



Introducción a la Visión Computacional y Procesamiento Gráfico

I Unidad

Ms. Ing. Liz Sofía Pedro H.



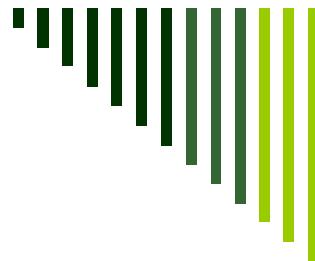
Pre-requisitos

□ Requeridos:

- Computación Gráfica II(Obligatorio).
- Algebra Lineal.
- Estructura de datos.
- Estadística (Deseable).

□ Programación: MATLAB, JAVA.

□ No incluye: Photoshop.

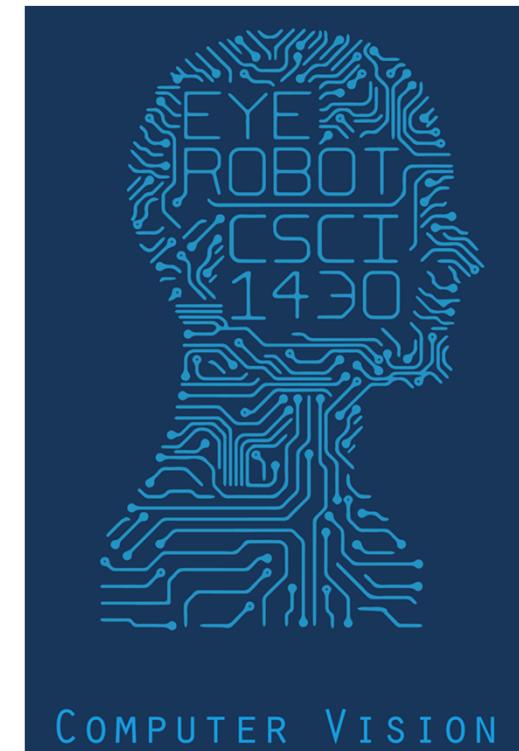
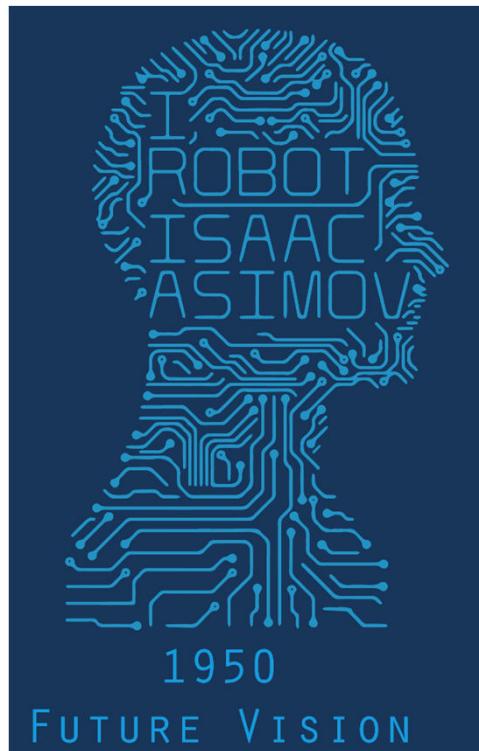


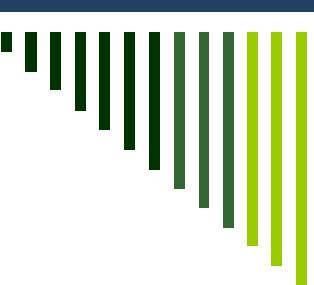
Contenidos.

1. Introducción
2. Sistemas de Visión Computacional
3. Procesamiento Digital de Imágenes.
 1. Operaciones estadísticas.
 2. Binarización.



1. INTRODUCCIÓN





¿Han usado Visión Computacional?

¿Dónde?

¿Cómo?



1.1. Historia

- El mundo físico percibido como tridimensional.
- Desde la antigüedad existe el deseo de representarlo con imágenes planas 2D, como por ejemplo: pinturas.
- Thales de Mileto (640 AC - 548 AC), predijo un eclipse y logró medir tamaño de pirámides empleando proyecciones.
- Los griegos consideraban que los ojos eran emisores de luz y no receptores.



1.1. Historia (Cont.)

- Los pintores italianos del Renacimiento fueron los primeros en entender la formación de las imágenes
- Alrededor de 1413, el arquitecto, pintor y escultor Filippo Brunelleschi inventó la perspectiva.
- Sus principales obras: La Catedral Santa María de Fiore (cúpula es la más grande del mundo con más de 50m de diámetro)



1.1. Historia (Cont.)

- A partir de esta época se empieza a considerar el *punto de fuga*, en el que líneas paralelas que se alejan del observador convergen en un punto.





1.1. Historia (Cont.)

- Albrecht Dürer (1471-1528) : la máquina de perspectiva, como apoyo a los pintores.

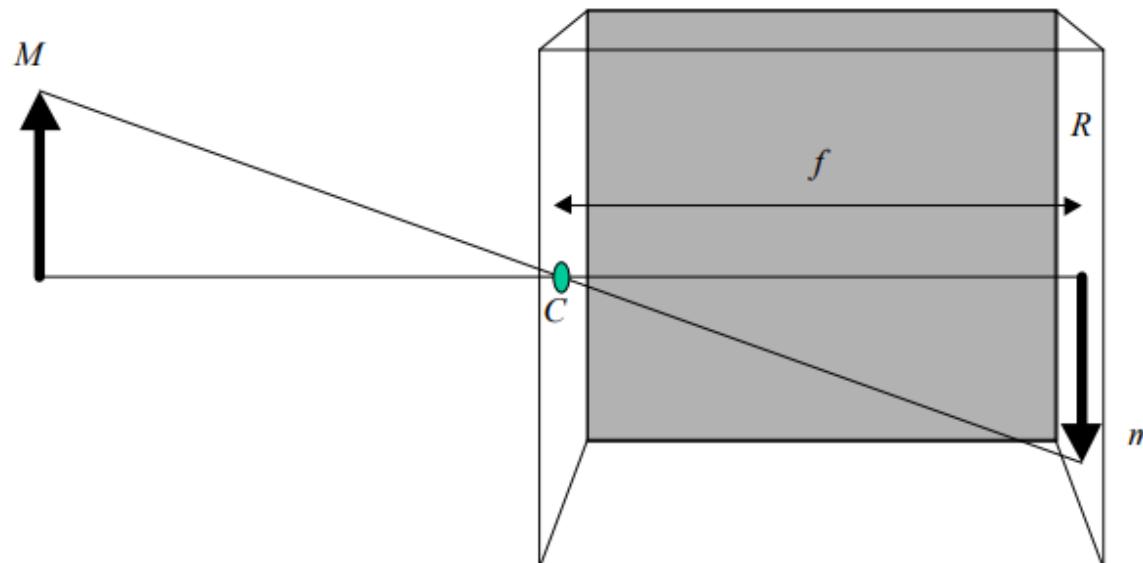


Fig. 34. Albrecht Dürer, "A Draftsman Making a Perspective Drawing of a Woman." Woodcut. From *Underweysung der Messung*. All rights reserved, The Metropolitan Museum of Art, Gift of Felix M. Warburg, 1918, (18.58.3 [recto]).



1.1. Historia (Cont.)

- En 1545 el astrónomo Germina Frisius: Cámara oscura.





1.1. Historia (Cont.)

- En 1826 el químico francés Niepce, obtuvó la primera fotografía, colocando una superficie fotosensible dentro de una *cámara oscura* para fijar la imagen.
- En 1838 el químico francés Daguerre (1787-1851) hizo el primer proceso fotográfico práctico empleando una placa fotográfica que era revelada con vapor de mercurio y fijada con trisulfato de sodio.





1.1. Historia (Cont.)

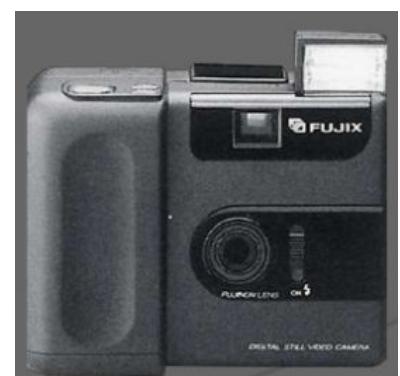
- En 1886, se inventó la película fotográfica (final de la fotografía primitiva).
- En 1888, salió a la venta la primera cámara que automatizaba el proceso, KODAK.





1.1. Historia (Cont.)

- Entre 1960 y 1970, las primeras patentes de la fotografía digital (Willard Boyle y George Smith).
- Steve Sasson y la primera cámara digital KODAK.
- 1988, La primera cámara que registraba imágenes en una computadora.
- Años después, devendrá el apogeo de la cámara analógica.



TAREA!



1.1. Historia (Cont.)

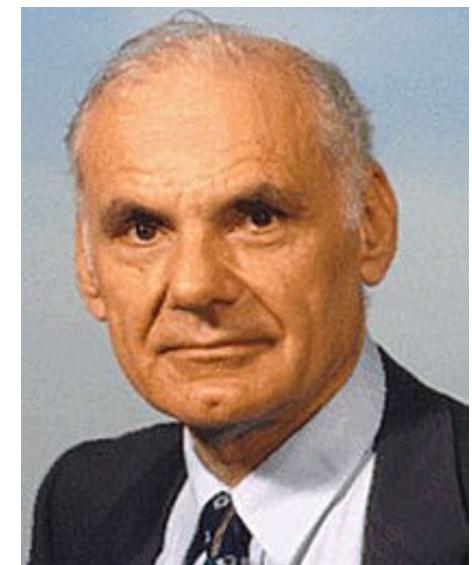
- En 1957, Perceptron
- En 1959, Russell Kirsch y sus colegas desarrollaron un aparato que permitía transformar imágenes en cuadrículas de números.
- La fundación del Laboratorio de Inteligencia Artificial en MIT en 1959.





1.1. Historia (Cont.)

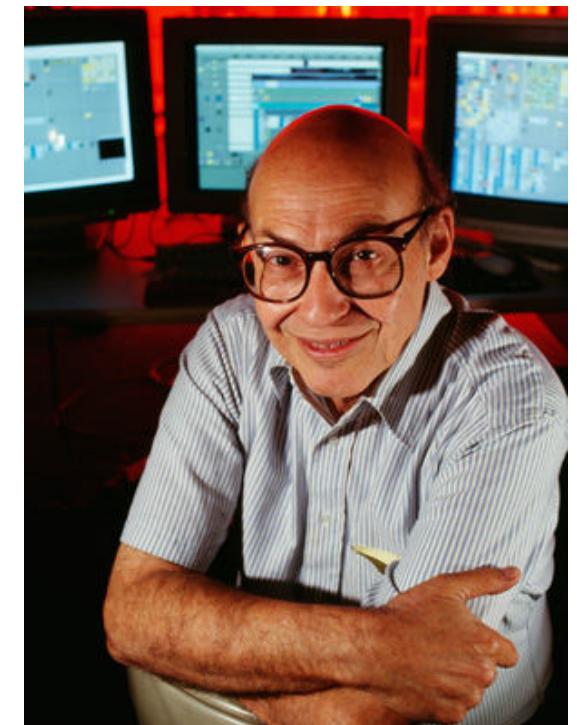
- La tesis de Lawrence Roberts , “*Machine perception of three-dimensional solids* (1963) el cual partiendo de imágenes 2D permite obtener modelos 3D de dichos objetos, es considerada uno de los precursores de la moderna visión por computadora.





1.1. Historia (Cont.)

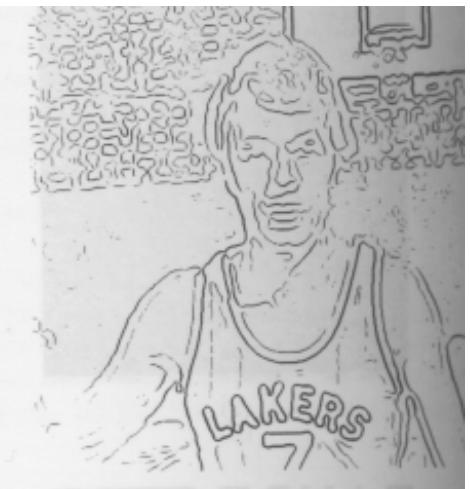
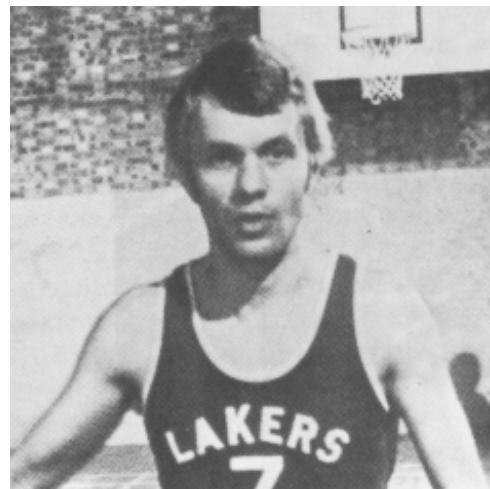
- 1966: Marvin Minsky le asignó una tarea específica a un estudiante universitario.
“Pasa el verano vinculando una cámara a una computadora. Luego, pida a la computadora que describa lo que ve”.
- 1966: Marvin Minsky le asignó una tarea específica a un estudiante universitario.





