International Journal of Mathematics and
Computer Science, **18**(2023), no. 1, 85–97

Principal Component Analysis for Zero Watermarking Technique

Saja Abdulameer Kahdim, Areej M. Abduldaim

Mathematics and Computer Applications
Applied Sciences Department
University of Technology
Baghdad 10001, Iraq

email: as.20.26@grad.uotechnology.edu.iq,
Areej.M.Abduldaim@uotechnology.edu.iq

(Received September 3, 2022, Accepted October 5, 2022)

Abstract

The power of mathematics in computer science appears in each step of designing algorithms or analyzing the results, not to mention all the other details. In this paper, the idea of the importance and effectiveness of mathematics, particularly linear algebra, is highlighted. To build a zero watermarking algorithm, the principal component analysis (PCA), considered one of the known decomposition methods, is used uniquely (that is, without any transform or other decomposition method) for extracting the crucial information from the concealment images. The watermark is XORed with the features chosen from this crucial information to generate the secret share. Using PCA only, the results were very satisfactory and the proposed algorithm proved its resistance to different attacks through the robustness metric NC.

1 Introduction

Linear algebra methods have drawn the attention of many researchers and appeared in many of their works in computer science. This is due to the characteristics of these algebraic methods in the ability to know the important

Key words and phrases: Principal Component Analysis, Zero Watermarking, Eigenvalues and Eigenvectors.
AMS (MOS) Subject Classifications: 15A18, 94A08.
ISSN 1814-0432, 2023, http://ijmcs.future-in-tech.net

locations within the data to be processed. In general, the strength of most of these methods is their stability as these methods are mostly based on the extraction of eigenvalues and eigenvectors. In particular, signal processing took the largest share using these method and, more accurately, image processing benefited from the advantages of these methods in different topics, including watermark techniques. Depending on the procedure used to embed the watermark into the concealment image, these techniques are divided into two. The watermarking technique means that the watermark image will be embedded into the image pixels, changing the original image quality while in the zero watermarking technique the crucial information is extracted to be used in the embedding steps keeping the original quality of the image. The famous algebraic decomposition frequently used is the singular value decomposition (SVD).

The unique singular values found using the SVD gives this algebraic method particular importance. In addition, the Hessenberg Decomposition (Hess) started being used recently. For watermarking technique, Pan et al. [1] used both previous algebraic methods to propose a steganography approach of double matrix decomposition image with multiple region coverage for solving the issue of inadequate extraction capability of the stego-images under interferences or attacks. Otherwise, with optimization algorithms suggested by Nazir et al. [2]. For zero watermarking techniques, they are also effective for medical images using double-tree complex wavelet transformed multi-level discrete cosine transform (MDCT) [3]. The Hess decomposition method is often optimized without any transformation in the YCbCr space for more robustness and security in [4]. Otherwise, without optimization in the same space [5].

There are, of course, endless possibilities to use various algebraic methods with many mathematical tools such as transformations, optimization algorithms, among others [6, 7]. The Discrete Wavelet Transforms (DWT) and the Lift Wavelet Transforms (LWT) have been widely used in both watermarking and zero watermarking techniques, in addition to the Discrete Cosine Transform (DCT), to improve the quality of the results regarding the security and the invisibility of the hidden watermark. Alzahrani [8] enhanced the invisibility and the robustness of digital image watermarking based on DWT-SVD while and improved upon the SVD-based color image watermarking by incorporating integer wavelet transform (IWT) and chaotic maps [9]. In [10, 11], the authors used first the SVD with a genetic algorithm (GA) without using any other popular transform and later designed an optimized image watermarking technique depending on the Generalized Singular Value

Decomposition (GSVD) method and Binary Gravitational Search Algorithm (BGSA).

