

Artificial Intelligence

Lecture 08 – Visión por Computadora
en 180 min

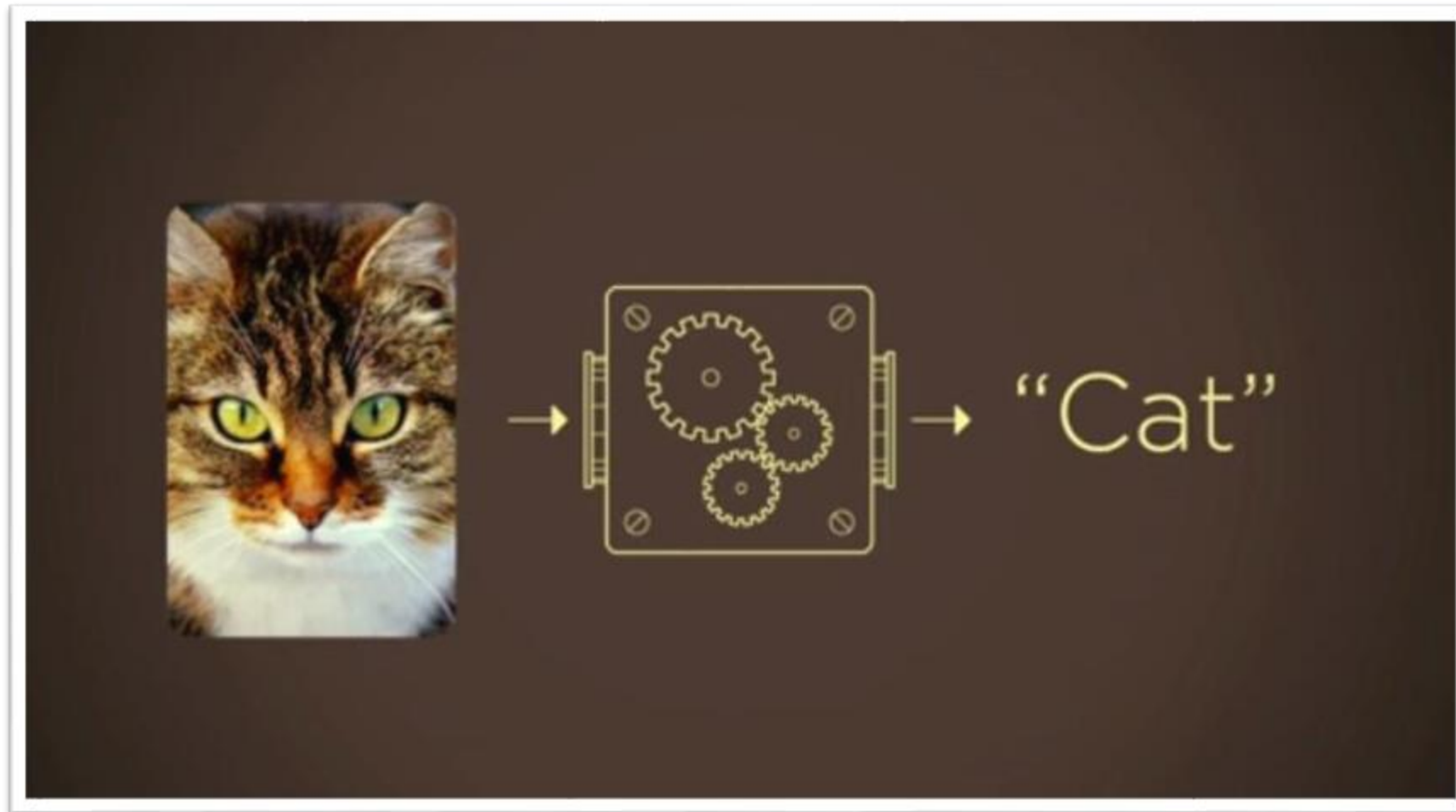


Objetivos

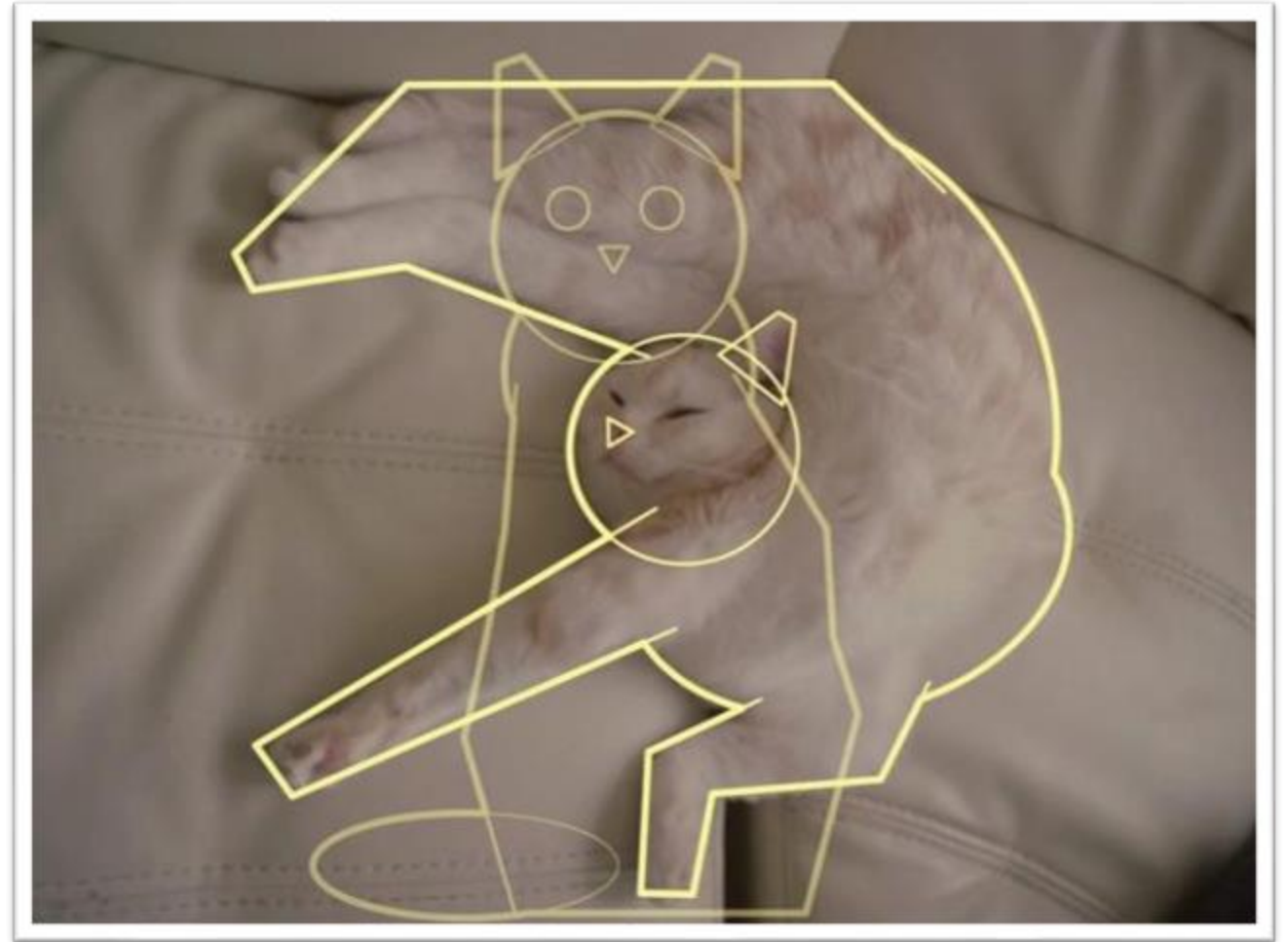
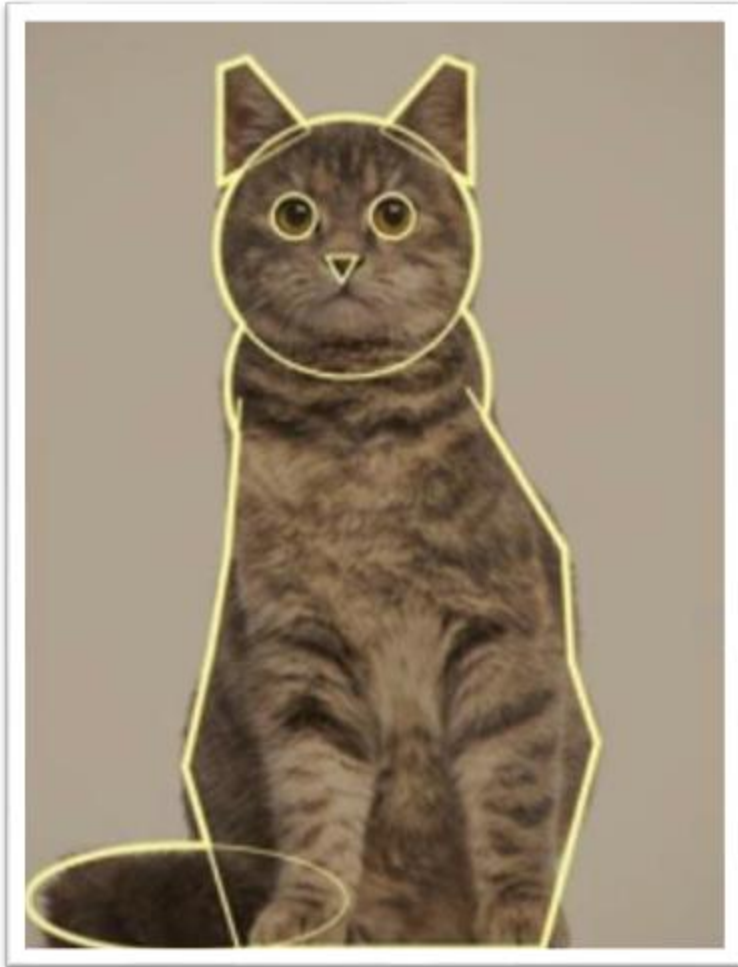
- I.Describir la estructura y las etapas de un sistema de visión artificial.
- II.Estudiar las técnicas fundamentales de las diferentes etapas de un sistema de visión artificial.
- III.Aplicar las técnicas y algoritmos descritos en un problema específico de la vida real.



