
KLASIFIKACIJA SLIKA BAZIRANA NA TENZORIMA

MATEMATIČKO MODELIRANJE PRETRAŽIVAČA

Ana Parać

Prirodoslovno-matematički fakultet
Matematički odsjek
Sveučilište u Zagrebu
Zagreb, Hrvatska
anaparac8@gmail.com

Sara Pužar

Prirodoslovno-matematički fakultet
Matematički odsjek
Sveučilište u Zagrebu
Pula, Hrvatska
sara.puzar@gmail.com

Petra Rožić

Prirodoslovno-matematički fakultet
Matematički odsjek
Sveučilište u Zagrebu
Zagreb, Hrvatska
petra.rozic01@gmail.com

Paula Vujasić

Prirodoslovno-matematički fakultet
Matematički odsjek
Sveučilište u Zagrebu
Zagreb, Hrvatska
paula.vujasic@gmail.com

SAŽETAK

U ovom radu predstavljamo algoritam za klasifikaciju slika. Metoda se bazira na smanjenju dimenzionalnosti tenzora koji predstavlja skup slika. Smanjenje dimenzionalnosti postignuto je HOOI metodom. Za testiranje algoritma korištena je baza AT&T te smo kreirale bazu vlastitih fotografija.

Keywords HOOI · Klasifikacija slika · HOSVD

1 Opis problema

1.1 Uvod

Automatska klasifikacija slika jedno je od najvažnijih područja analize slika. Za unaprijed definirane klase i danu nepoznatu sliku cilj je odrediti kojoj klasi pripada. Radi se o problemu nadziranog učenja budući da su slike iz skupa za treniranje unaprijed klasificirane.

Problem kojim se bavimo u ovom seminaru baziran je na prepoznavanju lica, odnosno za danu ekspresiju lica trebamo odrediti o kojoj se osobi radi. U našem problemu svaka je osoba predstavljena jednom klasom. Izazov je što među različitim klasama ima puno zajedničkih karakteristika (radi se o ljudskom licu), a unutar jedne klase ima puno varijacija jer ista osoba radi različite ekspresije lica. Za rješavanje ovoga problema razvijeno je više tehnika, a mi smo ga riješile pomoću tehnike koju ćemo opisati u nastavku, Higher-order Orthogonal Iteration (HOOI) metode.

Algoritam primjenjujemo na bazi podataka od AT&T Laboratories Cambridge. Baza sadrži 40 osoba, gdje svaka osoba ima 10 različitih ekspresija lica (iz profila, s osmijehom, sa/bez naočala...). Ukupno, radi se o 400 slika dimenzija 112×92 . Te slike slažemo u 40 tenzora dimenzija $112 \times 92 \times 10$. Nakon kratkog uvoda o tenzorima predstavljamo HOOI metodu.

1.2 Matematičko modeliranje problema

U uvodu matematičkog modeliranja problema donosimo kratki pregled tenzora i HOSVD-a.

Definicija 1 (Tensor). Tenzor reda N je produkt od N vektorskih prostora

$$A \in \mathbf{R}^{I_1 \times I_2 \times \dots \times I_N},$$

element tenzora A označavamo sa $A_{i_1 i_2 \dots i_N}$, pri čemu je $0 < i_n \leq I_n, n = 1, 2, \dots, N$

Za tenzore trećeg reda vrijede sljedeće definicije:

Definicija 2 (Niti u modu n). Niti u modu n tenzora $A \in \mathbf{R}^{I_1 \times I_2 \times I_3}$ dio su tenzora dobiven fiksiranjem svih indeksa osim jednog, n -tog indeksa.

Predstavljene su vektorom

$$V_n \in \mathbf{R}^{I_n}.$$

Proces razvijanja niti iz tenzora naziva se vektorizacija tenzora.

Definicija 3 (Odsječci u modu n). Odsječci u modu n tenzora $A \in \mathbf{R}^{I_1 \times I_2 \times I_3}$ dio su tenzora dobiven fiksiranjem n -tog indeksa. Ovisno o modu, odsječke nazivamo redom horizontalni, lateralni ili frontalni odsječak.

Predstavljene su matricama

$$A_{(n)} \in \mathbf{R}^{I_j \times I_k},$$

gdje $j, k \neq n$. Proces prikazivanja tenzora pomoću matrica naziva se matricizacija tenzora.