1.2. Enfoques de visión artificial

- Los investigadores se dieron cuenta de que era necesario abordar imágenes del mundo real para lo cual, era necesario investigar las tareas de visión "de bajo nivel", como la detección de bordes y segmentación.





1.2. Enfoques de visión artificial (Cont.)

- Para abordar el complejo problema de la visión humana se consideraron tres enfoques.
- ***Un primer enfoque*** (Azriel Rosenfeld), el desarrollo técnicas empíricas basadas en criterios de tipo matemático que aproximaban estos problemas de estimación de bordes y líneas empleando distintos criterios.
 - El uso simultáneo de operadores de distinto tamaño
 - El gran inconveniente :no poder proponer métodos para la evaluación de los distintos algoritmos.
 - Ha seguido su desarrollo hasta nuestros días .



1.2. Enfoques de visión artificial (Cont.)

- ***El segundo enfoque***, profundizar en el problema reduciendo su alcance a un mundo de juguete.
 - Un mundo de bloques blancos mate iluminados sobre fondo negro.
 - Los bloques podían tener cualquier forma, siempre que todas sus superficies fueran planas y todos sus bordes rectos.
 - La simplificación impuesta en el modelo se había hecho con la esperanza de que permitiera extrapolar dichos conocimientos a mundos más complejos, pero no ocurrió y hubo que estudiar un nuevo enfoque.



1.2. Enfoques de visión artificial (Cont.)

- ***El tercer enfoque***, Horn (1975, 1977) establece modelos de cálculo que expresan la formación de la imagen a través de ecuaciones diferenciales que relacionan los valores de intensidad de la imagen con la geometría de la superficie.
 - La iluminación, la geometría, la reflectancia de la superficie y el punto de vista del observador actúan de forma conjunta para crear los valores de intensidad medidos en la imagen.



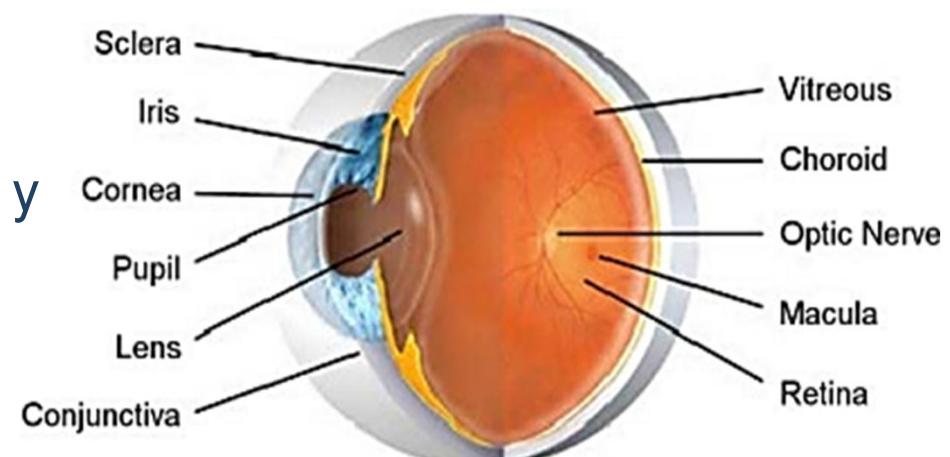
1.3. Visión Humana.

- Es la capacidad de interpretar el entorno circundante utilizando la luz en el espectro visible reflejado por los objetos en el entorno.
- Es diferente de la agudeza visual, que se refiere a qué tan claramente ve una persona.
- Se conoce como percepción visual, vista, vista o visión.



1.3.1. Ojo Humano.

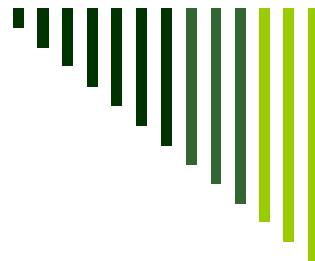
- El ojo humano es un órgano que reacciona a la luz y la presión proporcionando una imagen 3D en movimiento.
- Los conos y bastones en la retina permiten la percepción de la luz, incluida la diferenciación del color y la percepción de la profundidad.
- Puede diferenciar entre aproximadamente 10 millones de colores.





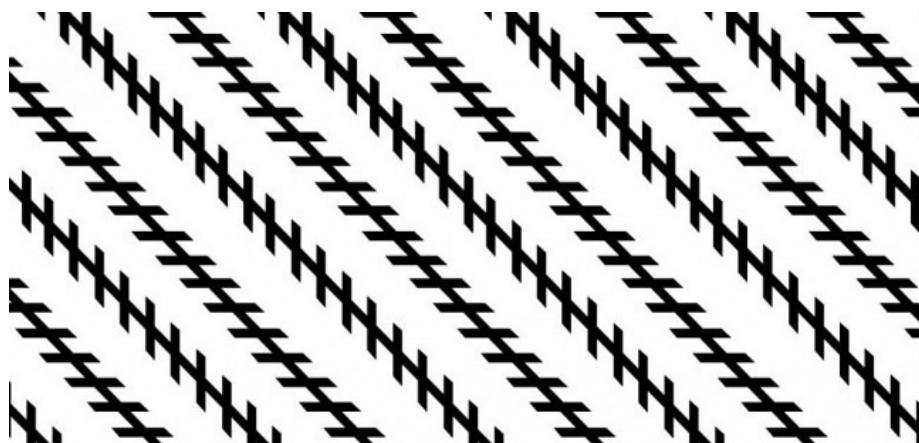
1.3.2. Problemas de percepción visual.

- Visión humana es adecuada para ciertas tareas pero no infalible.
- La visión humana está sujeta a ilusiones, ambigüedades e inconsistencias.



1.3.2. Problemas de percepción visual (Cont.)

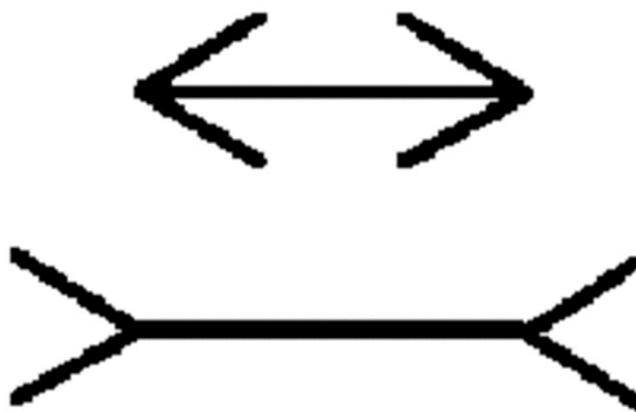
- ***Las ilusiones visuales*** consisten en ver lo que no es, parece ser algo que no es real.



Ilusión de Zöllner (1860)



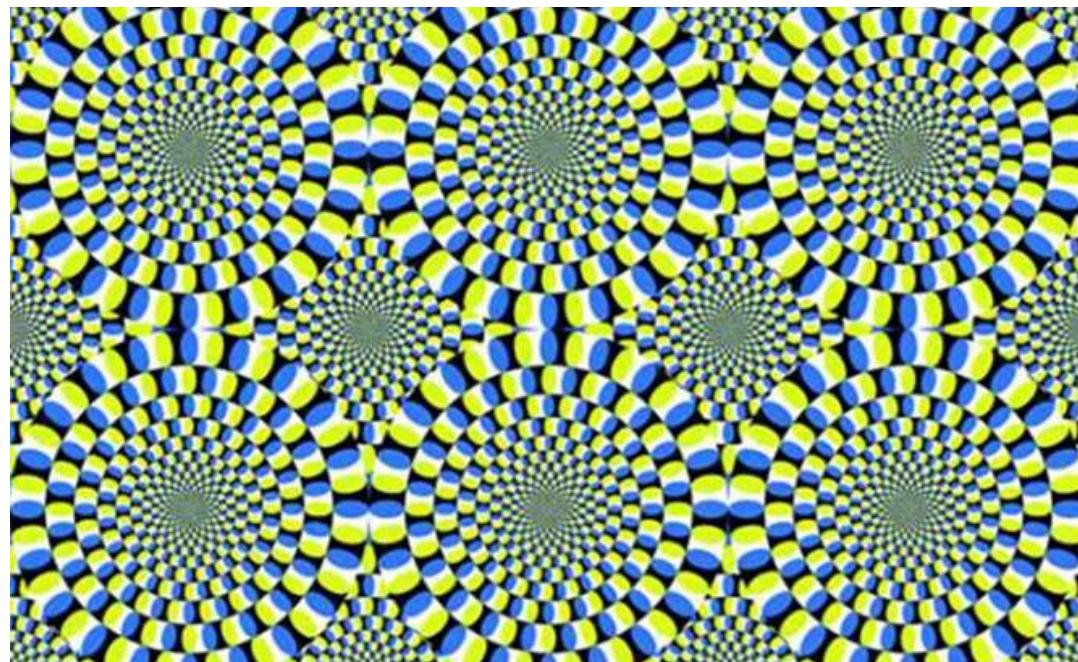
1.3.2. Problemas de percepción visual (Cont.)



Ilusión de Müller–Lyer (1889)



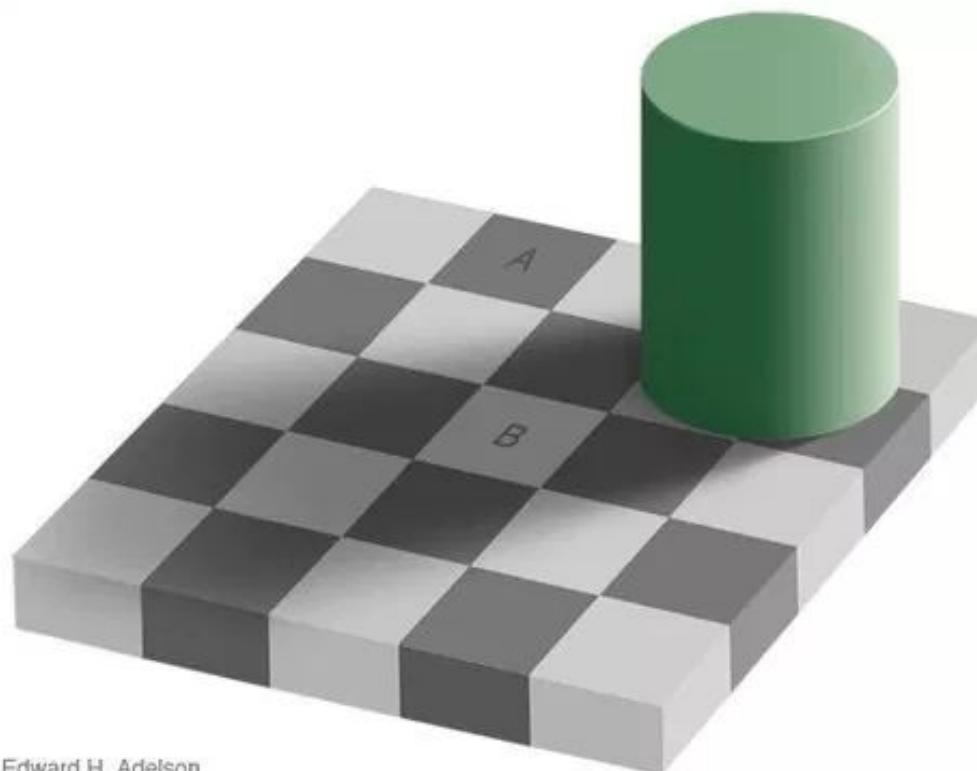
1.3.2. Problemas de percepción visual (Cont.)



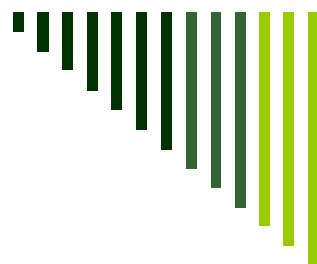
Akiyoshi Kitaoka



1.3.2. Problemas de percepción visual (Cont.)



Edward H. Adelson



1.3.2. Problemas de percepción visual (Cont.)

- ***La ambigüedad*** consiste en que una figura puede tener más de una posible interpretación.
- Es decir, múltiples interpretaciones pueden coexistir o una puede dominar a la(s) otra(s).