In general, the Principal Component Analysis (PCA) algebraic decomposition is not commonly used in watermarking techniques, especially in zero watermarking. Video and image watermarking techniques have been studied by several authors like Anjaneyulu [12] who used hybrid DWT-PCA to protect the copyright of images with the Ant Colony Optimization Technique. Also, PCA can be utilized for identifying the suitable bands that are utilized for watermark hiding to design a reversible and blind technique of digital watermarking [13]. In 2022 for COVID-19, a PCA has been computed between the cover and mark image for developing a scheme of COVID-19 data embedding that is capable of guaranteeing the copyright protection and security of the medical images [14]. The only zero watermarking technique that utilized PCA with DWT to build a robust image zero-watermarking algorithm is found in [15].

On the other hand, many ideas were put forward and various other algebraic decomposition methods were used. Abduldaim, Abdulrahman, and Tahir [16] developed a watermarking technique combining the QR algebraic decomposition method and new discrete wavelets derived from Hermite polynomials for the obtained Discrete Hermite Wavelet Transformation (DHWT). The algebraic matrix decomposition method LU in [17] was employed successfully with Mamdani fuzzy inference system (FIS) on the original image to take out a controller factor for a trade-off between imperceptibility and robustness. Wu et al. [18] presented a scheme of medical image zero-watermarking depending on color visual cryptography and Bi-Dimensional Empiric Mode Decomposition (BEMD). This presented scheme efficiently resolves the issue of inadequate robustness concerning the conventional watermarking schemes under extensive attacks.

In this work, the PCA is adopted without any other mathematical tool for the first time in the zero watermarking scheme. This is to accentuate the mathematical power (algebraic decomposition methods) in image processing, particularly in zero watermarking schemes.

The organization of this paper as follows. In section 2, we recall the background about the Principle Component Analysis decomposition method (PCA). In section 3, we give the suggested zero watermarking techniques as embedding and restoring steps of the included watermark. In section 4, we illustrate the results of the experiments and the measurement values of the robustness and imperceptibility are calculated to judge the proposed technique. In section 5, we compare the proposed technique with another based on PCA. In section 6, we end our paper with a conclusion.

2 Principle Component Analysis (PCA)

Principal component analysis (PCA) is the procedure of finding values named by the principal components and changing the basis using them on the data. The first few principal components are chosen for their effectiveness and ignoring the remaining ones. The principal components represent the eigenvectors of the extracted covariance matrix. Consequently, the principal components are often computed by eigendecomposition of the given covariance matrix or singular value decomposition of the data matrix.

Step 1: Let I be a data set ($n \times n$ Image matrix)

$$I = \begin{bmatrix} C_1 & C_2 & \dots & C_n \\ \downarrow & \downarrow & \dots & \downarrow \\ X_{11} & X_{12} & \dots & X_{1n} \\ X_{21} & X_{22} & \dots & X_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ X_{n1} & X_{n2} & \dots & X_{nn} \end{bmatrix}$$

where C_j's represent the columns of the matrix I, $j = 1, 2, \cdots, n$.

Step 2: Compute the mean of each column:

$$\bar{C_j} = \frac{\Sigma_{i=1}^{n} X_{ij}}{n} \tag{2.1}$$

where $j = 1, 2, ..., n$.

Step 3: Find the Covariance Matrix: Calculate the covariance between any two columns C_j and C_{j+1} to obtain the covariance matrix of the matrix I ($cov(I)$):

$$cov(C_j, C_{j+i}) = \frac{\Sigma_{i=1}^{n}(X_{ij} - \bar{C_j})(X_{i(j+i)} - \bar{C}_{j+i})}{(n-1)} \tag{2.2}$$

$$var(C_j, C_j) = \frac{\Sigma_{i=1}^{n}(X_{ij} - \bar{C_j})(X_{ij} - \bar{C_j})}{(n-1)} \tag{2.3}$$

where $j = 1, 2, ..., n$.

