AMRITA VISHWA VIDYAPEETHAM | School of Engineering
AMRITAPURI | BENGALURU | CHENNAI | COIMBATORE

Recommendation System & Image Processing using Singular Value Decomposition.

Team 7

CH.EN.U4AIE21055 - Surya Santhosh V R
CH.EN.U4AIE21054 - Sriram S
CH.EN.U4AIE21028 - Mithilesh A
CH.EN.U4AIE21041 - Roshan George
CH.EN.U4AIE21008 - Balaji M
CH.EN.U4AIE21025 - Logeshwar B

Introduction

Recommendation System:

Implementation of singular value decomposition (SVD) based on collaborative filtering in the task of movie recommendation. For simplicity, the MovieLens 1M Dataset has been used. This dataset has been chosen because it does not require any preprocessing as the main focus of this article is on SVD and recommender systems.

Collaborative Filtering

The assumption of this approach is that people who have liked an item in the past will also like the same in future. This approach builds a model based on the past behavior of users In this way, the model finds an association between the users and the items.

Introduction

Image Processing :

A two-dimensional function f (x, y) (2-D image) can be used to describe an image. The amplitude of f at any pair of (x, y) equals the grey level of the 3D image.

The term "digital picture" refers to an image where x, y, and the amplitude value of f are finite, discrete numbers. Picture elements, or pixels, are the name for the finite set of digital values.

Usually, a two-dimensional array or real number matrix is used to store the pixels in computer memory. Individual 2-D photos are combined to create color images. By processing each of the three image components separately, many image processing algorithms for monochrome photos can be extended to color images (3-D).

Objective

This research aims to use the "Singular Value Decomposition (SVD)" method of linear algebra to mid-level image processing, particularly in the areas of picture reduction and recognition.

The process involves splitting a matrix A into three new matrices, U, S, and V, where U and V are orthogonal matrices and S is a diagonal matrix.

The study also exhibits how to employ SVD approach for image processing in the domain of face recognition (FR) under various terms k of singular value and the outer product expansion of image matrix A for image compression.
The assumption that the matrix A in this project has m lines and n columns (m x n) is made purely for convenience; the results are still valid if n x m [8].

In this research, Python is utilized as the programming and experimentation platform due to its great performance in merging computation, visualization, and programming.

Modules, exceptions, dynamic typing, and very high-level dynamic data types and classes are all included in the Python programming language.

Theory of SVD

Singular Value Decomposition plays an important role matrix decomposition that's why it is mainly used in image processing. If you decompose a matrix A you will be getting three matrices as result

$A = U\Sigma V^T$

Where U is Left Singular vectors and V is Right Singular vectors . These both are orthonormal matrices, S is a diagonal matrix with singular values in the matrix.

Steps to find these values are

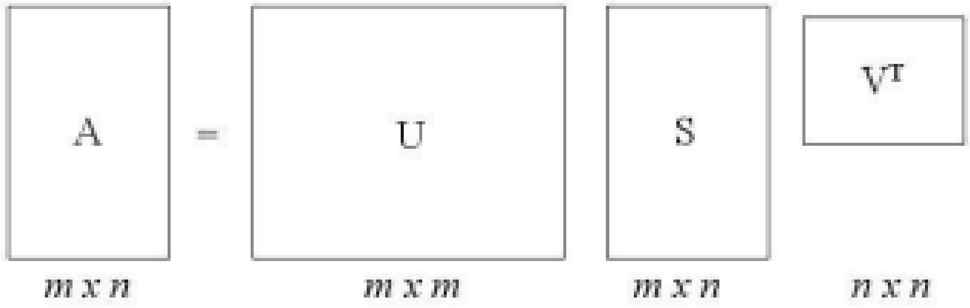
- Find the product of A and A^T
- Find the eigen value and eigen vector of the dot product
- Change it into orthonormal vector
- The Eigen values after orthonormal vector as a single matrix is said V
- Find the singular values by taking the sqrt of eigen values and substitute in Σ
- Now $AV\Sigma^T = U$
- By these steps we can find U V and S

Theory of SVD

Let’s say we have a matrix A with m rows and n columns. Then the A can be factorized into three matrices:

$$A = USV^T$$

Illustration of Factoring A to U∑VT:



Where Matrix U is an m × m orthogonal matrix And matrix V is an n × n orthogonal matrix.

