

Određivanje starosti osobe sa slike

Damjanović, Tamara

Undergraduate thesis / Završni rad

2016

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **Josip Juraj Strossmayer University of Osijek, Faculty of Electrical Engineering, Computer Science and Information Technology Osijek / Sveučilište Josipa Jurja Strossmayera u Osijeku, Fakultet elektrotehnike, računarstva i informacijskih tehnologija Osijek**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:200:913527>

Rights / Prava: [In copyright](#) / [Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2024-05-16**

Repository / Repozitorij:

[Faculty of Electrical Engineering, Computer Science and Information Technology Osijek](#)



DIGITALNI AKADEMSKI ARHIVI I REPOZITORIJ

SVEUČILIŠTE JOSIPA JURJA STROSSMAYERA U OSIJEKU
FAKULTET ELEKTROTEHNIKE, RAČUNARSTVA I
INFORMACIJSKIH TEHNOLOGIJA

Sveučilišni studij

ODREĐIVANJE STAROSTI OSOBE SA SLIKE

Završni rad

Tamara Damjanović

Osijek, 2016.

Sadržaj

1.	UVOD	1
1.1.	Zadatak završnog rada	2
2.	OPIS OSNOVNIH METODA OBRADE SLIKE	3
2.1.	Proces starenja ljudskog lica.....	3
2.2.	Postojeći pristupi rješavanju problema procjene starosne dobi	3
2.3.	Opis metoda korištenih pri određivanju starosne dobi	5
2.3.1.	Analiza glavnih komponenti (engl. <i>Principal Component Analysis</i>).....	5
2.3.2.	Linearna diskriminantna analiza (engl. <i>Linear Discriminant Analysis</i>)	8
2.3.3.	k-NN (engl. <i>k-Nearest Neighbors</i>).....	11
2.3.4.	Stroj s potpornim vektorima (engl. <i>Support Vector Machine</i>).....	13
3.	PROGRAMSKA IZVEDBA.....	16
4.	ZAKLJUČAK	19
	LITERATURA.....	20
	SAŽETAK.....	22
	ABSTRACT	23
	ŽIVOTOPIS	24

1. UVOD

Ljudske karakteristike, posebice karakteristike lica (starost, spol, etničko podrijetlo, osobni identitet), oduvijek su predmet istraživanja raznih znanosti s različitim ciljevima. Razvoj računalne grafike i računalnog vida uvećao je zanimanje za rješavanje problema vezanih uz ovu tematiku, s ciljem primjenjivanja rješenja u raznim aplikacijama i sustavima – najvećim dijelom sigurnosnim, ali i forenzičkim, biometrijskim, medicinskim, ali i u reklamnoj industriji te industriji zabave (filmovi, video spotovi). Postoje dvije različite grane unutar računalne znanosti koje se bave pitanjima sinteze i analize slika lica: sinteza starosnih karakteristika (engl. *age synthesis*) – izdvajanje slika lica s obzirom na starosnu dob te primjena učinaka pomlađivanja, procjena starosne dobi (engl. *age estimation*) – svrstavanje osoba sa slike u dobni razred ili točno određivanje starosti. Problem određivanja starosti osobe sa slike nije prisutan u tolikoj mjeri kao proces sinteze starosnih karakteristika [1]. Sam problem razlaže se na dva procesa: prikaz određene slike lica (razlaganje slike lica na više dijelova) te primjena tehnika određivanja starosne dobi.

1.1. Zadatak završnog rada

Potrebno je izraditi sustav koji je u mogućnosti na temelju ulazne slike lica odrediti starost osobe i svrstati je u jednu od četiri kategorije:

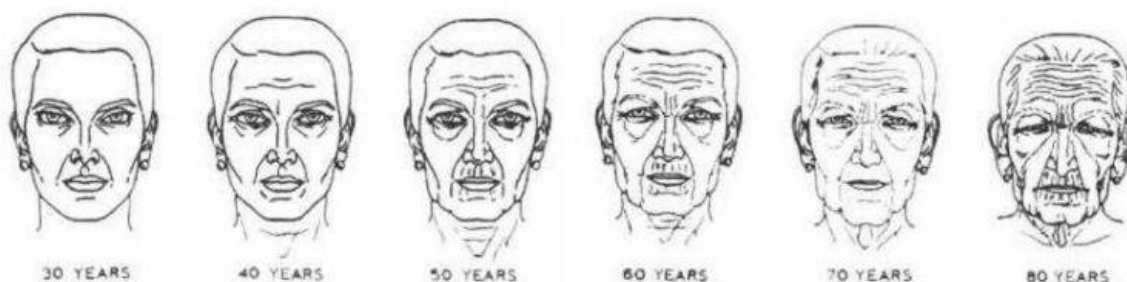
- djeca do 13 godina
- adolescenti (13-21g.)
- odrasle osobe (21-60g.)
- starije osobe (više od 60 g.)

Zbog navedenog razloga potrebno je upoznati se s osnovnim metodama obrade slike koji će u tome pomoći i ukratko ih opisati (analiza glavnih komponenti, linearna diskriminantna analiza, potporni vektorski stroj, algoritam k -najbližih susjeda...). Krajnji cilj je implementirati i testirati rješenje u programskom jeziku po izboru.

2. OPIS OSNOVNIH METODA OBRADJE SLIKE

2.1. Proces starenja ljudskog lica

Ljudsko starenje je dug proces, naizgled jedinstven svakoj osobi. No, moguće je izdvojiti neke opće promjene ljudskog lica tijekom procesa starenja – promjene u samoj strukturi kostiju (rast i pomicanje) te promjene vidljive na koži (pojava bora, slabljenje mišića).



Sl. 2.1. Skica starenja ljudskog lica [1]

Od rođenja do odraslog doba najznačajniji je rast kostiju ljudske lubanje. Kao posljedica rasta, čelo se nagiba i smanjuje, ostavljajući prostor za razvoj očiju, nosa, ušiju i usta. Obrazi se izdužuju, a brada postaje „isturenija“. U odrasloj dobi promjene lubanje nisu toliko vidljive koliko promjena kože – gubitak kolagena i gravitacija uzrokuju stanjivanje i tamnjenje kože, kao i gubitak elastičnosti te pojavu bora [2].

