МАТЕМАТИЧКИ ОМНИБУС, 1 (2017), 115 – 123

КОМПРЕСИРАЊЕ ДИГИТАЛНИ СЛИКИ СО ПРИМЕНА НА SVD РАЗЛОЖУВАЊЕ НА МАТРИЦИ

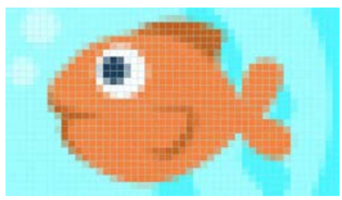
Филип Николовски [1]

1. ВОВЕД

Технолошкиот напредок во последните неколку децении овозможи луѓето да имаат лесен, брз и едноставен пристап до огромни количества информации. Ова не се однесува само на пристапот до информации, туку и на креирањето на информации во смисла на: слики, видео записи, документи и сл. Така, економичното складирање на споменатите информации почнува да игра сѐ поважна улога.

Особено чест и популарен начин на креирање нова содржина се сликите. Во овој труд под *(дигитална) слика* ќе го подразбираме кое било нумеричко претставување на дводимензионална слика со помош на компјутер. Според начинот на кои се зачувани, постојат два вида слики: *векторски* и *растерни*. Векторските слики се во суштина датотеки кои содржат математички опис на она што треба да се види, а задача на компјутерот при прикажувањето е да им пристапи на овие информации и да ја исцрта сликата. Растерните слики, пак, претставуваат *матрица од пиксели*, т.е. точки (види Слика 1.). За секоја од овие точки се чува информација за нејзината боја. Во овој случај компјутерот има задача на екранот да ја прикаже бојата која му одговара на секој пиксел и така да ја формира конечната слика.

Двата типа имаат свои предности и недостатоци. Векторските слики се посоодветни за математички и технички цртежи, а заради начинот на генери-рање на графиката овозможуваат зголемување и намалување на големината на сликата без загуба на квалитет; од друга страна, пак, овој тип слики се генерално несоодветни за зачувување фотографии. Растерните слики се секако природен избор кога е во прашање начинот на зачувување фотографии, но постојат два големи проблеми со овој тип. Имено, колку повеќе пиксели има сликата, толку повеќе информации треба да се зачуваат; дополнително, информациите се зачувани дискрет-



Слика 1. Пример за растерна слика.

но, па секое зголемување или намалување на големината на сликата резултира со загуба на квалитет.

Во овој труд ќе се фокусираме само на растерните слики и ќе разгледаме еден метод кој овозможува овие слики да може да ги зачуваме на поекономичен начин (т.е. така што тие заземаат помалку меморија), а притоа да нема забележлива загуба на квалитет (т.е. загубата да е незабележлива за човечкото око). Во продолжение трудот е организиран на следниот начин. Во вториот дел формално се воведуваат некои од основите поими кои се користат, како и математичкиот апарат со чија помош се анализираат овие поими. Во третиот дел е даден конкретен пример на кој се објаснува изложениот метод, додека во четвртиот дел е даден заклучокот.

2. МАТЕМАТИЧКИ АПАРАТ

Веќе споменавме дека растерните слики може да ги претставиме со помош на матрица. Нека сликата е претставена со помош на матрица $A=[a_{ij}]$ од ред $m \times n$. Тогаш пикселите на сликата се претставени преку елементите на матрицата, а димензиите на матрицата се нарекуваат *резолуција на сликата*. Елементите главно, но не исклучиво, содржат информации за бојата на пикселите. Како последица, на матрицата со која е претставена сликата може да се применат резултати од линеарната алгебра. На пример, сите графички трансформации на сликата може да се направат преку соодветни матрични трансформации. Претставена преку матрица на ваков начин, сликата зафаќа $m \cdot n$ мемориски места. Во продолжение ќе разгледаме еден начин на разложување (факторизација) на матрицата кој може да го искористиме за да ја намалиме меморијата потребна за зачувување на сликата, а притоа на одреден начин да го контролираме квалитетот на новодобиената слика. Од многуте достапни матрични разложувања, ќе го искористиме *разложувањето по сингуларни вредности* (анг. *singular value decomposition – SVD*). Ја имаме следнава теорема [2, стр. 109]:

