

UNIVERZA V LJUBLJANI
FAKULTETA ZA RAČUNALNIŠTVO IN INFORMATIKO

Matija Mlinar

Kategorizacija starosti iz slik obrazov

DIPLOMSKO DELO
NA UNIVERZITETNEM ŠTUDIJU

Mentor: prof. dr. Franc Solina
Somentor: dr. Borut Batagelj

Ljubljana, 2009



Št. naloge: 01563/2009

Datum: 05.04.2009

Univerza v Ljubljani, Fakulteta za računalništvo in informatiko izdaja naslednjo nalogo:

Kandidat: **MATIJA MLINAR**

Naslov: **KATEGORIZACIJA STAROSTI IZ SLIK OBRAZOV**
AGE CLASSIFICATION FROM FACIAL IMAGES

Vrsta naloge: Diplomsko delo univerzitetnega študija

Tematika naloge:

Kandidat naj pregleda tematiko kategorizacije oseb v starostne razrede oziroma določanja točne starosti na podlagi slike obraza. V okviru svoje diplomske naloge naj tudi izdelava sistem, ki bo omogočal kategorizacijo starosti. Implementirane rešitve naj preizkusi na podatkovni zbirki oseb ter predstavi rezultate. Na podlagi preizkusa naj predlaga najboljšo metodo oziroma njihovo kombinacijo.

Mentor:

prof. dr. Franc Solina

Dekan:

prof. dr. Franc Solina



Rezultati diplomskega dela so intelektualna lastnina Fakultete za računalništvo in informatiko Univerze v Ljubljani. Za objavlanje ali izkoriščanje rezultatov diplomskega dela je potrebno pisno soglasje Fakultete za računalništvo in informatiko ter mentorja.

Besedilo je oblikovano z urejevalnikom besedil \LaTeX .

Namesto te strani **vstavite** original izdane teme diplomskega dela s podpisom mentorja in dekana ter žigom fakultete, ki ga diplomant dvigne v študentskem referatu, preden odda izdelek v vezavo!

IZJAVA O AVTORSTVU

diplomskega dela

Spodaj podpisani/-a Matija Mlinar,

z vpisno številko 63010099,

sem avtor/-ica diplomskega dela z naslovom:

Kategorizacija starosti iz slik obrazov

S svojim podpisom zagotavljam, da:

- sem diplomsko delo izdelal/-a samostojno pod mentorstvom prof. dr. Franca Soline in somentorstvom dr. Boruta Batagelja
- so elektronska oblika diplomskega dela, naslov (slov., angl.), povzetek (slov., angl.) ter ključne besede (slov., angl.) identični s tiskano obliko diplomskega dela
- soglašam z javno objavo elektronske oblike diplomskega dela v zbirki "Dela FRI".

V Ljubljani, dne 2.7.2009

Zahvala

Zahvaljujem se mentorju asistentu dr. Borutu Batagelju za pomoč pri izdelavi diplomske naloge. Zahvaljujem se teti Andreji Ferreira, stari mami Berti Šubic, sestri Eriki Mlinar, da so prispevale slike za eksperimentiranje in mami Tanji Mlinar, ki je ravno tako prispevala slike in vse skupaj tudi predhodno uredila. Zahvaljujem se tudi vsem ostalim, ki so kakorkoli pripomogli k uspešni izvedbi diplomskega dela.

Kazalo

Povzetek	1
1 Uvod	3
1.1 Področja uporabe	3
1.2 Opis nadaljevanja	4
2 Sorodna dela	5
3 Avtomatična kategorizacija	11
3.1 Lokalizacija značilnic	11
3.2 Normalizacija slik obrazov	13
3.3 Kodiranje obrazov v podprostor	14
3.3.1 Analiza glavnih komponent	15
3.3.2 Linearna diskriminantna analiza	16
3.3.3 Dvo-dimenzionalna linearna diskriminantna analiza	17
3.3.4 Statistični modeli oblike in videza	18
3.4 Metode za kategorizacijo	20
3.4.1 Gaussovi modeli	20
3.4.2 Polinomski modeli	21
3.4.3 Najbližji sosedi	23
3.4.4 Metoda podpornih vektorjev	23
3.4.5 Odločitveno drevo	24
3.4.6 Več-stopenjska kategorizacija	25
4 Eksperimentiranje	27
4.1 Podatkovna baza	27
4.2 Potek eksperimenta	28
4.3 Eksperimentalni rezultati	30

5 Zaključek	43
A Numerični rezultati eksperimentov	46
Seznam slik	52
Seznam tabel	53
Literatura	54

Seznam uporabljenih kratic in simbolov

2DLDA - Two Dimensional LDA - Dvo dimenzionalna LDA.

AAM - Active Appearance Models - Modeli aktivnega videza - Modeli, ki se uporabljajo za iskanje značilnic na fleksibilnem objektu, v tem primeru na obrazu.

AGES - AGing pattErn Subspace - Podprostor starostnih vzorcev. Metoda, ki je definirana v [8].

AgeS - Oznaka se uporablja za najboljšo na starost specifično metodo oziroma nastavitev. To je dvo-stopenjska klasifikacija z NN in KM.

AppS - Oznaka se uporablja za najboljšo na videz specifično metodo oziroma nastavitev. To je klasifikacija z NN v skupino podobnih obrazov (skupine so zgrajene z algoritmom k-means) in SVM za klasifikacijo v starostni razred.

AS - Appearance Specific - Specifičen na videz. Označuje na videz specifično metodo, ki je definirana v [12].

DT - Decision Tree - Odločitveno drevo.

GM - Gaussov Model.

HCI - Human-Computer Interaction - Interakcija človek-računalnik.

KM - Kvadratni Model.

LDA - Linear Discriminant Analysis - Linearna diskriminantna analiza. Meto-

da, ki pretvori podatke v podprostor tako, da so razredi, v katere so razdeljeni podatki, najbolj ločeni.

NN - Nearest Neighbours - Najbližji sosedi. Metoda strojnega učenja, kjer se za kategorizacijo novega primera vzamejo najbližji primeri iz učne množice.

PCA - Principal Component Analysis - Analiza glavnih komponent. Metoda, ki pretvori podatke v podprostor tako, da so med seboj najbolj ločeni.

PM - Polinomski Model.

SAM - Statistical Appearance Models - Statistični modeli videza. Statistični modeli oblike in videza, ki so jih definirali Edwards in sodelavci [5].

SVM - Support Vector Machines - Metoda podpornih vektorjev - Metoda strojnega učenja, ki v učni množici loči primere v dva razreda s hiperravnino tako, da je rob med razredoma maksimalen.

WAS - Weighted Appearance Specific - Otežen specifičen na starost. Označuje oteženo na videz specifično metodo, ki je definirana v [13].

Povzetek

Pri interakciji med ljudmi ima informacija, ki jo razberemo z obraza, pomembno vlogo. Če bi bili tudi računalniki zmožni zanesljivo razbrati informacijo z obrazov, bi bila interakcija z računalnikom prijaznejša za uporabnika. Kljub temu da ima pomembno funkcijo pri interakciji tudi starost, razpoznavanje starosti z obraza še ni bilo izčrpno raziskano. Za veliko aplikacij bi bilo dovolj, da bi bili računalniki zmožni zanesljivo kategorizirati sliko obraza v nekaj starostnih obdobjih. Na svetovnem spletu bi bilo na primer dovolj, da računalnik loči otroka od odrasle osebe za dostop do spletnih strani z neprimerno vsebino.

Pri eksperimentiranju sem starosti umetno razdelil na razrede s širino 5, 10 in 15 let in opazoval, kakšna je klasifikacijska točnost pri različnih metodah za učinkovito predstavljanje slik obrazov z manj dimenzijami (PCA, LDA, 2DLDA, statistični modeli oblike in videza) in pri različnih metodah za klasifikacijo slik obrazov v starostne razrede (Gaussovi modeli, polinomski modeli, najbližji sosedi, metoda podpornih vektorjev, odločitveno drevo). Uporabljal sem podatkovno bazo FG-NET Aging Database, kjer obrazne slike vsebujejo skoraj vse možne variacije. Izkazalo se je, da ima oblika obraza, ki je poleg videza obraza vgrajena v statistične modele oblike in videza, pomembno prednost pred metodami, ki so osnovane le na videzu. Pri klasifikacijskih metodah se je najbolje obnesla metoda SVM. Pri štirih starostnih razredih s širino 15 let sem dobil najvišjo klasifikacijsko točnost 73%, kar nakazuje, da bi pri naravnejši razdelitvi starostnih razredov in bolj specifičnih slikah obrazov lahko uporabili kategorizacijo starosti iz slik obrazov tudi v realnih aplikacijah.

Ključne besede:

analiza obraza, PCA, LDA, modeli oblike in videza, AAM, klasifikacija starosti, Mahalanobisova razdalja, SVM

Abstract

Information which we comprehend from faces plays an important role in interaction between humans. If computers would be capable of reliably recognizing information from faces, interaction with computers could be more user friendly. Despite the fact that the age of a person also plays an important role in the interaction between humans, age recognition was not extensively studied so far. For many applications it would be enough if computers could reliably classify facial images into a few age categories such as separating children and adults when accessing the world wide web.

In experiments, I made age groups using 5, 10 and 15 year intervals and observed classification accuracy when applying different methods for effective presentation of facial images with less dimensions (PCA, LDA, 2DLDA, statistical shape and appearance models) and when applying different methods for classification of age groups (Gaussian models, polynomial models, nearest neighbours, support vector machines, decision tree). The FG-NET Aging Database where facial images display almost all possible variations was used in experiments. As it was shown the shape of face which is also built into the statistical shape and appearance models has an important advantage over methods which are based only on appearance. Best performance of classification methods was achieved with the SVM method. Using four age groups with 15 year age intervals the highest classification accuracy of 73% was achieved. This indicates that age classification from facial images when using a more naturally defined age groups and more specific facial images could be applied in real applications.

Key words:

face analysis, PCA, LDA, shape and appearance models, AAM, age classification, Mahalanobis distance, SVM

Poglavje 1

Uvod

Obrazne slike vsebujejo ogromno količino informacije. Človeku, ki je prikazan na sliki, lahko določimo njegovo identiteto, čustveno stanje, etničnost, spol, starost in orientacijo glave. Ta informacija ima pomembno vlogo pri neposrednem komuniciranju med ljudmi. Med interakcijo lahko ljudje v realnem času zelo natančno prepoznamo in interpretiramo obraze in obrazne geste in se na osnovi tega primerno odzivamo.

Trenutni trendi v informacijski tehnologiji diktirajo izgradnjo sistemov, ki bi lahko prav tako dobro uporabljali informacijo razbrano iz obraznih slik pri interakciji človek-računalnik (HCI) [14]. Avtomatično izločanje večine informacije z obrazov je bilo že obsežno raziskano. Avtomatična kategorizacija starosti obraza pa je kljub temu, da ima starost pri interakciji pomembno vlogo, relativno neraziskan, nov problem. Razlogov za to je več. Avtomatično ocenjevanje starosti je specifičen problem, saj celo človeški opazovalec težko natančno določi starost obraza na sliki. Če opazujemo populacijo ljudi vidimo, da pri staranju veljajo globalne zakonitosti, vsi ljudje se namreč staramo. Poleg tega pa obstajajo tudi lokalne zakonitosti, saj je staranje pri posameznem človeku odvisno od več faktorjev, ki so poleg dednega zapisa še zdravje, način življenja, vremenski pogoji, itd. Raziskovanje staranja otežuje tudi dejstvo, da je staranje dolgotrajen, nekontroliran proces in je težko zbrati primerne učne podatke.

1.1 Področja uporabe

Obstaja vrsta področij, kjer bi se sistem za razpoznavanje starosti lahko uporabljal.

- **interakcija med računalnikom in človekom, prilagojena starosti osebe:** Če bi računalniki znali določiti starost osebe, bi lahko prilagodili tako uporabniški vmesnik kot tudi tip komunikacije. Poleg standardnega sistema HCI bi se lahko tak sistem uporabljal v kombinaciji s sistemom za varen dostop do interneta in bi mladoletnim osebam preprečil dostop do internetnih strani z neprimerno vsebino. Naslednji primer, kjer bi komponenta za določanje starosti imela pomembno vlogo, je sistem za pametno oglaševanje, ki ga razvijajo v laboratoriju za računalniški vid na naši fakulteti za računalništvo in informatiko [17]. Sistem bo omogočal prilagoditev oglasov glede na osebo, ki oglase gleda. Glede na ocenjeno starostno obdobje, se bi osebi prikazali le oglasi, ki ustrezajo ocenjenemu starostnemu obdobju, za katere je večja verjetnost, da bodo v tej osebi vzbudili zanimanje.
- **indeksiranje v bazi slik obrazov po starosti:** Avtomatična kategorizacija starosti se bi lahko uporabljala pri iskanju slik po podatkovni bazi obrazov. Uporabniki bi imeli možnost v digitalnih foto albumih navesti interval starosti oseb na slikah obrazov. Vrnile se bi le tiste slike, ki ustrezajo podanim starostim.
- **razvijanje sistemov za simulacijo slike obraza pri podani starosti [13]:** Problem simulacije slike obraza pri podani starosti je inverzen problemu ocenjevanja starosti iz slike obraza. Metodologija potrebna za razvoj sistema za avtomatično kategorizacijo lahko predstavlja izhodišče za razvoj sistema za simulacijo slike obraza pri podani starosti. Tak sistem bi zelo koristil policiji za iskanje pogrešanih otrok.
- **razumevanje procesa staranja pri ljudeh:** Nazadnje bodo dognanja na področju ocenjevanja starosti iz slike obraza lahko v neprecenljivo pomoč psihologom, ki preučujejo, kako ljudje zaznavamo starost.

1.2 Opis nadaljevanja

Preostali del diplomskega dela je organiziran na naslednji način: V poglavju 2 so na kratko opisana sorodna dela na področju kategorizacije starosti iz slik obrazov. V poglavju 3 je opisana moja rešitev problema avtomatične kategorizacije starosti iz slik obrazov. V poglavju 4 je opisana podatkovna baza slik, ki sem jo uporabljal, nastavitve parametrov pri vodenju eksperimenta in eksperimentalni rezultati ter ugotovitve. V poglavju 5 je orisano, kako bi izgledala realna aplikacija za kategorizacijo starosti iz slik obrazov.

Poglavje 2

Sorodna dela

Klasifikacija obraznih slik v starostne razrede je lahko narejena na podlagi geometrijskih razmerij med obraznimi značilnicami in na podlagi gub, ki se pojavijo na obrazu, ko se oseba stara. Dojenčki in mladi otroci imajo drugačno strukturo kosti kot odrasli ljudje. Na primer, pri odraslih ljudeh se oči nahajajo višje na obrazu kot pri otrocih, nos se premakne bližje k ustom, stran od oči, ko se oseba stara. Ko oseba postane odrasla, se struktura kosti ne spreminja več.

Kwon in Lobo [9] sta prva raziskovalca, ki sta preučevala klasifikacijo obraznih slik v starostne razrede. Na slikah sta najprej detektirala obrazne značilnice in gube. Na podlagi razmerij med lokacijami obraznih značilnic in na podlagi števila gub sta določila starostno obdobje osebe na sliki. Slike sta klasificirala v tri razrede: dojenčki, mladi odrasli in starejši odrasli. Eksperimentirala sta s petimi dojenčki, petimi mladimi odraslimi in petimi starejšimi odraslimi. Klasifikacija je bila uspešna za vseh 15 obrazov.

Iga in sodelavci [10] so klasificirali obraze v pet starostnih skupin: 15-24, 25-34, 35-44, 45-54 in 55-64 let. Klasifikacija je temeljila na obrazni teksturi, na lokacijah obraznih značilnic, na Gaborjevih značilnicah, na barvi kože in na laseh. Na naštetih obraznih lastnostih so naučili SVM klasifikatorje. Za končno klasifikacijo v starostne razrede se je odločalo na podlagi glasovanja teh klasifikatorjev. V nadaljevanju so pred klasifikacijo starosti razpoznali še spol in imeli ločene SVM klasifikatorje za oba spola. Za eksperiment so zbrali 101 sliko z različnimi ljudmi, starimi od 22 do 66 let. Klasifikacijska točnost je bila 58,4%, pri ročno označenih točkah pa 61,4%.

