Seminarski rad iz Linearne Algebre
Prepoznavanje lica

Nemanja Saveski 2023/3163

6. februar 2024.

U ovom seminarskom radu ćemo se baviti prepoznavanjem lica putem algoritma za izdvajanje sopstvenih lica (*eng. eigenfaces*). Ovo je veoma poznat problem u kompjuterskoj viziji, a algoritmi koji se koriste su raznovrsni u odnosu na to šta želimo da postignemo. Neki algoritmi su namenjeni za korišćenje u realnom vremenu, pa samim tim oni zahtevaju brža izračunavanja, ali samim tim postoji i veća šansa da se dogodi greška pri klasifikaciji, dok neki drugi prave manju grešku, ali zahtevaju veće računarske resurse jer koriste kompleksnije proračune. Svakako, uočavanje lica, bilo u realnom vremenu ili van realnog vremena, deli se na deo koji se bavi detekcijom lica i deo koji se bavi preproznavanjem lica. Dakle, detekcija lica služi da vidimo da li uopste postoji osoba na slici, a prepoznavanje se bavi određivanjem identiteta osobe koja je na slici. Uglavnom je prvo potrebno uraditi preprocesiranje slika koje koristimo u našem algoritmu, ali zbog jednostavnosti, kao i zbog toga što je zadatak da pravimo algoritam samo za prepoznavanje lica, koristićemo bazu sa preprocesiranim slikama. Baza se zove *Labeled Faces in the Wild* (*LFW*) [1] i sastoji se od oko 13k slika poznatih ličnosti veličine **64x64**. Uzorak slika iz baze je predstavljen na slici 1.

Slika 1: LFW Uzorak

Algoritam ćemo bazirati na radu Turka i Pentlanda [2], koji koristi PCA (*eng. Principal Component Analysis*) i singularnu dekomopoziciju – SVD (*eng. Singular Value Decomposition*). Osnovna ideja je predstaviti podatke na kompaktan način pomoću sopstvenih vektora u ovom slučaju sopstvenih lica.

Singularna dekompozicija kaže da za svaku kompleksnu matricu A_{mxn} može da se izvrši dekompozicija:

$$A_{mxn} = U_{mxm}\Sigma_{mxn}V^T_{nxn}, \tag{1}$$

gde su matrice U i V unitarne matrice sačinjene od singularnih vektora matrice A. U je sačinjena od levih singularnih vektora, dok je matrica V sačinjena od desnih singularnih vektora matrice A. Takođe, važi da je:

$$U^{-1} = U^T, \quad V^{-1} = V^T \tag{2}$$

Matrica Σ je matrica singularnih vrednosti A i ona je dijagonalna. Singularna dekompozicija može da se napiše i na sledeći način:

$$Av_j = \sigma_j u_j \tag{3}$$

Da bismo odredili matrice V, Σ i U posmatraćemo proizvode AA^T i A^TA:

$$A^TA = (U\Sigma V^T)^T(U\Sigma V^T) = V\Sigma^T U^T U\Sigma V^T = V\Sigma^2 V^T \tag{4}$$

$$AA^T = (U\Sigma V^T)(U\Sigma V^T)^T = U\Sigma^2 U^T \tag{5}$$

Pošto su obe matrice AA^T i A^TA kvadratne (mxm i nxn), možemo se podsetiti da dekompozicija matrice neke B (ako je dijagonizabilna) na sopstvene vrednosti i vektore se vrši na sledeći način:

$$B = Q\Lambda Q^{-1}, \tag{6}$$

gde je Q matrica sopstvenih vektora, dok je Λ dijagonalna matrica odgovarajućih sopstvenih vrednosti. Upravo ovaj oblik odgovara obliku iz jednačina (4) i (5). Dakle, vektori u_j su sopstveni vektori matrice AA^T, vektori v_j su sopstveni vektori matrice A^TA, a σ_j^2 su sopstvene vrednosti matrice A^TA.

