Chapter I

Introduction
CHAPTER 1

INTRODUCTION

1.1 Background

In the era of ever-increasing need for efficient data transmission and storage, image compression techniques have gained popularity as means to reduce the data rate to be within the channel capacity. The science of obtaining a compact representation of an image while maintaining all the necessary information is referred to as image compression. This can be achieved by reducing the statistical redundancy and perceptual irrelevancy contained in the image, which contribute to the total number of bits required to represent them. In general compression consists of taking a stream of symbols and transforming them into codes. If the compression is effective, the resulting stream of codes will be smaller than the original symbols. The decision to output a certain code for a certain symbol or set of symbols is based on a model. The model is simply a collection of data and rules used to process input symbols and determine which codes to output. Image data compression techniques can be broadly classified into two types, namely (i) Lossless compression and (ii) Lossy compression.

Lossless compression is also referred to as reversible, entropy or noiseless coding [Cosm95] as it enables complete recovery of the original image from the compressed data. The main drawback of this method is low compression ratio. Some of the common lossless coding schemes are Run-Length coding, Huffman coding and Arithmetic coding [Alle92]. Lossy compression is also referred to as irreversible compression. Its main advantage is higher compression ratio. However, lossy compression introduces some error in the data by eliminating some information and the original data cannot be recovered completely. In fact, there is no strong limit on how much compression can be obtained using a lossy compression scheme. It depends on how much degradation in image quality an application can tolerate.
A typical transform coding consists of encoder and decoder as shown in figures 1.1 and 1.2 respectively. The encoder consists of three closely connected components namely Source Encoder, Quantizer, and Entropy Encoder. Compression is accomplished by applying a linear transform to decorrelate the image data, quantizing the resulting transform coefficients, and entropy coding the quantized values. Then the resultant compressed image is transmitted to the receiver. At the receiver side, the decoder performs the entropy decoding by the simple operation of table look up, dequantization and inverse transformation in order to get the approximated original image.

![Figure 1.1: Encoder](image1)

![Figure 1.2: Decoder](image2)
1.1.1 Sub image selection

In a transform coding scheme, the first step is to subdivide the image into non-overlapping blocks that are processed in the raster scan fashion from left to right and from top to bottom. A significant factor affecting the transform coding and computational complexity is sub image size. In most applications, images are subdivided so that the correlation between adjacent sub-blocks is reduced to some acceptable level so that the sub-block size is an integer power of 2. The latter condition simplifies the computation of the sub image transform. In general, both the level of compression and computation complexity increases as the sub-block size increases.

1.1.2 Source encoder

Source encoder is an image conversion process that transforms an image from spatial domain to frequency domain by applying a linear transformation. This transformation process aims to decorrelate the pixels of each sub-block or to pack as much information as possible into the smallest number of transform coefficients. Over the years, a variety of linear transformations have been developed which include Discrete Fourier Transform (DFT), Discrete Cosine Transform (DCT), Discrete Wavelet Transform (DWT) and many more, each with its own advantages and disadvantages.

1.1.3 Quantizer

Quantization forms the heart of any lossy compression technique. Quantization involves representation of a set of fine resolution data by a coarse resolution approximation. The image under analysis is generally divided into a number of non-overlapping blocks and any one of the unitary transformations is applied. The resulting transform coefficients are then quantized to yield a higher compression ratio. A quantizer simply reduces the number of bits needed to store the transformed coefficients by reducing the precision of those values. Since this is a many-to-one mapping, it is a lossy process and is the main source of compression process.
Quantization can be broadly classified into two types, namely Scalar Quantization (SQ) and Vector Quantization (VQ). SQ involves processing the input samples individually whereas VQ involves processing the input samples in groups. SQ involves mapping of each individual input to an output using some distortion measure. VQ involves mapping a group of inputs into a set of well defined vectors using some distortion measure. In Shannon’s Rate Distortion Theory [Alle92], it has been proved that better results would be obtained by processing a block of data as vectors instead of processing the data individually as scalars because vectors make good use of the statistics of the signal.