Wilhelm in sodelavci [19] so raziskovali prepoznavanje identitete, starosti, spola in mimike obraza. Starost so klasificirali v pet starostnih skupin med 10 in 60 leti. Najprej so obraze ročno označili s točkami, da so jih lahko normalizirali. Potem so jih pretvorili v podprostor z metodo ICA (Independent Component Analysis) in metodo AAM (Active Appearance Models). Za klasifikacijo so uporabili večnivojski perceptron (MLP), mrežo RBF, najbližjega soseda (NN), mrežo GLVQ. Najbolje se je obnesel AAM model skupaj z večnivojskim perceptronom. Klasifikacijska točnost je bila nekaj čez 40%. Po učinkovitosti, ne pa toliko po klasifikacijski točnosti (nekaj nad 30%), se je dobro odrezala tudi metoda ICA skupaj z NN. Kljub temu, da so bili rezultati slabi, so opazili, da bi bili pri uporabi samo dveh kategorij mladi in starejši ljudje ločeni zelo uspešno. Eksperimente so izvajali na bazi slik s 70 različnimi ljudmi in 7 slikami za vsakega človeka. Izrazi obrazov na slikah so bili nevtralni, osvetljenost slik je bila spremenljiva, orientacija obrazov je bila največ malo zamaknjena od pravokotnega pogleda iz slike.

Ueki in sodelavci [18] so pretvorili slike obrazov v podprostor v dveh korakih. Najprej so na njih izvedli metodo 2DLDA in nato še metodo LDA. Uporabljali so bazo z 5500 različnimi osebami japonskega porekla. Posamezna oseba v bazi je imela od 1 do 14 slik, tako da je bilo vseh slik skupaj 26222. Starost oseb je bila v intervalu od 3 do 85 let. Za klasifikacijo so uporabili Gaussove modele (enačba 2.1), ki so se jih naučili z EM algoritmom. Eksperimentirali so z uporabo starostnih intervalov 5, 10 in 15 let. Klasifikacijske točnosti pri uporabi posameznih intervalov so bile 46.3%, 67.8% in 78.1%. Za primerjavo so ponovili postopek še na slikah pretvorjenih s PCA in na slikah pretvorjenih z LDA. S tem so pokazali da prednost metode 2DLDA+LDA ni samo v tem, da je za razliko od PCA robustna na osvetljenost in da za razliko od LDA ne pride do singularnosti matrike lastnih vektorjev (v primeru da je učnih podatkov premalo), ampak daje tudi boljše rezultate.

$$P(x|c_i) = \frac{1}{(2\pi)^{\frac{d}{2}} |\Sigma|^{\frac{1}{2}}} \exp\left[-\frac{1}{2}(x - \mu)^t \Sigma^{-1}(x - \mu)\right] \quad (2.1)$$

Poleg klasifikacije v starostne razrede, je možno ocenjevati dejansko starost osebe. Lanitis, Taylor in Cootes [13] so za ocenjevanje uporabili linearno, kvadratno in kubično funkcijo. Koeficiente funkcij so poiskali z genetskim algoritmom. Boljše rezultate sta dala kvadratna in kubična funkcija. Razlika med njima je bila tako majhna, da je zaradi enostavnosti bolje uporabiti kvadratno funkcijo. Zgradili so globalno funkcijo za ocenjevanje starosti in funkcijo za vsako osebo posebej. Pri ocenjevanju starosti novega obraza so

pri več funkcijah uporabili funkcijo tiste osebe iz učne množice, ki je bila novi osebi najbolj podobna. Ta način ocenjevanja starosti so poimenovali ocenjevanje specifično na videz. Za ocenjevanje starosti nove slike obraza so uporabili tudi več funkcij. Funkcije so otežili glede na razdaljo med videzom novega obraza in videzom osebe iz učne množice. Ta način so poimenovali oteženo na videz specifično ocenjevanje. Eksperiment so izvajali na podatkovni bazi s 500 obraznimi slikami s 60 različnimi ljudmi pri starosti do 30 let. Učno in testno množico so sestavili na 3 načine: V prvem primeru so vse slike obravnavali enako in izvedli eksperiment pusti-eno-zunaj (leave-one-out), v drugem primeru so ločili slike glede na osebe in izvedli učenje po principu pusti-eno-osebo-zunaj (leave-one-person-out), v tretjem primeru so v testni množici imeli 65 slik oseb z drugačnim etničnim izvorom kot v učni množici. Za vse tri primere učne in testne množice in za vse tri načine ocenjevanja starosti iz slik obrazov s kvadratno funkcijo so rezultati prikazani v tabeli 2.1.

	Globalna funkcija	Na videz specifična funkcija	Otežena na videz specifična funkcija
primer 1	3.94	1.91	3.07
primer 2	6.21	7.43	5.04
primer 2	5.57	8.78	5.36

Tabela 2.1: Rezultati pri [13], kjer so eksperimentirali z globalno funkcijo in na videz specifičnimi funkcijami za oceno starosti iz obraznih slik. Eksperimentirali so tudi z različnimi nastavitvami učnih in testnih množic (primer1, primer2, primer3). Števila v tabeli predstavljajo srednjo absolutno napako ocene starosti v letih.

Lanitis [11] je preučeval uporabnost celotnega obraza (vključeni so tudi lasje), notranji del obraza (zajema oči in usta), spodnji del obraza (zajema usta) in zgornji del obraza (zajema oči) pri ocenjevanju starosti. Eksperimentiral je tudi s tremi različnimi resolucijami obraznih slik. 250 slik je uporabil za učenje in 80 slik za testiranje. Osebe na slikah so bile stare od 0 do 35 let. Na zgornjem delu obraza je bila ocena starosti najboljša. Srednja napaka je bila 3,83 let. Ne veliko slabše, s srednjo napako 4,05 let, se je odrezal notranji del obraza. Celoten obraz je dal slabše rezultate, kar pomeni, da imajo lasje zaradi različnih variacij negativen vpliv na določanje starosti. Spodnji del obraza pa ne vsebuje dovolj informacije za natančnejše ocenjevanje starosti. Razlike v resoluciji ne vplivajo bistveno na rezultat, kar pomeni, da na zazna-

vanje starosti bolj vpliva celovita struktura obraza kot pa posamezni obrazni detajli.

Lanitis, Draganova in Christodoulou [12] so eksperimentirali s štirimi različnimi metodami strojnega učenja: kvadratna funkcija, najkrajša razdalja, več-nivojski perceptron (MLP) in nenadzorovana nevronska mreža - mreža SOM (Self Organising Map). Koeficienti kvadratne funkcije so bili naučeni z genetskim algoritmom, za razdaljo pa so uporabili Mahalanobis razdaljo. Eksperimentirali so tudi z več-stopenjskim kategoriziranjem: direktna kategorizacija, na starost specifična kategorizacija, na videz specifična kategorizacija, na videz in na starost specifična kategorizacija. Pri direktni kategorizaciji se je starost določila direktno po eni izmed zgornjih štirih metod strojnega učenja. Pri kategorizaciji specifični na starost se je obrazne slike najprej kategoriziralo v enega od treh starostnih razredov: 0-10, 11-20, 21-35 let. Pri kategorizaciji specifični na videz so bili obrazi najprej kategorizirani v eno od skupin s podobnimi obrazi. Pri kategorizaciji specifični na starost in na videz so bili obrazi najprej kategorizirani v skupino s podobnimi obrazi, nato v enega od treh starostnih razredov in šele potem je bila izračunana dejanska starost. Eksperimentirali so na obraznih slikah 40 ljudi, starih od 0 do 35 let. Vseh slik skupaj je bilo 400. Najbolje, s srednjo napako 3,82 let, se je obnesla dvostopenjska klasifikacija s kvadratno funkcijo.

Lanitis in sodelavci [11, 12, 13] so v vseh svojih delih za obdelavo obraznih slik pred kategorizacijo uporabili statističen obrazni model [5]. Model omogoča kompaktno in reverzibilno kodiranje obrazov. Zgrajen je na podlagi statistične analize oblike in variacije intenzitete v reprezentativnem vzorcu. Najbolj pomembna lastnost teh modelov je, da kodirajo obraz v majhno število parametrov, imenovanih obrazni parametri. Na sliki 2.1 je slika obraza in njegova rekonstrukcija pri učenju modela na 35 obraznih slikah. Obraz je kodiran s 23 parametri.

Geng, Smith-Miles in Zhou [8] so svojo metodo poimenovali AGES (aGing pattErn Subspace). Tudi ti raziskovalci so razvili sistem, ki oceni dejansko starost in ne starostni razred. Njihova metoda je najmlajša med tu predstavljenimi in daje najboljše rezultate. Ideja njihove metode je modeliranje starostnega vzorca (aging pattern), ki je definiran kot zaporedje obraznih slik posameznika, razvrščenih v časovnem zaporedju. Starostni vzorci vseh oseb tvorijo podprostor starostnih vzorcev. Za nov, še neznan obraz se v podprostoru starostnih vzorcev najprej izbere primeren starostni vzorec, to je tisti



Slika 2.1: Originalna slika obraza (levo) in rekonstrukcija notranjega dela originalne slike obraza (desno) s statističnim modelom oblike in videza. Za učenje modela je bilo uporabljenih 35 obraznih slik. Obraz je kodiran s 23 parametri videza. Originalna slika (levo) je pomanjšana.

vzorec, ki nov obraz najboljše rekonstruira. Pozicija nove slike v izbranem starostnem vzorcu predstavlja starost osebe na novi sliki.

Za učenje so uporabljali bazo FG-NET Aging Database [7], ki je opisana v nadaljevanju. Pojavi se problem, ker iz same baze ni možno sestaviti kompletnih starostnih vzorcev. Pri vsaki osebi iz baze namreč manjkajo slike za določena leta. Problem so rešili z zelo domiselno rešitvijo. Manjkajočih slik so se naučili z iterativnim apliciranjem metode PCA na starostnih vzorcih.

V fazi učenja so iz vsake slike izločili značilnice z modelom videza [5], tako kot Lanitis in sodelavci. Starostni vzorec je bil tako predstavljen z vektorjem parametrov videza. Manjkajoča mesta v starostnem vzorcu so zapolnili s povprečnimi vrednostmi starostnih vzorcev. Na njih so potem naredili PCA in dobili začetno transformacijsko matriko. Transformacijsko matriko so uporabili za rekonstrukcijo manjkajočih delov v starostnih vzorcih. Na posodobljenih starostnih vzorcih so spet uporabili metodo PCA, itn. Iterativni postopek so ponavljali dokler napaka ni padla pod naprej definiran prag.

V času testiranja, ko je na vhodu ena sama slika se mora na izhodu pojaviti ena sama ocena za starost. Za dosego tega je slika oz. izločeni parametri postavljena na vsa možna mesta v starostnem vzorcu. S tem dobimo toliko starostnih vzorcev, kolikor je mest oz. starosti v njem. Pravilen je tisti starostni vzorec in s tem starost osebe na testni sliki, ki pri rekonstrukciji v podprostoru starostnih vzorcev vrne najmanjšo napako.

Metodo so testirali po principu pusti-eno-osebo-zunaj (Leave-One-Person-

Out - LOPO), kjer so v vsakem ciklu testiranja slike ene osebe v testni množici vse ostale slike pa so v učni množici. Poleg njihove metode so preizkusili tudi druge metode, ki so se na tem področju najbolj obnesle. Rezultati so prikazani v tabeli 2.2. Njihova metoda, s srednjo napako 6.22 let, ni samo prehitela vseh ostalih metod, ampak tudi človeške ocenjevalce.

Metoda	MAE
AGES	6.77
AGES+LDA	6.22
WAS	8.06
AAS	14.83
kNN	8.24
BP	11.85
C4.5	9.34
SVM	7.25
LjudjeA	8.13
LjudjeB	6.23

Tabela 2.2: Srednja absolutna napaka (MAE) pri [8], kjer so poleg njihove metode AGES preizkusili še oteženo na videz specifično metodo (WAS - Weighted Appearance Specific) [13], na videz in na starost specifično metodo (AAS - Appearance and Age Specific) [12], k-najbližjih sosedov (kNN - k-Nearest Neighbours), nevronska mrežo z vzvratnim razširjanjem (BP - Back Propagation), C4.5 odločitveno drevo (C4.5) in metodo podprtih vektorjev (SVM). Za primerjavo so starost ocenjevali tudi ljudje. V primeru LjudjeA so bila ljudem prikazana le sivinjska področja obraza (enako temu kar dobi algoritem na vhodu), medtem ko so bile v primeru LjudjeB ljudem prikazane celotne barvne slike.

Poglavje 3

Avtomatična kategorizacija

Za vsako avtomatično kategorizacijo obraza je potrebno obraz na sliki najprej detektirati. Za tem pride na vrsto lokalizacija značilnic na obrazu. Na podlagi odkritih značilnic se lahko obraz najprej poravna oziroma normalizira, v vsakem primeru pa se informacija o odkritih značilnicah rabi za dodatno obdelavo obraza v obliko, ki je primerna za kategorizacijo po kriteriju, ki nas zanima. V mojem primeru je ta kriterij starost.

3.1 Lokalizacija značilnic

Značilnice oziroma točke na obrazu se lahko odkrije z modeli aktivnega videza (Active Appearance Models - AAM) [6]. Za učenje modela AAM je potrebno nekaj slik ročno označiti in jih posredovati algoritmu. Na sliki 3.1 je slika obraza in varianta, kako označiti obraz s točkami. Vse slike v učni množici morajo biti označene na enak način. Trenutno sta na svetovnem spletu na voljo dve zbirki orodij za sestavljanje modelov AAM. To sta *am_tools* [1] in *AAM-API* [2]. Orodje *am_tools* sestavlja zbirka ukazov, ki se jih lahko požene v ukazni vrstici. *AAM-API* je za razliko od *am_tools* odprtokodno in ga sestavlja knjižnica *C++* funkcij in tudi nekaj prevedenih programov za poganjanje v ukazni vrstici. Kmalu bo na voljo še tretje javno orodje za AAM modele *AAMTools* [3] napisano v Matlabu. V primerjavi z *am_tools* in *AAM-API* bo omogočalo gradnjo izboljšanih različic modelov AAM.

Z orodjema *am_tools* in *AAM-API* se da enostavno zgraditi modele AAM in jih prilagati na neoznačenih obraznih slikah, vendar v primeru velikega števila variacij učnih slik postanejo preveč splošni in se na novo sliko obraza zelo slabo prilagajo. Problem predstavljata predvsem različne orientacije in različne osvetljenosti obrazov na slikah. Na temo robustnejših AAM modelov je bilo



Slika 3.1: Slika obraza označena s 68 točkami na enak način kot so označene slike v podatkovni bazi FG-NET Aging Database.

narejenih že veliko raziskav (npr. [4]), vendar trenutno še ni javno na voljo programov, ki bi omogočali direktno gradnjo takih modelov. To je prvi razlog, da sem se omejil na predpostavko, da so točke na obrazu že lokalizirane.

Če predpostavimo, da so obrazi približno na sredini slike, jih ni potrebno detektirati ampak se lahko direktno poišče točke z AAM modeli. V nasprotnem primeru je obraze potrebno detektirati, drugače iskanje točk z AAM modeli lahko divergira. Podatkovna baza FG-NET Aging Database, ki sem jo uporabljal, ima obraze približno na sredini slike, zato nisem iskal algoritmov za detekcijo obraza. Tudi točke so na obraznih slikah že označene, kar je drugi razlog, da sem se pri kategorizaciji starosti iz slik obrazov večinoma posvetil postopku kategorizacije starosti iz slik obrazov od lokaliziranih točk naprej.

Zaradi slabega prileganja AAM modelov na nov obraz niti nisem raziskoval, koliko detekcija točk zmanjša klasifikacijsko točnost kategorizacije starosti, ampak sem predpostavil, da so točke optimalno najdene. Tudi v večini sorodnih delih so v rezultatih predpostavili, da so točke optimalno najdene.

3.2 Normalizacija slik obrazov

Vsaj za metode osnovane samo na videzu (PCA, LDA, 2DLDA) je potrebno obraze predhodno normalizirati. Ker so v podatkovni bazi tudi črno-bele slike, sem vse slike najprej pretvoril v sivinske. Nato sem slike rotiral tako, da sta očesi pri vseh slikah na isti višini. Kot rotacije sem pri posamezni sliki izračunal s formulo:

$$\alpha = \arctan \frac{y_l - y_r}{x_l - x_r}, \quad (3.1)$$

kjer je (x_l, y_l) sredina levega očesa (x_r, y_r) pa sredina desnega očesa. Sredino očesa sem izračunal iz označenih točk na levem robu očesa in na desnem robu očesa. Sliko sem v Matlabu rotiral z ukazom:

```
imrotate(image, -alpha, 'bilinear'),
```

kjer je `image` matrika slikovnih elementov, `alpha` je vrednost kota izračunana po formuli 3.1, `bilinear` pa je metoda interpolacije. Primer rotacije slike je prikazan na sliki 3.2. Zaradi rotacije slike sem moral popraviti tudi koordinate označenih točk. Koordinatni sistem točk sem najprej premaknil v center slike, jih v tem koordinatnem sistemu s pomočjo polarnih koordinat rotiral in potem premaknil nazaj v prvotni koordinatni sistem. Prvotni koordinatni sistem sem moral še malo popraviti zaradi črnega roba, ki je nastal pri rotaciji. Točkam sem tako prištel še polovico krajše katete trikotnika nastale črne praznine.