2.2. Postojeći pristupi rješavanju problema procjene starosne dobi

Procjena starosne dobi može se smatrati problemom klasifikacije (engl. *classification problem*), ali i regresijskim problemom (engl. *regression problem*), prema [2]. Ako se svaki starosni razred smatra klasom koja ima svoja jedinstvena obilježja u odnosu na druge klase, tj. starosne razrede, problemu se pristupa koristeći klasifikacijske algoritme. S druge strane, problem se definira kao regresijski ako se na godine gleda kao na niz brojeva (ne postoje razredi). Kwon i Lobo (1999.) su prvi pokušali razviti algoritam za određivanje starosne dobi. Koristili su geometrijske omjere lica

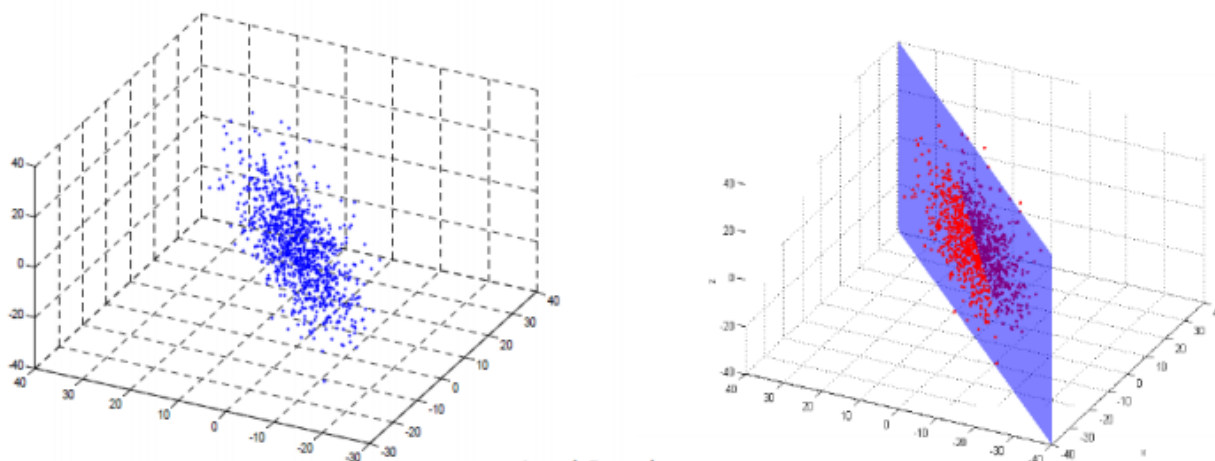
karakteristične određenoj starosnoj skupini kao i količinu bora. Uspjeli su klasificirati lica na lica beba, odraslih osoba i starijih osoba. Lanitis (2002.,2004.) je funkciju starenja prikazao u obliku kvadratnih funkcija, tzv. *Active Appearance Model*. Geng (2007.) je koristio slike iste osobe kroz godine kako bi odredio obrazac starenja. Ova se metoda pokazala učinkovitijom od prije spomenutih. Fu i Huang (2008.) su za razliku od Genga koristili slike različitih osoba iste starosti kako bi definirali obrazac starenja. Na taj je način učinkovitost povećana. Guo (2008.) je koristio tzv. regresijski stroj s potpornim vektorima (engl. *Support Vector Machine Regressor*) za povezivanje kodirane slike lica sa starosnom dobi. Wang (2009.) je pomoću klasifikacijske metode, koja primjenjuje izlazne kodove za ispravljanje pogrešaka (engl. *Error-Correcting Output Codes - ECOC*) na određenim značajkama lica, uspio kategorizirati osobe u četiri dobne skupine (djeca, adolescenti, odrasli, stariji). Za razliku od većine opisanih istraživanja koja su promatrala ljudsko lice u cjelini, Suo je hijerarhijski podijelio lice na tri dijela – prvu razinu predstavlja općeniti prikaz lica, druga razina se temelji na pojedinačnim regijama lica, a treća na detaljima kao što su bore i početna crta rasta kose. Eksperimentalni rezultati ukazuju na važnost detalja (treće razine). Točno procjenjivanje starosne dobi zahtjevan je zadatak za ljude, stoga se razvijanjem računalne metode želi smanjiti mogućnost greške i automatizirati sam proces, koji bi imao raznovrsnu primjenu. Upravo se tu javljaju razni izazovi, od problema kako detektirati lice na slici te kako sama poza lica, svjetlosni učinci, kvaliteta fotografije te mnoge druge „sitnice“ utječu na samu procjenu, do nekih specifičnijih problema kao što su određivanje dobnih razreda (često se javljaju greške u procjenjivanju kod krajnjih godina razreda koji obuhvaćaju odrasle pojedince), ovisnost o vanjskim čimbenicima (životni stil pojedinca, zdravstvena skrb, psihologija, estetska kirurgija) te možda jedan od većih problema – postojeće baze podataka. Kako bi algoritam postao učinkovitiji, potrebno ga je „trenirati“ (strojno učenje) i testirati. Postoji nekoliko baza podataka, neke od njih su javne, poput *FG-NET*-a (sastoji se od 1002 visoko-rezolucijske crno-bijele fotografije ili u boji, u rasponu od 0 do 69 godina, s kronološkim fotografijama pojedinca), *MORPH*-a (nastala je na Sveučilištu Sjeverna Karolina u Wilmingtonu za potrebe biometrijskih aplikacija, a sastoji se od dva albuma – prvi album sadrži 1724 fotografije lica, a drugi 20000), Gallagherove web kolekcije (engl. *Gallagher's Web-Collected Database*) - sadrži 5080 fotografija skupljenih putem *Flickr.com* web tražilice.

2.3. Opis metoda korištenih pri određivanju starosne dobi

Prilikom određivanja starosne dobi subjekta koristi se nekoliko metoda, svaka od njih s određenim ciljem. U nastavku slijedi opis te značaj svake od njih.

2.3.1. Analiza glavnih komponenti (engl. *Principal Component Analysis*)

U znanstvenim istraživanjima obrada podataka koji se dobivaju iz eksperimenata važna je komponenta jer usmjerava samo istraživanje i utječe na konačne zaključke istoga. Često se promatra više parametara (npr. brzina, tlak zraka, masa, visina) te je potrebno odrediti koji parametri utječu na samo ponašanje i djelovanje određenog sustava, odnosno koji su bitni, a koji nisu [3]. Upravo je to temelj analize glavnih komponenata, poznate i pod nazivom Karhunen-Loeve proširenje (engl. *Karhunen-Loeve expansion*). Matematička definicija je: statistička metoda kod koje se niz varijabli za koje postoji vjerojatnost korelacije ortogonalnom transformacijom prebacuje u linearni niz varijabli koje međusobno ne koreliraju (sl. 2.2.). One se nazivaju glavne komponente, a njihov broj obično je jednak broju varijabli (može biti i manji). Ovu metodu osmislio je prvi Karl Pearson 1901.godine, kasnije ju je neovisno razvio i imenovao Harold Hotelling 1930.godine. Ima široku primjenu, najčešće prilikom vizualizacije podataka, kompresije slika, u latentnoj semantičkoj analizi (engl. *Latent Semantic Analysis*) i kao dio određenih algoritama, poput Kleinbergovog HITS (engl. *Hyperlink-Induced Topic Search*) algoritma i *PageRank* algoritma kojeg koristi Google tražilica.



