

Compresión de datos



Compresión de datos

- Reducir el tamaño de un archivo para que ocupe menos espacio de almacenamiento o para que su transferencia a través de una red sea más rápida.



Compresión de datos

- Todo archivo es una secuencia de bytes, y por lo tanto de números. La idea es aplicar métodos que permitan, en ciertas condiciones, reducir ese tamaño.
- El algoritmo de compresión puede ser de **sin pérdida**, o **con pérdida**.



Compresión de datos sin pérdida

- Decimos que la compresión ha sido **sin pérdida** cuando puede extraerse, desde el archivo comprimido, exactamente la misma información que antes de la compresión.



Compresión de datos sin pérdida

- La compresión sin pérdida es reversible: aplicando el algoritmo inverso, o de descompresión, siempre puede volverse a la información de partida.



Compresión de datos con pérdida

- Cuando la compresión se hace **con pérdida**, no existe un algoritmo de descompresión que recupere la información original; es decir, no existe un algoritmo inverso.



Compresión de datos con pérdida

- El resultado de la compresión con pérdida de un archivo es otro archivo del cual ya no puede recuperarse la misma información original, pero que de alguna manera sigue sirviendo a los fines de la usuaria.



Compresión de datos con pérdida

- La pérdida de información es intencional, y es la usuaria quien ha elegido descartar esa información porque no es necesaria.



Compresión de datos con pérdida

- Por ejemplo, al utilizar un servicio de streaming de video o audio, muchas veces se nos da la oportunidad de elegir la calidad del audio o del video.
- Para disminuir la calidad del audio o sonido, se aplican métodos de compresión con pérdida.



Compresión de imagen con pérdida: Reducción de color

- Si la imagen tiene $\textit{ancho} \times \textit{alto}$ pixels, y la profundidad de color es de n bits, el archivo sin su cabecera mide $\textit{ancho} \times \textit{alto} \times n$ bits.



Compresión de imagen con pérdida: Reducción de color

- Si la imagen tiene $\text{ancho} \times \text{alto}$ pixels, y la profundidad de color es de n bits, el archivo sin su cabecera mide $\text{ancho} \times \text{alto} \times n$ bits.
- Una forma sencilla de compresión con pérdida, que no modifica la resolución, es la reducción de la profundidad de color de una imagen.



Compresión de imagen con pérdida: Reducción de color

- Por ejemplo una imagen a cuatro colores.



Compresión de imagen con pérdida: Reducción de color

- Por ejemplo una imagen a cuatro colores.
- Si sacamos un bit, ¿en cuánto se achica la imagen?



Compresión de imagen con pérdida: Reducción de color

- Por ejemplo una imagen a cuatro colores.
- Si sacamos un bit, ¿en cuánto se achica la imagen?
- El nuevo archivo, sin su cabecera, mide $\text{ancho} \times \text{alto} \times (n - 1)$, o sea, $\text{ancho} \times \text{alto}$ bits mas corto que el original.



Compresión sin pérdida: Codificación por long. de secuencia

- Debemos identificar secuencias de elementos de un mismo valor, computar su longitud, y en lugar de la secuencia, colocar el coeficiente de longitud y el valor que corresponde.



Compresión sin pérdida: Codificación por long. de secuencia

- Debemos identificar secuencias de elementos de un mismo valor, computar su longitud, y en lugar de la secuencia, colocar el coeficiente de longitud y el valor que corresponde.
- La efectividad de este método de compresión depende de la redundancia presente en el material original.



Codificación por longitud de secuencia

- Primeramente fijamos la cantidad de bits que ocuparán los coeficientes de longitud de secuencias.



Codificación por longitud de secuencia

- Primeramente fijamos la cantidad de bits que ocuparán los coeficientes de longitud de secuencias.
- Si los coeficientes son muy pequeños...



Codificación por longitud de secuencia

- Primeramente fijamos la cantidad de bits que ocuparán los coeficientes de longitud de secuencias.
- Si los coeficientes son muy pequeños...
- Si los coeficientes son muy grandes...



Codificación por longitud de secuencia

- Ejemplo: 10 10 11 11 11 11 11 11 11 10
10 10 10 10 10 00...
- Tiene una secuencia de dos pixels con valor 10, siete pixel son valor 11, seis pixels con valor 10, etc.
- ¿Cuántos bits necesitamos para el coeficiente?



Codificación por longitud de secuencia: Ejemplo

- Si tenemos estos datos de la imagen (sin cabecera): 10 10 11 11 11 11 11 11 11 10 10 10 10 10 10 00...
- Los datos de la imagen se comprimirán como 010 10 111 11 110 10....
- Los primeros **treinta bits** de los datos de imagen han quedado comprimidos a **quince bits**.



¿Consultas?