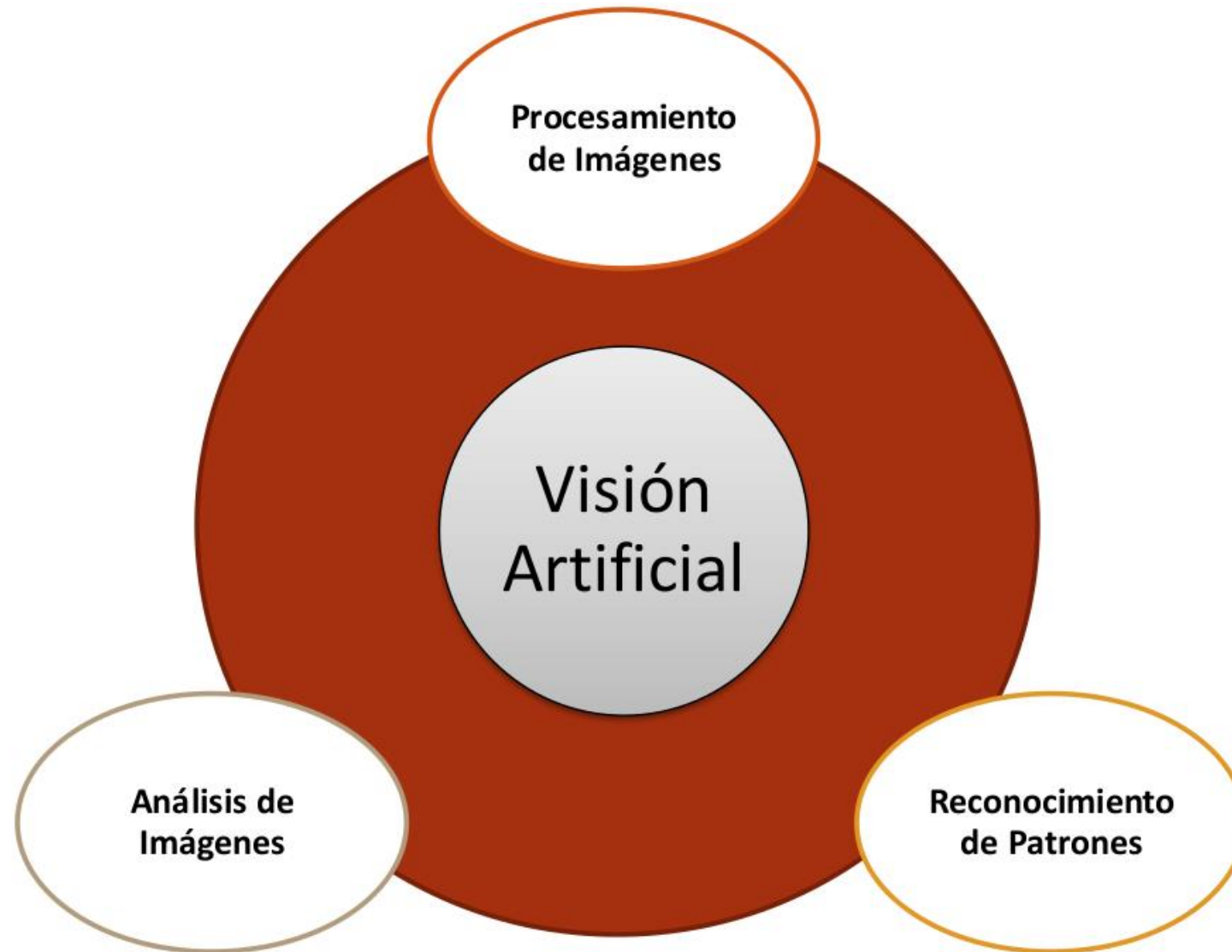
¿QUÉ ES LA VISIÓN ARTIFICIAL?



¿QUÉ ES LA VISIÓN ARTIFICIAL?



¿QUÉ ES LA VISIÓN ARTIFICIAL?



¿QUÉ ES LA VISIÓN ARTIFICIAL?

I. PROCESAMIENTO DE IMÁGENES:

- Es un proceso mediante el cual se toma una imagen y se produce una versión modificada de esta imagen

II. ANÁLISIS DE IMÁGENES:

- Proceso mediante el cual a partir de una imagen se obtiene una medición de los objetos en la imagen.

III. RECONOCIMIENTO DE PATRONES:

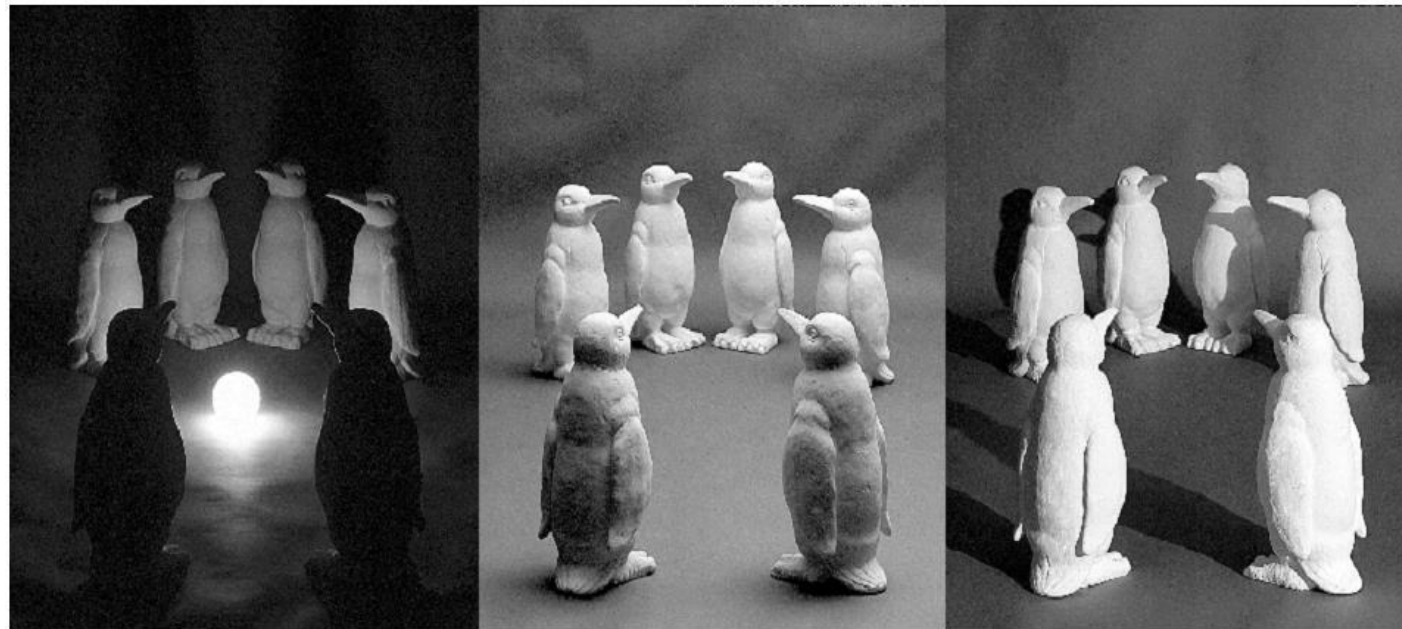
- Proceso mediante el cual a partir de una imagen se obtiene una medición, y se hace una interpretación y/o se toma decisión.



PROBLEMAS EN LA VISIÓN ARTIFICIAL

Dotar a las Máquinas con la capacidad de “ver” NO es una tarea fácil

1. Cambios de Iluminación



PROBLEMAS EN LA VISIÓN ARTIFICIAL



Dotar a las Máquinas con la capacidad de “ver” NO es una tarea fácil

2. Cambios en la Escala

¿Podríamos considerar que estos objetos son "iguales" a pesar de su tamaño diferente?

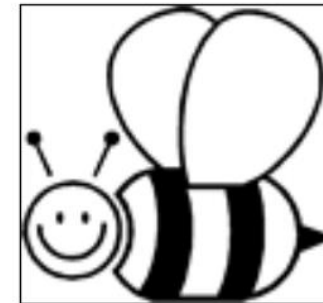


PROBLEMAS EN LA VISIÓN ARTIFICIAL

Dotar a las Máquinas con la capacidad de “ver” NO es una tarea fácil

3. Deformación

¿Podríamos considerar que estos objetos son "iguales" a pesar de que hay una deformación?



PROBLEMAS EN LA VISIÓN ARTIFICIAL

Dotar a las Máquinas con la capacidad de “ver” NO es una tarea fácil

4. Oclusión de objetos



PROBLEMAS EN LA VISIÓN ARTIFICIAL



Dotar a las Máquinas con la capacidad de “ver” NO es una tarea fácil

5. Movimiento

A pesar del movimiento, ¿Podríamos identificar los objetos?



PROBLEMAS EN LA VISIÓN ARTIFICIAL

Dotar a las Máquinas con la capacidad de “ver” NO es una tarea fácil

6. Distinción de Objetos



ETAPAS EN UN SISTEMA DE VISION ARTIFICIAL



ADQUISICIÓN DE LA IMAGEN

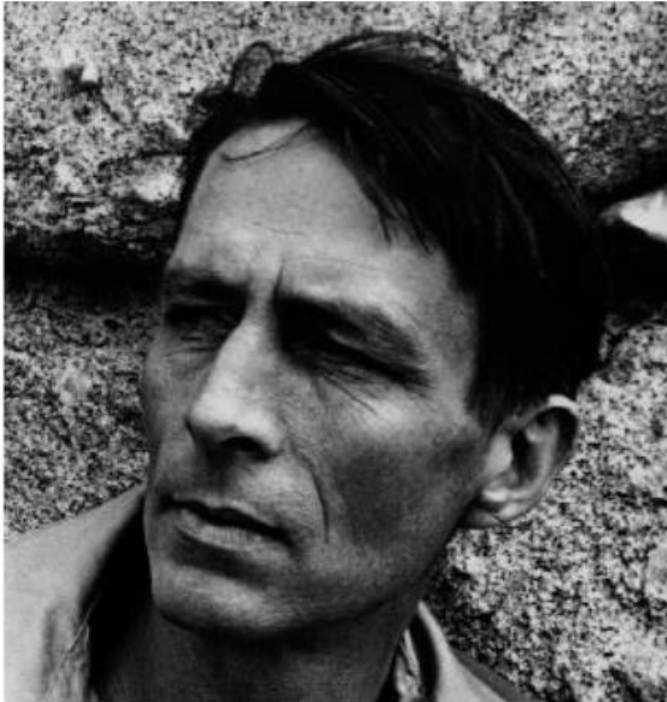


Imagen Original



Cada 3 Pixels



Cada 10 Pixels



PREPROCESAMIENTO

Consiste, en general, en tomar un arreglo de pixels como entrada y producir otro arreglo de pixels como salida, el cual de alguna manera representa una mejora del arreglo original.