Sl. 2.2. Prikaz PCA metode nad skupom podataka [4]

Matematički gledano, analiza glavnih komponenata temelji se na prikazu podataka pomoću matrice; s obzirom da je linearna metoda, moguće je koristiti razne matematičke operacije proizašle iz teorije linearne algebre. Početna matrica X je dimenzija $m \times n$, gdje m (redak matrice) predstavlja broj raznih vrsta mjerenja (različitih parametara), a n (stupac matrice) broj mjerenja, tj. broj subjekata/objekata nad kojima je mjerenje izvršeno. Nadalje, traži se ortonormalna (ortogonalna i normirana) matrica P gdje će vrijediti slijedeća jednakost:

$$Y = PX \quad (2-1)$$

Jednadžba (2-1) predstavlja promjenu baze. Promjena baze omogućuje izdvajanje najvažnijih značajki iz niza podataka. X predstavlja originalni niz podataka, odnosno početnu matricu, Y prikaz podataka, a P je transformativni faktor. Upravo su redci matrice P set novih baza vektora za prikaz stupaca početne matrice X , odnosno redci matrice P predstavljaju glavne komponente (engl. *principal components*) matrice X . Promjenom baze ne utječe se na podatke, već samo na njihov prikaz – reprezentaciju, jer podaci, odnosno vektori (ako se podaci promatraju kao vektori), samo se projiciraju na vektore baze. Geometrijski gledano matrica P predstavlja rotaciju i produženje matrice X . Matrica P mora biti takva da vrijedi slijedeća jednadžba:

$$S_Y = \frac{1}{n-1} YY^T \quad (2-2)$$

S_Y predstavlja dijagonaliziranu matricu kovarijanci. Kovarijanca je mjera linearne ovisnosti dviju varijabli. Matrica kovarijanci je kvadratna (broj redaka jednak je broju stupaca, tj. dimenzije su joj $m \times m$), simetrična matrica čiju dijagonalu čine varijance (procjena preciznosti mjernog postupka u statistici koja se često koristi umjesto standardnog odstupanja) određenog tipa mjerenja (parametara), a ostali elementi (izvan dijagonale) kovarijance između različitih tipova mjerenja (parametara). S obzirom da je cilj smanjiti redundanciju (zalihost), svaka varijabla bi u što manjoj mjeri trebala korelirati s drugima, odnosno kovarijance između različitih mjerenja trebale bi biti jednake nuli (svi elementi koji nisu na dijagonali jednaki su nuli). Stoga, varijance opisuju najvažnija svojstva, tj. omogućuju izdvajanje signala od šumova, a uklanjanjem redundancije dobiva se dijagonalizirana matrica. Zato se PCA može definirati kao linearna transformacija u novi koordinatni sustav u kojem najveća varijanca leži na prvoj koordinati te se naziva prva glavna komponenta, druga najveća varijanca leži na drugoj koordinati, i tako dalje. Y^T predstavlja transponiranu matricu matrice Y – redci matrice Y jednaki su stupcima matrice Y^T . Nadalje, moguće je zamijeniti Y u jednadžbi (2-2) s jednakošću iz jednadžbe (2-1).

$$S_Y = \frac{1}{n-1} YY^T = \frac{1}{n-1} (PX)(PX)^T = \frac{1}{n-1} PXX^T P^T = \frac{1}{n-1} (XX^T)P^T = \frac{1}{n-1} PAP^T \quad (2-3)$$

Vidljivo je da je XX^T zamijenjen s matricom A , koja je simetrična što dokazuje teorem, za bilo koju matricu A vrijedi da su A i $A^T A$ i AA^T simetrični. Nadalje, prema slijedećem teoremu, neka se matrica A može ortogonalno dijagonalizirati ako postoji matrica, npr. E , takva da vrijedi slijedeća jednakost:

$$A = EDE^T \quad (2-4)$$

gdje je D dijagonalna matrica koja sadrži svojstvene vrijednosti matrice A , a E matrica svojstvenih vektora matrice A poredanih u stupce. Prije spomenuta matrica P bira se tako da svaki njen redak svojstveni vektor XX^T . Stoga, ako se u jednakosti (2-4) izvrši zamjena $P = E^T$ te ako se primijeni teorem da je inverz ortogonalne matrice jednak transponiranoj matrici, tj. $P^{-1} = P^T$, slijedi:

$$S_Y = \frac{1}{n-1} PAP^T = \frac{1}{n-1} P(P^T DP)P^T = \frac{1}{n-1} (PP^T)D(PP^T) = \frac{1}{n-1} (PP^{-1})D(PP^{-1}) = \frac{1}{n-1} D \quad (2-5)$$

Može se zaključiti da su glavne komponente matrice X svojstveni vektori XX^T , odnosno redci matrice P , a dijagonalna vrijednost matrice S_Y predstavlja varijancu matrice X izraženu pomoću vrijednosti p_i . Nakon određivanja svojstvenih vektora i svojstvenih vrijednosti, odnosno matrica E i D , stupci se poredaju od većih prema manjima, ali tako da se jednakost među stupcima dviju matrica ne mijenja te se odabire set svojstvenih vektora koji će poslužiti kao vektori baze na koje će se projicirati podaci. Svojstveni vektor v , odnosno ne-nul rješenje neke matrice A je ono koje zadovoljava jednadžbu:

$$Av = \lambda v \quad (2-6)$$

gdje λ predstavlja svojstvenu vrijednost. Svojstvena se vrijednost odabire tako da vrijedi:

$$\det(A - \lambda I) = 0 \quad (2-7)$$

PCA ima brojne prednosti koje se u određenim slučajevima mogu smatrati nedostacima, npr. PCA je neparametarska analiza te ne ovisi o korisniku, što u određenim slučajevima može usporiti cijeli algoritam (ako je korisnik iskustvom stekao određeno znanje).

