

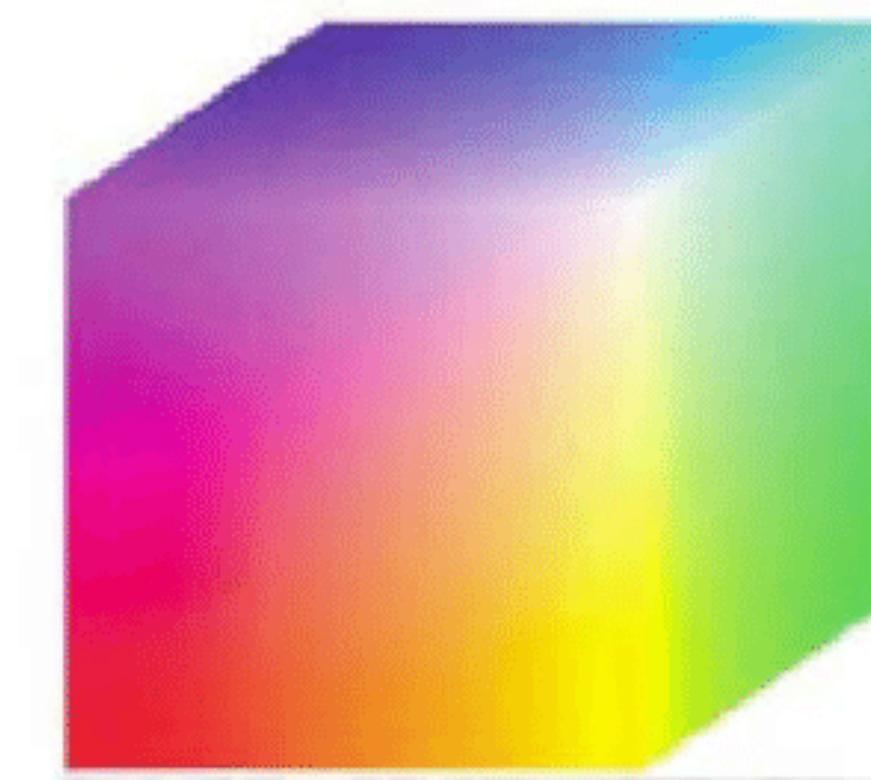
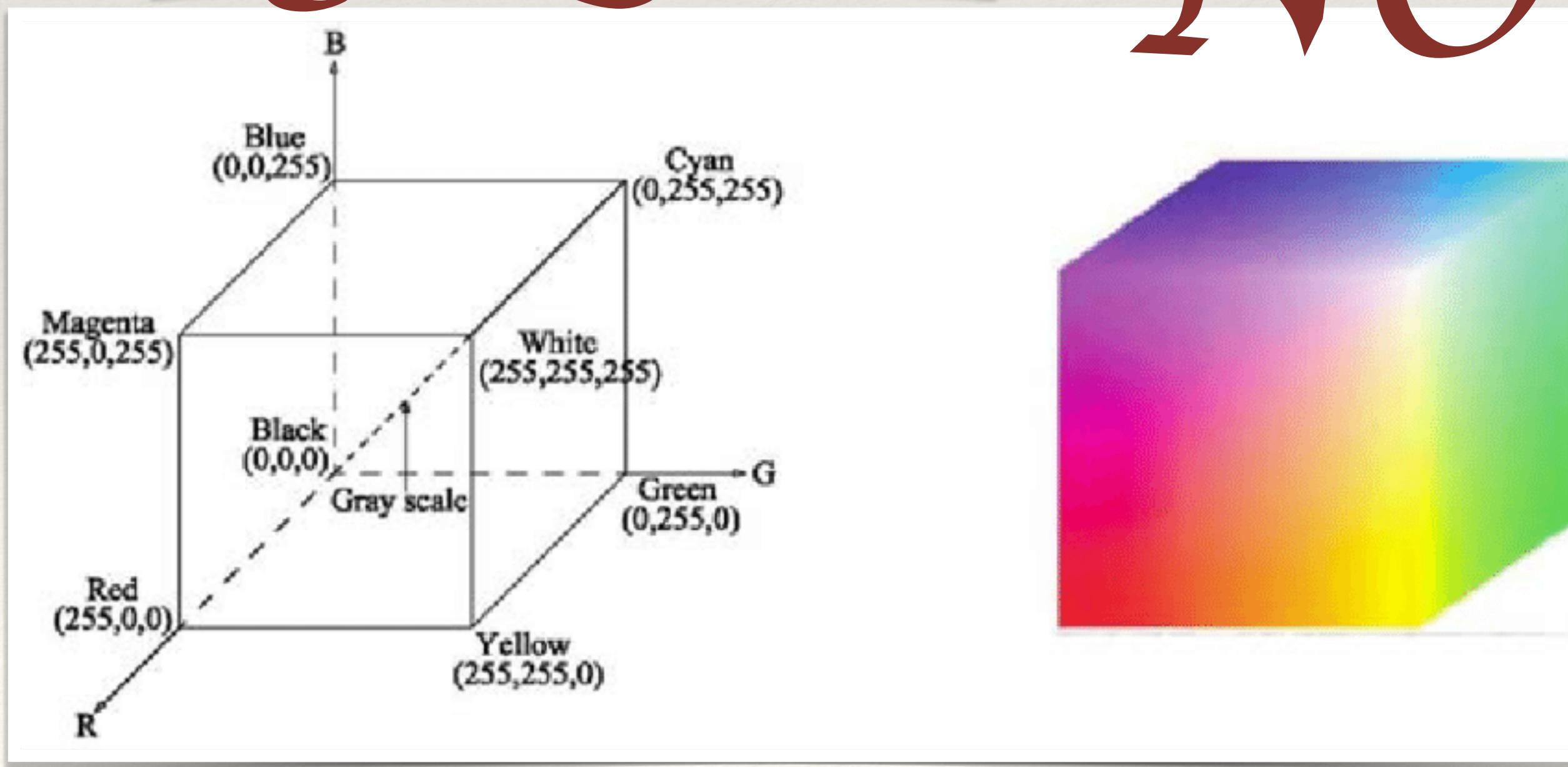