Imagen Original

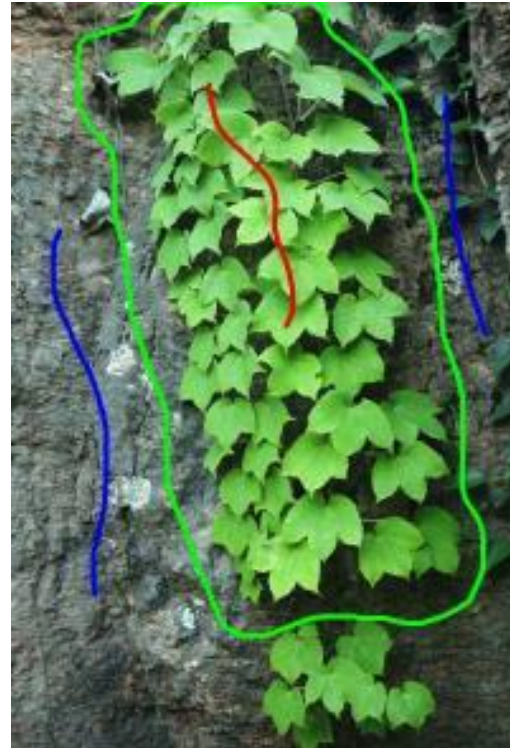
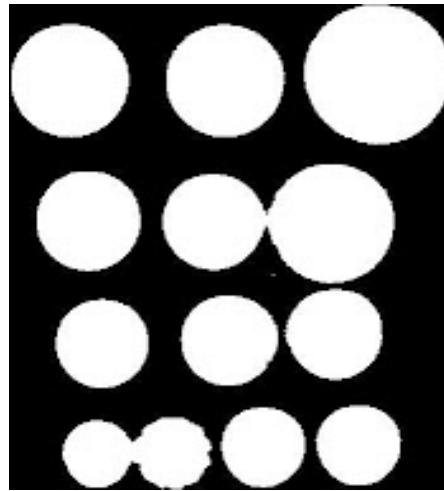


Eliminar ruido



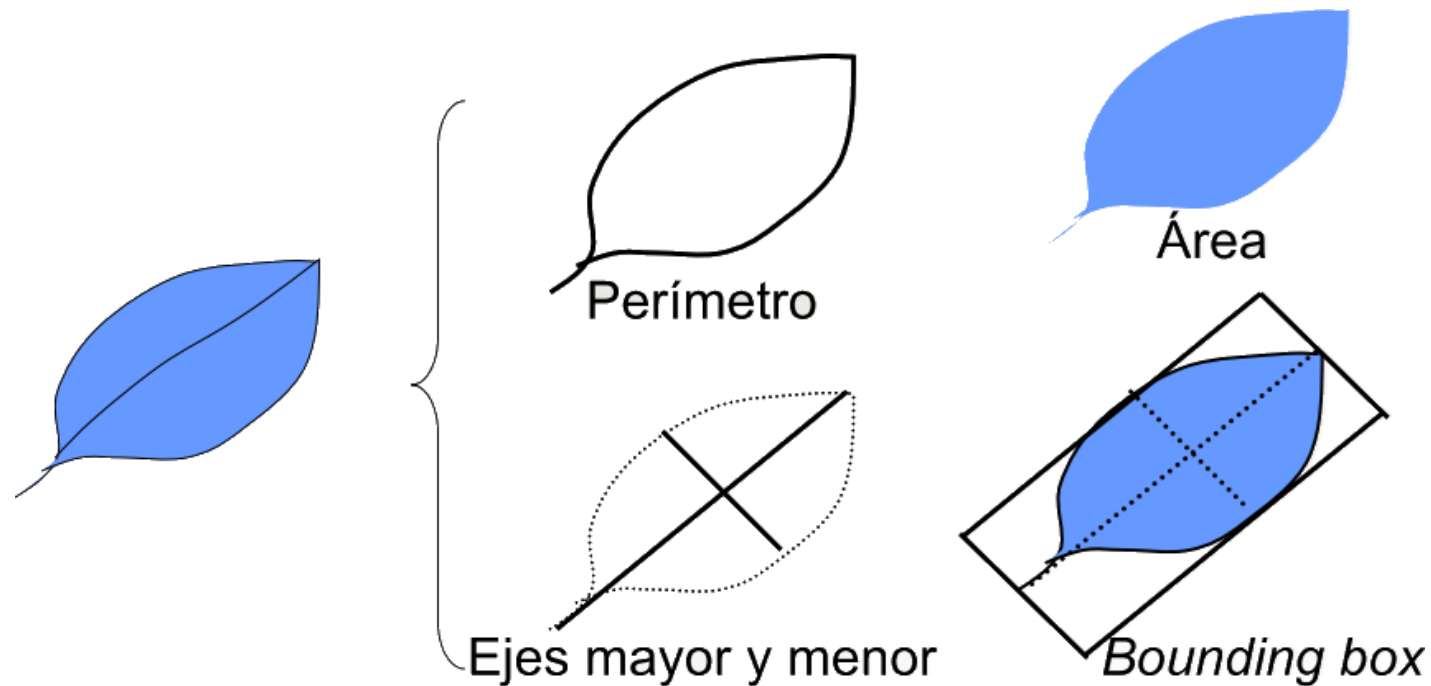
SEGMENTACIÓN

Se divide la imagen en regiones homogéneas que corresponden con los objetos contenidos en ella.

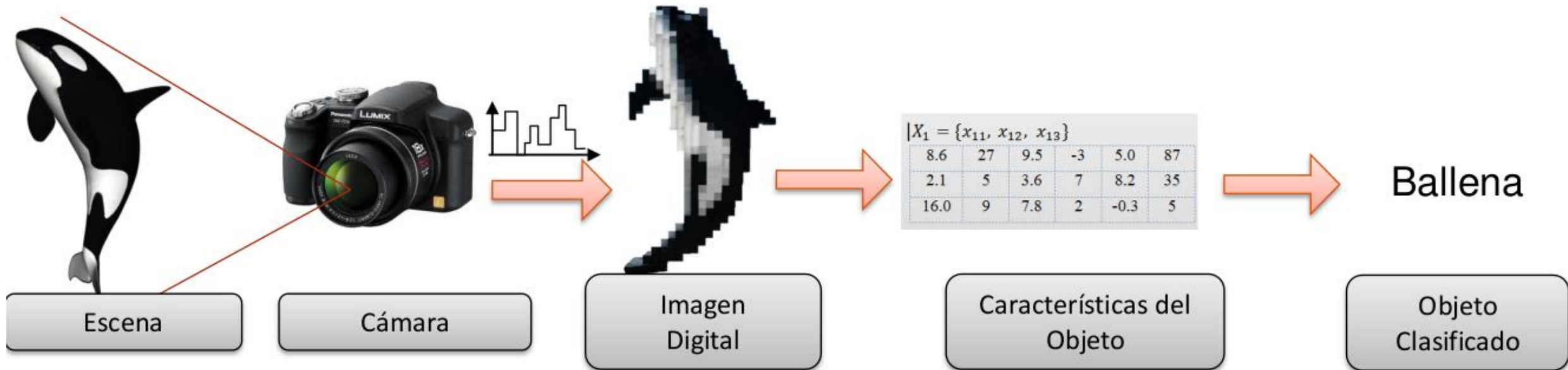


EXTRACCIÓN DE CARACTERÍSTICAS

Se obtienen medidas de características de los objetos segmentados, por ejemplo, características de color, de textura y/o de forma (área, perímetro, número de agujeros, ...)

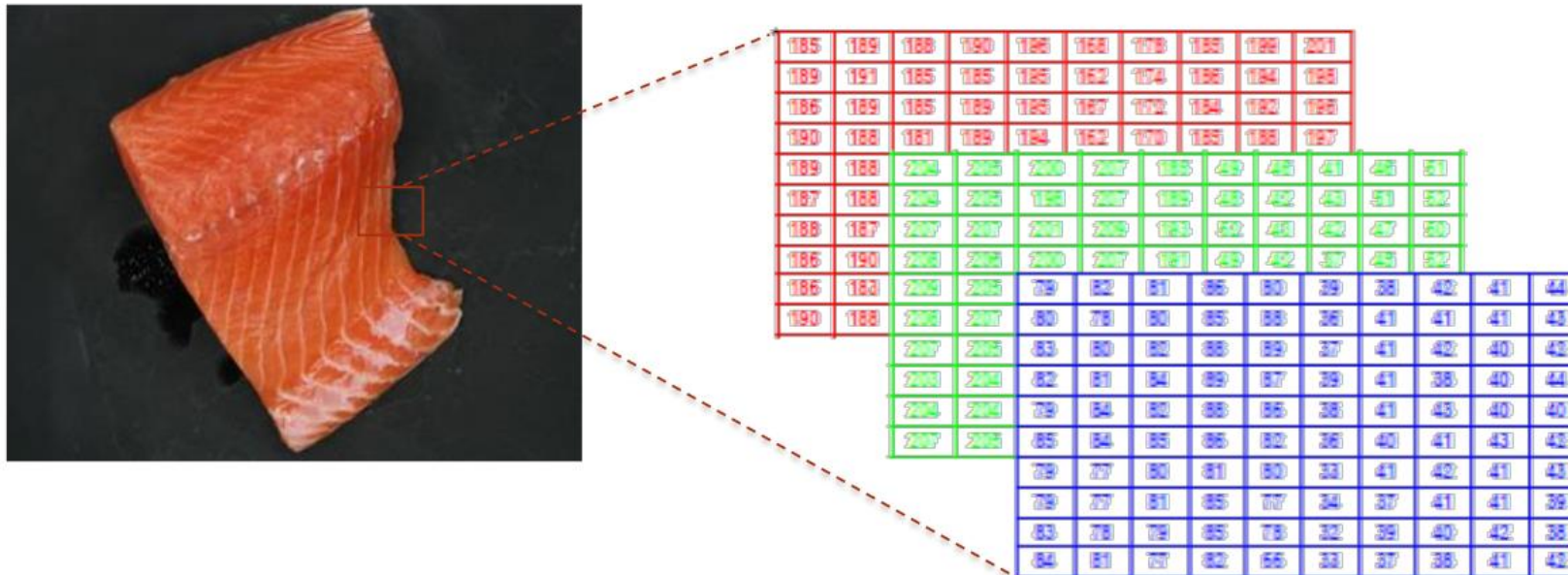


EL RESULTADO



LA IMAGEN DIGITAL

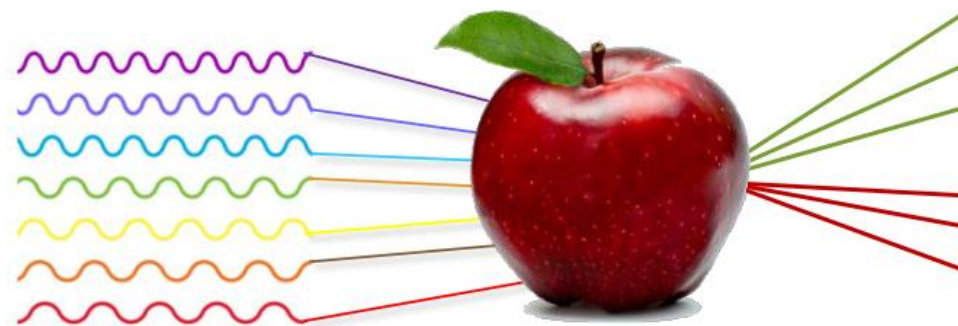
Las imágenes a color se representan usando tres canales: uno para el rojo, uno para el verde y otro para el azul (Canal RGB)



MODELOS DE COLOR

Es una especificación de un sistema de coordenadas en el que cada color está representado por un único punto. Existen varias representaciones o modelos de color. Estos modelos los podemos dividir en dos clases:

1. Unos son los modelos que están más orientados a los equipos, por ejemplo las cámaras, monitores y televisores, a los que llamaremos Modelos Sensoriales.
2. Otros son los modelos que se asemejan más a la percepción humana y que, en general, están orientados al procesamiento de imágenes y visión, éstos se denominan Modelos Perceptuales.



