

IG CS Data Representation Assess 3

Created by HardyWen

1.4 Measuring Data Storage

Data measuring units

Name	Symbol	Value (base 2)
kibibyte	KiB	2^{10}
mebibyte	MiB	2^{20}
gibibyte	GiB	2^{30}
tebibyte	TiB	2^{40}
pebibyte	PiB	2^{50}
exbibyte	EiB	2^{60}

- What might be in the test
 1. Ordering the units from the lowest to the highest
 2. Transformation between units

Image Size

- width * height * color depth(in byte) * number of images
- E.g.,

Calculate the file size of 10 images with resolution 100*150, using a color depth of 16. Give your answer in KiB.

Size of a single image = $100 * 150 * (16 / 8) = 30000$ bytes

Size of the file = $10 * 30000 = 300000$ bytes

= $300000 / 1024 = 292.97$ KiB

Remember to convert color depth into bytes !

Recording Size

- sample rate(in hertz) * sample resolution(in byte) * length of soundtrack(in second) * number of channels
 - Stereo: 2 channels
 - Mono: 1 channel
- E.g.,

Calculate the file size of a 2 minutes recording, with sample rate of 44,100 hz and sample resolution of 8 bits, using stereo, give your answer in MiB.

$$\begin{aligned}\text{Size of the recording} &= 44100 * (8 / 8) * 120 * 2 \\ &= 10584000 \text{ bytes} = 10584000 / 1024 / 1024 \text{ MiB} = 10.1 \text{ MiB}\end{aligned}$$

Remember to convert sample resolution into bit and time into seconds!

1.5 File Compression

- Compression: a method that uses an algorithm to reduce the size of a file

Advantages of Compression

- Might be in the test
1. Less storage space required
 2. Less time to transmit the file from one device to another
 3. Less time to download & upload
 4. Less bandwidth required to transmit the file over the internet

Types of Compression

- Lossy File Compression
 - **Permanently** delete the data
 - Delete the useless content
 - Perceptual Music Shaping: a process that is used in lossy compression that removes sounds that are not normally heard by the human ear
 - Quality & Content loss
 - Data cannot be reconstructed
 - **Used for image**
- Lossless File Compression
 - **Temporarily** alter data
 - No data & content loss
 - Data can be reconstructed

- **Used for text**

- Because losing any single content from the text may result in ambiguity, for example, lossy compress "Father" may result in the word "Fer".

- What might be in the test?

- Compare lossy and lossless file compression
- Which compression method is used to compress text and why
- Which compression method suits the situation given
- What might be carried out in a lossy compression for an audio recording

Compression Algorithm

- **RLE (run length encoding)**

- An algorithm that groups together repeating patterns and indexes them.
- It is used to lossless file compression
- Sometimes may result in bigger file size !
- Remember to offer examples when defining RLE !!!

- What might be in the test?

- Use RLE to compress an image or a text
- Definition of RLE