

Nghiên cứu phương pháp phân tích SVD và ứng dụng trong xử lý ảnh

■ **ThS. NGUYỄN THU HƯỜNG; TS. NGUYỄN HIẾU CƯỜNG**
Trường Đại học Giao thông vận tải

TÓM TẮT: SVD (Singular Value Decomposition) là một phương pháp để phân tích một ma trận bất kỳ thành tích của ba ma trận thành phần. Phân tích SVD thường dùng để trích chọn các đặc trưng bền vững và có nhiều ứng dụng trong lĩnh vực xử lý ảnh nói chung và xử lý ảnh trong ngành GTVT nói riêng. Trong bài báo này, chúng tôi nghiên cứu về phân tích SVD, ứng dụng của nó trong nén ảnh và xác thực ảnh. Chúng tôi cũng so sánh việc sử dụng phân tích SVD với phương pháp biến đổi sang miền tần số bằng DCT (Discrete Cosine Transform) trong các ứng dụng trên.

TỪ KHÓA: SVD, DCT, nén ảnh, xác thực ảnh.

ABSTRACT: SVD (Singular Value Decomposition) is a method to decompose a matrix into the product of three matrices. SVD analysis is often used to extract robust features and has many applications in the field of image processing in general, as well as image processing in the transportation industry in particular. In this paper, we study SVD and its applications in image compression and image forensics. We also compared using SVD with using DCT (Discrete Cosine Transform) in the above applications.

KEYWORDS: SVD, DCT, image compression, image forensics.

1. ĐẶT VẤN ĐỀ

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... Các nghiên cứu về xử lý ảnh cũng đã và đang có nhiều ứng dụng trong lĩnh vực GTVT, như trong các hệ thống camera giám sát, kiểm soát ra vào bãi đỗ xe, xác định lưu lượng xe, xác định các điểm tắc nghẽn giao thông, xác thực các hồ sơ TNGT... [1,2,3].

Nén ảnh là một kỹ thuật mã hóa các ảnh số hóa nhằm giảm số lượng các bit dữ liệu cần thiết để biểu diễn ảnh. Mục đích là giảm kích thước, do đó giảm dung lượng lưu

trữ và thời gian để truyền tải ảnh, trong khi vẫn đảm bảo được chất lượng của ảnh. Có hai cách nén ảnh là nén không mất thông tin (lossless compression) và nén có mất thông tin (lossy compression). Trong thực tế, để có thể nén được nhiều, người ta thường áp dụng các phương pháp nén có mất mát thông tin. Để đảm bảo việc truyền tải thông tin được nhanh chóng, giảm độ trễ, nhiều ứng dụng xử lý ảnh trong lĩnh vực giao thông như (camera giám sát, nhận dạng biển số...) đều cần được nén [1,3].

Ảnh giả mạo là ảnh không có thật, có được do sự nguy tạo bởi các quá trình xử lý ảnh hoặc các phần mềm máy tính. Chỉnh sửa ảnh nói riêng và các nội dung đa phương tiện (hình ảnh, video, âm thanh...) nói chung có thể làm phong phú cuộc sống hàng ngày của con người. Tuy nhiên, nó cũng có thể bị lợi dụng với những động cơ khác nhau, chẳng hạn như lý do chính trị, lợi ích tài chính bất hợp pháp... Với các phần mềm chỉnh sửa ảnh rất mạnh và phổ biến hiện nay như Photoshop, CorelDraw, GIMP... người ta có thể dễ dàng thực hiện việc giả mạo hoặc tổng hợp hình ảnh để tạo nên các bức ảnh giả mà gần như không thể nhận biết bằng mắt thường. Ví thể, xác định tính chân thực của ảnh là rất cần thiết, nhất là trong những lĩnh vực như nghiên cứu khoa học, luật pháp, thương mại điện tử, quân sự, GTVT... [4]. *Hình 2.1* là một ví dụ về việc ảnh biển số xe có thể bị thay đổi. Ví dụ, trong *Hình 2.1* người ta đã sử dụng một số kỹ thuật để làm thay đổi hình ảnh biển số xe. Rõ ràng, với những kỹ thuật như vậy, người ta có thể dễ dàng thay đổi các minh chứng trong các hồ sơ về bảo hiểm hay các vụ TNGT.