Теорема 1. (за разложување по сингуларни вредности)
Нека $A=[a_{ij}]$ е произволна $m \times n$ матрица и нека $m \geq n$. Тогаш постојат матрици U од ред $m \times n$ и V од ред $n \times n$ за кои важи $UU^T = U^T U = I_m$ и $VV^T = V^T V = I_n$ (т.е. U и V се ортогонални, а

I_p *е единична матрица од ред* p *) и постои дијагонална матрица* S *од ред* $n \times n$ *:*

$$S = \mathrm{diag}(\sigma_1, \sigma_2, ..., \sigma_n) = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_n \end{bmatrix}$$

каде што $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n \geq 0$*, при што важи:*

$$A = USV^T \ .$$

Броевите $\sigma_1, \ldots, \sigma_n$ се нарекуваат *сингуларни вредности*, а записот $A = USV^T$ – *разложување по сингуларни вредности*. Да забележиме дека за секоја матрица A, ова разложување секогаш постои и сингуларните вредности се еднозначно определени. Ако $m < n$, тогаш разложувањето се дефинира во смисла на A^T.

Вистинската моќ на Теоремата 1 се состои во следниов запис на разложувањето по сингуларни вредности. Нека со $u_1, \ldots, u_n$ и $v_1, \ldots, v_n$ се означени вектор-колоните на матриците U и V, соодветно:

$$U = [u_1 \mid u_2 \mid \ldots \mid u_n]$$

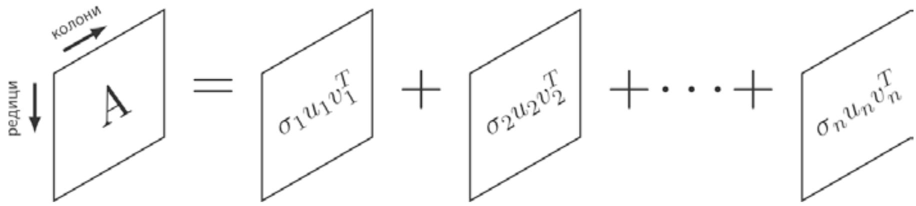
и

$$V = [v_1 \mid v_2 \mid \ldots \mid v_n].$$

Тогаш матрицата A може да се претстави како:

$$A = USV^T = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \ldots \sigma_n u_n v_n^T \ .$$

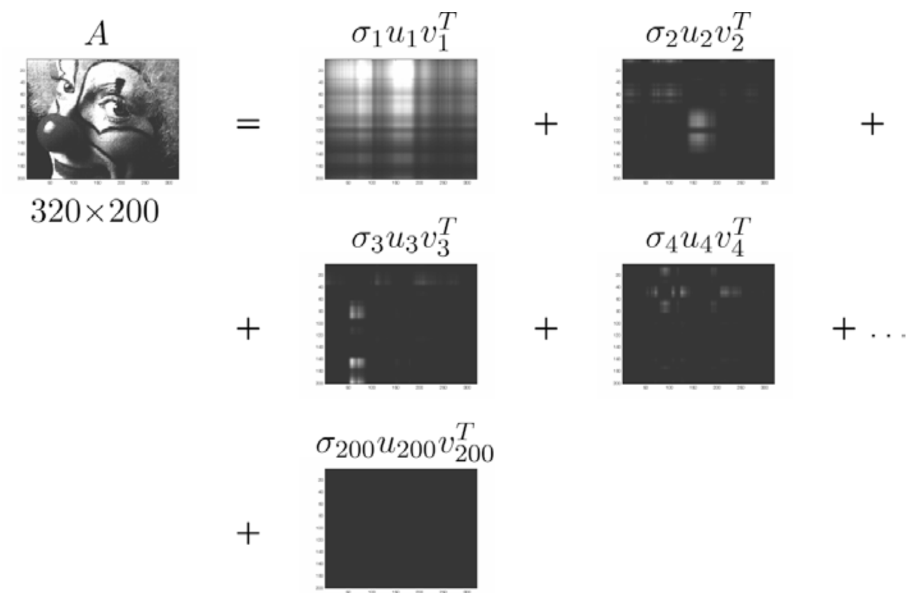
Секој од собироците на десната страна претставува матрица од ред $m \times n$ со ранг единица, помножен со соодветна сингуларна вредност. Оттука заклучуваме дека секоја слика може да се претстави како збир од слики чии матрици имаат ранг единица и нив ќе ги нарекуваме *елементарни слики*. Понатаму, колку една сингуларна вредност е поголема, толку нејзината соодветна елементарна слика е „поважна“, додека елементарните слики на кои им одговараат сингуларни вредности близу до нулата се „помалку важни“. Така, ако големината на сингуларните вредности се сфати како „интензитет“ на соодветните елементарни слики, секоја слика може



Слика 2. Визуелно претставување на SVD разложувањето на матрица ([1]).