1.1.4 Entropy encoder

An entropy encoder further compresses the quantized values losslessly to give better overall compression. It uses a model to accurately determine the probabilities so that the resultant output code stream will be smaller than the input stream. The most commonly used entropy encoders are the Huffman encoder and the arithmetic encoder, although for applications requiring fast execution, simple Run-Length Encoding (RLE) has proven very effective. An overview on various entropy encoding techniques is reported in [Alle92] and [Nels95]. It is important to note that a properly designed quantizer and entropy encoder are absolutely necessary along with optimum signal transformation to get the best possible compression.

In general, most compression algorithms are either based on transformation first, followed by scalar quantization and coding or by direct VQ of the original image, skipping the transformation part. The main drawback of any transform technique is edge degradation and artifacts in the texture region of the reconstructed image. The objective of this research work is to obtain a common framework for supporting both edge-based as well as texture analysis methods which can be used to establish a new transform coding scheme using a set of orthogonal polynomials that preserves texture and edges in the reconstructed image.
Image compression has been widely investigated and many algorithms have been proposed [Siko05] and [Egge99]. Comparison between algorithms is often based on two aspects: compression ratio and reconstruction quality. The image compression schemes can be categorized into four subgroups according to the processing element. These subgroups are: pixel-based, block-based, subband-based, and region-based.

In transform coding, on the other hand, a linear transformation is used to map the current block into the transform domain producing uncorrelated coefficients. The energy is concentrated in few transform coefficients, typically the lowest in frequency, which are quantized and efficiently coded to perform compression. The transformation can be generally described by

$$F = T f$$
$$f = T^\top F = T^\top F$$

where $T$ is the transformation matrix and $f$ and $F$ are the original and transformed blocks respectively. $T$ is a unitary transformation and hence, the transpose is the inverse. The previous equation governs all transform coding techniques. The transformed coefficients $F$ are quantized, depending on the quantization table used, and sent in a zigzag order. In fact, the performance of any block-based transformation is upper-bounded by the optimum or KL transform. However, due to the computational complexity of KL transform and its data dependency, DCT is preferred [Jain81], [Alkh07], and [Pono07]. There is a wealth of literature on applying transform coding in image compression, like Hartley transform [Sund06], fuzzy transforms [Di07], Łukasiewicz transform [Di07a], and 3-D matrix transform [Zhan08].

Subband coding (wavelets) [Lin96] differs from block-based techniques in performing the transformation on the whole image rather than part of it. However, some techniques operate on large blocks. Hence, it has less blocking artifacts; however, the reconstructed image tends to be blurry. Nevertheless, its performance is much better than traditional block-based techniques [Kaur06] and [Brun07]. The subbands are constructed through successive filtering-downsampling (upsampling at the decoder) [Lin08]. This technique can be viewed as performing block processing in the frequency domain.
Traditional transform coding techniques are saturated as far as compression ratio is concerned [Gilg90] and [Kaup94]. Region or segmentation based techniques have been suggested to exceed this barrier [Kaup98] and to support new multimedia services [Sale99]. The term “second generation” is often used to indicate their superior performance over the previously mentioned schemes [Cerm94], [Sale94] and [Siko05]. It has been shown in [Bigg88] and [Ran95] that at higher rates, reconstruction quality of region-based techniques exceeds that of DCT.

1.2 Orthogonal polynomials framework for image compression

In literature several transforms have been employed for the compression of images, of which Discrete Cosine Transform (DCT) is popular. However, the DCT based compression scheme poses problems of slow decay rate of DCT coefficients and introduction of blocking artifacts. Hence in chapter 2 of the thesis, a new image coding framework is proposed, based on orthogonal polynomials transformation.

Using the set of orthogonal polynomials a complete set of basis operators are constructed and a polynomial transform coding which separates the signals due to low level features from noise by proposing statistical testing procedures is described. The proposed transform coding has low computational complexity because it is configured as an integer transform. The proposed transform also has very good compression ability or separation-ability. The degree of this separation-ability of the transform which is based on the statistical testing procedures is robust because it is no longer an intuitive notion as in the other transform coding schemes.

