

ĐẠI HỌC QUỐC GIA TP. HỒ CHÍ MINH
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN

KHOA KHOA HỌC MÁY TÍNH



BÁO CÁO ĐỒ ÁN CUỐI KỲ

**Đề tài: Image Compression use Singular Value
Decomposition and Kmeans Clustering**

GVHD: Đỗ Văn Tiến

Lớp: CS232.N21

Sinh viên thực hiện: Hồ Hồng Hà - 20520480

TP. Hồ Chí Minh, ngày 20 tháng 6 năm 2023

Mục lục

I. Tổng quan đề tài	3
II. Quy trình thực hiện.....	3
1. K-means clustering	3
2. Singular Value Decomposition	4
3. Phương pháp đánh giá.....	7
III. Kết quả thực nghiệm và đánh giá.....	7
1. K-means clustering	7
2. Singular Value Decomposition	9
3. Nhận xét.....	11
Tài liệu tham khảo	11

I. Tổng quan đề tài

- Ảnh số ngày càng được sử dụng nhiều vì có những ưu điểm vượt trội so với ảnh truyền thống (ảnh chụp bằng phim). Trong những năm gần đây, lĩnh vực xử lý ảnh số đã được nghiên cứu mạnh mẽ và có nhiều ứng dụng trong thực tế. Một số hướng nghiên cứu quan trọng trong xử lý ảnh có thể kể đến là nén ảnh, xác thực ảnh, nhận dạng ảnh...
- Nén hình ảnh là một kiểu nén dữ liệu được áp dụng cho hình ảnh kỹ thuật số mà làm giảm chất lượng của hình ảnh đến mức có thể chấp nhận được. Nén ảnh nhằm giảm chi phí lưu trữ và truyền tải chúng.
- Trong bài toán này em sẽ thực hiện tìm hiểu về nén ảnh sử dụng hai phương pháp là K-means clustering và Singular Value Decomposition

II. Quy trình thực hiện

1. K-means clustering

- K-means clustering (phân cụm K-means) là thuật toán cơ bản nhất trong Unsupervised learning.
- Trong thuật toán k-Means mỗi cụm dữ liệu được đặc trưng bởi một tâm (centroid). tâm là điểm đại diện nhất cho một cụm và có giá trị bằng trung bình của toàn bộ các quan sát nằm trong cụm. Chúng ta sẽ dựa vào khoảng cách từ mỗi quan sát tới các tâm để xác định nhãn cho chúng.
- Cụ thể các bước của thuật toán k-Means được tóm tắt như sau:
 1. Khởi tạo ngẫu nhiên k tâm cụm $\mu_1, \mu_2, \dots, \mu_k$
 2. Xác định nhãn cho từng điểm dữ liệu c_i dựa vào khoảng cách tới từng tâm cụm:

$$c_i = \arg \min_j \|\mathbf{x}_i - \mu_j\|_2^2$$

- $\|\mathbf{x}\|_2^2$ là bình phương của norm chuẩn bậc 2, kí hiệu là L_2 , norm chuẩn bậc 2 là một độ đo khoảng cách thường được sử dụng trong machine learning, được tính theo công thức khoảng cách Euclid.

3. Tính toán lại tâm cho từng cụm theo trung bình của toàn bộ các điểm dữ liệu trong một cụm:

$$\mu_j := \frac{\sum_{i=1}^n \mathbf{1}(c_i = j) \mathbf{x}_i}{\sum_{i=1}^n \mathbf{1}(c_i = j)}$$

4. Lặp lại bước 2 đến khi không thể thay đổi được vị trí hạt nhân nữa (việc gán dữ liệu vào từng cluster ở bước 2 không thay đổi so với vòng lặp trước nó).

- K-means clustering trong nén ảnh:

○ Quy trình thực hiện:

- Bước 1: Load ảnh input
- Bước 2: Làm phẳng hình ảnh
- Bước 3: Thực hiện K-means
- Bước 4: Thay thế các giá trị bằng các giá trị Centroid tương ứng
- Bước 5: Dựng lại ảnh từ mảng đã thực hiện qua K-means

○ Với hình ảnh ban đầu, mỗi pixel sử dụng 3 kênh màu RGB, mỗi kênh cần 8 bit để lưu trữ do đó cần 24 bit cho mỗi pixel. Từ đó tính được kích thước ảnh N pixel trước khi nén là $24 \cdot N$. Sử dụng K-means Clustering để tìm k cụm (số màu) sẽ đại diện cho các màu tương tự của nó. K-màu này sẽ là các điểm trung tâm. Sau khi dùng K-means Clustering thì chỉ số của mỗi pixel cần $\log_2 k$ bit để lưu trữ. Do đó kích thước ảnh sau khi nén là $24 \cdot k + N \cdot \log_2 k$.

