RV College of Engineering
Experiential Learning
Synopsis

25-2024

Title of the Project

[Huffman + LZW Hybrid encryption]

Team Details

.Sl.No	USN	Name
1	**1RV24CS121**	**K.SAI JASWANTH REDDY**
2	**1RV24CS114**	**K.V.DAIWIK**
3	**1RV24CS111**	**VIJAY SAI JETTY**

Mentor

Name:Prof.Mekhala Vinod Purohit

Designation:Assistant Professor

Department:Computer Science and engineering

1. Introduction

Data compression is an essential technique in computer science used to reduce the size of data for storage and transmission. The Huffman + LZW Hybrid Compression project combines two well-known lossless algorithms—Lempel-Ziv-Welch (LZW) and Huffman coding—to develop a highly efficient compression system. This hybrid approach improves upon the limitations of individual compression methods, offering better compression ratios and practical relevance to formats like ZIP and multimedia data encoding.

2. Problem Definition

2.1. Problem Statement

The growing volume of digital data demands efficient compression to optimize storage space and speed up data transmission. Existing single compression methods either focus on pattern substitution (LZW) or statistical encoding (Huffman), but using them independently limits the compression performance. This project aims to design and implement a hybrid compression system that combines the benefits of both algorithms to achieve superior compression and maintain lossless data recovery.

2.2. Background Information

Huffman coding assigns shorter binary codes to frequently occurring symbols based on their frequency, optimizing prefix codes for variable-length encoding. LZW dynamically builds a dictionary of substrings found in the data, replacing repeated patterns with dictionary indices to reduce redundancy. Previous works show that combining dictionary-based and frequency-based methods can greatly enhance compression efficiency. The DEFLATE algorithm, widely used in ZIP files, employs a similar hybrid approach. However, implementing this combination from first principles requires careful design of data structures like hash tables and priority queues and managing bit-level encoding effectively.

3. Objectives

3.1. Primary Objectives

- Develop a hybrid lossless compression system combining LZW and Huffman algorithms.
- Implement efficient data structures for dynamic dictionary creation and Huffman tree construction.
- Achieve higher compression ratios than standalone Huffman or LZW compression.
- Ensure accurate lossless decompression, recovering original data exactly.

3.2. Secondary Objectives

- Demonstrate practical applicability by testing on various data types (text, images).
- Optimize runtime performance for encoding and decoding.
- Document the algorithm design, implementation details, and performance evaluation.
- Lay groundwork for further research or integration into real-world compression utilities.

4. Methodology

4.1. Approach

The project implements two sequential compression phases:

- **Phase 1 (LZW Encoding):** Input data is processed to build a dynamic dictionary of substrings. Each occurrence of a substring is replaced by a numeric dictionary index.
- **Phase 2 (Huffman Encoding):** Frequencies of these dictionary indices are computed, and a Huffman tree is constructed to assign optimal variable-length codes. The indices are then re-encoded using these codes, resulting in the final compressed binary output.

A decoding process reverses these steps precisely to recover the original input data.

Flowchart:
(Input Data) → [LZW Encoder] → (Dictionary Indices) → [Huffman Encoder] → (Compressed Output)
(Compressed Output) → [Huffman Decoder] → (Dictionary Indices) → [LZW Decoder] → (Original Data)

4.2. Procedures

- Research and review of Huffman and LZW algorithms and related literature.
- Design data structures (hash tables, priority queues) and plan encoding/decoding processes.
- Implement LZW encoder and decoder modules.
- Implement Huffman tree building, encoder, and decoder.
- Integrate both algorithms into a hybrid pipeline.
- Test on sample data sets; measure compression ratios and correctness.
- Optimize code and write detailed documentation and report.

5. Expected Outcome

- A fully functional hybrid compression software prototype implementing combined LZW and Huffman encoding and decoding.
- Demonstrated improvement in compression ratios over single-algorithm methods.
- Enhanced understanding of dynamic dictionary management, prefix coding, and bit-level data handling.
- Potential for journal or conference publication documenting the hybrid method's design and performance.
- Contribution to compression technology through education and possible enhancement of real-world utilities.

- The project offers significant potential for academic publication and patent filing by demonstrating a novel, efficient, and lossless hybrid compression algorithm that outperforms standalone Huffman or LZW methods, with practical applications in digital communication, image compression, and data storage.
- Enabling further research on hybrid and multi-phase compression techniques or integration into existing frameworks.