Then the covariance matrix will be:

$$cov(I) = \begin{bmatrix} var(C_1,C_1) & cov(C_1,C_2) & \dots & cov(C_1,C_n) \\ cov(C_2,C_1) & var(C_2,C_2) & \dots & cov(C_2,C_n) \\ \vdots & \vdots & \ddots & \vdots \\ cov(C_n,C_1) & cov(C_n,C_2) & \dots & var(C_n,C_n) \end{bmatrix}$$

Step 4: Find the Eigenvalues $\lambda_1 > \lambda_2 > ... > \lambda_n$ and the Eigenvectors $U_j = [u_{1j}, u_{2j}, ...u_{nj}]$, $j = 1, 2, \ldots, n$ of $cov(I)$. We emphasize. that the eigenvectors are unit eigenvectors as it is very important for PCA.

Step 5: Construct a new matrix of eigenvectors: Arrange the eigenvectors by eigenvalues, maximum to minimum. This provides the components according to importance.

$$V = \begin{bmatrix} \lambda_1 & \lambda_2 & ... & \lambda_n \\ \downarrow & \downarrow & ... & \downarrow \\ U_1 & U_2 & ... & U_n \\ \downarrow & \downarrow & ... & \downarrow \\ u_{11} & u_{12} & ... & u_{1n} \\ u_{21} & u_{22} & ... & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ u_{n1} & u_{n2} & ... & u_{nn} \end{bmatrix}$$

Step 6: Deriving the feature matrix: The components with lesser significance are truncated. Choose the eigenvectors that we want to keep from the list of eigenvectors. Finally, form a PC matrix with these eigenvectors in the columns.

$$PC = \begin{bmatrix} u_{11} & u_{12} & ... & u_{1k} \\ u_{21} & u_{22} & ... & u_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ u_{n1} & u_{n2} & ... & u_{nk} \end{bmatrix}$$

where $k \leq n$.

3 The Suggested Technique Based on PCA

In this work, all images used are of size 256×256. The concealment images and the watermark are converted from RGB space to grayscale space. Each host image is divided into 4×4 non-overlapping blocks. The PCA is applied to each block as an algebraic transform to get the important information of the concealment images and then generate the features matrix bits.

3.1 Features Extraction and Embedding Process

The PCA is applied to decompose each block of size 4×4 of the concealment image to obtain the eigenvectors matrices V_i. Hence, choose the first column from each V_i matrix to be principal components PC_i to generate the features matrix and then the secret share. These details are explained in Figure 1. Following are the steps of the embedding process:

Step 1: Insert the $n \times n$ concealment image and switch to grayscale.

Step 2: Partition the resulting image into 4×4 non-overlapping blocks.

Step 3: Apply the PCA method on each block in step 2 to configure the matrices V_i.

Step 4: Choose the first column from each V_i to be the PC_i to form the feature Matrix F.

Step 5: Convert the matrix F to a binary matrix B as follows:

$$\begin{array}{ll} B(i,j) = 1 & \text{if } F(i,j) \geq F(i,j+1) \\ B(i,j) = 0 & \text{otherwise} \end{array} \tag{3.4}$$

where $1 \leq i \leq 256$ and $1 \leq j \leq 255$

Step 6: Insert and switch the watermark image W to binary to obtain the matrix BW.

Step 7: XORing B and BW matrices to generate the secret share K.

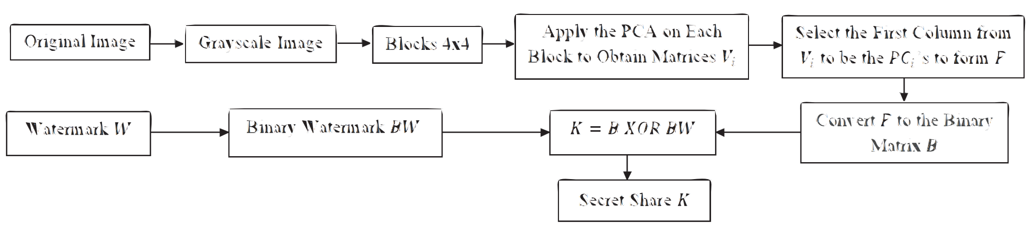


Figure 1: The Diagram of Embedding Process

3.2 Restoring Process

The restoration of the included watermark is illustrated in Figure 2. The exhaustive steps are given as follows:

Step 1: Insert the $n \times n$ the concealment image and switch to grayscale.

