

Implementation Of A Hybrid Color Image Compression Technique Using Principal Component Analysis And Discrete Tchebichef Transform

S. Elamparuthi¹, N. Puviarasan²

¹PG Department of Computer Science Government Arts College, Chidambaram, Tamilnadu, India

²Professor, Division of Computer and Information Science Annamalai University, Annamalai Nagar, Tamilnadu, India

¹elam.dod@gmail.com, ²npuvin2410@yahoo.in

Article History: Received: 10 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 28 April 2021

Abstract—The proliferation of scientific and technological demands of high-resolution multimedia content has extremely increased the data volume in enterprise data centers and servers on the internet. The characterization of high-resolution pictures creates a great challenge to transfer files with the immensely colossal volume of image data over communication networks. The uploading/downloading time of large images has always been a key problem on the Internet. In addition to data communication issues, high-resolution photo consumes larger storage capacity. Therefore, compression is an almost inevitable process towards the reduction of the transmission time and/or storage capacity requirements of images. Principle component analysis (PCA) and Discrete Tchebichef Transform (DTT) algorithms are often employed for image compression in several references. In this paper, we propose a hybrid color image compression approach based on PCA and DTT algorithms (PCADTT), which integrates the benefits of both PCA and DTT algorithms. Our hybrid approach exploits (i) PCA to reduce the dimensionality of the image; and (ii) DTT algorithm to enhance the image quality. The proposed technique has been assessed and related to a compression method that integrates DTT with a singular value decomposition (SVD) scheme (DTTSVD) using different performance measures including compression time (CT), compression ratio (CR), structural similarity index measure (SSIM), peak signal-to-noise ratio (PSNR), and universal quality index (UQI). The experimental results reveal that our proposed method outperforms the existing method for all kinds of image content at a high compression ratio with lower computational complexity and retaining the quality of image data.

Keywords— compression ratio; discrete tchebichef transform; hybrid method; image compression; principal component analysis

1. Introduction

The amount of data produced and communicated by public service industries, nonprofit sectors, business organizations, and scientific research, has augmented immeasurably [1]. Each day our cyber world generates approximately 2.5 quintillion bytes of data (i.e. 1 quintillion byte = 1 billion gigabytes) [2]. International Data Corporation expects that our global datasphere—the digital data we generate, capture, imitate, and consume—will increase from around 40 zettabytes of data in 2019 to 175 zettabytes in 2025 (1 zettabyte = 1 trillion gigabytes) [3]. With this irresistible quantity of intricate and a multiplicity of data pouring from any device, any-time, and anywhere, there is undeniably an era of *Big Data* – a phenomenon also called the *Data Deluge*. These data comprise textual content (i.e. unstructured, semi-structured, and structured), to multimedia content (e.g. audio, images, and videos) on heterogeneous platforms such as sensors networks, social media websites, machine-to-machine communications, the internet of things, and numerous safety-critical cipher physical systems [4].

In this advanced cyber world with the data deluge, transferring a plethora of multimedia products at every second, contributing to circumscribed bandwidth consumption and storage space. Image compression is the preeminent way to create applications to characterize a picture with smaller number of bits without degrading its quality. This can be realized by removing irrelevancies and redundancies that present in the images. It converts the actual image to its compressed form by the recognition and application of patterns present in the image data. These applications play a pivotal role in several scientific domains such as communication, medical imaging, astrophysics, satellite, videoconferencing, etc. [5, 6].

Generally, image compression approaches are categorized into two distinct types, which are lossless and lossy methods [7]. Lossless methods compress images without loss of essential data but have low CR [8]. In lossless compression, complete data fidelity is assured after reconstruction, but overall CR is restricted between 2:1 and 3:1. These techniques hold the data integrity throughout the entire compression and decompression has been employed as an essential process in the application of satellite communications [9], remote sensing [10, 11],

medical imaging [12], etc. Lossy compression methods can compress images with great CRat the cost of information integrity [13]. In lossy compression, the reconstructed (compressed) image comprises some deprivation related to the actual one but it is nearly close to it. This method has been generally employed in mobile devices, World Wide Web, digital cameras, etc. [14].

Primarily, digital images are pigeonholed by three different redundancies: coding redundancy, spatial redundancy, and psychovisual redundancy [15]. The compression techniques utilize these redundancies to reduce the dimensionality of pictures [16]. The coding redundancy also known as statistical redundancy refers to the utilization of adjustable length code to compare the statistics of the input picture. Spatial redundancy reveals a fact that the gray value of one pixel may be partially measured by values of other pixels. Psychovisual redundancy relies on the human insight of the image data [17]. Eliminating psychovisual redundancy includes noise in reconstructed pictures; therefore, this process is circumvented in the lossless compression method.

There are several methods have been proposed to compress the image as well as to enhance its quality. The transform domain-based compression is the most extensively employed approach for lossy image compression that converts the pixels from one domain into another domain to generate some coefficients, in which the characterization is very normal and consequently more compressed image. This characterization enables a set of coefficients to provide the majority of the energy in the pictures whereas others are probably infinitesimal or zero. It employs a revocable and linear transform to relate the picture into coefficients which are quantized and compressed subsequently. An ideal transformation relates the utmost information into only a few coefficients. In the quantization process, the coefficients that hold the minimum information are discarded. The transform coding decomposes an $N \times N$ size of input picture into many $n \times n$ nonoverlapping blocks (i.e., sub-images). Then, each block is individually transformed to produce $(N/n)^2$ arrays with the size of $n \times n$. Generally, the correlation among the neighboring pixels in an image is maximum. Transform compression method exploits this correlation to realize a higher compression ratio. It realizes greater efficiency due to three aspects as given below:

1. This type of compression method is a block-based method in which a block of image data is processed instead of only one element.
2. The quantization leads to eliminating the correlation between the pixels of each block.
3. Only selected transformed coefficients are quantized and transferred to the destination side; accordingly, greater compression ratios can be obtained.

The application of the transformation-based compression method poses more challenges than the deployment of the predictive compression method and therefore it is not desirable for applications that demand less complexity and cost [18]. There are numerous transformation techniques found in the literature which are employed in image compression including Karhunen-Loeve Transform (KLT) (also called PCA) [19], Discrete Fourier Transform (DFT) [20], Discrete Cosine Transform (DCT) [21], Singular Value Decomposition (SVD) [22], DTT [23], Set Partitioning in Hierarchical Trees (SPIHT) [25], Discrete Wavelet Transform (DWT) [24], and so on.

The rudimentary idea of the hybrid compression approaches is to integrate spatial domain with transform domain methods to enhance the compression performance as well as the quality of reconstructed images with reduced noise factor. One more option is implementing two transformation methods on given images. The integration of two methods enable improved representations of given pictures, improved handling of curved shapes, and better-quality compressed images. The overall benefits of the hybrid methods are increasing the visual quality and reducing noisedata and artifacts in output images.