Prilikom određivanja starosti osobe sa slike, analiza glavnih komponenti predstavlja prvi podsustav, a slika ulaz u taj sustav. Analogno objašnjenju principa PCA, može se zaključiti kako se svaka slika promatra kao vektor dimenzije N^2 , odnosno kao točka u prostoru dimenzija N^2 te se daljnjim

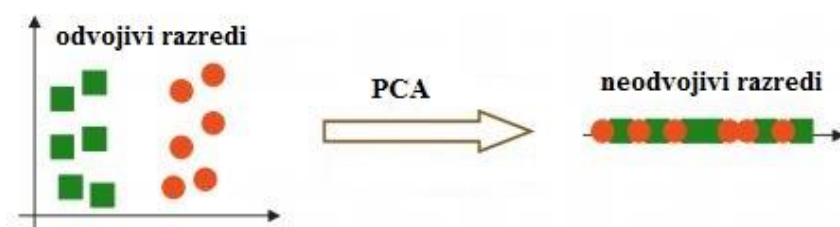
obrađivanjem određuje M-dimenzionalni potprostor čiji vektori baze odgovaraju najvećoj varijanci u prvotnom prostoru slike [5]. Taj novi potprostor naziva se prostor lica, a glavne komponente predstavljaju svojstvena lica (sl.2.3.).



Sl. 2.3. *Primjer svojstvenih lica* [6]

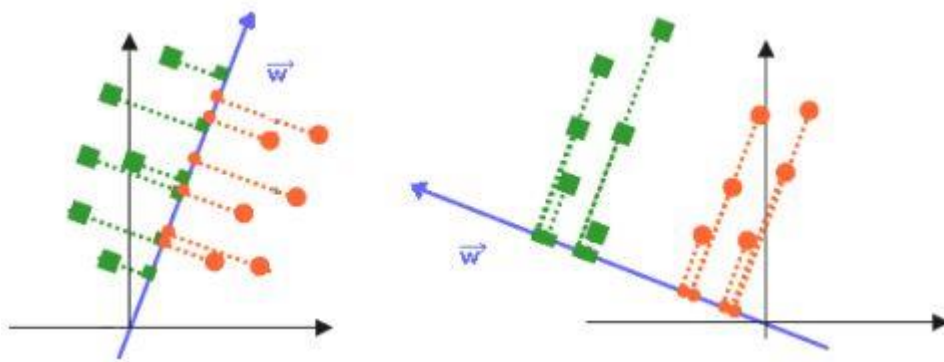
2.3.2. Linearna diskriminantna analiza (engl. *Linear Discriminant Analysis*)

Linearna diskriminantna analiza, poznata i pod nazivom Fisherova linearna diskriminanta (prema engleskom znanstveniku Ronaldu Fisheru koji ju je osmislio), je metoda koja se koristi u statistici, prepoznavanju obrazaca te strojnom učenju kako bi se pronašla linearna kombinacija značajki koje predstavljaju razliku između dvije ili više klasa objekata ili događaja [7]. Ovaj se matematički postupak, kao i prethodno opisana analiza glavnih komponenata (engl. *PCA*), koristi za reduciranje dimenzionalnosti. Razlika je u tome što linearna diskriminantna analiza nastoji povećati odvojivost razreda u prostoru reducirane dimenzije i time ubrzati proces klasifikacije.



Sl. 2.4. *Prikaz djelovanja analize glavnih komponenata na različite razrede (otežana klasifikacija)* [8]

S obzirom da je eksperimentalno dokazano da primjena obiju metoda u procesu prepoznavanja lica i određivanja starosne dobi daje najbolje rezultate, svojstveni vektori, tj. svojstvena lica, dobivena analizom glavnih komponenti dalje se obrađuju linearnom diskriminantnom analizom. Cilj je pronaći svojstvena lica koja najviše pridonose razlici dobnih grupa kako bi se izgradio model pomoću kojeg bi se najbolje mogla predvidjeti dobna skupina nepoznate osobe na slici. Matematički rečeno, potrebno je pronaći orijentaciju pravca na koji se projiciraju uzorci x_i d - dimenzionalnosti pri čemu je $i = 1, 2, \dots, n$, ali tako da su projicirani uzorci odvojivi (sl.2.5.).



Sl. 2.5. Prikaz projekcije uzoraka na dva različita pravca – desna slika prikazuje bolju razdvojenost projekcija crvenih i zelenih točaka (klasa) [8]

Radi prikaza matematičkog postupka Fisherove linearne diskriminante na jednoj klasi (može se primijeniti na više klasa) pretpostavit će se da postoji skup koji sadrži n d – dimenzionalnih uzoraka $x_1, x_2, \dots, x_n$ od čega n_1 uzoraka čini podskup D_1 (označeni s w_1) i n_2 uzoraka podskupa D_2 (označeni s w_2). Linearna kombinacija komponenti vektora $\vec{x}$ tvori se:

$$y_i = \vec{w}^T \vec{x}_i \quad (2-8)$$

gdje je $i=1, 2, \dots, n$. Traži se vektor $\vec{w}$ za koji su projekcije uzoraka odvojive. Mjeru odvojivosti predstavlja razlika srednjih vrijednosti projiciranih uzoraka. Kako bi se postigla odvojivost, potrebno je odrediti linearnu funkciju $\vec{w}^T \vec{x}$ za koju je kriterijska funkcija $J(\vec{w})$ maksimalna.

$$J(\vec{w}) = \frac{|\tilde{m}_1 - \tilde{m}_2|^2}{\tilde{s}_1^2 + \tilde{s}_2^2} \quad (2-9)$$

U jednadžbi (2-9) $\tilde{m}_1$ i $\tilde{m}_2$ predstavljaju srednje vrijednosti projiciranih uzoraka, prema izrazu

$$\tilde{m}_i = \frac{1}{n_i} \sum_{y \in Y_i} y = \frac{1}{n_i} \sum_{y \in Y_i} w^T x \quad (2-10)$$

, a $\tilde{s}_1^2 + \tilde{s}_2^2$ ukupnu mjeru raspršenosti podataka unutar jednog razreda. Iz izraza (2-9) vidljivo je da razlika srednjih vrijednosti treba biti što veća, a raspršenost projiciranih uzoraka unutar jednog razreda što manja da bi kriterijska funkcija $J(\bar{w})$ bila maksimalna. Kako bi se pronašla optimalna vrijednost $\bar{w}$, kriterijska funkcija $J(\bar{w})$ mora se eksplicitno izraziti kao funkcija $\bar{w}$. Potrebno je definirati matricu S_i i matricu S_w .