2. Singular Value Decomposition

- SVD (Singular Value Decomposition) là một phương pháp phân tích ma trận, có nhiều ứng dụng trong thực tế, như nén ảnh, phát hiện ảnh giả mạo, thủy văn số,

nhận diện khuôn mặt... Điểm đặc biệt của SVD là nó có thể áp dụng được trên bất kỳ ma trận thực nào.

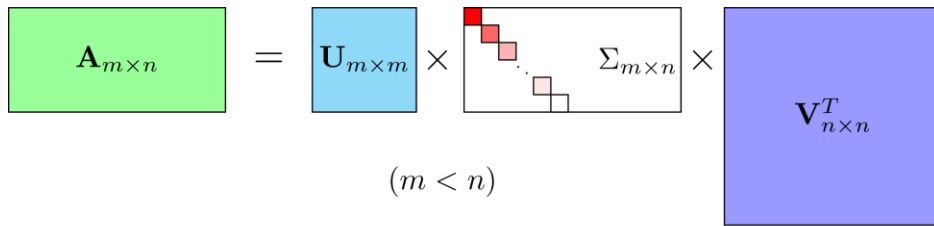
- Một ma trận $A_{m \times n}$ bất kỳ đều có thể phân tích thành dạng:

$$A_{m \times n} = U_{m \times m} \times \Sigma_{m \times n} \times (V_{n \times n})^T$$

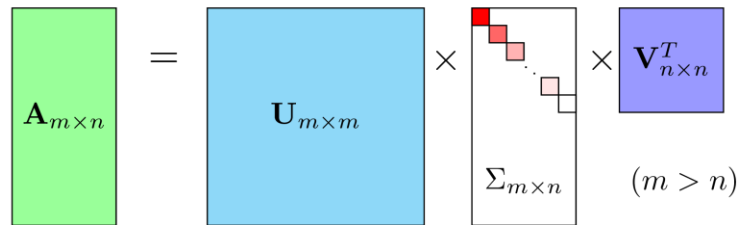
- o Trong đó:

- U và V là các ma trận trực giao.
- Σ là ma trận chéo không vuông có giá trị giảm dần trên đường chéo và là các giá trị không âm.

- o SVD cho ma trận A khi: $m < n$ (hình trên), và $m > n$ (hình dưới). Σ là một ma trận đường chéo với các phần tử trên đó giảm dần và không âm. Màu đỏ càng đậm thể hiện giá trị càng cao. Các ô màu trắng trên ma trận này thể hiện giá trị 0.



$$A_{m \times n} = U_{m \times m} \times \Sigma_{m \times n} \times V_{n \times n}^T \quad (m < n)$$



$$A_{m \times n} = U_{m \times m} \times \Sigma_{m \times n} \times V_{n \times n}^T \quad (m > n)$$

- Để tính SVD của một ma trận, chúng ta sử dụng module linalg (Linear algebra) của numpy.
- Truncated SVD:
 - o Chú ý rằng trong ma trận Σ , các giá trị trên đường chéo là không âm và giảm dần. Thông thường, chỉ một lượng nhỏ giá trị mang giá trị lớn, các giá trị còn lại thường nhỏ và gần 0. Chúng ta chỉ lấy ra k dòng và cột

tương ứng với top các giá trị riêng lớn nhất của ma trận A từ U và V^T .
 Khi đó sẽ thu được ma trận xấp xỉ của ma trận A:

$$A \approx U_k \times \Sigma_k \times (V_k)^T$$

The diagram illustrates the SVD decomposition of matrix A into U, Sigma, and V^T, and its truncated approximation A-hat. The top part shows the full decomposition: A (m x n) = U (m x m) · Σ (m x n) · V^T (n x n). The Σ matrix is shown with a blue box highlighting the top k singular values and a red diagonal line indicating the truncation of the remaining values. The bottom part shows the truncated approximation: A-hat (m x n) = U (m x k) · Σ (k x k) · V^T (k x n). The dimensions of each matrix are labeled below them.