Po rotaciji sem sliki obraza odrezal neuporabno ozadje. Na levi strani sem sliki odrezal vse, kar se je nahajalo levo od najbolj leve označene točke $(x_{min}, y_{x_{min}})$, na desni pa vse, kar se je nahajalo desno od najbolj desne označene točke na obrazu $(x_{max}, y_{x_{max}})$. Na vrhu sem odrezal vse, kar se je nahajalo višje od točke z minimalno vertikalno koordinato $(x_{y_{min}}, y_{min})$, na dnu pa sem odrezal vse kar se je nahajalo nižje od točke z maksimalno vertikalno koordinato $(x_{y_{max}}, y_{max})$. Primer odrezane slike je prikazan na sliki 3.3. V Matlabu sem za opisan postopek uporabil ukaz:

```
imcrop(image, [xmin, ymin, xmax-xmin, ymax-ymin]).
```



Slika 3.2: Primer rotacije slike obraza na podlagi koordinat oči: levo originalna slika obraza, desno rotirana slika obraza.

Označene točke na obrazih sem popravil tako, da sem jih zmanjšal za odrezani levi in zgornji rob. V primeru, ko sem za eksperimentiranje uporabil celoten notranji del obraza, sem obraz obrezal na podlagi vseh označenih točk. Pri eksperimentiranju z zgornjim delom notranjega obraza (obraz od obrvi do konice nosa) pa sem točke, ki so se nahajale pod nosom izbrisal in obraz obrezal na podlagi točk, ki so ostale.

Nazadnje sem slike še pretvoril na isto širino z ukazom:

```
imresize(image,[NaN,imwidth]),
```

kjer je `image` matrika slikovnih elementov, `imwidth` pa širina slike. Zaradi oznake `NaN` se višina slike poveča oziroma pomanjša sorazmerno s širino. Ker morajo imeti slike za metode osnovane na videzu tudi isto višino, sem na dnu matrik slikovnih elementov po potrebi dodal ničle.

3.3 Kodiranje obrazov v podprostor

Ko so točke na obrazu odkrite, je potrebno obraze preslikati v prostor z manj dimenzijami. Kodirani obrazi v podprostoru so tako primerni za kategorizacijo. Pri metodah, ki so osnovane na videzu je najbolj razširjena in enostavna analiza glavnih komponent (PCA), zato sem tudi jaz to metodo najprej uporabil. Poleg te metode sem uporabil še linearno diskriminantno analizo (LDA), ki je za metodo PCA druga najbolj razširjena metoda. Problem metode LDA je,



Slika 3.3: Primer rezanja slike obraza na podlagi optimalnih vertikalnih in horizontalnih označenih točk: levo rotirana slika obraza, v sredini odrezana slika celotnega notranjega dela obraza, desno odrezana slika zgornjega dela obraza.

da zahteva veliko število podatkov, drugače lahko matrike lastnih vektorjev, ki jih je potrebno sestaviti, postanejo singularne. Za premagovanje te ovire se pred metodo LDA lahko zmanjša dimenzijo podatkov z uporabo metode PCA ali dvo-dimenzionalne linearne diskriminantne analize (2DLDA). 2DLDA deluje enako kot LDA le da je uporabljena samo po eni dimenziji slik, samo po vrsticah ali samo po stolpcih, in se zaradi mnogo manjše dimenzije ne more izroditi, tako kot LDA. Skupaj sem torej uporabil 4 metode osnovane na videzu: PCA, LDA, PCA+LDA, 2DLDA+LDA. Poleg metod osnovanih na videzu sem modeliral slike obrazov tudi s hibridno metodo, ki združuje obliko in videz (SAM, SAM+LDA). Iz slik obrazov sem tvoril statistične modele oblike in videza [5]. V nadaljevanju so na kratko opisane metode in ukazi v Matlabu za PCA, LDA, 2DLDA in za statistične modele oblike in videza (SAM).

3.3.1 Analiza glavnih komponent

Metoda PCA iz podatkov tvori take vektorje, da podatke v smislu reprodukcije, kar najboljše opišejo. Ti vektorji so lastni vektorji kovariančne matrike C , ki

jo tvorimo iz podatkov z enačbo:

$$C = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(x_i - \bar{x})^T,$$

kjer je x_i vektorizirana slika, \bar{x} vektorizirana povprečna slika, n pa je število vseh slik. Novo vektorizirano sliko x_n pretvorimo v podprostor PCA tako, da jo odštejemo od povprečja učnih slik \bar{x} in pomnožimo z lastnimi vektorji:

$$y_n = W^T(x_n - \bar{x}),$$

kjer je W matrika lastnih vektorjev. V Matlabu se PCA izvede z ukazom:

$$[W, Y] = \text{princomp}(X),$$

kjer je X matrika vektorjev slik, ki so postavljeni v matriko X po vrsticah. Pri uporabljanju ukazov v Matlabu je treba biti vedno pozoren, da se vektorje podatkov vstavlja v matrike kot vrstice, kar je ravno obratno kot pri enačbah v literaturi, kjer so vektorji podatkov v matrikah postavljeni kot stolpci. Zato se novo sliko v Matlabu transformira v PCA prostor z ukazom:

$$y_n = (x_n - \text{mean}(X, 1)) * W$$

Ker sem za izvajanje metode LDA, ki je ni v Matlabu, našel na svetovnem spletu Matlab orodja za zmanjševanje dimenzij [16], ki so dostopna na povezavi [15], sem tudi metodo PCA večinoma izvajal s temi orodji. Metodo PCA sem v Matlabu tako lahko izvajal z ukazom:

$$[Y, \text{mapping}] = \text{pca}(X, \text{ndim}),$$

kjer je pri tem ukazu potrebno navesti še število dimenzij ndim in je število lastnih vektorjev, ki so vrnjeni v strukturi `mapping` v polju `M` (`mapping.M`), že zmanjšano na ndim .

3.3.2 Linearna diskriminantna analiza

Za razliko od metode PCA, ki poišče takšne lastne vektorje, ki slike med seboj najboljše ločijo, poišče metoda LDA takšne lastne vektorje, ki med seboj najboljše ločijo posamezne razrede. Sestaviti je potrebno matriko razpršenosti v razredih S_w in matriko razpršenosti med razredi S_b :

$$S_w = \frac{1}{n} \sum_{j=1}^c \sum_{x \in c_j} (x - \bar{x}_j)(x - \bar{x}_j)^T,$$

$$S_b = \frac{1}{n} \sum_{j=1}^c n_j (\bar{x}_j - \bar{x})(\bar{x}_j - \bar{x})^T,$$

kjer je n_i število slik v razredu c_j , n je število vseh slik, \bar{x}_j je povprečje vektorjev slik v razredu c_j , \bar{x} pa je povprečje vektorjev vseh učnih slik. Matrika lastnih vektorjev W se lahko sestavi s pomočjo največjih lastnih vrednosti matrike $S_b S_w^{-1}$. Novo vektorizirano sliko x_n se pretvori v podprostor LDA z enačbo: $y_n = W^T x_n$. V Matlabu sem s pomočjo orodij [15] za analizo LDA uporabljal ukaz:

```
[Y,mapping] = lda(X,labels,ndim),
```

kjer je parameter `labels` vektor oznak starostnih razredov, `ndim` pa dimenzija pretvorjenih slik, ki mora biti manjša od števila razredov c . Za pretvorbo nove slike v podprostor LDA je potrebno v Matlabu uporabiti stavek:

```
yn=xn*mapping.M.
```

3.3.3 Dvo-dimenzionalna linearna diskriminantna analiza

Za metodo 2DLDA slik ni treba pretvoriti v vektorje, ampak se jih pri računanju lastnih vektorjev posreduje kar v dvo-dimenzionalni obliki. Matriki S_w in S_b sta definirani skoraj enako kot pri LDA:

$$S_w = \frac{1}{n} \sum_{j=1}^c \sum_{X \in c_j} (X - \bar{X}_j)(X - \bar{X}_j)^T,$$

$$S_b = \frac{1}{n} \sum_{j=1}^c n_j (\bar{X}_j - \bar{X})(\bar{X}_j - \bar{X})^T,$$

kjer je X matrika slikovnih elementov učne slike, \bar{X} je povprečje matrik slikovnih elementov vseh učnih slik, \bar{X}_j pa je povprečje matrik slikovnih elementov v razredu c_j . Matriko W se ravno tako kot pri LDA tvori na podlagi lastnih vrednosti produkta $S_b S_w^{-1}$. Pri preslikavi se nove matrike slikovnih elementov X_n ne pretvori v vektor, ampak se jo kar tako množi z preslikovalno matriko W : $Y_n = X_n W$. V Matlabu sem spremenil funkcijo `lda` za izvajanje metode 2DLDA in jo poimenoval `lda2d`. Glavna stvar, ki sem jo spremenil v funkciji `lda` je, da sem ukaz `C = cov(X)` zamenjal z ukazom `C = cov2d(X3D)`. Funkcija `cov` v Matlabu iz vhodne matrike X izračuna kovariančno matriko C po enačbi:

$$C = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(x_i - \bar{x})^T,$$

kjer je x_i i -ta vrstica matrike X , \bar{x} je povprečje vrstic matrike X , n pa je število vrstic matrike X . Funkcijo `cov2d` sem napisal tako, da iz polja matrik $X3D$, kjer tretja dimenzija določa zaporedno številko matrike, izračuna matriko C po enačbi:

$$C = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})(X_i - \bar{X})^T,$$

kjer je X_i i -ta matrika v polju matrik $X3D$, \bar{X} je povprečje matrik polja matrik $X3D$, n pa je število matrik v polju matrik $X3D$.

3.3.4 Statistični modeli oblike in videza

Pri metodi statističnih modelov oblike in videza se na obliki, to je na koordinatah točk, najprej uporabi metodo PCA. Obliko obrazov se nato ukrivi v povprečno obliko vseh obrazov. Nato se še na videzu, to je na teksturi, ki je znotraj konveksne ovojnice, ki jo določajo točke, uporabi metodo PCA. Ker sta oblika in videz obraza do neke mere odvisna se PCA naredi še na parametrih, ki so dobljeni s PCA na obliki in na parametrih, ki so dobljeni s PCA na teksturi. Obraze sem lahko kodiral z orodji, ki sem jih našel na spletu. Ta orodja so orodja za gradnjo modelov AAM. Modeli AAM so s statističnimi modeli oblike in videza neposredno povezani. Tudi za učenje modelov AAM je potrebno najprej sestaviti statistične modele oblike in videza in potem dodati informacijo, kako in v kateri smeri se ti modeli lahko spreminjajo, da se jih lahko prilega na nove obraze.

Za učenje modelov oblike in videza sem uporabil orodje *am_tools* [1]. V ukazni vrstici je potrebno za učenje modela uporabiti ukaz:

```
am_build_apm face,
```

kjer je `face` vhodna datoteka s končnico *smd*, v kateri so specifikacije za učenje modela oblike in videza. Primer datoteke *smd* je prikazan na sliki 3.4. V datoteki so podatki o lokacijah, kjer so shranjene slike in točke, razni parametri za učenje modela in katere datoteke *pts* in *jpg* se uporabijo za učenje. Za izračun parametrov oblike in videza je potrebno v ukazni vrstici uporabiti še ukaz:

```
am_get_params -i face.smd -d ..\params\,
```

kjer je `face.smd` spet datoteka *smd* s specifikacijami za učenje modela oblike in videza. Argument `..\params\` označuje pot do direktorija, kamor se za vsako sliko zapišejo 3 datoteke s končnicami *b_shape*, *b_tex* in *b_app*. V datoteko

```

(1)  model_name: face
(2)  model_dir: ./
(3)  parts_file: face
(4)  image_dir: ../Images/
(5)  points_dir: ../Points/
(6)
(7)  shape_aligner: align_similar_2d
(8)  shape_modes: { min: 0 max: 30 prop: 0.98 }
(9)  tex_modes: { min: 0 max: 40 prop: 0.99 }
(10) combined_modes: { min: 0 max: 30 prop: 0.99 }
(11) params_limiter: mdpm_box_limits
(12) {
(13) sd_limits: 3
(14) }
(15) n_pixels: 10000
(16) colour: Grey // Alternatives: Grey,RGB,...
(17) tex_sampler: vapm_triangle_sampler<vxl_byte>
(18) tex_aligner: align_linear_1d
(19) shape_wts: EqualVar
(20)
(21) tex_model: pca
(22) pyr_builder: gauss_byte
(23) points_pyr_builder: Same
(24) max_im_pyr_levels: 5
(25)
(26) min_level: 0
(27) max_level: 4
(28)
(29) training_set:
(30) {
(31) 001A02.pts : 001A02.jpg
(32) 001A05.pts : 001A05.jpg
(33) 001A08.pts : 001A08.jpg
(34) 001A10.pts : 001A10.jpg
(35) 001A14.pts : 001A14.jpg
(36) 001A16.pts : 001A16.jpg
(37) 001A18.pts : 001A18.jpg
(38) 001A19.pts : 001A19.jpg
(39) 001A22.pts : 001A22.jpg
(40) }

```

Slika 3.4: Primer *smd* datoteke. V datoteki so podatki o: imenu modela, ki se bo zgradil in bo zapisan v datoteko s končnico *apm* (1); direktoriju, kamor se bo datoteka *apm* zapisala (2); imenu datoteke s končnico *parts*, ki ima podatke o delih obraza (3) (ta datoteka ni obvezna); direktoriju, kjer se nahajajo učne slike (4); direktoriju, kjer se nahajajo datoteke *pts* z označenimi točkami (5); raznih parametrih za gradnjo modela (7-27); slikah, ki sestavljajo učno množico in kako so datoteke *pts* povezane s slikami (29-40). Razen za število pikslov, ki se uporabijo za modeliranje videza (15), sem za vse parametre za učenje modela pustil privzete vrednosti.

b_shape se zapišejo parametri, ki so izračunani iz oblike, v datoteko *b_tex* se zapišejo parametri, ki so izračunani iz tekstore, v datoteko *b_app* pa se

zapišejo parametri, ki nosijo informacijo o obliki in teksturi. Jaz sem uporabljal samo datoteko *b_app*, saj je v parametrih, ki so v tej datoteki, približno toliko informacije kot v parametrih v datotekah *b_shape* in *b_tex* skupaj. Parametre je bilo potrebno tako le še prebrati iz datoteke *b_app* in se jih je lahko posredovalo naprej klasifikatorju. V prvi različici parametrov nisem dodatno spreminjal, v drugi različici pa sem na parametrih uporabil še metodo LDA, da bi se nepomembne variacije še dodatno zadušile (SAM+LDA).

3.4 Metode za kategorizacijo

Zadnji korak pri kategorizaciji starosti iz slik obrazov je učenje modela, ki kategorizira kodiran obraz v starostni razred. Eksperimental sem z naslednjimi modeli: Gaussovi modeli naučeni z EM algoritmom, polinomski modeli, najbližji sosedi z Mahalanobisovo razdaljo, metodo podpornih vektorjev, odločitvenim drevesom, več-stopenjsko klasifikacijo specifično na starost ali na videz. Kategorizacijo sem izvajal na slikah obrazov, ki so imele zmanjšano dimenzijo po eni izmed metod, ki so opisane v prejšnem razdelku. Spremenljivke kodiranih obrazov, ki nastopajo v tem razdelku imajo oznako x ali X (ker so vhod v sistem za kategorizacijo), kljub temu da so bile v prejšnem razdelku označene z y ali Y (ker so bile izhod iz sistema za kodiranje v podprostor).

3.4.1 Gaussovi modeli

Pri metodah za kodiranje slik obrazov v podprostor sem najprej eksperimental z metodo PCA, zato sem pri klasifikaciji kodiranih slik obrazov v starostne razrede najprej uporabil Gaussove modele (enačba 2.1), naučene z EM algoritmom. V članku [18] so namreč Gaussovi modeli uporabljeni skupaj z metodami PCA, LDA in 2DLDA+LDA. Potrebno je bilo zgraditi toliko modelov, kolikor je bilo razredov. Za razrede s širino 5 let, kjer je bilo vseh razredov skupaj 9, je bilo torej potrebno zgraditi 9 modelov. Nato sem novo sliko obraza prilagal na vse modele. Model, ki je vrnil največjo verjetnost, je določal razred, v katerega je bila klasificirana nova slika obraza. Za gradnjo modelov sem v Matlabu uporabil zaporedje ukazov:

```
for i=1:nclass
    model{i} = gmdistribution.fit(X(labels == i,:),1);
end,
```

kjer je `nclass` število vseh starostnih razredov, `gmdistribution` je Matlab razred Gaussovih mešanih modelov (ta razred nima nobene povezave s starostnimi razredi, ampak je mišljen kot programski konstrukt za kreiranje objektov), `fit` je funkcija, ki poišče parametre Gaussovega modela, `X(labels == i, :)` pa so učne slike, ki so v starostnem razredu `i`. Drugi argument pri klicu funkcije `fit` posreduje funkciji število komponent oziroma vrhov okoli katerih so po normalni porazdelitvi porazdeljeni podatki. Pri slikah obrazov sem predpostavil, da so v posameznem starostnem razredu kodirani obrazi porazdeljeni okoli enega idealnega navideznega obraza. Kategorizacija novega kodiranega obraza `newx` se v Matlabu izvede z naslednjim zaporedjem ukazov:

```
for i=1:nclass
    p(i) = pdf(model{i},newx);
end
[maxp,newlabel] = max(p),
```

kjer je `p(i)` verjetnost, ki jo vrne model `i`, `pdf` je funkcija gostota verjetnostne porazdelitve, ki prilega nove podatke na model in vrne verjetnost. Zadnja vrstica s funkcijo `max` še izračuna maksimalno vrednost vektorja `p` in s tem oznako starostnega razreda `newlabel`.