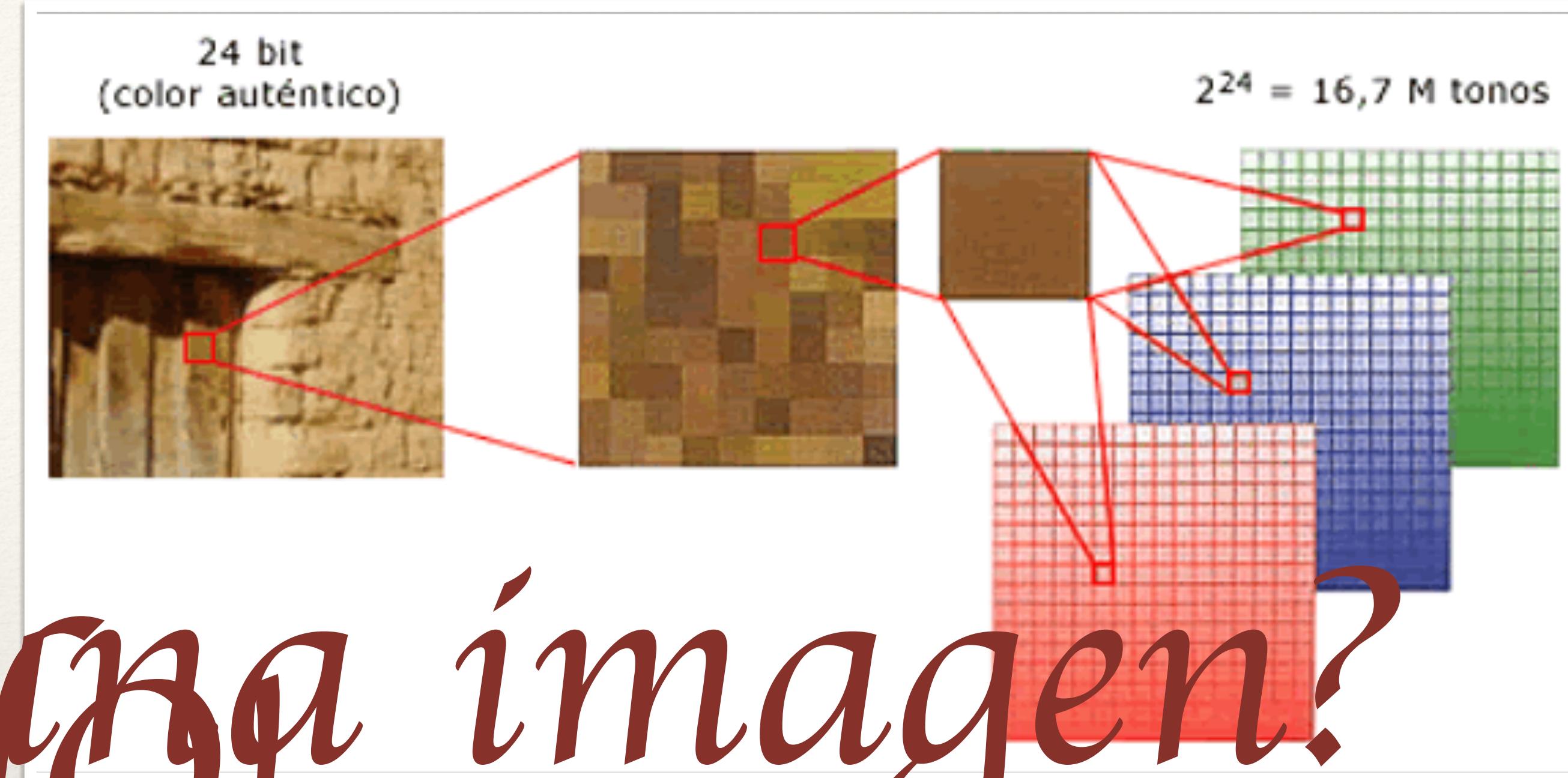
Física computacional 2: Ph.D. Santiago Echeverri Arteaga