Theory of SVD

Here, S is an m × n diagonal matrix with singular values (SV) on the diagonal.

$$S = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_r & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & \sigma_{r+1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & \sigma_n \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix}$$

For i = 1, 2, …, n, sigma are called Singular Values (SV) of matrix A. It can be proved that

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$$

Methodology

Image Processing

SVD Approach for Image Compression

Compression is achieved by the removal of three basic data redundancies:

1) coding redundancy, which is present when less than optimal;
2) interpixel redundancy, which results from correlations between the pixels;
3) psychovisual redundancies, which is due to data that is ignored by the human visual

The singular values of a matrix decrease quickly with increasing rank. This property helps us in compressing matrix data by eliminating the small singular values or higher ranks.

A is the outer product expression

$$A = USV^T = \sum_{i=1}^{r} \sigma_i \mathrm{u}_i \mathrm{v}_i^T$$

$$A = \sigma_1 \mathrm{u}_1 \mathrm{v}_1^T + \sigma_2 \mathrm{u}_2 \mathrm{v}_2^T + \cdots + \sigma_r \mathrm{u}_r \mathrm{v}_r^T$$

Methodology

Recommendation System

User	M1	M2	M3	M4	M5	M6	M7
A	4			5	1		
B	5	5	4				
C				2	4	5	
D		3					3

Jacket Similarity = No.of movies rated in common/total number of movies rated

Sim(A,B)=1/5   Sim(A,C)=½

Sim(A,B) < Sim(A,C)

Cosine similarity = Ra.Rb/|Ra||Rb|

Sim(A,B) = 4*5+0*5+4*0+5*0+1*0/sqrt[(4)^2+(5)^2+(1)^2]*sqrt[(5)^2+(4)^2+(5)^2]

Sim(A,B)=0.38   Sim(A,C)=0.32

Sim(A,B) >Sim(A,C)

If the angle is less then the similarity is more

Normalize = R – mean(Row)

Methodology

After Normalization

User	M1	M2	M3	M4	M5	M6	M7
A	2/3			5/3	-7/3		
B	1/3	1/3	-2/3				
C				-5/3	1/3	4/3	
D		0					0

Normalize = R – mean(Row)
Sim(A,B)=0.09   Sim(A,C)=-0.58
Sim(A,B)>Sim(A,C)
Her, (A,C) is in negative it represents only the direction but the difference is less in (A,B) so, It is more similar than (A,C)

Methodology

M1	M2	M3	M4	M5	
1	1	1	0	0	
3	3	3	0	0	Sci-fi
4	4	4	0	0	
5	5	5	0	0	
0	2	0	4	4	
0	0	0	5	5	Rom
0	1	0	2	2	

=

0.13	0.02	-0.01
0.14	0.07	-0.03
0.55	0.09	-0.04
0.68	0.11	-0.05
0.15	-0.59	-0.65
0.07	-0.73	-0.67
0.07	-0.29	0.32

U

X

12.4	0	0
0	9.5	0
0	0	1.3

Sigma

X

0.56	**0.59**	**0.56**	**0.56**	**0.09**
0.12	-0.02	-0.02	0.12	0.69
0.55	0.09	-0.80	0.40	0.09