3.4.2 Polinomski modeli

V delih [11, 13, 12] so dobili dobre rezultate kar s prileganjem kvadratne funkcije na učne parametre oblike in videza. Pri kvadratni funkciji sta kodirana slika obraza in starost osebe na sliki povezana z naslednjo enačbo:

$$age = c + W_1 x^T + W_2 (x^2)^T, \quad (3.2)$$

kjer je `age` dejanska starost osebe, `x` je vektor s kodirano sliko obraza, `W1` in `W2` sta vektorja uteži in `c` je potreben prištevek. Za kategorizacijo nove slike je potrebno na učnih slikah poiskati vektorje uteži `W1` in `W2` in konstanto `c`. V statistiki se model 3.2 imenuje linearna regresija. V Matlabu je v zbirki orodij za statistiko za linearno regresijo na voljo funkcija `regress`, ki prilega model na podatke po metodi najmanjših kvadratov. Funkcijo sem klical na naslednji način:

```
W = regress(age, [ones(n,1),X,X.^2]),
```

kjer je `W` vektor uteži, v katerem so po vrsti speti odmik `c` in vektorja uteži `W1` in `W2`; `n` je število slik obrazov v učni množici; v matriki `[ones(n,1),X,X.^2]` so

speti vektor enic (`ones(n, 1)`), ki se rabi za izračun odmika c , matrika kodiranih slik X in matrika kvadratov posameznih elementov kodiranih slik $X.^2$; `age` je vektor, ki za vsako sliko v X vsebuje starost osebe na sliki. Funkcija `regress` za rešitev linearne regresije kliče operator *backslash* (`\`) za deljenje iz leve strani, zato dobimo enak rezultat če v Matlabu uporabimo ukaz:

$$W = [\text{ones}(n, 1), X, X.^2] \setminus \text{age}.$$

Na podlagi izračunanih uteži in odmika W ocenimo starost osebe na novi sliki obraza z ukazom:

$$\text{newage} = [1, \text{newx}, \text{newx}.^2] * W,$$

kjer je `newx` kodirana slika novega obraza, `newage` pa je ocenjena starost novega obraza. Pri testiranju je bilo potrebno zaradi primerljivosti z drugimi metodami še pretvoriti starost v starostni razred.

Ko sem imel enkrat sestavljen algoritem za izračun uteži kvadratne funkcije, sem zaradi enostavnosti preizkusil še linearno funkcijo, kubično funkcijo in funkciji oz. polinoma četrte stopnje in pete stopnje. Enačbe in pripadajoči ukazi v Matlabu za izračun uteži enačbe so po vrsti navedeni spodaj:

$$\text{age} = c + W_1 x^T, \quad (3.3)$$

$$W = \text{regress}(\text{age}, [\text{ones}(n, 1), X]),$$

$$\text{age} = c + W_1 x^T + W_2 (x^2)^T + W_3 (x^3)^T, \quad (3.4)$$

$$W = \text{regress}(\text{age}, [\text{ones}(n, 1), X, X.^2, X.^3]),$$

$$\text{age} = c + W_1 x^T + W_2 (x^2)^T + W_3 (x^3)^T + W_4 (x^4)^T, \quad (3.5)$$

$$W = \text{regress}(\text{age}, [\text{ones}(n, 1), X, X.^2, X.^3, X.^4]).$$

$$\text{age} = c + W_1 x^T + W_2 (x^2)^T + W_3 (x^3)^T + W_4 (x^4)^T + W_5 (x^5)^T, \quad (3.6)$$

$$W = \text{regress}(\text{age}, [\text{ones}(n, 1), X, X.^2, X.^3, X.^4, X.^5]).$$

3.4.3 Najbližji sosedi

Za novo sliko obraza se lahko izračuna razdalje do vseh slik v učni množici. Izbran je tisti starostni razred, v katerem se nahajajo najbližje slike obrazov. Pri najbližjih sosedih sem za mero bližine vzela Mahalanobisovo razdaljo, ki je definirana z enačbo:

$$d = (x_{new} - x_m)C^{-1}(x_{new} - x_m), \quad (3.7)$$

kjer je x_{new} vektor, ki vsebuje kodirano novo sliko; x_m je vektor s povprečnimi kodiranimi učnimi slikami obraza za dani starostni razred; C pa je kovariančna matrika parametrov obraza iz danega razreda. Značilnost Mahalanobisove razdalje je, da je invariantna na velikost (scale-invariant) in je zato bolj robustna kot ostale razdalje. Razdaljo je potrebno izračunati za vsak razred posebej, zato je v Matlabu za izračun potrebna zanka:

```
for i=1:nclass
    d(i) = mahal(newx,X(labels == i,:));
end
[mind,newlabel] = min(d);
```

kjer je `nclass` število starostnih razredov, `d(i)` je razdalja vektorja nove slike `newx` do učnih slik `X(labels == i,:)` v razredu `i`, `mahal` je funkcija za izračun Mahalanobisove razdalje, funkcija `min` poišče minimalno vrednost vektorja `d` in s tem oznako razreda za novo sliko `newlabel`. Na krajši način dobimo v Matlabu enak rezultat z diskriminantno analizo (funkcija `classify`), kjer za tip diskriminantne funkcije uporabimo 'mahalanobis':

```
newlabel = classify(newx,X,labels,'mahalanobis');
```

3.4.4 Metoda podpornih vektorjev

Na področju avtomatične analize obraza je pogosto uporabljena metoda podpornih vektorjev (SVM). Pri kategorizaciji starosti iz slik obrazov je uporabljena v delu [10]. Metoda loči učne primere v dva razreda s hiperravnino tako, da je rob ob hiperravnini maksimalen. V Matlabu je funkcija za učenje po metodi SVM del zbirke orodij za bioinformatiko (bioinformatics toolbox). Njena oblika je naslednja:

```
model = svmtrain(X,labels,'Method','LS');
```

V matriki `X` so kodirane slike obrazov. V vektorju `labels` so oznake starostnih razredov za posamezen kodiran obraz v `X`. Argumenta `'Method'` in `'LS'` posredujeta klicani funkciji informacijo o metodi za iskanje ločevalne hiperravnine; `'LS'` pomeni, da naj se uporabi metoda najmanjših kvadratov. V privzetem načinu tretjega in četrtega parametra ni potrebno posredovati klicani funkciji in se za iskanje ločevalne hiperravnine uporabi metoda kvadratičnega programiranja (Quadratic Programming). Problem iskanja hiperravnine je na danih podatkih težak, zato se izvajanje z metodo kvadratičnega programiranja ni končalo z uspešnim rezultatom ampak z opozorilom, da je potrebno v nastavitvah povečati maksimalno število iteracij za iskanje modela. Med alternativama povečanje maksimalnega števila iteracij ali izbira druge metode za izračun ločevalne hiperravnine sem raje izbral zadnjo, ker se je izvedla v razmeroma kratkem času z uspešno zgrajenim modelom.

Kodiran obraz nove slike `newx` se kategorizira v starostni razred z oznako `newlabel` s funkcijo `svmclassify`:

```
newlabel = svmclassify(model,newx).
```

Metoda podpornih vektorjev deluje na dveh razredih, starostnih razredov pa je več, zato sem po vrsti ločeval starosti v dva razreda, dokler nisem dobil razredov z zelenimi širinami. Ker se obrazi ljudi v začetku hitreje spreminjajo in ker je v podatkovni bazi FG-NET več mladih obrazov kot starih, sem razrede ločeval po vrsti od mlajšega do starejšega. Pri razredih s širino 10 let sem tako najprej ločil slike obrazov v razred do 9 leta in razred od 10 let naprej. Za ta dva razreda sem naučil prvi model. Potem sem ločil razred od 10 let naprej v razred od 10 do 20 let in v razred od 20 let naprej in sem naučil model na tej ravni. Postopek se je potem ponavljal, dokler se tudi zadnji razred ni več mogel ločiti. Število zgrajenih modelov je bilo na koncu enako številu razredov manj ena. Za kategorizacijo novega obraza je potrebno enako število iteracij kot za učenje. Najprej se novo sliko obraza kategorizira po modelu na prvem nivoju. Če je ta kategoriziran v razred do 9 let je to že rezultat, drugače ga je potrebno kategorizirati z modelom na naslednjem nivoju.

Preizkusil sem tudi drugo alternativo, kjer sem pričel ločevati starosti v dva razreda na sredini med končnimi razredi. Pri končnih razredih s širino 10 let se je ločevanje tako pričelo z razredoma do 30 let in od 30 let naprej. Pri tem načinu sem napisal rekurzivno funkcijo za učenje SVM modelov.

3.4.5 Odločitveno drevo

Preizkusil sem tudi znano metodo strojnega učenja - odločitveno drevo. V Matlabu se odločitveno drevo `t` zgradi z ukazom:

```
t = classregtree(X,labels,'method','classification'),
```

kjer je X matrika kodiranih slik obrazov, v vektorju `labels` so oznake za razrede za vsako sliko iz X , parametra `'method'` in `'classification'` posredujeta informacijo, da naj se zgradi odločitveno drevo. Novo sliko obraza `newx` se klasificira v starostni razred z oznako `newlabel` z ukazom:

```
newlabel = t(newx).
```

3.4.6 Več-stopenjska kategorizacija

Podobni obrazi se starajo na enak način, zato je dobro, da se pri kategorizaciji nove slike obraza upoštevajo le podobni obrazi iz učne množice. Ta problem sem poskusil rešiti na podoben način kot v raziskavah [13, 12]. Za upoštevanje subjektivnega faktorja pri staranju je v teh raziskavah nova slika obraza najprej klasificirana v skupino podobnih obrazov in potem je na podlagi skupine podobnih obrazov ocenjena starost nove slike obraza. Prvi način je, da pri vsaki osebi iz učne množice posebej modeliramo staranje [13] in za novo sliko obraza poiščemo nekaj najbližjih oseb iz učne množice. Drugi način je, da modeliramo staranje pri podobnih slikah obrazov ne glede na osebo, ki ji te slike pripadajo [12].

Za modeliranje staranja pri vsaki osebi posebej sem uporabil regresijo. Za vsako osebo iz učne množice je bilo potrebno izračunati koeficiente kvadratne enačbe 3.2. Pri novi sliki obraza se je moralo iz učne množice najprej izbrati eno ali več oseb, ki imajo najbolj podobne slike obrazov. Ker je bilo slik, ki jih je imela posamezna oseba v učni množici, manj kot je največja dimenzija kodirane slike, se ni dalo uporabiti Mahalanobisove razdalje ampak je bilo potrebno uporabiti Evklidsko. Po izračunu najbližjih oseb se je lahko uporabilo njihove funkcije staranja za oceno starosti nove slike.

Pri drugem načinu sem z enakim iterativnim postopkom, kot v raziskavi [12], tvoril množice podobnih slik obrazov. V tej raziskavi so se najprej naučili modela za kategorizacijo starosti na celotni učni množici. Z naučenim modelom so klasificirali vse učne slike. Vse slike, ki so bile dovolj pravilno klasificirane (znotraj mej dovoljene napake) so dali v prvo skupino podobnih obrazov. Model za kategorizacijo starosti so potem zgradili na množici nepravilno klasificiranih slik. Učne slike obrazov, ki so bile sedaj pravilno klasificirane so dali v drugo skupino podobnih obrazov. Postopek so nadaljevali tako dolgo dokler ni bila večina učnih slik klasificiranih pravilno (običajno 95%). S tem so dobili 5 do 7 množic podobnih obrazov. Jaz sem pri izbiri pravilno klasificiranih učnih slik obrazov uporabil mejo dovoljene napake 6 let. Za kategorizacijo

starosti sem uporabil kvadratno funkcijo, ki se je pri [12] najbolj obnesla. Za novo sliko obraza je bilo zopet potrebno najprej izbrati množico slik z najbolj podobnimi obrazi. Množice slik podobnih obrazov so bile večje kot pri modeliranju starosti pri vsaki osebi posebej, zato sem lahko uporabil Mahalanobisovo razdaljo.

Množice podobnih obrazov pa sem se poskusil naučiti še na drug način - z uporabo algoritma *k*-means, ki razdeli učno množico primerov v *k* skupin tako, da vsak primer pripada skupini z najbližjim povprečjem. Postopek klasifikacije nove slike obraza je enak kot če so množice podobnih obrazov zgrajene na način [12]. V Matlabu sem za gradnjo skupin na tak način uporabil klic funkcije `kmeans`:

```
cluster = kmeans(X,k,'distance','correlation'),
```

kjer je `cluster` vektor, ki vsebuje oznake skupin podobnih obrazov, `X` je matrika učnih slik obrazov, `k` je število skupin, argumenta `'distance'` in `'correlation'` posredujeta funkciji zahtevo za uporabo razdalje, ki upošteva tudi korelacijo podatkov in je bolj robustna kot evklidska razdalja. Skupaj sem pri klasifikaciji specifični na videz preizkusil 5 različnih nastavitev.

V članku [12] je poleg dvo-stopenjske kategorizacije specifične na videz (klasifikacija v skupino podobnih obrazov in potem ocena starosti) omenjena tudi dvo-stopenjska kategorizacija specifična na starost, zato sem preizkusil tudi več-stopenjsko kategorizacijo, kjer se postopoma razpolavljajo starostni razredi. Zaradi narave SVM je ta postopek že sam po sebi vključen v kategorizacijo s SVM, jaz pa sem ga preizkusil še pri regresiji in najbližjih sosedih. Postopek razpolavljanja razredov je že opisan pod naslovom Metoda podpornih vektorjev. Pri regresiji in najbližjih sosedih sem preizkusil oba načina razpolavljanja: iterativnega, kjer se razpolavlja starosti v dva razreda od mlajšega do starejšega razreda in rekurzivnega, kjer se razpolavljanje začne približno pri povprečni starosti in nadaljuje v dve smeri. Preizkusil sem tudi dvo-stopenjsko klasifikacijo specifično na starost, kjer se pri prvi klasifikaciji najprej klasificira v združene sosedne razrede in potem v drugi klasifikaciji klasificira v enega od združenih razredov. Pri dvo-stopenjski klasifikaciji sem preizkusil dvakratno klasifikacijo z najbližjimi sosedi (NN, NN) in klasifikacijo kjer sem pri prvi klasifikaciji uporabil najbližje sosede v drugi pa kvadratni model (NN, KM). Skupaj sem pri klasifikaciji specifični na starost preizkusil 6 različnih nastavitev. Ker rezultati niso bili boljši ali pa skoraj neopazno boljši, nisem preizkušal še kombinirane klasifikacije specifične na videz in specifične na starost.

Poglavje 4

Eksperimentiranje

V tem poglavju je opisana podatkovna baza, ki sem jo uporabljal pri eksperimentiranju z metodami za kategorizacijo starosti iz slik obrazov, kako sem eksperiment izvajal in kakšni so bili eksperimentalni rezultati.

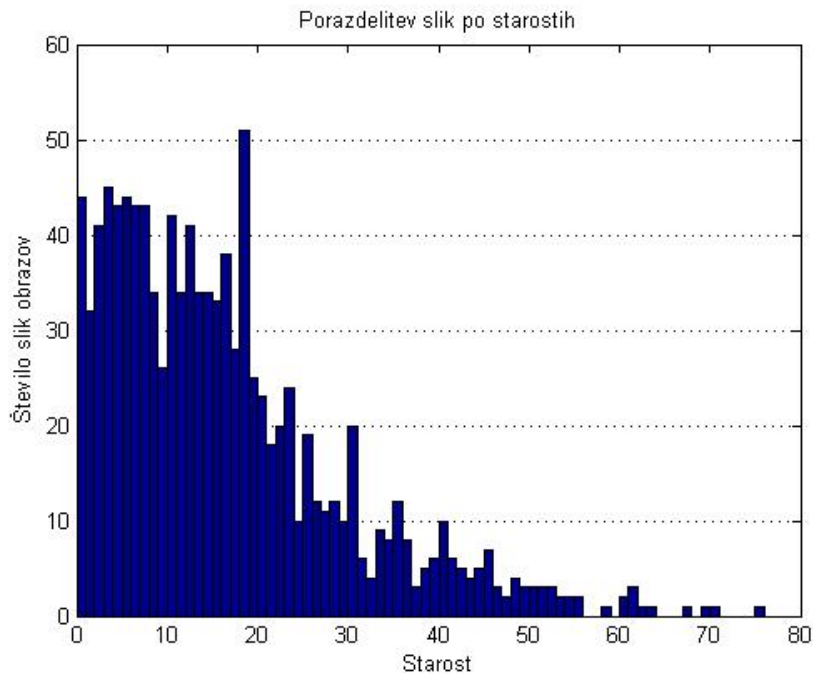
4.1 Podatkovna baza



Slika 4.1: Primeri slik obrazov iz podatkovne baze FG-NET Aging Database. Zgornja vrsta slik: naključno izbrane slike iz podatkovne baze, spodnja vrsta slik: zaporedje slik iz baze za osebo z zaporedno številko 67.