Step 2: Partition the resulting image into 4×4 non-overlapping blocks.

Step 3: Apply the PCA method on each block in step 2 to configure the matrices V_i.

Step 4: Choose the first column from each V_i to be the PC_i to form the feature Matrix F.

Step 5: Convert the matrix F to a binary matrix B as follows:

$$\begin{array}{ll} B(i,j) = 1 & \text{if } F(i,j) \geq F(i,j+1) \\ B(i,j) = 0 & \text{otherwise} \end{array} \tag{3.5}$$

where $1 \leq i \leq 256$ and $1 \leq j \leq 255$

Step 6: Input the secret share K and apply XOR logical operation between B and K to get the watermark.

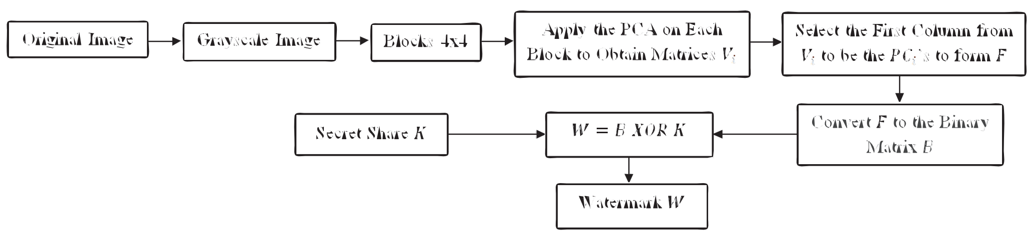


Figure 2: The Diagram of Restoring Process

4 The Experiential Results of Performing PCA

In this section, the results before and after putting images under attacks are given and explained in three cases: the thoroughness of the suggested technique, the impact and the importance of PCA, and the results of performing the PCA after attacks.

4.1 The Thoroughness of the Suggested Technique

Some experiments are performed, before exposing the images to any attack, to measure the robustness and imperceptibility of the suggested technique. The proposed zero watermarking scheme is examined on four grayscale images of size 256 × 256. A binary image of size 256 × 256 is used as the watermark image. Figure 3 shows the concealment images and the watermark used in this work.

Lenna Image Baboon Image Girl Image Peppers Image Watermark

Figure 3: The Concealment Images and the Watermark Used in this Work

From a mathematical point of view, to check the proposed technique and to locate whether the final results are reasonable or not the NC measurement is used (the realistic value of NC=1). Before attacking the concealment images, the NC=1 after performing the proposed techniques. See Table 1.

Table 1: NC Values with No Attacks

images	Lena	Girl	Baboon	Peppers
NC	1	1	1	1

4.2 The Impact and the Importance of PCA

The main aim is to elicit the feature matrix using the algebraic method PCA to hide the binary watermark in such a way that the same included watermark is attained exactly after the process of restoring. The proposed embedding technique used the PCA only, without any public transformation, on each block after dividing the concealment image into 4×4 non-overlapping blocks to decompose each block into high frequencies (the remarkable information of the image) represented by the matrices PC_i's consisting of the eigenvectors which correspond to the highest eigenvalues while lower frequencies (the least important information in the image) are represented by the eigenvectors which correspond to the lowermost eigenvalues.

The algebraic decomposition method works almost as a transformation because both transform the image into high frequencies and low frequencies. Consequently, the embedded watermark can be extracted exactly the same as the original watermark image depending on the strength and stability properties of the eigenvectors and the eigenvalues. Moreover, the process of restoring the watermark from the concealment images that were attacked gave good and passable results.

