CHAPTER 1

INTRODUCTION

There have been extensive effort and continued growth of interest in the development of picture coding systems during the past three decades [1-7]. Two factors have contributed to the growth of the field: (i) a tremendous expansion of the number of applications ranging from space exploration to madicine, involving transmission and storage of pictorial information and (ii) the development of cheaper and more powerful hardware for implementing complex coding schemes.

Applications which demand picture transmission and storage have been increasing steadily. Entertainment television has proliferated through broadcast and cable media. Besides broadcast television, there is a large amount of point-to-point transmission of pictorial information taking place today. For example, terrestrial television networks and satellite links transmit live programmes as well as other picture material around the world. With rising cost of travel, business establishments have started taking advantage of the video teleconferencing services [5]. Communication of graphics has grown in many business applications. Facsimile transmission of newspapers and printed material has now become widespread. Satellites are continuously beaming to earth weather photographs and earth-resource pictures. Video systems are used in space exploration and in medical applications. There are also a number of military applications like reconnaissance, control of remotely piloted vehicles, etc., where video systems are used [8].

Although picture information is originally analog, there are many advantages in digitizing it for transmission, storage and processing [5-10]. A digital signal can be transmitted at a lower Signal-to-Noise Ratio (SNR) for the same quality of the reproduced picture. It can be easily multiplexed, regenerated or encrypted and offers greater flexibility. It can be easily stored in computer storage media and retrieved easily. Complex manipulations of the signal which are vital in picture processing can be done more easily in the digital domain than in the analog domain. The resulting signal may be converted to analog form for transmission over an analog channel, if such a need arises.

With the advent of LSI technology, the digital integrated circuits for processing and storage are becoming cheaper every year. The prospect of VLSI parallel prcessors [11] and large RAM memory [12] seem to make complex manipulations of picture data in real-time economically viable in the near future. Digital representation of picture signal permits efficient use of the more reliable and cheaper digital integrated circuits for processing and storage.

The major diadvantage of having the picture material in digital form is that it requires a very large transmission bandwidth or storage space. For example, a video signal of broadcast quality has a bandwidth of over 5.0 MHz in analog domain. If it is sampled at Nyquist rate and digitized using 8 bits/sample, the bit-rate of the resulting digital signal will be around 80 Mb/s, requiring a transmission bandwidth of atleast 40 MHz (assuming PSK modulation of 2 bits per Hz) [4,13]. The aim of efficient coding is, therefore, to narrow the transmission bandwidth as much as possible. Such compression can reduce the cost of transmission in commercial and space applications. It permits more parallel television channels to be transmitted through a communication link and/or a reduction of transmitted power [1]. It can reduce the susceptibility to interference in military applications [8]. It makes digital storage of the picture signals for computer retrieval much more viable and compact.

Another area of application of efficient coding is where large storage of picture material is required. For example, the number of X-ray pictures in hospitals, picture data bases like enigineering drawings and finger prints, etc., are very large and their storage and retrieval is extremely difficult [9]. Application of efficient digitisation will remedy the situation by reducing the storage requirements.

The task of digitising picture signals resulting in bandwidth compression can be divided into three stages [9,15] : (i) A reversible operation, where the signal is transformed into a more efficient form by removing or reducing redundancy. Examples of this stage are, generation of prediction error signal and conversion of picture signal into transform domain. (ii) An irreversible operation, where the accuracy of representation is reduced within the required picture quality objective (in general this operation is quantization). (iii) A reversible encoding operation to reduce the number of bits that must be transmitted, provided the symbols from the quantizer have a non-uniform probability distribution. For example, a Huffman code which assigns shorter code words to symbols that occur more frequently and longer code words to symbols that occur less frequently, will reduce the number of bits required to transmit the symbol sequence [13]. While the second stage is a necessary part of a coding scheme, the other two stages may or may not be used depending upon the complexity permitted and the amount of compression required.

