**Topic: Lossy and lossless file compression**

Reading Time: 20 mins

**·        Note\* Highlight important/core points while reading**

·        Read the content and write the answers given in the document in your words, to get the solid grip on topic.

**Lossy and Lossless File Compression**

File compression is a technique used to **reduce the size of a file** while maintaining its usability. It helps in **saving storage space, reducing transmission time, and improving system performance**.

There are **two types** of file compression:

1. **Lossy Compression**
2. **Lossless Compression**

**1. Lossy Compression**

Lossy compression **removes some data permanently** to reduce file size. This is commonly used for **multimedia files** such as images, audio, and video.

**How Lossy Compression Works**

* Unnecessary details are **discarded**, reducing file size.
* The compressed file **cannot be restored** to its original quality.

**Examples of Lossy Compression Algorithms**

|  |  |  |
| --- | --- | --- |
| **Compression Type** | **File Type** | **Example Format** |
| **Audio** | Sound files | **MP3 (MPEG-3)** |
| **Video** | Video files | **MP4 (MPEG-4)** |
| **Image** | Pictures | **JPEG (Joint Photographic Experts Group)** |

**Lossy Compression in Different File Types**

**A. Image Compression (JPEG)**

* JPEG compression works by:
  1. **Reducing colour depth** – The human eye is **less sensitive to colour variations** than brightness differences.
  2. **Dividing the image into 8×8 pixel blocks** and removing unnecessary details.

**B. Audio Compression (MP3 & MP4)**

* MP3 and MP4 use **perceptual encoding** to remove data that the human ear cannot detect:
  1. **Removing sounds outside the human hearing range**.
  2. **Perceptual music shaping** – If two sounds are played at the same time, only the **louder sound** is kept.

**2. Lossless Compression**

Lossless compression **reduces file size without losing any data**. The original file can be **perfectly reconstructed** from the compressed version.

**How Lossless Compression Works**

* **Identifies repeating patterns** in data.
* Stores these patterns **more efficiently** without removing any data.

**Examples of Lossless Compression Algorithms**

|  |  |  |
| --- | --- | --- |
| **Compression Type** | **File Type** | **Example Format** |
| **Text & Documents** | Text files, software | **ZIP, RAR** |
| **Images** | Pictures | **PNG** |
| **Audio** | High-quality sound files | **FLAC** |

**Run-Length Encoding (RLE) – A Lossless Compression Method**

RLE is a simple **lossless** compression algorithm that reduces file size by **replacing repeated data** with a shorter representation.

**How RLE Works:**

* It **identifies long sequences of the same data**.
* The repeating string is **stored as two values**:
  1. **The number of times the data is repeated**.
  2. **The actual data value**.

**Example of RLE on Text Data**

Original text:

AAAAABBBBCCCCCDDD

Compressed using RLE:

5A4B5C3D

* Instead of storing each letter separately, RLE stores **the letter + the count**.

**When is RLE Useful?**

* **Effective for images with large areas of the same colour**.
* **Not effective for files with lots of unique data**, like normal text documents.

**Comparison of Lossy vs. Lossless Compression**

|  |  |  |
| --- | --- | --- |
| **Feature** | **Lossy Compression** | **Lossless Compression** |
| **Data Loss** | Yes (some data is removed) | No (all data is preserved) |
| **File Size** | Much smaller | Larger than lossy, but still reduced |
| **Can be Reversed?** | No, original file cannot be restored | Yes, original file can be restored |
| **Best for** | Images, music, videos | Text, software, high-quality audio/images |

**Real-Life Examples of Compression**

|  |  |  |
| --- | --- | --- |
| **Scenario** | **Compression Type** | **Reason** |
| **Streaming music on Spotify** | **Lossy (MP3)** | Saves bandwidth and loads faster |
| **Downloading a ZIP file** | **Lossless (ZIP)** | Keeps all data intact |
| **Uploading photos to a website** | **Lossy (JPEG)** | Reduces file size for faster loading |
| **Sending a document via email** | **Lossless (ZIP, PDF)** | Reduces file size while keeping text unchanged |

**A-Rated Questions/Answers By Examiner**

**Q1: What is the main difference between lossy and lossless compression?**

**Answer:**

* **Lossy compression removes data permanently**, reducing file size but **decreasing quality**.
* **Lossless compression retains all data**, allowing files to be **fully restored** to their original state.