V

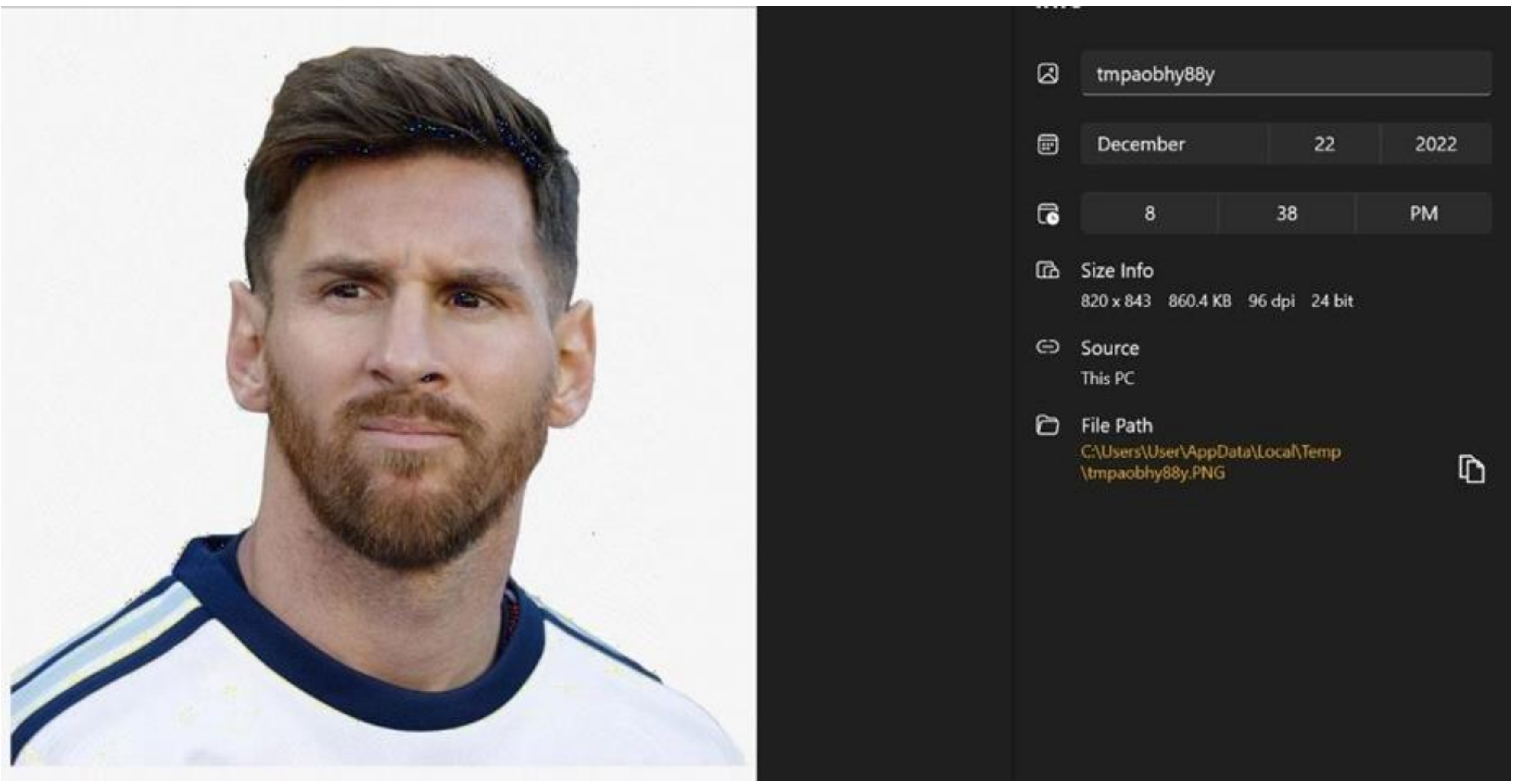
If you need to recommend a movie for user with q = [5,0,0,0,0] then you have to multiply with the Vi i.e. q*Vi

Then the result will be [2.8(Sci-fi) 0.6(Rom)]

