CHAPTER - 1

INTRODUCTION

CHAPTER - 1

INTRODUCTION

1.1 REVIEW OF TRANSFORM METHODS OF PICTURE PROCESSING

Orthogonal transforms have found many applications in digital signal processing. The pace of activity in this area has been greatly accelerated due to the impact of high speed digital computers. In particular, the emergence of mini and microcomputers, has acted as a catalyst to the general field of data processing [1 - 7]. The main applications of orthogonal transforms relate to image processing, speech processing, feature selection and pattern recognition, character recognition, digital filtering, spectral analysis, data compression etc. Orthogonal transforms have been used to process various types of data including speech, sonar, radar, biological, biomedical, forensic sciences, satellite TV pictures, weather forecasts, electron micrograph etc. [8 - 11].

Image processing involves spatial filtering, image coding, image restoration and enhancement etc. The high energy compaction property of transform techniques has been taken advantage of in reducing bandwidth, improving tolerance to error and achieving sample and bit rate reduction [7].

Image coding systems may be classified in a number of ways. The coding systems are classified as to the type of images, pictorial or non-pictorial [2]. Some coding systems are designed to process only binary valued images i.e., computer graphics data, or pure black and white data [12], while others accept continuous tone images.

As is well known, digital coding methods may be divided into three types -

(a) methods that utilize point processing,
(b) techniques based on spatial coding concept, and
(c) hybrid coding, combining (a) and (b).

Those utilizing point processing include pulse code modulation (PCM) [13], differential pulse code modulation (DPCM) [14] and predictive image coding [14-15]. In the basic PCM coding system the continuous signal is quantized. Each quantized sample is assigned a constant length group of bits called a code word. Various quantization schemes have [16] been developed in recent years for reduction of quantizing errors.

In a predictive image coding system the value of each scanned pixel is predicted based on some previous history of scanned elements. Then the predictive estimate is subtracted from the actual pixel value and the "difference pixel" is quantized, encoded and transmitted. At the receiver the quantized received signal is used for the reconstruction of the image signal. Various schemes for bandwidth reduction in such cases have been reported in the literature [15,17,18].

The basic DPCM concept has also been incorporated by forming the prediction signal from the linear combination of several previously scanned signals and from the previous lines. Both theoretical and experimental studies have shown that mean square error measure and subjective appearance can then be improved [18].

In digital image coding techniques based on spatial coding concept, a sequence of pixel values is considered rather than the value of a single pixel. This may be considered as an extension of the concept of predictive coding. Interpolative coding systems are based on approximation techniques whereby a sequence of pixel values is fitted by a continuous function [19].

The transform coding image processing technique is indirect and is similar to feature selection schemes [6]. A unitary transformation is performed on the image data to produce a set of transform coefficients which are then quantized and coded for transmission. Transform coding has proved to be an effective way of image coding for monochrome and multispectral images, for both still pictures and real time television [20-23]. Common types of unitary operators are :
Fourier transform, even and odd cosine transform, sine transform, Hadamard transform, Haar transform, Slant transform, discrete cosine transform etc. [1,2].

The concept of coding and transmitting two-dimensional Fourier transforms of monochrome image rather than the image itself was introduced in 1968 [8]. The basic concept of Fourier transform coding is that for most natural images, many of the transformed coefficients are of relatively low magnitude and therefore can be discarded [8]. Hadamard transform coding has been utilized in place of Fourier transform with a considerable decrease in the computational requirement [24].

Investigations on discrete Karhunen-Loeve (KLT) and Haar transforms for image coding have also been carried out [2]. KLT provides minimum mean square error coding performance but requires statistical knowledge about image source and does not possess a fast computational algorithm. A Haar transform [25] has the attribute of an extremely efficient computational algorithm but usually results in relatively large coding errors. Slant transform (developed on slant vector) has been utilized by Pratt et al. [26] and has been found to be quite efficient for effectively representing gradual brightness changes in an image. Ahmed [27] has used cosine transform which possesses a fast algorithm and approaches the efficiency of KLT for Markov process image data. Sine transform with similar properties has been suggested by Jain [28].

The Singular Value Decomposition (SVD) [2], [29-30] transform is a two-dimensional unitary transform based on the singular value decomposition of the image matrix.