**Q2: Explain how MP3 compression reduces file size.**

**Answer:**  
MP3 compression works by:

* Removing **inaudible sounds** (e.g., very high/low frequencies).
* **Perceptual music shaping** – if two sounds play together, the **quieter one is removed**.
* Reducing the **bitrate** to decrease file size.

**Q3: Why is JPEG compression effective for images?**

**Answer:**  
JPEG compression reduces file size by:

1. **Reducing colour depth** – human eyes are **less sensitive to small colour changes**.
2. **Breaking images into 8×8 pixel blocks** and removing **unnecessary details**.
3. **Using lossy compression** to discard subtle colour variations.

**Q4: How does Run-Length Encoding (RLE) work?**

**Answer:**  
RLE works by **replacing repeated data** with a shorter code:

* Example:

AAAAAABBBB → 6A4B

* It is **only effective when there are long repeated patterns**.

**Q5: Why is lossless compression important for text files?**

**Answer:**

* Text files must be **perfectly restored** after compression.
* Lossless compression (e.g., **ZIP**) **keeps all data intact**.
* Ensures **accuracy** in important documents.

### Write your Answers on your Notebook and Verify it on Next Screen

**Q6: Why is lossy compression not suitable for medical images or legal documents?**

**Q7: What are the advantages and disadvantages of using lossy compression?**

### ****Q8: How does lossless compression achieve file size reduction without losing data?****

**Q9: Why does lossy compression work well for audio and images but not for text files?**

**Q10: Why does an MP3 file sound different from a WAV file?**

**6. Answer:**

* **Medical images (X-rays, MRIs) need full accuracy** for diagnosis, so losing any detail could lead to incorrect medical assessments.
* **Legal documents must remain exactly as they were written**, so any removed data could change their meaning.
* **Lossless compression (PNG, ZIP, PDF) is preferred** to ensure full restoration of original data.

**Answer:** **Lossy compression removes data permanently, which can be risky for medical images and legal documents where accuracy is critical.**

**7. Answer:**  
**Advantages:**

* **Much smaller file sizes**, making storage and transfer faster.
* **Ideal for multimedia (MP3, JPEG, MP4)** where slight quality loss is acceptable.
* **Efficient bandwidth usage** for streaming services like YouTube and Spotify.

**Disadvantages:**

* **Data is permanently lost**, meaning quality cannot be fully restored.
* **Recompressing the file multiple times reduces quality further** (e.g., saving a JPEG multiple times).
* **Not suitable for files requiring full accuracy**, like software programs or spreadsheets.

**Answer:** **Lossy compression significantly reduces file size but causes irreversible quality loss, making it unsuitable for critical data.**

**8. Answer:**

* **Uses algorithms like Run-Length Encoding (RLE) and Huffman Coding** to store data more efficiently.
* **Finds repeating patterns** and replaces them with shorter representations.
* **No data is lost**, allowing the file to be restored to its exact original state.

**Example:**  
Original: **WWWWWWBBBWW**  
Compressed (RLE): **6W3B2W**

**Answer:** **Lossless compression reduces file size by efficiently storing patterns without removing any data, ensuring the original file can be restored.**

**9. Answer:**

* **Human perception plays a role in audio and images**, meaning minor data loss isn’t noticeable.
* **In text files, every character matters**, so losing even one letter could change the meaning of a document.
* **Lossy compression is great for reducing multimedia file sizes**, but **lossless compression is needed for text-based files** to preserve accuracy.

**Answer:** **Lossy compression is effective for multimedia because humans don’t notice small losses, but text files require lossless compression to maintain accuracy.**

**10. Answer:**

* **MP3 is a lossy format**, meaning it removes audio details to make the file smaller.
* **MP3 removes inaudible frequencies**, especially very high and very low sounds.
* **MP3 reduces the bitrate**, which affects the overall sound quality.
* **WAV is uncompressed**, meaning it stores all the original sound data, resulting in a larger file with better quality.

**Example:**

* **WAV (Uncompressed)** → 50 MB file with full sound quality.
* **MP3 (Compressed)** → 5 MB file with some lost details but still clear sound.

**Answer:** **MP3 compression removes some sound data to reduce file size, while WAV keeps all original audio, leading to better quality but larger files.**