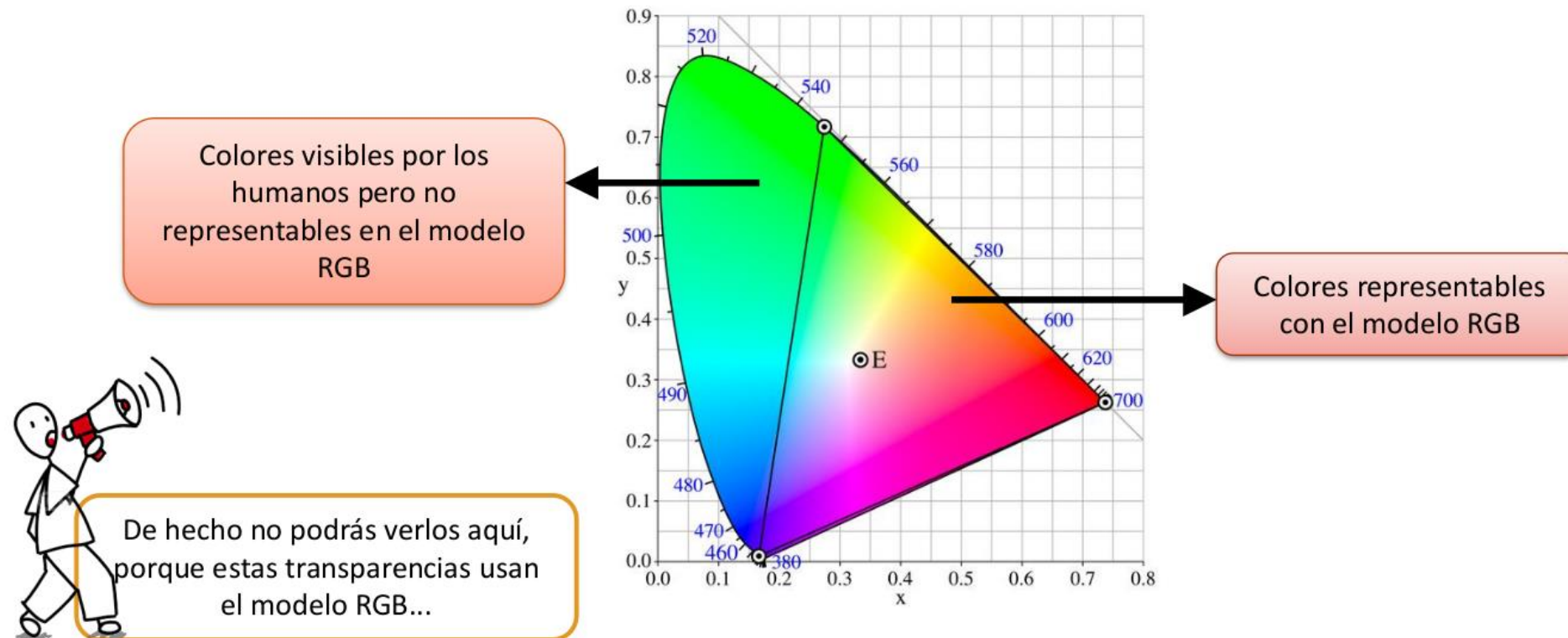
Luz Blanca



MODELOS DE COLOR SENSORIALES

Existen 4 modelos comúnmente utilizados

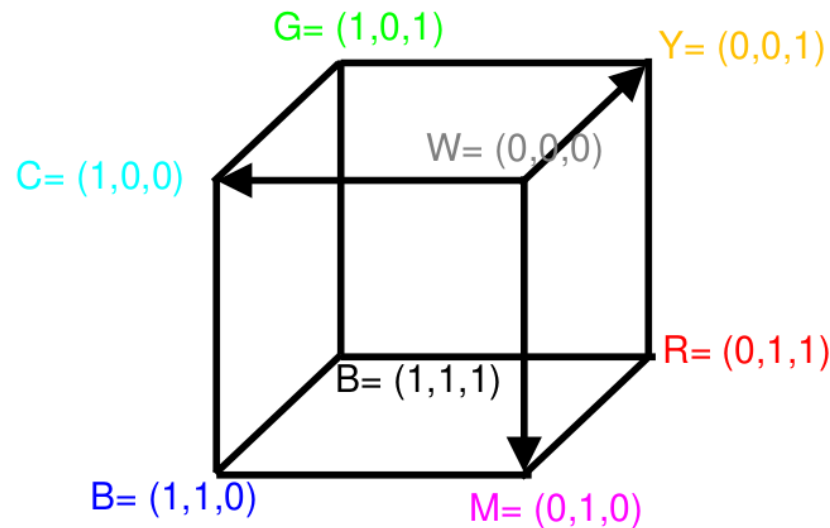
1. RGB: Es un modelo aditivo, que debe su nombre a las iniciales de los tres colores primarios: **Red** (rojo), **Green** (verde) y **Blue** (azul).



MODELOS DE COLOR SENSORIALES

Existen 4 modelos comúnmente utilizados

2. CMY: es usado ciertas aplicaciones, como por ejemplo en pintura e impresión de imágenes. El modelo CMY (o CMYK) está basado en un modelo sustractivo y en la práctica, no llega a producir negro, sino una especie de gris.



Conversión RGB → CMY:

$$\begin{aligned}C &:= 255 - R \\M &:= 255 - G \\Y &:= 255 - B\end{aligned}$$

Conversión CMY → RGB:

$$\begin{aligned}R &:= 255 - C \\G &:= 255 - M \\B &:= 255 - Y\end{aligned}$$

1 o 255?



MODELOS DE COLOR SENSORIALES

Existen 4 modelos comúnmente utilizados

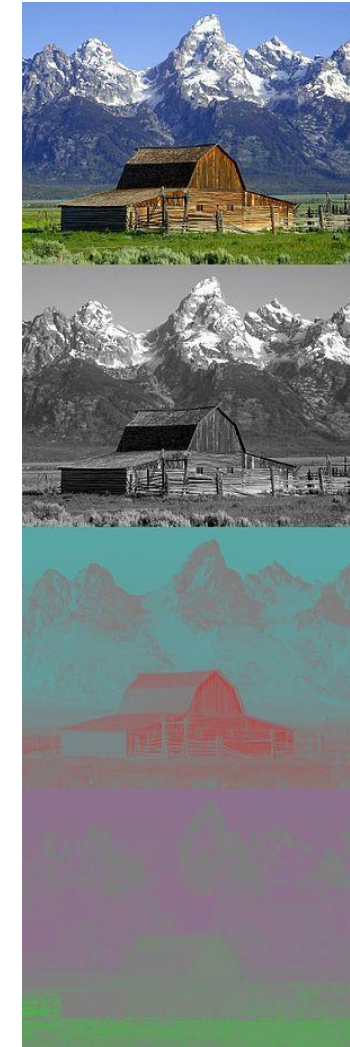
3. YIQ: separa la información de intensidad o luminancia (Y) de la información de color (I, Q). Es un modelo que antiguamente usaba el estándar de televisión NTSC.

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.274 & -0.322 \\ 0.211 & -0.523 & 0.312 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$



$$Y = 0.299R + 0.587G + 0.114B$$

La componente Y es el único componente usado por los televisores en blanco y negro