A unitary transformation of a (N x N) pixel image into (N x N) transform coefficients requires of the order of N^4

computation operations including addition and multiplication. Any transform to be useful must have a fast computational algorithm, The key to this is the ability to subdivide the computational task to a series of computational steps so that the partial results from the initial steps be utilized in the subsequent steps. Examples are the butterflies used in the case of FFT and Hadamard transform [1].

The basic premise of a two-dimensional image coding system is the clustering of the image energy in a few transform domain components; the two-dimensional transform has an energy distribution more suitable for coding than the spatial domain representation [31]. The transform domain components are selected based on the relative magnitude by either Zonal sampling, or Threshold sampling technique [2].

In Zonal sampling the reconstruction is made with a subset of transform samples lying in certain prespecified geometric zones, usually the low-frequency coefficients. For analogue transmission the amplitude of each component in the zone is transmitted, while for digital transmission each component in a zone is quantized and assigned a binary code word. The number of quantization levels is usually made proportional to the estimated variance of the components and the number of code bits made proportional to its expected probability of occurrence [32].

With threshold sampling the image reconstruction is made with a subset of the samples that have amplitudes larger than a specified threshold. Since the locations of the significant samples must be communicated, threshold sampling is usually employed only in digital links.

Two-dimensional transform coding provides a lower mean square error at low bit rates but at the cost of complexity of implementation. A differential predictive (DPCM) coder, on the other hand, is very simple to implement. Hybrid transform-DPCM image coding methods have been proposed which possess many of the theoretical attributes of two-dimensional transform coding methods [33].

It is known that the utility and effectiveness of a particular transform depends on the type of the signal, in particular its statistical properties [28]. Image processing systems make use of the correlation and spectral character of the images, continuous or discretized. In deterministic image representation, the image is considered to be three-dimensional. The spatial energy distribution is a function of spatial coordinates x and y and also time t [34].

Digital image processing is based on the conversion of continuous image field to an equivalent digital form obtained by spatial sampling of the image and quantizing them. For the purpose of representation of discrete images it is often helpful to regard the sampled array as an (N1 x N2) array element matrix.

It is convenient to convert the image matrix to vector form by column (or row) scanning and then stringing the vectors together in a long vector. The vector representation of data matrix [F] is given by the stacking operation. The moments of a discrete image process may be conveniently expressed in vector space form

$$E\left\{[F]\right\} = E\left\{[f(n_1 , n_2)]\right\},$$

where $E\{\cdot\}$ is the expectation operator.

The correlation function of the image array is given by

$$R(n_1,n_2,n_3,n_4) = E\left\{f(n_1,n_2)\ f^*(n_3,n_4)\right\}$$

where * stands for conjugate.

If the image array is wide sense stationary then the correlation function can be expressed as

$$R(n_1,n_2,n_3,n_4) = R(n_1 - n_3 , n_2 - n_4)$$

In stochastic representation of continuous images one describes the process in terms of the joint probability densities and corresponding auto-covariance spectrum. An image field is often modeled as the sample of first order Markov process for which the correlation between the points of the image field is proportional to their geometrical separation.

The quantizer used in most digital work is the companding quantizer which consists of a cascade of non-linear transformations on an uniform quantizer where the non-linear function is so arranged as to minimize the distortion on decoding [32].

The objective of image coding is to code an image for storage or transmission with as few code symbols as possible that maintains the quality of the coded image at some acceptable level. There are two aspects of image quality - image fidelity and image intelligibility [2]. Image fidelity compares the departure of the processed image from a standard image. A common fidelity measure is the mean square error (MSE) defined elsewhere in the thesis. Entropy and rate distortion functions are also used in establishing performance bound for coding and restoration.

1.2 SCOPE OF THE THESIS

It is desirable to study the relative efficiency of different transforms in handling real time data, and also if new transforms can be derived providing improved performance. A problem encountered in bandwidth compression is that the quality of the picture is degraded unless some of the discarded coefficients are extrapolated [35]. Most transform techniques are applicable only if the random field may be viewed as the output of a discrete linear filter driven by a white noise source. Such a model would never show sharp variations of intensity. It is therefore necessary to devise special techniques for reproducing the edge information . It should be mentioned that forward and inverse transforms including selection scheme are equivalent to a linear filter which causes smearing of the images. Efforts must therefore be made for removing, at least partially, the effect of such filtering through use of appropriate restoration techniques.