Visión computacional

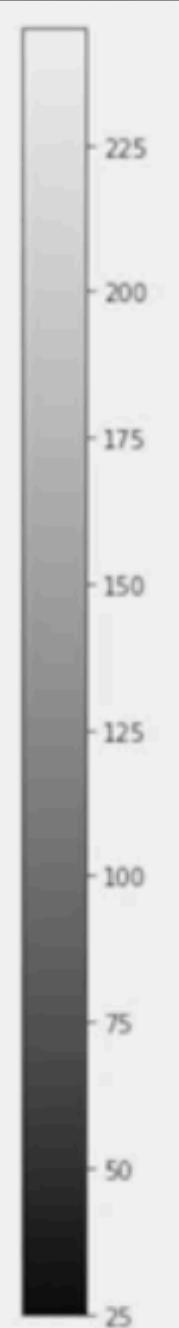
Iniciemos en el mundo del procesamiento de imágenes. ¿Para qué? ¿Qué es una imagen? ¿Habrá una analogía entre imagen y un cristal? ¿Hay alguna similitud entre una imagen y el tensor energía-momento-deformación?



¿Qué es una imagen?



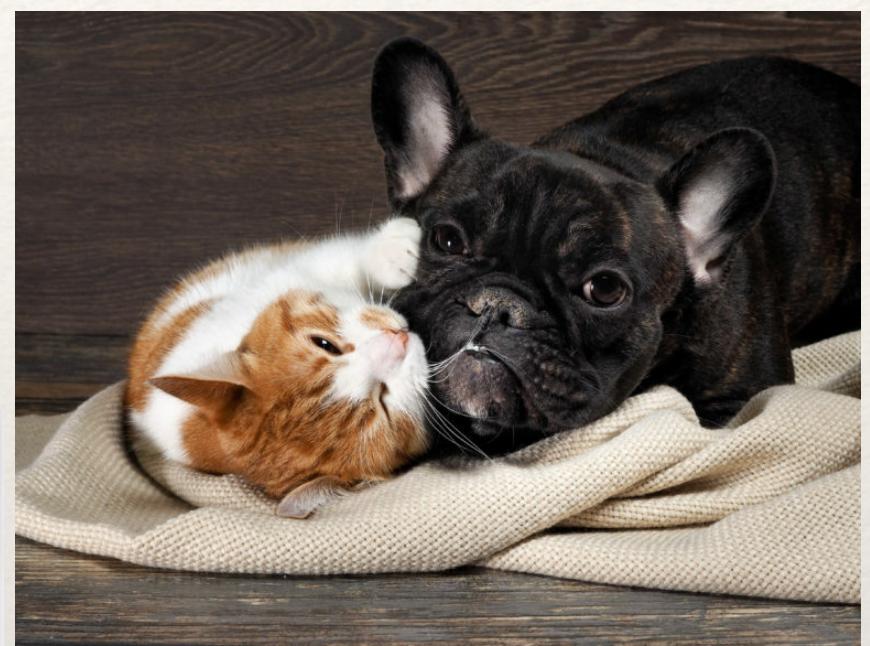
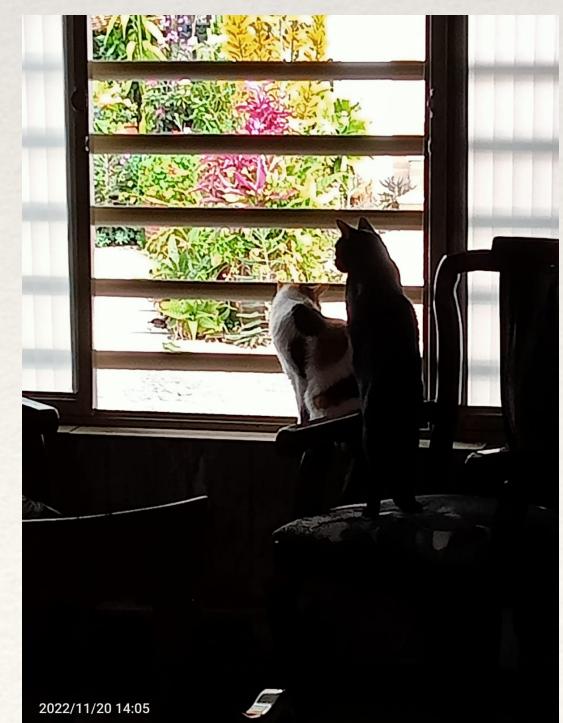
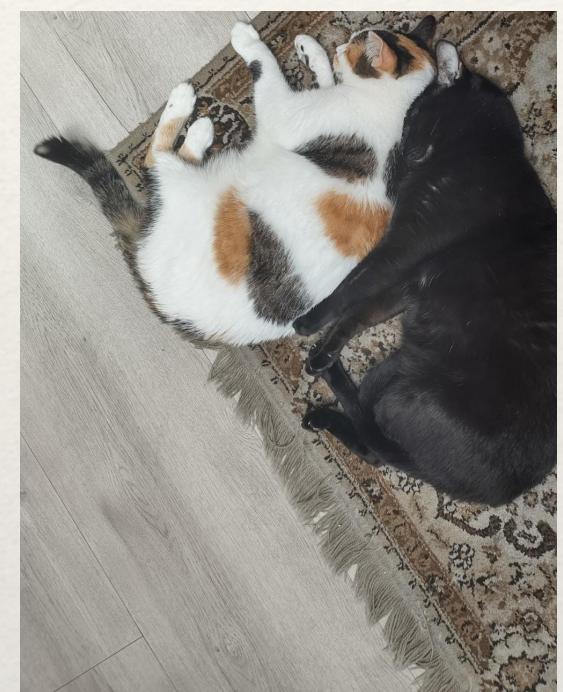
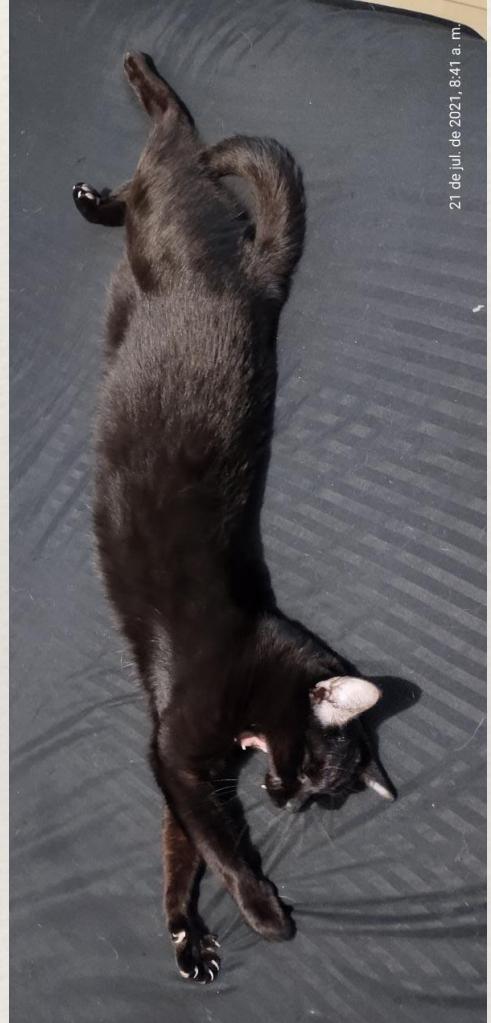
| |
|-----|
| 0 |
| 1 |
| 2 |
| 3 |
| : |
| 500 |





¿Que retos tiene clasificarlas?

