I. Introduction

Image coding has been an active area of research for many years. It is defined as the process in which the number of bits needed to represent an image is minimised subject to some quality criterion. Transform image coding has been found to be an efficient means by which to achieve this minimisation [1,2,4,7]. Thus, this thesis will not investigate image coding in general, since it is a very wide field, but a thorough investigation of all aspects regarding *transform image coding* will be given[1].

Intraframe or single image coding of video images has several applications including reconnaissance, remote sensing, and is used in interframe coding for encoding of difference images. The transmission of images over a bandwidth limited channel places constraints on the time in which the images can be transmitted. Lower average bit rates result in less data to be transmitted and consequently faster transmission rates.

Image coding techniques can be classified as either noiseless or minimum distortion. In noiseless image coding, the image is reconstructed perfectly from the coded information. The compression ratio that can be achieved, from information theory, is limited to the entropy of the image and is typically more than two bits per pel (pixel or picture element). This is not sufficient since the average bit rate

1 For a more general discussion of image coding the reader is referred to W.K.Pratt "Digital Image Processing" John Wiley & Sons 1978.

required by most applications is less than one bit/pel. Compression ratios of this order can only be achieved at some expense in fidelity. Shannon introduced the concept of the "rate distortion function" specifically to give a mathematical framework for this kind of problem. The rate distortion function determines the minimum channel capacity required to transmit data with a specified distortion percentage subject to a suitable distortion measure.

The desired bit rates, that the algorithms will have to achieve, is between 2.0 and 0.5 bits/pel. This means that in most cases the algorithms will introduce a certain amount of distortion, and thus can be categorised under the rate distortion theory. The rate distortion function is rarely used in image coding since no representative statistical model exists for images and the definition of a suitable distortion criteria is troublesome. However, it can be used to find an absolute upper bound on the achievable performance of any image coder for a given set of images. Using the rate distortion theory, images will be generated that should give one a subjective idea of what the optimum codec will be able to achieve. These images can then be used to determine the subjective efficiency of a specific codec.

Most applications require that the reconstructed image be as close as possible to the original with little or no artifacts. This raises the question of the determination of the *quality* or *fidelity* of the image. In the rate distortion theory, the problem is one of finding a suitable distortion criteria. In most applications the user of the video information will be a human interpreter. This places a subjective image criteria on the evaluation of the quality of a reconstructed image. In section VI the question of image fidelity is discussed briefly. The conclusion reached in that section is that a human visual weighted mean square error gives an acceptable image fidelity measure, although the mean square error itself will also be used for comparison purposes. The use of the human visual system in the image coding environment has been widely publicised [4,7]. However, it has been found that care should be taken in the application of these models since they depend on certain physical parameters, such as the viewing angle, that the user might want to adapt. The viewing angle might change considerable if the sections of the image is magnified. In general the incorporation of the model of the human visual system will increase the performance of the image coder since it provides additional information regarding the image coding system [1,2,7]. The model of the human visual system developed by Mannos and Sakrison [16] were used in the experiments.

The thesis was started with an investigation into the different types of distortion that is generated during the coding process. The origins of the distortions were categorised according to quantisation and block processing noise. Of the different sources of noise the block processing noise, generated by the independent processing of blocks, was found to generate the most subjective objectionable distortion, namely the *block-effect*. An efficient solution to this problem, using an overlapping transform, was given by Cassereau [61]. The specific transform was called the lapped orthogonal transform (LOT). Another technique that reduces the block-effect and also improves overall coder performance is the use of a visual error criterium in the assignment of bits for quantisation. A feasible method to incorporate the visual system into transform coding was given by Eggerton [56].

The final goal of the image coding system is the transmission of the coded image over a transmission channel. Therefore the effect of channel errors had to be kept in mind in the design of the image codec's. Since channel errors are more serious when the data is represented efficiently, the addition of a controlled form of redundancy, in the form of error detection and correction schemes, is necessary [3]. The evaluation of different error coding schemes did not form part of the thesis but the effect of channel errors on the different types of data, i.e. quantisation information and coded coefficients, present in the encoded stream were investigated.

A. Thesis Structure

It was decided to divide the thesis into problem statement and problem solving sections. The thesis starts with a general background to transform image coding, which include a simplistic but relatively efficient image codec. The sources of errors in this basic transform image codec is investigated and presented. The next sections presents possible solutions to the identified errors. A diagrammatical representation of the structure of the thesis is shown in figure 1.

