

# Taller Visión Artificial

José Carlos Rangel Ortiz, Ph.D.  
*jose.rangel@utp.ac.pa*

FISC-UTP

18 de octubre de 2023

## Simposio Nacional de Computación FISC 2023



# Outline

## 1 Introducción a la Visión Artificial

- Introducción a la Visión por Computadora
- Concepto de Imagen

## 2 Introducción a la Visión Artificial



Section 1.1

Introducción a la Visión Artificial

Introducción a la Visión por Computadora

# Introducción a la Visión por Computadora I

## Introducción a la Visión Artificial

- Desde su concepción se considera uno de las entradas que mayor información puede proporcionar a un sistema.
- Por ende, es de suma importancia el desarrollo de métodos que permitan extraer la mayor cantidad de información de este tipo de entrada.
- La VA busca emular la visión humana y producir de una imagen de entrada, una salida de alto nivel, tal como lo haría una persona



# Introducción a la Visión por Computadora II

## Introducción a la Visión Artificial

- Como ciencia se apoya en otras para lograr su cometido, a la vez que es utilizada por otras áreas para cumplir metas globales.
- En la actualidad es una de las ramas de la IA que es más notada por las personas.
- En el mundo cotidiano está presente en sistemas de vigilancia, cámaras fotográficas hasta en coches y robots autónomos, en los cuales cumple una función primordial.



# Introducción a la Visión por Computadora III

## Introducción a la Visión Artificial

### Definición

Se puede definir la “Visión Artificial” como un campo de la “Inteligencia Artificial” que, mediante la utilización de las técnicas adecuadas, permite la obtención, procesamiento y análisis de cualquier tipo de información especial obtenida a través de imágenes digitales.



# Introducción a la Visión por Computadora IV

## Introducción a la Visión Artificial

- La visión artificial, es un área de estudio multidisciplinaria, encargada de trabajar con imágenes (videos o imágenes estáticas).
- Obteniendo de estas entradas, datos significativos para tomar decisiones, como las que realizan las aplicaciones tradicionales en las computadoras
- Para la visión artificial o la implementación de este sentido en las computadoras, la entrada “cruda” de datos, se reduce a no más que a una matriz de números, representado de forma discreta algunos de los ejemplares de rayos de luz captados en un momento determinado.



# Introducción a la Visión por Computadora V

## Introducción a la Visión Artificial

- Este área, involucra varias técnicas y métodos para hacer posible que la computadora tenga, en cierto aspecto, un sentido de la vista similar al del ser humano.
- En la actualidad la Visión por Computadora confía plenamente en algoritmos de *Machine Learning* y *Deep Learning* para la construcción de aplicaciones de visión artificial.
- También se compone de varias tecnologías que trabajan de manera conjunta como lo son: Gráficos por Computadora, Procesamiento de Imágenes, Procesamiento de Señales, Tecnología de Sensores, Matemáticas e incluso Física.



Introducción a la Visión Artificial

## Section 1.2 | Concepto de Imagen

# Formulación de una Imagen I

## Concepto de Imagen

- Una imagen puede ser descrita como una función 2D,  $f(x, y)$
- $(x, y)$  representan las coordenadas espaciales
- El valor de  $f$  en un punto  $(x, y)$  es proporcional al brillo o nivel de gris de la imagen (**Imagen en escala de Grises**).
- Cuando  $(x, y)$  y el brillo ( $f$ ) son cantidades discretas finitas, la imagen es llamada **digital**.
- Por lo tanto  $f(x, y)$  tomaría los siguientes valores:
  - $x \in [0, h - 1]$ , donde  $h$  representa la altura de la imagen
  - $y \in [0, w - 1]$ , donde  $w$  representa la anchura de la imagen
  - $f(x, y) \in [0, L - 1]$  , donde  $L = 256$  para una imagen de 8 – bit



# Formulación de una Imagen II

## Concepto de Imagen

- **Imagen de color** se puede representar de la misma manera, pero utilizando 3 funciones para representar los valores de los colores rojo, verde y azul.
- Cada una de estas funciones seguirá la misma formulación de  $f(x, y)$  definida para la escala de grises.
- Se suelen identificar con las letras  $R$ ,  $G$ ,  $B$  y sus formulas pueden ser  $fR(x, y)$ ,  $fG(x, y)$  y  $fB(x, y)$  respectivamente.



