow, R. Cipolla, Feature-based human face detection, *Image and Vision Computing*, no. 15, pp. 713-735, 1997.
- [17] A. L. Yuille, P. W. Hallinan, D. S. Cohen, Feature Extraction from Faces Using Deformable Templates, *International Journal of Computer Vision*, vol. 8, no. 2, pp. 99-111, 1992.
- [18] Internet (neka j osnovnih naslovov):
- <http://www.cs.rug.nl/~peterkr/FACE/face.html> - Face Recognition Home Page,
 - <http://pics.psych.stir.ac.uk/> - PICS Image Database,
 - <http://www.tele.ucl.ac.be/M2VTS/> - M2VTS Face Database,
 - <http://www.cogs.susx.ac.uk/users/davidy/teachvision/vision0.html> - Computer Vision Teach Files,
 - <http://www.ius.cs.cmu.edu/demos/facedemo.html> - CMU's Face Detector Demo.