Ảnh thật Ảnh giả
Hình 1.1: Ảnh biển số xe giả mạo

Một bức ảnh bao gồm tập hợp các điểm ảnh, nên dữ liệu ảnh thường rất lớn và dễ bị tác động. Chẳng hạn, chỉ cần chỉnh ảnh hơi sáng lên một chút là giá trị tất cả các điểm ảnh sẽ bị thay đổi. Do đó, để đảm bảo tính hiệu quả và bền vững, trong rất nhiều ứng dụng xử lý ảnh, người ta cần phải tìm cách trích ra những đặc trưng của ảnh. Có một số hướng chính để xác định các đặc trưng, như qua phân tích ma trận điểm ảnh, biến đổi ảnh từ miền không

gian sang miền tần số, hay xác định các điểm cốt yếu (SIFT, SURF...) [5]. Trong bài báo này, chúng tôi nghiên cứu một phương pháp phân tích ma trận là SVD và áp dụng nó trong các ứng dụng về nén ảnh và xác thực ảnh.

2. PHÉP PHÂN TÍCH SVD

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 [6].

Xét một ma trận A có m hàng và n cột (với $n \leq m$), có hạng là r và $r \leq n \leq m$, phép biến đổi SVD phân tích ma trận A thành ba ma trận U, S, V như sau:

$$A = USV^T$$

Trong đó:

- U - Ma trận trực giao cấp m:

$$U = [u_1, u_2, u_3, \dots, u_r, u_{r+1}, \dots, u_m]$$

Với các vector u_i ($i = 1, 2, \dots, m$), tạo thành một tập trực chuẩn:

$$u_i^T u_j = \delta_{ij} = \begin{cases} 1, 2, \dots, i = j \\ 0, \dots, i \neq j \end{cases}$$

- V - Ma trận trực giao cấp n:

$$V = [v_1, v_2, v_3, \dots, v_r, v_{r+1}, \dots, v_n]$$

Với các vector v_i ($i = 1, 2, \dots, n$) tạo thành một tập trực chuẩn:

$$v_i^T v_j = \delta_{ij} = \begin{cases} 1, 2, \dots, i = j \\ 0, \dots, i \neq j \end{cases}$$

Như vậy, ma trận U là ma trận gồm các vector riêng trái của A, ma trận V là ma trận gồm các vector riêng phải của A.

S là ma trận đường chéo, với mỗi phần tử trên đường chéo là một trị riêng của A. Ký hiệu các giá trị trên đường chéo chính của S là δ_i , các giá trị này có độ lớn được sắp xếp theo thứ tự như sau:

$$\delta_1 \geq \delta_2 \geq \delta_3 \geq \dots \geq \delta_r > 0$$

Các giá trị còn lại δ_i ($i = r+1, \dots, n$) đều bằng 0. Khi áp dụng SVD vào ma trận điểm ảnh, "năng lượng" của ảnh được tập trung vào các phần tử ở hàng đầu, cột đầu của các ma trận U, V và các phần tử đầu tiên trên đường chéo của S. Các phần tử này có tính ổn định rất cao, vì vậy có thể sử dụng chúng để làm các đặc trưng bền vững, ứng dụng trong nhiều việc khác nhau.

3. MỘT SỐ ỨNG DỤNG CỦA SVD

3.1. Ứng dụng nén ảnh

Nén ảnh thực hiện được là do một thực tế: thông tin trong bức ảnh không phải là ngẫu nhiên mà có trật tự, có cấu trúc. Vì thế, nếu bóc tách được tính trật tự, cấu trúc đó thì sẽ biết phần thông tin nào quan trọng nhất trong bức ảnh để biểu diễn và truyền đi với số lượng ít bit hơn so với ảnh gốc mà vẫn đảm bảo tính đầy đủ cần thiết của thông tin. Ở bên nhận, quá trình giải mã sẽ tổ chức, sắp xếp lại được bức ảnh xấp xỉ gần chính xác so với ảnh gốc nhưng vẫn thỏa mãn chất lượng yêu cầu. Nén ảnh (nén mất thông tin) đạt được bằng cách loại bỏ các phần ít quan trọng trong ảnh đã được số hóa. Những phần tử "dư thừa"

này có thể là dư thừa thông tin về không gian, dư thừa về cấp xám hay dư thừa về thời gian.

