CHAPTER 1

INTRODUCTION

As images are getting larger in size, storage and transmission of them take much more storage space and transmission time on communication mediums, necessitating compression of images in many applications of computer vision ranging from medical to video phone. Image coding aims to reduce as much as possible the number of bits necessary to represent the image while preserving the quality and intelligibility required for the given application. The level quality and intelligibility required vary widely, depending on the application. In some applications like medical imaging, it is important that reconstructed image is the exact duplicate of the original. Such image-coding technique that preserves all information in the image and allows exact reconstruction of the original one is said to be information-preserving. Some applications such as video phone, where the exact reconstruction is not the primal goal, do not require information-preserving methods. They allow small amount of degradation on the reconstructed image. Such techniques are called information-lossy. For information-lossy techniques, it is important to describe the amount and the type of degradation at the reconstructed image [1]. Unfortunately, there is no subjective way of measuring the degradation introduced by the method. When the visual system is human, then it seems that best test measure is human brain. In other case, when the visual system is another visual machine system, the amount of degradation, deteriorating the task performed by machine vision system, is measured by reduction in performance and requires extensive examination on the system.

A typical image coding environment is shown in Figure 1.1. Coding has two parts as depicted in the figure. Purpose of the first part, called source coder, is to encode the digital data such that encoded data occupies less space. Second part, called channel coder, transforms the bit streams encoded by source coder into a form more suitable for transmission over a communication channel through appending some error detection/correction bits. At the receiver, transmitted image is reconstructed by image coder. In this study, we are interested in only source coding and decoding.

Image coding techniques can be classified into *three* groups according to what they actually code. In waveform image-coding techniques an image itself or some simple variation of that such as difference in intensity at consecutive pixels is coded. In transform-based image-coding techniques, an image is transformed into another domain such as frequency domain and transform coefficients are coded. In model-based image-coding techniques, an image is modeled and model parameters are coded. Waveform image coding techniques has used information theory end its results.

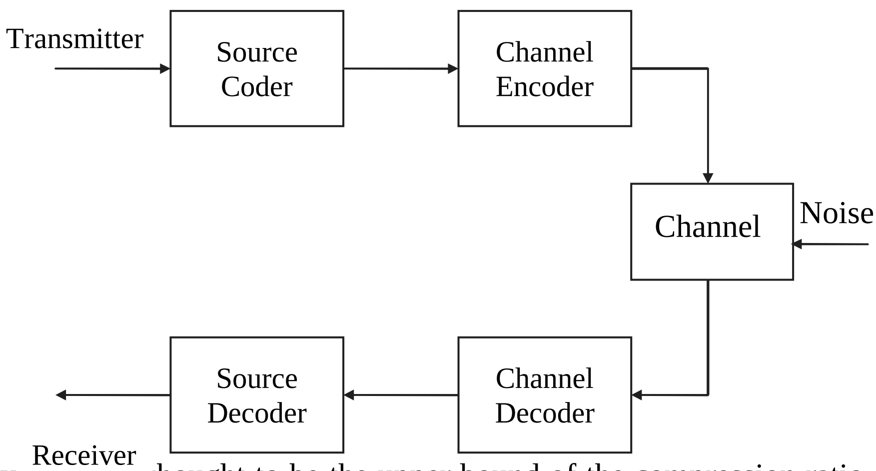

Figure 1.1 Typical environment for Image Coding

The entropy has been thought to be the upper bound of the compression ratio of any entropy codin ... ceed.

$$H(u)=\sum_{i=1}^{N} P(u=u_i)\cdot\log_2\left(\frac{1}{P(u=u_i)}\right)$$

It is clear that images are not random and locally homogeneous. That means gray levels are similar over a small neighbor. First-order entropy is defined as

$$H(u_k|u_{k-1})=\sum_{i=1}^{N}\sum_{j=1}^{N} P(u=u_i,u=u_j)\cdot\log_2\left(\frac{1}{P(u=u_i|u=u_j)}\right)$$

where $P(u=u_i|u=u_j)$ is the conditional probability. This is considered as the average information content of u_k in the case that u_{k-1} is known. Second and higher order entropy can be defined similarly. It can be shown that Huffman coding for higher order entropy approaches to the first-order entropy irregularly.

Waveform coding methods based on the decorrelation of images followed by quantization and entropy coding (Huffman Coding). Decorrelation is obtained by prediction (i.e. Differential PCM). Compression ratio is about 4:1 for waveform coding techniques.

In transform coding methods, an image is transformed into such a domain which is much more appropriate for coding. The transforms have the common property called *energy compaction* which states that transformation coefficients are distributed in a small fraction of the supporting domain. Compression ratio is about 20:1 for transform-based coding.

Actually, entropy of the image is unknown and depends mostly on the model used for the source. Third class of image-coding techniques are *model-based* approaches to the image coding problem. Model parameters are extracted and coded, at the receiver side reconstruction is obtained by using these model parameters. There are two groups of approaches in the class. One group of method makes use of local operators, convolving the image with impulse response of 2-D filters or filter banks. These filters or filter banks are designed so as to extract local feature. Then these local features are combined to obtain the messages to be coded. Nonstationary predictive coding is an example of this group of methods. Second approach uses completely different method. They utilize edges and textures to model the image. First step in this method is to detect the edges or segment the image into regions. Contour or region boundary locations, brightness at edges or regions and contrast are coded. Reconstruction of the original image is obtained by solving a *diffusion* (*heat*) equation. Compression ratio for model-based coding is dependent heavily on the model and model precision.

In this study, a model-based image coding method called “centipede model” is developed for compression of images using primitives based on edge location, intensity, contrast and scale. Edges, correspond to object boundaries where sharp changes occur due to some physical aspect of an image such as surface reflectance or illumination, are detected by using Generalized Edge Detector (GED). GED [2] introduces a $\lambda\tau$–space representation of images and consequently edges. On the $\lambda\tau$–plane, a point associate to two filters denoted by $R(\lambda,\tau)$ and its first-order derivative $G(\lambda,\tau)$ whose shapes are determined by τ and scales are determined by λ. One can obtain well-known edge filters by setting these parameters appropriately. Then edges are traced by following adjacent edge elements to produce distinct contours. Intensities on edges and contrasts are modeled by polynomials. We have shown that using these parameters and minimizing hybrid energy functional do not yield a powerful representation of the original image. It results in a blurry reconstructed image with many artifacts. We utilize edge scales in the form of widths to overcome the problem. Width is defined as the distance at which the difference in intensity along the normal direction is lower than a given threshold.

Coding of all these parameters is a bit-consuming operation. Since they change smoothly in a small neighborhood, they can be approximated well by polynomials, instead of model parameters, the coefficients of the polynomials are coded. Edge locations are coded by constructing differential chain code followed by Huffman coding and end points are coded in the form of difference between lexicographically ordered points. A reliable approximation to the original image from the sparse information is obtained by solving the hybrid energy functional which constructs $\lambda\tau$-space, where λ represents smoothness of the image and τ represents the continuity of the image.

The thesis is organized as the following. In chapter two, a classification for image coding methods is presented based on what they code and particularly, existing model-based approaches are investigated. "Centipede Model" is explained at the end of this chapter. Chapter three develops edge detection, contour tracing, filtering and coding methods. Chapter four goes on the extraction of the centipede model (width and contrast). Since the parameters are modeled by polynomials, curve fitting with polynomials is introduced by minimizing MSE. Quantization and coding of polynomial coefficients are given at the end of the chapter. Experimental results are discussed and conclusions are drawn in chapter five.