Mujer joven / Anciana



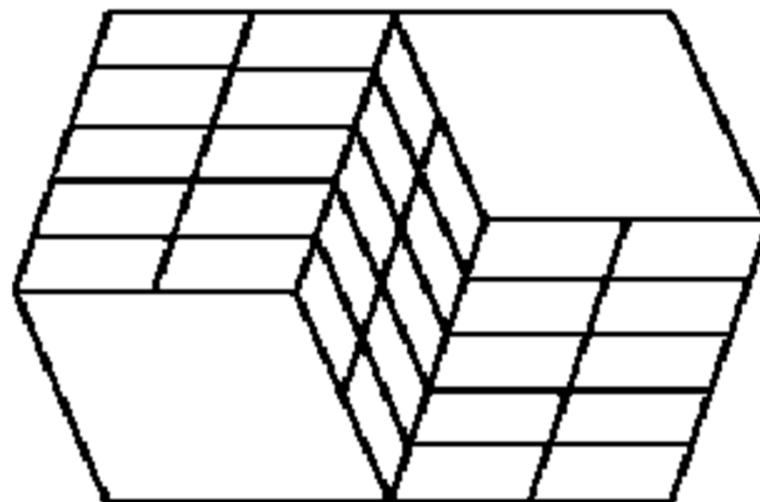
1.3.2. Problemas de percepción visual (Cont.)



Jarrón / Caras



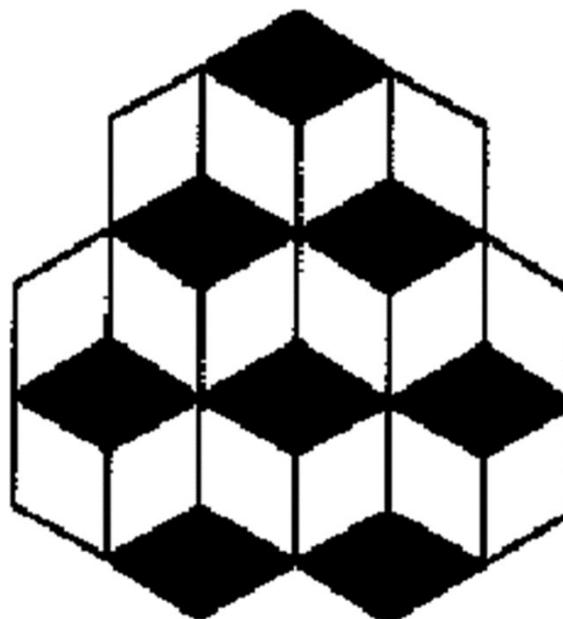
1.3.2. Problemas de percepción visual (Cont.)



Hacia dentro / Hacia fuera



1.3.2. Problemas de percepción visual (Cont.)



Seis cubos / Siete cubos



1.3.2. Problemas de percepción visual (Cont.)

- ***Las inconsistencias*** consisten en figuras que aunque una figura no sea una ilusión, ni ambigua; pero esta es completamente irrealizable en el sentido de que no podemos construir el objeto 3D percibido en el espacio real 3D.



1.3.2. Problemas de percepción visual (Cont.)

- ¿Inferimos de las imágenes en la retina más de lo que de verdad está soportado por la geometría y física de la formación de imágenes?





1.3.2. Problemas de percepción visual (Cont.)

- Cada imagen es una imagen de algo sólo para aquel que conoce como leerla, y que está capacitado con la ayuda de la imagen a formar una idea de los objetos.
- Como resultado los humanos podríamos ver lo que no hay y no ver lo que hay.
- ***¿Le permitiríamos a una máquina hacer lo mismo?***



1.4. Visión Computacional.

- La función principal de la visión es reconocer y localizar objetos en el ambiente mediante procesamiento de imágenes.
 - *Visión es saber que hay y dónde mediante la vista* (Aristóteles).
 - *Visión es recuperar de la información de los sentidos (vista) propiedades válidas del mundo exterior* (Gibson).
 - *Visión es un proceso que produce a partir de las imágenes del mundo exterior una descripción que es útil para el observador y que no tiene información irrelevante* (Marr).



1.4. Visión Computacional (Cont.)

- La visión de computadora está construyendo algoritmos que pueden entender el contenido de imágenes y usarla para otras aplicaciones.
- El objetivo de la visión artificial es desarrollar programas que sean capaces de interpretar imágenes y vídeo, obteniendo información sobre ellas.
- Estudio de procesos que permiten reconocer y localizar objetos en el ambiente mediante procesamiento digital de imágenes, para entenderlos y construir máquinas con capacidades similares.

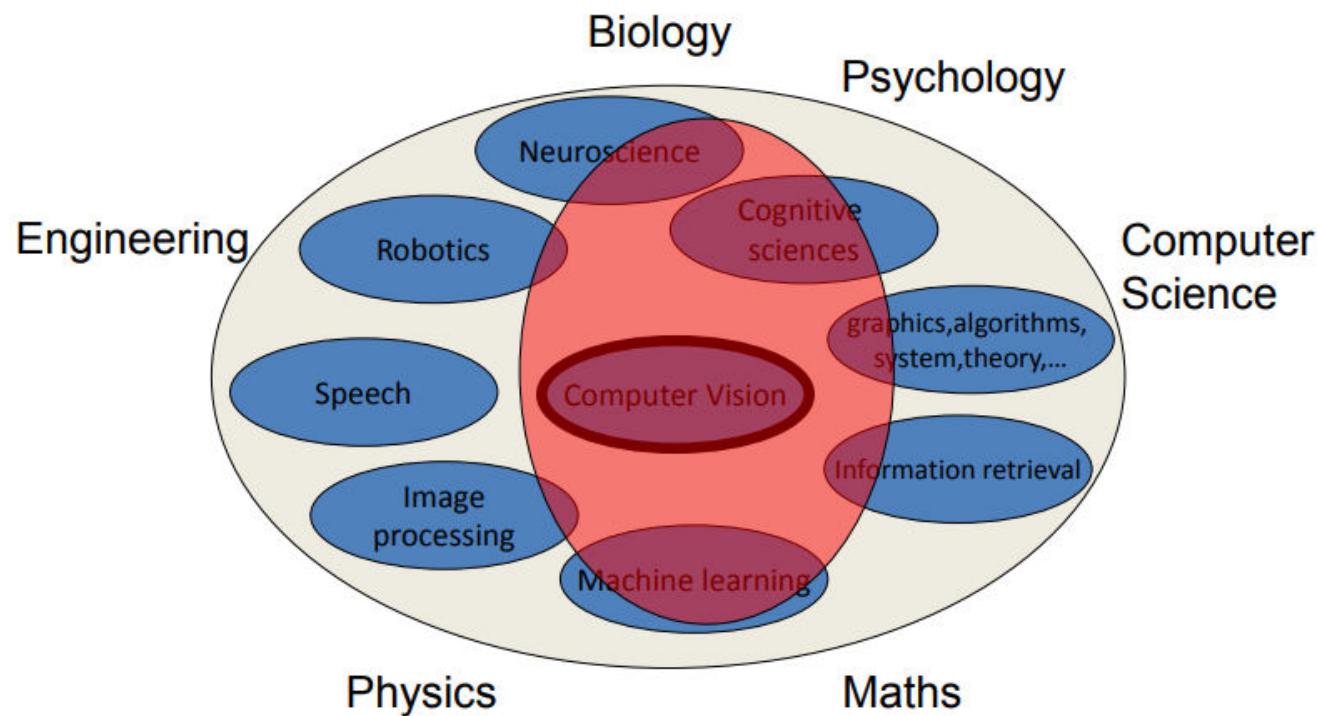


1.4. Visión Computacional (Cont.)

- Ventajas de la visión computacional en relación a la visión humana:
 - Mejor midiendo magnitudes físicas.
 - Mejor para la realización de tareas rutinarias.
 - Mejor en tareas de bajo nivel de proceso.
- Mientras que la visión humana normalmente posee una mejor adaptación a situaciones imprevistas, empleando conocimiento previo.



1.4. Visión Computacional (Cont.)





2. SISTEMAS DE VISIÓN COMPUTACIONAL



2.1. Sistemas de visión computacional

- Los dos pilares del sistema de visión artificial son:
 - Sistema de formación de las imágenes
 - Sistema de procesamiento de imágenes



2.1. Sistemas de visión computacional (Cont.)

- *Sistema de formación de las imágenes:*
 - *Subsistema de iluminación.* Conjunto de artefactos que producen radiación electromagnética que incidirá sobre los objetos a visualizar. Por ejemplo, lámparas, pantallas fotográficas, filtros de luz, láseres, etc.
 - *Subsistema de captación.* Son los transductores que convierten la radiación reflejada luminosa en señales eléctricas. Por ejemplo las cámaras CCD, no sólo en el espectro visible, sino que van desde la radiación gamma hasta la radiofrecuencia o microondas, dando paso a sensores de ultrasonidos, sonar, radar, etc.



2.1. Sistemas de visión computacional (Cont.)

- *Subsistema de adquisición.* La señal eléctrica procedente de las cámaras forman la señal de vídeo. Hay una tendencia creciente a que su naturaleza sea de tipo digital, pero todavía existen muchas señales de vídeo de carácter analógico.
Para ser tratadas hay que muestrearlas y cuantificarlas. Ambas tareas son realizadas por las tarjetas de adquisición (*frame grabbers*), y se depositan en el bus de expansión del computador.



2.1. Sistemas de visión computacional (Cont.)

- *Sistema de procesamiento de imágenes.* La imagen será procesada para transformarla en información de alto nivel, para vincularla a múltiples periféricos receptores de información y vincularse con el sistema de procesamiento de las imágenes.
 - *Subsistema de procesamiento.* Suele ser un computador o un clúster de computadores, dependiendo de las necesidades; partiendo de una representación digital de las imágenes la cual será procesada hasta alcanzar otro tipo de información de más alto nivel.

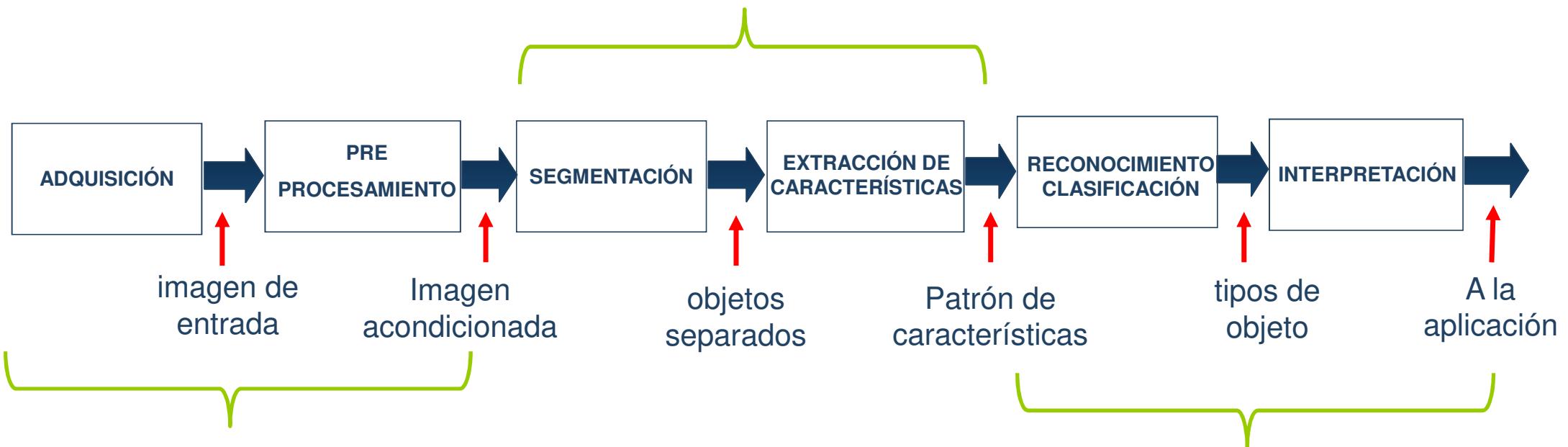


2.1. Sistemas de visión computacional (Cont.)

- *Subsistemas de periféricos.* Conjunto de elementos receptores de la información de alto nivel. Puede ser un monitor de altas prestaciones gráficas, una impresora sacando las características, etc.



2.2. Etapas





2.2. Etapas (Cont.)

1. *Captación de la imagen*: Proceso por el cual se obtiene una imagen visual.
2. *Preprocesamiento*: Conjunto de técnicas que permiten reducir el ruido y realzar los detalles.
3. *Segmentación*: Proceso que permite dividir a una imagen en objetos de interés.
4. *Extracción de características*: Proceso que obtiene las características necesarias para diferenciar un tipo de objeto de otro como el tamaño y forma.



2.2. Etapas (Cont.)

5. *Reconocimiento y clasificación:* Proceso que identifica a los objetos de una escena. Por ejemplo, diferentes tipos de piezas en un tablero de ajedrez.
6. *Interpretación:* Proceso que asocia un significado a un conjunto de objetos reconocidos.