In this paper, we propose a new hybrid color image compression method that integrates PCA and DTT algorithms. Even though PCA provides better CR, the quality of the reconstructed image is poor. Hence, the DTT algorithm is used with PCA to improve the quality of the output image as well as to achieve a more compact image. First, the input image is divided into blocks using PCA. Each sub-image is employed as a sample vector. PCA chooses a covariance matrix of bigger singular values equivalent to the Eigenvector to perform image compression. Then the dimension of the image is reduced further by means of the DTT technique.

The systematic organization of the manuscript is given as follows: In Section II, we summarize the previous research works on image compression techniques. In Section III, we discuss our proposed hybrid color image compression technique in detail. In Section IV, we describe the implementation details of our work with experimental results. Finally, we conclude our paper in Section V.

2. Related Works

Of late, the image processing professionals have been dynamically involved in the designing of compression models. With the advent of multimedia products, complicated issues have surfaced that need a deeper understanding

and substantial investigation. Digital image compression, with their ability to establish valued insights for effective communication and storage system, have recently gained intensive attention from both researchers and academics. This section discusses some of the topical methods developed for image compression. PCA is a method to decrease the higher-dimensional space presented by mathematical functions [26, 27]. It extracts the basic features of a linear system using the SVD technique [28]. This technique has been extensively used to remove noise elements in image recognition, digital signal processing, solving classification problems, etc. [29 – 33]. Santo implemented PCA to reduce the dimension of medical images [34]. Other PCA based digital image compression methods can be found in [35 - 39].

Vaish and Kumar proposed a new image compression technique using PCA and Huffman Coding [40]. The dimension of the given picture is first reduced by PCA, a limited number of the principal components (PCs) are employed to reproduce the output image, whereas the other less important PCs are discarded. The reproduced image is further quantized to decrease contouring, before following to a smaller number of PCs are employed for image reproduction. Finally, the Huffman coding is used to eliminate coding redundancy in the quantized image. This method is tested on different images and results are compared with JPEG2000. Visual results and comparative analysis reveal that the proposed method outperforms the JPEG2000.

Yadav and Nagmode proposed a hybrid image compression method using PCA and DCT (i.e., PCA-DCT), in which PCA is used to calculate the dissimilarities, similarities, and feature vector in the form of the residual picture. Then, DCT is implemented to reduce the dimension of the given picture [41]. In order to increase the compression efficiency further, Mei et al. developed an integrated technique called folded-PCA in combination with JPEG2000 [42]. This method calculates the covariance matrix by folding the spectral vector into a matrix. Then, the Eigenvectors are employed to find PCs that can characterize features of the whole picture. JPEG2000 is used to reduce the dimension of the picture further.

Xiao et al. presented an image compression method using DTT and matrix factorization theory. In this work, the $N \times N$ transform matrix was factorized into $N + 1$ single-row elementary reversible matrices with minimum rounding errors [43]. Actually, for efficient lossless compression, integer DTT (iDTT) was developed to realize integer to integer mapping. Furthermore, a series of investigations are carried out, and the results confirmed that the applied iDTT method not only provides a greater CR than the iDCT method, however it was compatible with the widely used JPEG standard.

Kishk et al. introduced an integral image compression method that combines DWT and PCA. This method is based on implementing PCA on the wavelet coefficients of the 3D input images to enhance the quality of the compressed images while attaining higher CR. The wavelet coefficients of the specific image are weighted and reordered before implementing the PCA algorithm. The PCA is applied to each sub-band separately to improve the CR. The quality of the compressed 3D images and given input images are measured [44]. Senapati et al. developed a DTT-based hybrid method, where DTT is combined with SPIHT to compress images [45]. Also, suitable perceptual weights are used to enhance the quality of the compressed image. The authors compare the system performance with several state-of-the-art compression schemes. Nevertheless, the above mentioned approaches cannot provide a finite resolution and need a substantial amount of processing time to compress the given image.

3. Proposed System

In this section, we propose a hybrid color image compression approach based on PCA - DTT algorithms, which integrates the benefits of both PCA and DTT. The proposed method is implemented in order to realize improved compression performance and to get better-quality reconstructed images. The number of features extracted from DTT is relatively high, which can be reduced to a manageable low-dimensional space by eliminating the irrelevant features in the given image using PCA.

3.1 Principal Component Analysis

PCA is traditionally used as a dimension reduction technique and is a useful tool to visualize high dimensional data in a manageable low-dimensional space [46, 47]. It exploits an orthogonal transformation to convert a set of y (perhaps correlated) perceived variables (i.e., features) into another set of x uncorrelated features (i.e., PCs). The principal components are uncorrelated linear functions of the initially perceived features that consecutively exploit variance such that the first PC implies that the axis along which the observed data show the maximum variance; the second PC implies that the axis that is orthogonal to the first PC and along which the perceived data show the second largest variance; the third PC implies that the axis that is orthogonal to the first two PCs and along which the perceived data show the third-largest variance, and so forth. Thus, y orthogonal dimensions of data variability are captured in y PCs and the amount of variability that each PC accounts for gathered entire data variation. The key objective of PCA is to capture as much variation as possible in the first few PCs. It is, therefore, often the case

that the first x ($x \ll y$) PCs hold possibly valuable information in the perceived data, and the rest hold variation mostly due to noise [27].

More precisely, let b_{ij} represent a real-valued observation of the j th feature made on the i th subject, where $i = 1, 2, \dots, n$ and $j = 1, 2, \dots, y$. Assume that the n observations are organized in n rows of a $n \times y$ data matrix B with columns related to y features. We normalize columns of B to have zero mean and unit standard deviation and save the resultant values in a data matrix A , that is, the elements a_{ij} of A are calculated by Equation (1).

$$a_{ij} = \frac{(b_{ij} - \bar{b}_j)}{S_j} \quad (1)$$

where \bar{b}_j and S_j are the mean and standard deviation of the j th column of B , correspondingly. The PCA can be carried out using the SVD method of A ; that is, the $n \times y$ matrix A of rank $r \leq \min(n, y)$ is decomposed as given in Equation (2).

$$A = K\mathcal{D}L^T \quad (2)$$

where K is an $n \times r$ orthonormal matrix ($K^TK = I_r$), \mathcal{D} is an $r \times r$ diagonal matrix comprising r non-negative singular values in descending order of magnitude on the diagonal and L is a matrix with orthonormal columns ($L^TL = I_r$). The sample correlation matrix of A by M can be expressed as follows

$$M = \frac{1}{n-1} A^T A = L\delta L^T \quad (3)$$

$n - 1$

where δ is an $r \times r$ diagonal matrix encompassing r non-zero positive singular values (i.e., eigenvalues) $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_r)^T$ of matrix M on the diagonal in descending order of magnitude. It follows that the r columns of matrix L comprise the eigenvectors of $A^T A$ and hence are the desired directions of variation. The derived set of r PCs are calculated by

$$C = AL \quad (4)$$

It is noteworthy to mention that the matrix K comprises normalized PCs in its columns and is a scaled version of C , which is provided additionally in Equation (1). To see this, multiply Equation (1) on the right by L to obtain Equation (5).

