

A Study of Various Image Compression Techniques

Sukanya, Gagan kumar

Department of Computer Science and Engineering, Modern Institute of engineering & technology, Mohri, Kurukshetra , Haryana, India

Abstract: This paper addresses the area of image compression as it is applicable to various fields of image processing. On the basis of evaluating and analyzing the current image compression techniques this paper presents the Principal Component Analysis approach applied to image compression. PCA approach is implemented in two ways - PCA Statistical Approach & PCA Neural Network Approach. It also includes various benefits of using image compression techniques.

Keywords: Image processing, Image Compression, Digital Image

I. INTRODUCTION

Image

An image is essentially a 2-D signal processed by the human visual system. The signals representing images are usually in analog form. However, for processing, storage and transmission by computer applications, they are converted from analog to digital form. A digital image is basically a 2-Dimensional array of pixels.

Images form the significant part of data, particularly in remote sensing, biomedical and video conferencing applications. The use of and dependence on information and computers continue to grow, so too does our need for efficient ways of storing and transmitting large amounts of data.

Image Compression

Image compression addresses the problem of reducing the amount of data required to represent a digital image. It is a process intended to yield a compact representation of an image, thereby reducing the image storage/transmission requirements. Compression is achieved by the removal of one or more of the three basic data redundancies:

1. Coding Redundancy
2. Interpixel Redundancy
3. Psychovisual Redundancy

Coding redundancy is present when less than optimal code words are used. Interpixel redundancy results from correlations between the pixels of an image. Psychovisual redundancy is due to data that is ignored by the human visual system (i.e. visually non essential information). Image compression techniques reduce the number of bits required to represent an image by taking advantage of these redundancies. An inverse process called decompression (decoding) is applied to the compressed data to get the reconstructed image. The objective of compression is to reduce the number of bits as much as possible, while keeping the resolution and the visual quality of the reconstructed image as close to the original image as possible. Image

compression systems are composed of two distinct structural blocks : an encoder and a decoder.

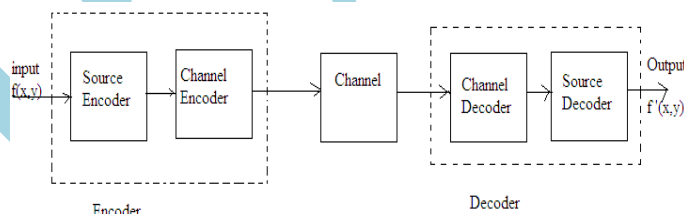


FIGURE 1

Image $f(x,y)$ is fed into the encoder, which creates a set of symbols from the input data and uses them to represent the image. If we let n_1 and n_2 denote the number of information carrying units (usually bits) in the original and encoded images respectively, the compression that is achieved can be quantified numerically via the compression ratio,

$$C_R = n_1 / n_2$$

As shown in the figure, the encoder is responsible for reducing the coding, interpixel and psychovisual redundancies of input image. In first stage, the mapper transforms the input image into a format designed to reduce interpixel redundancies. The second stage, quantizer block reduces the accuracy of mapper's output in accordance with a predefined criterion. In third and final stage, a symbol decoder creates a code for quantizer output and maps the output in accordance with the code. These blocks perform, in reverse order, the inverse operations of the encoder's symbol coder and mapper block. As quantization is irreversible, an inverse quantization is not included.

Benefits Of Compression

- It provides a potential cost savings associated with sending less data over switched telephone network where cost of call is really usually based upon its duration.
- It not only reduces storage requirements but also overall execution time.

- It also reduces the probability of transmission errors since fewer bits are transferred.
- It also provides a level of security against illicit monitoring.

II. IMAGE COMPRESSION TECHNIQUES

The image compression techniques are broadly classified into two categories depending whether or not an exact replica of the original image could be reconstructed using the compressed image. These are:

1. Lossless technique
2. Lossy technique

Lossless compression technique

In lossless compression techniques, the original image can be perfectly recovered from the compressed (encoded) image. These are also called noiseless since they do not add noise to the signal (image). It is also known as entropy coding since it uses statistics/decomposition techniques to eliminate/minimize redundancy. Lossless compression is used only for a few applications with stringent requirements such as medical imaging.

