

Illustrate the physical - based & probability band models used in lossless compression tech. How are they used to evaluate efficiency?

A Lossless compression reduces file size without losing information. The original data can be perfectly reconstructed. Based on modeling tech., it is lossless.

1. physical Models.

2. Probability-based Models.

→ Physical Models.

Not exploit the structural data. They ignore spatially separated regions with

repetition of patterns in an redundancy by replacing shorter references.

Working Principle:

Data is stored in strings.
Replaced run with pointers / length or dict ref.
Reconstruct original data during decoding.

Techniques

a) Run length Encoding (RLE).

stores repeated symbols as:
symbol + count

Eg) AAAA → A4

Used in bitmap img & fax transmission

b) Lempel-Ziv (LZ77/LZ78)

Uses sliding window or dict.

Repeated patterns are encoded as (offset, length)

Used in ZIP & GZIP

advantage: simple, fast.

Limitations: inefficient if repetition is low.

2. Probability-based models.

use statistical freq of symbols
symbols core assignment shorter codes,
statistical redundancy
overlapping Principle.

calculate symbol probability
Assign variable-length codes.
minimize avg code length.

Techniques

a) Huffman coding

construct binary tree using freq
generate prefix-free codes
Used in large & audio compression.

b) Arithmetic coding

represent entire message as a
provide higher compression ratio.

Adv.: Better compression ratio.

Limitation: more complex

Dyn. model remove redundancy
while probability-based model use symbol freq
to assign optimal codes.

function has obs &uffman

using repeated patterns,

use symbol freq