$$C = AL = K\mathcal{D} \quad (5)$$

In practice, the first $x \ll r$ major PCs are of much interest since they represent the majority of the data variation. Without unnecessary loss of information, the size of C may thus be reduced from y to x , that is

$$\bar{C} = A\bar{L} \quad (6)$$

where \bar{L} is a $y \times x$ matrix that comprises the first x columns of L and \bar{C} comprises the first x PCs in its columns. The set of first x PCs is a lower-dimensional characterization of a y -dimensional dataset and can be applied to expose trends and patterns in the data, which is possibly the most predominant application of PCA. Additionally, the most useful PCs can be employed in further data analyses. From original features, the x PCs (although useful because they represent most of the data variation) are difficult to infer [48], and frequently misleading or imaginary interpretations are happened [27]. The noise level in datasets owing to high dimensionality often conceals valuable interpretable patterns [49]. Generally, the PCs are generated in an unsupervised method so that may not be optimum for further studies like regression. In this case, it is assumed that only a limited number of features can reveal reproducible patterns in the perceived data and feature selection (instead of extraction) techniques are generally selected.

A lower rank approximation of A can be calculated by Equation (7).

$$\bar{A} = \bar{C}\bar{L}^T \quad (7)$$

which is the best approximation of A in the least-squares sense by a matrix of rank x [50, 51]. The value of x is selected by means of existing heuristic options including a plot of Eigenvalues in descending order of magnitude, called a scree-plot, which selects the least value of x for which the scree plot surpasses a predefined threshold μ [0,1]. Generally, the value of μ ranges from 0.7 to 0.9 [27]. A value of $\mu = 0.85$ implies that the selected x PCs describe at least 85% of the cumulative variance in the perceived data. PCA executes the following steps to compress an input image:

1. Determine the vector from the input image matrix.
2. Calculate the covariance matrix.

3. Singular values and Eigenvectors are calculated by solving the characteristics equation. Each Eigenvector should be standardized.
4. From the standardized Eigenvectors, the transformation matrix is constructed.
5. Derive the transform of the given image.
6. Input values are reproduced from the transformed coefficients.
7. Reduce the dimensionality of the input image matrix.

3.2 Discrete Tchebichef transform

DTT algorithm has been widely used to enhance the reconstruction quality of the conventional image compression methods. This work implements an effective lossy compression technique based on DTT to yield a better-quality reproduced picture for the expected CR. The DTT is a new transform that exploits the Tchebichef moments to derive a basis matrix. DTT is derived from the orthonormal Tchebichef polynomials [52]. For an image of dimension $D_x \times D_y$, the forward DTT of order $r + s$ is defined as

$$T_{rs} = \sum_{x=0}^{D_x-1} \sum_{y=0}^{D_y-1} t_r(x)t_s(y) g(x, y) \quad (8)$$

where $r, s = 0, 1, \dots, D - 1$. The inverse transform of DTT is defined by

$$g(x, y) = \sum_{r=0}^{D-1} \sum_{s=0}^{D-1} t_r(x)t_s(y)T_{rs} \quad (9)$$

where $x, y = 0, 1, \dots, D - 1$. From (8) and (9), $t_r(x)$ and $t_s(y)$ are r th- and s th-order Tchebichef polynomials, correspondingly. Generally, q th-order Tchebichef polynomial is defined using the following recurrence relation as

$$t_q(x) = (\eta_1(x) + \eta_2)t_{q-1}(x) + \eta_3 t_{q-2}(x) \quad (10)$$

$$\eta_1 = \frac{2}{q} \sqrt{\frac{4q^2 - 1}{D^2 - q^2}} \quad (11)$$

$$\eta_2 = \frac{1 - D}{q} \sqrt{\frac{4q^2 - 1}{D^2 - q^2}} \quad (12)$$

$$\eta_3 = \frac{q-1}{q} \sqrt{\frac{2q+1}{2q-3}} \sqrt{\frac{4q^2 - 1}{D^2 - q^2}} \quad (13)$$

The initial values of $t_\square(\square)$ for $\square = 0, 1$ is defined as

$$t_0(x) = \frac{1}{\sqrt{q}} \quad (14)$$

$$t_1(x) = \frac{2x + 1 - D}{\sqrt{3}} \sqrt{\frac{1}{D(D^2 - 1)}} \quad (15)$$

Equation (16) can be expressed using a series representation involving matrices as follows

$$g(x, y) = \sum_{r=0}^{D_x-1} \sum_{s=0}^{D_y-1} \vartheta_{rs} T_{rs} \quad (16)$$

where $x, y = 0, 1, \dots, D - 1$, and ϑ_{rs} is known as the basis matrix. The basis matrix ϑ_{rs} can be defined as follows

$$\vartheta_{rs} = [\begin{matrix} t_r(0)t_s(0) & t_r(0)t_s(1) & \dots & t_r(0)t_s(7) \\ t_r(1)t_s(0) & t_r(1)t_s(1) & \dots & t_r(1)t_s(7) \\ \dots & \dots & \dots & \dots \\ t_r(7)t_s(0) & t_r(7)t_s(1) & \dots & t_r(7)t_s(7) \end{matrix}] \quad (17)$$

Hence, the DTT of a square image

$$I = \{g(x, y)\}_{x, y=0}^{D-1} \quad (18)$$

as in (12) can be viewed as the projection of the image I on the basis image ϑ_{rs} , which is the product of the vectors t_r and t_s , where $t_r = t_r(0)t_r(1) \dots t_r(D-1)$ and $t_s = t_s(0)t_s(1) \dots t_s(D-1)$. Equation (18) can be written as

$$\vartheta_{rs} = [t_r]'[t_s] \quad (19)$$

In other words, Tchebichef transform T_{rs} estimates the correlation between the image I and basis image ϑ_{rs} . It stores a high positive value if there is a strong similarity between them. It shows that when the order of the transform is increased, the basis images are changed from low spatial frequency to high spatial frequency. This proves that there will be neither large variation in the dynamic range of transformed values nor numerical instabilities that occur for large values of D .

3.3 Hybrid color image compression approach

In the proposed approach, a hybrid technique is developed that integrates the PCA and DTT techniques to compress color images. The block diagram representation of the proposed method is depicted in Figure 1.