да се претстави преку „слоеви“ со одреден интензитет при што секој слој е производ на скалар и матрица со ранг единица. Ова е прикажано на Слика 2. Следниов пример го илустрира ваквото раслојување на вистинска слика.

Пример 1. Сликата `clown.mat` на која се однесува овој пример е една од стандардните датотеки кои се дел од програмскиот пакет MATLAB ([2]). На Слика 3 таа е означена како матрица A која има димензии 320×200 и е од ранг 200. Покрај целосната слика, дадени се првите четири и последната, двестота, елементарна слика. Од прикажаното се гледа каква разлика постои помеѓу елементарните



Слика 3. „Раслојување“ на дигитална слика со помош на SVD разложување на нејзината матрица.

слики.

3. КОМПРЕСИРАЊЕ СЛИКИ СО ПОМОШ НА SVD РАЗЛОЖУВАЊЕ

Согласно Теорема 1, сингуларните вредности на матрицата се подредени во опаѓачки редослед долж главната дијагонала на матрицата S во SVD разложувањето. Тоа значи дека колку индексот на сингуларната вредност е помал, дотолку поважна е елементарната слика која ѝ одговара. Природно се поставува прашањето: што ќе се случи со првичната слика ако, едноставно, занемариме еден дел од помалку важните слоеви? Секако, очекуваме мемориските капацитет кој е потребен за зачувување на сликата да се намали, бидејќи во суштина земаме само дел(ови) од првичната слика; понатаму, вака добиениот резултат би бил само *приближно* претставување на првичната слика бидејќи изоставуваме некои нејзини елементарни слики. Во продолжение ги разгледуваме нештата од поформално гледиште.

Нека A е матрицата на оригиналната слика за која претпоставуваме дека е од ранг n. Нека $1 \leq k < n$ е индекс и да дефинираме матрица A_k на следниов начин:

$$A_k = \sigma_1 u_1 v_1^T + \ldots + \sigma_k u_k v_k^T = \sum_{i=1}^{k} \sigma_i u_i v_i^T .$$

Тогаш A_k е матрица од ранг најмногу k како збир на k матрици од ранг единица. Во смисла на првичната слика, матрицата A_k ѝ одговара на сликата која е добиена со комбинација на елементарните слики од A кои им одговараат на k-те најголеми сингуларни вредности на A; во смисла на матрици, A_k ја *апроксимира* A со одредена точност и притоа A_k е матрица од ранг k која најдобро ја апроксимира A. Формално, ја имаме следнава теорема [2, стр. 110-111].

Теорема 2. (апроксимација на матрица со матрица од помал ранг) *Нека A е матрица од ранг n и нека $1 \leq k < n$. Матрицата која е од ранг k и е „најблиску“ до A, т.е. е решение на задачата на оптимизација:*

$$\min\{ \|A - B\|_2 \;\big|\; B \in \mathbb{R}^{m\times n}, \operatorname{rank}(B) = k \}$$

е дадена со:

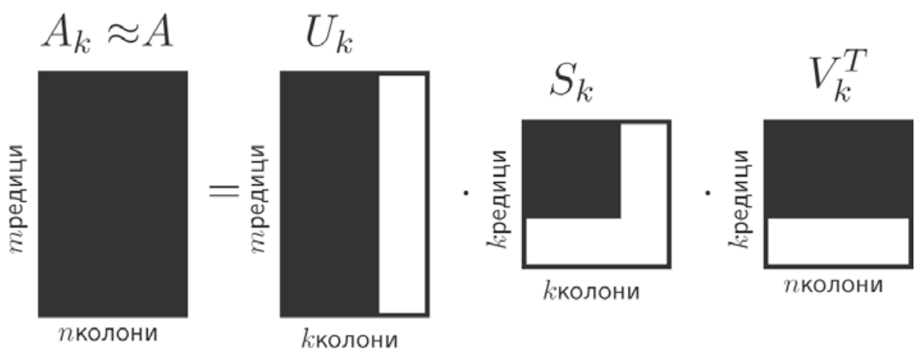
$$A_k = \sigma_1 u_1 v_1^T + \ldots + \sigma_k u_k v_k^T = \sum_{i=1}^{k} \sigma_i u_i v_i^T$$