A part of the study described in this work is devoted to finding some new transform. It is known that the KLT operators are calculated from the eigen vectors of the covariance matrix of image data. KLT coefficients are therefore time varying. A simple approach is to assume that the data are generated by a first order Markov process. It is known that the covariance matrix of first order Markov process is of Toeplitz from [1]. It has been shown that Discrete Cosine Transform (DCT) elements can be derived by making use of this property [27-28]. We have used the circulant matrix which is a special form of Toeplitz matrix corresponding to Markov process to find the values of the transform elements of a proposed new transform.

A second transform was found by orthogonalizing the covariance matrix. In one variation of this scheme the element amplitudes of the row vectors were made to assume only ternary values. Some transforms derived by combining the above with Hadamard transform are also studied [36]. The relationship between statistical picture model and the transforms which best handles the corresponding signal is discussed. Results of these studies are described in Chapter 2. These show that mean square error comparable to that obtainable in DCT is achievable with some of the new transforms devised.

In order to achieve bandwidth compression in image processing by orthogonal transform technique it is necessary to retain and transmit only a few transform coefficients and discard the others. The quality of the picture can be improved and the error in reconstruction could be reduced if the components discarded in the process of transmission could be reconstructed from the received components. Such extrapolation is possible only if there is finite correlation between the received transform domain coefficients. This correlation depends on the nature of the image and the transform being used [35,37].

If simple models of the image, such as two-dimensional Markovian, are appropriate for the picture, one may formulate simple simultaneous equations for extrapolation of discarded components [38]. This technique is described in Chapter 3. An extrapolator can also be used in the transmitter in an essentially predictive encoding mode in which the actual transform domain components are compared with the values extrapolated locally at the transmitter and the difference transmitted [38]. It is shown that the use of such a technique results in considerable reduction in the number of bits per pixel and mean square error.

An important part in the design of the hardware for an extrapolator involves setting up of the simultaneous equations relating the transform domain covariances. Considerable simplification results if specific use is made of the expected correlation properties of a particular transform. Fortunately, many transforms e.g., Hadamard, Slant, DCT, possess symmetry in the transform coefficients which can be used in the estimation and calculation of transform domain covariances. A method based on symmetry property of the transform operators to directly locate the vanishing terms in the transform domain covariances is described in Chapter 4, [39]. A scheme which uses transform coding in the predictive mode based on the utilization of the correlation properties of the transform domain coefficients of neighbouring blocks is also discussed. The selection of the transform domain coefficients for predictive transform coding is made by utilizing symmetry property of the transform components [40].

Bandwidth compression by transform method is feasible only if picture elements are highly correlated. When the picture cannot be modelled as a simple Markov process the reconstructed picture shows considerable errors, particularly when there is an abrupt change in image intensity [41,42]. Schemes based on separate coding of smooth correlated regions and edge information for one- and two-dimensional signals are investigated in Chapter 5. For bandwidth compression Hadamard transform is used and for the extraction of edge information Haar transform and difference coding schemes have been used. Results show an improvement in mean square error and error probability distribution with small increase in transmission bandwidth requirements.

The problem of degradation or blurring of pictures because of linear filtering and the implementation of some restoration schemes have been discussed in Chapter 6. Any transform coding scheme which discards transform components and uses reduced number of bits for representation for the purpose of bandwidth reduction suffers from signal degradation. The blurring that results due to an equivalent low pass filtering depends on the location of the picture elements in the data block. Severe distortion seems to occur when there are rapid changes of the amplitude.

An attempt has been made to find equivalent filter which represents the effect of bandwidth compression. As a first step in finding the appropriate restoration filter, a study has been made on the effect of pre-emphasis filters which emphasize the high-frequency components before the input is applied to the transform system. Based on the spectral estimation of the transform system, feed forward processors have also been developed to be used in the receiver for removing the picture blur. The performance of the transform coding scheme both in the case of maximum variance selection scheme and variable bit allocation with these filters has been studied in detail.

Linear feed forward filters are however limited in their efficacy in restoration. A study has been made on non-linear restoration schemes based on maximum entropy. The constraints that have been used in the restoration scheme are based on the quantum optical modelling of picture formation [43]. The formulation is similar to Frieden's developments based on Jaynes' techniques [44] of maximum entropy estimation of positive probability density function where the most likely object distribution is found subject to specified constraints such as the values of the image at some specified points [45].

Conclusions and scope of further work are included in Chapter 7.