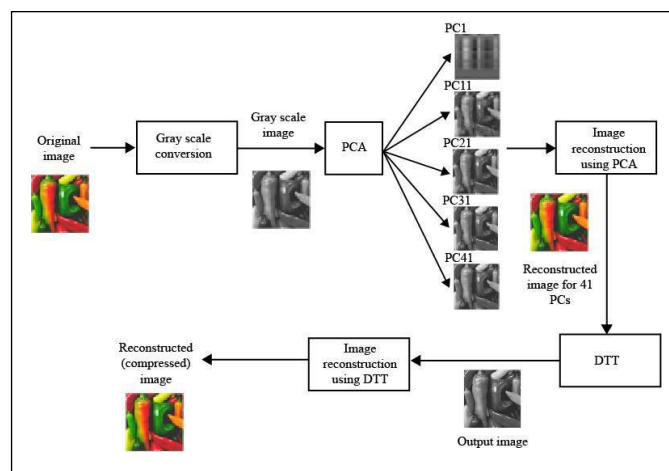


Figure 1: Overview of proposed hybrid image compression method

Our PCA-DTT technique implemented in two phases as follows:

1. First, the dimension of the input picture is reduced by PCA. It relates the majority of the image data in the first few PCs itself. Dimensionality reduction is carried out using the first few PCs of the given picture. Subsequently, the picture can be reproduced with the first few coefficients. PCA is considered an ideal compression technique in terms of its energy compaction feature. Even though PCA provides better compression ratio, the other quality parameters including PSNR, SSIM, and UQI values of the reconstructed image are poor. Accordingly, these quality metrics can be improved using the DTT algorithm.

2. In the second phase, DTT is used for image compression which provides improved PSNR, SSIM, and UQI values as compared to PCA. Accordingly, the quality of the reproduced picture regained from the PCA can be further enhanced using the DTT method. In this phase, DTT is employed to the output image obtained from PCA to generate transformed coefficients. Subsequently, the compression can be performed through the transformed coefficient values using DTT.

4. Results and Discussion

In this section, we demonstrate the efficiency of our compression method using the results obtained from the experimentation. The PCA-DTT technique is implemented and evaluated using MATLAB 2017a running in a system with a Core-i7 processor, 32 GB RAM, and Windows 10 operating system. Different images including animals, human faces, historical images, objects, vegetables with different formats (i.e., bmp, png, tiff, and jpg) and different sizes (i.e., 512×512, 256×256, and 120×80) are used as inputs. For all images, we use a block size of 8×8. The quality of the reproduced images is assessed by calculating the PSNR, SSIM, and UQI values. The performance of the PCA-DTT system is evaluated with respect to CR and CT.

4.1 Compression Performance Assessment

The performance of image compression methods can be evaluated in many aspects. We assess the compression performance of our method using the metrics including CR and CT. The compression ratio is defined as the proportion of the total number of bits essential to save the input picture and the total number of bits essential to save the reproduced picture.

$$CR = \frac{\text{Number of bits in uncompressed image}}{\text{Number of bits in a compressed image}} \quad (20)$$

The compression time is the time taken by the proposed technique to compress an image. This quality parameter demonstrates the computational complexity and the speed of the compression process.

4.2 Image Quality Assessment

In order to enumerate the visually observed variances between the input and the reconstructed images, many quality metrics have been considered [53]. The difference in the compressed image from the original image in lossy compression is called distortion. Other terms like quality and fidelity are also used to denote the variation between input and compressed image. The widely used measure for this purpose is PSNR. It is used to find the observed errors perceptible to the human vision. This ratio is usually calculated in terms of the logarithmic decibel scale (dB). It is defined as the proportion between the peak signal power and the humiliating noise power that distresses the fidelity of its characterization. The higher value of PSNR indicates the improved quality of the reproduced image. The PSNR is calculated as follows

$$PSNR = 10 \log \left(\frac{255^2}{MSE} \right) \quad (21)$$

The value of SSIM determines the resemblance between two pictures and can be considered as the quality of one of the pictures being related when the other one is taken as the reference [54]. The rudimentary tenet of measuring SSIM is that the structural distortion can be related to observed picture quality and is measured as:

$$SSIM(o, r) = \frac{(2\mu_o\mu_r + c_o)(2\sigma_{or} + c_r)}{(\mu_o^2 + \mu_r^2 + c_o)(\sigma_o^2 + \sigma_r^2 + c_r)} \quad (22)$$

where o and r represent the original and the reconstructed pictures, μ_o and μ_r are mean values of the luminance in the original and reconstructed picture correspondingly. The parameters σ_o and σ_r are the equivalent standard deviations of the luminance, and σ_{or} is the cross-covariance. The factors c_o and c_r are the contrast values of the original and reconstructed pictures respectively. The value of SSIM can range from -1 to +1; SSIM = +1 is realized only when the reconstructed image is identical to the original image. SSIM does not consider the attributes of the human visual system but was nevertheless shown to be sturdily associated with the individual picture quality metrics. Estimation of SSIM also creates a map, which gives the value of picture quality over space thus making it possible to relate various regions in the pictures and to perceive their resemblances.

Another important quality measure to evaluate compression algorithm is known as the UQI [55]. It is calculated by modeling the picture distortion in terms of luminance distortion, contrast distortion, and loss of correlation. The value of UQI for each sub-image can be estimated as:

$$UQI = \frac{4\sigma_{or}\mu_o\mu_r}{(\mu_o^2 + \mu_r^2)(\sigma_o^2 + \sigma_r^2)} \quad (23)$$

where μ_o and μ_r are the mean value of the original and the reconstructed pictures. The parameters σ_o and σ_r are the standard deviations of the original and reconstructed pictures and σ_{or} is the covariance.

4.3 Implementation of PCA

Figure 2 displays the original picture (Figure 2(a) and 2(b)) and feature-reduced pictures for different PCs (Figure 2(c) to 2(h)) after applying PCA. The input image of size 512×512 which is shown in Figure 2(a) has experimented with PCA. Figure 2(b) is the equivalent gray scale image. Figures 2(c) - 2(h) are the feature-reduced images for 1 PC, 11 PCs, 21 PCs, 31 PCs and 41 PCs respectively.