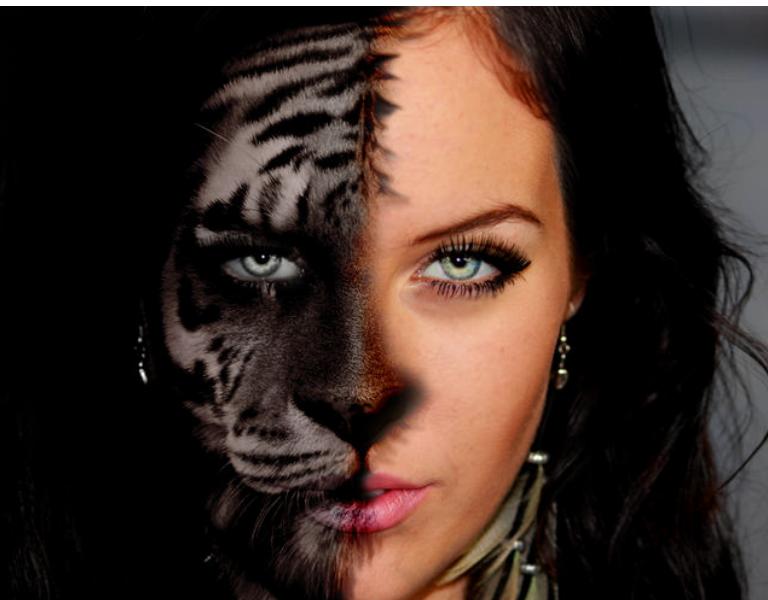
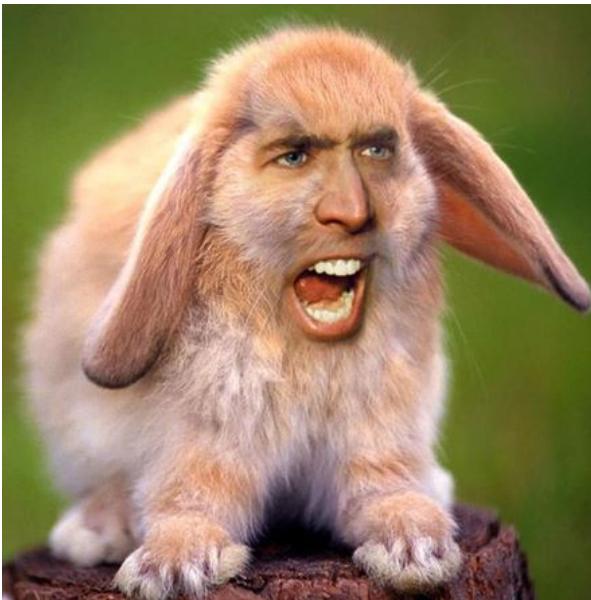
2.3. Hardware

- *Tarea.* Investigar los diversos tipos de hardware involucrados en un sistema de visión computacional.



2.4. Aplicaciones.

□ Efectos Especiales



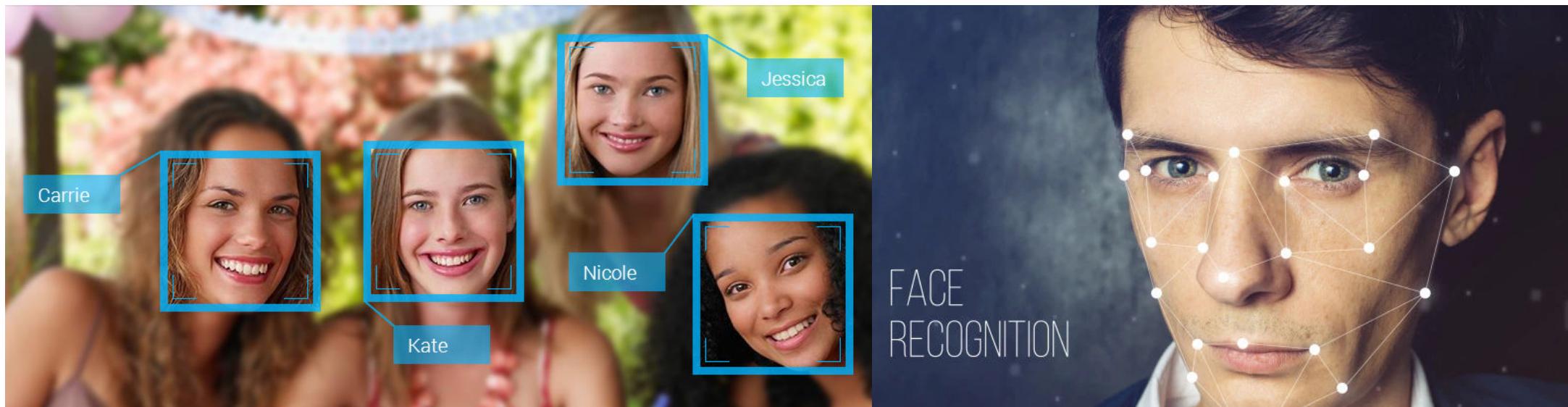
2.4. Aplicaciones (Cont.)





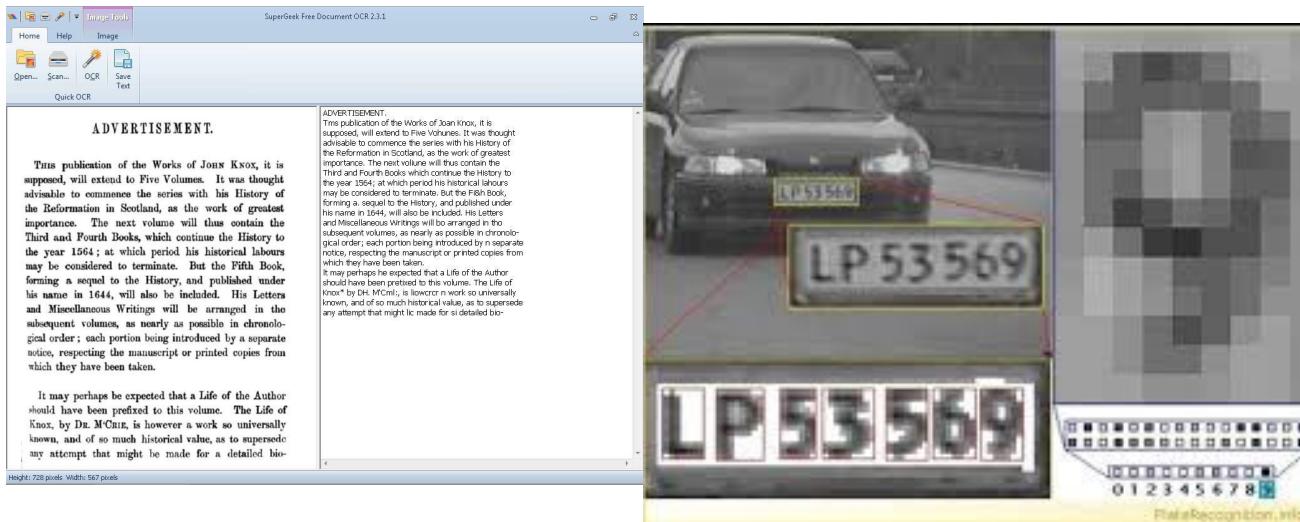
2.4. Aplicaciones (Cont.)

□ Reconocimiento de rostros



2.4. Aplicaciones (Cont.)

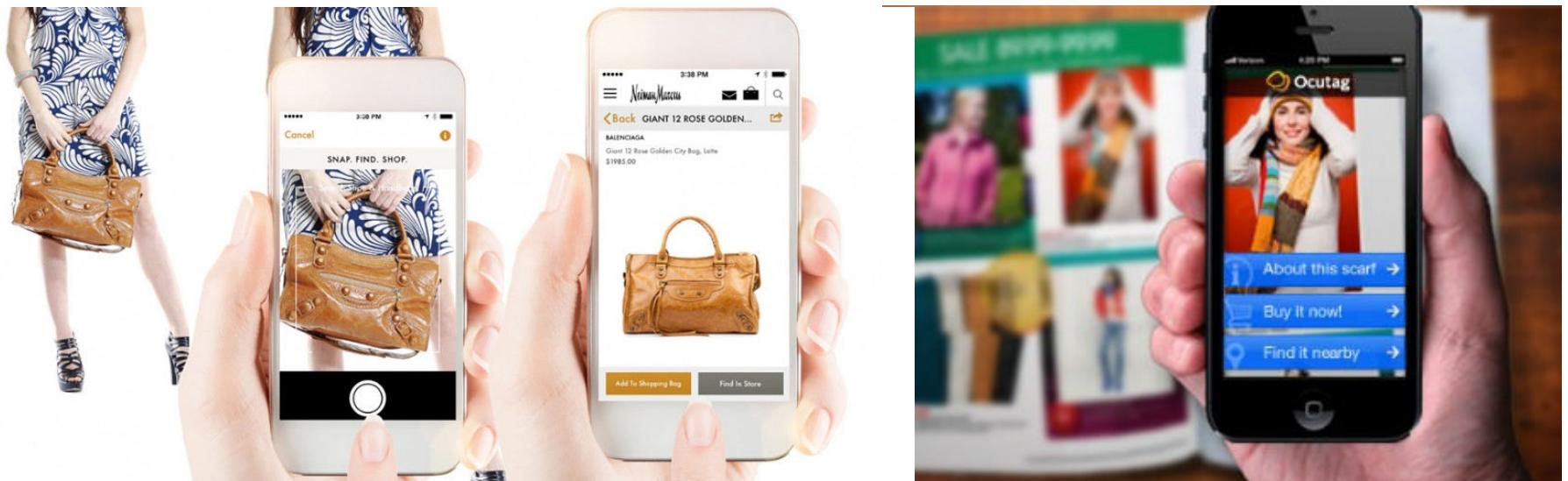
□ Reconocimiento de caracteres ópticos (OCR)