# Formulación de una Imagen III

## Concepto de Imagen

- Para **Imágenes en blanco y negro** se sigue la aproximación de una sola función.
- La única diferencia radica en que en este caso el valor de la función solo puede ser 0 (negro) o 255 (blanco)



# Formulación de una Imagen IV

## Concepto de Imagen

- Las **imágenes digitales** se pueden considerar como representaciones de escenas reales ya que utilizan cantidades discretas finitas.
- Para imágenes en escala de grises e imágenes en blanco y negro, se tiene una sola muestra para cada par  $(x, y)$ .
- Para imágenes a color se tendrán 3 muestras para cada par de coordenadas.



# Píxeles I

## Concepto de Imagen

- Un **Pixel** es la mínima unidad que conforma una imagen digital.
- Cada imagen esta compuesta por un gran conjunto de estos elementos.
- Se corresponde con cada uno de los pares de coordenadas  $(x, y)$  con las cuales se formula una imagen.



# Píxeles II

## Concepto de Imagen

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	119	166	225	217	179	110	0	0	0	0	0	0	0
0	0	0	0	0	81	0	0	0	32	169	94	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	101	213	30	0	0	0	0	0	0
0	0	0	0	0	0	0	113	187	149	31	0	0	0	0	0	0	0
0	0	0	0	0	127	238	226	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	67	91	95	135	213	124	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	18	214	62	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	15	227	57	0	0	0	0	0	0
0	0	0	0	0	0	0	73	147	229	135	0	0	0	0	0	0	0
0	0	0	0	0	46	162	190	241	215	170	49	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0



# Píxeles III

## Concepto de Imagen

- Cada unidad tendrá asociada un valor producido por la función de color de la imagen.
- Utilizando la notación  $ANCHO \times ALTO \times CANALES$  ( $w \times h \times channel$ ) se suelen expresar las características básicas de un archivo de imagen.



## Canales de Color I

## Concepto de Imagen

Tomando en cuenta lo anterior, una imagen tendrá asociado un modelo de color. Este se puede ver de forma similar a la formulación de un imagen, pero en cada canal se representará un solo color.

# Resolución I

## Concepto de Imagen

- Si una imagen tiene una resolución de  $800 \times 1200$ , esta es una matriz/grid/malla con 800 columnas y 1200 filas.
- La cantidad de píxeles de una imagen se calcula:  $800 \times 1200 = 960\,000$ .
- Esta cantidad de píxeles, no indica las dimensiones físicas de la imagen.
- La resolución es el número de píxeles por pulgada que contiene y se expresa en *ppp* puntos por pulgada o en inglés *DPI* (*dots per inch*).
- Una imagen con resolución de *300 dpi* indica que tiene 90 000 píxeles en un área de impresión de 1 pulgada cuadrada.