$$S_i = \sum_{x \in D_i} (x - m_i)(x - m_i)^T \quad (2-11)$$

m_i u izrazu (2-11) predstavlja srednju vrijednost zadanih uzoraka, prema jednadžbi:

$$m_i = \frac{1}{n_i} \sum_{x \in D_i} x \quad (2-12)$$

$$S_w = S_1 + S_2 \quad (2-13)$$

Nadalje, ako se u izraz (2-14), koji predstavlja mjeru raspršenosti za pojedine uzorke, uvrste izrazi (2-8) i (2-10), dobije se:

$$\tilde{s}_i^2 = \sum_{y \in Y_i} (y - \tilde{m}_i)^2 = \sum_{x \in D_i} (w^T x - w^T m_i)^2 = \sum_{x \in D_i} w^T (x - m_i)(x - m_i)^T w = \sum_{x \in D_i} w^T S_i w \quad (2-14)$$

Stoga je ukupna mjera raspršenosti jednaka:

$$\tilde{s}_1^2 + \tilde{s}_2^2 = w^T S_i w \quad (2-15)$$

Analogno, razlika srednjih vrijednosti projiciranih uzoraka je:

$$(\tilde{m}_1 - \tilde{m}_2)^2 = (w^T m_1 - w^T m_2)^2 = w^T (m_1 - m_2)(m_1 - m_2)^T w = w^T S_B w \quad (2-16)$$

U jednadžbi (2-16) matrica S_B je matrica raspršenosti između klasa. Na kraju, kriterijska funkcija $J(\bar{w})$ izražena pomoću parametra $\bar{w}$:

$$J(\vec{w}) = \frac{\vec{w}^T S_B \vec{w}}{\vec{w}^T S_W \vec{w}} \quad (2-17)$$

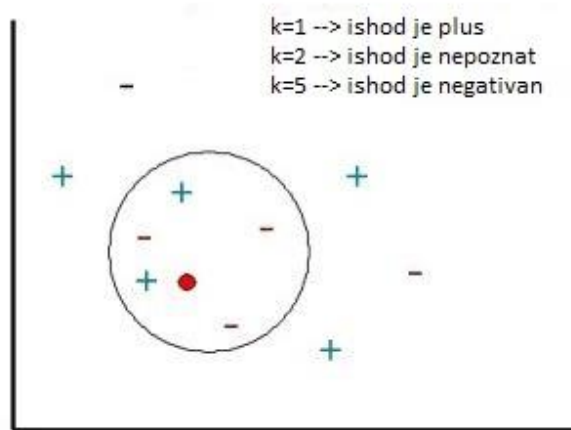
Maksimum bilo koje funkcije, tako i kriterijske, određuje se izjednačavanjem prve derivacije s nulom. Na kraju se dobiva izraz:

$$\lambda S_W w = S_B w \quad (2-18)$$

Izraz (2-18) predstavlja problem pronalaska svojstvenih vrijednosti λ . S obzirom da je u pitanju linearna diskriminantna analiza (Fisherova linearna diskriminanta), te se vrijednosti nazivaju Fisherovima vrijednostima.

2.3.3. k-NN (engl. *k-Nearest Neighbors*)

k -NN je model koji se temelji na nizu objekata (primjera) za koje je poznata određena vrijednost, tj. pripadnost klasi. Svaki se objekt sastoji od određenih podataka kod kojih je niz neovisnih vrijednosti povezan s nizom ovisnih ishoda. Ovisne i neovisne varijable mogu biti uzastopne ili kategoričke. O tome ovisi primjena ovog algoritma, kao i krajnji rezultat – za uzastopne, ovisne varijable smatra se regresijom, u ostalim slučajevima klasifikacijom [9]. Prilikom određivanja pripadnosti nekog slučaja određenoj kategoriji (klasi) na temelju već poznatih ishoda, pronalaze se primjeri k klasa koje su najbliže upitu, dok se kod regresijskih problema pomoću ovoga algoritma određuje nepoznata vrijednost objekta koja predstavlja prosjek vrijednosti k najbližih susjeda; upravo tome algoritam duguje svoj naziv. Na slici 2.6. prikazano je određivanje klase nepoznatog uzorka (označenog crvenom točkom) s obzirom na broj susjednih uzoraka koji se uzimaju u obzir (vrijednost k). Vidljivo je da odabir vrijednosti k utječe na konačan ishod. Prilikom odabira vrijednosti jedan ($k=1$), uzima se u obzir najbliži susjed, to je u ovom slučaju plus te se pretpostavlja da je i nepoznati uzorak plus. Ako na određivanje klase nepoznatog uzorka utječu dva najbliža susjeda ($k=2$), ishod je nepoznat jer je prvi najbliži susjed plus, a drugi minus. Analogno tome, kad se gleda 5 najbližih susjeda (označen krug), ishod će biti negativan. Iz primjera se jasno vidi da je odabir vrijednosti k ključan za uspješno učenje klasifikatora, a time i za svaku novu klasifikaciju.



Sl. 2.6. Prikaz određivanja klase kojoj pripada nepoznati uzorak (crvena točkica) s obzirom na pripadnost susjednih uzoraka određenim klasama [9]

Kako bi se odredila vrijednost k , koristi se algoritam unakrsne potvrde (engl. *cross-validation*). Uzorci se podijele u nekoliko v segmenata te se za neku izabranu vrijednost k primjenjuje k -NN algoritam na nekom v segmentu kako bi se procijenila greška, odnosno prilikom klasifikacije, točnost (postotak točno klasificiranih uzoraka). Postupak se ponavlja za svaki v -ti segment te se na kraju određuje prosječna vrijednost pogreške. Proces se ponavlja za više vrijednosti k , a odabire se ona vrijednost za koju je prosječna pogreška najmanja. S obzirom na duljinu ovog postupka te potrebne resurse, najbolje je postaviti vrijednost k . U ovom slučaju (prilikom određivanja starosti osobe sa slike) koristit će se najjednostavniji algoritam – vrijednost k bit će jednaka jedan. Udaljenost nepoznatog uzorka do najbližeg susjeda može se odrediti na više načina. Najčešće se koristi standardna Euklidova udaljenost (najkraća udaljenost između dvije točke u n -dimenzionalnom prostoru):

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2} \quad (2-19)$$

Koristi se i Minkowski udaljenost (udaljenost u normaliziranom vektorskom prostoru koja se smatra generalizacijom Euklidove i Manhattan udaljenosti, prema [10]):

$$\left(\sum_{i=1}^n |x_i - y_i|^p \right)^{\frac{1}{p}} \quad (2-20)$$

gdje su: $(x_1, x_2, \dots, x_i)$ i $(y_1, y_2, \dots, y_i) \in \mathbb{R}^N$.