2.4. Aplicaciones (Cont.)

□ Mobile visual search





2.4. Aplicaciones (Cont.)

- Self-driving cars



2.4. Aplicaciones (Cont.)

□ Automatic checkout





2.4. Aplicaciones (Cont.)

- Augmented reality and virtual reality

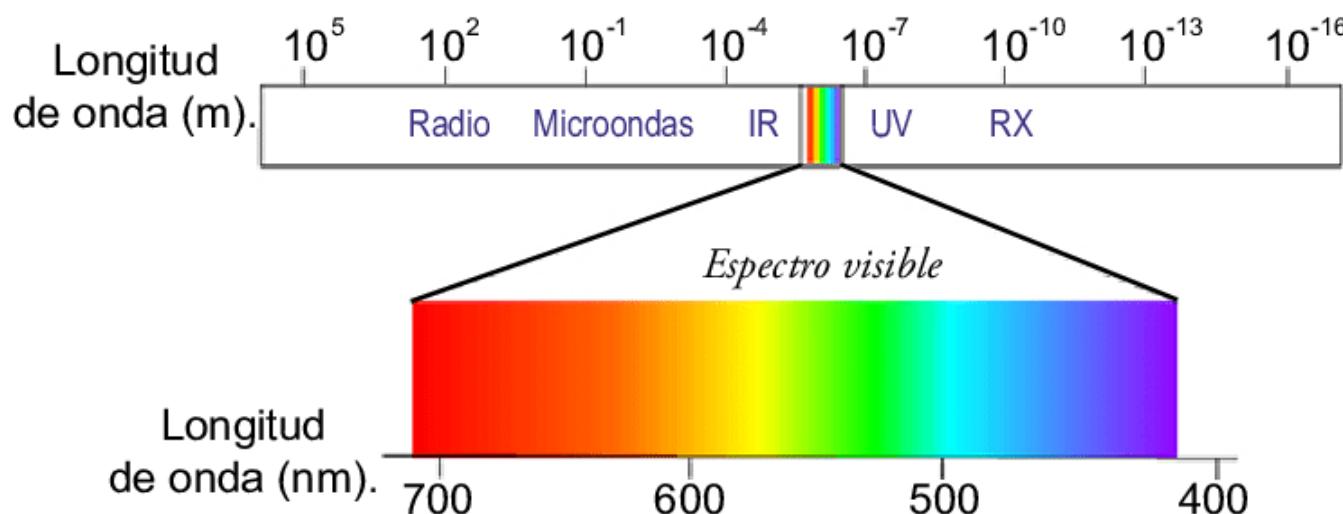




3. IMÁGENES DIGITALES

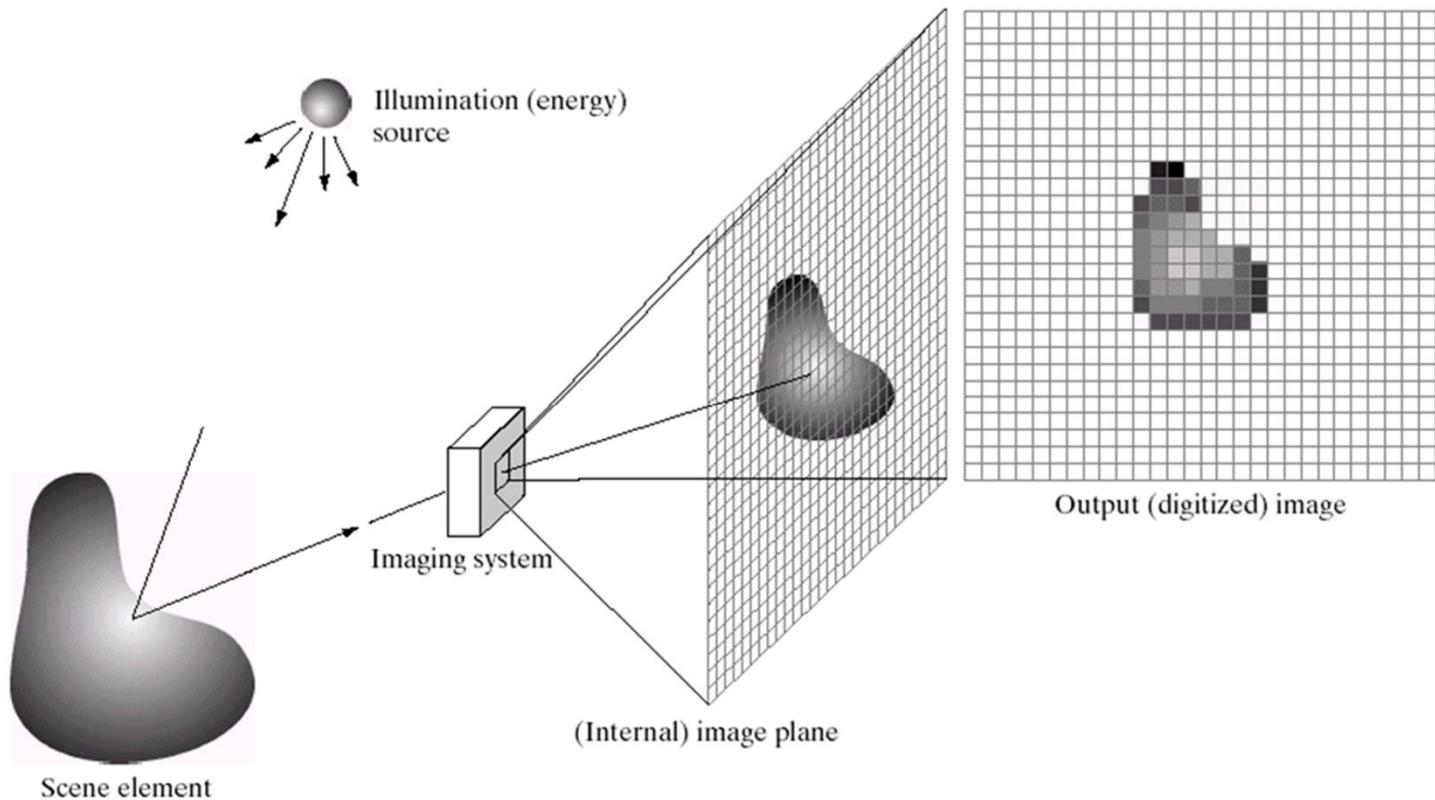


3.1. El espectro electromagnético



- Imágenes basadas en radiación del espectro EM.
- Cada partícula contiene energía.

3.3. Formación de la imagen





3.2. Formación de la imagen (Cont.)

- El proceso de formación de la imagen se da en dos partes:
 - Geometría: Determina donde se localizará cada punto de escena en el plano de la imagen.
 - La física de la luz determina el brillo de un punto en el plano de la imagen

$$f(x, y) = i(x, y)r(x, y)$$

donde i es iluminación y r es reflectancia



3.2. Formación de la imagen (Cont.)

$0 < i(x, y) < \infty$; Iluminación o intensidad de la luz incidente

$0 < r(x, y) < 1$; Reflectancia o reflexividad del objeto.

En situaciones reales: $L_{min} \leq f(x, y) \leq L_{max}$

- Una imagen es una función bidimensional, $f(x, y)$ donde x e y son coordenadas espaciales.
- El valor de $f(x, y)$ en (x, y) es proporcional al brillo de la imagen en ese punto.



3.2. Formación de la imagen (Cont.)

- *Ejemplo.* Primera fotografía digitalizada.





3.2. Formación de la imagen (Cont.)

- Muestreo. Mide el valor de una imagen en un numero finito de puntos. (Resolución)
 - Efecto muestreo.





3.2. Formación de la imagen (Cont.)

- *Cuantización.* Imposibilidad de tener un rango infinito de valores.
 - Efecto Cuantización $f(x, y)$





3.3. Imagen digital

- Una imagen análoga es una función bidimensional continua $f(s, t)$.
- Una imagen digital es una función discreta vista como $f(x, y)$ donde (x, y) son coordenadas espaciales.
- Una imagen digital emplea muestreo y cuantización.
- El valor de $f(x, y)$ en (x, y) es proporcional al brillo de la imagen en ese punto.



3.3. Imagen digital (Cont.)

- Una imagen digital se puede representar como una colección de puntos dispuestos en un arreglo matricial, cada uno con un nivel de gris determinado por $f(x, y)$.
- Los elementos de dicha matriz son conocidos como pixels.
- Un pixel es conocido como picture element, la menor unidad de mismo color en una imagen.
- Un pixel en escala de grises tiene valores de pixel entre 0 y 255.



3.3. Imagen digital (Cont.)

$$I = \begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{m1} & \cdots & x_{mn} \end{bmatrix}$$

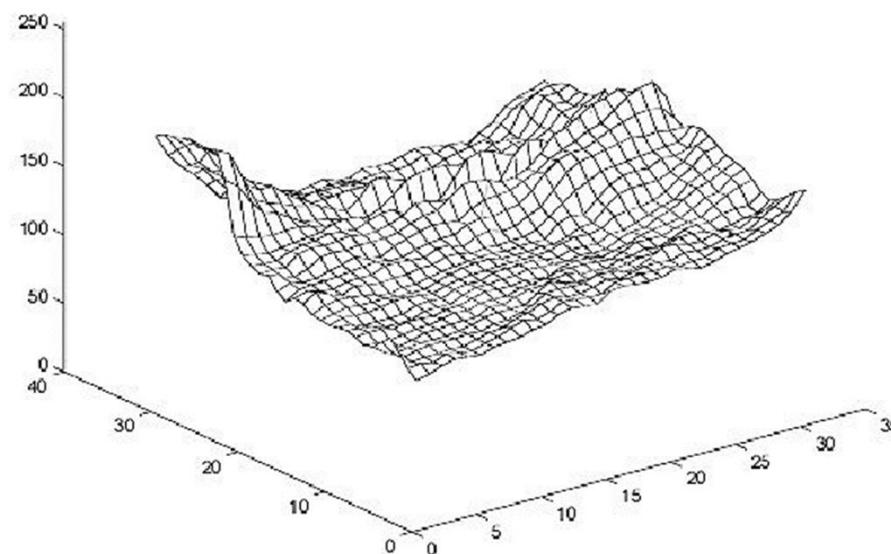
En algunos casos, es necesario considerer a la imagen como una función de variables continuas.

$$I: R \times R \rightarrow [0, M]$$

Donde $I(x, y)$ es un valor de intensidad

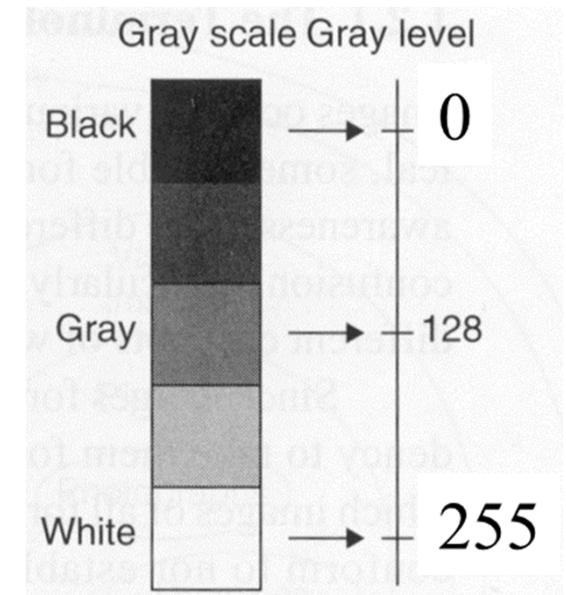


3.3. Imagen digital (Cont.)



3.3. Imagen digital (Cont.)

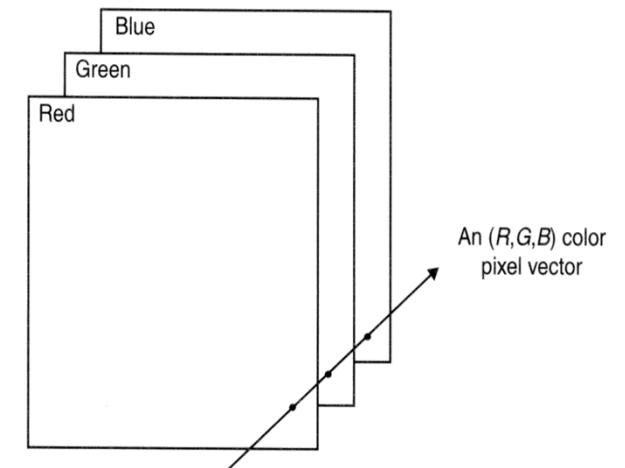
- Una imagen digital es un arreglo de $m \times n$.
- Las intensidades discretas se representan en intensidades en potencia de 2
 - $L = 2^k$
 - El intervalo de intensidad es $[0, L - 1]$
 - Almacenamiento= $m \times n \times k$ bits





3.3. Imagen digital (Cont.)

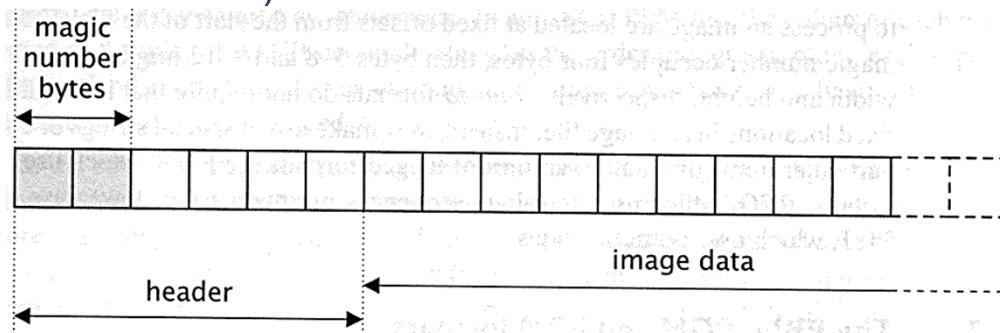
- *Imágenes de Color.* Las imágenes en color se componen de tres canales de color (rojo, verde y azul) que se combinan para crear la mayoría de los colores que podemos ver.





3.4. Formatos de archivo de imagen

- Muchos formatos de imagen se adhieren al modelo simple.
 - El encabezado contiene al menos el ancho y el alto de la imagen.
 - La mayoría de los encabezados comienzan con una firma o un "número mágico" (secuencia corta de bytes para identificar el formato de archivo)





3.4. Formatos de archivo de imagen (Cont.)

- Algunos formatos de imagen
 - GIF (Graphic Interchange Format) -
 - PNG (Portable Network Graphics)
 - JPEG (Joint Photographic Experts Group)
 - TIFF (Tagged Image File Format)
 - PGM (Portable Gray Map)
 - FITS (Flexible Image Transport System)
 - RAW



3.5. Histograma de una imagen

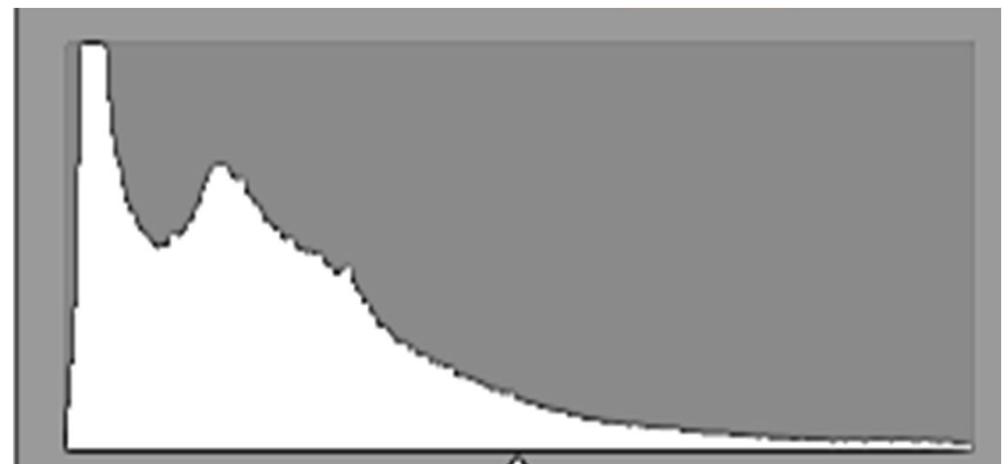
- Sea I una imagen en su representación matricial entonces su histograma será un vector h de $M + 1$ elementos.

$$h = \{ h(i) \mid h(i) = \sum_{I(x,y)=i} 1 , \quad 0 \leq i \leq 255 \}$$

- $h(i)$ representa la frecuencia de aparición de el tono i de gris en la imagen
- *Ejemplo.* Si $h(150) = 39$ significa que hay 39 píxeles de tono de gris 150.