- ❖ Solapamiento, cambios de iluminación, diferentes fondos, deformación



¿Qué es una imagen? ¿Que relación existe con los campos escalares?

- ❖ Sean los conjuntos

$D = \{(x, y, z) : x \in \mathbb{Z}; x \in [0, N_{pixels}]; y \in \mathbb{Z}; y \in [0, M_{pixels}]; z \in \{0, 1, 2\}\}$ con N_{pixels} la cantidad de pixeles de una imagen horizontalmente y $y \in [0, M_{pixels}]$ la cantidad de pixeles de una imagen verticalmente. Además sea el conjunto $C = \{x : x \in \mathbb{Z}; x \in [0, 255]\}$. Una imagen a color con profundidad de 8bits la podemos ver como un campo escalar $D \rightarrow C$

- ❖ https://es.wikipedia.org/wiki/Profundidad_de_color

- ❖ ¿Qué estamos haciendo al normalizar una imagen al rango $[0, 1]$

¿Qué es una imagen? ¿Que relación existe con los campos escalares?

- ❖ Consideremos un campo escalar $F : \mathbb{Z}^3 \rightarrow \mathbb{Z}$, cuya imagen sobre una bola abierta es acotada
- ❖ Es posible generar (si limitamos el dominio a esa bola abierta) un isomorfismo a un campo escalar $E : \mathbb{Z}^3 \rightarrow \mathbb{Z}^+$
- ❖ ¿Qué relación tiene esto con una imagen?
- ❖ Es posible encontrar un isomorfismo entre cada campo F y cada campo E , es posible traducir cualquier campo escalar de la física a una imagen (de profundidad arbitraria)

Recordemos algo de cálculo

Campo: Sistema algebraico en el cual las operaciones llamadas adición y multiplicación se pueden realizar y cumplen las propiedades: asociativa, conmutativa y distributiva de la multiplicación respecto de la adición, además de la existencia de inverso aditivo, de inverso multiplicativo y de un elemento neutro para la adición y otro para la multiplicación, los cuales permiten efectuar las operaciones de sustracción y división (excepto la división por cero).

Espacio vectorial: Un espacio vectorial $_nE$ sobre un cuerpo conmutativo K es un conjunto no vacío V , dotado de dos operaciones para las cuales será cerrado y cumple 8 axiomas (Véase su cuaderno de vectorial). La primera de estas operaciones se llama adición vectorial y la segunda operación, producto por escalar. Los elementos de este conjunto se llaman vectores y a los elementos del cuerpo se le llaman escalares.

Base: Una base de $_nE$ será denotada como e_i y los vectores como x^i . Un vector cualquiera se escribirá así $x = \sum_{i=1}^n x^i e_i = x^i e_i$, por lo que dada una base un vector puede ser completamente determinado por su imagen sobre esa base.

Recordemos algo de cálculo

Forma Lineal: Sea un espacio vectorial ${}_nE$, se llama forma lineal f (o covector) a toda aplicación lineal ${}_nE \rightarrow K$. El espacio vectorial de las formas lineales se llama espacio dual ${}_nE^*$ (En cuántica a los elementos del el espacio ${}_nE$ se le llaman kets y a las formas se les llama bras. En álgebra lineal los vectores columna / fila eran los vectores / covectores)

Tensor: Sean ${}_nE$ y ${}_nE^*$ sobre K y $p, q \geq 0$ dos números enteros positivos dados. Se llama tensor T sobre ${}_nE$ (p -veces covariante y q -veces contravariante) de rango $p + q$ a toda forma $(p + q)$ -lineal T sobre ${}_nE$:
 $T \in ({}_nE^*)^p \times ({}_mE)^q : T(f_1, \dots, f_p, x^1, \dots, x^q) \in K$

Tensor métrico: Tensor que permite definir distancias y ángulos (el mismo ds^2).

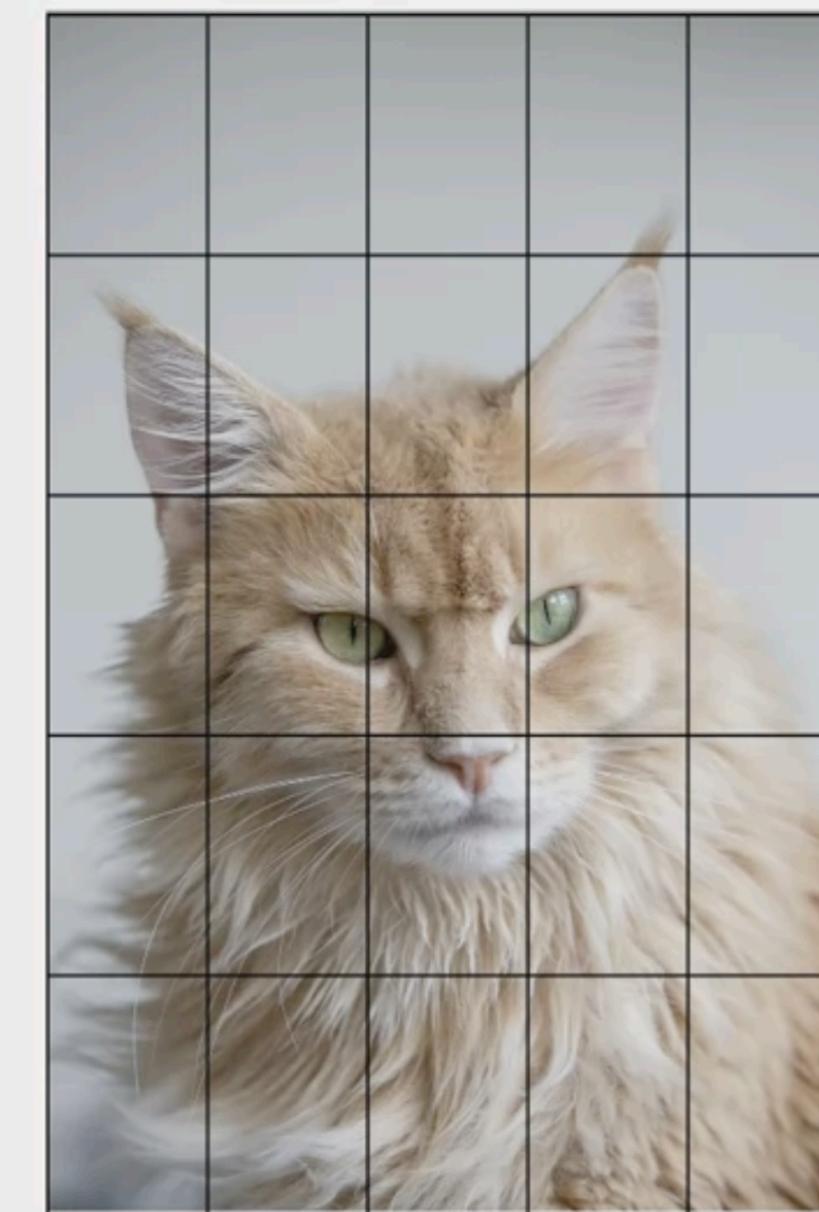
Imagen: Tensor de rango 3 definido sobre el cuerpo \mathbb{Z}^+