Phân tích SVD với đặc tính xác định được các phần tử có sự tập trung "năng lượng" cao của ảnh nên có thể được sử dụng trong nén ảnh. Khi một hình ảnh được chuyển đổi bằng SVD, nó không bị nén. Tuy nhiên, do dữ liệu có dạng trong đó giá trị số ít đầu tiên có một lượng lớn thông tin hình ảnh, chúng ta chỉ có thể sử dụng một vài giá trị lớn nhất để thể hiện hình ảnh thì kết quả đạt được cũng sẽ không có nhiều khác biệt so với bản gốc [7], cụ thể:

$$A = USV^T = \sum_{i=1}^r \delta_i u_i v_i^T = \delta_1 u_1 v_1^T + \delta_2 u_2 v_2^T + \dots + \delta_r u_r v_r^T$$

Khi cần nén một ảnh A, chúng ta chỉ sử dụng một số ít, k phần tử (với $k \leq r$) và loại bỏ đi các phần tử còn lại. Ma trận chính của hạng k có được bằng cách trích xuất các hạng đó sau các chuyển đến hạng k đầu tiên và ảnh tương ứng với A_k có thể vẫn gắn với hình ảnh gốc:

$$A_k = \delta_1 u_1 v_1^T + \delta_2 u_2 v_2^T + \dots + \delta_k u_k v_k^T$$

Tổng dung lượng cho A_k sẽ là $k*(m+n+1)$. Để đo hiệu suất của phương pháp nén hình ảnh SVD, chúng ta có thể tính toán hệ số nén, có thể được tính bằng tỷ lệ nén $C_r = m*n/(k*(m+n+1))$. Như vậy, giá trị C_r càng lớn thể hiện độ nén càng cao và đồng thời chất lượng ảnh cũng sẽ kém hơn.

3.2. Ứng dụng phát hiện ảnh giả mạo

Trong phần này, chúng tôi trình bày về việc xác thực ảnh đã qua chỉnh sửa cắt-dán. Hầu hết các thuật toán phát hiện ảnh bị cắt-dán đều dựa trên phân tích đặc trưng các khối con của ảnh. Có nhiều hướng tiếp cận chính để xác định các thuộc tính riêng của mỗi khối, thường gồm các bước tương tự như lược đồ tương tự nhau. Dựa trên ý tưởng ban đầu của Kang và Wei [8], chúng tôi áp dụng phân tích SVD cho từng khối trên miền không gian ảnh, sau đó lấy một số giá trị đầu tiên trên đường chéo chính của ma trận đường chéo làm giá trị đặc trưng. Chúng tôi cũng so sánh với thuật toán phổ biến nhất sử dụng biến đổi DCT [9]. Thuật toán gồm các bước chính như sau:

Bước 1: Phân hoạch ảnh A thành các khối chờm nhau có kích thước 8x8.

Bước 2: Áp dụng phép phân tích SVD trên từng khối ảnh 8x8, sau đó lấy một số giá trị đầu tiên trên đường chéo của ma trận S làm các giá trị đặc trưng.

Bước 3: Đặc trưng của mỗi khối được sắp vào thành một hàng của một ma trận, sau đó các hàng được sắp xếp lại theo thứ tự từ điển.

Bước 4: Dựa trên các hàng giống nhau, xác định các khối giống nhau tương ứng.