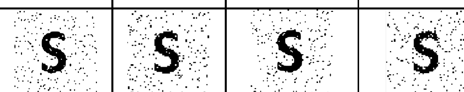
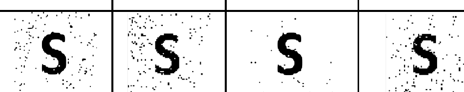
4.3 The Results of Performing PCA after Attacks

Testing the imperceptibility and robustness of the proposed scheme without utilizing any common transform shows that the results of the values of NC and PSNR after restoring the watermark from the attacked images are acceptable and reasonable as shown Table 2. On the other hand, Table 2 gives the restored watermark image after the concealment images are attacked.

Table 2: PSNR and NC Values after Attacks

Attacks	**Lenna**		**Girl**		**Baboon**		**Peppers**	
	PSNR	**NC**	**PSNR**	**NC**	**PSNR**	**NC**	**PSNR**	**NC**
Salt and Pepper Noise	26.761	0.970	60.715	0.972	27.187	0.972	27.019	0.976
JPEG Compression	60.708	0.981	60.715	0.981	60.709	0.995	60.143	0.984
Hist. Equalization	19.074	0.920	10.120	0.930	16.665	0.945	20.584	0.911
Gaussian low pass filter	28.073	834	32.191	0.845	25.186	0.829	28.265	0.862
Motion Filter	22.058	0.726	24.454	0.713	21.260	0.699	21.095	0.722
Median Filter	26.544	0.829	30.496	0.842	24.599	0.828	26.568	0.857
Gaussian noise	37.649	0.834	37.665	0.814	37.665	0.926	37.685	0.842
Poisson noise	27.256	0.762	30.428	0.758	26.997	0.824	27.287	0.764
Average filter	28.072	0.834	32.193	0.845	25.186	0.829	28.265	0.862

Table 3: The Restoration of the Watermark after Attacks

Attacks	**Lena**	**Girl**	**Baboon**	**Peppers**
Salt and Pepper				
Restored Watermark				
JPEG Compression				
R.W.				
Gaussian Noise				
R.W.				
Histogram Equalization				
R.W.				
Motion Filter				
R.W.				
Average Filter				
R.W.				
Med Filter				
R.W.				

5 Comparison

In this work, the PCA has been applied to transform the concealment images algebraically into a frequency domain to create the master secret without any other transform. The results are more than acceptable depending on the values obtained. To our knowledge, the only study with this interest in PCA is [15], but the Discrete Wavelet Transform (DWT) is used with the algebraic decomposition PCA under BER metric (the number of deferent bits between two images). The differences are explained in Table 4:

Table 4: Comparison Between the Proposed Technique and [15] Technique under BER Metric

Attacks	Lena Image	
	The Proposed Technique	[15] Technique
Salt pepper noise	0.04	0.04
JPEG compression	0.0336	0.0000
Gaussian noise	0.28	0.07
Median filter	0.2272	0.0049

6 Conclusion

In this paper, we proposed an effective technique of zero-watermarking. The algebraic Principal Component Analysis decomposition method (PCA) aws adopted which has not been used in zero watermarking techniques before withoutan auxiliary transform. The proposed technique relied on PCA without using any transform to extract the features matrix from the original color images (with size 256x256) for creating the master secret, which is being kept by a third party, in order to protect the rights of property from tampering. The PCA utilization was effective with the significance and strength of the results being so close (if not better) to the optimal results achieved from using a popular transformation of extracting the image features optimally. The proposed technique gave good flexibility against different types of common attacks.

References

[1] P. Pan, Z. Wu, C. Yang, B. Zhao, Double-Matrix Decomposition Image Steganography Scheme Based on Wavelet Transform with Multi-Region Coverage, Entropy, 24, **246,** (2022), 1–21.

[2] H. Nazir, I. S. Bajwa, M. Samiullah, W. Anwar, M. Moosa, Robust Secure Color Image Watermarking Using 4D Hyperchaotic System, DWT, HbD, and SVD Based on Improved FOA Algorithm, Security and Communication Networks,(2021).