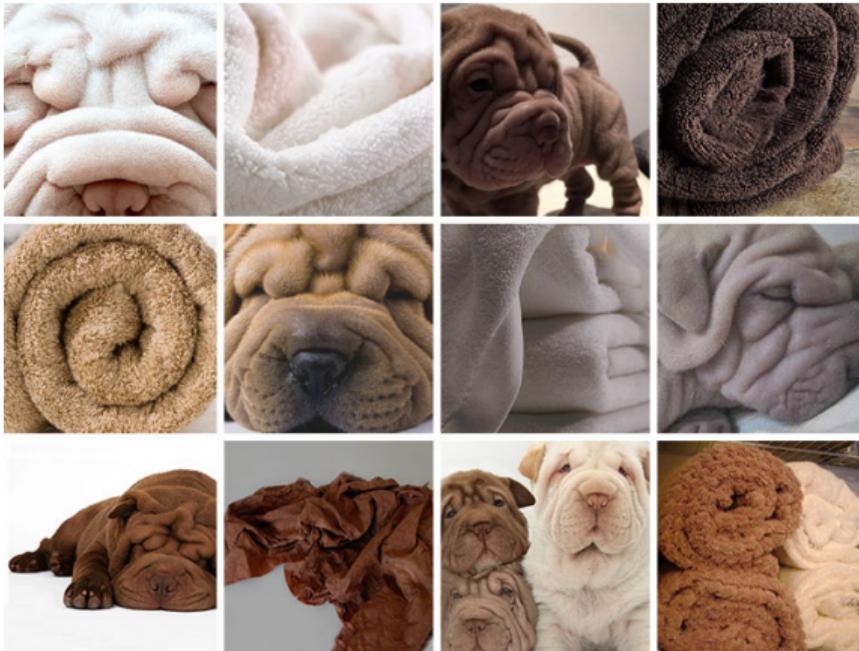
# Problemática del Trabajo con Imágenes I

## Concepto de Imagen



# Problemática del Trabajo con Imágenes II

## Concepto de Imagen



# Problemática del Trabajo con Imágenes III

## Concepto de Imagen



# Problemática del Trabajo con Imágenes IV

## Concepto de Imagen



# Problemática del Trabajo con Imágenes V

## Concepto de Imagen



# Problemática del Trabajo con Imágenes VI

## Concepto de Imagen



# Outline

1 Introducción a la Visión Artificial

2 Introducción a la Visión Artificial

- OpenCV



Introducción a la Visión Artificial

## Section 2.9 | OpenCV

# OpenCV I

## Introducción a la Visión Artificial

- Es una librería *open source* de visión por computador, análisis de imagen y con algunas herramientas de aprendizaje automático.
- Para ello dispone de infinidad de algoritmos que permiten, con sólo escribir unas pocas líneas de código, identificar rostros, reconocer objetos, clasificarlos, detectar movimientos de manos, etc.
- Multiplataforma
- Puede programarse con C, C++, Python, Java, Matlab, entre otros.
- Se ha convertido en la librería por defecto para el desarrollo de aplicaciones con Visión Artificial.