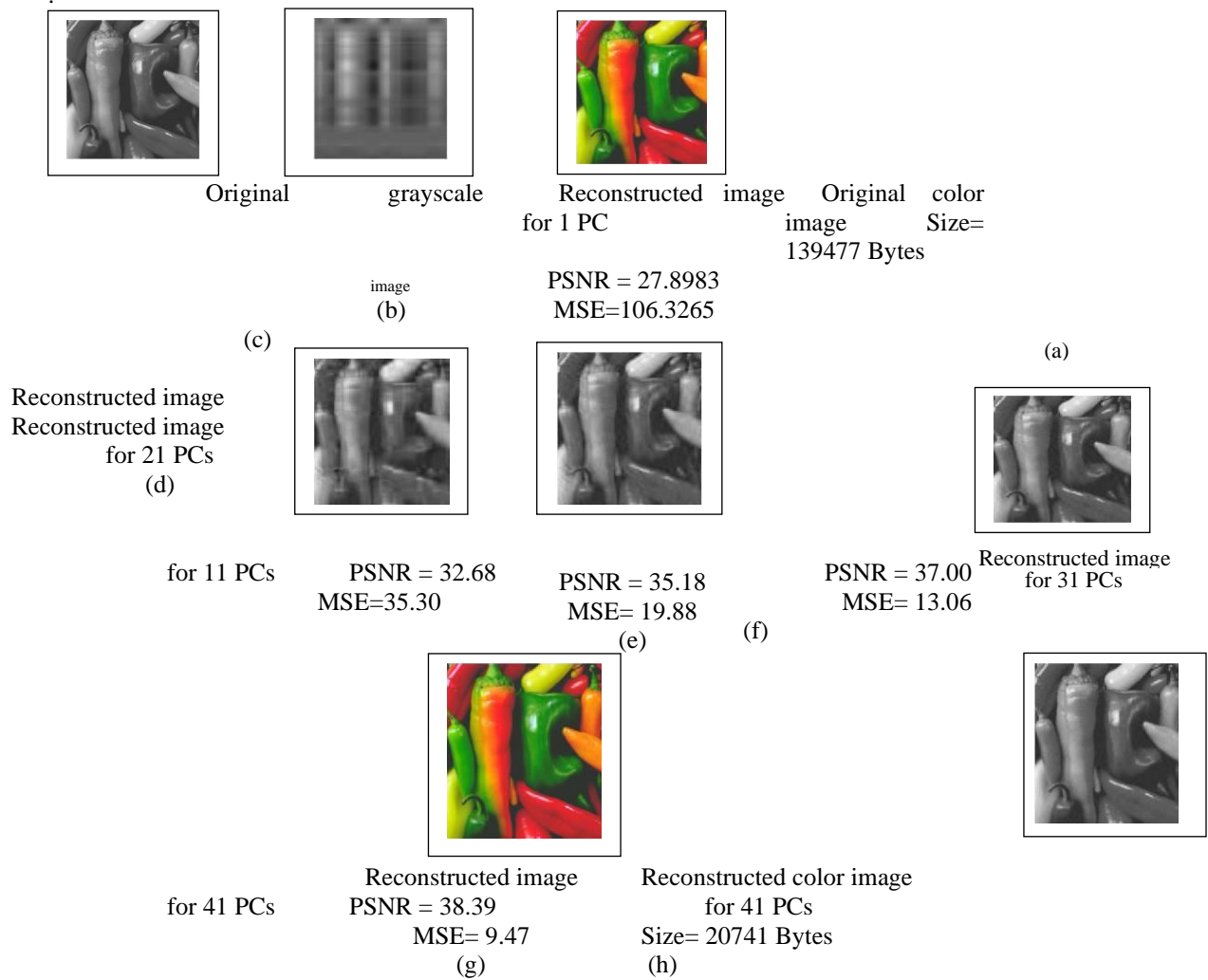


Figure 2: Input and output images of the PCA algorithm

Figure 3 shows the screen plot (i.e., the variances of the extracted PCs as a function of the component index [27]). In this plot, the eigenvalues are sorted from larger to smaller. This will have significant inferences, in selecting a subset of PCs to find the representative behavior of the image data. The graph levels off after PC5, demonstrating that factors 6–8 represent comparatively slight additional variance. Hence, five PCs are considered.

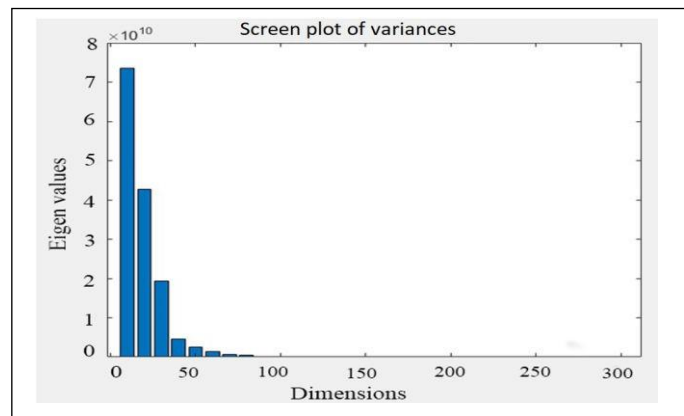


Figure 3: Screen-plot of variance in PCA algorithm

4.4 Implementation Of DTT

The compression efficiency gained from the PCA is not found to be reasonable. Therefore, the quality of the compressed picture gained from the PCA can be further enhanced by the DTT technique that provides superior PSNR value. DTT is applied to the picture reproduced by PCA in order to obtain the transformed coefficients.

Subsequently, the compression can be performed by applying the transformed coefficient values. Figure 3 displays input image (Figure 3(a)) and output images (Figure 3(b) and 3(c)) of DTT algorithm.

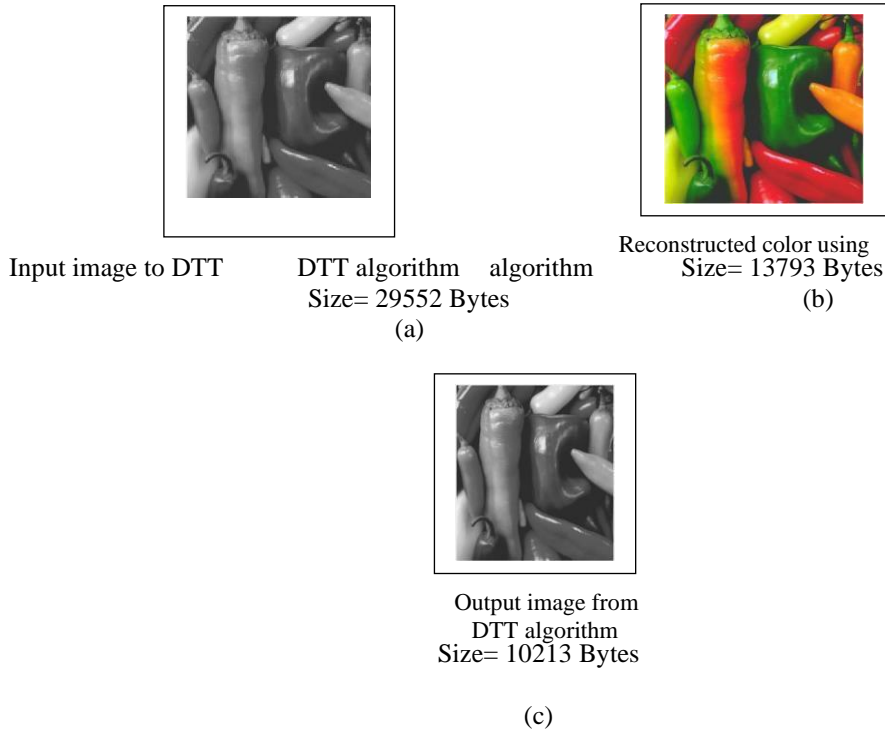


Figure 3: Input and output images of the DTT algorithm

4.5 Implementation Of Hybrid Color Image Compression Technique

The PCA-DTT method is evaluated with different pictures as shown in Figure 4. The reproduced output pictures are displayed in Figure 5.