Pri eksperimentiranju sem uporabljal podatkovno bazo FG-NET Aging Database [7]. Baza vsebuje 1002 slike obrazov, ki pripadajo 82 različnim osebam. Starosti so v intervalu od 0 do 69 let. Primer slik iz baze prikazuje slika 4.1. Skupaj s slikami, ki so v formatu jpg, sta za vsako sliko v bazi še datoteka pts, ki vsebuje lokacije za 68 označenih točk na obrazu in datoteka dat, ki vsebuje specifikacije za posamezno sliko. Na sliki 3.1 je primer slike označene s točkami na enak način kot so označene slike v podatkovni bazi FG-NET Aging Database. Slike obrazov v podatkovni bazi imajo poleg variacije starosti še

množico drugih variacij kot je orientacija glave, osvetljenost, obrazni izraz, zakrivanje, itd. Od vsakega, ki uporablja to bazo, se pričakuje, da prispeva nekaj novih slik za drugi del baze B, ki bo na razpolago, ko se bo zbralo dovolj slik. Tako sem tudi sam zbral in s točkami označil 69 slik, ki pripadajo 5 različnim osebam (starosti so v intervalu od 0 do 75 let), in jih poslal vzdrževalcem podatkovne baze FG-NET Aging Database. Vseh slik skupaj, na katerih sem izvajal eksperiment je bilo tako 1071. Kako so slike iz baze FG-NET Aging Database in slike, ki sem jih zbral jaz, porazdeljene po starostih, prikazuje graf 4.2.



Slika 4.2: Porazdelitev mojih slik in slik iz podatkovne baze FG-NET Aging Database po starostih.

4.2 Potek eksperimenta

Pri eksperimentiranju z različnimi metodami sem uporabil princip pusti-eno-osebo-zunaj (leave-one-person-out). Podatkovna baza, ki je opisana v prejšnjem podrazdelku, je tako zahtevala 87 ponovitev učenja in testiranja. Zaradi omejenosti števila slik v posameznem starostnem razredu sem moral določene

razrede združiti, da je bilo možno testirati vse metode. Mahalanobisova razdalja namreč zahteva vsaj toliko podatkov, kakršna je dimenzija podatka, pri Gaussovih modelih pa mora biti število podatkov v razredu večje kot je dimenzija podatka. Slike obrazov so imele največjo dimenzijo (39), ko so bile kodirane s PCA analizo, se pravi, da je moralo biti v vsakem razredu vsaj 40 slik obrazov. Za PCA analizo sem izbral dimenzijo 39 prav zaradi tega, ker je bilo število razredov pri vseh 87 ponovitvah testiranja 4, 5 in 9 za širino razreda 15, 10 in 5 let po vrsti. Če bi pri analizi PCA obdržal več kot 39 glavnih komponent pri npr. širini razreda 15 let že ne bi mogel sestaviti 4 razrede za vsako ponovitev učenja ampak nekje le 3 in klasiﬁkacijskih točnosti ne bi mogel primerjati. Porazdelitev učnih slik obrazov po prilagojenih starostnih razredih s širino 5, 10 in 15 let je prikazana v tabelah 4.1.

Oznaka razreda	Obdobje let	Število slik
1	0-4	205
2	25-9	190
3	10-14	185
4	15-19	175
5	20-24	95
6	25-29	64
7	30-34	47
8	35-44	64
9	45-75	46

Oznaka razreda	Obdobje let	Število slik
1	0-9	395
2	10-19	360
3	20-29	159
4	30-39	81
5	40-75	76

Oznaka razreda	Obdobje let	Število slik
1	0-14	580
2	15-29	334
3	30-44	111
4	45-75	46

Tabela 4.1: Porazdelitev slik po starostnih razredih s širino 5, 10 in 15 let. Zaradi majhnega števila slik obrazov pri starejših letih, so nekateri razredi združeni

Kot sem napisal je bilo pri metodi PCA zmanjšanje dimenzije slike obraza najmanjše. Poleg tega sem se moral sam odločiti koliko dimenzij bom obdržal. Če število učnih slik ne bi bilo omejeno, bi lahko v skrajnem primeru obdržal

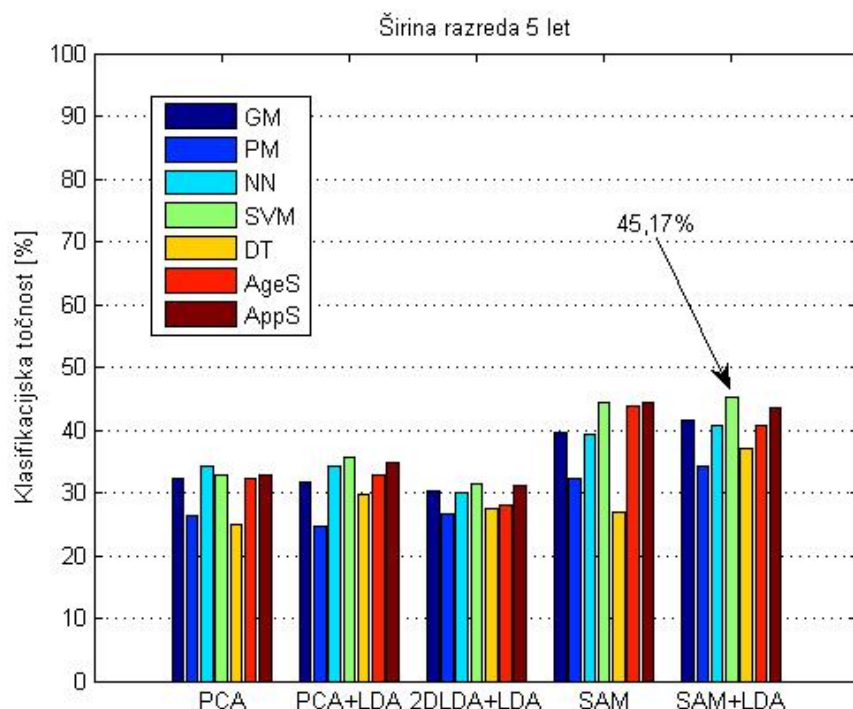
vse dimenzije (za sliko z velikostjo 40x40 slikovnih elementov je dimenzija 1600), kar je enako, kot če dimenzije slike ne bi zmanjševal in bi za klasifikacijo v starostne razrede uporabil vse slikovne elemente. Vse druge metode za kodiranje obrazov v podprostor so zmanjšale slike na manjše število dimenzij. Pri metodi LDA in metodah PCA+LDA, 2DLDA+LDA, SAM+LDA, kjer nastopa LDA v drugi fazi, je dimenzija slik omejena s številom razredov. Pri razredih s širino 5, 10 in 15 let je bila tako dimenzija kodirane slike 8, 4 in 3, gledano po vrsti. Pri metodi statističnih modelov oblike in videza je bila slika obraza zmanjšana na 30 dimenzij oz. na 30 parametrov videza.

4.3 Eksperimentalni rezultati

Na grafih 4.3, 4.4, 4.5 so prikazane klasifikacijske točnosti pri opisanih metodah za kodiranje slik obrazov v podprostor in pri opisanih metodah za kategorizacijo starosti pri širinah starostnih razredov 5, 10 in 15 let po vrsti. Slike obrazov so poravnane po postopku, ki je opisan v razdelku 3.2 Normalizacija slik obrazov. Resolucija poravnanih slik obrazov je 40x47 slikovnih elementov. Tabela s klasifikacijskimi točnostmi se nahaja v dodatku A (tabela A.1).

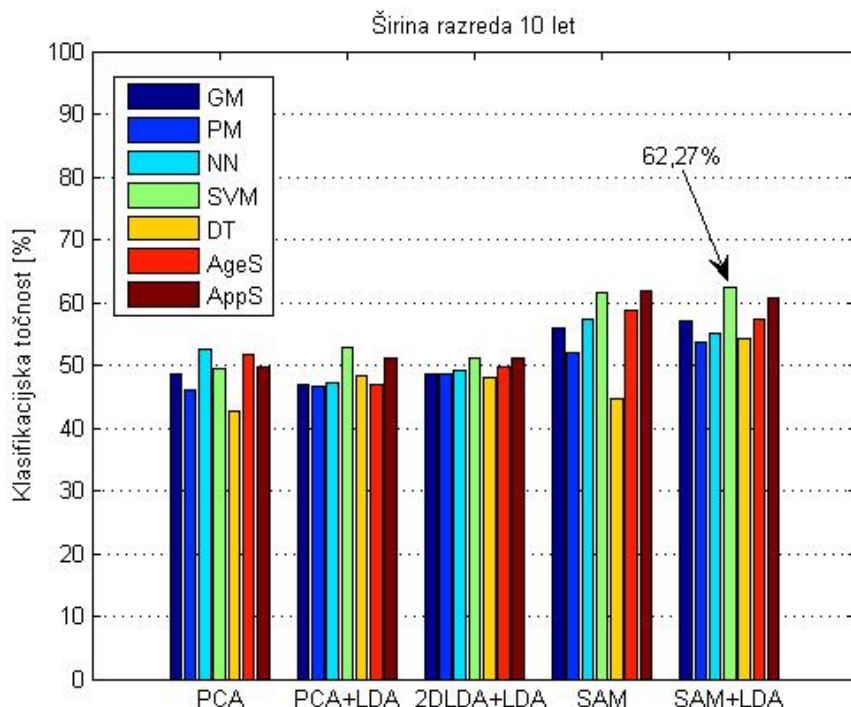
Pri glavnih rezultatih, ki jih prikazujejo grafi 4.3, 4.4 in 4.5, sem primerjal naslednje metode za kodiranje obrazov v podprostor: PCA, PCA+LDA, 2DLDA+LDA, SAM, SAM+LDA in naslednje klasifikacijske metode: Gaussov model (GM), naučen z EM algoritmom; polinomski model (PM), rešen z regresijo; najbližji sosedi (NN), kjer je razdalja Mahalanobisova; metoda podpornih vektorjev (SVM); odločitveno drevo (DT); najboljša na starost specifična metoda (AgeS); najboljša na videz specifična metoda (AppS). Za najboljšo na starost specifično metodo se je izkazala dvo-stopenjska klasifikacija z najbližjimi sosedi in polinomskim modelom. Za najboljšo na videz specifično metodo pa se je izkazala klasifikacija z najbližjimi sosedi v skupino podobnih obrazov (skupine so bile zgrajene z algoritmom k-means) in SVM za končno klasifikacijo starosti.

Pri širini razreda 5 let je bila najvišja klasifikacijska točnost 45,17% dosežena na metodi SAM+LDA in SVM. Pri širini razreda 10 let je bila najvišja klasifikacijska točnost 62,27% dosežena pri enaki kombinaciji metod. Pri širini razreda 15 let je bila najvišja klasifikacijska točnost 72,8% dosežena pri SAM in klasifikaciji specifični na videz (k-means, SVM). Kot se vidi iz grafov je metoda SVM najbolj zanesljiva, ravno tako tudi modeli oblike in videza (SAM in SAM+LDA) povsod vrnejo najboljše rezultate.



Slika 4.3: Klasifikacijske točnosti kategorizacije starosti iz slik obrazov za glavne metode za širino razreda 5 let. Na abscisni osi so grupirane metode za kodiranje slik obrazov v podprostor (PCA, PCA+LDA, 2DLDA+LDA, SAM, SAM+LDA). V posamezni grupi so na abscisni osi po vrsti razvrščene klasifikacijske metode: Gaussov model (GM), polinomski model (PM), najbližji sosedi (NN), metoda podpornih vektorjev (SVM), odločitveno drevo (DT), najboljša na starost specifična nastavitev (AgeS) in najboljša na videz specifična nastavitev (AppS).

Kot sem že omenil rabi metoda LDA dovolj veliko število učnih primerov, da matrika razpršenosti ne postane singularna in se postopek zmanjševanja dimenzije podatkov pravilno izvede. Na slikah obrazov z resolucijo 40x47 je dimenzija podatkov 1880, kar je več kot število slik v učni množici v nekaterih ponovitvah učenja, zato LDA ne uspe. Zaradi tega sem LDA preizkusil na slikah z velikostjo 20x24. Zaradi primerjave sem izvedel na manjših slikah še ostale metode. Klasifikacijske točnosti metode LDA so prikazane na grafu 4.6. Zaradi primerjave so zraven prikazane še klasifikacijske točnosti metode 2DLDA+LDA. Klasifikacijske točnosti vseh metod na slikah z resolucijo 20x24

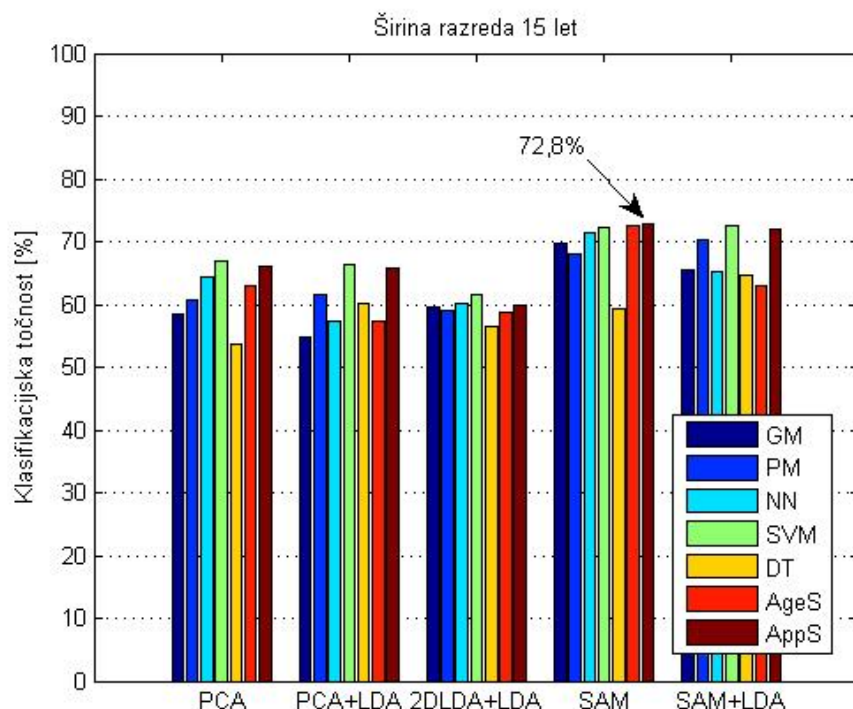


Slika 4.4: Klasifikacijske točnosti kategorizacije starosti iz slik obrazov za glavne metode za širino razreda 10 let.

so zapisane v tabeli A.2. Na grafu in v tabeli se vidi, da se metoda LDA obnese slabše kot ostale metode.

Kot se vidi iz grafov 4.3, 4.4, 4.5 in 4.6 dobimo pri vseh treh različnih širinah razredov najboljše rezultate s statističnimi modeli oblike in videza (SAM), kar pomeni, da informacija o obliki pri kategorizaciji starosti iz slik obrazov pride zelo prav. Pri dodatni obdelavi modelov SAM z LDA se rezultati v splošnem ne izboljšajo ampak so pri vseh metodah za klasifikacijo bolj podobni. Pri vseh metodah za klasifikacijo se lahko opazi, da naknadna obdelava z metodo LDA v splošnem ne izboljša rezultata, kakor se je pričakovalo, ampak ga za večino metod celo poslabša. Razlog je verjetno v temu, da je na voljo premalo komponent (dimenzija je premajhna), da bi klasifikator natančneje napovedal starost. Pri širini razredov 15, kjer dobimo 4 razrede, je dimenzija kodirane slike obraza po uporabi metode LDA samo 3.