- Truncated SVD trong nén ảnh:

○ Quy trình thực hiện:

- Bước 1: Load ảnh input
- Bước 2: Tách từng kênh màu Red, Green, Blue
- Bước 3: Thực hiện tính SVD trên từng kênh màu
- Bước 4: Áp dụng Truncated SVD với từng kênh màu
- Bước 5: Gộp các kênh màu lại để tái tạo hình ảnh

- Để lưu trữ ảnh sử dụng Truncated SVD chỉ cần lưu trữ các ma trận U_k có kích thước $m \times k$, Σ_k kích thước $k \times k$, $(V_k)^T$ kích thước $k \times n$. Số phần tử phải lưu là $k \times (m + 1 + n)$.

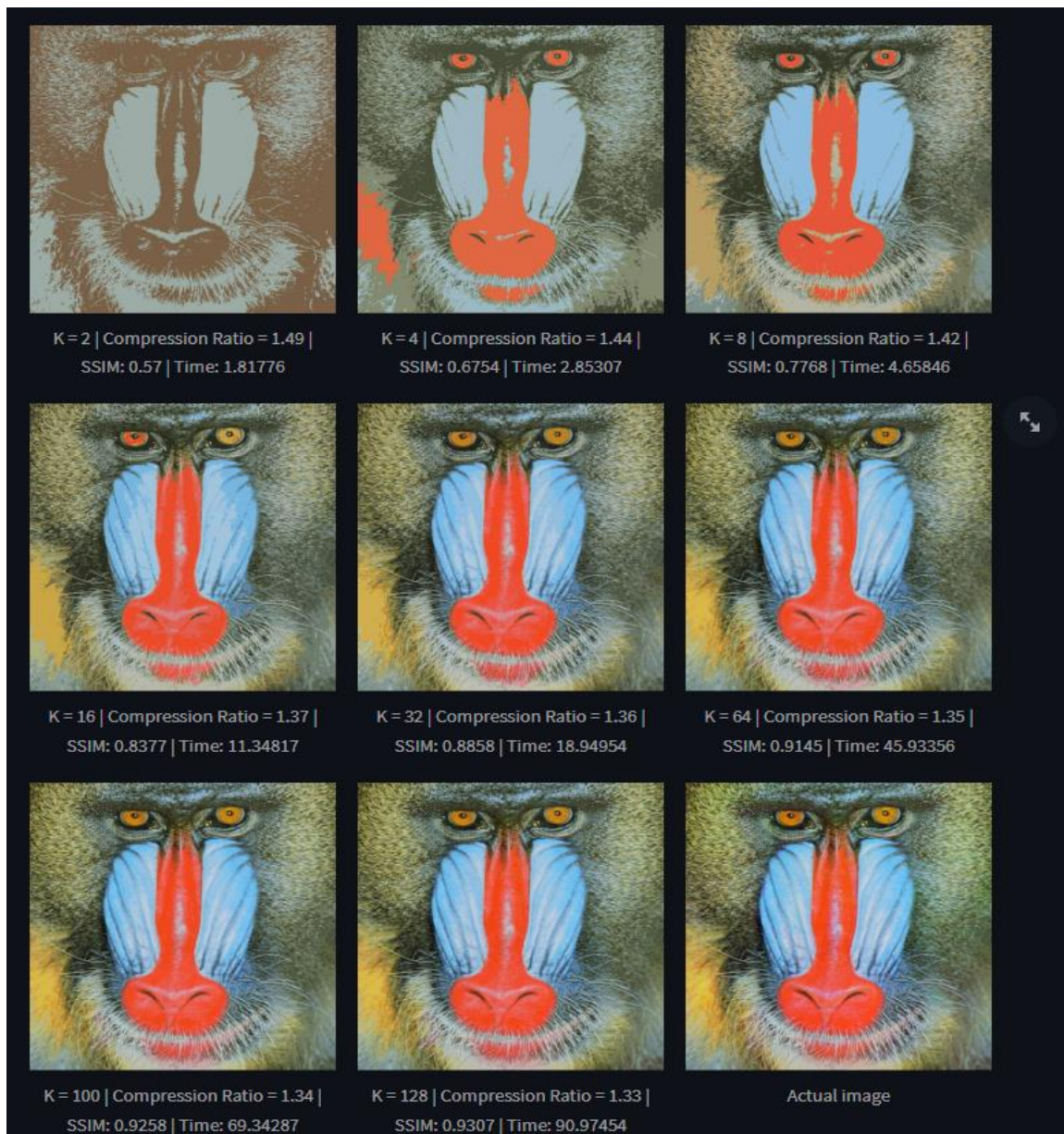
- Tỷ lệ nén là $\frac{mn}{k(m+1+n)}$

3. Phương pháp đánh giá

- Tỷ lệ nén: Tỷ lệ giữa kích thước file trước và sau khi nén.
- Structural Similarity Index Measurement(SSIM): được dùng để đánh giá độ tương đồng của hai hình ảnh.
- Thời gian thực hiện.