Definicija 4 (Tuckerov produkt u modu n). Neka je tenzor $\tau \in \mathbf{R}^{I \times J \times K}$ i matrice $U_1 \in \mathbf{R}^{\hat{I} \times I}$, $U_2 \in \mathbf{R}^{\hat{J} \times J}$ i $U_3 \in \mathbf{R}^{\hat{K} \times K}$. Tada je Tuckerov produkt u n -modu definiran slijedećim izrazom:

$$1\text{-mod} \quad (\tau \bullet_1 U_1)_{\hat{i}, j, k} = \sum_{i=1}^I \tau_{i, j, k} A_{\hat{i}i}$$

$$2\text{-mod} \quad (\tau \bullet_2 U_2)_{\hat{j}, i, k} = \sum_{j=1}^J \tau_{i, j, k} B_{\hat{j}j}$$

$$3\text{-mod} \quad (\tau \bullet_3 U_3)_{\hat{k}, i, j} = \sum_{k=1}^K \tau_{i, j, k} C_{\hat{k}k}$$

Dekompozicija tenzora na singularne vrijednosti također se naziva multilinearna dekompozicija na singularne vrijednosti ili HOSVD.

Teorem 5 (HOSVD). Tenzor trećeg reda $\tau \in \mathbf{R}^{I \times J \times K}$ može se reprezentirati kao produkt:

$$\tau = \mathcal{B} \bullet_1 U_1 \bullet_2 U_2 \bullet_3 U_3$$

gdje su $U_1 \in \mathbf{R}^{I \times I}$, $U_2 \in \mathbf{R}^{J \times J}$ i $U_3 \in \mathbf{R}^{K \times K}$ ortogonalne matrice. Za podtenzore jezgrenog tenzora $\mathcal{B} \in \mathbf{R}^{I \times J \times K}$, $\mathcal{B}_{i=\alpha}^1 \in \mathbf{R}^{J \times K}$, $\mathcal{B}_{j=\alpha}^2 \in \mathbf{R}^{I \times K}$ i $\mathcal{B}_{k=\alpha}^3 \in \mathbf{R}^{I \times J}$ vrijede slijedeća svojstva:

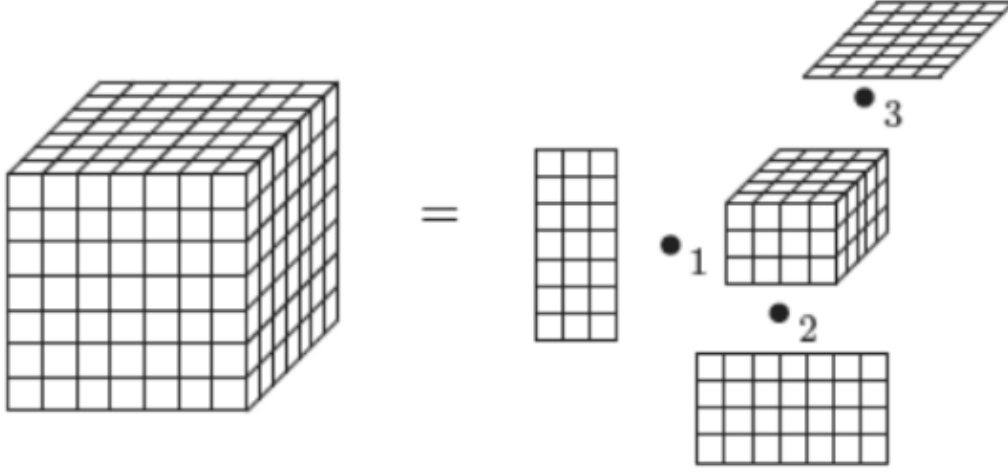
- potpuna ortogonalnost

$$\langle \mathcal{B}_{i_n=\alpha}^n, \mathcal{B}_{i_n=\beta}^n \rangle = (\sigma_\alpha^{(1)})^2 \delta_{\alpha, \beta}, \quad \alpha, \beta = 1, \dots, I_n$$

- padajući poredak singularnih vrijednosti

$$\|\mathcal{B}_{i=1}^n\|_F \geq \|\mathcal{B}_{i=2}^n\|_F \geq \dots \geq \|\mathcal{B}_{i=I_n}^n\|_F \geq 0$$

gdje $\|\mathcal{B}_{i=\alpha}^n\|_F = \sigma_\alpha^{(n)}$ za $\alpha = 1, \dots, I_n$.



Tenzor reda 3 i odsječci u modu 1, 2 i 3

Jezgreni tenzor \mathcal{B} možemo dobiti kao :

$$\mathcal{B} = \tau \times_1 U_1^T \times_2 U_2^T \times_3 U_3^T.$$

2 Metoda rješavanja

Mnoga rješenja problema analize slike oslanjaju se na reduciranje visokih dimenzija tenzora. Redukcija dimenzija je nužna zbog velike količine podataka koji imaju vrlo malo značajnih doprinosa. Trenutno je razvijeno nekoliko metoda za aproksimaciju niskog ranga u više dimenzija. Neke od njih su: *truncated* HOSVD, HOOI, alternirajući SDP (Semidefinite Programming). U ovom radu koristimo algoritam baziran na HOOI metodi za najbolju- (R_1, R_2, R_3) aproksimaciju tenzora niskoga ranga.