EXPERIMENTATIONS AND RESULTS

Result of Experimentations for Image processing

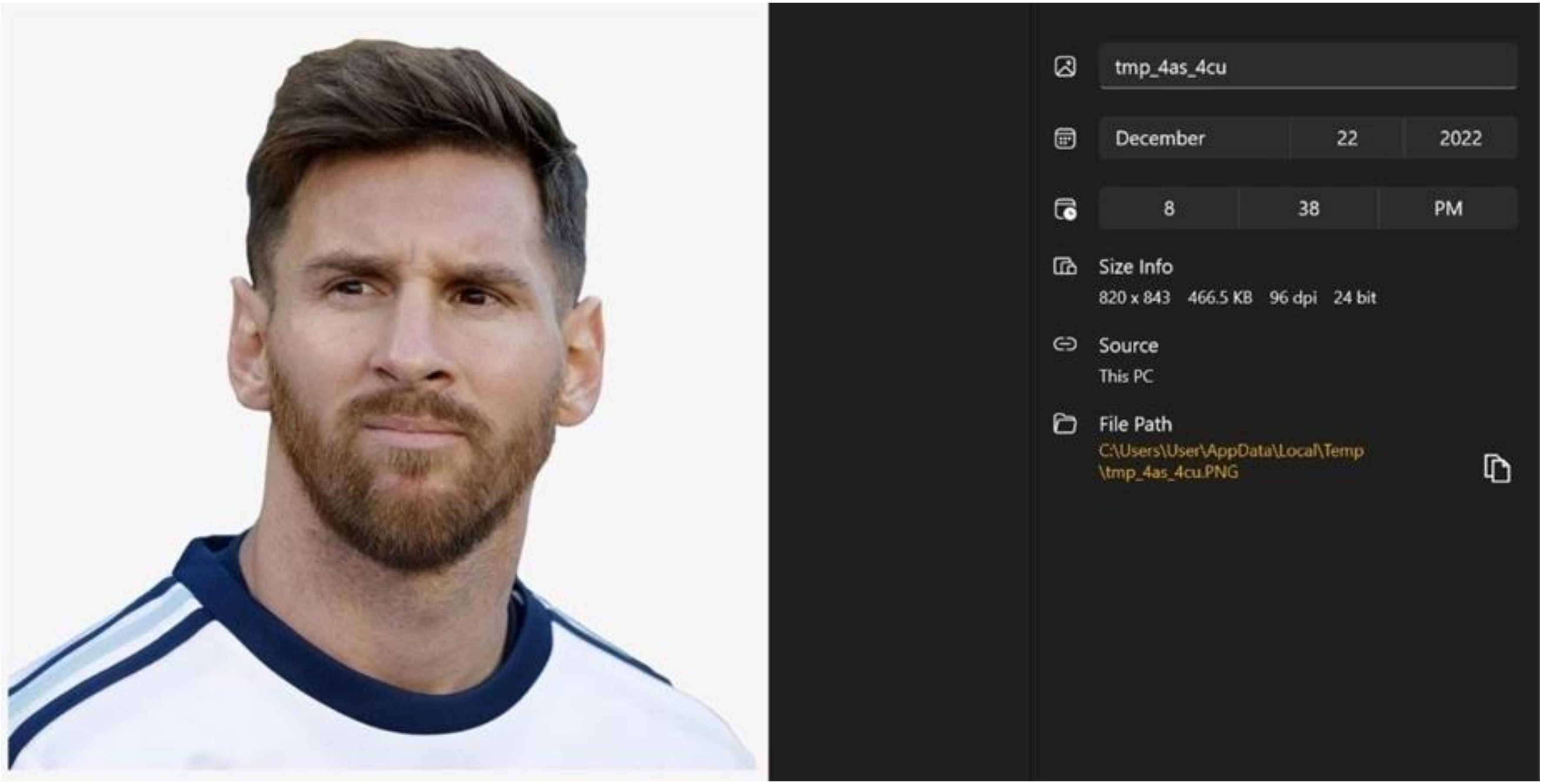
Before Compression:



EXPERIMENTATIONS AND RESULTS

Result of Experimentations for Image processing

After Compression:



EXPERIMENTATIONS AND RESULTS

Result of Experimentations for Recommendation System

```
Recommendations for Contender, The (2000):

Contender, The (2000)
Nurse Betty (2000)
Almost Famous (2000)
Wonder Boys (2000)
Best in Show (2000)
Meet the Parents (2000)
Saving Grace (2000)
Remember the Titans (2000)
Duets (2000)
Requiem for a Dream (2000)
```

CONCLUSION AND FUTURE WORK

This project has applied technique of linear algebra “singular value decomposition (SVD)” to digital image processing. Two specific areas of image processing are investigated and tested. Basis on the theory and result of experiments, we found that SVD is a stable and effective method to split the system into a set of linearly independent components, each of them is carrying own data (information) to contribute to the system, Thus, both rank of the problem and subspace orientation can be determined.