3.5. Histograma de una imagen





3.6. Acumulación de una imagen

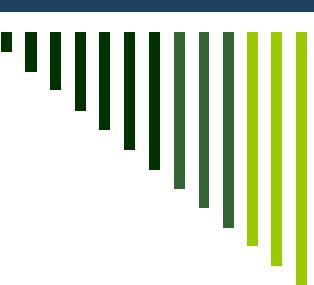
- Sea I una imagen en su representación matricial entonces su histograma será un vector H de $M + 1$ elementos.

$$H(i) = \sum_{j=0}^{i-1} h(j)$$

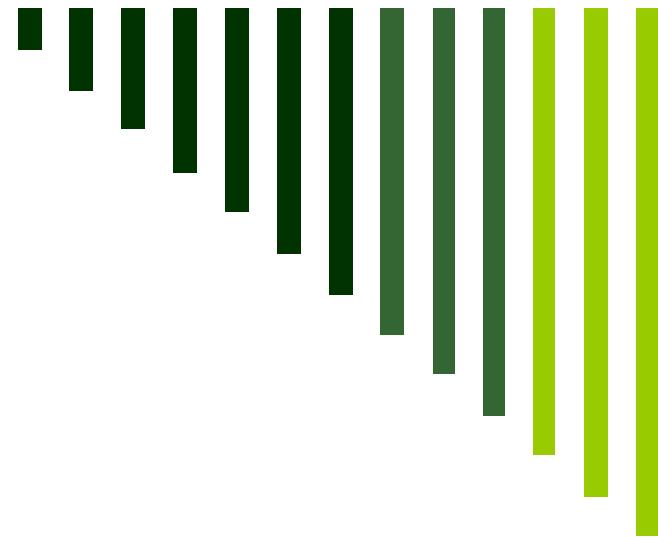
- $H(i)$ representa la frecuencia de aparición de tonos de gris menores que i .
- *Ejemplo.* Si $H(150) = 120$ significa que hay 120 pixeles de tono de gris 150 menores.



4. PROCESAMIENTO DIGITAL DE IMÁGENES



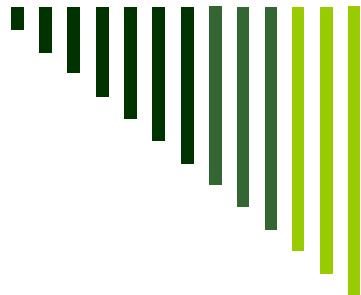
GRACIAS..



Introducción a la Visión Computacional y Procesamiento Gráfico

I Unidad

Ms. Ing. Liz Sofía Pedro H.



Pre-requisitos

- Requeridos:
 - Computación Gráfica II(Obligatorio).
 - Algebra Lineal.
 - Estructura de datos.
 - Estadística (Deseable).
- Programación: MATLAB, JAVA.
- No incluye: Photoshop.

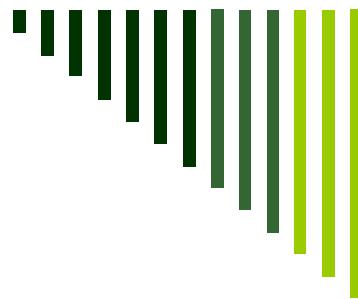


Contenidos.

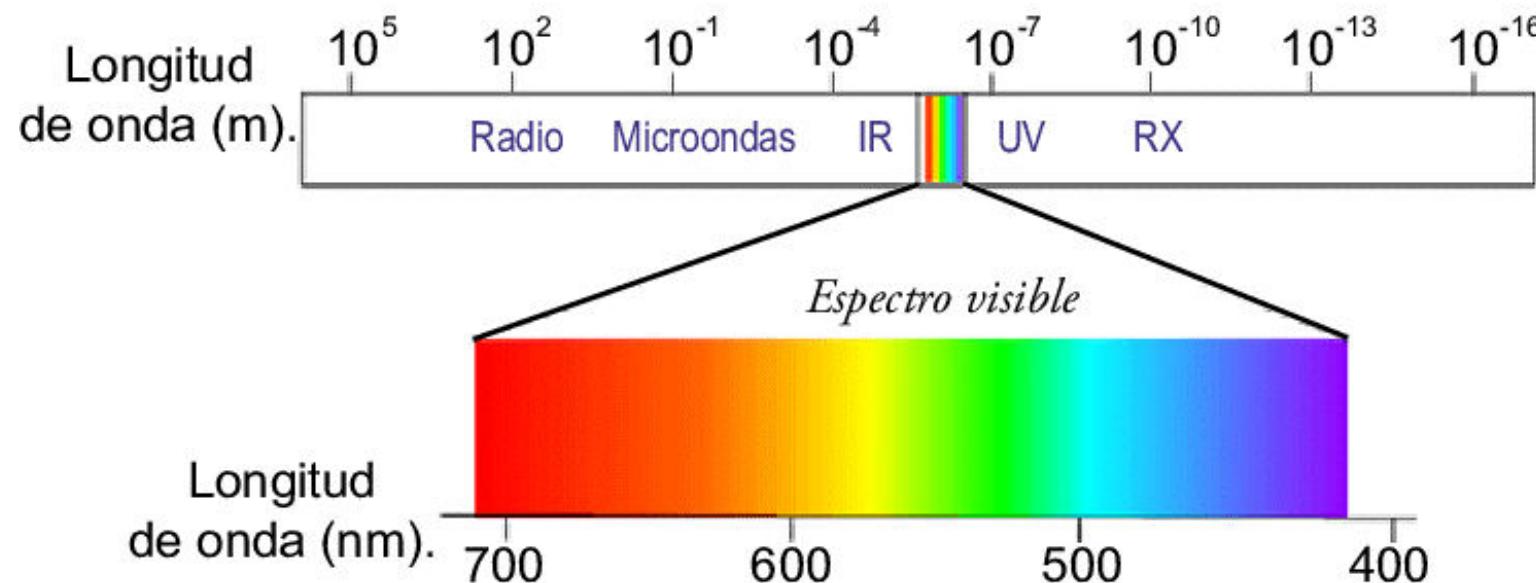
1. Introducción
2. Sistemas de Visión Computacional
3. Procesamiento Digital de Imágenes.
 1. Operaciones estadísticas.
 2. Binarización.



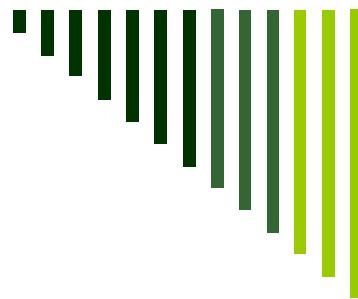
3. IMÁGENES DIGITALES



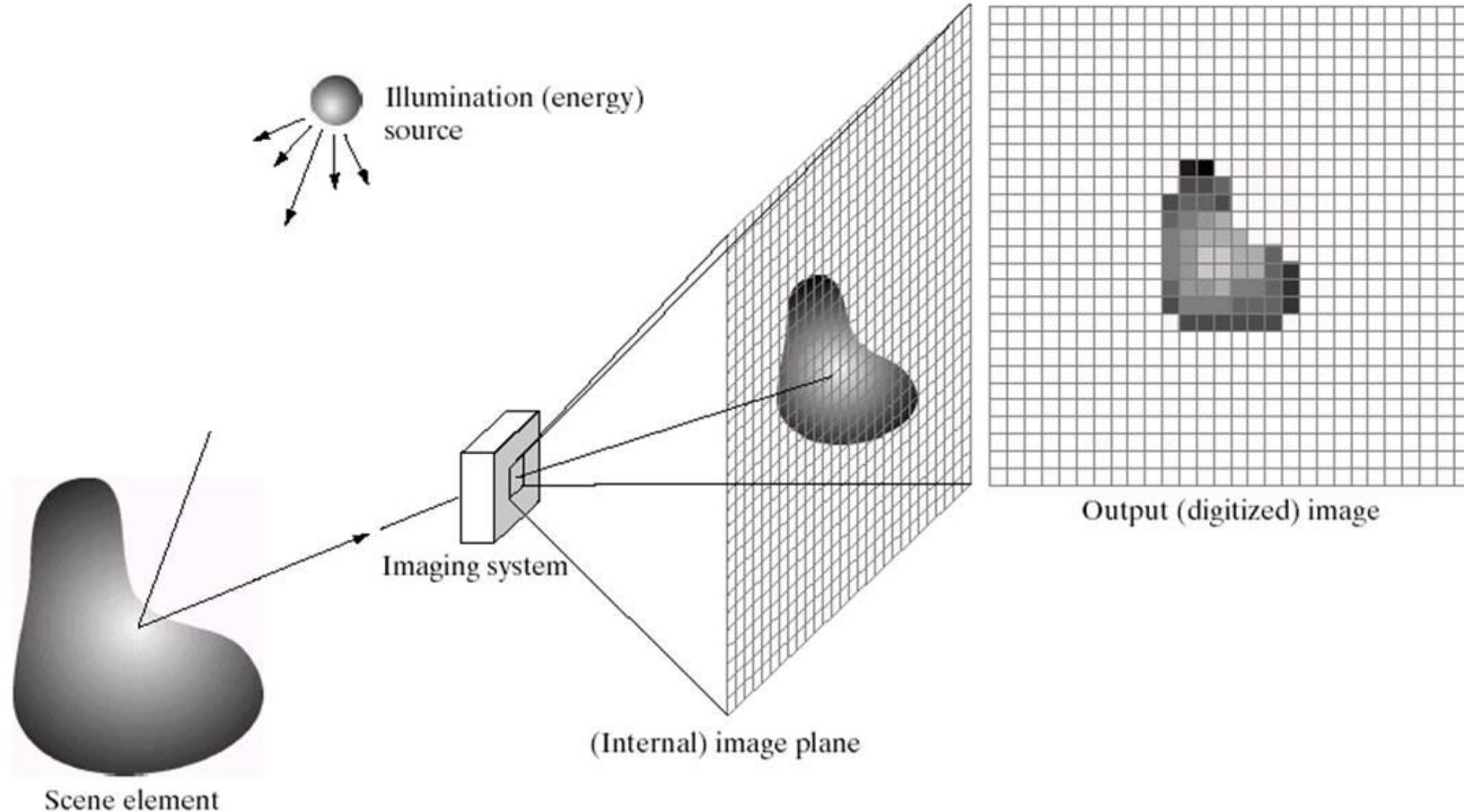
3.1. El espectro electromagnético

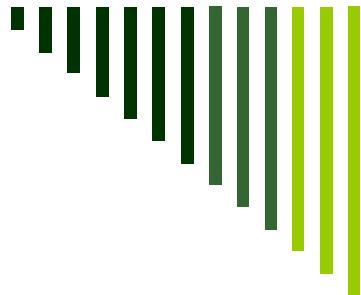


- Imágenes basadas en radiación del espectro EM.
- Cada partícula contiene energía.



3.3. Formación de la imagen





3.2. Formación de la imagen (Cont.)

- El proceso de formación de la imagen se da en dos partes:
 - Geometría: Determina donde se localizará cada punto de escena en el plano de la imagen.
 - La física de la luz determina el brillo de un punto en el plano de la imagen

$$f(x, y) = i(x, y)r(x, y)$$

donde i es iluminación y r es reflectancia



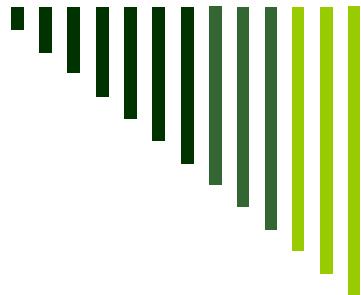
3.2. Formación de la imagen (Cont.)

$0 < i(x, y) < \infty$; Iluminación o intensidad de la luz incidente

$0 < r(x, y) < 1$; Reflectancia o reflexividad del objeto.

En situaciones reales: $L_{min} \leq f(x, y) \leq L_{max}$

- Una imagen es una función bidimensional, $f(x, y)$ donde x e y son coordenadas espaciales.
- El valor de $f(x, y)$ en (x, y) es proporcional al brillo de la imagen en ese punto.



3.2. Formación de la imagen (Cont.)

- *Ejemplo.* Primera fotografía digitalizada.