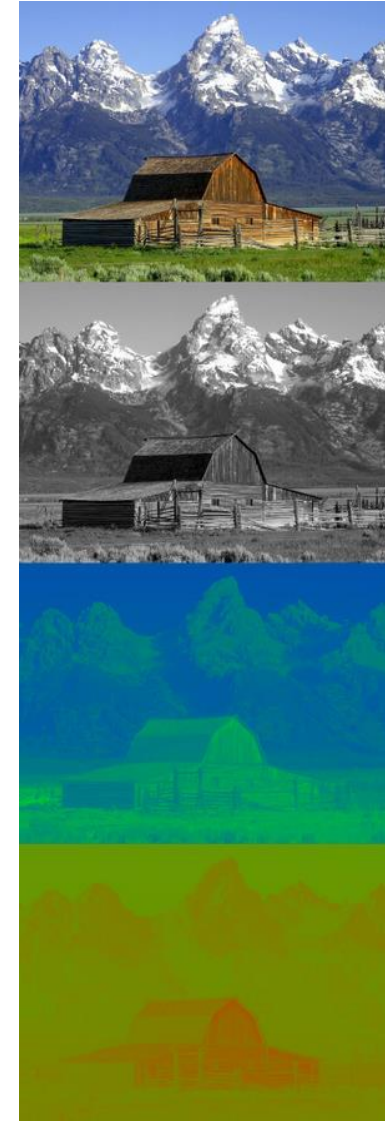


MODELOS DE COLOR SENSORIALES

Existen 4 modelos comúnmente utilizados

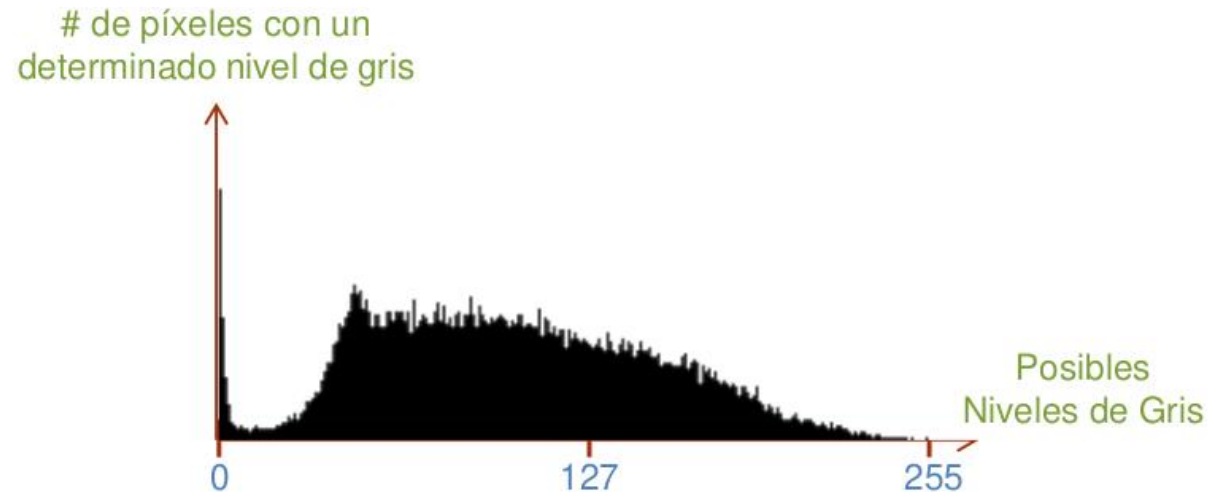
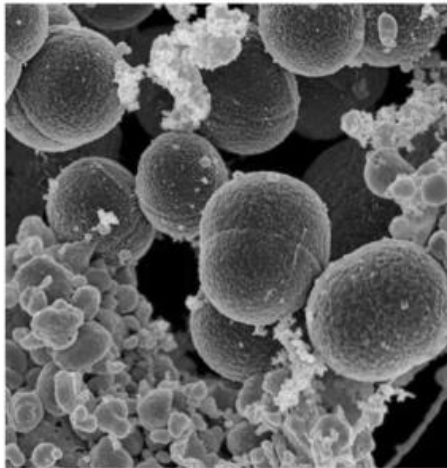
4. YUV: permite utilizar un ancho de banda reducido para los componentes de crominancia, de esta forma, hace que los errores de transmisión o las imperfecciones de compresión se oculten más eficientemente a la percepción humana que usando una representación RGB "directa".

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0,299 & 0,587 & 0,114 \\ -0,147 & 0,289 & 0,436 \\ 0,615 & -0,515 & -0,100 \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$



EL HISTOGRAMA DE UNA IMAGEN

El histograma de una imagen presenta la frecuencia de ocurrencia de los niveles de gris en la imagen, es decir, determina la distribución de frecuencias de los niveles de gris en la imagen.

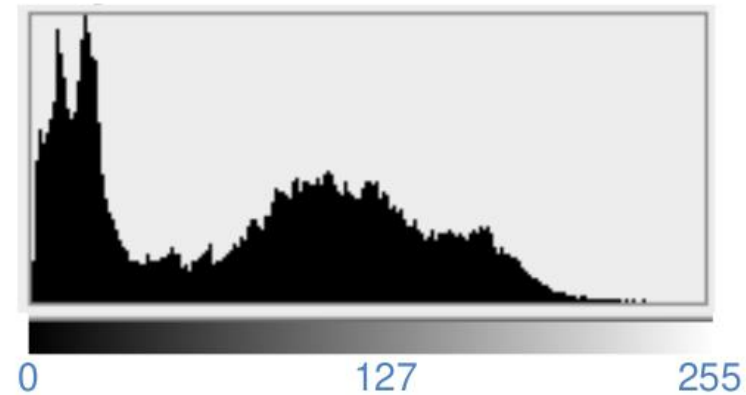


Son un elemento importante en la etapa de Pre-procesamiento pues este nos ayuda a comprender y a determinar qué transformaciones usar para mejorar la calidad de una imagen.

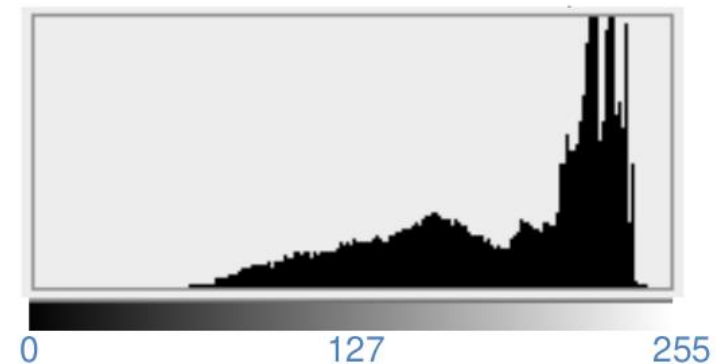


EL HISTOGRAMA DE UNA IMAGEN

Una imagen oscura con falta de luz:

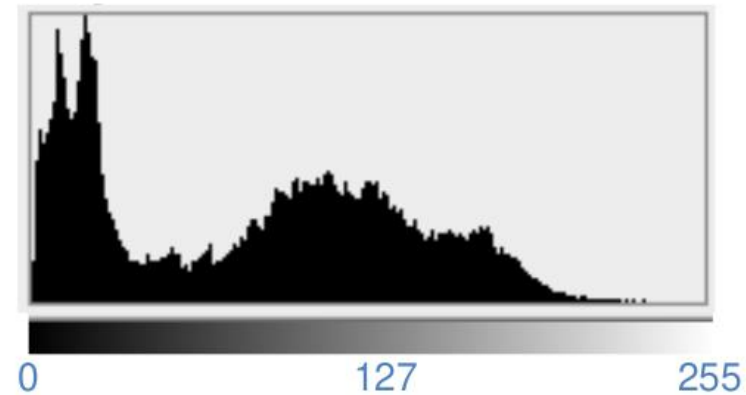


Una imagen muy clara con exceso de brillo:

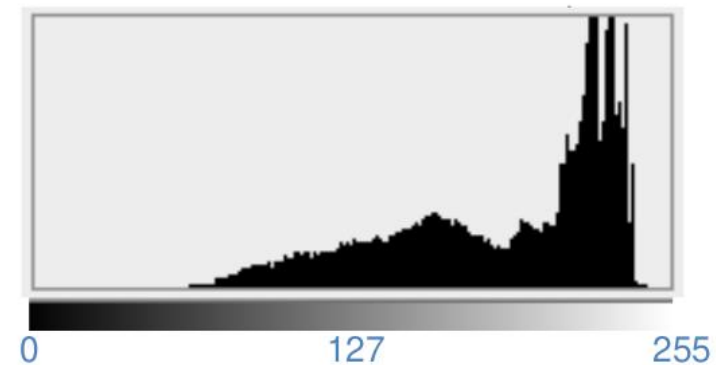


EL HISTOGRAMA DE UNA IMAGEN

Una imagen oscura con falta de luz:



Una imagen muy clara con exceso de brillo:



TRANSFORMACIONES LINEALES

La forma general de una transformación lineal es la siguiente:

$$G(x,y) = a * I(x,y) + b$$

Con base en esta ecuación tenemos que:

- Si $a = 1$ y $b = 0$ entonces $g(x, y) = f(x, y)$ (Identidad)
- Si $a = 1$ y $b > 0$, el nivel de gris se aumenta en b unidades (Suma)
- Si $a = 1$ y $b < 0$, el nivel de gris se disminuye en b unidades (Resta)
- Si $a > 1$, se produce un incremento del contraste (Multiplicación)
- Si $0 < a < 1$, se reduce el contraste (División)



TRANSFORMACIONES LINEALES

La Suma aumenta el brillo de las imágenes

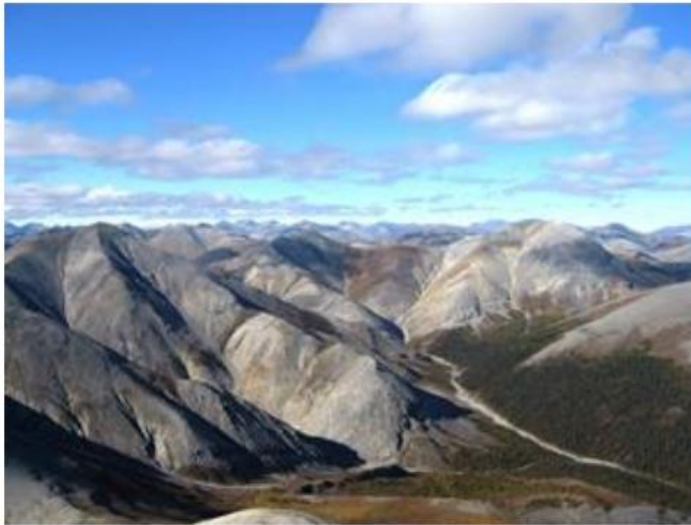


+ 80 =

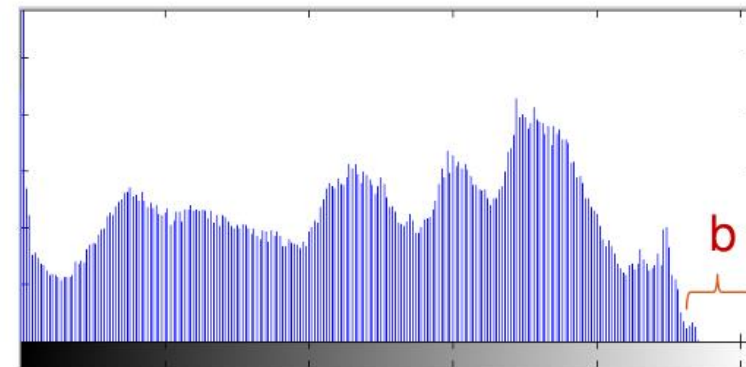
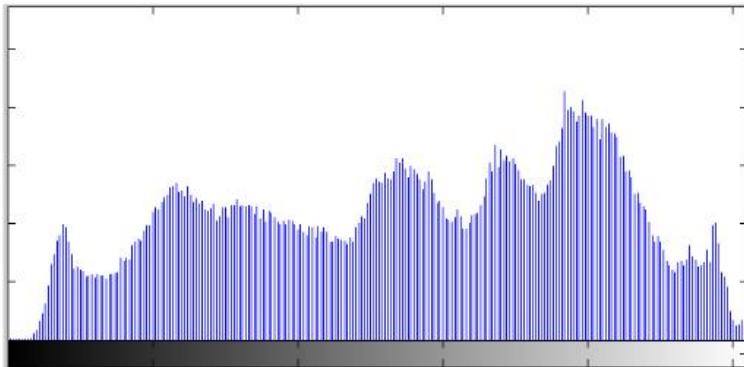