Operaciones sobre imágenes: Campo vectorial del espacio de los tensores de rango 3 de dimensión n sobre el cuerpo \mathbb{Z}^+ al espacio de los tensores de rango 3 de dimensión m sobre el cuerpo \mathbb{Z}^+

¿Qué sería un video?

¿Es posible hacer visión computacional con lo que aprendimos en Física computacional 1?

Importancia del pre-procesamiento de las imágenes y su aplanamiento en tareas de visión computacional



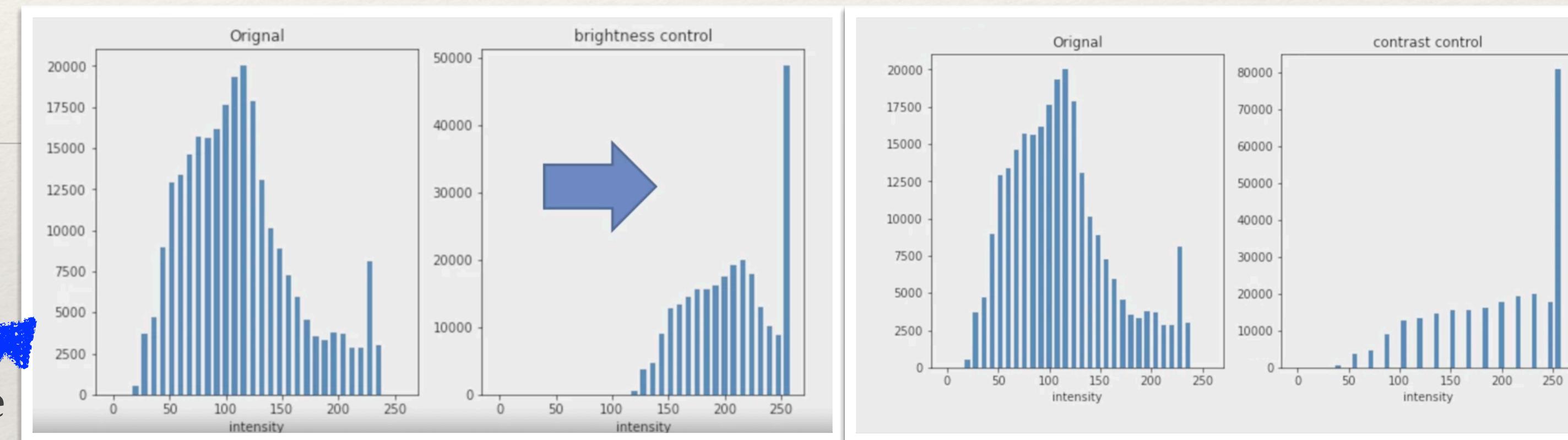
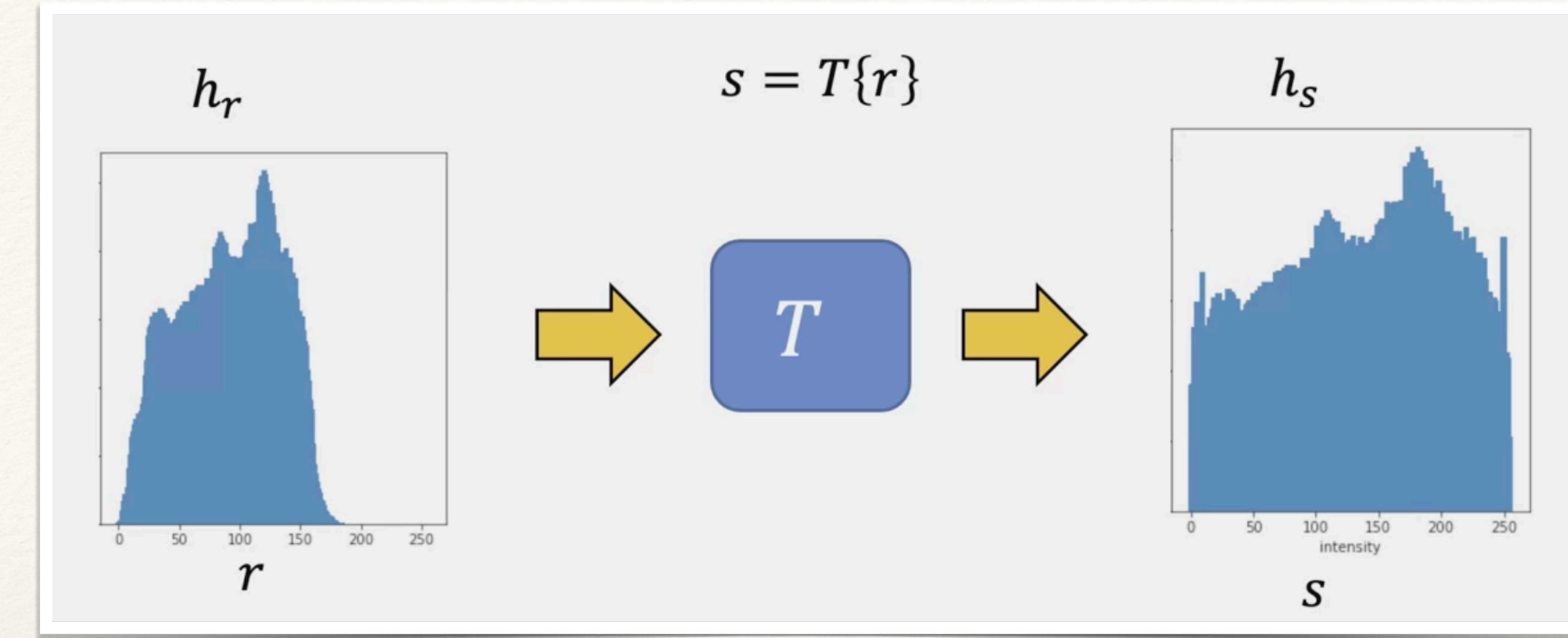
IBM Developer

