BIBLIOGRAPHICAL SURVEY FOR A NOVEL APPROACH TOWARDS DEVELOPMENT OF A HYBRID APPROACH OF IMAGE CODING USING NEURAL NETWORK AND WAVELET TRANSFORM

Geetha H S¹, Rehna V J²

¹¹⁴th Semester, M.Tech(Electronics)
²Associate Professor, Department of Electronics and Communication Engineering,
HKBK College of Engineering, Bangalore, India

ABSTRACT

This paper deals with survey of a number of reference papers in order to develop a hybrid approach of image coding using neural network and wavelet transform, various methodologies for effectual image compression and different coding techniques.

Index terms: DWT (Discrete Wavelet Transform), Neural Network, DPCM (Differential Pulse Code Modulation).

1. INTRODUCTION

Image compression plays an important role in the field of communication and multimedia [14]. The image files can be very large and can occupy a large space in memory. A gray scale image that is 256 x 256 pixels have to store 65,536 elements, and a typical 640 x 480 colour image has nearly a million elements to store. Image data comprise of a significant portion of the multimedia data and they occupy the major portion of the communication bandwidth for multimedia communication. Therefore development of efficient techniques for image compression has become quite necessary. The basic objective of image compression is to find an image representation in which pixels are less correlated. The two fundamental principles used in image compression are redundancy and irrelevancy. Redundancy removes redundancy from the signal source and irrelevancy omits pixel values which are not noticeable by human eye.
Image compression standards bring about many benefits, such as: (1) easier exchange of image files between different devices and applications; (2) reuse of existing hardware and software for a wider array of products; (3) existence of benchmarks and reference data sets for new and alternative developments. The need for image compression becomes apparent when number of bits per image is computed resulting from typical sampling rates and quantization methods.

There are three types of redundancies:

(i) spatial redundancy, which is due to the correlation or dependence between neighbouring pixel values; (ii) spectral redundancy, which is due to the correlation between different color planes or spectral bands; (iii) temporal redundancy, which is present because of correlation between different frames in images. Image compression research aims to reduce the number of bits required to represent an image by removing the spatial and spectral redundancies as much as possible.

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. [9] They are:

1. Lossless technique
2. Lossy technique

Following techniques are included in lossless compression:

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

Lossy compression techniques includes following schemes:

1. Transformation coding
2. Vector quantization
3. Fractal coding
4. Block Truncation Coding
5. Sub-band coding

1.1 DESIGN METRICS

Mean square error
MSE for monochrome images

\[ \frac{1}{N^2} \sum_{i}^{N} \sum_{j}^{N} [X(i,j) - Y(i,j)]^2 \]

MSE for colour image

\[ \frac{1}{N^2} \sum_{i}^{N} \sum_{j}^{N} \left[ (r(i,j) - r^*(i,j))^2 + (g(i,j) - g^*(i,j))^2 + (b(i,j) - b^*(i,j))^2 \right] \]
Peak Signal to Noise Ratio (PSNR)

Peak signal to Noise Ratio\[8\] is the ratio between signal variance and reconstruction error variance. PSNR is usually expressed in Decibel scale. The PSNR is a most common measure of the quality of reconstructed image in case of image compression.

\[
PSNR = 10 \log_{10} \frac{255^2}{MSE}
\]

Here 255 represent the maximum pixel value of the image, when the pixels are represented using 8 bits per sample. PSNR values range between infinity for identical images, to 0 for images that have no commonality. PSNR is inversely proportional to MSE and compression ratio i.e PSNR decreases as the compression ratio increases.

Compression Ratio (CR)

Compression ratio\[8\] is defined as the ratio between the original image size and compressed image size.

\[
compression ratio = \frac{n1}{n2}
\]

Where n1 is original image size and n2 is compressed image size.

2. LOSSY COMPRESSION TECHNIQUE

2.1 TYPES OF TRANSFORMATION

2.1.1 Discrete Cosine Transform (DCT) Based Coding:

DCT gives an approximate representation of DFT considering only the real part of the series. For a data of N values, DCT's time complexity (amount of computational time) is of the order of Nlog2N similar to DFT. But DCT gives better convergence, as compared to DFT. A given image is divided into 8 x 8 blocks and forward discrete cosine transform (FDCT) is carried out over each block. Since the adjacent pixels are highly correlated, the FDCT processing step lays the foundation for achieving data compression. This transformation concentrates most of the signal in the lower spatial frequencies, whose values are zero (or near zero). These coefficients are then quantized and encoded (which we will discuss later) to get a compressed image. The decompression is obtained by applying the above operations in reverse order and replacing 'FDCT by inverse discrete cosine transform (IDCT).[6]

2.1.2 Discrete Wavelet Transform (DWT) Based Coding

Wavelets provide a basis set which allows one to represent a data set in the form of differences and averages, called the high-pass and low-pass coefficients, respectively. The number of data points to be averaged and the weights to be attached to each data point, depends on the wavelet one chooses to use. Usually, one takes N = 2n (where n is a positive integer), number of data points for analysis. In case of Haar wavelet, the level-l high-pass and low-pass coefficients are the nearest neighbour differences and nearest neighbour averages respectively, of the given set of data with the alternate points removed. Subsequently, the level-l low pass coefficients can again be written in the form of level-2 high-pass and low-pass coefficients, having one-fourth number of points of the original set. In this way, with 2n number of points, at the n\textsuperscript{th} level of decomposition, the low-pass will have only one point. For the case of Haar, modulo a normalization factor, the n\textsuperscript{th} level low-pass coefficient is the average of all the data points. In principle, an infinite choice of wavelets exists. The
choice of a given wavelet depends upon the problem at hand. Wavelets are the probing functions, which give optimal time-frequency localization of a given signal. Merits of DCT and DWT, [6].