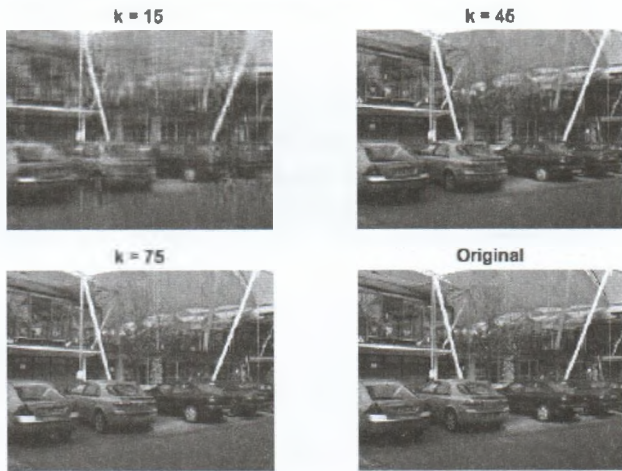
Bước 5: Nếu xác định được hai nhóm gồm một số đủ lớn các cặp khối tương ứng giống nhau thì có thể kết luận là ảnh đã bị chỉnh sửa.

4. KẾT QUẢ THỬ NGHIỆM

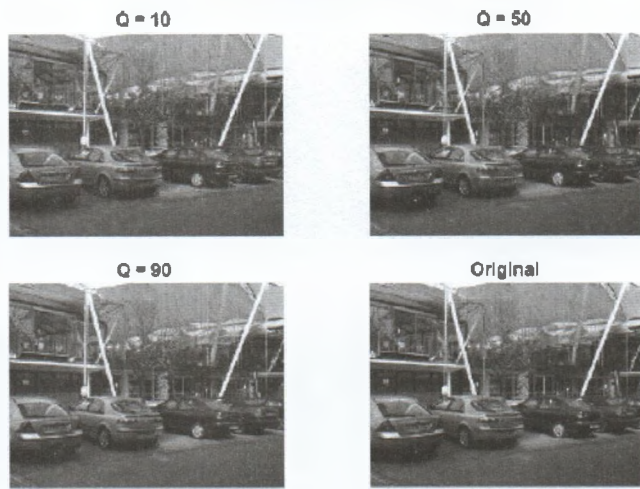
4.1. Nén ảnh

Để thử nghiệm hiệu quả của việc nén ảnh, chúng tôi chọn một ảnh gốc (original) là ảnh chưa được nén (Lena. bmp) có kích thước 512x512. Tiếp đó, chúng tôi thực hiện nén ảnh theo thuật toán trình bày ở mục 3.1 với các giá trị k khác nhau. Một số hình ảnh minh họa được thể hiện

trong Hình 4.1 và thống kê C_R với các giá trị k khác nhau thể hiện trong Bảng 4.1. Nhận thấy, khi thực hiện nén với k bằng 75 thì gần như không có sự khác biệt giữa ảnh nén và ảnh gốc, dù kích thước ảnh gốc lớn gấp 1,7 lần. Khi giảm giá trị k xuống 45 thì chất lượng ảnh vẫn chấp nhận được, dù kích thước ảnh nén giảm 3,7 lần so với ảnh gốc. Tuy nhiên, khi giảm giá trị k còn bằng 15 thì chất lượng ảnh sẽ khá thấp.



Hình 4.1: Một số hình ảnh nén sử dụng SVD



Hình 4.2: Một số hình ảnh nén sử dụng DCT

Nén ảnh cũng có thể được thực hiện thông qua phép biến đổi Cosin rời rạc (DCT - Discrete Cosine Transform). Hiện nay, DCT là thuật toán được sử dụng rộng rãi nhất trong một số định dạng hình ảnh và âm thanh, như JPEG và MP3. Do DCT biến đổi hình ảnh từ miền không gian sang miền tần số, nên có thể tách các phần của hình ảnh thành các khu vực tầm quan trọng khác nhau. Khi nén (ví dụ ảnh Lena) bằng thuật toán DCT với các mức độ nén khác nhau, các ảnh kết quả được thể hiện trong Hình 4.2 và thống kê C_R với các mức độ nén Q khác nhau thể hiện ở trong Bảng 4.2. Chúng ta có thể thấy, khi Q bằng 50 (nén ở mức 50%) thì chất lượng ảnh vẫn chấp nhận được, dù kích thước ảnh nén chỉ bằng gần 1/9 kích thước của ảnh gốc.