3.2. Formación de la imagen (Cont.)

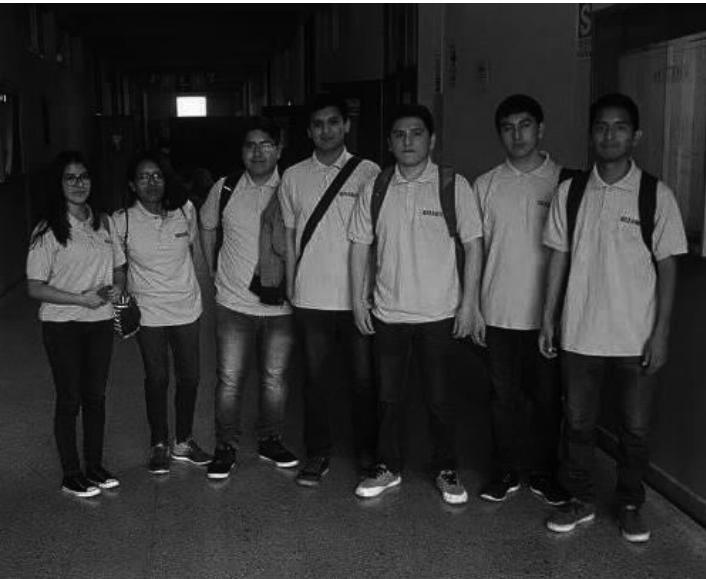
- Muestreo. Mide el valor de una imagen en un numero finito de puntos. (Resolución)
 - Efecto muestreo.





3.2. Formación de la imagen (Cont.)

- *Cuantización.* Imposibilidad de tener un rango infinito de valores.
 - Efecto Cuantización $f(x, y)$





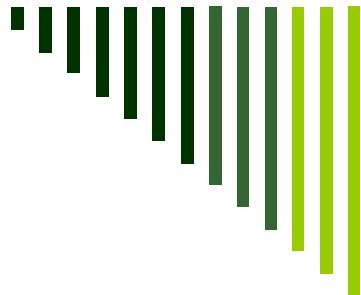
3.3. Imagen digital

- Una imagen análoga es una función bidimensional continua $f(s, t)$.
- Una imagen digital es una función discreta vista como $f(x, y)$ donde (x, y) son coordenadas espaciales.
- Una imagen digital emplea muestreo y cuantización.
- El valor de $f(x, y)$ en (x, y) es proporcional al brillo de la imagen en ese punto.



3.3. Imagen digital (Cont.)

- Una imagen digital se puede representar como una colección de puntos dispuestos en un arreglo matricial, cada uno con un nivel de gris determinado por $f(x, y)$.
- Los elementos de dicha matriz son conocidos como pixels.
- Un pixel es conocido como picture element, la menor unidad de mismo color en una imagen.
- Un pixel en escala de grises tiene valores de pixel entre 0 y 255.



3.3. Imagen digital (Cont.)

$$I = \begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{m1} & \cdots & x_{mn} \end{bmatrix}$$

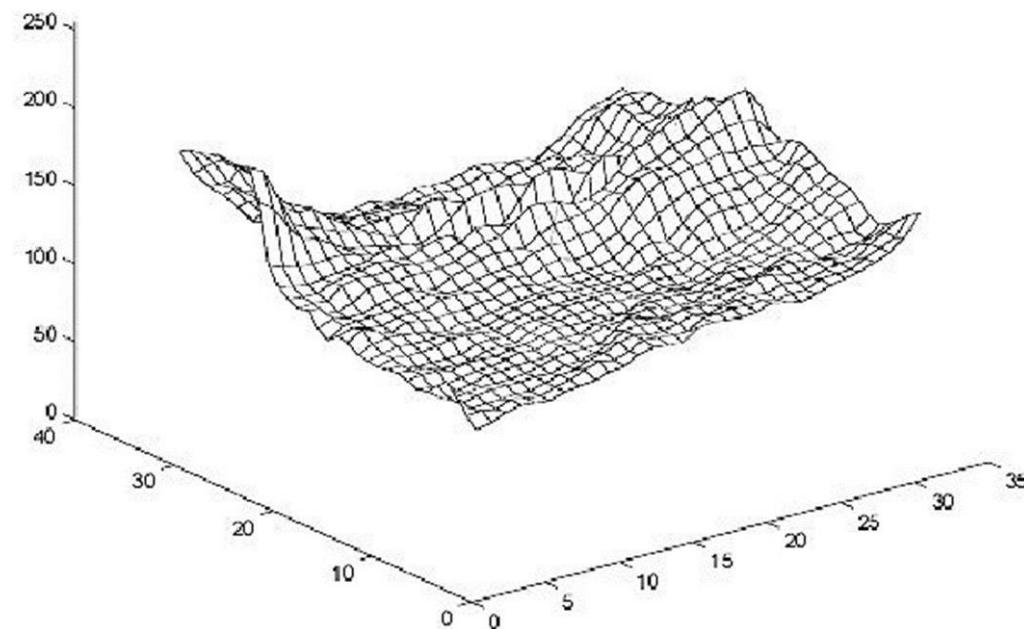
En algunos casos, es necesario considerer a la imagen como una función de variables continuas.

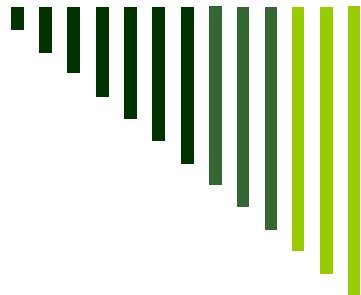
$$I: R \times R \rightarrow [0, M]$$

Donde $I(x, y)$ es un valor de intensidad



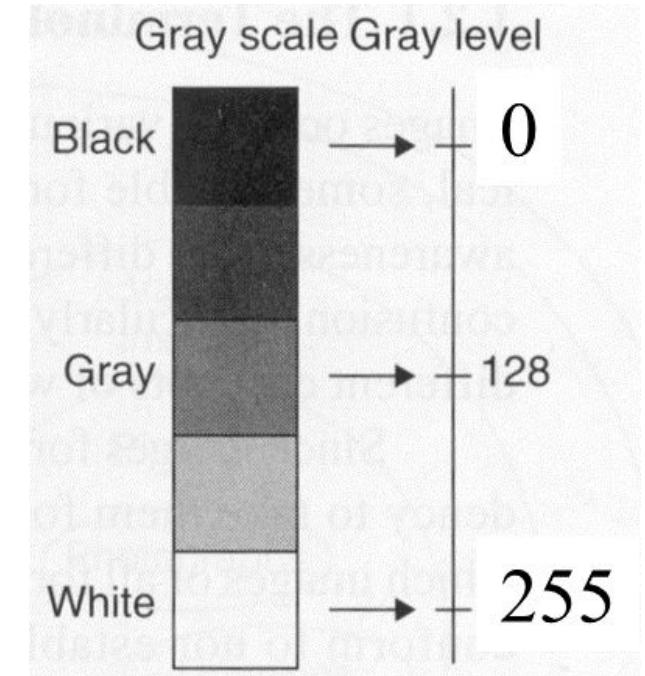
3.3. Imagen digital (Cont.)





3.3. Imagen digital (Cont.)

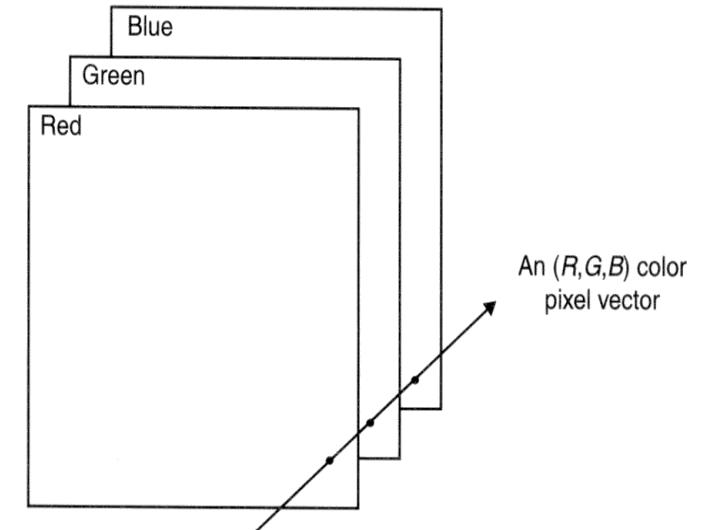
- Una imagen digital es un arreglo de $m \times n$.
- Las intensidades discretas se representan en intensidades en potencia de 2
 - $L = 2^k$
 - El intervalo de intensidad es $[0, L - 1]$
 - Almacenamiento= $m \times n \times k$ bits





3.3. Imagen digital (Cont.)

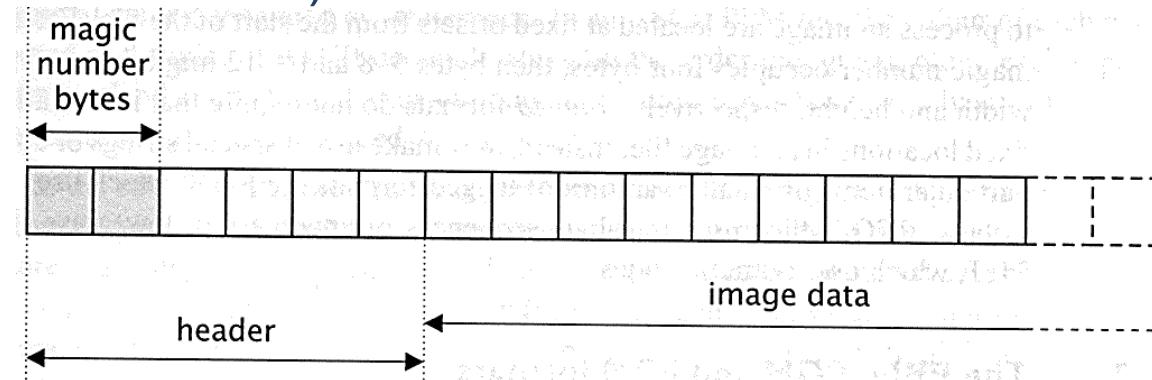
- *Imágenes de Color.* Las imágenes en color se componen de tres canales de color (rojo, verde y azul) que se combinan para crear la mayoría de los colores que podemos ver.

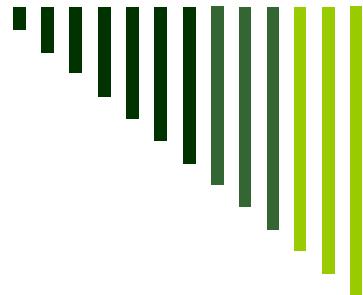




3.4. Formatos de archivo de imagen

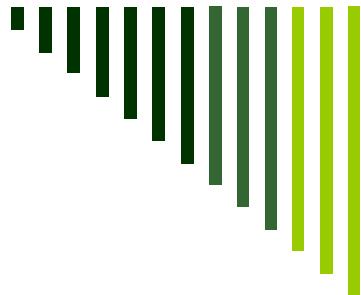
- Muchos formatos de imagen se adhieren al modelo simple.
 - El encabezado contiene al menos el ancho y el alto de la imagen.
 - La mayoría de los encabezados comienzan con una firma o un "número mágico" (secuencia corta de bytes para identificar el formato de archivo)





3.4. Formatos de archivo de imagen (Cont.)

- Algunos formatos de imagen
 - GIF (Graphic Interchange Format) -
 - PNG (Portable Network Graphics)
 - JPEG (Joint Photographic Experts Group)
 - TIFF (Tagged Image File Format)
 - PGM (Portable Gray Map)
 - FITS (Flexible Image Transport System)
 - RAW

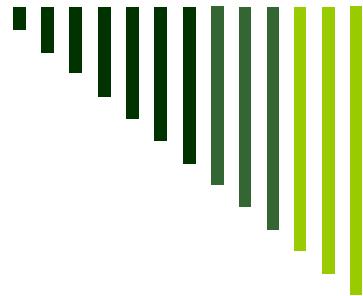


3.5. Histograma de una imagen

- Sea I una imagen en su representación matricial entonces su histograma será un vector h de $M + 1$ elementos.

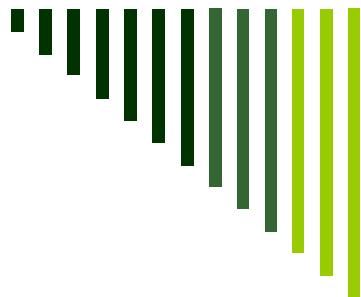
$$h = \{ h(i) \mid h(i) = \sum_{I(x,y)=i} 1 , \quad 0 \leq i \leq 255 \}$$

- $h(i)$ representa la frecuencia de aparición de el tono i de gris en la imagen
- *Ejemplo.* Si $h(150) = 39$ significa que hay 39 pixeles de tono de gris 150.