# OpenCV II

## Introducción a la Visión Artificial

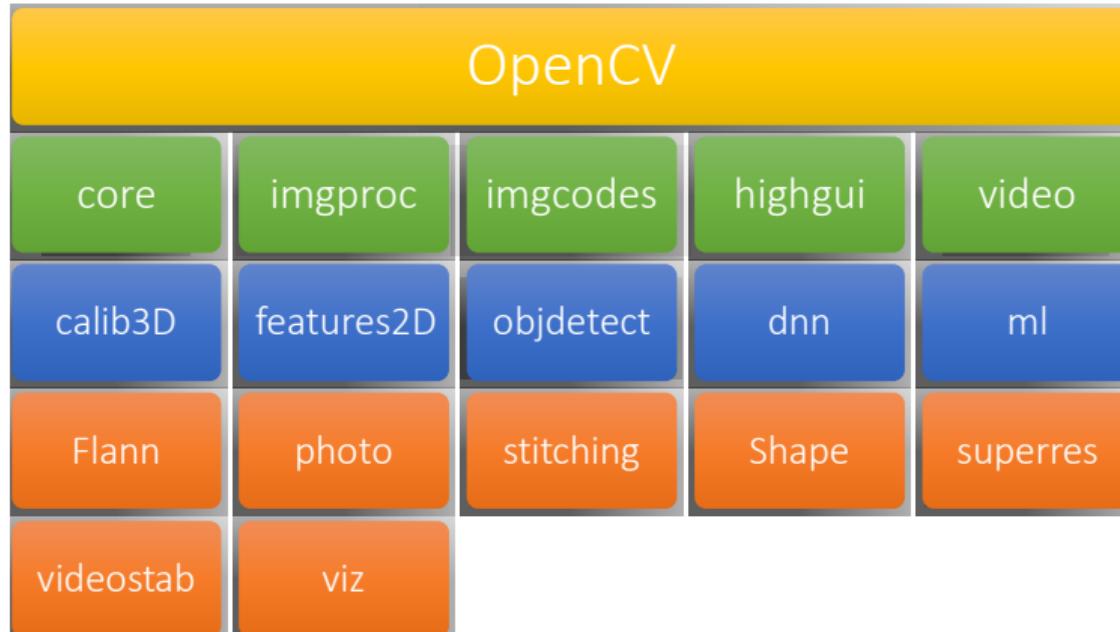
Incluye los siguientes modulos

- Procesamiento de Imágenes
- Análisis de Vídeo
- Calibración de Cámaras
- Reconstrucción 3D
- Características 2D
- Detección de Objetos
- Interfaces de alto nivel
- Vídeo I/O



# OpenCV III

## Introducción a la Visión Artificial



Dentro de sus aplicaciones se puede encontrar

- Interacción Humano Computador - HCI
- Identificación de Objetos
- Reconocimiento de Objetos
- Reconocimiento de Rostros
- Reconocimiento de Gestos
- Rastreo de Movimiento
- Procesamiento de Imágenes
- Robótica Móvil
- Realidad Aumentada
- Registro de imágenes de vista de calle
- Vigilancia Automatizada
- Análisis de Imágenes Médicas
- *Egomotion*
- ... entre otras



### Características

- Puede procesar al rededor de 30 frames por segundo
- Se puede integrar con diferentes entornos de desarrollo
- Optimizada para la ejecución
- Gran comunidad de desarrolladores que también dan soporte
- Soporte para el *Machine Learning*
- Cuenta con más de 2500 algoritmos optimizados y del estado del arte en visión computacional



# Convoluciones I

## Introducción a la Visión Artificial

- El filtrado de imágenes consiste en la modificación de las matrices correspondientes a la imagen digitalizada mediante el procedimiento que se halla diseñado.
- Uno de los procedimientos más utilizado es la llamada **convolución de matrices**.
- Convolución es el tratamiento de una matriz por otra que se llama “**kernel**” .



# Convoluciones II

## Introducción a la Visión Artificial

- Son útiles para extraer la información relevante de una imagen.
- Es un procesamiento matemático sobre los píxeles de una imagen, con la intención de resaltar ciertos tipos de relaciones.
- Matemáticamente una **Máscara de Convolución** o filtros espaciales, son ventanas (o matrices) con valores dados por el tipo de relación que se desea resaltar.



# Convoluciones III

## Introducción a la Visión Artificial

- \* Un ejemplo típico de estas ventanas puede ser el siguiente, el cual es un filtro especial de pasabajos.

$$\frac{1}{9} \times \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$



# Convoluciones IV

## Introducción a la Visión Artificial

- \* El proceso de aplicar un filtro espacial a una imagen se explica mediante el proceso de convolución, la cual esta dada por la siguiente ecuación:

$$f(x, y) * g(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\alpha, \beta)g(x - \alpha, y - \beta)d\alpha d\beta$$

- \* Donde:
  - $x$  y  $y$  son las coordenadas del pixel
  - $\alpha$  y  $\beta$  son variables para controlar el desplazamiento de la matriz de convolución



