Procesamiento de Imágenes

Fecha: 29-Julio-2020
Inicio: 1:02pm

Compresión de imagen: Comprimir una imagen es reducir los datos redundantes e irrelevantes de la imagen con la menor pérdida posible, para permitir su almacenamiento o transmisión de forma eficiente.

$$A = \begin{bmatrix} & & \\ & & \\ & & \end{bmatrix}_{m \times n}$$

Total de m·n pixeles.

Comprimir →

← Reconstruir

Valor codificado

$$\boxed{x = [x_1, x_2, \ldots, x_p]}$$

donde $p \ll m \cdot n$

↳ mucho menor

→ reconstruir la imagen A

En el curso, explicaremos un método para comprimir imágenes llamado JPEG, que utiliza la transformada discreta de coseno de 2 dimensiones (DCT-2D)

JPEG:

- Joint Photographic Experts Group
- Es el nombre de un comité de expertos que creó un estándar de compresión y codificación de imágenes.
- Al inicio JPEG fue considerado un método de compresión, pero después fue considerado un formato de archivo. Los archivos de este tipo suelen almacenarse con la extensión .jpg.
- El formato JPEG utiliza un algoritmo de compresión de pérdida para reducir el tamaño de los archivos de imágenes. Esto significa que al descomprimir o visualizar la imagen no se obtiene la imagen original, sino una aproximación.
- Este algoritmo se basa en el fenómeno del ojo humano de detectar más fácilmente los valores de baja frecuencia.
- Este método usa dos conceptos primordiales: DCT-2D y Cuantización.

DCT-2D Sea $A \in \mathbb{R}^{m \times m}$. Entonces la DCT-2D de A se define como $C \in \mathbb{R}^{m \times m}$ tal que:

$$C(i,j) = \frac{1}{\sqrt{2m}} \cdot C(i) \cdot C(j) \cdot \sum_{x=0}^{m-1} \sum_{y=0}^{m-1} A(x,y) \cdot \cos\left(\frac{(2x+1) \cdot i\pi}{2m}\right) \cdot \cos\left(\frac{(2y+1) j \cdot \pi}{2m}\right)$$

donde: $C(K) = \begin{cases} 1/\sqrt{2} & \text{si } K = 0 \\ 1 & \text{si } K > 0. \end{cases}$

Tarea: Investigar la inversa de DCT-2D.

Cuantización: Es una técnica de compresión con pérdida que consiste en comprimir un rango de valores a un grupo reducido de valores. En este caso, necesitamos una matriz de cuantificación Q. Entonces para cuantificar una matriz A, se realiza la siguiente operación:

Matriz Original

$$D_{i,j} = \text{redondear}\left(\frac{A_{i,j}}{Q_{i,j}}\right)$$

Matriz Cuantificada

Matriz de cuantificación.

Algoritmo JPEG (V1)

Parte 1: Compresión.

Los pasos del algoritmo JPEG son los siguientes:

Paso 0: Cargar matriz en formato double, donde las entradas toman valores del conjunto $\{0, 1, 2, \ldots, 255\}$. Sea esta matriz $A \in \mathbb{R}^{m \times m}$.

Paso 1: Restar a cada una de las entradas el valor 128 (ya que la DCT-2D está diseñada para trabajar con valores en $]-128, 127]$). Sea esta matriz $M \in \mathbb{R}^{m \times m}$, donde:

$$M_{i,j} = A_{i,j} - 128$$

Paso 2: Calcular la DCT-2D de M: Sea esta matriz $D \in \mathbb{R}^{m \times m}$, donde

$$D = dct2(M).$$

Paso 3: Utilizando una matriz de cuantificación $Q \in \mathbb{R}^{m \times m}$, obtendremos la matriz cuantificada $C \in \mathbb{R}^{m \times m}$, donde:

$$C_{i,j} = \text{redondear}\left(\frac{D_{i,j}}{Q_{i,j}}\right)$$

Paso 4: Codificar la matriz C en un vector $X \in \mathbb{R}^{p}$, $p \ll m^2$, utilizando el método Zig-Zag.

Matriz de Cuantificación Q.

La matriz de cuantificación es una matriz que permite comprimir valores. Realizando estudios con la visión humana, investigadores obtuvieron que la siguiente matriz permite realizar una compresión del 50% de una imagen de 8×8.

```
Q50=[16 11 10 16 24 40 51 61;
     12 12 14 19 26 58 60 55;
     14 13 16 24 40 57 59 56;
     14 17 22 29 51 87 80 62;
     18 22 37 56 68 109 103 77;
     24 35 55 64 81 104 113 92;
     49 64 78 87 103 121 120 101;
     72 92 95 98 112 100 103 99];
```

Para variar la calidad de la cuantificación, se utiliza la siguiente fórmula:

$$Q_n = \begin{cases} \text{redondear}\left[\dfrac{(100-n)}{50} \cdot Q_{50}\right] & \text{si } n > 50 \\ \text{redondear}\left[\dfrac{50}{n} \cdot Q_{50}\right] & \text{si } n < 50 \end{cases}$$

Q_7

donde n representa la calidad de compresión, $n = 1, 2, \ldots, 100$

Método de Zig-Zag

Este método se utiliza para almacenar en un vector los valores de la matriz A, cuando los valores diferentes de 0 se encuentran agrupados en la parte superior izquierda de esta:

$$A = \begin{pmatrix} 1 & 4 & 3 & 0 & 0 \\ 2 & 5 & 3 & 0 & 0 \\ 4 & 3 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}_{5\times 5}$$

→ Tiene 9 valores diferentes de 0.

→ 25 entradas.

Codificación

$$x = [\,1\ 4\ 2\ 4\ 5\ 3\ 0\ 3\ 3\ 1\,]$$ → 9 entradas.

Decodificación

$$\begin{pmatrix} 1 & 4 & 3 & 0 & 0 \\ 2 & 5 & 3 & 0 & 0 \\ 4 & 3 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Nota: • En general, una matriz $A \in \mathbb{R}^{m\times m}$, tiene $2\cdot m - 1$ diagonales.

Par   Impar

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix}$$