$$\mathbf{x} = \begin{bmatrix} \cdot & \cdot & \cdot \\ \vdots & \vdots & \vdots \\ \cdot & \cdot & \cdot \end{bmatrix}$$

Transformaciones de pixeles

- ❖ Su efecto se puede evidenciar en el histograma
- ❖ **Tomar el negativo:** Reversar los colores de la imagen

$$g_{i,j} = 255 - f_{ij}$$
- ❖ **Contraste y brillo:** $g_{ij} = \alpha f_{ij} + \beta$ donde α es el control de contraste y β el control de brillo
- ❖ **Equalización de histograma:** Uso del histograma para ajustar brillo y contraste
- ❖ **Thresholding an image (Umbralizando):** Se le asigna 255 si es mayor que cierto umbral y se deja negro si no.



$$f(x, y) \rightarrow g(x', y')$$

Transformaciones geométricas: Afines

- ❖ Escalado: $x' = ax$
- ❖ Corrimiento: $x' = x + t_x$
- ❖ Cizallado: $x' = by$
- ❖ En resumen una transformación afín es:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix}$$

❖ Rotación: $\begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}$

| | | | | | | | | | | | | | | | | |
|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|---------|
| f[0,0] | f[0,1] | f[0,2] | f[0,3] | f[0,4] | f[0,5] | g[0,0] | g[0,1] | g[0,2] | g[0,3] | g[0,4] | g[0,5] | g[0,6] | g[0,7] | g[0,8] | g[0,9] | g[0,10] |
| f[1,0] | f[1,2] | f[1,2] | f[1,3] | f[1,4] | f[1,5] | g[1,0] | g[1,1] | g[1,2] | g[1,3] | g[1,4] | f[1,5] | g[1,6] | g[1,7] | g[1,8] | g[1,9] | g[1,10] |
| f[2,0] | f[2,2] | f[2,2] | f[2,3] | f[2,4] | f[2,5] | g[2,0] | g[2,1] | g[2,2] | g[2,3] | g[2,4] | g[2,5] | g[2,6] | g[2,7] | g[2,8] | g[2,9] | g[2,10] |
| f[3,0] | f[3,1] | f[3,2] | f[3,3] | f[3,4] | f[3,5] | g[3,0] | g[3,1] | g[3,2] | g[3,3] | g[3,4] | g[3,5] | g[3,6] | g[3,7] | g[3,8] | g[3,9] | g[3,10] |
| f[4,0] | f[4,1] | f[4,2] | f[4,3] | f[4,4] | f[4,5] | g[4,0] | g[4,1] | g[4,2] | g[4,3] | g[4,4] | g[4,5] | g[4,6] | g[4,7] | g[4,8] | g[4,9] | g[4,10] |
| f[5,0] | f[5,1] | f[5,2] | f[5,3] | f[5,4] | f[5,5] | g[5,0] | g[5,1] | g[5,2] | g[5,3] | g[5,4] | g[5,5] | g[5,6] | g[5,7] | g[5,8] | g[5,9] | g[5,10] |