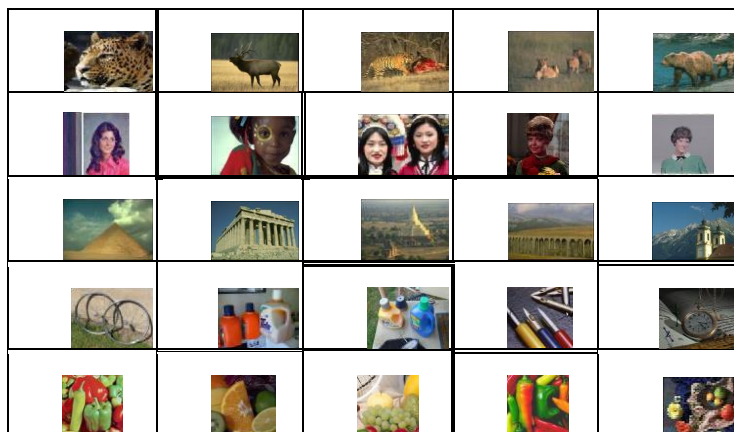


Figure 4: Input images used to evaluate PCA-DTT algorithm

In order to illustrate the effectiveness of the PCA-DTT technique over the DTT-SVD method, an extensive comparison in terms of different metrics, namely, CR, CT, PSNR, SSIM, and UQI are carried out and the results are given in Table 1. From a comprehensive study of the results observed, it can be stated that our

hybridtechniqueprovidesbetter results over the DTT-SVD method,hencethe proposed method can beconsideredbeingconsistent and robust.

Table 1: Comparison of performance measures between DTT-SVD and PCA-DTT

Performance measures		CR	CT(Sec)	PSNR	SSIM	UQI
Animal Image	DTT-SVD	7.30	4.86	29.25	0.86	0.69
	PCA-DTT	8.20	4.20	31.02	0.93	0.77
Human face	DTT-SVD	6.10	4.41	30.07	0.86	0.70
	PCA-DTT	7.90	3.92	31.54	0.95	0.84
Historical image	DTT-SVD	7.30	3.40	32.99	0.89	0.79
	PCA-DTT	8.00	3.10	33.43	0.92	0.89
Object image	DTT-SVD	9.00	5.36	29.03	0.91	0.68
	PCA-DTT	10.40	4.21	32.40	0.97	0.84
Vegetable image	DTT-SVD	9.40	4.53	28.57	0.90	0.69
	PCA-DTT	11.00	3.21	33.15	0.98	0.80

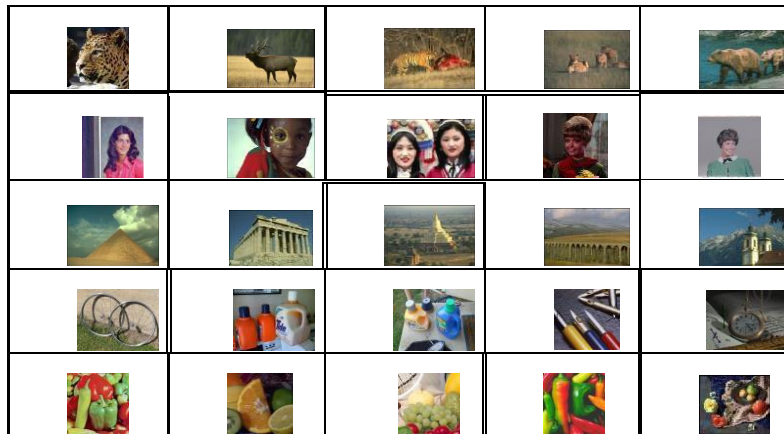


Figure 5: Output images obtained from PCA-DTT algorithm

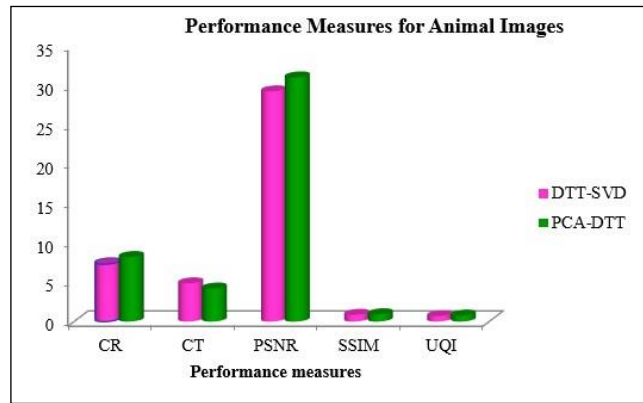


Figure 6: Performance measures for Animal Images

The results gained from our PCA-DTT technique in terms of performance metrics are displayed in Figures 6 to 10. From the graphs, it is found that the proposed method achieved improved performance as well as quality metrics. More precisely, it can be observed that the PCA-DTT method provides an improved compression ratio, which is superior to that of the DTT-SVD. The average CR gained by the PCA-DTT method and DTT-SVD is 8.20 and 7.30 respectively. Therefore, PCA-DTT can be considered as a better technique.

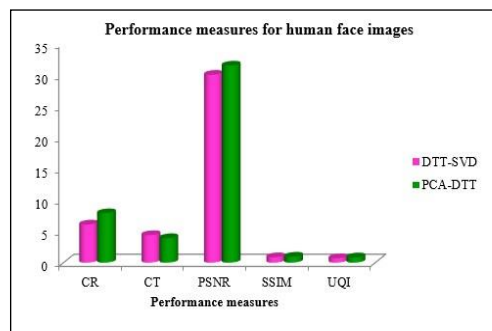


Figure 7: Performance measures for human face image

Additionally, to prove the superiority of the PCA-DTT technique in terms of computational complexity the compression time is considered. The average time taken for image compression by the PCA-DTT technique and DTT-DVD is 3.92 sec and 4.41 sec respectively. Hence, it is concluded that the computational complexity associated with PCA-DTT is less as compared to DTT-SVD.

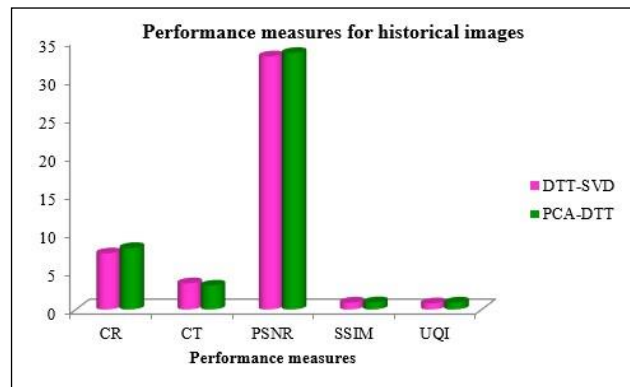


Figure 8: Performance measures for historical image

The quality of output images obtained from the PCA-DTT technique has also been evaluated and compared in terms of PSNR with a similar statistical setup. It is observed that the proposed method outperforms the DTT-SVD technique with respect to PSNR. The proposed method attains a higher average PSNR (33.43) than DTT-SVD (32.99).