2.1 Higher-Order Orthogonal Iteration

Higher-Order Orthogonal Iteration (HOOI) se zasniva na metodi alternirajućih najmanjih kvadrata (ALS) i služi za optimizaciju sljedećeg problema:

$$\min_{\mathcal{B}, U_1, U_2, U_3} \|\mathcal{T} - \mathcal{B} \bullet_1 U_1 \bullet_2 U_2 \bullet_3 U_3\|_F = \max_{U_1, U_2, U_3} \|\mathcal{B} \bullet_1 U_1^T \bullet_2 U_2^T \bullet_3 U_3^T\|_F, \quad (1)$$

pri čemu su $U_1 \in \mathbf{R}^{I \times k_1}$, $U_2 \in \mathbf{R}^{J \times k_2}$ i $U_3 \in \mathbf{R}^{K \times k_3}$ matrice niskog ranga, a $\mathcal{B} \in \mathbf{R}^{k_1 \times k_2 \times k_3}$ reducirani jezgreni tenzor.

Prikažimo HOOI algoritam za tenzor trećeg reda, budući da u modeliranju problema koristimo tenzor trećeg reda.

Ulaz: $\mathcal{T} \in \mathbf{R}^{I_1 \times I_2 \times I_3}$

Izlaz: $\hat{\mathcal{T}} \in \mathbf{R}^{I_1 \times I_2 \times I_3}$, procjenitelj najbolje- (R_1, R_2, R_3) aproksimacije od \mathcal{T}

1. Inicijalizacija početnih vrijednosti U_1, U_2, U_3

2. Iteriranje do konvergencije:

- $\hat{\mathcal{U}}_{k+1}^{(1)} = \mathcal{A} \times_2 U_k^{(2)T} \times_3 U_k^{(3)T}$
maksimiziramo $\|\hat{\mathcal{U}}_{k+1}^{(1)} \times_1 U^{(1)T}\|$ po $U^{(1)} \in \mathbf{R}^{I_1 \times \mathbf{R}_1}$, pri čemu vrijedi $U^{(1)T} U^{(1)} = I$
 $U_{k+1}^{(1)} = U_{max}^{(1)}$

- $\hat{\mathcal{U}}_{k+1}^{(2)} = \mathcal{A} \times_1 U_k^{(1)T} \times_3 U_k^{(3)T}$
maksimiziramo $\|\mathcal{U}_{k+1}^{(2)} \times_2 U^{(2)T}\|$ po $U^{(2)} \in \mathbf{R}^{I_2 \times \mathbf{R}_2}$, pri čemu vrijedi $U^{(2)T} U^{(2)} = I$
 $U_{k+1}^{(2)} = U_{max}^{(2)}$
 - $\hat{\mathcal{U}}_{k+1}^{(3)} = \mathcal{A} \times_1 U_k^{(1)T} \times_2 U_k^{(2)T}$
maksimiziramo $\|\mathcal{U}_{k+1}^{(3)} \times_3 U^{(3)T}\|$ po $U^{(3)} \in \mathbf{R}^{I_3 \times \mathbf{R}_3}$, pri čemu vrijedi $U^{(3)T} U^{(3)} = I$
 $U_{k+1}^{(3)} = U_{max}^{(3)}$
3. $\mathcal{B} = \mathcal{A} \times \{U^T\}$, pri čemu $\{U\}$ označava matrice dobivene u zadnjoj iteraciji
 4. $\hat{\mathcal{T}} = \mathcal{B} \times \{U\}$

3 Opis implementacije algoritma

Provođenje algoritma odvojeno je u dvije ključne faze, *offline* fazu i *online* fazu.

U *offline* fazi provodi se učitavanje podataka te potom i *učenje* na tim podacima. Pri učitavanju, slike su podijeljene u tenzore na temelju neke zajedničke karakteristike, u našem slučaju, sve slike jedne osobe čine jedan tenzor. Potom koristeći HOOI metodu određujemo aproksimacijske tenzore B^μ za svaku osobu μ .

U *online* fazi testiramo algoritam na nepoznatim slikama.

3.1 Pseudokod

Učitavamo niz slika koje spremamo u tenzore s obzirom na pripadnost nekoj osobi. Zatim na svaki tenzor primjenjujemo HOOI algoritam te dobivamo aproksimacije nižeg ranga početnih tenzora B^μ . Svaki tenzor B^μ se sastoji od k matrica, gdje je k broj ekspresija.