Sada ćemo malo više pričati o PCA. Glavna ideja iza PCA je u redukciji dimenzija, tj. želimo da predstavimo podatke na kompaktan način bez velikog gubitka informacije. Hoćemo da pronađemo ortogonalnu transformaciju M koja će originalne podatke X da prevede u prostor Y u kome su obeležja nekorelisana (kovarijaciona matrica matrice Y je dijagonalna).

$$Y = XM \tag{7}$$

Dakle, primenom PCA nad podacima pozicioniramo se u koordinatni sistem čiji su osnovni vektori pravca vektori koji nose najveću varijabilnost podataka. Zbog toga, za M se uzima matrica sopstvenih vektora kovarijanse obeležja ili matrica singularnih vektora originalnih merenja X. Možemo da pišemo:

$$M = [\Phi_1, ..., \Phi_n], \tag{8}$$

gde su vektori Φ_i sopstveni vektori koji odgovaraju i-toj po veličini sopstvenoj vrednosti kovarijacione matrice X, Σ_X. Izdvajanjem prvih **p** Φ_i vektora (gde je $p < n$) smanjujemo dimenziju naših podataka sa n na p:

$$M_{red} = [\Phi_1, ..., \Phi_p], \;\rightarrow\; Y_{red} = XM_{red} \tag{9}$$

Kako da znamo na koliko dimenzija treba redukovati prostor obeležja? Posmatraju se sortirane sopstvene vrednosti matrice Σ_X ili X^TX:

$$\Sigma_X M = M\Lambda, \tag{10}$$

dakle Λ je dijagonalna matrica sa sortiranim sopstvenim vrednostima i bira se onaj broj dimenzija koji obezbeđuje da uticaj sortiranih komponenti:

$$\frac{\sum_{i=1}^{p} \lambda_i}{\sum_{i=1}^{n} \lambda_i} \tag{11}$$

bude veći od **0.9** ili **0.95**. Veličina se naziva i kumulativna objašnjena varijansa, jer pokazuje koliko naš algoritam dobro "razume"varijabilnost podataka kao i koliko dobro može da rekonstruiše ulaz ako mu nije data celokupna informacija. Da je ona kojim slučajem manja, recimo oko 0.7, algoritam ne bi uspeo lepo da rekonstruiše, tj. izgubilo bi se dosta informacije. Naravno, ne treba da bude ni mnogo blizu 1, jer onda nismo obavili osnovni zadatak, a to je redukcija dimenzija. Ako želimo da uradimo PCA uz pomoć SVD-a, za transformacionu matricu M ćemo usvojiti desne singularne vektore:

$$Y = XV \;\; (X = U\Sigma V^T), \tag{12}$$

jer dobijamo za kovarijacionu matricu novog prostora:

$$\begin{aligned} \Sigma_Y &= \tfrac{1}{m-1} Y^T Y = \tfrac{1}{m-1} V^T X^T X V = \tfrac{1}{m-1} V^T V \Sigma_X^T U^T U \Sigma_X V^T V \\ \Sigma_Y &= \tfrac{1}{m-1} \Sigma_X^2, \end{aligned} \tag{13}$$

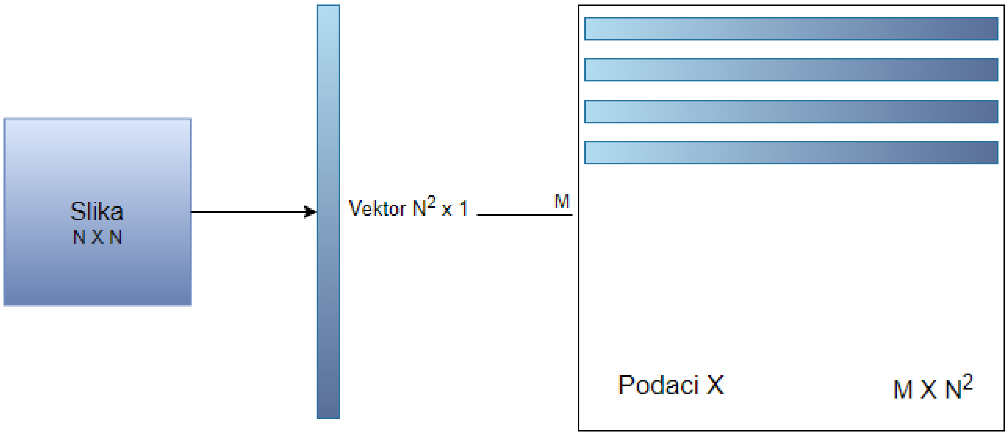
što dalje implicira da su singularne vrednosti σ_i (jednačina 3) povezane sa sopstvenim vrednostima λ_i:

$$\lambda_i = \frac{1}{m-1}\sigma_i^2 \tag{14}$$

Naravno bitno je napomenuti da ovo sve važi ako je matrica X centrirana u smislu da je oduzeta srednja vrednost po uzorku od svakog uzorka.

Sada ćemo preći na adaptaciju ovih metoda u zadatku prepoznavanja lica.

Na slici 2 je dat način prikaza podataka u računaru. Dakle, svaku sliku NxN stanjimo u jedan vektor N^2x1, a zatim sve te vektore pakujemo u matricu jedan do drugog, gde dobijamo matricu X veličine MxN^2 (jer imamo M slika u bazi).