TRANSFORMACIONES LINEALES

La Resta disminuye el brillo de las imágenes



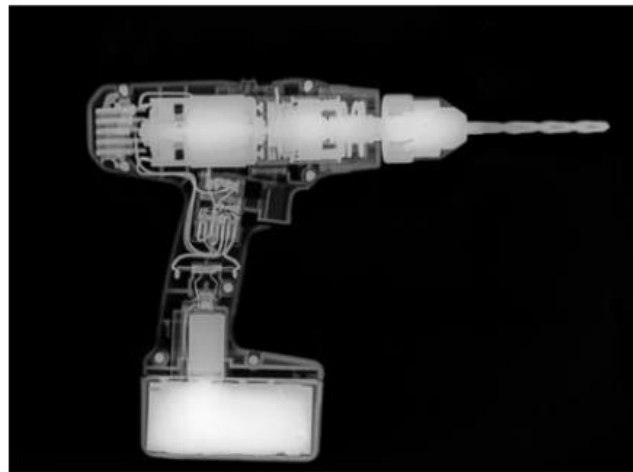
+ (-20) =



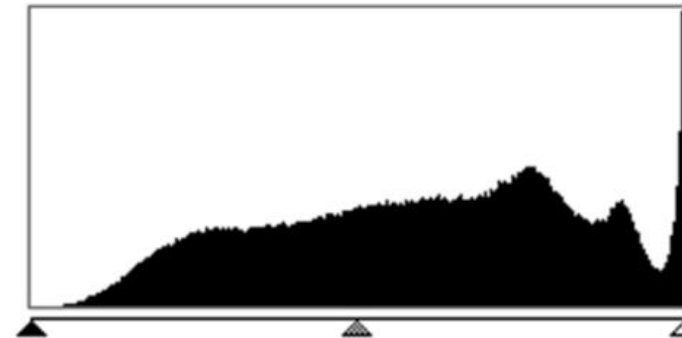
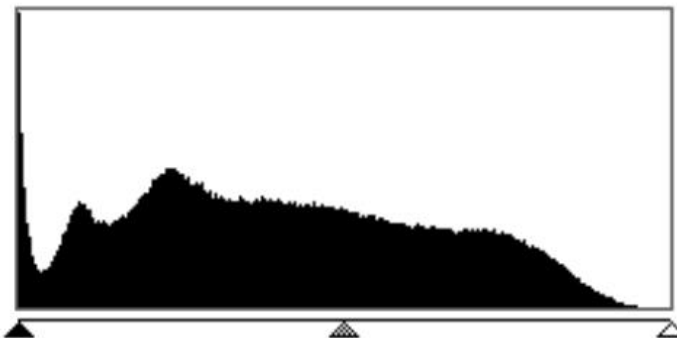
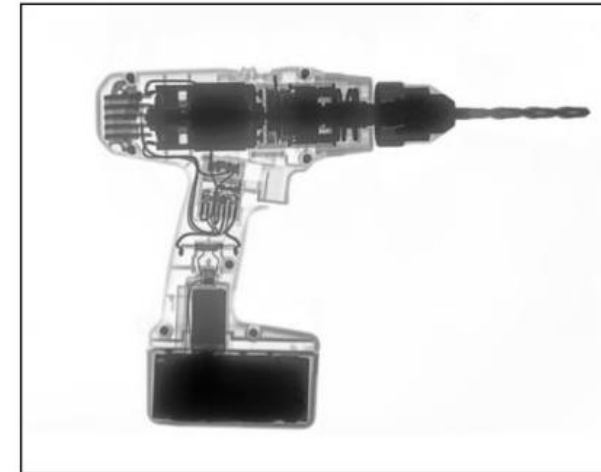
TRANSFORMACIONES LINEALES

El Negativo de una imagen es considerado un caso particular de la resta en el que se resta a toda la imagen el mayor valor de intensidad posible de la imagen.

255 -



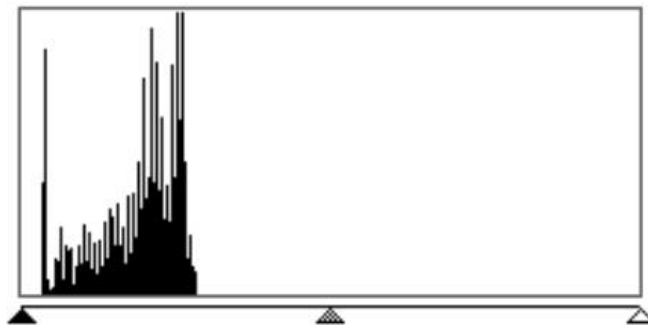
=



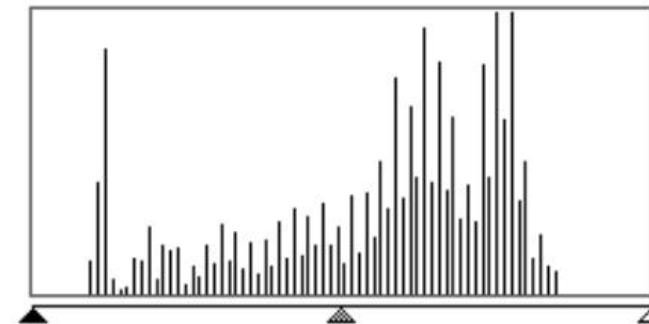
TRANSFORMACIONES LINEALES

La Multiplicación es una operación que permite “estirar” el histograma de una imagen.

$$G(x,y) = a * I(x,y)$$



$$* 3 =$$



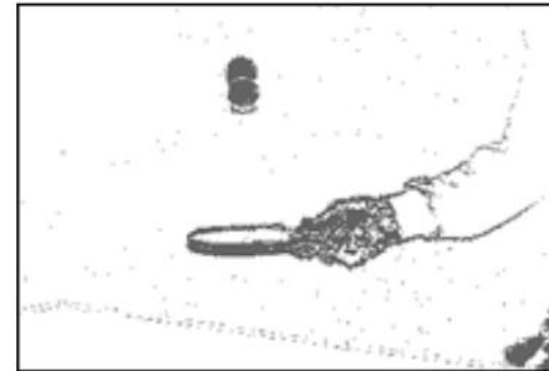
TRANSFORMACIONES LINEALES

La Suma de Imágenes es usada, por ejemplo, para resaltar los bordes de los objetos:

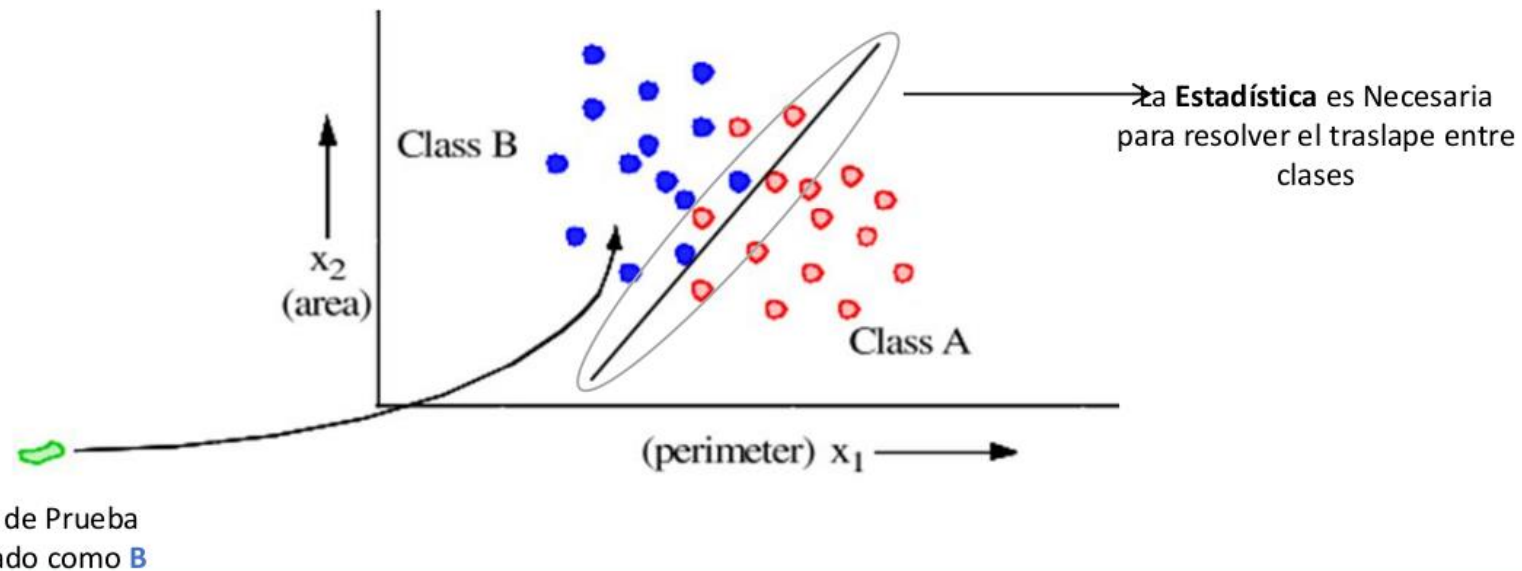
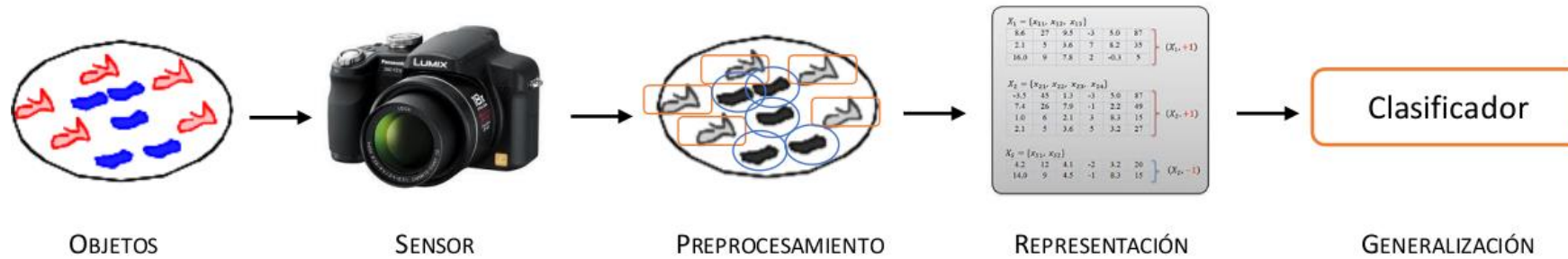


TRANSFORMACIONES LINEALES

Una aplicación de la resta es en la detección de movimiento entre dos imágenes, útil en la compresión de video y en el seguimiento de objetos.