# Convoluciones V

## Introducción a la Visión Artificial

- \* De igual manera el resultado de la convolución para un pixel de la imagen resultado se puede definir utilizando la siguiente formula:

$$p_{conv}(x, y) = \sum_{s=-a}^a \sum_{t=-b}^b g(s, t) f(x + s, y + t)$$

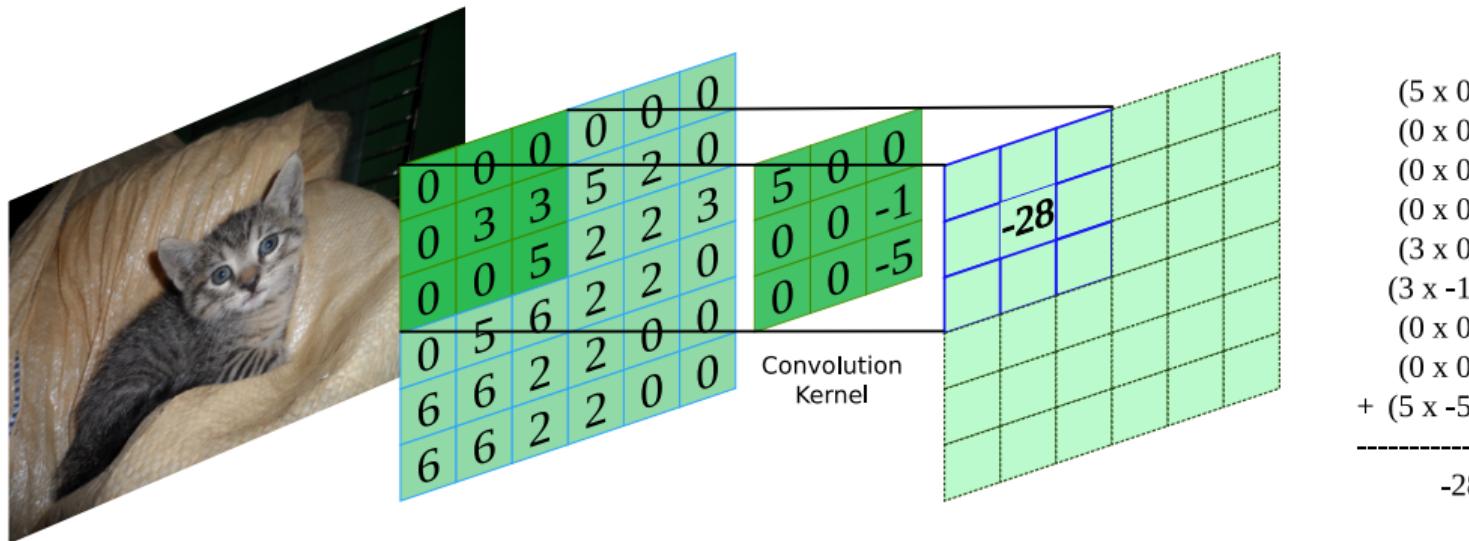
- \* Por lo cual se puede decir que el resultado para un pixel, es la sumatoria de los productos de los valores del filtro ( $g$ ) por los valores de los píxeles cercanos al este.