Với cùng mức độ chất lượng ảnh, thuật toán dựa trên DCT có khả năng nén cao hơn nhiều so với sử dụng SVD.

Tuy nhiên, do đặc tính nén theo từng khối 8×8 , nên khi nén ảnh JPEG (sử dụng DCT) với mức độ nén cao thì ảnh có hiện tượng bị chia khối (blocking artifact).

Bảng 4.1. Thống kê khả năng nén ảnh bằng SVD với các giá trị k khác nhau

k	5	15	25	45	55	65	75
C_R	25,6	8,5	5,1	3,7	2,8	2,3	1,7

Bảng 4.2. Thống kê khả năng nén ảnh bằng DCT với các giá trị Q khác nhau

Q	10	20	30	40	50	60	70
C_R	23,3	15,7	12,4	10,7	9,4	8,3	7,1

4.2. Phát hiện ảnh giả mạo

Để chuẩn bị dữ liệu thử nghiệm cho khả năng phát hiện ảnh giả mạo cũng như sự bền vững trước các phép tấn công, trước hết chúng tôi chọn ngẫu nhiên một tập gồm 1.000 ảnh, với kích thước mỗi ảnh là 512×318 hoặc 318×512 từ 1.338 ảnh trong cơ sở dữ liệu UCID (Uncompressed Colour Image Database) của Schaefer và Stich [10]. Đây là một cơ sở dữ liệu ảnh chuẩn, có thể tải về miễn phí và được sử dụng nhiều trong các nghiên cứu về xử lý ảnh.

Để tạo tập ảnh bị chỉnh sửa, chúng tôi lấy các ảnh gốc, thực hiện cắt - dán ngẫu nhiên theo các vùng hình vuông có kích thước là 64 và 32 trên ảnh. Chúng tôi áp dụng áp dụng các thuật toán phát hiện giả mạo dựa trên SVD và DCT trên tập ảnh giả và kết quả nhận được cho thấy cả hai thuật toán đều có khả năng phát hiện giả mạo là 100% (Bảng 4.3). Khi áp dụng các thuật toán trên vào tập ảnh thật, vẫn có một số trường hợp xác định là ảnh giả (phát hiện nhầm). Trên Bảng 4.3 cũng thể hiện tỷ lệ phát hiện nhầm của các thuật toán sử dụng SVD và DCT lần lượt là 7,3% và 6,1%.

Để che giấu những dấu hiệu chỉnh sửa nhằm gây khó khăn cho các thuật toán phát hiện, ảnh sau khi sửa có thể được hậu xử lý bằng các thao tác xử lý ảnh (gọi là các phép tấn công) như nén JPEG, thêm nhiễu, làm mờ... Một thuật toán được coi là bền vững trước một phép tấn công, nếu nó có thể phát hiện việc chỉnh sửa ngay cả khi ảnh chỉnh sửa đã bị tấn công.

Bảng 4.3. Tỷ lệ phát hiện ảnh đã qua chỉnh sửa [%]

	Ảnh giả	Ảnh thật
SVD	100	7,3
DCT	100	6,1

Bảng 4.4. Tỷ lệ phát hiện ảnh chỉnh sửa sau khi thêm nhiễu Gauss [%]

	Gauss-40	Gauss-30	Gauss-25
SVD	34,2	0	0
DCT	97,5	92,4	91,6

Bảng 4.5. Tỷ lệ phát hiện ảnh chỉnh sửa sau khi nén JPEG [%]