Picture coding algorithms can be broadly classified into three categories : (i) Spatial domain coding, where the correlation present in the input picture signal is utilised to encode a single picture element (pel) or a group of pels at a time in a straight forward manner. Predictive coding systems like Differential Pulse Code Modulation (DPCM) and Delta Modulation (DM) belong to this category. Other examples of spatial domain coding are vector quantization [16,17] and sub-band coding [18]. (ii) Transform coding, where the bandwidth reduction is achieved by an energy saving transformation of the picture signal into another domain such that maximum information is compacted into a small number of samples [9,14,19]. Examples of this category are : Discrete Fourier Transform (DFT), Discrete Cosine Transform (DCT) and Hadamard Transform [19-21]. The transform coefficients are quantized with sufficient accuracy to get the required picture quality. (iii) Hybrid coding, which is a combination of the first two methods [22]. Typical examples of this category are DFT/DPCM coders and DCT/DPCM coders.

Transfrom coding systems achieve superior performance at low bit-rates. They distribute coding degradation in a manner less objectionable to the human viewer, show less senstivity to signal statistics and are less vulnerable to channel noise. But these systems are inherently complex in terms of memory requirement and the number of operations required per pel and produce delay in transmission. On the other hand predictive coding systems achieve a good performance at higher bit-rates. They do not require large memory and are relatively simple to implement in hardware. However, predictive coding systems are sensitive to variations in signal statistics and tend to propagate of channel errors on the transmitted picture [7,9]. In hybrid coding systems the strong points of the two methods are combined to obtain an improved coding performance. Their performance is close to that of transform coding at moderate bit-rates, while the complexity lies between that of tranform coding and predictive coding [22].

Predictive coding is the most widely used technique for encoding picture signals in real-time applications because of the minimum memory requirement and the ease of hardware implementation. Here, using the correlation present in the input picture signal the redundancy in the input is removed by predicting the present sample as accurately as possible from the past reconstructed (transmitted) samples. The present pel and its predicted value are used to generate a new signal (prediction residual) which is substantially decorrelated. The decorrelated signal is quantized and transmitted. In a DPCM coder the difference between the present pel and its predicted value (predcition error) is quantized and transmitted. DM is a special case of DPCM, where the picture signal is oversampled and the prediction error is quantized into two levels.

For an improved performance of a predictive coding system, it is important to design an efficient predictor and an efficient quantizer. An efficient predictor would reduce the variance of the prediction residual to a minimum value and would adapt to the variations in the picture statistics. Adaptive prediction has not received much attention because of the difficulty involved in real-time measurement of the local statistics of the picture that affect the prediction process. An efficient quantizer would match the statistics of the prediction residual and produce minimum distortion in the reconstructed signal. For example, in a DPCM system the probability distribution of the prediction error is fairly well behaved, so that an optimum quantizer can be designed to match its Probability Density Function (PDF) [21]. But the variance of the prediction error may change with the local statistics of the picture. An efficient quantizer should adapt to these variations in the variance of the prediction error.

The performance of a predictive coding system is measured by the amount of distortion introduced in the reconstructed signal. Since predcition is a reversible process, no distortion is introduced in the reconstructed signal by this process. Quantization, however, is an irreversible process and is the source of distortion in any predictive coding system. In applications where the reconstructed picture is 'viewed' by a machine, the quantizer is designed to produce minimum Mean Square Error (MSE) in the reconstructed signal. But, nearly all pictures are viewed and judged by human observers. It is of great importance, therefore, to design the quantizer on the basis of such psychovisual criteria that introduce minimum perceptible error in the reconstructed picture signal. The most important property of the human visual system that is exploited in the design of the quantizers is the variation of the sensitivity of the eye to impairments introduced in the picture signal.

The two broad categories of the sensitivity variations of the human eye to the noise introduced in the picture that are used in the design are [9,15,24]:

(i) Picture independent variations viz the visibilty of noise decreases with increasing frequency of noise.

(ii) Picture dependent variations viz the visibility of noise decreases with (a) increasing luminance level (luminance dependence) and (b) inceasing picture detail (detail dependence).