SELECCIÓN DE CARACTERÍSTICAS



SELECCIÓN DE CARACTERÍSTICAS

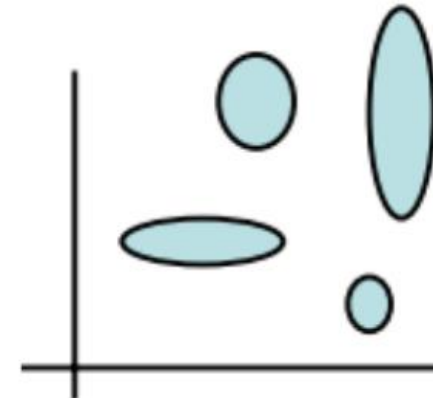
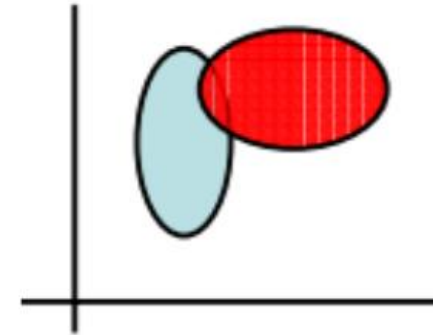
BUENAS REPRESENTACIONES:

1. Específicas a las Clases:

- Diferentes clases deberían estar representados en diferentes posiciones en el espacio de representación.

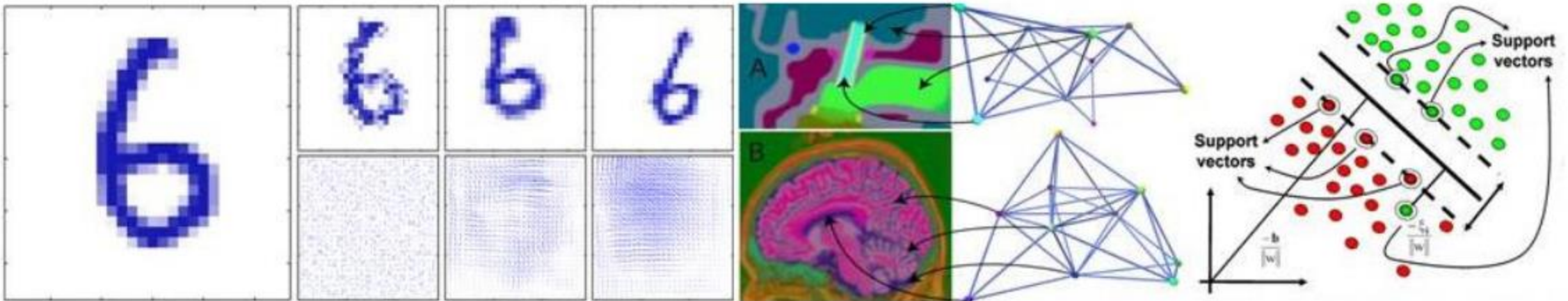
2. Deben ser Compactas:

- Cada clase debe estar representada en un pequeño conjunto de dominios finitos.



RECONOCIMIENTO DE PATRONES

El **Reconocimiento de Patrones** es la última etapa dentro de un sistema de visión artificial, en la que a partir de las características encontradas, los posibles objetos se **CLASIFICAN** en dos o más clases.



Clasificar (o reconocer) significa, en este contexto, asociar a clases (o prototipos) una serie de elementos (u objetos). Esta asociación se realiza en base a las características o propiedades de los objetos.



RECONOCIMIENTO DE PATRONES

Si los descriptores elegidos son adecuados, objetos similares tendrán patrones próximos en el espacio de características.



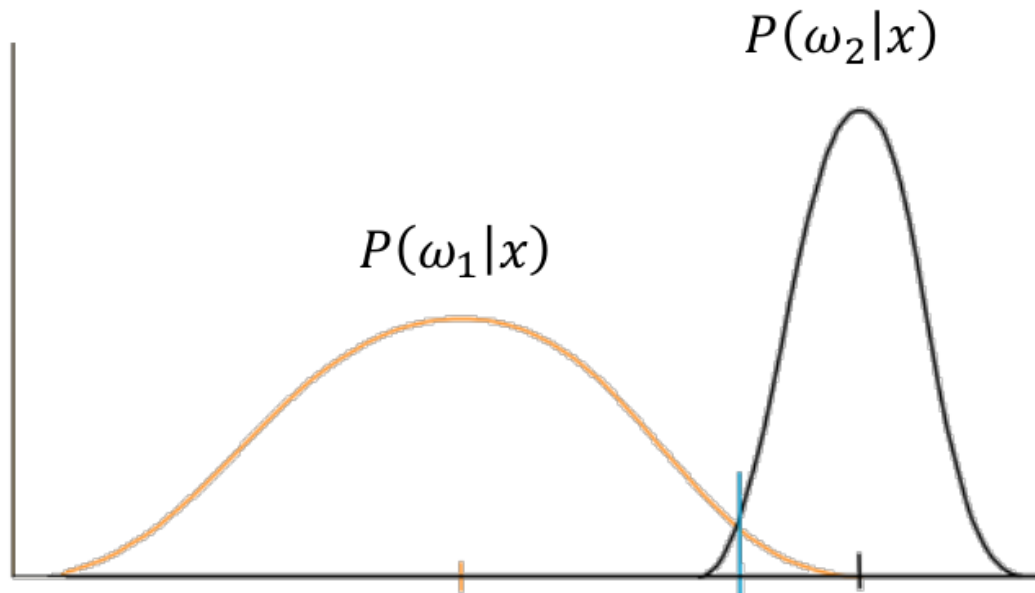
- Patrones que describen objetos de una misma clase, presentan características similares.
- Patrones que describen objetos de diferentes clases presentan características diferenciadas.



TEOREMA DE BAYES

Usar la teoría de la probabilidad para clasificar el objeto en la clase que tenga mayor probabilidad posteriori.

$$P(\omega_i|x) = \frac{p(x|\omega_i)P(\omega_i)}{p(x)}$$



$P(\omega_i)$ = Probabilidad de que en la población haya un objeto de clase ω_i

$p(x|\omega_i)$ = Probabilidad de que en la clase ω_i se de un vector de características x

$P(\omega_i|x)$ = Probabilidad de que el objeto de vector de características x pertenezca a la clase ω_i

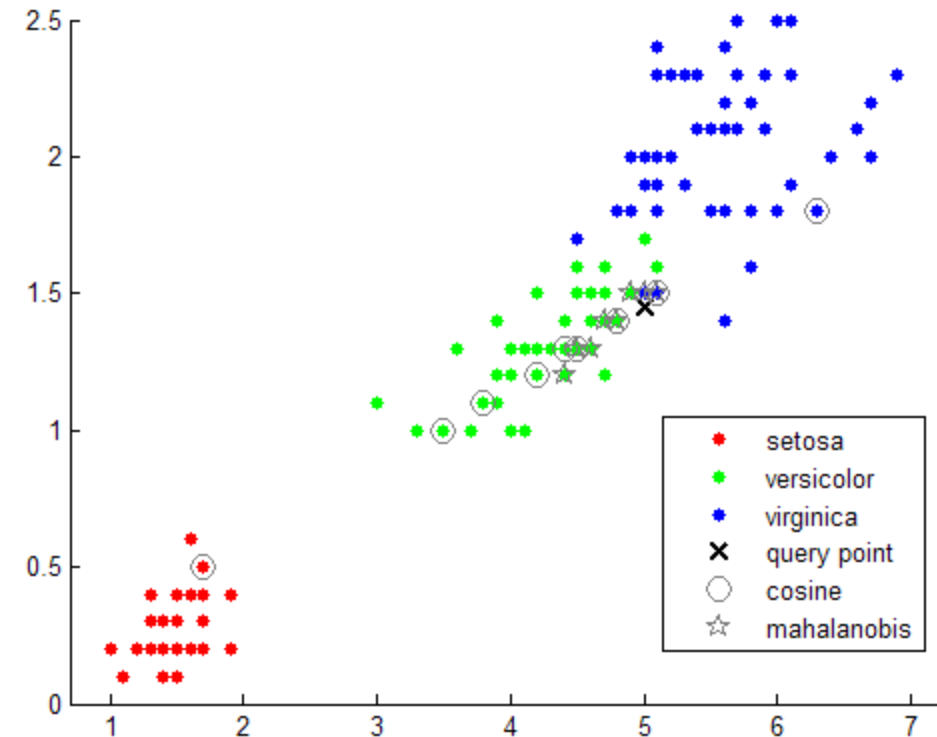
$$g(x) = \begin{cases} 1 & \text{si } P(\omega_1|x) > P(\omega_2|x) \\ 2 & \text{en otro caso} \end{cases}$$



K VECINOS MÁS CERCANOS

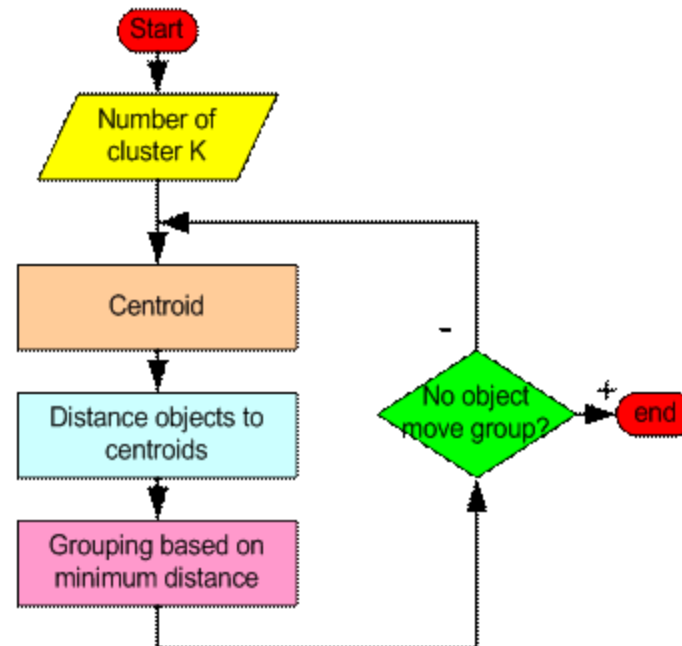
La idea básica del método considera la utilización de un conjunto de vecinos para etiquetar el nuevo objeto. Esta regla basa su operación en el supuesto de considerar a los patrones cercanos, como aquellos que tienen la mayor probabilidad de pertenecer a la misma clase.

Así el algoritmo asigna la etiqueta de clase que tengan la mayoría de los k - vecinos.