	Jpeg-50	Jpeg-40	Jpeg-10
SVD	59,5	53,2	39,7
DCT	96,8	96,5	85,4

Để đánh giá độ bền vững của các phương pháp phát hiện, từ bộ ảnh giả, chúng tôi tạo ra các bộ ảnh bị tấn công. Chúng tôi lấy các ảnh từ tập ảnh giả và thêm nhiễu Gauss với các cường độ khác nhau là 40 decibel (dB), 30 dB và 25 dB để tạo ra các tập ảnh bị tấn công, ký hiệu là Gauss-40, Gauss-30 và Gauss-25. Tương tự, chúng tôi lấy các ảnh từ tập ảnh giả và nén JPEG với chỉ số chất lượng nén khác nhau (50%, 40% và 10%) để tạo ra các tập ảnh bị tấn công, ký hiệu là Jpeg-50, Jpeg-40 và Jpeg-10. Kết quả thử nghiệm được thể hiện trên *Bảng 4.4* và *Bảng 4.5*. Kết quả này cho thấy, thuật toán sử dụng SVD không bền vững trước các phép tấn công thêm nhiễu và nén JPEG.

5. KẾT LUẬN

Bài báo đã trình bày về phép biến đổi SVD và một số ứng dụng của nó trong xử lý ảnh và trong GTVT. Trong ứng dụng thứ nhất, chúng tôi áp dụng SVD để nén ảnh. Các thử nghiệm với các hệ số khác nhau, đồng thời so sánh với nén ảnh JPEG cho thấy, tuy không bằng JPEG khi nén ở các mức độ cao, nhưng SVD cũng nén ảnh rất tốt, đồng thời không bị hiệu ứng chia khối khi nén ở mức cao. Ứng dụng thứ hai là về phát hiện ảnh giả mạo. Kết quả thử nghiệm trên nhiều ảnh cho thấy, thuật toán dùng SVD có thể hoạt động tốt khi không có tấn công. Trong trường hợp có tấn công bằng nhiễu Gauss hoặc nén JPEG thì phương pháp dùng DCT bền vững hơn.

Lời cảm ơn: Nghiên cứu này được tài trợ bởi Trường Đại học GTVT trong Đề tài mã số T2020-CN-003.

Tài liệu tham khảo

- [1]. G. Zhang, J. Wang, C. Yan and S. Wang (2019), *Application research of image compression and wireless network traffic video streaming*, J. Vis. Commun. Image Represent., vol.59, pp.168-175.
- [2]. G. Mahantesh, M. Vedanth, K. G. Pooja and D. Kulala (2020), *Traffic Density Estimation in Traffic System Using Image Processing*, no.1, pp.1-5.
- [3]. Y. Wiseman (2017), *Real-time monitoring of traffic congestions*, IEEE Int. Conf. Electro Inf. Technol., pp.501-505.
- [4]. A. Piva (2013), *An Overview on Image Forensics*, ISRN Signal Process., vol.2013, pp.1-22.
- [5]. G. Kumar and P. K. Bhatia (2014), *A detailed review of feature extraction in image processing systems*, Int. Conf. Adv. Comput. Commun. Technol. ACCT, pp.5-12.
- [6]. J. Hourigan and L. Mcindoo (2014), *The Singular Value Decomposition*, Introd. to Gr. Penetrating Radar Inverse Scatt. Data Process., vol.9781118305, pp.229-241.
- [7]. L. Cao (2006), *Singular Value Decomposition Applied To Digital Image Processing*, pp.1-15.

[8]. X. Kang and S. Wei (2008), *Identifying Tampered Regions Using Singular Value Decomposition in Digital Image Forensics*, 2008 Int. Conf. Comput. Sci. Softw. Eng., vol.3, pp.926-930.

[9]. J. Fridrich, D. Soukal and J. Lukas (Dec. 2003), *Detection of Copy-Move Forgery in Digital Images*, Proc. Digit. Forensic Res. Work.

[10]. G. Schaefer and M. Stich (2004), *UCID - An Uncompressed Colour Image Database*, Proc. SPIE, Storage Retr. Methods Appl. Multimed.

Ngày nhận bài: 11/4/2021

Ngày chấp nhận đăng: 25/5/2021

**Người phản biện: TS. Bùi Ngọc Dũng
TS. Lại Mạnh Dũng**