Any one of the three aspects or a combination thereof can be used in designing the quantizer that produces minimum perceptible distortion in the reproduced picture.

This thesis deals with the study and design of some new predictive coders and quantizers that introduce minimum perceptible distortion in the reconstructed picture and are used in certain spatial domain coding systems. A novel predictive coding system called Ratio Pulse Code Modulation (RPCM) system has been proposed. In the RPCM coder, a function of the ratio of the present sample of the input picture signal and its predicted value is quantized and transmitted. It has been shown in the thesis that the system has the following advantages over the DPCM system : (i) It has a much larger dynamic range than an ADPCM system using a Jayant type adaptive quantizer [25]. (ii) It exploits the luminance dependent visual sensitivity effect to produce, subjectively, a better picture than the ADPCM coder for the same SNR of the reconstructed signal. While it produces an SNR that is almost equal to that of the ADPCM coder at the same bit-rate, it distributes the reconstruction error in such a way that areas of higher luminance level have more error and areas of lower luminance level have less error. This error distribution results in subjectively better quality pictures.

A new scheme where the RPCM coder uses a special type of quantizer called Centre-Clipping Quantizer (CCQ), has been proposed. It has been found that accurate quantization of the high amplitude portions of the prediction residual of the RPCM coder is necessary for achieving good visual reproduction of the picture at low bit-rates. Moreover, even if the prediction residual is severely centre clipped, very little visual distortion is introduced in the reconstructed picture. In our scheme, the prediction residual is centre clipped by setting it to zero if its absolute value is less than a certain threshold. The centre-clipped prediction residual is then quantized by a multilevel quantizer. The entropy of the output of the CCQ is very small because of the centre-clipping operation. The quantized signal is, therefore, entropy coded to reduce the final bit-rate in the channel. It has been shown that at an effective bit-rate of 1.58 bit/pel, the overall subjective quality of the picture reproduced by the scheme is almost equal to that of the ADPCM coder at 3 bit/pel.

Further, another scheme where a DPCM coder uses a two-dimensional (2-D) polar quantizer, has been proposed and its performance studied. Here, the prediction error signals of two adjacent lines of the picture are grouped together into a pair. Each pair of the prediction errors is then transformed into a polar format. The magnitude and the phase of the transformed signal are quantized separately using optimum quantizers that produce minimum MSE. The PDFs of the magnitude and phase of this polar signal are more suitable for quantization than that of the prediction errors themselves. This scheme produces about 20 percent saving in transmission bandwidth over a conventional DPCM coder, in terms of both the SNR of the reconstructed signal and the subjective quality of the reproduced picture.

Finally, an attempt has been made to design a subband coding system for picture signals, where the frequency dependent visual sensitivity variation is exploited to produce minimum perceptible distortion in the reproduced picture. In this coder, a raster scanned picture signal is divided into a number of subbands and each subband is coded by an ADPCM or an APCM coder. The bit allocation to each subband has been done in such a way that the masking of the quantization error spectrum by the combined effect of the signal spectrum and the frequency dependent visual sensitivity is maximum. The subjective quality of the reconstructed picture of the subband coder is better than that of the full-band ADPCM coder for bit-rates ranging from 1 bit/pel to 3 bit/pel where the study has been made.

The organisation of the thesis is as follows : In Chapter 2, we have reviewed the three major categories of picture coding techniques mentioned earlier in this chapter. In addition, we have reviewed certain work in speech coding which are relevant to the work reported in this thesis. In Chapter 3, we have described the basic RPCM system and presented simulation resuslts to substantiate the merits of the system. In Chapter 4, we have described the CCQ and presented results of coding picture signal using RPCM-CCQ combination. In Chapter 5, we have described the DPCM coder with the 2-D polar quantizer and explained a procedure for designing the polar quantizer. We have presented simulation results for various bit-rates. Chapter 6 deals with different aspects of the design of a subband coder for picture signal and presents the corresponding simulation results. Finally, in Chapter 7, we have drawn conclusions on the performance of each scheme and provided suggestions for future work.