Illustrate the physical and probability based models used in lossless compression highlighting their applications. ?

In lossless data compression, modeling is used to estimate symbol probabilities so that encoding can be done efficiently without losing any information two major modeling approaches are Physical model and Probability-based models.

1. Physical models (source / mechanistic models)

Physical models are based on the actual generation mechanism of the data. They exploit domain-specific knowledge about how the data is produced, rather than relying purely on statics.

Characteristics

- use explicit knowledge of the source of data
- Highly accurate for a specific

- usually combined with entropy coding

Examples

Image compression: modeling spatial correlation between neighboring pixels

Audio compression: modeling sound production and wave propagation

Video compression: modeling object motion and temporal redundancy

Applications

- Medical image storage
- Satellite imagery
- Scientific and sensor data
- Audio and video archival systems

Advantages

- High compression efficiency.
- Very accurate predictions

Limitations

- Difficult to design
- Not general-purpose
- Require deep domain knowledge.

2. Probability-Based models (Statistical models)

Probability-based models estimate the probability of symbols based on their frequency of occurrence in the data stream.

Key idea

Symbols that occur more frequently are assigned shorter codes while rare symbols get longer codes.

Characteristics

- Data-driven and adaptive
- Simpler and more general than physical models.
- Can be static or adaptive.

Common Probability models

Zero-order model (each symbol independent)

- First order model (depends on previous symbol)

High order Context models.

common algorithms

- Huffman coding
- Arithmetic coding
- Adaptive coding techniques

Applications

- Text file compression
- Executable and binary compression of files and data bases
- General-purpose compressors

Advantages

- simple and flexible
- no domain knowledge required
- suitable for many data types.

Limitations

- less efficient than physical models for structured data
- Performance depends on accuracy of probability estimation

Conclusion

- Physical models are ideal when the data source is well understood and high compression efficiency is required.
- Probability-based models are preferred for general purpose lossless compression due to their simplicity and adaptability
- Modern compressors often combine both models to achieve better performance.