Following techniques are included in lossless compression:

1. Run length encoding
2. Huffman encoding
3. LZW coding
4. Area coding

Lossy compression technique

Lossy schemes provide much higher compression ratios than lossless schemes. Lossy schemes are widely used since the quality of the reconstructed images is adequate for most applications. By this scheme, the decompressed image is not identical to the original image, but reasonably close to it.

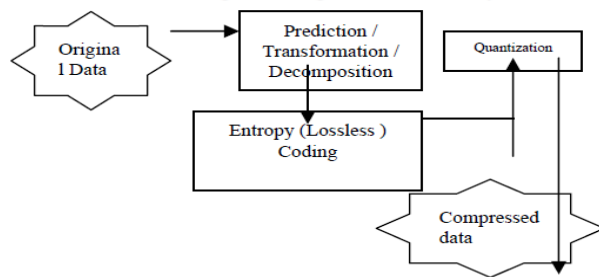


Figure: Outline of lossy image compression

As shown above the outline of lossy compression techniques. In this prediction - transformation - decomposition process is completely reversible. The quantization process results in loss of information. The entropy coding after the quantization step, however, is lossless. The decoding is a reverse process. Firstly, entropy decoding is applied to compressed data to get the quantized data. Secondly, dequantization is applied to it & finally the inverse transformation to get the reconstructed image.

Major performance considerations of a lossy compression scheme include:

1. Compression ratio
2. Signal - to - noise ratio
3. Speed of encoding & decoding.

Lossy compression techniques includes following schemes:

1. Transformation coding
2. Vector quantization
3. Fractal coding
4. Block Truncation Coding
5. Subband coding

LOSSLESS COMPRESSION TECHNIQUES

Run Length Encoding

This is a very simple compression method used for sequential data. It is very useful in case of repetitive data. This technique replaces sequences of identical symbols (pixels), called runs by shorter symbols. The run length code for a gray scale image is represented by a sequence $\{ V_i, R_i \}$ where V_i is the intensity of pixel and R_i refers to the number of consecutive pixels with the intensity V_i as shown in the figure. If both V_i and R_i are represented by one byte, this span of 12 pixels is coded using eight bytes yielding a compression ratio of 1: 5

Huffman Encoding

This is a general technique for coding symbols based on their statistical occurrence frequencies (probabilities). The pixels in the image are treated as symbols. The symbols that occur more frequently are assigned a smaller number of bits, while the symbols that occur less frequently are assigned a relatively larger number of bits. Huffman code is a prefix code. This means that the (binary) code of any symbol is not the prefix of the code of any other symbol. Most image coding standards use lossy techniques in the earlier stages of compression and use Huffman coding as the final step.

LZW Coding

LZW (Lempel- Ziv - Welch) is a dictionary based coding. Dictionary based coding can be static or dynamic. In static dictionary coding, dictionary is fixed during the encoding and decoding processes. In dynamic dictionary coding, the dictionary is updated on fly. LZW is widely used in computer industry and is implemented as compress command on UNIX.

Area Coding

Area coding is an enhanced form of run length coding, reflecting the two dimensional character of images. This is a significant advance over the other lossless methods. For coding an image it does not make too much sense to interpret it as a sequential stream, as it is in fact an array of sequences, building up a two dimensional object. The algorithms for area coding try to find rectangular regions with the same characteristics. These regions are coded in a descriptive form as an element with two points and a certain structure. This type of coding can be highly effective but it bears the problem

of a nonlinear method, which cannot be implemented in hardware. Therefore, the performance in terms of compression time is not competitive, although the compression ratio is.