Algorithm 1: Offline

Data: Niz slika

Result: Matrice B_k^μ

```

1 učitavanje slika u tenzore  $\tau$ ;
2 for  $i \leftarrow 0$  to broj tenzora do
3   |  $B_k \leftarrow \text{hooi}(\text{tenzor } \tau_i)$ ;
4 end
```

Pseudokod funkcije HOOI (*High-Order-Orthogonal-Iteration*).

Algorithm 2: HOOI

Data: Tenzor τ veličine $I_1 \times I_2 \times I_3$ i ciljani rang izlaznih matrica

Result: Tenzor \mathbf{B} veličine $I_1 \times I_2 \times I_3$ zadanog ranga

```

1 inicijaliziraj  $U_1, U_2, U_3$  pomoću svd-a na prvim  $R_n$  stupaca matrice  $\mathbf{U}$ ;
2 for  $i \leftarrow 0$  to konvergencija do
3   | for  $n \leftarrow 1$  to 3 do
4     |  $\tilde{\mathbf{U}} = \tau \times_{-n} \{U_k^T\}$ ;
5     |  $\mathbf{W} = \text{svds}(\tilde{\mathbf{U}}, \text{dimenzija } n)$ ;
6     |  $U_t = \mathbf{W}'$ 
7   | end
8   |  $\mathbf{U} = \mathbf{W}$ ;
9   |  $\lambda = \tau \times \{U^T\}$ ;
10  |  $\mathbf{B} = \lambda \times \{U\}$ ;
11 end
```

Nepoznatu sliku učitavamo te potom računamo udaljenosti u odnosu na svaku matricu B_μ^k , aproksimacijskih tenzora B_μ . Sliku klasificiramo na temelju minimalne udaljenosti od aproksimacije. Udaljenost računamo kao Frobeniusovu normu razlike matrice nepoznate slike i aproksimacijske matrice.

Algorithm 3: Online

Data: Nepoznata slika

Result: Broj μ koji označava klasu osobe

```

1 učitavanje slike u matricu  $X$ ;
2 for  $i \leftarrow 1$  to broj tenzora do
3   | udaljenosti  $\leftarrow \text{norm}(X - \text{matrica } B_i^k)$ ;
4 end
5  $\mu \leftarrow$  indeks na kojem se nalazi  $\min(\mu)$ ;

```

4 Rezultati testiranja

Skup podataka preuzet je sa stranice <https://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>.

Podaci na kojima smo trenirale i testirale algoritam strukturirani su u 40 mapa koje sadrže po 10 slika osoba koje prikazuju različite ekspresije ljudi.

4.1 Testiranje na preuzetim podacima

Od cjelokupnog skupa podataka izdvojile smo po tri slike za svaku osobu za testiranje algoritma, dok smo za treniranje koristile ostalih sedam slika po osobi. Veličina slika je 112×92 piksela te smo ih zbog toga spremile u tenzore dimenzije $112 \times 92 \times 7$.

Algoritam je testiran u više verzija koje se razlikuju u rangovima R_1, R_2, R_3 tenzora B^μ dobivenih u *offline* fazi. U sljedećoj tablici prikazani su neki od dobivenih rezultata. Točnost je izračunata kao omjer točno klasificiranih slika u odnosu na ukupan broj slika.

Točnost algoritma na preuzetim podacima	
Rangovi R_1, R_2, R_3	Točnost
$R_1 = 17, R_2 = 17, R_3 = 1$	81.67
$R_1 = 17, R_2 = 17, R_3 = 2$	92.50
$R_1 = 16, R_2 = 10, R_3 = 4$	93.33
$R_1 = 5, R_2 = 5, R_3 = 4$	94.17

4.2 Izrada vlastitog skupa podataka i testiranje

Nakon treniranja i testiranja algoritma na preuzetom setu podataka, fotografirale smo 10 različitih slika nekoliko osoba kako bismo testirale algoritam i na novim podacima, snimljenim u drugačijim uvjetima. Slikama smo potom prilagodile dimenziju na potrebnih 112×92 piksela.

Algoritam je uspješno klasificirao sve slike.

Literatura

- [1] R. Sigurdson, C. Navasca: *Randomized Tensor-based Algorithm for Image Classification*, 1985.
- [2] T. G. Kolda, B. W. Bader: *MATLAB Tensor Classes for Fast Algorithm Prototyping*, 2004.
- [3] B. N. Sheehan, Y. Saad: *Higher Order Orthogonal Iteration of Tensors (HOOI) and its Relation to PCA and GLRAM*, 2007.