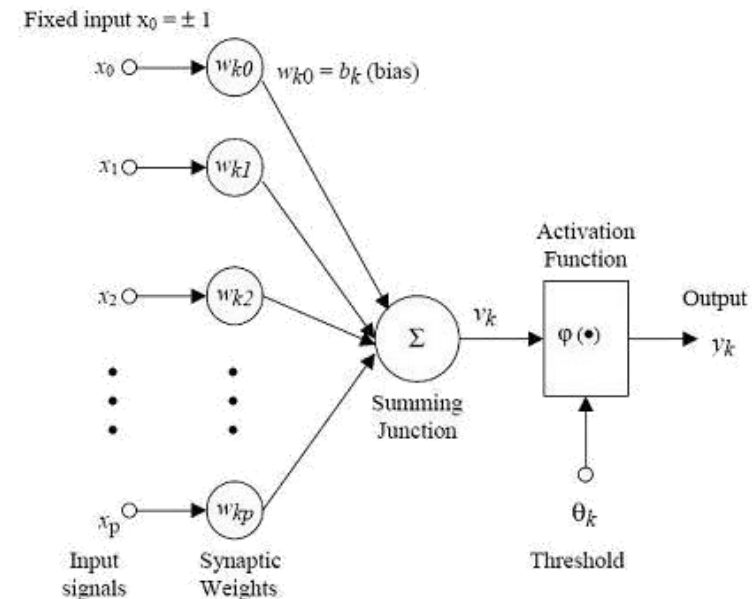
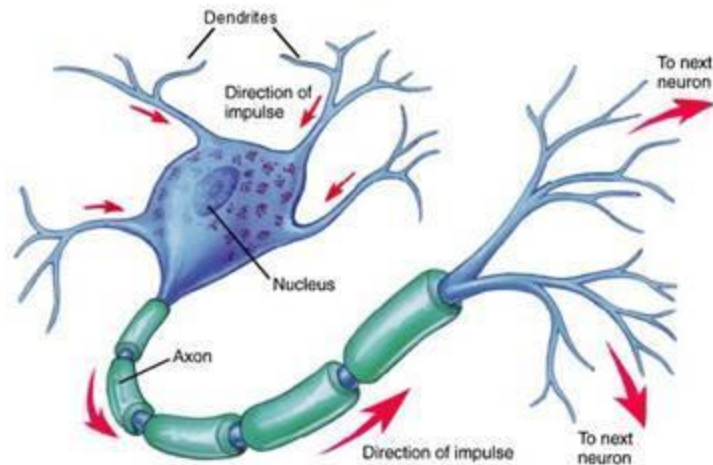
CLUSTERING

Los algoritmos de clustering o agrupamiento intentan dividir el conjunto de datos de entrenamiento en k grupos, de acuerdo con un criterio de cercanía que se define en términos de una función de distancia, como la Euclidiana, la Manhattan o la de Mahalanobis



REDES NEURONALES

- Por su capacidad de aprendizaje las neuronas de los organismos biológicos se han estudiado para su aplicación en sistemas de aprendizaje automático.
- Al igual que las neuronas biológicas están conectadas, las redes de neuronas artificiales están formadas por elementos sencillos de cómputo interconectados según diferentes modelos.



REDES NEURONALES

El Perceptrón, en su forma básica, consiste en una neurona que es capaz de aprender una función discriminante lineal v_k , que permite dividir a dos conjuntos de entrenamiento linealmente separables.

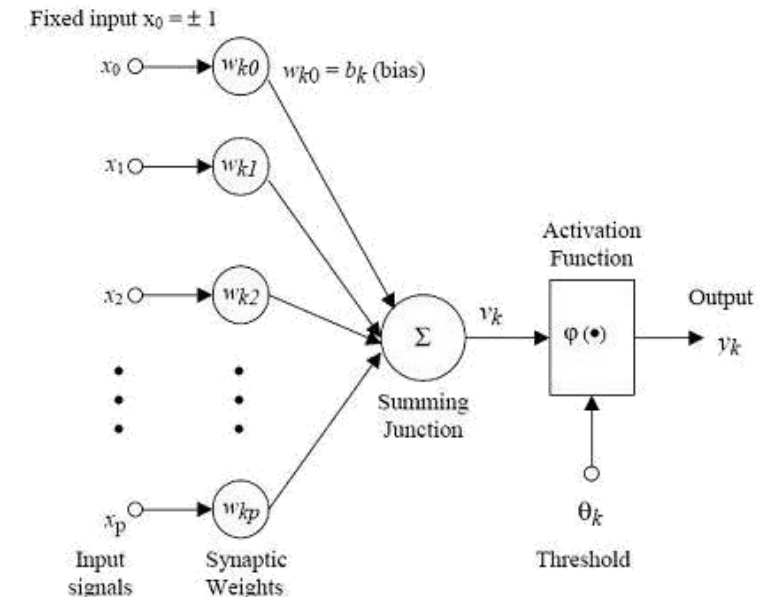
Su respuesta consiste en una suma ponderada de sus entradas que representa la ecuación de un hiperplano en el espacio p-dimensional.

$$v_k = \sum_{j=1}^p w_{kj} x_j$$

A la salida se aplica una función de activación $\Phi(v)$ (escalón, sigmoide, etc) que indica si se activa o no la neurona.

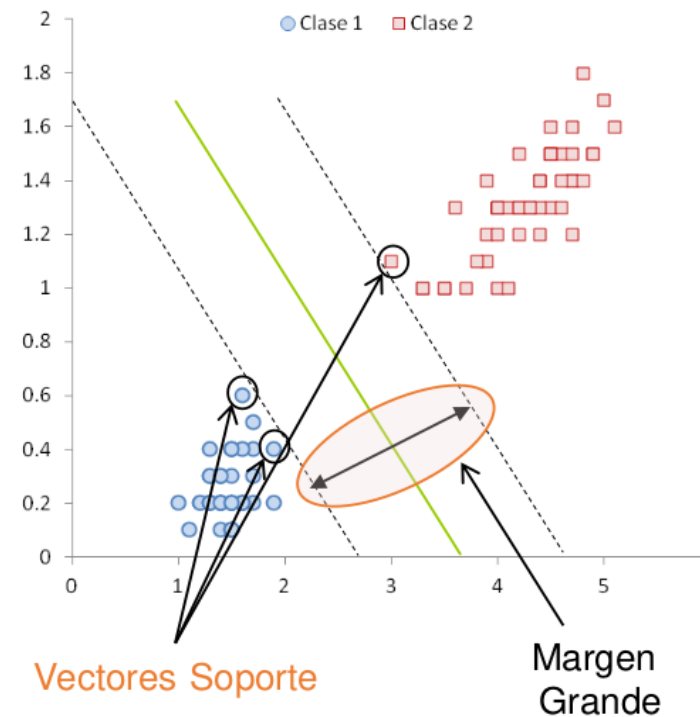
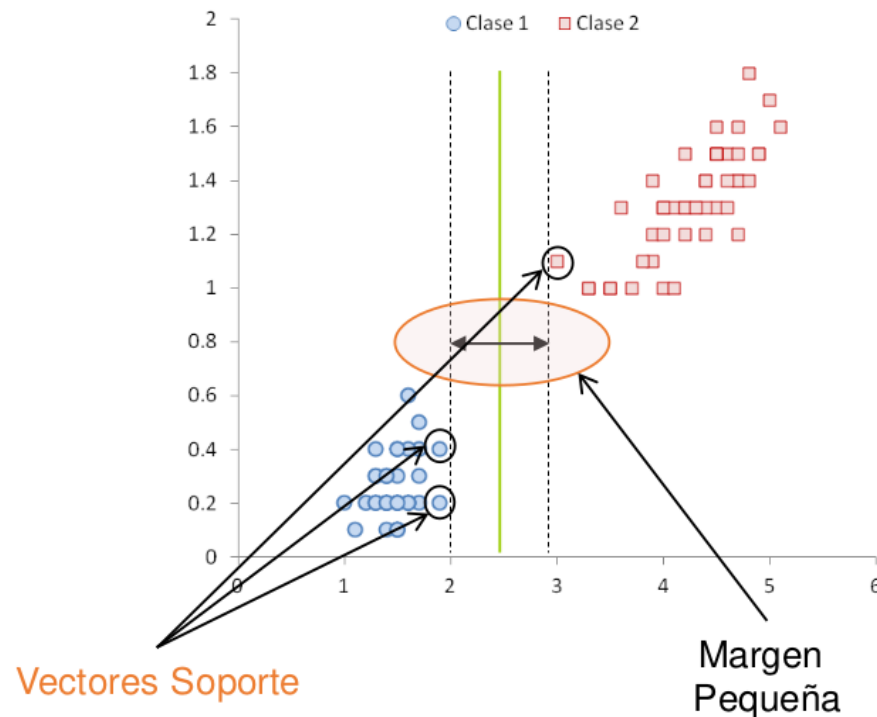
$$\varphi(v) = \begin{cases} 1 & \text{if } v \geq 0 \\ 0 & \text{if } v < 0 \end{cases}$$

$$\varphi(v) = \tanh\left(\frac{v}{2}\right) = \frac{1 - \exp(-v)}{1 + \exp(-v)}$$



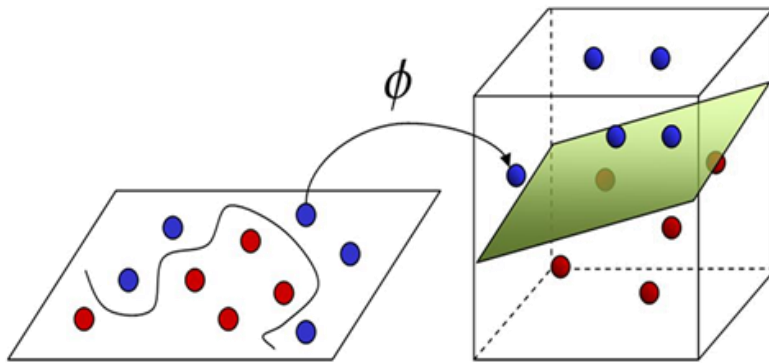
MÁQUINAS DE VECTORES SOPORTE

Las SVM son un tipo de clasificadores de patrones basados en técnicas estadísticas de aprendizaje y están a la cabeza de los métodos de clasificación por permitir construir fronteras de decisión flexibles, y su buena capacidad de generalización.



MÁQUINAS DE VECTORES SOPORTE

La **Clasificación NO Lineal** con una SVM realiza una transformación del espacio de entrada a otro de dimensión más alta, en el que los datos son separables linealmente.



Lineal: $K(x_i, x_j) = x_i \cdot x_j$

Polinómico: $K(x_i, x_j) = (x_i \cdot x_j + 1)^d$

Gausiano: $K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right)$

Al introducir un kernel, los parámetros alpha del vector w se hayan así:

$$\tilde{L}(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j \mathbf{x}_i^T \mathbf{x}_j = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j k(\mathbf{x}_i, \mathbf{x}_j)$$



BIBLIOGRAFÍA

- Digital Image Processing, Gonzalez and Woods, 2008
- Image Processing, Analysis and Machine Vision, Sonka and Hlavac, 2008
- Visión por Computador. Imágenes Digitales y Aplicaciones, Pajares y de la Cruz, 2008
- Computer Vision: Algorithms and Applications. Richard Szeliski. 2010.
- Feature extraction & image processing. Nixon, M. & Aguado, A. (2008).
- Image Processing: The Fundamentals – Second Edition, Petrou, M. & Petrou, C. (2010).
- Programming Computer Vision with Python. Jan Erik Solem. 2012.