при што $\|A - A_k\|_2 = \sigma_{k+1}$, *каде со* $\|\cdot\|_2$ *е зададена стандардната евклидска норма.* A_k се нарекува *компактно разложување по сингуларни вредности* и се бележи со $A_k = U_k S_k V_k^T$ каде што:

$$\begin{aligned} U_k &= [u_1 \mid \ldots \mid u_k] \\ S_k &= \mathrm{diag}(\sigma_1, \ldots, \sigma_k) \\ V_k &= [v_1 \mid \ldots \mid v_k]. \end{aligned}$$

Овој резултат е визуелно прикажан на Слика 4.

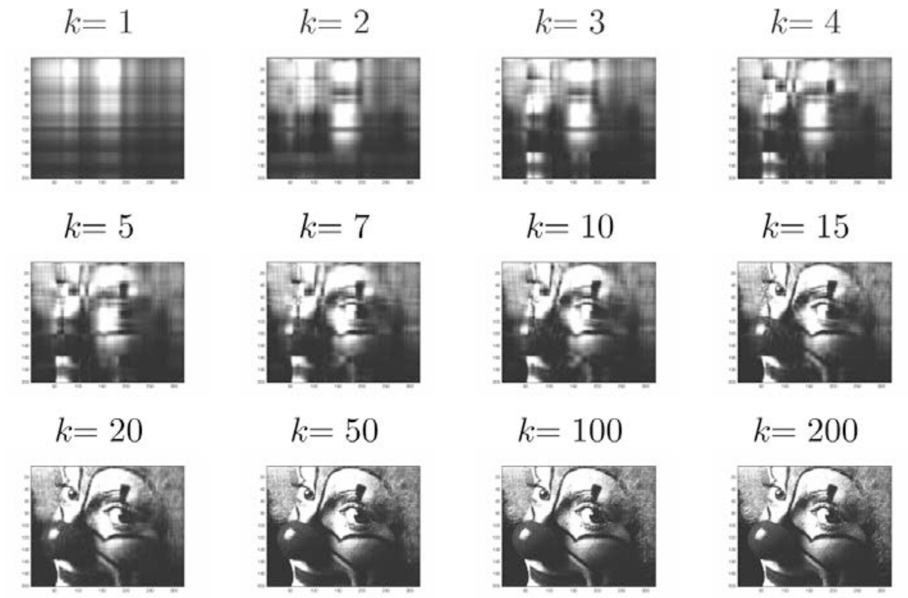
□



Слика 4. Конструкција на матрицата A_k од Теорема 2 ([3]).

Согласно Теорема 2, за конструкција на матрицата A_k потребни ни се: $m \cdot k$ мемориски единици за вектор-колоните $u_1, \ldots, u_k$ и $n \cdot k$ единици за вектор-колоните $\sigma_1 v_1, \ldots, \sigma_k v_k$, односно вкупно $(m+n) \cdot k$ единици. Со $\frac{m+n}{mn} \cdot k$ е даден односот на мемориските капацитети на сликите/матриците A и A_k; ако овој однос е помал од единица, тогаш матрицата A_k зафаќа помалку меморија од матрицата A, што значи дека сме успеале да извршиме компресија на оригиналната слика. Овој однос се нарекува *коефициент на компресија*.

Пример 2. Кај сликата од Пример 1, $m = 320$ и $n = 200$. За секој k, $1 \le k \le 200$, коефициентот на компресија изнесува $520k / 64000$,



Слика 5. Реализации на сликите кои одговараат на матриците A_k за различни вредности на k.

т.е. приближно $k / 123$. Оттука следи дека сликата може да се компресира за сите $k < 123$, односно дека A_k ќе има поголем мемориски капацитет од A за сите $k = 123, \dots, 200$. Во Табела 1 се прикажани коефициентите на компресија за неколку вредности на k заедно со процентот од меморијата на првичната матрица, додека на Слика 5 се прикажани соодветните реализации на сликата за одбрани вредности на k.