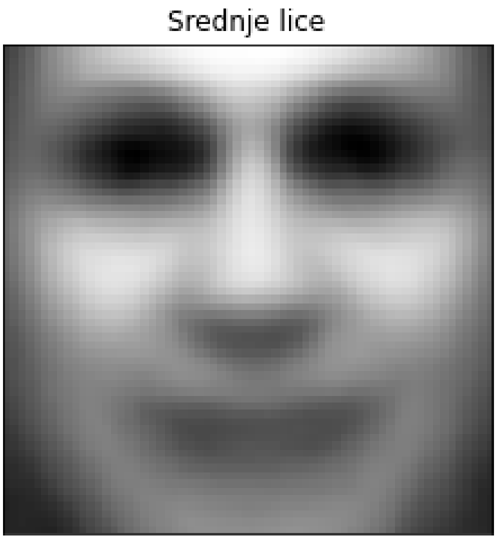


Slika 2: Podaci

Posto se kaže da su obeležja u izlaznoj matrici nekorelisana, to može da nas intuitivno povuče da probamo nekako da dijagonalizujemo matricu X. Prvo ćemo naći srednju vrednost po dimenziji M cele matrice X i oduzeti taj vektor od svake slike u X, dakle centriramo naše podatke. Za srednje lice:

$$\bar{X} = \frac{1}{M} \sum_{i=1}^{M} x_i \tag{15}$$

dobijamo rezultat na slici 3



Slika 3: $\bar{X}$

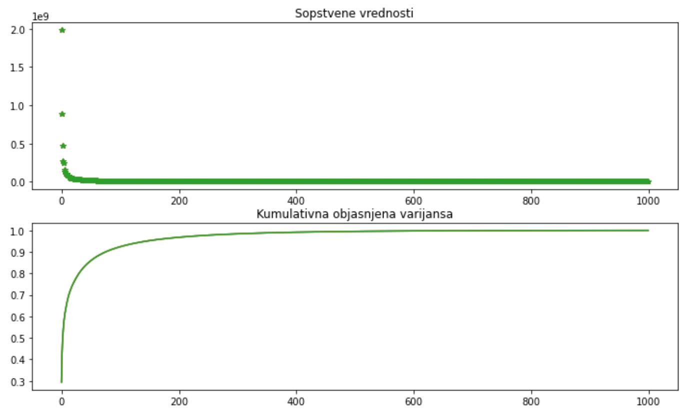
dok centrirana lica računamo:

$$\tilde{x}_i = x_i - \bar{X} \tag{16}$$

Primeri centriranih lica su na slici 4.