LOSSY COMPRESSION TECHNIQUES

Transformation Coding

In this coding scheme, transforms such as DFT (Discrete Fourier Transform) and DCT (Discrete Cosine Transform) are used to change the pixels in the original image into frequency domain coefficients (called transform coefficients). These coefficients have several desirable properties. One is the energy compaction property that results in most of the energy of the original data being concentrated in only a few of the significant transform coefficients. This is the basis of achieving the compression. Only those few significant coefficients are selected and the remaining are discarded. The selected coefficients are considered for further quantization and entropy encoding. DCT coding has been the most common approach to transform coding. It is also adopted in the JPEG image compression standard.

Vector Quantization

The basic idea in this technique is to develop a dictionary of fixed-size vectors, called code vectors. A vector is usually a block of pixel values. A given image is then partitioned into non-overlapping blocks (vectors) called image vectors. Then for each in the dictionary is determined and its index in the dictionary is used as the encoding of the original image vector. Thus, each image is represented by a sequence of indices that can be further entropy coded.

Fractal Coding

The essential idea here is to decompose the image into segments by using standard image processing techniques such as color separation, edge detection, and spectrum and texture analysis. Then each segment is looked up in a library of fractals. The library actually contains codes called iterated function system (IFS) codes, which are compact sets of numbers. Using a systematic procedure, a set of codes for a given image are determined, such that when the IFS codes are applied to a suitable set of image blocks yield an image that is a very close approximation of the original. This scheme is highly effective for compressing images that have good regularity and self-similarity.

Block truncation coding

In this scheme, the image is divided into non overlapping blocks of pixels. For each block, threshold and reconstruction values are determined. The threshold is usually the mean of the pixel values in the block. Then a bitmap of the block is derived by replacing all pixels whose values are greater than or equal (less than) to the threshold by a 1 (0). Then for each segment (group of 1s and 0s) in the bitmap, the reconstruction value is determined. This is the average of the values of the

corresponding pixels in the original block.

2.4.5 Sub band coding
In this scheme, the image is analyzed to produce the components containing frequencies in well-defined bands, the sub bands. Subsequently, quantization and coding is applied to each of the bands. The advantage of this scheme is that the quantization and coding well suited for each of the sub bands can be designed separately.

III. PCA (Principal Component Analysis)

Introduction

In statistics, PCA is a technique for simplifying a dataset by reducing multidimensional datasets to lower dimensions for analysis. PCA is a standard technique commonly used for data reduction in statistical pattern recognition and signal processing. PCA has been called one of the most valuable results from applied linear algebra. It is used abundantly in all forms of analysis from neuroscience to computer graphics, because it is a simple non-parametric method of extracting relevant information from confusing datasets. PCA is also called the KARHUNEN-LOEVE Transform (KLT, named after Kari Karhunen & Michel Loeve) or the HOTELLING Transform. Its general objectives are:

1. Data reduction
2. Interpretation

There are basically two approaches for performing PCA. They are classical statistical method and artificial neural network method.

PCA Classical Statistical Method

It involves finding eigen values and corresponding eigen vectors of the data set using covariance matrix. The corresponding eigen values of the matrix gives an indication of amount of information the respective principal components represent. The methodology for calculating principal component is given by the following algorithm. Let X_1, X_2, \dots, X_m are the sub images of dimension N . The corresponding algorithm is described as follows:

Computation of the global mean (\bar{X}) from sub images.

$$\bar{X} = 1 / M \sum X_i$$

Subtraction of the mean from each sub image to generate the mean removed image. $0_i = X_i - \bar{X}$

Formation of the matrix using mean removed sub image of ($M \times N$) dimension

$$A = [0 \ 1 \ 02 \ \dots \dots \dots \ 0M \]$$

4. Computation of the sample covariance matrix (C) of dimension ($N \times N$)

Computation of the Eigen values of the covariance matrix.