Primjenjuje se i Mahalanobisova udaljenost (mjera različitosti dva slučajna vektora iste distribucije s matricom kovarijanci S , prema [11]):

$$d(\bar{x}, \bar{y}) = \sqrt{(\bar{x} - \bar{y})^T S^{-1} (\bar{x} - \bar{y})} \quad (2-21)$$

S obzirom da se pretpostavlja da su objekti koji se nalaze u blizini slični, algoritam k -najbližih susjeda može se modificirati uvođenjem težinskih faktora (w) za svaki od k susjeda, ovisno o udaljenosti od upita y , kao što je i vidljivo u jednadžbi (2-22).

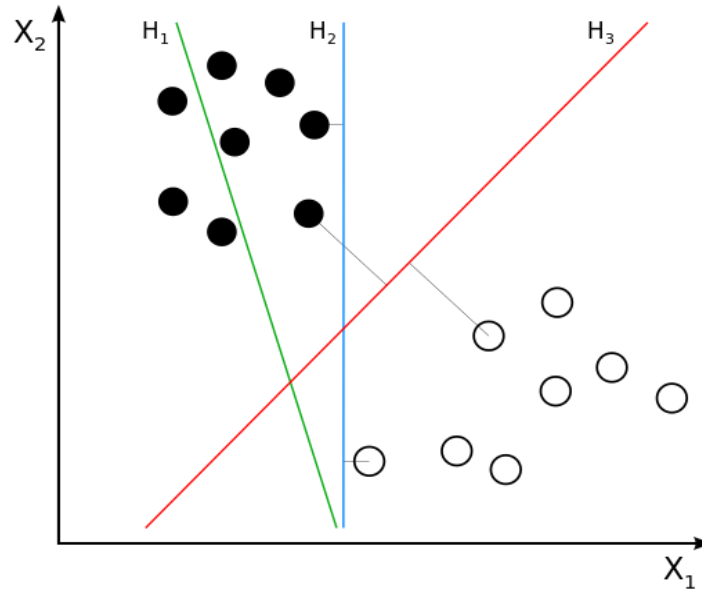
$$w(x, y_i) = \frac{\exp(-d(x, y_i))}{\sum_{i=1}^k \exp(-d(x, y_i))} \quad (2-22)$$

Prilikom učenja klasifikatora zbog sličnosti izmjerenih podataka svaka se postojeća klasa „natječe“ s ostalima kako bi pobijedila ili bila najbližija nepoznatom uzorku i tako pridonijela klasifikaciji. Iz tog razloga, k -NN algoritam smatra se algoritmom „natjecanja“. Osim toga, pripada skupini lijenih algoritama jer ne stvara odluku o klasifikaciji sve do novog upita, drugim riječima, radi sve u zadnji tren. Glavni nedostatak algoritma k -najbližih susjeda je osjetljivost na sve značajke, bez obzira na njihov broj (dimenziju prostora) i značaj za krajnju funkciju – tzv. „prokletstvo dimenzionalnosti“ [12]. Moguća rješenja ovoga problema su rastezanje ili stiskanje osi euklidskog prostora ili smanjenje dimenzionalnosti. Prethodne dvije metode – analiza glavnih komponenti i linearna diskriminativna analiza omogućile su izdvajanje najbitnijih obilježja, tj. smanjenje dimenzionalnosti.

2.3.4. Stroj s potpornim vektorima (engl. *Support Vector Machine*)

Stroj s potpornim vektorima je jedan od algoritama za nadgledano strojno učenje - svaki se novi uzorak klasificira u dva razreda značajki (binarna klasifikacija) određenih na temelju prethodno klasificiranih objekata [13]. Boser, Guyon i Vapnik predstavili su ga 1992. godine. Osim u linearnoj klasifikaciji, algoritam je učinkovit i kod nelinearne klasifikacije zahvaljujući tzv. jezgrenom triku (engl. *kernel trick*) koji omogućava klasifikaciju podataka koji nemaju određene dimenzije u vektorskom prostoru, primjerice niz DNK ili proteina [14]. Stroj s potpornim vektorima temelji se na odabiru hiperravnine ili niza hiperravnina u visoko dimenzionalnom ili beskonačno dimenzionalnom prostoru koja će biti najudaljenija od najbližih uzoraka obje klase (potpornih vektora), tj. koji ima

najveću marginu, kao što je prikazano na slici 2.7. Time se mogućnost pogreške prilikom klasifikacije smanjuje.



Sl. 2.7. Treća hiperravnina (H_3) predstavlja najbolji odabir jer najbolje razdvaja najbliže uzorke [13]

Hiperravnina je definirana linearnom funkcijom oblika:

$$f(x) = w^T x + b \quad (2-23)$$

gdje $w^T x$ predstavlja potporni vektor, definiran jednađbom (2-24), a b nagib.

$$w^T x = \sum_i w_i x_i \quad (2-24)$$

Prema izrazu (2-24) potporni vektor definiran je skalarnim umnoškom (umnožak iznosa prvoga i drugoga vektora te kosinusa kuta između njih). U slučajevima kad nije moguće linearno odvojiti uzorke u prostoru određenih dimenzija, primjenjuje se nelinearna klasifikacija. Uzorci se projiciraju u prostor većih dimenzija uz pomoć nelinearne funkcije $\phi(x)$. Iz jednađbe (2-25) vidljivo je da je računanje skalarnog ili unutarnjeg produkta vektora otežano, odnosno nemoguće kad je riječ o prostoru beskonačnih dimenzija.

$$f(x) = w^T \phi(x) + b \quad (2-25)$$

Kako bi se to izbjeglo, uvedene su tzv. jezgrene funkcije. Postoji nekoliko standardnih oblika istih:

- linearna:

$$K(x_i, x_j) = x_i^T \cdot x_j \quad (2-26)$$

- polinomna:

$$K(x_i, x_j) = (x_i^T \cdot x_j + a)^b \quad (2-27)$$

- Gaussova (engl. Radial Basis Function):

$$K(x_i, x_j) = \ell^{-\gamma \|x_i - x_j\|^2}, \gamma > 0 \quad (2-28)$$

- racionalna kvadratna:

$$K(x_i, x_j) = 1 - \frac{\|x_i - x_j\|^2}{\|x_i - x_j\|^2 + \tau} \quad (2-29)$$

- sigmoidalna:

$$K(x_i, x_j) = \tanh(\kappa x_i^T \cdot x_j + a) \quad (2-30)$$

Prilikom određivanja starosti osobe sa slike potrebno je više od dva razreda, koliko omogućuje stroj s potpornim vektorima, no uvođenjem dodatnih algoritama broj razreda se može povećati.