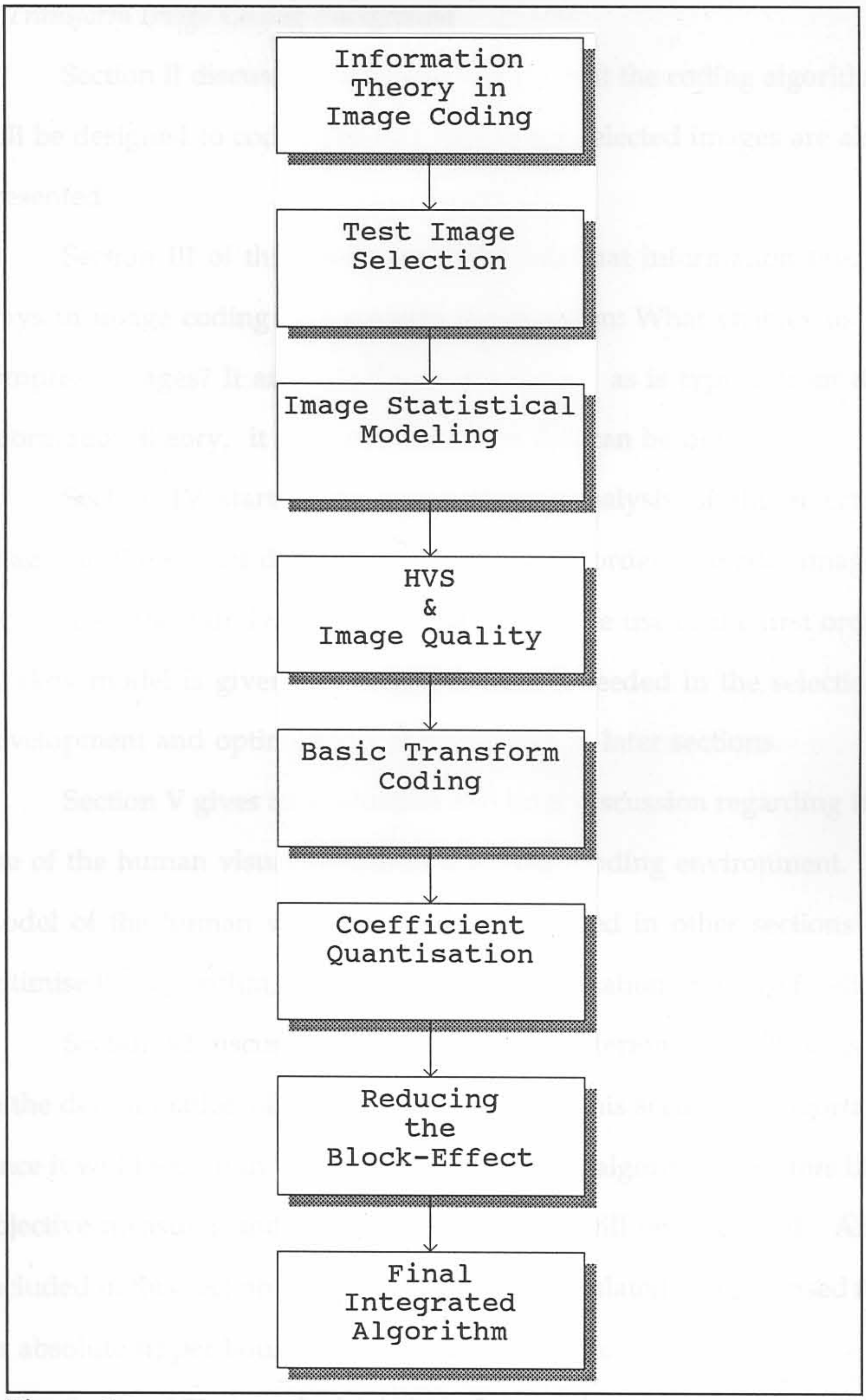

Figure 1 Diagrammatic representation of the thesis structure.

1. Transform Image Coding Background

Section II discusses the type of images that the coding algorithm will be designed to code. The originals of the selected images are also presented.

Section III of this thesis states the role that information theory plays in image coding and answers the question: What enables us to compress images? It answers the question but, as is typical from the information theory, it does not state how this can be done.

Section IV starts with the statistical analysis of the selected images in the spatial domain. This is done in order to model images for mathematical derivations. The reason for the use of the first order Markov model is given. This model will be needed in the selection, development and optimisation of transforms in later sections.

Section V gives an evaluation and brief discussion regarding the use of the human visual system in the image coding environment. A model of the human visual system will be used in other sections to optimise the algorithm, and also in the determination of image fidelity.

Section VI discusses the image quality criterion that will be used in the determination of coded image quality. This section is important since it will be used in the synthesis of the final algorithm. Factors like objective measures and subjective evaluation will be discussed. Also included in this section is the rate distortion simulated images, used for an absolute upper bound on codec performance.

Section VII gives the basic transform coder structure that will be used as a starting point. The functions of the basic structures is investigated and sources of errors is identified and discussed.

2. Codec Optimisation

The following sections are aimed at optimising the image codec, in order to minimise the observed coding-error.

Section VIII looks at improving quantisation by using statistical models, making the quantisation adaptive, and using source coding.

Section IX looks at methods for reducing the block effect at low bit rates. This includes using the human visual system model, spatially variant filtering and overlapping transforms.

Section X puts the different subsections together and evaluates the system's overall performance, including performance in the presence of channel errors.

Section XI contains the concluding remarks regarding the results achieved in all the sections.