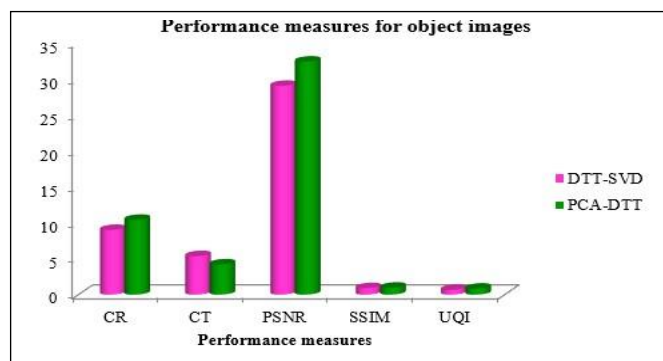


Figure 9: Performance measures for object image

From a comprehensive study of the output images obtained, it can be concluded that the proposed method provides better results over DTT-SVD with respect to both SSIM and UQI. The average SSIM achieved by the PCA-DTT method and DTT-DVD is 0.97 and 0.91 correspondingly. Also, the average UQI achieved by the PCA-DTT method and DTT-DVD is 0.84 and 0.68 correspondingly.

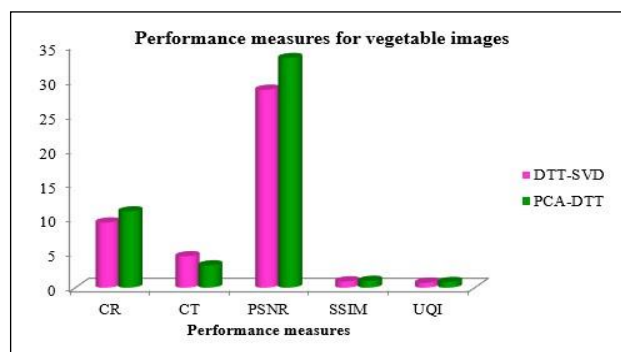


Figure 10: Performance measures for vegetable image

5. Conclusion

The increasing applications of digital imaging in various scientific and technological domains make it more difficult to process huge volumes of multimedia products. The processing, transmission, and storage of digital images in their original form are very costly. Hence, data compression has become an inevitable part of image processing towards the reduction of the transmission time, communication bandwidth, and storage space requirements of images. This paper presents a lossy compression method, which integrates PCA and DTT algorithms. This approach employs the PCA technique to reduce the dimensionality of the image and the DTT algorithm to enhance the quality of the reconstructed image. The proposed approach is implemented using MATLAB and its sturdiness is verified against the DTT-SVD image compression algorithm as per quantifying the measuring factors such as CR, CT, PSNR, SSIM, and UQI. The proposed method overcomes the restrictions about image representation confronted by the prior compression methods. The experimental results reveal that our PCA-DTT method outdoes the existing DTT-SVD method for different types of image contents at a high CR with lower computational complexity and retaining the quality of input digital image.

References

- A. Agarwal R., Dhar V. Editorial – big data, data science, and analytics: the opportunity and challenge for is research Information Systems Research, 25 (3) (2014), pp. 443-448
- B. Dobre C., Xhafa F. Intelligent services for big data science Future Generation Computer Systems, 37 (2014), pp. 267-281
- C. IDC, The Digitization of the World from Edge to Core, November 2018.
- D. Nagalakshmi, K., Gomathi, N., Criticality-cognizant Clustering-based Task Scheduling on Multicore Processors in the Avionics Domain, International Journal of Computational Intelligence Systems, 11 (1), pp. 219 – 237, 2018

- E. Garcia-Vilchez F, Munoz-Mari J, Zortea M, Blanes I, Gonzalez-Ruiz V, Camps-valls G, Plaza A, SerraSagrasta J. On the impact of lossy compression on hyperspectral image classification and unmixing. *IEEE Geosci Remote Sens Lettr* 2011; 8: 253-257.
- F. Glatard T, Lartizien C, Gibaud B, da Silva R.F, Forestier G, Cervenansky F, Alessandrini M, Benoit-Cattin H, Bernard O, Camarasu-Pop S, Cerezo N, Clarysse P, Gaignard A, Hugonnard P, Liebgott H, Marache S, Marion A, Montagnat J, Tabary J, Friboulet D. A virtual imaging platform for multi-modality medical image simulation. *IEEE Transact Med Imag* 2013; 32: 110-118.
- G. Singh, M., Kumar, S., Singh, S. & Shrivastava, M. (2016). Various Image Compression Techniques: Lossy and Lossless. *International Journal of Computer Applications*, 142(6), 23-26. doi: 10.5120/ijca2016909829.
- H. Joshi, M.A. *Digital Image Processing: An Algorithm Approach*; PHI Learning Pvt. Ltd.: New Delhi, India, 2006
- I. X. Hou, M. Han, C. Gong and X. Qian, SAR complex image data compression based on quadtree and zerotree coding in discrete wavelet transform domain: A comparative study. *Neurocomputing*, 148 (2015): 561-568.
- J. L. Zhang, L. Zhang, D. Tao, X. Huang and B. Du, Compression of hyperspectral remote sensing images by tensor approach. *Neurocomputing*, 147 (2015): 358-363.
- K. C. Shi, J. Zhang and Y. Zhang, Content-based onboard compression for remote sensing images. *Neurocomputing*, 191 (2016): 330-340.
- L. M. J. Weinberger, G. Seroussi, and G. Sapiro, The LOCO-I lossless image compression algorithm: Principles and standardization into JPEG-LS, *IEEE Trans. Image Processing*. 9 (8) (2000) 1309–1324.
- M. Hussain, A.J.; Al-Fayadh, A.; Radi, N. Image compression techniques: A survey in lossless and lossy algorithms. *Neurocomputing* 2018, 300, 44–69.
- N. AbuBaker, A., Eshtay, M., AkhoZahia, M., “Comparison Study of Different Lossy Compression Techniques Applied on Digital Mammogram Images”, *International Journal of Advanced Computer Science and Applications*, 7(12), 2016, pp. 149-155.
- O. A. Avramovic, Lossless compression of medical images based on gradient edge detection, 19th Telecommunications forum Tel. (Serbia, 2011) 22-24.
- P. Golomb, S. Run-length encodings (Corresp). *IEEE Trans. Inform. Theory* 1966, 12, 399–401.
- Q. S. Grgic, M. Grgic and B. Z. Cihlar, Performance analysis of image compression using wavelets, *IEEE Trans. on Industrial Electronics* 48 (2001) 682-695.
- R. L. Rabiner and R. Schafer. *Digital Processing of Speech Signals*. Prentice-Hall, Inc., Englewood Cliffs, NJ, 1971.
- S. P. A. Wintz, “Transform Picture Coding”, *Proceedings IEEE*, Vol. 60, pp. 809-820, July 1972.
- T. H. C. Anderws and W. K. Pratt, “Fourier Transform Coding of Images”, *Hawaii Int. Conf. Sys. Sci.*, Western periodicals Company, North Hollywood, California, pp. 677- 679, 1968.
- U. N. Ahmed, T. Natarajan, and K. R. Rao, “Discrete Cosine Transform”, In: *IEE Transactions on Computers*, Vol. 23, pp. 90-93, 1974.
- V. Rufai AM, Anbarjafari G, Demirel H. Lossy image compression using singular value decomposition and wavelet difference reduction. *Digital Signal Process* 2014; 24: 117-123.
- W. R. Mukundan, “Image Analysis by Tchebichef Moments”, *IEEE Trans. on Image Proc.*, vol.10, no.9, pp.1357-1364, 2001
- X. Said A, Pearlman WA. A new, fast and efficient image codec based on set partitioning in hierarchical trees. *IEEE Transact Circuit Syst Video Technol* 1996; 6: 243-250.
- Y. Antonini M, Barlaud M, Mathieu P, Daubechies I. Image coding using wavelet transform. *IEEE Transact Image Process* 1992; 1: 205-220.
- Z. B. Moore, "Principal component analysis in linear systems: Controllability, observability, and model reduction," *IEEE Transactions on Automatic Control*, vol. 26, no. 1, 17-32, 1981.
- AA. I. Jolliffe, *Principal Component Analysis*, Springer, 2002.
- BB. J. Tzeng, "Split-and-combine singular value decomposition for large-scale matrix," *Journal of Applied Mathematics*, 2013
- CC. R. Petrolis, and A. Krisciukaitis, "Multi stage Principal component analysis based method for detection of fetal heart beats in abdominal ECGs," *Computing in Cardiology Conference 2013, Zaragoza*, 2013, pp. 301 - 304.
- DD. H. Shi, "Application of principal component analysis to general contracting risk assessment," *2009 ISECS International Colloquium on Computing, Communication, Control, and Management, Sanya*, 2009, pp. 53-
- EE. 56.