# Convoluciones VI

## Introducción a la Visión Artificial

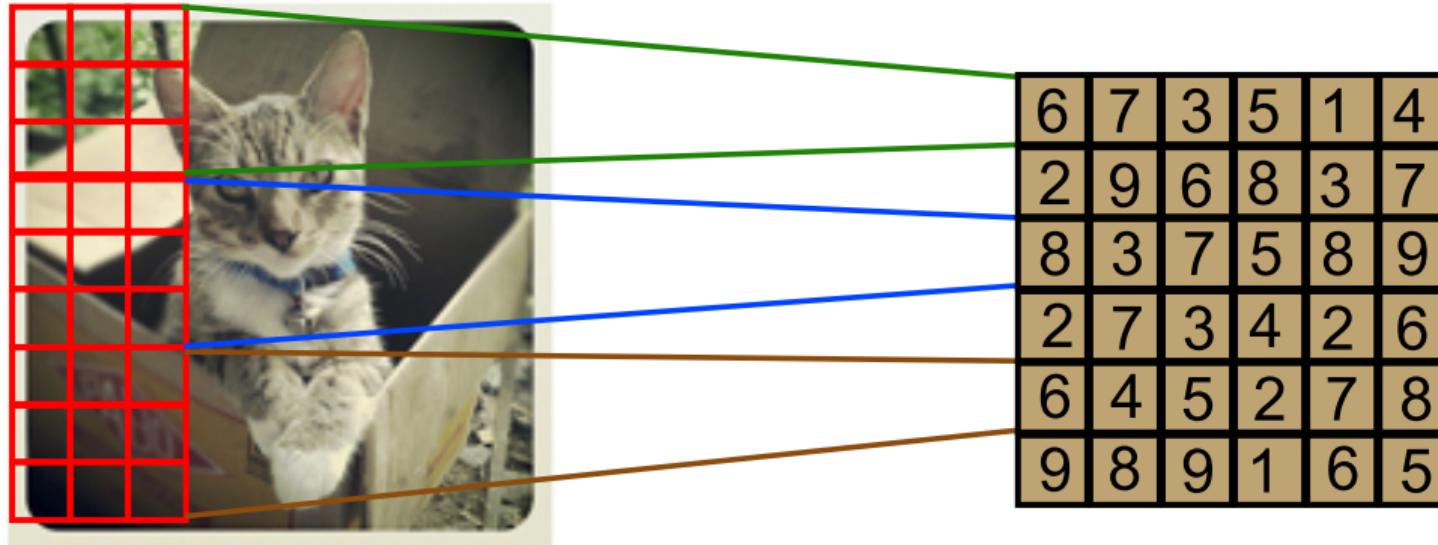
- De manera gráfica una convolución se podría representar de la siguiente manera



# Convoluciones VII

## Introducción a la Visión Artificial

- Los píxeles de la imagen resultante se originarían según el siguiente esquema.



# Convoluciones VIII

## Introducción a la Visión Artificial

### Ejemplo de Convolución

- Teniendo la imagen  $f(x, y)$  y el filtro  $g(x, y)$ , respectivamente.

$$f(x, y) = \begin{bmatrix} 17 & 24 & 1 & 8 & 15 \\ 23 & 5 & 7 & 14 & 16 \\ 4 & 6 & 13 & 20 & 22 \\ 10 & 12 & 19 & 21 & 3 \\ 11 & 18 & 25 & 2 & 9 \end{bmatrix} \quad g(x, y) = \begin{bmatrix} 1/9 & 1/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \end{bmatrix}$$



# Convoluciones IX

## Introducción a la Visión Artificial

### Ejemplo de Convolución

- El resultado de la Convolución  $f * g$  sería el siguiente

$$f * g = \begin{bmatrix} 11 & 10 & 12 \\ 11 & 13 & 15 \\ 13 & 15 & 14 \end{bmatrix}$$

- \* Se trabajan con imágenes discretizadas, por lo tanto, cuando se obtiene un valor decimal, este se trunca.



# Convoluciones X

## Introducción a la Visión Artificial

### Ejemplo de Convolución

- Realizando los cálculos para el elemento de la fila 2, columna 1 de la imagen resultado, se obtendría lo siguiente

$$\begin{bmatrix} 23 & 5 & 7 \\ 4 & 6 & 13 \\ 10 & 12 & 19 \end{bmatrix} * \begin{bmatrix} 1/9 & 1/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \end{bmatrix}$$

$$\frac{23 + 5 + 7 + 4 + 6 + 13 + 10 + 12 + 19}{9} = \frac{99}{9} = 11$$



# Convoluciones XI

## Introducción a la Visión Artificial

- Ejemplos de máscaras de filtros más utilizadas.