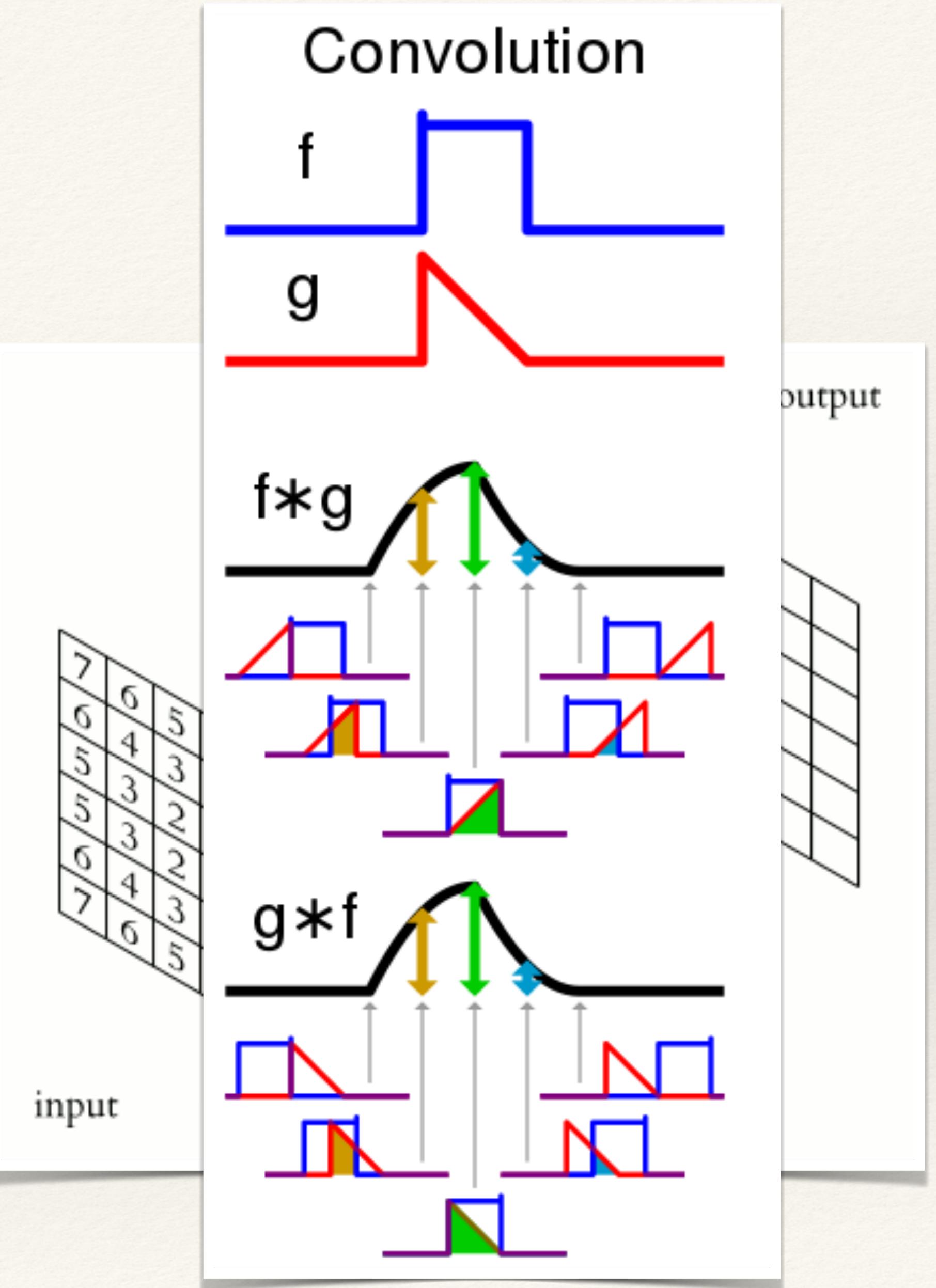
| | | | | | | | | |
|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| f[0,0] | f[0,1] | f[0,2] | f[0,3] | f[0,4] | f[0,5] | g[0,1] | g[0,2] | g[0,3] |
| f[1,0] | f[1,2] | f[1,2] | f[1,3] | f[1,4] | f[1,5] | g[1,1] | [1,2] | g[1,3] |
| f[2,0] | f[2,2] | f[2,2] | f[2,3] | f[2,4] | f[2,5] | g[2,1] | f[2,2] | g[2,3] |
| f[3,0] | f[3,1] | f[3,2] | f[3,3] | f[3,4] | f[3,5] | g[3,1] | f[3,2] | g[3,3] |
| f[4,0] | f[4,1] | f[4,2] | f[4,3] | f[4,4] | f[4,5] | g[4,1] | f[4,2] | g[4,3] |
| f[5,0] | f[5,1] | f[5,2] | f[5,3] | f[5,4] | f[5,5] | g[5,1] | f[5,2] | g[5,3] |

| | | |
|--------|--------|--------|
| g[0,1] | g[0,2] | g[0,3] |
| g[1,1] | [1,2] | g[1,3] |
| g[2,1] | f[2,2] | g[2,3] |
| g[3,1] | f[3,2] | g[3,3] |
| g[4,1] | f[4,2] | g[4,3] |
| g[5,1] | f[5,2] | g[5,3] |

Les llamaremos filtros, Kernels o operadores de SOBEL

Convolución

- ❖ Expresa cómo la forma de uno es modificada por el otro. Se define como la integral del producto de las dos funciones después de que una se refleje sobre el eje y y se desplace
- ❖ Si tengo $f(x = p)$ y quiero desplazar la función $g(x)$ a ese valor debo hacer $g(x - p)$
- ❖ A **funciones continuas**: $(f * g)(t) = \int_{-\infty}^{\infty} f(\tau)g(t - \tau)d\tau$
- ❖ A **funciones discretas**: $(f * g)[n] = \sum_{m=-\infty}^{\infty} f[m]g[n - m]$
- ❖ A **secuencias discretas**: $(f * g)[n] = \sum_{m=-M}^{M} f[m]g[n - m]$



Les llamaremos filtros, Kernels o operadores de SOBEL

Convolución

- ❖ Low pass filter: Suaviza la imagen quitando el ruido

Filtro promedio: $\begin{bmatrix} \frac{1}{N} & \dots & \frac{1}{N} \\ \dots & \dots & \dots \\ \frac{1}{N} & \dots & \frac{1}{N} \end{bmatrix}$