III. Kết quả thực nghiệm và đánh giá

1. K-means clustering



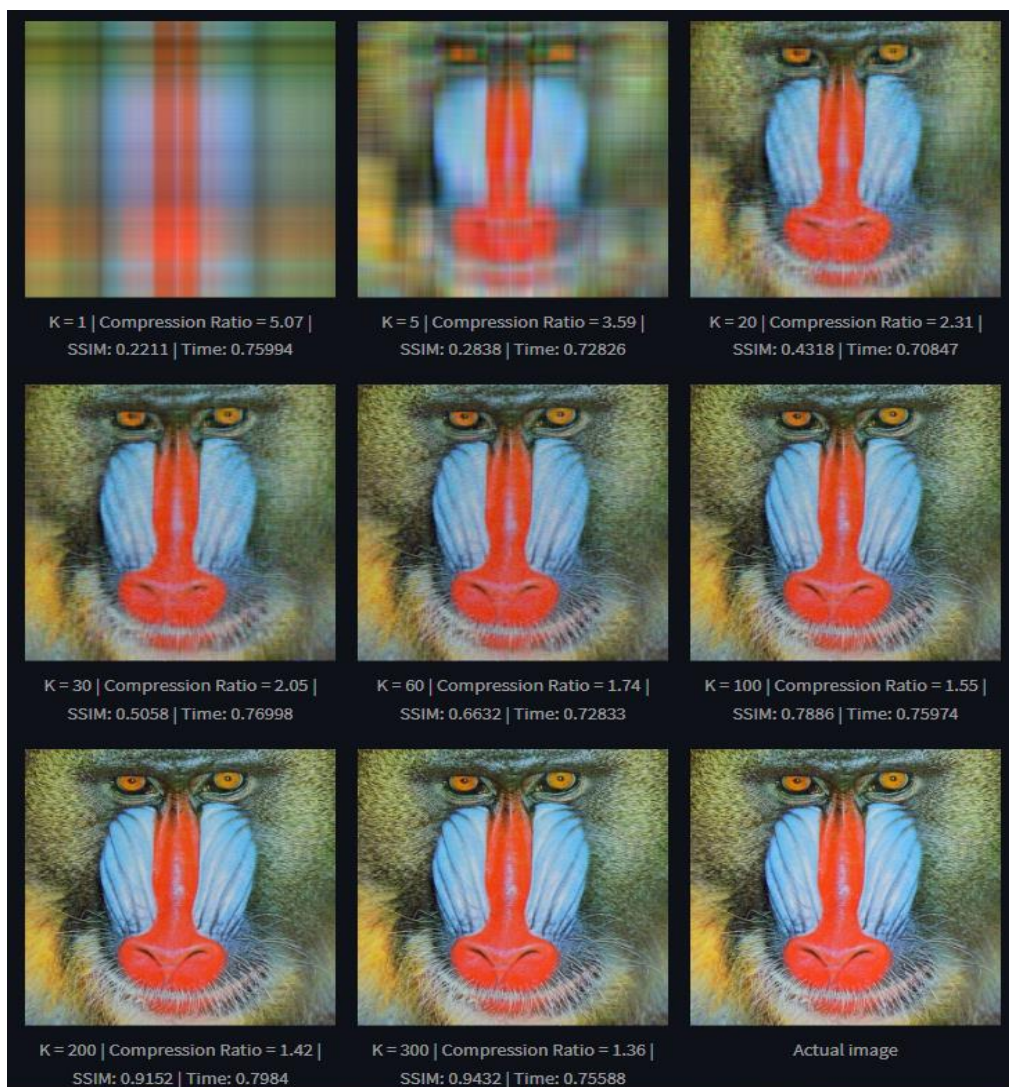


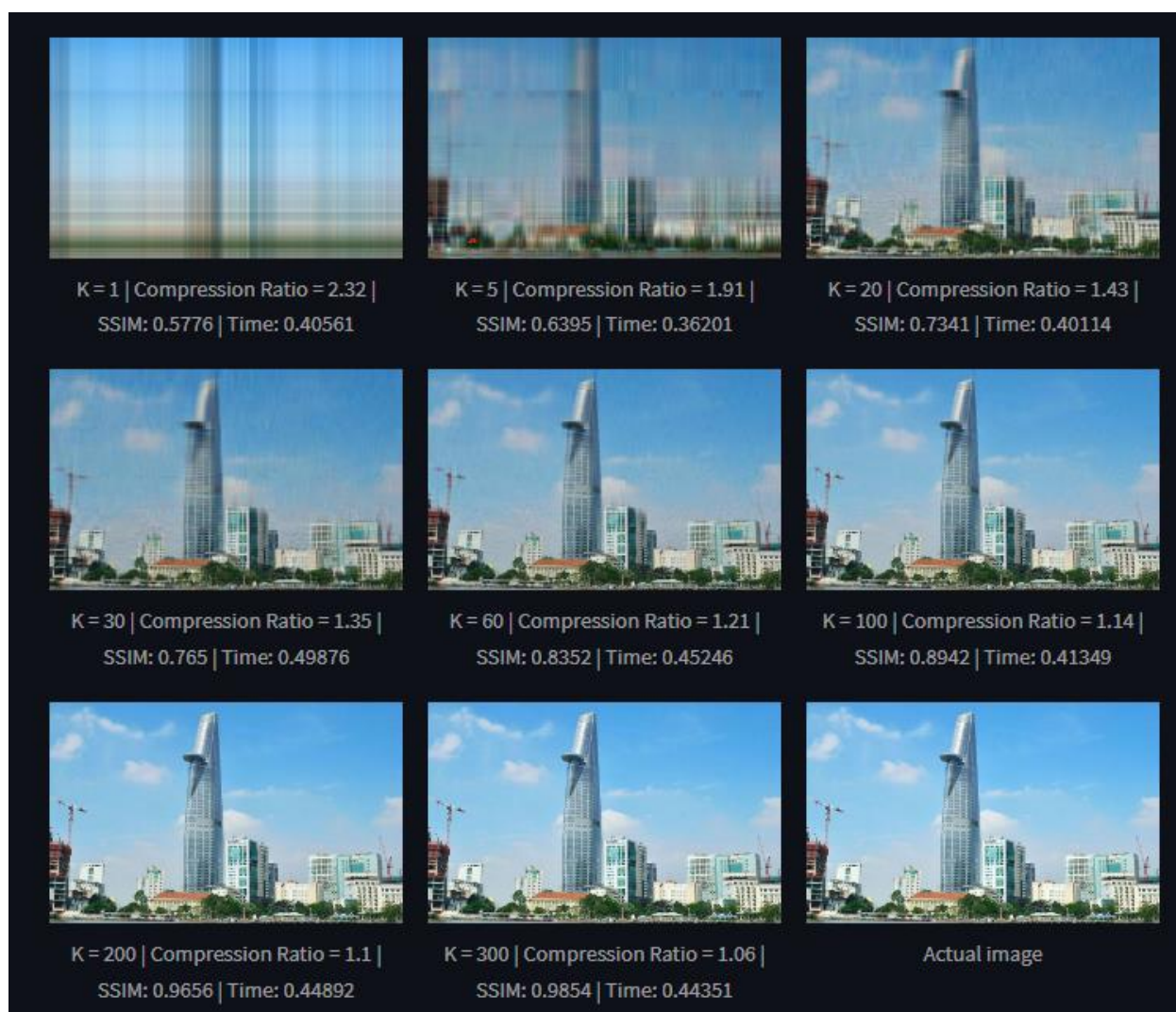
	K-Means								
	K = 2			K = 32			K = 100		
	Tỉ lệ	SSIM	Time	Tỉ lệ	SSIM	Time	Tỉ lệ	SSIM	Time
Hình 1	1.49	0.57	1.81776	1.36	0.8858	18.94954	1.34	0.9258	69.34287
Hình 2	1.4	0.7862	0.68158	1.13	0.9446	12.1239	1.09	0.9669	46.13479
Hình 3	1.72	0.4816	0.69169	1.04	0.9301	11.53847	1.03	0.9633	60.13652
Hình 4	1.39	0.5896	3.54361	1.12	0.9079	58.23279	1.09	0.9544	190.13032
Hình 5	2.81442	0.5184	1.75634	2.17	0.814	37.8604	2.18	0.8841	141.86119