Enfoque

$$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$

Desenfoque

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

Realce de Bordes

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

Repujado

$$\begin{bmatrix} -2 & -1 & 0 \\ -1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix}$$



# Convoluciones XII

## Introducción a la Visión Artificial

Detección de Bordes

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Filtro de Tipo Sobel

$$\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

Filtro de Tipo Sharpen

$$\begin{bmatrix} 1 & -2 & 1 \\ -2 & 5 & -2 \\ 1 & -2 & 1 \end{bmatrix}$$

Filtro Norte

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & -2 & 1 \\ -1 & -1 & -1 \end{bmatrix}$$

Filtro Este

$$\begin{bmatrix} -1 & 1 & 1 \\ -1 & -2 & 1 \\ -1 & 1 & 1 \end{bmatrix}$$

Filtro de Tipo Gauss

$$\begin{bmatrix} 1 & 2 & 3 & 1 & 1 \\ 2 & 7 & 11 & 7 & 2 \\ 3 & 11 & 17 & 11 & 3 \\ 2 & 7 & 11 & 7 & 1 \\ 1 & 2 & 3 & 2 & 1 \end{bmatrix}$$

# Filtros No Lineales I

## Introducción a la Visión Artificial

- **Filtro de Máximo:** Selecciona el mayor valor dentro de una ventana ordenada de valores de nivel de gris.

**Ventaja** Elimina el ruido pimienta (píxeles negros).

**Inconveniente** Funciona solo con el ruido pimienta y tiende a aclarar la imagen

- **Filtro de Mínimo:** Selecciona el menor valor dentro de una ventana ordenada de valores de nivel de gris.

**Ventaja** Elimina el ruido sal (píxeles blancos).

**Inconveniente** Funciona solo con el ruido es sal y tiende a oscurecer la imagen



# Filtros No Lineales II

## Introducción a la Visión Artificial

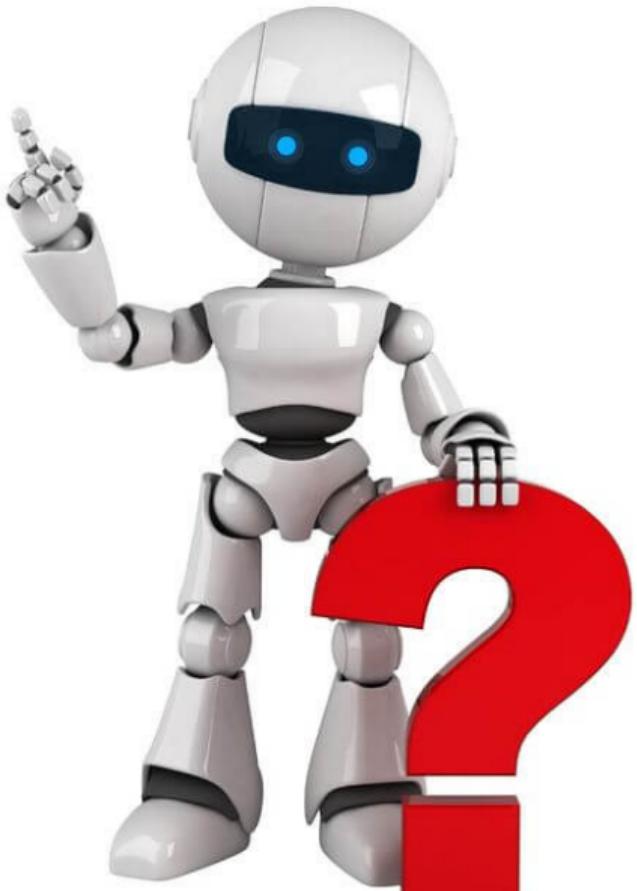
### Ejemplo de Aplicación del Filtro de Máximos

Original



MAX/Dilate





*Muchas Gracias por su atención.*

¿Preguntas?

Información de contacto:

✉: jose.rangel@utp.ac.pa

🐦: jcarlos2289

㏌: jose-carlos-rangel-ortiz-75482721

# Taller Visión Artificial

José Carlos Rangel Ortiz, Ph.D.  
*jose.rangel@utp.ac.pa*

FISC-UTP

18 de octubre de 2023

## Simposio Nacional de Computación FISC 2023