[3] T. Huang, Jia Xu, Y. Yang, B. Han, Robust Zero-Watermarking Algorithm for Medical Images Using Double-Tree Complex Wavelet Transform and Hessenberg Decomposition, Mathematics, **10,** (2022), 1154.

[4] N. S. Mohammed, A. M. Abduldaim, Algebraic Hessenberg Decomposition Method Optimized by Genetic Algorithm for Zero Watermarking Technique, Int. J. Math. Comput. Sci., **16,** no. 4, (2021), 1497–1514.

[5] N. S. Mohammed, A. M. Abduldaim, Algebraic Decomposition Method for Zero Watermarking Technique in YCbCr Space, Engineering and Technology Journal,**4,**(2022), 605–616.

[6] H. B. Abdul Wahab, A. J. Abdul-Hossen, S. A. Kadhom, Encrypted Image Watermark in Audio Files Using Homogenous Deffie-Hellman with Chebyshev Polynomial, Engg. and Tech. Journal, **34,** (2016).

[7] L. F. Jalil, M. M. Laftah, Text Hiding in 3D Object, Engg. Tech. J., **34,** (2016).

[8] W. H. Alshoura, Z. Zainol, J. M. Alawida, An FPP-resistant SVD-based image watermarking scheme based on chaotic control, Alexandria Engineering Journal, (2022), 61.

[9] A. Alzahrani, Enhanced Invisibility and Robustness of Digital Image Watermarking Based on DWT-SVD, Applied Bionics and Biomechanics, (2022).

[10] N. S. Mohammed, A. M. Abduldaim, Algebraic Decomposition Method Utilized in Optimized Zero Watermarking Technique, First Babylon International Conference on Information Technology and Science, (2021).

[11] A. M. Abduldaim, A. K. Faraj, Combining Algebraic GSVD and Gravitational Search Algorithm for Optimizing Secret Image Watermark Sharing Technique, Int. J. Math. Comput. Sci.,**17,** no. 2, (2022), 753–774.

[12] S. Anjaneyulu, T. Ramashri, Hybrid Discrete Wavelet Transform - Methods of Principal Components Based Video Watermarking For Copyright Protection Using Ant Colony Optimization Techniques, Global J. Pure Appl. Math., **13,** no. 2,(2017), 795–815.

[13] R. Thomas, M. Sucharitha, Reversible Color Image Watermarking Scheme Based on Contourlet Transform and Principal Component Analysis, International Conference on Electrical Energy Systems, (2021), 641–644.

[14] O. P. Singh, A. K. Singh, A. K. Agrawal, H. Zhou, SecDH, Security of COVID-19 images based on data hiding with PCA, Computer Communications, **191,**(2022), 368–377.

[15] X. Leng, J. Xiao, Y. Wang, A Robust Image Zero-Watermarking Algorithm Based on DWT and PCA, Communications in Computer and Information Science, Springer, Berlin, Heidelberg, 2012, 289.

[16] A. M. Abduldaim, A. A. Abdulrahman, F. S. Tahir, The effectiveness of discrete Hermite wavelet filters technique in digital image watermarking, Indonesian Journal of Electrical Engineering and Computer Science, **25,** no. 3, (2022), 1392–1399.

[17] R. I. Sabri, A. M. Abduldaim, J. Waleed, Mamdani FIS Combined with LU Decomposition Method and Two-Level LWT for Image Watermarking Technique, 3rd International Conference on Engineering Technology and its Applications, IICETA-IEEE, (2020), 12–17.

[18] Deyang Wu et al., Color Zero-Watermarking Algorithm for Medical Images Based on BEMD-Schur Decomposition and Color Visual Cryptography, Security and Communication Networks, (2021), Article ID 7081194.

[19] S. P. Mishra, U. Sarkar, S. Taraphder, S. Datta, D. P. Swain1, R. Saikhom, S. Panda, M. Laishram, Multivariate Statistical Data Analysis-Principal Component Analysis (PCA), International Journal of Livestock Research, **7** no. 5,(2017), 60–78.