1. Time Complexity
Time complexity (broadly speaking, amount of computational time) of DCT is of $O(N\log_2 N)$ while many wavelet transforms can be calculated with $O(N)$ operations. More general wavelets require $O(N\log_2 N)$ calculations, same as that of DCT.

2. Blocking Artifacts
In DCT, the given image is sub-divided into 8 x 8 blocks. Due to this, the correlation between adjacent blocks is lost. This result is noticeable and annoying, particularly at low bit rates. In DWT, no such blocking is done and the transformation is carried over the entire image.

3. Compression Performance
The DCT based JPEG-93 compressor performs well for a compression ratio of about 25:1. But the quality of image rapidly deteriorates above 30:1; while wavelet based coders degrade gracefully, well beyond ratios of 100:1.

![Figure 2.1.1](image)

The different transforms provided different resolutions of time and frequency

2.2 Vector Quantization
The basic idea in this technique is to develop a dictionary of fixed-size vectors, called code vectors. A given image is partitioned into non-overlapping blocks (vectors) called image vectors.
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. [2].

2.3 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 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.

2.4 Block Truncation Coding

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.

2.5 Subband Coding

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.

3. LOSSLESS COMPRESSION TECHNIQUE

3.1 Run Length Encoding

This technique replaces sequences of identical 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 \( V_i \).

3.2 Huffman Coding

This technique for coding symbols based on their statistical occurrence frequencies. 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 frequent are assigned a relatively larger number of bits. Huffman code is a prefix code. 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 earlier stages of compression and use Huffman coding as the final step.

3.3 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.

3.4 Area Coding

This technique 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 twodimensional 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.

4. NEURAL NETWORK

ANNs that are used to model real neural networks, and study behaviour and control in animals and machines, but also there are ANNs which are used for engineering purposes, such as pattern recognition, forecasting, and data compression.[5]

1. Back Propagation Neural Network[10]

The neural network is designed with three layers, one input layer, one output layer and one hidden layer. The input layer and output layer are fully connected to the hidden layer. Compression is achieved by designing the number of neurons at the hidden layer, less than that of neurons at both input and the output layers. Image compression is achieved by training the network in such a way that the coupling weights scale the input vector of N-dimension into a narrow channel of K-dimension with K less than N, at the hidden layer and produce the optimum output value which makes the quadratic error between input and output minimum. The basic back-propagation network is further extended to construct a hierarchical neural network by adding two more hidden layers into the existing network.


The basic back-propagation network is further extended to construct a hierarchical neural network by adding two more hidden layers into the existing network. All three hidden layers are fully connected. Nested training algorithm is proposed to reduce the overall neural network training time. The neuron weights are maintained the same throughout the image compression process. Adaptive schemes are based on the principle that different neural networks are used to compress image blocks with different extent of complexity. The basic idea is to classify the input image blocks into a few subsets with different features according to their complexity measurement. A fine-tuned neural network then compresses each subset. Prior to training, all image blocks are classified into four classes according to their activity values which are identified as very low, low, high and very high activities. The network results in high complexity.

3. Multi-layer Feed Forward Artificial neural Network[10]

The network is designed in a way such that N will be greater than Y, where N is input layer/output layer neurons and Y is hidden layer neurons. Divide the training image into blocks. Scale each block and apply it to input layer and get the output of output layer. Adjust the weight to minimize the difference between the output and the desired output. Repeat until the error is small enough. The output of hidden layer is quantized and entropy coded to represent the compressed image.


Basic multilayer perception (MLP) building unit is a model of artificial neuron. This unit computes the weighted sum of the inputs plus the threshold weight and passes this sum through the activation function usually sigmoid. In a multilayer perception, the outputs of the units in one layer form the inputs to the next layer. The weights of the network are usually computed by training the network using the back propagation.
CONCLUSION

For image compression, loss of some information is acceptable. The purpose of wavelet transform is to change the data from time-space domain to time-frequency domain which makes better compression results. Among all of the above lossy compression methods, vector quantization requires many computational resources for large vectors; fractal compression is time consuming for coding; predictive coding has inferior compression ratio and worse reconstructed image quality than those of transform based coding. So, transform based compression methods are generally best for image compression. For transform based compression, JPEG compression schemes based on DCT (Discrete Cosine Transform) have some advantages such as simplicity, satisfactory performance, and availability of special purpose hardware for implementation.[4] However, because the input image is blocked, correlation across the block boundaries cannot be eliminated. In many applications, wavelet-based schemes achieve better performance than other coding schemes like the one based on DCT. Since there is no need to block the input image and its basis functions have variable length, wavelet based coding schemes can avoid blocking artifacts. Wavelet based coding also facilitates progressive transmission of images. Huffman coding is the better lossless technique compared to other technique for image compression. This scheme used to remove the redundant bits.[2]

REFERENCES

1. “A study of various image compression techniques”, Sonal, Dinesh Kumar.
4. “Advanced Digital Signal Processing” Tutorial Recent Development of Image Compression Techniques Chia-Jung Chang National Taiwan University.
5. “Artificial Neural Networks for Beginners”, Carlos GershensonC.Gershenson sussex.ac.uk”.
6. “Wavelets: Applications to Image Compression-I”, Sachin P Nanavati has joined the Scientific and Engineering Computing Group of CDAC, Pune. He enjoys trekking and bird watching. Prasanta KPanigrahi is currently with the Quantum Information and Quantum Optics Division at PRL.