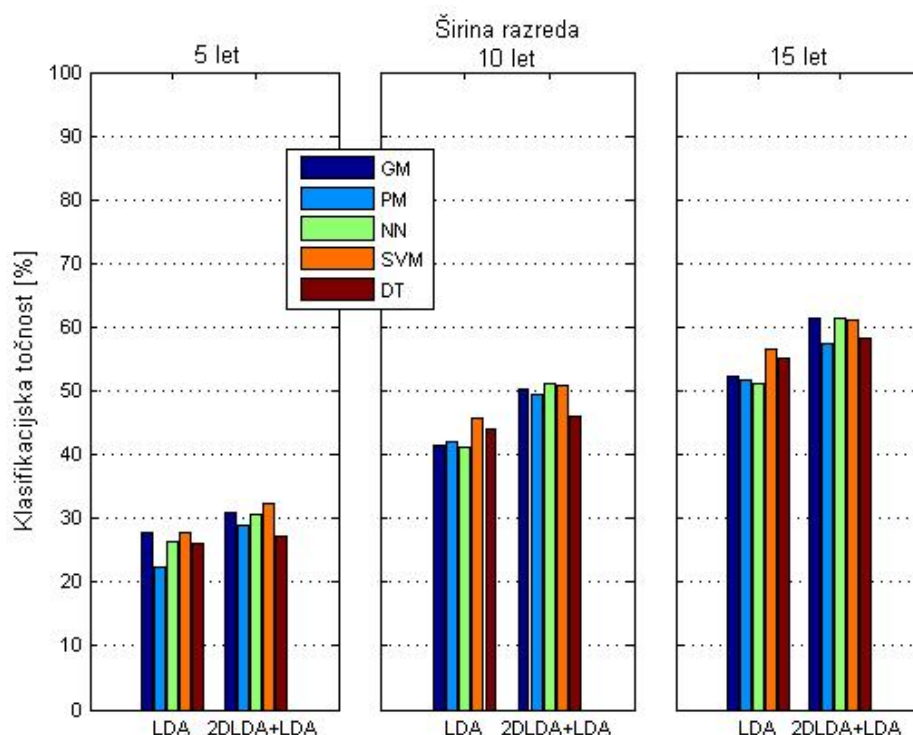
Pri metodi SAM (grafi 4.3, 4.4 in 4.5) se zelo slabše obnese le odločitveno



Slika 4.5: Klasifikacijske točnosti kategorizacije starosti iz slik obrazov za glavne metode za širino razreda 15 let.

drevo. Razlog tiči v dimenziji kodirane slike, ki je pri SAM večja kot pri metodah, kjer je naknadno uporabljena metoda LDA. Tudi pri PCA je ugotovitev za odločitveno drevo podobna. Rezultate odločitvenega drevesa se bi verjetno dalo popraviti s spreminjanjem parametrov pri gradnji, vendar se na to nisem osredotočil, ker so bile druge metode bolj obetavne.

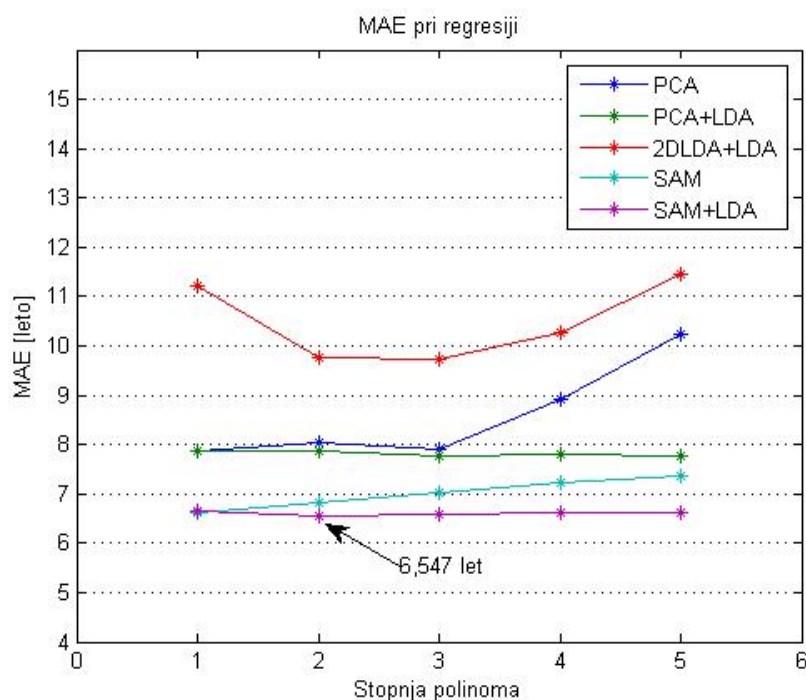
Posebej je narejena primerjava polinomskih modelov z različnimi stopnjami. Ti modeli v osnovi ocenijo kar dejansko starost osebe, zato se lahko naredi primerjavo med njimi s srednjo absolutno napako v letih (Mean Absolute Error - MAE) med ocenjeno in pravo starostjo osebe na sliki. Rezultati so v tabeli A.3 in na grafu 4.7). Opazil sem, da med njimi ni bistvene razlike in da z večanjem stopnje in s tem kompleksnosti nič ne pridobimo. Podobno so ugotovili tudi v raziskavi [13]. Le med linearnim in kvadratnim modelom je bila v tej raziskavi razlika velika. Uporabljali so sicer bolj kompleksen postopek za iskanje koeficientov. Ker najboljši rezultat (6.547 let) tudi pri meni vrne



Slika 4.6: Klasifikacijska točnost metode LDA v primerjavi z metodo 2DLDA+LDA pri resoluciji slik 20x24 slikovnih elementov.

kvadratni model, sem pri primerjavi glavnih metod na grafih 4.3, 4.4 in 4.5 in v tabeli A.1 prikazal klasifikacijsko točnost kvadratnega modela. Pri srednji absolutni napaki moram omeniti to, da je za PCA in SAM enaka za vse širine starostnih razredov. Pri metodah, ki uporabljajo tudi LDA pa je za širino razreda 5, 10 in 15 let različna, ker LDA kodira slike obrazov v podprostor na podlagi razredov. Na grafu 4.7 je pri metodah PCA+LDA, 2DLDA+LDA, SAM+LDA prikazana samo MAE za širino razreda 10 let, v tabeli A.3 pa so zapisane vse MAE.

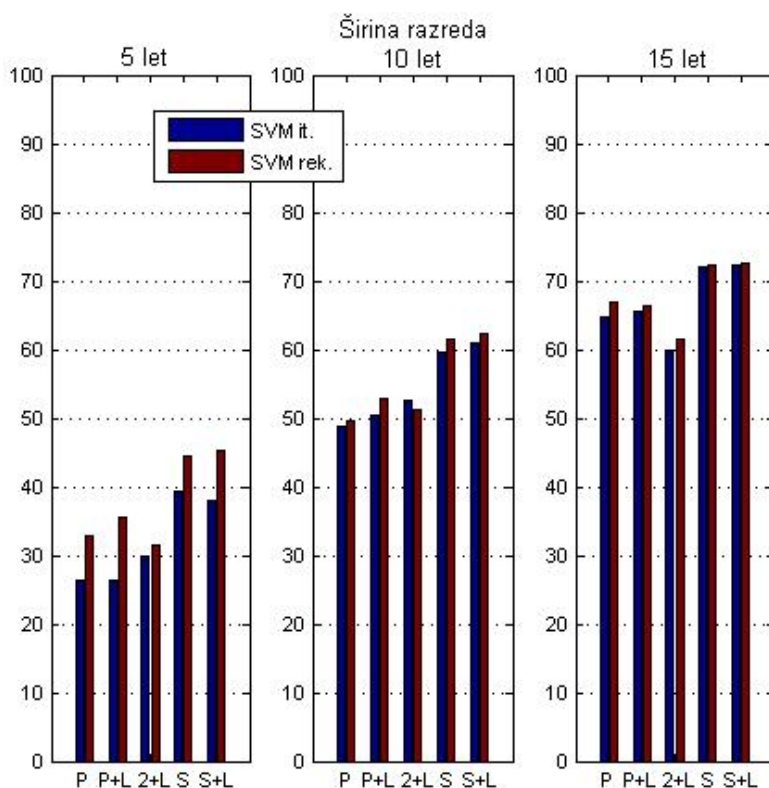
Pri metodi SVM sem opisal dve možnosti za učenje klasifikacijskega modela. Kot se vidi iz grafa 4.8 se v večini primerov bolje obnese rekurziven postopek, kjer prično ločevati razrede v dva razreda na povprečni meji razredov in ločevanje nadaljujemo v levo in desno smer. Klasifikacijske točnosti so zapisane v tabeli A.4.



Slika 4.7: Srednje absolutne napake pri polinomskih modelih od 1. do 5. stopnje. Medtem ko je za metodi PCA in SAM MAE enak pri vseh različnih širinah razredov, se ostale metode, ki so osnovane na metodi LDA ali imajo v drugem koraku metodo LDA razlikujejo. MAE za te metode je vzeta za širino razreda 15 let.

Tudi za klasifikacijo specifično na videz in klasifikacijo specifično na starost sem v glavnih rezultatih prikazal le klasifikacijske točnosti z najboljšimi nastavitvami. Na grafih 4.9 in 4.10 je prikazana primerjava različnih nastavitvev za več-stopenjsko klasifikacijo specifično na starost in klasifikacijo specifično na videz. Nisem dobil toliko boljših rezultatov kot raziskovalci v [12]. Pri večini nastavitvev sem dobil še slabše rezultate kot pri eno-stopenjski klasifikaciji.

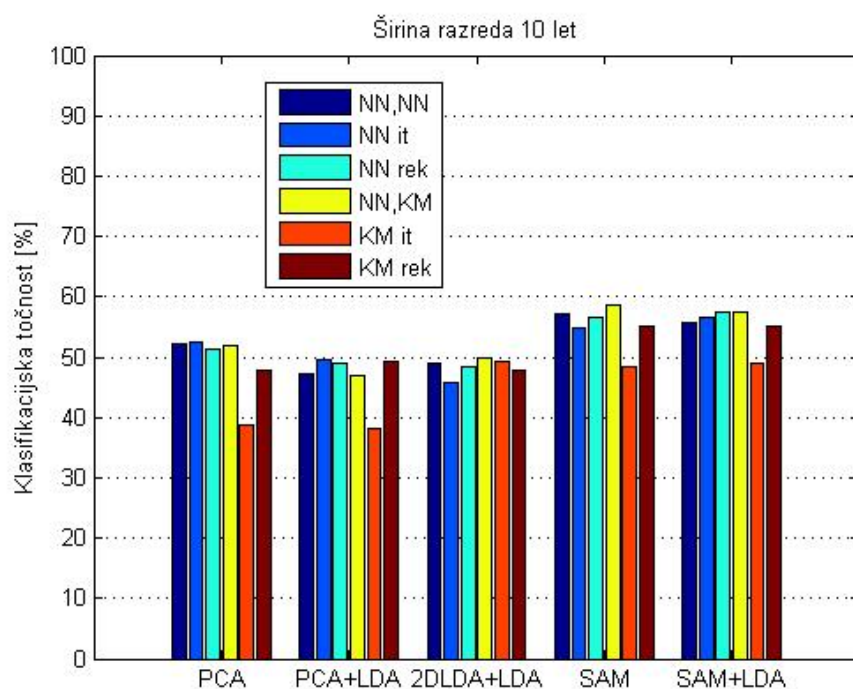
Pri klasifikaciji specifični na starost (graf 4.9) se je najbolje obnesla dvo-stopenjska klasifikacija z najbližjimi sosedi in polinomskim modelom (NN,KM). Pri širini razreda 5 let (43,88%) in 15 let (72.64%) je bila zelo blizu najboljše rešitve, ampak metode SVM, ki tudi deluje na več-stopenjski način, ni prehitela. Čeprav skupen rezultat s klasifikacijo specifično na starost ni bil izboljšan pa je bila malo izboljšana klasifikacijska točnost za metode, ki



Slika 4.8: Klasifikacijska točnost za dva načina SVM. Pri prvem načinu, ki deluje iterativno, sestavljamo razrede od mlajšega do starejšega. Pri drugem načinu, ki deluje rekurzivno, pa začnemo razrede sestavljati na povprečni meji med razredi. Namesto celotnih okrajšav za metode podprostora so na abscisni osi podane samo začetne črke okrajšav.

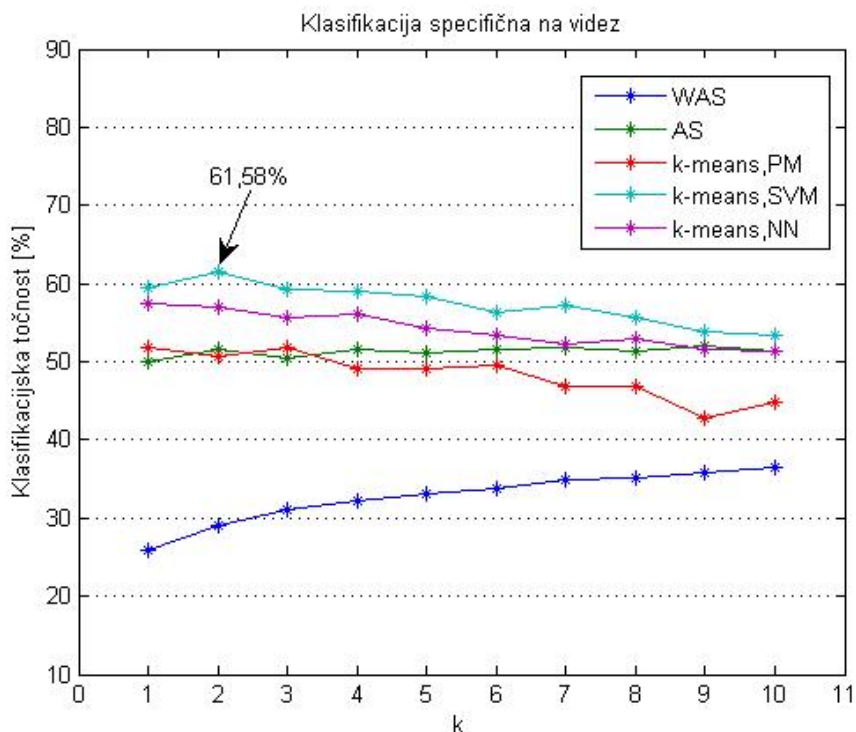
nastopajo v več-stopenjski klasifikaciji specifični na starost. Izboljšava je bila bolj opazna pri raziskovalcih [12], ki so sicer ocenjevali dejansko starost. Klasifikacijske točnosti različnih nastavitvev več-stopenjske klasifikacije specifične na starost so zapisane v tabeli A.5.

Čeprav veliko raziskovalcev omenja, da je pomembno obravnavati specifičnost staranja posamezne osebe, meni tega v klasifikator ni uspelo vgraditi tako, da bi se rezultati bistveno izboljšali. Klasifikacijske točnosti pri klasifikaciji specifični na videz pri širini razreda 10 let za modele oblike in videza pri vrednostih parametra k od 1 do 10 so prikazane na sliki 4.10. Numerične vrednosti so zapisane v tabeli A.6.



Slika 4.9: Klasifikacijska točnost pri klasifikaciji specifični na starost. Uporabil sem šest različnih nastavitvev: dvo-stopenjska klasifikacija z najbližjimi sosedi (NN,NN), več-stopenjska klasifikacija z NN z razpolavljanjem razredov od mlajšega do starejšega (NN it.), več-stopenjska klasifikacija z NN z razpolavljanjem razredov od povprečne meje med razredi (NN rek.), dvo-stopenjska klasifikacija z najbližjimi sosedi in kvadratnim modelom (NN, KM), več-stopenjska klasifikacija z KM z razpolavljanjem razredov od mlajšega do starejšega (KM it.), več-stopenjska klasifikacija z KM z razpolavljanjem razredov od povprečne meje med razredi (KM rek.).