Од приложеното на Слика 5 се гледа дека компресијата резултира со намалување на квалитетот на новодобиените слики кој е особено изразен за мали вредности на k ($k < 10$). Веќе за $k = 20$ сликата е доволно „чиста“, додека за $k = 50, 100, 200$ разлика помеѓу првичната слика и компресираните слики не е воочлива. Како мерка за квалитетот на компресираната слика може да ја искористиме оценката $\|A - A_k\|_2 = \sigma_{k+1}$ од Теорема 2, земајќи ја за доволно квалитетна компресираната слика A_m за чиј индекс важи:

$$m = \min_{k=1,\dots,n} \{\sigma_{k+1} \le \tau\}$$

за однапред зададена вредност на толеранција τ. Да забележиме дека согласно Табела 1, сликата добиена за $k=50$ зафаќа двојно помалку меморија од сликата добиена за $k=100$, додека сликата за $k=200$ зафаќа повеќе меморија од првичната слика и во овој случај не може да зборуваме за компресија ($k \geq 123$).

k	$\approx k$ / 123	%
1	0,008	0,8%
3	0,024	2,4%
5	0,041	4,1%
10	0,081	8,1%
15	0,122	12,2%
20	0,163	16,3%
50	0,406	40,6%
100	0,813	81,3%
150	1,219	121,9%
200	1,625	162,5%

Табела 1. Коефициент на компресија за одбрани вредности на *k*.

4. ЗАКЛУЧОК

Во овој труд презентиравме еден начин на кој со примена на резултати од линеарна алгебра може да се изврши компресија на дигитални слики, т.е. намалување на нивниот мемориски капацитет. Ова го направивме со помош на разложувањето на матрица по сингуларни вредности, го применивме на конкретен пример и дадовме неколку реализации на различни степени на компресија. Иако од примерите јасно се гледаат ефектите на компресијата и во намалување на меморискиот капацитет, но и во загуба на квалитет, сепак постојат неколку практични предизвици кои треба да се имаат предвид при примена на опишаниот пристап. Имено, дадената постапка на компресија се заснова на разложување на матрица по сингуларни вредности, но во овој труд само претпоставивме дека таквото разложување е веќе достапно без да се осврнеме на проблемите кои може да настанат при неговата конструкција, на пример: колку димензиите на матрицата влијаат на исплатливоста на конструкцијата, кои методи на конструкција може или е пожелно да се користат, итн.

Во секој случај, разложувањето по сингуларни вредности се покажува како силна алатка во анализа на структурата на матриците. Во таа смисла тоа може да се искористи за „отфрлање“ на некои непожелни својства, секако доколку е јасно дека (или како) тие зависат од сингуларните вредности. Во овој контекст, природно се наметнува идејата дека со помош на ова разложување и своевидна „компресија“, може да се направи обид за филтрирање на „шум“ кој ги „загадува“ основите податоци кои ги содржи матрицата. Ова би било особено корисно во оние случаи каде шумот има силно влијание врз пресметките; еден таков случај се итеративните процеси каде што грешките кои се резултат од пресметките се пренесуваат во последователните итерации и на тој начин ја намалуваат пресметковната ефикасност на процесот, може значително да го зголемат потребниот број на итерации за решавање на поставената задача, па дури и да го „срушат“ целиот процес.

ЛИТЕРАТУРА

[1] H. C. Andrews, C. L. Patterson, *Singular value decompositions and digital image processing*, IEEE Transactions on Acoustics, Speech and Signal Processing, 24(1) (1976), 26 – 53.

[2] J. W. Demmel, *Applied Numerical Linear Algebra*, SIAM, 1997.

[3] R. L. Burden, J. D. Faires, *Numerical Analysis 9th ed.*, Brooks/Cole, Cengage Learning, 2010.

[4] G. H. Golub, C. F. Van Loan, *Matrix Computations 4th ed.*, Johns Hopkins University Press, 2013.

[1] Меѓунардни Училишта НОВА,
Прашка 27, 1000 Скопје, Р. Македонија
e-mail: filipnikolovski@gmail.com

Примен: 24.05.2017
Поправен: 17.07.2017
Одобрен: 18.07.2017