- FF. G. Tzimiropoulos, S. Zafeiriou, and M. Pantic, "Principal component analysis of image gradient orientations for face recognition," 2011 IEEE International Conference on Automatic Face & Gesture Recognition and Workshops, Santa Barbara, CA, 2011, pp. 553-558.
- GG. Li, H., "Accurate and efficient classification based on common principal components analysis for multivariate time series," *Neurocomputing*, vol. 171, pp. 744-753, 2016.
- HH. P. Kamencay, T. Trnovszky, M. Benco, R. Hudec, P. Sykora and A. Satnik, "Accurate wild animal recognition using PCA, LDA and LBPH," 2016 ELEKTRO, Strbske Pleso, 2016, pp. 62-67.
- II. R. D. Santo, "Principal component analysis applied to digital image compression," *Einstein*, vol. 10, no. 2, pp. 135-139, 2012.
- JJ. Q. Du and James E. Fowler, "Low-complexity principal component analysis for hyperspectral image compression," *International Journal of High Performance Computing Applications*, vol. 22, no. 4, pp. 438 - 448, 2002.
- KK. C. Lee, S. Youn, T. Jeong, E. Lee and J. Serra-Sagrìstà, "Hybrid Compression of Hyperspectral Images Based on PCA With Pre-Encoding Discriminant Information," in *IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 7, pp. 1491-1495, July 2015.
- LL. C. W. Wang and J. H. Jeng, "Image compression using PCA with clustering," 2012 International Symposium on Intelligent Signal Processing and Communications Systems (ISPACS), New Taipei, 2012, pp. 458-462.
- MM. A. Vaish and M. Kumar, "WDR coding based image compression technique using PCA," 2015 International Conference on Signal Processing and Communication (ICSC), Noida, 2015, pp. 360-365.
- NN. S. T. Lim, D. F. W. Yap and N. A. Manap, "Medical image compression using block-based PCA algorithm," *Computer, Communications, and Control Technology (I4CT)*, 2014 International Conference on, Langkawi, 2014, pp. 171-175.
- OO. A. Vaish and M. Kumar, "A new Image compression technique using principal component analysis and Huffman coding," 2014 International Conference on Parallel, Distributed and Grid Computing, Solan, 2014, pp. 301-305, doi: 10.1109/PDGC.2014.7030760.
- PP. R. J. Yadav and M. Nagmode, "Compression of hyperspectral image using PCA–DCT technology," in *Innovations in Electronics and Communication Engineering*, H. Saini, R. Singh, and K. Reddy, Eds., pp. 269–277, Springer, Singapore (2018).
- QQ. S. Mei et al., "Low-complexity hyperspectral image compression using folded PCA and JPEG2000," in *IEEE Int. Geosci. and Remote Sens. Symp.*, IEEE, pp. 4756–4759 (2018).
- RR. Xiao B, Lu G, Zhang Y, Li W and Wang G, Lossless image compression based on integer discrete Tchebichef transform. *Neurocomputing* 214: 587–593, 2016.
- SS. S. Kishk, H. Eldin, M. Ahmed, and H. Helmy, "Integral Images Compression using Discrete Wavelets and PCA," vol. 4, no. 2, pp. 65–78, 2011.
- TT. Senapati R.K, Pati U.C., Mahapatra K.K., A Novel Hybrid HVS Based Embedded Image Coding Algorithm Using DTT and SPIHT, *International Conference on Devices and Communications*, IEEE, 2011, pp. 1-5
- UU. Jolliffe IT, Cadima J. Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*. 2016;374(2065):20150202.
- VV. Ringner M. What is principal component analysis? *Nature Biotechnology*.2008;26(3):303.
- WW. Zou H, Hastie T, Tibshirani R. Sparse principal component analysis. *Journal of Computational and Graphical Statistics*. 2006;15(2):265–286.
- XX. Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science*. 1999;286(5439):531– 537.
- YY. Eckart C, Young G. The approximation of one matrix by another of lower rank. *Psychometrika*. 1936;1(3):211–218.
- ZZ. Saporta G, Niang N. Principal component analysis: application to statistical process control. *Data Analysis*. 2009; p. 1–23
- AAA. S. Prattipati, S. Ishwar, M. N. S. Swamy and P. K. Meher, "A fast 8×8 integer Tchebichef transform and comparison with integer cosine transform for image compression," 2013 IEEE 56th International Midwest Symposium on Circuits and Systems (MWSCAS), Columbus, OH, 2013, pp. 1294-1297.
- BBB. Lin, W., Jay Kuo, C.-C. (2011): "Perceptual visual quality metrics: A survey", *Journal of Visual Communication and Image Representation*, vol 22, no 4, 2011, pp. 297–312
- CCC. Wang, Z., Bovik, A. C., Sheikh, H. R., Simoncelli, E. P. (2004): "Image quality assessment: From error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600-612.
- DDD. Bovik Z. and Alan C., "A Universal Image Quality Index," *IEEE Signal Processing Letters*, vol. 9, no. 3, pp. 81-84, 2002.