Pri klasifikaciji specifični na videz se otežena na videz specifična metoda (Weighted Appearance Specific - WAS) obnese veliko slabše kot eno-stopenjska metoda s polinomskim modelom. Slabo se obnese pri vseh širinah razredov, ni važno kakšen k (število najbližjih oseb, ki se upoštevajo pri kategorizaciji starosti) vzamemo. Preizkusil sem k od 1 do 10. Na grafu 4.10 se vidi, da se z večanjem k klasifikacijska točnost počasi povečuje (pri $k = 10$ je enaka 36,47%), vendar ne pride niti blizu klasifikacijski točnosti eno-stopenjskega kvadratnega modela (51,87%). Slabe rezultate lahko razložim s tem, da se pri WAS sestavijo



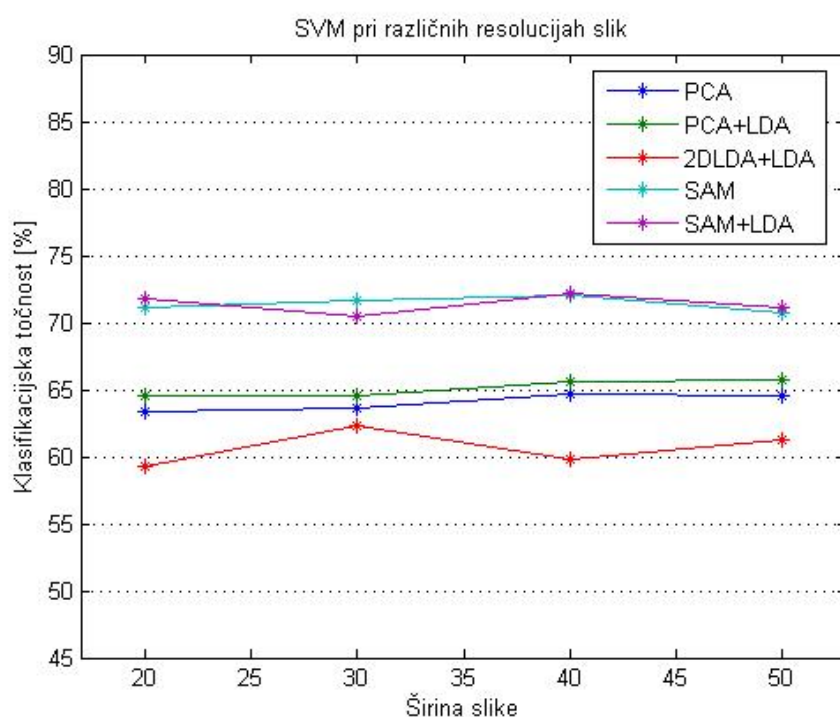
Slika 4.10: Klasifikacijska točnost pri klasifikaciji specifični na videz pri modelih oblike in videza za širino razreda 10 let. Prikazano je, kako vpliva parameter k na klasifikacijsko točnost. Pri WAS k pomeni število najbližjih oseb, ki so se upoštevale pri kategorizaciji; pri AS k pomeni mejo dovoljene napake pri sestavljanju množic podobnih slik obrazov; pri vseh kombinacijah z k-means pa k pomeni število množic podobnih slik obrazov.

individualni kvadratni modeli, ki so zaradi majhnega števila slik obrazov, ki jih ima posamezna oseba, nezanesljivi. Zaradi majhnega števila slik obrazov pri računanju najbližjih oseb tudi ne moremo uporabiti Mahalanobisove razdalje ampak Evklidsko, ki je manj robustna. V raziskavi [13] so imeli manjšo bazo z manj osebami, razpon starosti je bil bolj strnjen, zaradi česar so bili modeli bolj zanesljivi.

Klasifikacija specifična na videz z iterativnim postopkom, ki je opisan v [12] (na grafu 4.10 označeno z AS), da skoraj neopazno boljše rezultate. Klasifikacijska točnost je običajno le malo boljša za dopustno napako od 6 do 9 let pri ocenjevanju starosti na učnih slikah, kjer dobimo od 6 do 10 množic podobnih obrazov. Na primer, za širino razreda 10 let je klasifikacijska točnost boljše le

pri dopustni napaki 9 let ($k=9$) in znaša 52,12%, kar je le malo boljše kot pri eno-stopenjski klasifikaciji s kvadratnim modelom, kjer znaša 51,87%. Pri [12] so uspešno zgradili od 5 do 7 skupin podobnih obrazov in dobili opazno boljše rezultate. Razlog za to se spet verjetno skriva v tem, da sem jaz uporabljal zahtevnejšo podatkovno bazo slik obrazov.

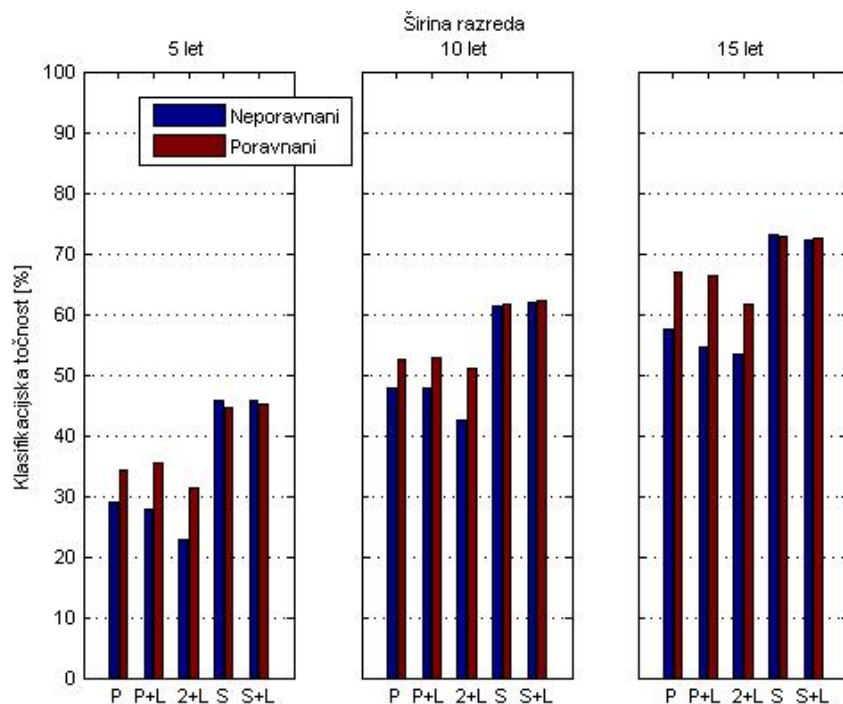
Algoritem k-means v splošnem ne izboljša rezultata ampak klasifikacijska točnost, z večanjem števila skupin podobnih obrazov (k), v večini primerov pada. Poveča se le pri SVM za $k=2$ za širino razreda 10 let (graf 4.10) in za širino razreda 15 let.



Slika 4.11: Vpliv resolucije slike na klasifikacijsko točnost pri metodi SVM. Primerjane so resolucije slik 20x24, 30x35, 40x47 in 50x58 slikovnih elementov.

Metode sem primerjal na različnih resolucijah slik in tako kot avtor v [11] opazil, da ni velikih razlik v klasifikacijski točnosti, kar pomeni, da na klasifikacijo manjši obrazni detajli ne vplivajo veliko. Na grafu 4.11 je prikazana primerjava klasifikacijskih točnosti za resolucije slik 20x24, 30x35, 40x47 in 50x58 slikovnih elementov pri različnih metodah za kodiranje slik obrazov v

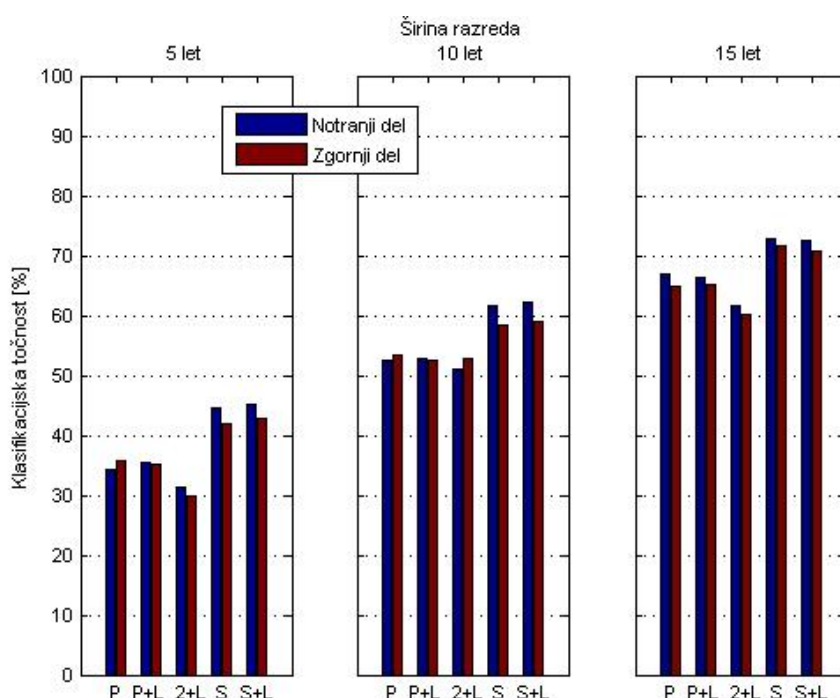
podprostor in pri metodi SVM za kategorizacijo starosti.



Slika 4.12: Klasifikacijska točnost pri neporavnanih in pri poravnanih slikah za širino razreda 5, 10 in 15 let po vrsti. Za akronime metod za kodiranje obrazov v podprostor (PCA, PCA+LDA, 2DLDA+LDA, SAM, SAM+LDA) so prikazane, zaradi pomanjkanja prostora, le začetne črke. Za vsako od metod za kodiranje obrazov v podprostor je prikazana le klasifikacijska metoda, ki je vrnila najboljše rezultate.

Metode sem primerjal tudi na poravnanih in na neporavnanih obrazih. Pri statističnih modelih oblike in videza ni bistvene razlike, saj metoda sama vsebuje normalizacijo. Za vsako sliko se namreč vzame le slikovne elemente znotraj konveksnega območja označenih točk, poleg tega se pri vsaki sliki vzame isto število slikovnih elementov znotraj tega konveksnega območja. Kot sem pričakoval so pri metodah PCA, PCA+LDA, 2DLDA+LDA razlike velike, saj se v podprostor pretvori celotna slika obraza ne samo slikovni elementi v območju označenih točk. Na grafu 4.12 je primerjana klasifikacijska točnost na neporavnanih slikah in klasifikacijska točnost na poravnanih slikah. Ker

pri nepravilnih slikah zavzema slika obraza manjše področje sem primerjal poravnane slike z resolucijo 40x47 in nepravilne slike z resolucijo 50x74. Za vsako metodo za kodiranje obrazov v podprostor je prikazana le tista klasifikacijska metoda, ki se je najbolje obnesla. Klasifikacijske točnosti vseh metod pri nepravilnih slikah obrazov so zapisane v tabeli A.7. V tej tabeli se lahko vidi, katere klasifikacijske metode so se najbolje obnesle. Za najboljše klasifikacijske metode na poravnanih slikah lahko pogledate v osnovno tabelo s klasifikacijskimi točnostmi (tabela A.1) ali na osnovne grafe s klasifikacijskimi točnostmi (grafi: 4.3, 4.4 in 4.5).



Slika 4.13: Primerjava klasifikacijske točnosti pri notranjem delu obraza (celoten obraz brez las) in pri zgornjem delu obraza. Za vsako od metod za kodiranje obrazov v podprostor je prikazana le klasifikacijska metoda, ki je vrnila najboljše rezultate..

V članku [11] je avtor ugotovil, da je za ocenjevanje starosti najpomembnejši zgornji del obraza okoli oči. Jaz pa sem pri slikah obrazov, ki sem jih porezal pod nosom in nad obrvmi (zgornji del obraza), dobil slabše rezultate

kot na slikah obrazov, ki sem jih porezal na koncu brade (notranji del obraza). Za lažjo predstavo si lahko primer slik notranjega dela obraza in zgornjega dela obraza še enkrat ogledate na sliki 3.3. Klasifikacijska točnost za kategorizacijo starosti iz zgornjega dela obraza in notranjega dela obraza je prikazana na grafu 4.13. Primerjava je narejena samo za najboljše klasifikacijske metode. Klasifikacijske točnosti pri zgornjemu delu obraza so za vse metode zapisane v tabeli A.8. Klasifikacijske točnosti za notranji del obraza so vzete iz osnovne tabele s klasifikacijsko točnostjo A.1

Poglavje 5

Zaključek

V končno aplikacijo sem vgradil iskanje značilnic z AAM modeli, statistične modele oblike in videza (SAM) za kodiranje slik obrazov v podprostor in kategorizacijo starosti po metodi SVM. Grafični vmesnik je prikazan na sliki 5.1. Za iskanje točk z AAM modeli sem uporabil preveden program, ki sem ga naložil skupaj z knjižnico *AAM-API*. Normalizacije obrazov nisem vgrajeval, ker so pri modelih SAM rezultati približno enaki.



Slika 5.1: Slika grafičnega vmesnika za kategorizacijo starosti iz slike obraza.

Sliko obraza je potrebno odpreti s klikom na gumb *Odpri sliko jpg*. Slika se prikaže na levi strani. Iskanje značilnic se izvede s klikom na gumb *Poišči točke*. Najdene točke se prikažejo na sliki. Ker so značilnice z AAM modeli zelo slabo lokalizirane, se lahko označene točke naloži s klikom na gumb *Odpri datoteko pts*. Kategorizacijo starosti se izvede s pritiskom na gumb *Določi starostni razred*, pri čemer se obdobje let, v katero je kategorizirana slika prikaže pod napisom *Starostni razred*.

Za popolno avtomatizacijo sistema bi bilo le še potrebno povezati program z zunanjo napravo za zajemanje slik, na primer s spletno kamero. Ko bi oseba pred napravo oziroma sam sistem rabil informacijo o starosti osebe, bi se avtomatično zajela slika obraza in izvedla kategorizacija starosti; na primer, ko bi uporabnik dostopal na spletno stran, ki zahteva preverjanje starosti. Prožilec kategorizacije starosti bi bil tako klik na spletno povezavo in ne zaporedje pritiskov na gumbe *Odpri sliko jpg*, *Poišči točke*, *Določi starostni razred*.

Vse to zgleda zelo enostavno, kakor da je rešitev dobra in je potrebno le zamenjati nekaj ukazov v programu, ampak ni tako. Največji problem in izziv za naprej je iskanje točk na obrazu. Modeli AAM zelo slabo najdejo točke, če jih uporabimo v osnovni različici. Zadnje čase je bilo na tem področju narejenih veliko raziskav, tako da že obstajajo izboljšani modeli AAM, ki so robustnejši na množico variacij, ki se lahko pojavijo na slikah obrazov. Kakšna bi bila klasifikacijska točnost, če ne bi izvajal eksperimenta na pravilno označenih obrazih. Jasno je, da bi bila kvečjemu nižja. Pri izboljšanih modelih AAM bi bila mogoče blizu najboljšega dobljenega rezultata.

V praksi se sistem ne bi najboljše obnesel, mogoče le na bolj specifičnih področjih. Če bi imeli opravka z manj splošnimi slikami obrazov, se bi klasifikacijska točnost dvignila. Na frontalnih obrazih so raziskovalci [18] dobili približno za 10 procentov višjo klasifikacijsko točnost (za moške slike obrazov, za ženske slike obrazov niti niso dobili veliko boljših rezultatov) z metodo 2DLDA+LDA. Na takih obrazih bi bila tudi razlika med metodami osnovanimi na videzu (PCA, LDA, 2DLDA) in modeli oblike in videza (SAM) manjša (če bi sploh bila). Za metode osnovane na videzu ni potrebno, da se najdejo vse točke na obrazu, ampak je dovolj, da se najde oči. V tem primeru bi se tako lahko izognili modelom AAM in bi uporabili enostavnejši in bolj uspešen detektor za lokalizacijo oči.

Rezultati, ki sem jih dobil so drugače primerljivi z rezultati enakih ali podobnih metod v sorodnih delih. V teh delih je narejena tudi primerjava, kako dobro bi človek ocenil starost iz slike obraza. Zaradi dejstva, da tudi ljudje ne moremo natančno oceniti starosti obraza, rezultati računalnikov niso slabi. Pri metodi AGES računalnik vrne celo boljše ocene kot človeški opazovalec.

Metode AGES, ki se trenutno obnese najboljše, nisem implementiral, ker sem jo našel razmeroma pozno. Z njo bi mogoče klasifikacijsko točnost dvignil še za nekaj procentov. Vprašanje je, če se da starost oceniti še boljše kot pri metodi AGES. Predvidevam, da ne bistveno boljše.

Dodatek A

Numerični rezultati eksperimentov

Š. r.	Metoda	GM	PM	NN	SVM	DT	AgeS	AppS
5	PCA	32.34	26.41	34.12	32.78	24.84	32.33	32.69
	PCA+LDA	31.76	24.79	34.12	35.49	29.73	32.92	34.65
	2DLDA+LDA	30.41	26.62	29.93	31.41	27.57	27.98	31.21
	SAM	39.56	32.38	39.41	44.45	26.76	43.88	44.50
	SAM+LDA	41.42	34.16	40.81	45.17	37.06	40.82	43.50
10	PCA	48.60	45.99	52.49	49.55	42.65	51.79	49.86
	PCA+LDA	46.98	46.60	47.13	52.91	48.20	46.93	51.01
	2DLDA+LDA	48.72	48.70	49.12	51.11	47.98	49.82	51.01
	SAM	55.81	51.87	57.34	61.45	44.53	58.76	61.71
	SAM+LDA	57.06	53.54	54.98	62.27	54.18	57.43	60.81
15	PCA	58.35	60.81	64.41	66.87	53.59	62.92	66.14
	PCA+LDA	54.87	61.48	57.26	66.45	60.23	57.27	65.71
	2DLDA+LDA	59.70	58.94	60.04	61.51	56.39	58.73	59.98
	SAM	69.73	67.94	71.47	72.17	59.39	72.64	72.80
	SAM+LDA	65.57	70.40	65.20	72.45	64.76	63.05	71.96

Tabela A.1: Klasifikacijska točnost v procentih pri glavnih metodah na slikah obrazov z resolucijo 40x47. Po vrsticah so metode za kodiranje slik obrazov v podprostor pri širini razreda 5, 10 in 15 let. Po stolpcih so metode za klasifikacijo.