3. PROGRAMSKA IZVEDBA

Programska izvedba nije bila uspješna, no slijedi opis ideje mogućeg programskog rješenja u Pythonu 2.7.12. S obzirom da se programski kod temelji na strojnom učenju (engl. *machine learning*), potrebno je pripremiti bazu slika koje će služiti za treniranje algoritma. Na internetu je dostupno nekoliko baza podataka ljudskih lica [15], kao što je već spomenuto u uvodu ovoga rada. Nadalje, kako bi algoritam mogao biti primjenjiv na slikama po izboru korisnika/korisnice te uopće koristan, koristi se besplatna biblioteka programskih funkcija (engl. *Open Source Computer Vision, OpenCV*). Biblioteka je prvotno napisana u programskim jezicima C i C++, no aktivno se razvija i za druge programske jezike, poput Pythona, Rubyja, Matlaba te se može pokretati na različitim operacijskim sustavima [16]. *OpenCV* sadrži preko 500 funkcija koje omogućavaju različite mogućnosti, primjerice detekciju ljudskog lica na slici/video, obradu slike. U nastavku slijedi opis programa koji omogućava detekciju lica sa slike, prema predlošku [17].

```
import cv2, os  
  
import numpy as np  
  
from PIL import Image
```

Programski kod 3.1. Kod za uključivanje Pythonovih modula

Na početku su uključeni Pythonovi moduli koji omogućavaju različite funkcije, primjerice *cv2* uključuje *OpenCV* biblioteku, *os* omogućuje korištenje funkcija koje ovise o operacijskom sustavu – primjerice pristup raznim datotekama, direktorijima, *numpy* je jedan od osnovnih paketa s raznim korisnim matematičkim funkcijama, a *PIL* (engl. *Python Imaging Library*) predstavlja Pythonovu biblioteku za manipulaciju slikama. Često se koristi uz *OpenCV* biblioteku. Za detekciju lica moguće je koristiti *Haar Cascade* koji je dio *OpenCV*-a.

```
cascadePath = "haarcascade_frontalface_default.xml"  
  
faceCascade = cv2.CascadeClassifier(cascadePath)
```

Programski kod 3.2. Kod koji omogućava detekciju lica

Slijedi dio programskog koda (programski kod 3.3.), tj. funkcija koja bi trebala omogućiti učitavanje slika iz direktorija, spremanje u listu te označavanje, što bi se kasnije moglo upotrijebiti za učenje klasifikatora. Funkcija vraća nazive slika te pripadne oznake. U primjeru je korištena baza podataka *Yale Facedatabase A* [18], koja se sastoji od slika 15 pojedinaca, dimenzija 320 x 243, slikanih u

različitim pozama, s različitim izrazima lica, pri različitom osvjetljenju. Baza nije previše obimna te se može koristiti za početna treniranja.

```
def get_images_and_labels(path):  
    image_paths = [os.path.join(path, f) for f in os.listdir(path)  
if not f.endswith('.sad')]  
  
    images = []  
    labels = []  
  
    for image_path in image_paths:  
        image_pil = Image.open(image_path).convert('L')  
  
        image = np.array(image_pil, 'uint8')  
  
        nbr =  
int(os.path.split(image_path)[1].split(".")[0].replace("subject",  
""))  
  
        faces = faceCascade.detectMultiScale(image)  
  
        for (x, y, w, h) in faces:  
            images.append(image[y: y + h, x: x + w])  
  
            labels.append(nbr)  
  
            cv2.imshow("Adding faces to traning set...", image[y: y  
+ h, x: x + w])  
  
            cv2.waitKey(50)  
  
    return images, labels  
  
path = '/home/pallas/projects/Age estimation/yalefaces'  
images, labels = get_images_and_labels(path)
```

Programski kod 3.3. Kod koji omogućava učitavanje slika iz direktorija te spremanje u listu

Prilikom daljnje obrade slika, prijedlog je pretvoriti ih u matrice nad kojima bi se vršila analiza glavnih komponenti (engl. *Principal Component Analysis*), a rezultati analize, glavne komponente, u ovom slučaju svojstvena lica (engl. *Eigenfaces*), predstavljaju ulazne varijable u linearnoj diskriminantnoj analizi (engl. *Linear Discriminant Analysis*), a izlazne varijable se mogu nazvati

Fisherovim vrijednostima. Nakon obrade podataka, slijedi klasifikacija istih. Prethodno opisane algoritme (algoritam k -najbližih susjeda te stroj s potpornim vektorima) moguće je naći u Pythonovoj biblioteci za strojno učenje, *scikit-learn*, koja se sastoji od programa za obradu podataka, klasifikacijske i regresijske probleme, pretprocesiranje podataka te mnogih drugih.

```
from sklearn.decomposition import PCA
from sklearn.discriminant_analysis import LinearDiscriminantAnalysis
from sklearn.neighbors import NearestNeighbors
from sklearn import svm
```

Programski kod 3.4. Kod koji omogućava primjenu gotovih programskih rješenja iz *scikit-learn* biblioteke

Osim u *scikit-learn* biblioteci, analiza glavnih komponenti se nalazi i u *matplotlib* biblioteci koja omogućava korištenje funkcija Matlab-a u Pythonu.

```
from matplotlib.mlab import PCA
```

Programski kod 3.5. Kod koji omogućava primjenu gotovog programskog rješenja analize glavnih komponenti iz *matplotlib* biblioteke

Programska izvedba nije bila uspješna, no dosadašnja literatura dokazala je da ovakav pristup daje dobre rezultate. U daljnjem proučavanju ove tematike nastojat će se primijeniti spomenuto rješenje uz moguća poboljšanja.

4. ZAKLJUČAK

Određivanje starosti osobe nije lak zadatak za čovjeka, a zbog povećane interakcije ljudi i računala te razvoja znanosti i industrije, taj se proces nastoji automatizirati. To je na neki način potaklo znanstvenike na istraživanje karakteristika ljudskog lica koje mogu pridonijeti razlikovanju starosne dobi. Zbog potrebe za izdvajanjem bitnih značajki, primjenjuju se razne statističke metode za obradu podataka. Ovdje su spomenute analiza glavnih komponenti (engl. *PCA*) te linearna diskriminantna analiza (engl. *LDA*). Analiza glavnih komponenti omogućuje prikazivanje podataka iz visoko dimenzionalnog prostora u niže dimenzionalnom prostoru, a primjena linearne diskriminantne analize (Fisherove linearne diskriminante) omogućuje razdvajanje značajki različitih razreda, što olakšava klasifikaciju. Opisani su algoritam *k*-najbližih susjeda (engl. *k-Nearest Neighbours*) i stroj s potpornim vektorima (engl. *Support Vector Machine*) koji se mogu primjenjivati u postupku klasifikacije. Algoritam *k*-najbližih susjeda klasificira nepoznati objekt s obzirom na već klasificirane objekte koji se nalaze blizu njega, čiji broj ovisi o vrijednosti *k*. Jedan je od lijenih algoritama – do same klasifikacije objekta dolazi po potrebi, a njegov glavni nedostatak – „prokletstvo dimenzionalnosti“ – otklonjen je primjenom spomenutih metoda za obradu podataka. Stroj s potpornim vektorima (engl. *Support Vector Machine*) jedan je od algoritama za nadgledano strojno učenje, a svaki objekt se klasificira u jedan od dva razreda, no moguće ga je prilagoditi za više klasa. Ideja mogućeg programskog rješenja je ukratko objašnjenja, no ostavlja se prostor za daljnje razrađivanje.