- ❖ Gradiente en x , G_x : Resultado del kernel

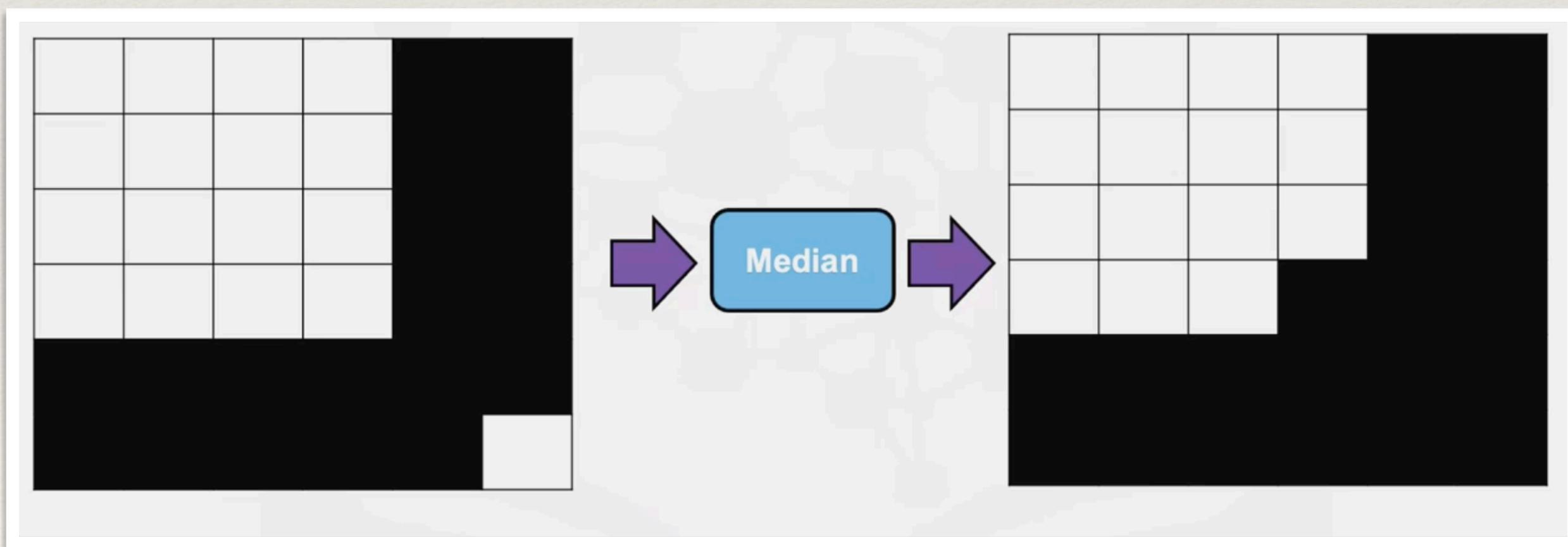
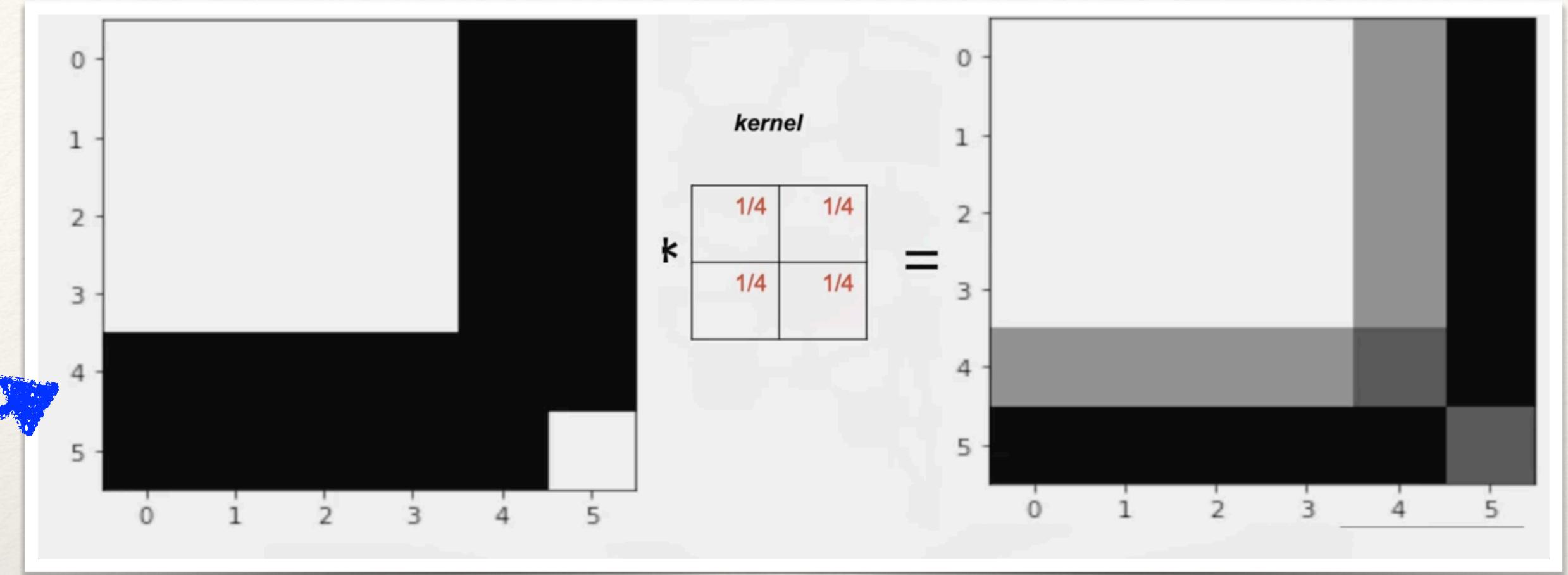
$$\begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix}$$

- ❖ Gradiente en y , G_y : Resultado del kernel

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix}$$

- ❖ Vector gradiente: $\mathbb{G} = [G_x, G_y]$ (G_x y G_y son matrices)

Filtro de mediana: $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$



Histograma de gradientes orientados

- ❖ Cuenta ocurrencias de orientaciones de gradientes en porciones localizadas de la imagen. Generando un histograma separado en cada una de éstas regiones
 - ❖ Encontrar los gradientes de la imagen (por filtrado). Si es una imagen de tres canales se toma por pixel el gradiente más alto entre los tres canales.
 - ❖ Dividir la imagen y por cada celda se calcula el histograma de la magnitud de los gradientes tomando los ángulos como bins
 - ❖ Se toman cuatro celdas y se normalizan. Se mueve una celda horizontalmente / verticalmente y se repite. Por cada celda grande se obtiene un vector.
 - ❖ Se concatenan todos los vectores

Tarea: Preparar una explicación de Scale-Invariant Feature Transform (SIFT) o de Speeded Up Robust Feature (SURF)