Š. r.	Metoda	GM	PM	NN	SVM	DT	AgeS	AppS
5	PCA	33.17	23.98	35.07	32.91	25.19	33.15	33.83
	LDA	27.82	22.39	26.23	27.65	26.12	25.09	28.09
	PCA+LDA	31.39	25.94	31.88	33.64	31.26	32.76	35.90
	2DLDA+LDA	31.04	28.95	30.69	32.23	27.27	31.29	31.98
	SAM	38.53	31.63	39.09	43.65	27.51	42.69	43.08
	SAM+LDA	40.02	32.39	39.88	44.18	40.06	39.91	43.36
10	PCA	50.85	44.92	53.27	50.18	41.46	51.57	51.60
	LDA	41.50	41.96	41.20	45.67	44.07	42.65	45.16
	PCA+LDA	46.78	47.54	46.52	52.29	46.49	47.08	52.54
	2DLDA+LDA	50.30	49.29	51.03	50.90	46.01	50.18	50.08
	SAM	55.10	51.85	57.17	60.44	46.10	57.32	58.42
	SAM+LDA	55.54	53.80	53.63	59.27	51.92	54.50	59.68
15	PCA	59.49	60.06	63.94	65.39	53.60	63.01	66.01
	LDA	52.32	51.85	51.09	56.62	55.13	49.77	56.75
	PCA+LDA	57.94	62.74	57.56	64.85	57.56	57.44	65.77
	2DLDA+LDA	61.30	57.40	61.30	61.02	58.26	57.95	59.59
	SAM	68.22	68.80	69.92	71.68	56.85	70.87	72.28
	SAM+LDA	65.73	69.54	64.66	71.83	65.94	64.00	70.53

Tabela A.2: Klasifikacijska točnost v procentih pri resoluciji slik obrazov 20x24. Po vrsticah so metode za kodiranje slik obrazov v podprostor pri širini razreda 5, 10 in 15 let. Po stolpcih so metode za klasifikacijo.

Polinom		1.st.	2.st.	3.st.	4.st.	5.st.
PCA		7.85	8.03	7.89	8.92	10.23
PCA+LDA	5	7.82	7.87	7.77	7.82	7.88
	10	7.82	7.75	7.75	7.72	7.78
	15	7.88	7.88	7.76	7.81	7.77
2DLDA+LDA	5	10.84	9.52	10.08	12.44	21.26
	10	10.77	9.04	8.95	9.85	13.18
	15	11.22	9.75	9.74	10.27	11.47
SAM		6.63	6.84	7.03	7.22	7.36
SAM+LDA	5	6.61	6.64	6.65	6.70	6.82
	10	6.63	6.56	6.52	6.52	6.53
	15	6.64	6.55	6.57	6.61	6.61

Tabela A.3: Srednja absolutna napaka v letih pri polinomskih modelih od 1. do 5. stopnje. Medtem ko je za metodi PCA in SAM srednja absolutna napaka (MAE) enaka pri vseh različnih širinah razredov, se ostale metode, ki so osnovane na metodi LDA ali imajo v drugem koraku metodo LDA, razlikujejo. Zaradi tega je pri metodah LDA MAE podan za vse tri širine razredov: 5, 10 in 15 let.

Š. r.	5 let		10 let		15 let	
Metoda	SVM it	SVM rek	SVM it	SVM rek	SVM it	SVM rek
PCA	26.37	32.78	48.72	49.55	64.63	66.87
PCA+LDA	26.44	35.49	50.51	52.91	65.60	66.45
2DLDA+LDA	29.91	31.41	52.66	51.11	59.81	61.51
SAM	39.34	44.45	59.51	61.45	72.07	72.17
SAM+LDA	38.07	45.17	61.03	62.27	72.20	72.45

Tabela A.4: Klasifikacijska točnost pri dveh načinih uporabe metode SVM.

Š.r.	Metoda	NN, NN	NN it	NN rek	NN, PM	PM it	PM rek
5	PCA	33.99	32.62	33.11	32.33	17.33	23.04
	PCA+LDA	33.15	34.42	32.80	32.92	17.26	25.31
	2DLDA+LDA	28.88	24.42	28.01	27.98	22.82	21.94
	SAM	40.61	34.26	39.83	43.88	24.96	27.89
	SAM+LDA	39.52	40.07	40.97	40.82	26.88	28.17
10	PCA	52.25	52.40	51.35	51.79	38.76	47.85
	PCA+LDA	47.13	49.57	49.09	46.93	38.05	49.26
	2DLDA+LDA	48.93	45.78	48.44	49.82	49.12	47.88
	SAM	57.16	54.89	56.56	58.76	48.39	54.99
	SAM+LDA	55.79	56.58	57.32	57.43	49.00	55.12
15	PCA	63.74	65.47	63.74	62.92	58.04	62.17
	PCA+LDA	56.69	59.05	56.69	57.27	60.57	63.47
	2DLDA+LDA	59.81	58.96	59.81	58.73	59.87	59.69
	SAM	70.97	70.89	70.97	72.64	64.93	71.03
	SAM+LDA	63.27	66.05	63.27	63.05	69.35	69.00

Tabela A.5: Klasifikacijska točnost pri klasifikaciji specifični na starost. Po vrsticah so metode za kodiranje slik obrazov v podprostor. Po stolpcih so več-stopenjske metode za klasifikacijo.

Š. r.	k	WAS	AS	kmeans		
				PM	SVM	NN
5	1	16.04	31.36	32.38	44.45	42.74
	2	17.21	31.83	30.83	44.41	42.02
	3	16.54	31.40	29.36	42.79	41.70
	4	17.65	32.16	30.38	42.39	40.40
	5	18.43	31.49	28.32	40.33	38.68
	6	18.71	32.78	30.58	36.27	38.40
	7	20.22	32.73	29.73	37.91	36.80
	8	20.29	32.17	28.89	35.78	37.76
	9	19.59	32.13	27.66	36.55	35.75
	10	19.72	31.96	26.49	32.87	33.88
10	1	25.92	50.10	51.87	59.51	57.34
	2	29.13	51.51	50.58	61.58	57.01
	3	31.15	50.41	51.81	59.26	55.67
	4	32.15	51.47	49.11	59.07	56.13
	5	33.14	51.20	49.14	58.31	54.26
	6	33.75	51.58	49.56	56.27	53.44
	7	34.80	51.72	46.74	57.13	52.24
	8	35.09	51.33	46.90	55.58	52.84
	9	35.81	52.12	42.82	53.86	51.68
	10	36.47	51.25	44.76	53.40	51.44
15	1	36.84	65.18	67.94	72.17	65.79
	2	39.48	65.95	67.26	72.72	67.62
	3	40.45	67.10	65.90	71.01	65.90
	4	44.15	67.25	64.99	70.92	65.13
	5	44.32	66.96	62.70	70.29	65.75
	6	45.25	67.66	63.60	70.09	65.39
	7	45.59	68.28	61.67	68.97	63.11
	8	47.29	67.59	62.12	67.77	62.69
	9	48.09	67.89	60.83	66.92	62.40
	10	48.05	67.40	60.54	64.79	62.85

Tabela A.6: Klasifikacijska točnost v procentih pri klasifikaciji specifični na videz pri širinah razredov 5, 10 in 15 let in pri vhodnem parametru k od 1 do 10. Slike so bile kodirane z modeli SAM. Uporabljene so bile naslednje metode: WAS (Weighted Appearance Specific) [13], AS (Appearance Specific) [12], k-means skupaj z polinomoskim modelom (kmeans,PM), metodo SVM (kmeans,SVM) in najbližjimi sosedi (kmeans,NN).

Š. r.	Metoda	GM	PM	NN	SVM	DT	AgeS	AppS
5	PCA	27.94	18.90	28.11	27.80	19.05	28.96	26.35
	PCA+LDA	21.62	19.88	23.06	27.81	24.24	25.41	27.14
	2DLDA+LDA	21.87	19.05	21.87	22.70	19.36	20.94	22.17
	SAM	39.19	33.47	38.77	44.96	29.92	42.57	45.80
	SAM+LDA	39.86	33.78	40.15	45.70	37.94	40.75	44.84
10	PCA	40.41	37.61	45.73	46.86	34.88	47.72	45.02
	PCA+LDA	38.79	39.86	39.91	47.19	41.24	39.71	47.72
	2DLDA+LDA	39.88	39.22	39.38	42.38	37.17	40.12	41.52
	SAM	56.83	53.20	57.21	61.35	47.01	57.39	60.91
	SAM+LDA	55.52	54.10	53.89	61.91	53.94	55.83	59.95
15	PCA	52.40	55.53	57.48	54.97	47.85	56.67	56.57
	PCA+LDA	44.92	53.56	45.00	54.52	47.69	47.12	54.10
	2DLDA+LDA	50.86	53.40	49.10	53.46	50.04	48.84	53.41
	SAM	68.98	68.65	70.22	72.51	57.22	73.15	71.95
	SAM+LDA	65.44	69.92	64.63	72.29	64.82	64.83	70.96

Tabela A.7: Klasifikacijska točnost v procentih pri uporabi neporavnanih slik obrazov z resolucijo 50x74.

Š. r.	Metoda	GM	PM	NN	SVM	DT	AgeS	AppS
5	PCA	35.68	22.52	35.47	35.15	26.79	34.18	34.66
	PCA+LDA	29.02	27.23	33.18	35.12	28.48	32.16	34.41
	2DLDA+LDA	28.58	25.93	28.42	29.87	24.27	28.10	29.94
	SAM	39.32	31.26	38.38	41.84	26.59	40.65	39.44
	SAM+LDA	39.49	29.05	38.93	42.92	35.78	39.97	42.36
10	PCA	47.91	42.14	51.57	53.28	40.16	51.05	53.43
	PCA+LDA	45.55	44.10	46.61	52.57	46.72	45.99	51.85
	2DLDA+LDA	50.48	48.10	49.44	52.69	46.81	50.52	52.21
	SAM	53.65	49.91	56.57	58.34	45.04	57.71	58.20
	SAM+LDA	54.92	50.12	54.88	59.10	51.32	57.06	56.56
15	PCA	57.40	60.34	63.42	64.39	51.19	64.07	64.78
	PCA+LDA	54.89	60.15	56.48	65.12	58.92	56.48	64.02
	2DLDA+LDA	60.13	58.58	59.70	60.11	57.69	57.85	59.63
	SAM	68.09	66.06	69.50	71.18	55.30	70.80	71.47
	SAM+LDA	63.78	68.02	64.32	70.83	64.65	62.65	69.73

Tabela A.8: Klasifikacijska točnost v procentih pri uporabi zgornjega dela slike obraza z resolucijo 40x27.

Slike

2.1	Obrazna slika in njena rekonstrukcija z modelom oblike in videza.	9
3.1	Slika obraza označena s točkami.	12
3.2	Rotacija slike obraza.	14
3.3	Rezanje slike obraza.	15
3.4	Primer <i>smd</i> datoteke.	19
4.1	Primeri slik iz baze FG-NET Aging Database.	27
4.2	Porazdelitev slik iz podatkovne baze po starostih.	28
4.3	Klasifikacijske točnosti glavnih metod (širina razreda 5 let).	31
4.4	Klasifikacijske točnosti glavnih metod (širina razreda 10 let).	32
4.5	Klasifikacijske točnosti glavnih metod (širina razreda 15 let).	33
4.6	Klasifikacijska točnost pri metodi LDA.	34
4.7	Srednja absolutna napaka v letih pri polinomskih modelih.	35
4.8	Klasifikacijska točnost pri dveh načinih SVM.	36
4.9	Klasifikacijska točnost pri klasifikaciji specifični na starost.	37
4.10	Klasifikacijska točnost pri klasifikaciji specifični na videz.	38
4.11	Klasifikacijska točnost pri različnih resolucijah slik.	39
4.12	Klasifikacijska točnost pri neporavnanih in pri poravnanih slikah.	40
4.13	Klasifikacijska točnost pri notranjem in pri zgornjem delu obraza.	41
5.1	Slika grafičnega vmesnika za kategorizacijo starosti.	43

Tabele

2.1	Rezultati pri [13].	7
2.2	Rezultati pri [8].	10
4.1	Porazdelitev slik po starostnih razredih.	29
A.1	Klasifikacijska točnost pri glavnih metodah.	46
A.2	Klasifikacijska točnost pri resoluciji slik 20x24.	47
A.3	Srednja absolutna napaka v letih pri polinomskih modelih.	48
A.4	Klasifikacijska točnost pri dveh načinih SVM.	48
A.5	Klasifikacijska točnost pri klasifikaciji specifični na starost.	49
A.6	Klasifikacijska točnost pri klasifikaciji specifični na videz.	50
A.7	Klasifikacijska točnost pri neporavnanih obraznih slikah.	51
A.8	Klasifikacijska točnost pri zgornjem delu slike obraza.	51

Literatura

- [1] am_tools. Dostopno na:
http://personalpages.manchester.ac.uk/staff/timothy.f.cootes/software/-am_tools_doc/download_win.html (29.4.09)
- [2] AAM-API. Dostopno na:
<http://www2.imm.dtu.dk/aam/> (29.4.09)
- [3] AAMTools. Dostopno na:
<http://cvsp.cs.ntua.gr/software/AAMtools/> (29.4.09)
- [4] AAM Fitting Algorithms. Dostopno na:
http://www.ri.cmu.edu/research_project_detail.html?project_id=448&-menu_id=261 (29.4.09)
- [5] G. J. Edwards, A. Lanitis, C. J. Taylor, T. F. Cootes, “Statistical Models of Face Images: Improving Specificity”, *Image and Vision Computing*, št. 16, zv. 3, str. 203–211, 1998.
- [6] G. J. Edwards, C. J. Taylor, T. F. Cootes, “Interpreting Face Images using Active Appearance Models”, v zborniku *3rd. International Conference on Face and Gesture Recognition*, Nara, Japonska, april 1998, str. 300-305.
- [7] FG-NET Aging Database. Dostopno na:
<http://www.fgnet.rsunit.com/> (17.4.2009)
- [8] X. Geng, K. Smith-Miles, Z.-H. Zhou, “Automatic Age Estimation Based on Facial Aging Patterns”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, št. 29, zv. 12, str. 2234-2240, 2007.
- [9] Y. H. Kwon, N. da Vitoria Lobo, “Age Classification from facial images”, *Computer Vision Image Understanding*, št. 74, zv. 1, str 1-21, 1999.

- [10] R. Iga, K. Izumi, H. Hayashi, G. Fukano, T. Ohtani, "A gender and age estimation system from face images", v zborniku *Annual conference on Society of Instrument and Control Engineers*, Fukui, Japonska, avg. 2003, str. 756-761.
- [11] A. Lanitis, "On the significance of different facial parts for automatic age estimation", v zborniku *14th Intl. Conference on Digital Signal Processing*, Santorini, Greece, julij 2002, strani 1027-1030.
- [12] A. Lanitis, C. Draganova, C. Christodoulou, "Comparing different classifiers for automatic age estimation", *IEEE Transactions on Systems, Man and Cybernetics*, št. 34, zv. 1, str. 621-628, 2004.
- [13] A. Lanitis, C. Taylor, T. Cootes, "Toward automatic simulation of aging effects on face images", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, št. 24, zv. 4, str. 442-455, 2002.
- [14] E. Makinen, *Face Analysis Techniques for Human-Computer Interaction*, Academic Dissertation, Department of Computer Sciences, University of Tampere, 2007.
- [15] Matlab Toolbox for Dimensionality Reduction. Dostopno na: http://ticc.uvt.nl/~lvdrmaaten/Laurens_van_der_Maaten/Matlab_Toolbox_for_Dimensionality_Reduction.html (9.5.09)
- [16] L.J.P. van der Maaten. An Introduction to Dimensionality Reduction Using Matlab. Technical Report MICC 07-07. Maastricht University, Maastricht, The Netherlands, 2007.
- [17] R. Ravnik, B. Batagelj, F. Solina, *Računalniški vid in digitalna karakterizacija za oglaševanje*, Fakulteta za računalništvo in informatiko, 2009.
- [18] K. Ueki, T. Hayashida, T. Kobayashi, "Subspace-based age-group classification using facial images under various lightening conditions," v zborniku *7th IEEE International Conference on Automatic Face and Gesture Recognition*, Sauthampton, UK, april 2006, str. 43-48.
- [19] T. Wilhelm, H. Bohme, H. Groos, "Classification of face images for gender, age, facial expression, and identity", v zborniku *15th International Conference on Artificial Neural Networks: Biological Inspirations*, 2005, str. 569-574.