1.3 Transform coding of monochrome images with orthogonal polynomials

The proposed coding framework described in the previous section utilizes statistical procedure to separate the signals due to low level features from the noise and provides good quality of reconstructed image. In order to get higher compression, in chapter 3 of the thesis, a new image coding technique is proposed, based on Orthogonal Polynomials Transformation (OPT) which is configured as an integer transform with suitability to use JPEG entropy coding.
To propose this orthogonal polynomials transformation, a class of orthogonal polynomials is utilized for designing a point spread operator. Then the polynomial basis operators of different widths are designed for decoding purpose. The image to be compressed is partitioned into non-overlapping blocks and the proposed transformation is applied on the image region under analysis to obtain the transform coefficients. These coefficients are then scale quantized and the resultant quantized coefficients are subjected to entropy coding as in JPEG baseline system. At the receiver, the decoder performs entropy decoding and scale dequantization to obtain the approximated transform coefficients. The transform blocks thus obtained are subjected to inverse transform using orthogonal polynomial basis operators.

1.4 Texture characteristics with orthogonal polynomials for image coding

The transform coding technique described in the previous section provides good quality reconstructed images at low bit rates. However, the texture regions in the image are not taken care while compression. Hence, in chapter 4 of the thesis, an image coding technique based on texture characteristics with orthogonal polynomials is proposed.

The proposed scheme is based on the model that represents textures using points spread operator relating to a linear system. In the proposed texture-based image coding scheme, the encoder first identifies textured regions, which are then analyzed to produce the model features. These are later transmitted to the decoder that produces a synthetic texture based on these features through the synthesis stage. Texture synthesis is also used to remove undesirable artifacts in image obtained after compression-decompression process. The proposed algorithm extends for this purpose by replacing artifacts in the textured backgrounds. This algorithm is easy to use and requires only a sample texture as input. The key advantage of this approach is that it can efficiently generate high quality textures.
1.5 Edge preserving image coding with orthogonal polynomials

The main drawback of the proposed image coding techniques described in the previous sections is edge degradation. Since edges contribute a lot to the perceptibility of an image, the degradation of edges is undesirable. To overcome the problem of edge degradation, a new edge preserving image coding technique with the orthogonal polynomials transform coefficients is proposed in chapter 5 of the thesis.

Existing image compression techniques are able to achieve good compression ratio. Higher compressions have been achieved using lossy techniques that remove visual information that is not perceived by the human eye. It can be observed that the higher compression ratio can only be achieved at the expense of higher distortion that degrades the quality of the reconstructed image. To meet the challenges of achieving high compression ratio with good quality of reconstructed picture, edge based compression schemes have been employed, so as to preserve high image quality. With this end in view three well known edge based coding schemes namely (1) Edge preserving image coding with Vector quantization (VQ), (2) Edge preserving compression using neural network, and (3) Edge preserving image compressions with wavelet transform are reported in the literature.

In this work, an algorithm that captures the locations of important edges with an edge detection step using orthogonal polynomials at the encoder is proposed. The edge information captured are transmitted and are effectively removed from the images during the forward transform coding and reinserted at the decoder during the inverse transform. The proposed method not only well preserve the object boundaries leading to improved performance but also achieves better compression when compared to existing methods.
1.6 Adaptive down sampling algorithm for image coding with orthogonal polynomials

A new low bit rate coding scheme with down sampling based on orthogonal polynomials is presented in Chapter 6 of the thesis. The proposed coding scheme achieves better coding quality by down sampling the image prior to compression and estimating the missing portion after decompression. This scheme is based on the adaptive decision of appropriate down sampling directions /ratios and quantization steps in order to achieve higher coding quality with low bit rates with the consideration of local visual significance. The full resolution image can be restored from the orthogonal polynomials based transform coefficients of the down sampled pixels so that the spatial interpolation required otherwise is avoided. The experiments have demonstrated better improvement in Peak Signal to Noise Ratio (PSNR) over the existing techniques before the critical bit rate. In addition, the adaptive mode decision makes the critical bit rate less image-independent and automates the switching coders in variable bit-rate applications.