1. Multimedia Compression Methods

Multimedia files like audio, video, and images are often very large in their uncompressed form. Compression is used to reduce the amount of data required to store or transmit these files, making storage more efficient and reducing bandwidth requirements for transmission. There are two main types of compression methods:

1. **Lossless Compression**: This preserves all the original data perfectly, meaning that the decompressed file is identical to the original. Common examples include **PNG** for images and **FLAC** for audio.
2. **Lossy Compression**: This reduces the file size by removing data that is less important or less noticeable to human perception, often sacrificing some quality in the process. **JPEG** for images and **MP3** for audio are popular lossy formats.

2. Lossy Compression Techniques in Multimedia

Lossy compression techniques are often used for **audio, video, and images**, aiming to remove data that is not perceptually significant to humans.

1. **Human Perception Optimization**: Lossy compression exploits limitations in the human eye and ear:
    1. For images, the human eye is **less sensitive** to subtle changes in **high spatial frequency** areas (fine details), which allows image compression methods like **JPEG** to reduce the size by dropping some fine-grained details.
    2. For audio, humans are more sensitive to **lower frequencies** and less sensitive to **higher frequencies**, which allows audio codecs like **MP3** to selectively discard less audible frequencies.

3. Audio Sampling Example

When audio is recorded (e.g., with a microphone), it needs to be **digitized** for storage. The **sampling rate** is the number of times per second that the audio signal is measured.

1. **CD Quality Audio**: Has a sampling rate of **44.1 kHz** (44,100 samples per second) and each sample is represented as a **16-bit** value. In stereo audio, two channels (left and right) are recorded, which means twice the amount of data is needed.
2. Example Calculation:
    1. **1 second of stereo CD quality audio**:

\[44100 \text{ samples/second} \times 16 \text{ bits/sample} \times 2 \approx 1.4 \text{ Megabits} \] This is a significant amount of data for just one second of sound, demonstrating why compression is essential.

4. Video Frame Compression Example

Video consists of a sequence of frames displayed rapidly to create the illusion of motion.

1. **Full HD Video**: A Full HD (1080p) frame has a resolution of **1920×1080** pixels, and with **32 bits** per pixel (which allows for true color with alpha transparency), one frame takes up:

$$ 1920 \times 1080 \times 32 \text{ bits} \approx 8 \text{ Megabytes} $$

1. A typical video displays **30 frames per second** (fps), meaning that **240 MB** of data would be required per second without compression. This is why video compression is vital to make streaming and storage practical.

5. Two-Dimensional Fourier Transform

The **Fourier Transform** is a mathematical operation that transforms a signal from the **time domain** (or spatial domain, for images) to the **frequency domain**.

1. In multimedia compression, **2D Fourier Transform** is used for images. It converts image pixels into frequency components. These frequency components can then be analyzed, and parts that are less significant to human perception can be discarded.
2. **Example**: Imagine an audio waveform where the amplitude is plotted over time. The Fourier Transform decomposes this waveform into its frequency components—essentially figuring out which notes (frequencies) are being played and how strong they are. This concept is applied to images as well, allowing more effective compression.

1. In practice, this means that areas of an image with **high-frequency details** (e.g., fine patterns) can be simplified or removed during compression without significantly impacting the perceived quality. This is what is exploited in compression standards like **JPEG** to achieve significant size reduction.

6. Fourier Transform Analogy with Music

The analogy mentioned in the text compares the **Fourier Transform** to recognizing musical notes in an audio recording:

1. Imagine you have a **mono recording** of music with different notes being played over time. The **Fourier Transform** is like figuring out which notes are being played (e.g., **C#**, **C**) during different time intervals.
2. This is similar to trying to write the **musical score** (sheet music) just by listening to the audio. By focusing only on the most important notes (the **note heads**), you would end up with a much more **compressed** version of the original sound, while still retaining most of the important information.

7. Human Sensory Limitations and Compression

1. **Vision**: The human eye is more sensitive to **low spatial frequencies** (smooth gradients) and less sensitive to **high spatial frequencies** (fine details or noise). This property is used in image and video compression to drop unnecessary detail in complex patterns, which most viewers won’t notice.
2. **Hearing**: The human ear is more sensitive to certain frequency ranges. **MP3 compression**

uses this by discarding audio data that is in ranges we typically cannot hear well.

Summary

Multimedia compression methods, particularly lossy ones, take advantage of **human sensory limitations** to reduce data size without noticeable loss in quality. Techniques such as **Fourier Transform** allow multimedia compression algorithms to identify and remove data that is less perceptible, thereby achieving substantial compression ratios. These methods make it possible to store and transmit multimedia content effectively, without overwhelming data storage capacities or requiring impractical bandwidth.