- Trên đây là một số kết quả thực nghiệm trên thuật toán Kmeans Clustering với từng giá trị K khác nhau, dựa vào kết quả trên có thể thấy rằng:
 - o Với giá trị K thấp nghĩa là số màu trong ảnh thấp nên ảnh thu được có tỉ lệ nén cao hơn và chất lượng ảnh kém hơn.

- Với giá trị K càng tăng thì số màu trong ảnh tăng lên do đó tỉ lệ nén giảm, ảnh thu được có chất lượng cao hơn và giống với ảnh gốc hơn.
- Với những giá trị K nhỏ như là 2, 4 thì ảnh thu được có sự khác biệt nhiều so với ảnh gốc, tuy nhiên vẫn có thể nhận ra được phần nào đó các chi tiết trong ảnh.
- Với những giá trị K lớn hơn như là 64, 100, 128 thì ảnh thu được gần như là giống với ảnh gốc.

2. Singular Value Decomposition





	SVD								
	K = 1			K = 30			K = 100		
	Tỉ lệ	SSIM	Time	Tỉ lệ	SSIM	Time	Tỉ lệ	SSIM	Time
Hình 1	5.07	0.2211	0.75994	2.05	0.5058	0.76998	1.55	0.7886	0.75974
Hình 2	2.32	0.5776	0.40561	1.35	0.765	0.49876	1.14	0.8942	0.41349
Hình 3	3.48	0.3431	0.50948	1.49	0.6253	0.53452	1.15	0.8027	0.54613
Hình 4	3.23	0.3508	8.73845	1.63	0.5419	4.70713	1.34	0.6743	5.00873
Hình 5	6.55	0.3345	3.37774	1.38	0.2927	3.23602	1.1	0.3392	2.89249

- Trên đây là một số kết quả thực nghiệm trên thuật toán SVD với từng giá trị K khác nhau, dựa vào kết quả trên có thể thấy rằng:

- Với giá trị K thấp nghĩa là số thành phần giữ lại ít nên ảnh thu được có tỉ lệ nén cao hơn và chất lượng ảnh kém hơn.
- Với giá trị K càng tăng nghĩa là số thành phần giữ lại nhiều hơn do đó tỉ lệ nén giảm, ảnh thu được có chất lượng cao hơn và giống với ảnh gốc hơn.
- Với những giá trị K nhỏ như là 1, 5 thì ảnh thu được xem như bị phá hủy hoàn toàn, không thể nhận ra được các chi tiết trong ảnh ban đầu.
- Với những giá trị K lớn hơn như là 30,60,100...thì ảnh thu được gần đã dần rõ ràng hơn và giống với ảnh gốc hơn.

3. Nhận xét

- Xét về thời gian chạy thì SVD vượt trội hoàn toàn so với Kmeans.
- Với các giá trị K thấp thì SVD có tỉ lệ nén cao hơn so với Kmeans. Tuy nhiên ảnh thu được bởi SVD gần như là bị phá hủy, không thể nhận ra chi tiết nào trong ảnh nữa, còn với Kmeans thì vẫn có thể nhận ra được một số chi tiết trong bức ảnh.
- Với các giá trị K cao hơn thì hai phương pháp có hiệu quả tương tự nhau.

Tài liệu tham khảo

1. <https://machinelearningcoban.com/2017/06/07/svd/>
2. <https://www.geeksforgeeks.org/image-compression-using-k-means-clustering/>