LITERATURA

- [1] Yun Fu, Guodong Guo, Thomas S. Huang: Age Synthesis and Estimation via Faces: A Survey, IEEE Transactions on Pattern Analysis And Machine Intelligence, No. 11, Vol. 32, November 2010.
- [2] Procjena starosne dobi osobe s obzirom na karakteristike lica (engl. *Facial Age Estimation*) http://www.scholarpedia.org/article/Facial_Age_Estimation, 17.06.2016.
- [3] J. Shlens: A tutorial on principal component analysis: Derivation, Discussion and Singular Value Decomposition, 2003., prva verzija
- [4] I. Batal, E. Strobl, M. Hauskrecht: CS 3750 Machine Learning, Lecture 10 – Principal Component Analysis, Singular Value Decomposition
- [5] Hlaing Htake Khaung Tin: Subjective Age Prediction of Face Images Using PCA, International Journal of Information and Electronics Engineering, No.3, Vol.2, May 2012.
- [6] P. Koruga, J. Ševa, M. Bača: A Review of Face Recognition Algorithms and Their Application in Age Estimation, Proceedings of the 22nd Central European Conference on Information and Intelligent Systems, Varaždin, 2011.
- [7] Linearna diskriminantna analiza (engl. *Linear Discriminant Analysis*) https://en.wikipedia.org/wiki/Linear_discriminant_analysis, 13.07.2016.
- [8] O. Veksler: CS434a/541a - Pattern Recognition, Lecture 8 http://www.csd.uwo.ca/~olga/Courses//CS434a_541a//Lecture8.pdf , 13.07.2016.
- [9] Algoritam k najbližih susjeda (engl. *k-Nearest Neighbors Algorithm*) <http://www.statsoft.com/textbook/k-nearest-neighbors>, 30.08.2016.
- [10] Minkowski udaljenost (engl. *Minkowski distance*) https://en.wikipedia.org/wiki/Minkowski_distance, 30.08.2016.
- [11] Mahalanobisova udaljenost (engl. *Mahalanobis distance*) https://en.wikipedia.org/wiki/Mahalanobis_distance, 30.08.2016.
- [12] Algoritam k najbližih susjeda (engl. *k-nearest neighbors algorithm*) https://en.wikipedia.org/wiki/K-nearest_neighbors_algorithm, 30.08.2016.

- [13] Stroj s potpornim vektorima (engl. *Support Vector Machine*)
https://en.wikipedia.org/wiki/Support_vector_machine, 08.09.2016.
- [14] A. Ben-Hur, J. Weston: A User's Guide to Support Vector Machines
- [15] Prepoznavanje lica, baze podataka (engl. *Face Recognition Homepage, Databases*)
<http://face-rec.org/databases/> , 20.09.2016.
- [16] G. Bradsky, A. Kaehler: Learning OpenCV – Computer Vision with the OpenCV Library, O'Reilly Media, Inc., Sjedinjene Američke Države, 2008.
- [17] Prepoznavanje lica koristeći Python i OpenCV (engl. *Face Recognition using Python and OpenCV*) <http://hanzratech.in/2015/02/03/face-recognition-using-opencv.html>, 20.09.2016.
- [18] Yale-ova baza slika (engl. *Yale Facedatabase A*)
<http://vision.ucsd.edu/content/yale-face-database> , 20.09.2016.

SAŽETAK

Karakteristike ljudskog lica oduvijek su predmet proučavanja društvenih znanosti. Razvojem tehnologije i industrije te zbog sve veće interakcije čovjeka i računala nastoje se izlučiti značajke ljudskog lica koje bi računalu omogućile razlikovanje identiteta, spola, raspoloženja i starosti pojedinaca. U ovom radu naglasak je upravo na razlikovanju starosne dobi pojedinaca. Primjenom određenih matematičkih postupaka (analiza glavnih komponenti, linearna diskriminantna analiza) smanjena je dimenzionalnost podataka, tj. izlučene su potencijalno bitne značajke za određivanje starosne dobi pojedinca. Algoritmi strojnog učenja primjenjuju se u postupku klasifikacije. Opisana su dva često korištena - algoritam k najbližih susjeda te stroj s potpornim vektorima. S obzirom da se računalni vid ubrzano razvija, poboljšanje već postojećih rješenja može se očekivati, kao i nova rješenja.

Ključne riječi: određivanje starosne dobi pojedinca, računalni vid, smanjenje dimenzionalnosti, analiza glavnih komponenti, linearna diskriminantna analiza, strojno učenje, algoritam k najbližih susjeda, stroj s potpornim vektorima

AGE ESTIMATION FROM FACE IMAGE

ABSTRACT

Characteristics of human face have always been studied in social sciences. Development of technology and industry, as well as increased interaction between human and machine, have increased the need for computer based estimation of identity, gender, mood and age of person. In this paper emphasis has been put on age estimation. Application of mathematical methods, such as principal component analysis (PCA) and linear discriminant analysis (LDA), has enabled dimensionality reduction, i.e. extraction of potentially significant features for determining age of individual. Machine learning algorithms are used for classification. Here have been described two frequently used algorithms – k -nearest neighbors (k -NN) and support vector machine (SVM). Due to rapidly developing of computer vision, improved solutions, as well as new ones, can be expected.

Keywords: age estimation from face image, computer vision, dimensionality reduction, principal component analysis (PCA), linear discriminant analysis (LDA), machine learning, k -nearest neighbors (k -NN), support vector machine (SVM)

ŽIVOTOPIS

Tamara Damjanović rođena je 1994. godine. Završila je Gimnaziju u Novoj Gradiški 2012. godine i trenutno je studentica treće godine sveučilišnog preddiplomskog studija računarstva na Fakultetu elektrotehnike, računarstva i informacijskih tehnologija u Osijeku. Govori engleskim jezikom i služi se talijanskim jezikom. Dobro vlada alatima Microsoft Office-a te programskim jezicima C, C++ i Python.

Potpis:
