**CHAPTER 1**

**INTRODUCTION**

In recent years, the need for data compression has steadily grown. More complex applications require considerable storage capacity and transmission bandwidth. For example, a typical digital satellite image requires 48 Mbits of storage space. There compression of data prior to its storage and saving that data in smaller memory space will economize memory and transmission capacity. studies show that We create 2.5 quintillion bytes of data every day, so much that 90% of data in the world today has been created in the last two years alone (*Source IBM*).

**1.1 BACKGROUND**

The main aim of image compression is to remove the redundancy from image in so as to allow the same image reconstruction at the receiver end. Data compression techniques can be classified in two categories, namely:

* **lossless compression technique (or reversible coding):** It enables an exact replica of original data to be reproduced after reconstruction.
* **lossy compression techniques (or nonreversible coding**): allows slight differences between the original and the reconstructed data. This type of technique allows a greater degree of compression. These techniques are useful in applications such as broadcast television, videoconference transmission etc.

For lossy type of compression, transform coding techniques like cosine transform, wavelet transform are very effective techniques, which give better results but it processes the data in serial manner and hence requires more time for processing.

Many lossy encoding techniques are capable of reproducing recognizable monochrome images from data that have been compressed by a factor more than 30:1 and images that are virtually indistinguishable from their originals at a compression factor from 10:1 to 20:1. This type of techniques is how ever suitable for untypical applications in some scientific and medical images.

A traditional approach to reduce the large amount of data would be to discard some images redundancy and introduce some noise after reconstruction. Here, advanced technologies to reduce the size of the storage device and improve the efficiency of transmission are needed. There are a number of useful methods to implement advanced data compression e.g. variable length coding is the simplest approach to lossless image compression which reduces only coding redundancy.

The most popular technique for removing coding redundancy is Huffman coding. When individually coding the symbols of an information source, Huffman coding yields the smallest possible number of code symbols per source symbol.

**1.2 STATEMENT OF THE PROBLEM**

Following the rapid development of ICT, more information is being processed, stored, and transmitted over networks. The need for data compression and transmission is increasingly becoming a significant topic in all areas of computing and communications

The Hutter Prize (2006) emphasizes “Compression is motivated by the fact that being able to compress well is closely related to acting intelligently, thus reducing the slippery concept of intelligence to hard file size numbers. In order to compress data, one has to find regularities in them, which is intrinsically difficult (many researchers live from analyzing data and finding compact models). So compressors beating the current "dumb" compressors need to be smarter”.

.

**1.3 AIM AND OBJECTIVES**

The aim of data compression is to reduce redundancy in stored or communicated data, thus increasing effective data density. Also to introduce the main issues in data compression and common compression techniques for text, audio, image and video data and to show the significance of some compression technologies.

The objectives of the subject are to:

* outline important issues in data compression
* describe a variety of data compression techniques
* implement the use of Artificial Neural Networks in data compression
* describe elementary techniques for Artificial Intelligence in data compression

**1.4 RESEARCH QUESTIONS**

There are two main research questions that will be investigated in this project. The first is whether the threshold of. The second is to ascertain if compression is relevant to Artificial Intelligence.

**1.5 SCOPE OF STUDY**

This project will consist of improving the current data rates to support future high speed applications, especially for use in multimedia services. Data rate is a function of various factors such as the Data compression and decompression time, size, transfer protocol etc.

**1.6 METHODOLOGY**

The Burrows-Wheeler Algorithm was published in the year 1994 by Michael Burrows and David Wheeler in the research report “A Block-sorting Lossless Data Compression Algorithm”. [22] This research report is based on an unpublished work by David Wheeler from the year 1983.

The Burrows-Wheeler Algorithm is used for data compression. It consists of several stages with each stage performed in succession. The Burrows-Wheeler Transform (BWT) is a way of permuting the characters of a string T into another string BWT (T). This permutation is reversible; one procedure exists for turning T into BWT(T) and another exists for turning BWT(T) back into T. The BWT has two main applications: **compression and indexing**.

When a character string is transformed by the BWT, none of its characters change value. The transformation permutes the order of the characters. The Burrows-Wheeler transform consists of a reversible transformation of the input string. Each stage is a block transformation of the input buffer data and forwards the output buffer data to the next stage. The stages are processed sequentially from left to right for compression; for decompression they are processed from right to left with the respective backward transformations. The BWT stage keeps the number of symbols during the transformation constant, except an additional index, which is created during the forward transformation. A useful side-effect of the Burrows-Wheeler transform is that it produces, as an artefact of the inverse transform, a list of all the substrings of the text in sorted order, making it possible to perform a binary search on any le encoded with the Burrows-Wheeler transforms an O (m log n) pattern matching algorithm. [1]

**1.7 SIGNIFICANCE OF THE STUDY**

This study will broaden the knowledge of compression techniques and the mathematical foundations of data compression, also to become aware of existing compression standards and some compression utilities available. To understand the concept of Neural Networks and their application in the world of computer science. To also benefit from the development of transferable skills such as problem analysis and problem solving and improving of programming skills by doing laboratory work for this subject.

**1.8 ORGANIZATION OF DISSERTATION**

The remainder of the thesis is structured as follows.

**Chapter 2** reviews the relationship between

**1.9 DEFINITION OF TERMS**

**REFERENCES**

The Hutter Prize (2006) online, available: <http://prize.hutter1.net/#motiv> [accessed 6 June 2016].

1. M. Burrows and D.J. Wheeler, A Block-sorting Lossless Data Compression Algorithm, Research Report 124, System Research Center, Digital System Research Center, Palo Alto, CA, 1994.