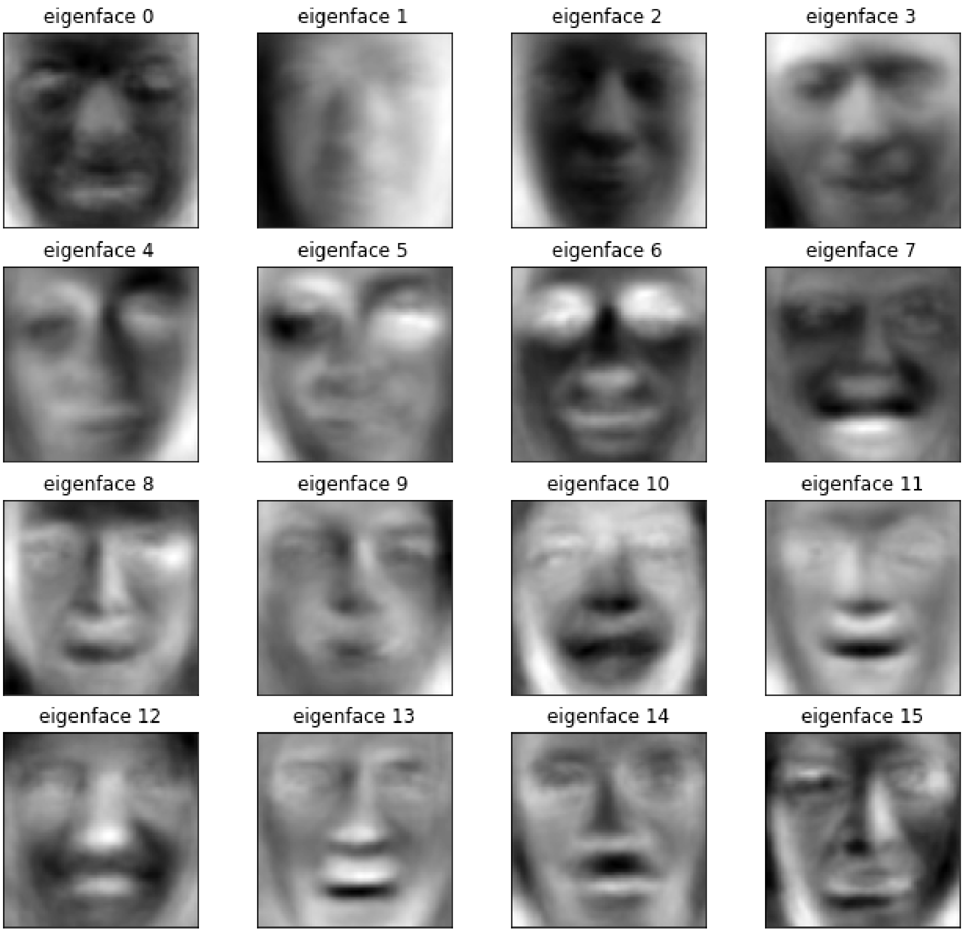
Slika 4: Centrirana lica

Kada uradimo SVD dobijamo sledeće sopstvene vrednosti i kumulativnu objašnjenu varijansu:



Slika 5: Sopstvene vrednosti i kumulativna objašnjena varijansa

Sopstveni vektori koji odgovaraju ovim sopstvenim vrednostima su u stvari sopstvena lica prikazana na slici 6.



Slika 6: Sopstvena lica

Za rekonstrukciju ćemo korisiti sva izdvojena sopstvena lica (u zavisnosti od toga koliko komponenti čuvamo, tj. kolika je dimenzija novodobijenog prostora):

Slika 7: Metod rekonstrukcije

Dalje možemo da pišemo ($\hat{x}_i$ je rekonstruisano lice):

$$\hat{x}_i = \bar{X} + \sum_{j=1}^{p} w_j v_j \tag{17}$$

$$\tilde{x}_i = \sum_{j=1}^{p} w_j v_j \rightarrow w_j = v_j^T \tilde{x}_i \tag{18}$$

Konačno dobijamo vektor težina po svakom sopstvenom licu za svaku sliku x_i:

$$W = V^T \tilde{X} \tag{19}$$

Na slikama 8, 9, 10 je prikazana rekonstrukcija početnih slika u slučaju **20**, **100** i **200** komponenti (sopstvenih lica).

Slika 8: 20 sopstvenih lica

Kao što možemo primetiti za 20 sopstvenih lica algoritam je veoma loše rekonstruisao lica, ali to je i očekivano zbog kumulativne objašnjene varijanse na slici 5. S druge strane za 100 sopstvenih lica, algoritam je malo poboljšan, ali za 200 je velika količina informacije očuvana, što je i očekivano ako pogledamo kumulativnu objašnjenu varijansu.

Slika 9: 100 sopstvenih lica

Slika 10: 200 sopstvenih lica

Konačno, dolazimo do dela gde ćemo zadati sliku algoritmu, a on treba da pogodi koje lice se nalazi na slici. U ovom delu ćemo projektovati naše novo ulazno lice u prostor sopstvenih lica i na osnovu toga izračunati težine tog ulaznog lica. Dalje, da bismo prepoznali koja osoba je na slici koristićemo sledeći kriterijum:

$$e_r = \min_{l} ||w_{novo} - w_l||_2, \ w_l \in W \tag{20}$$

Dakle, pretpostavljamo da je ulazno lice ono lice čije se težine (koordinate) u prostoru dimenzija najmanje razlikuju od težina ulazog lica (euklidska distanca).

Algoritam uspeva da prepozna osobu čak i ako je ona u drugom položaju:

Slika 11: Ista osoba druga slika

Naravno uvek postoji problem sa bazom podataka (slika), takav da ako želimo da ubacimo sliku osobe koju naš algoritam nije nikada "video"ranije, ne može pogoditi osobu koja se nalazi na slici. Ali, sa slike 12, možemo reći da iako postoje osobe koje algoritam nije nikada procesirao, veoma dobro uspeva da uhvati neke karakteristike lica (položaj glave, oči, usta, nos):

Slika 12: Nepoznata osoba

Literatura

[1] *LWF dataset*. URL: `https://conradsanderson.id.au/lfwcrop/`.

[2] Matthew Turk i Alex Pentland. "Eigenfaces for recognition". U: *Journal of cognitive neuroscience* 3.1 (1991.), str. 71–86.