Computation of Eigen values is performed by jacobian iteration method. $C : X_1 > X_2 > \dots > X_N$

Computation of the eigen vectors for the eigen values

$$C: u_1, u_2, \dots, u_n$$

7. Dimensionality reduction step.

Keep only the Eigen vectors corresponding to K largest eigen values. These Eigen values are called as "principal components".

The above said steps are needed to generate the principal components of the image. Corresponding eigen vectors are uncorrelated and have the greater variance. In order to avoid the components that have an undue influence on the analysis, the components are usually coded with mean as zero and variance as one. This standardization of the measurement ensures that they all have equal weight in the analysis.

PCA Neural Network

Artificial neural network are model that attempt to achieve performance via dense inter connection of simple computational elements. The most important property of a neural network is the ability to learn from its environment. PCA is a powerful linear block transform coding in which, an image is subdivided into non-overlapping blocks of $N \times N$ pixels which can be considered as N-Dimensional vectors with $N = n \times n$. A linear

Transformation, which can be written as an $M \times N$ - dimensional matrix W with $M < N$, is performed on each block with the M rows of W , w_i being the basis vectors of the transformation.

An adaptive principal component extraction (APEX) is used to decorrelate the principal components.

The main difference between this APEX architecture and the existing PCA networks lies in the additional lateral connections at the outputs of the network.

Applications of PCA in Computer Vision

- **Representation**

When using these sort of matrix techniques in computer vision, representation of images should be considered. A square, N by N image can be expressed as an N^2 - dimensional vector.

$$x = (x_1 \ x_2 \ x_3 \ \dots \ x_N)$$

Where the rows of pixels in the image are placed one after the other to form a one dimensional image. E.g. The first N elements $x_1 \dots x_N$ will be the first row of the image, the next N elements are the next row, and so on. The values in the vector are the intensity values of the image, possibly a single greyscale value.

- **PCA to find patterns**

Say we have 20 images Each image is N pixels high by N pixels wide. For each image we can create an image vector as described in the representation section. These all images can be put together in one big image-matrix like this :

ImageVec 1 ImageVec2
Image Matrix ImageVec20

Which gives a starting point for our PCA analysis. It turns out that these axes works much better for recognizing faces, because the PCA analysis has given the original images in terms of the differences and similarities between them. The PCA analysis has identified the statistical patterns in the data.

- **PCA for Image Compression**

If we have 20 images each with N^2 vectors and 20 dimensions. Now, PCA can be implemented on this set of data. 20 Eigen vectors will be obtained because each vector is 20 -dimensional. To compress the data, choose the data using only 15 Eigen vectors. This gives a final data set with only 15 dimensions, which has saved % of the space. However, when the original data is reproduced, the images have lost some of the information. This compression is said to be lossy because the decompressed image is not exactly the same as the original.

IV. CONCLUSION

This paper presents various types of image compression techniques. There are basically two types of compression techniques. One is Lossless Compression and other is Lossy Compression Technique. Comparing the performance of compression technique is difficult unless identical data sets and performance measures are used. Some of these techniques are obtained good for certain applications like security technologies. Some techniques perform well for certain classes of data and poorly for others. PCA (Principal Component Analysis) also found its applications as image compression. PCA can be implemented in two forms i.e. either statistical approach or neural network approach. The PCA Neural Network provides new way of generating codebook based on statistical feature of PCA transformational coefficients. It leads to less storage of memory and reduction of calculation.

REFERENCES

- [1]. Subramanya A, "Image Compression Technique," Potentials IEEE, Vol. 20, Issue 1, pp 19-23, Feb-March 2001,
- [2]. David Jeff Jackson & Sidney Joel Hannah, "Comparative Analysis of image Compression Techniques," System Theory 1993, Proceedings SSST '93, 25th Southeastern Symposium, pp 513-517, 7 -9 March 1993.
- [3]. Milos Klima, Karel Fliegel, "Image Compression Techniques in the field of security Technology: Examples and Discussion," Security Technology, 2004, 38th Annual 2004 Intn. Carnahan Conference, pp 278-284, 11-14 Oct., 2004