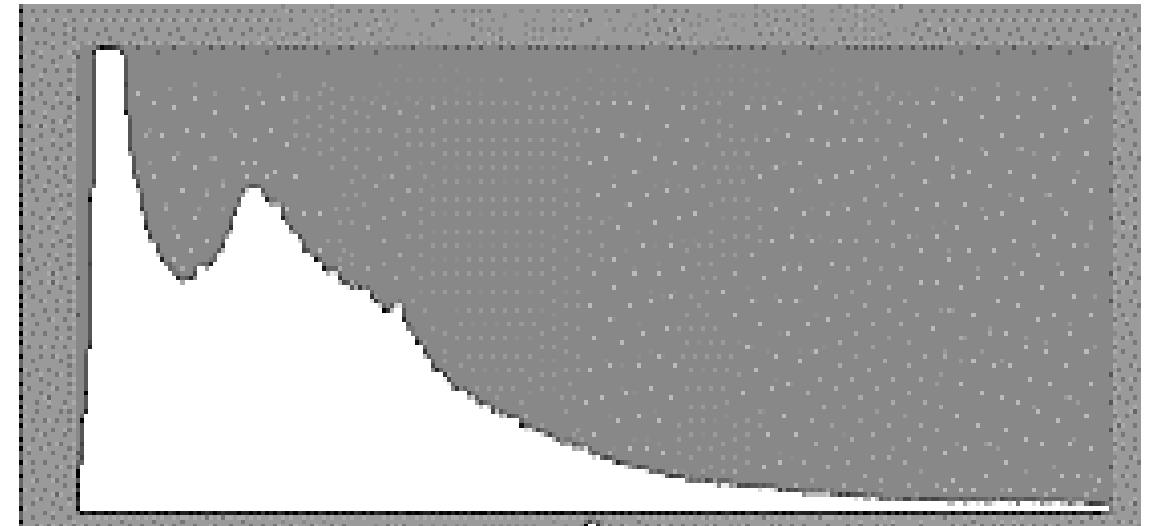


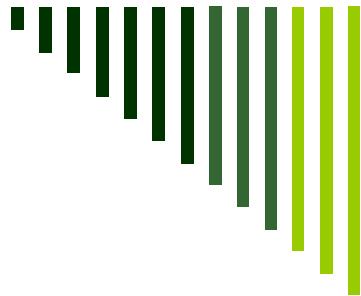
3.5. Histograma de una imagen (Cont.)

- Sea I una imagen en su representación matricial entonces su histograma será un vector h de $M + 1$ elementos.



3.5. Histograma de una imagen (Cont.)





3.6. Acumulación de una imagen

- Sea I una imagen en su representación matricial entonces su histograma será un vector H de $M + 1$ elementos.

$$H(i) = \sum_{j=0}^{i-1} h(j)$$

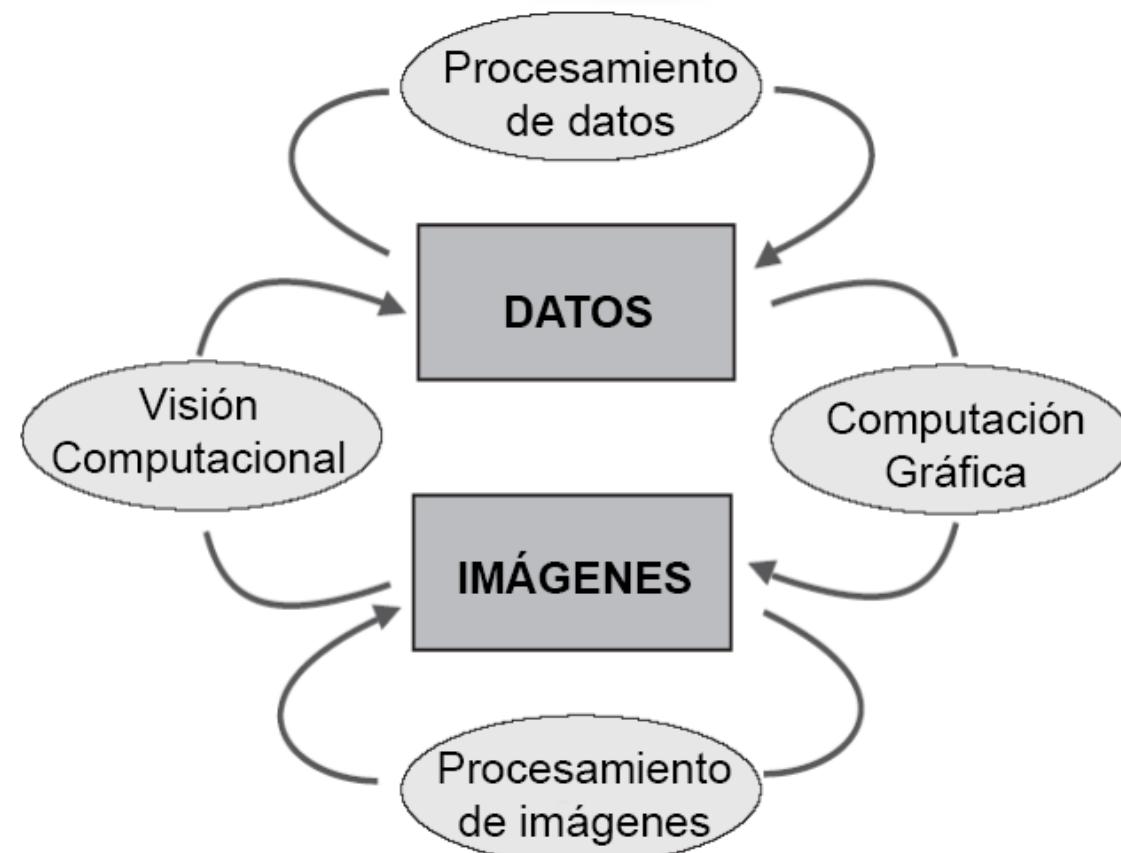
- $H(i)$ representa la frecuencia de aparición de tonos de gris menores que i .
- *Ejemplo.* Si $H(150) = 120$ significa que hay 120 pixeles de tono de gris 150 menores.

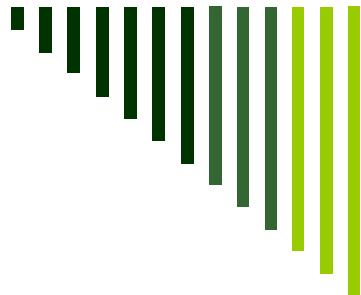


4. PROCESAMIENTO DIGITAL DE IMÁGENES



4.1. Introducción





4.1. Introducción (Cont.)

- *Lo que distinguió la visión por computadora del campo del procesamiento de imágenes digitales fue el deseo de recuperar la estructura tridimensional del mundo a partir de imágenes y usarlo como trampolín hacia la comprensión completa de la escena.*

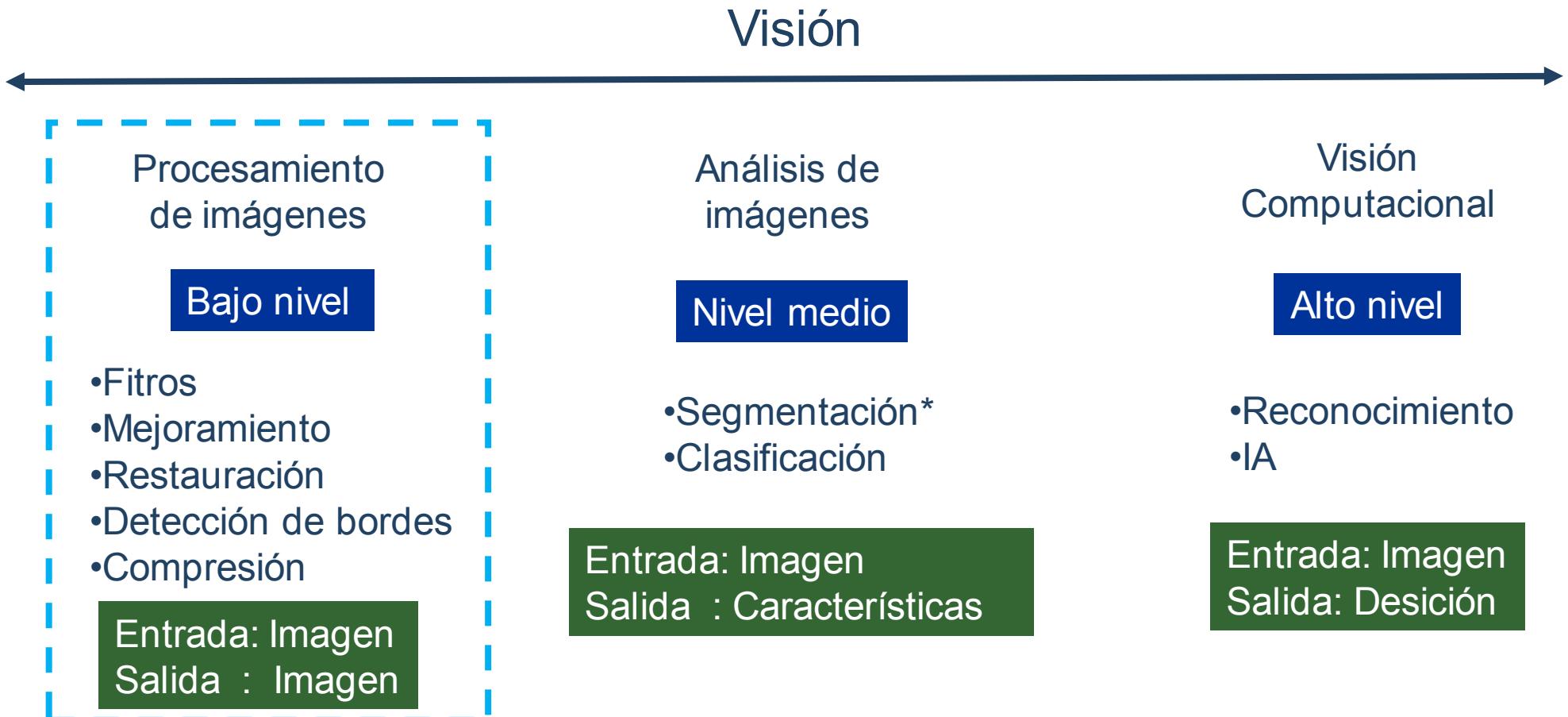


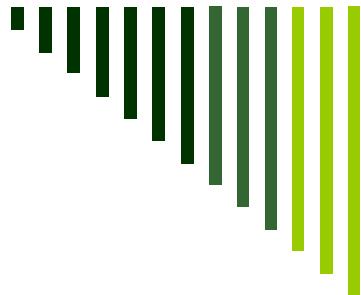
4.2. Definición

- El procesamiento digital de imágenes se centra en dos tareas principales
 - Mejora de la información pictórica para la interpretación humana.
 - Procesamiento de datos de una imagen para almacenamiento, transmisión y representación para que sean percibidos por una máquina autónoma.
- Existen discusiones sobre dónde termina el procesamiento de imágenes y campos como el análisis de imágenes y visión computacional.



4.2. Definición (Cont.)





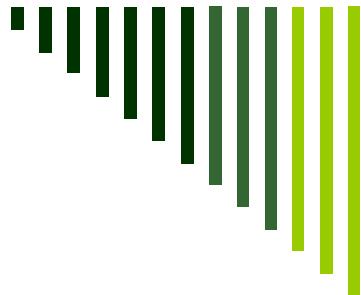
4.3. Historia

- 1920's: una de las primeras aplicaciones de imágenes digitales fue en la industria de los periódicos.
 - El servicio de transmisión de imágenes por cable de Bartlane.
 - Las imágenes fueron transferidas por cable submarino entre Londres y Nueva York.
 - Las imágenes se codificaron para la transferencia por cable y se reconstruyeron en el extremo receptor en una impresora telegráfica
 - 1920: 5 tonos
 - 1929 : 15 tonos de gris



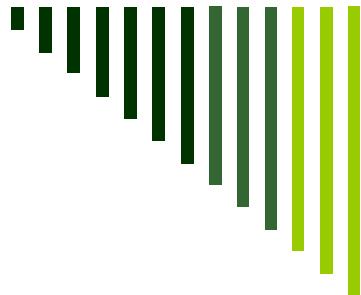
4.3. Historia (Cont.)

- 1950-1960 : verdadero inicio de PDI
 - Objetivo Inicial : mejorar calidad visual de imágenes espaciales
 - 1960's: Mejoras en la tecnología informática y el inicio de la carrera espacial.
 - 1964: Mejoró la calidad de las imágenes de la luna tomadas por la sonda Ranger 7, usándose posteriormente en otras misiones espaciales.



4.3. Historia (Cont.)

- 1970: Uso en aplicaciones médicas.
 - 1979: Sir Godfrey N. Hounsfield y el Prof. Allan M. Cormack inventan la tomografía, precursor de la tomografía axial computarizada (TAC).
- En 1990: el telescopio Hubble puede tomar imágenes de objetos muy distantes, pero las imágenes tenían fallas fueron apoyados por PDI.



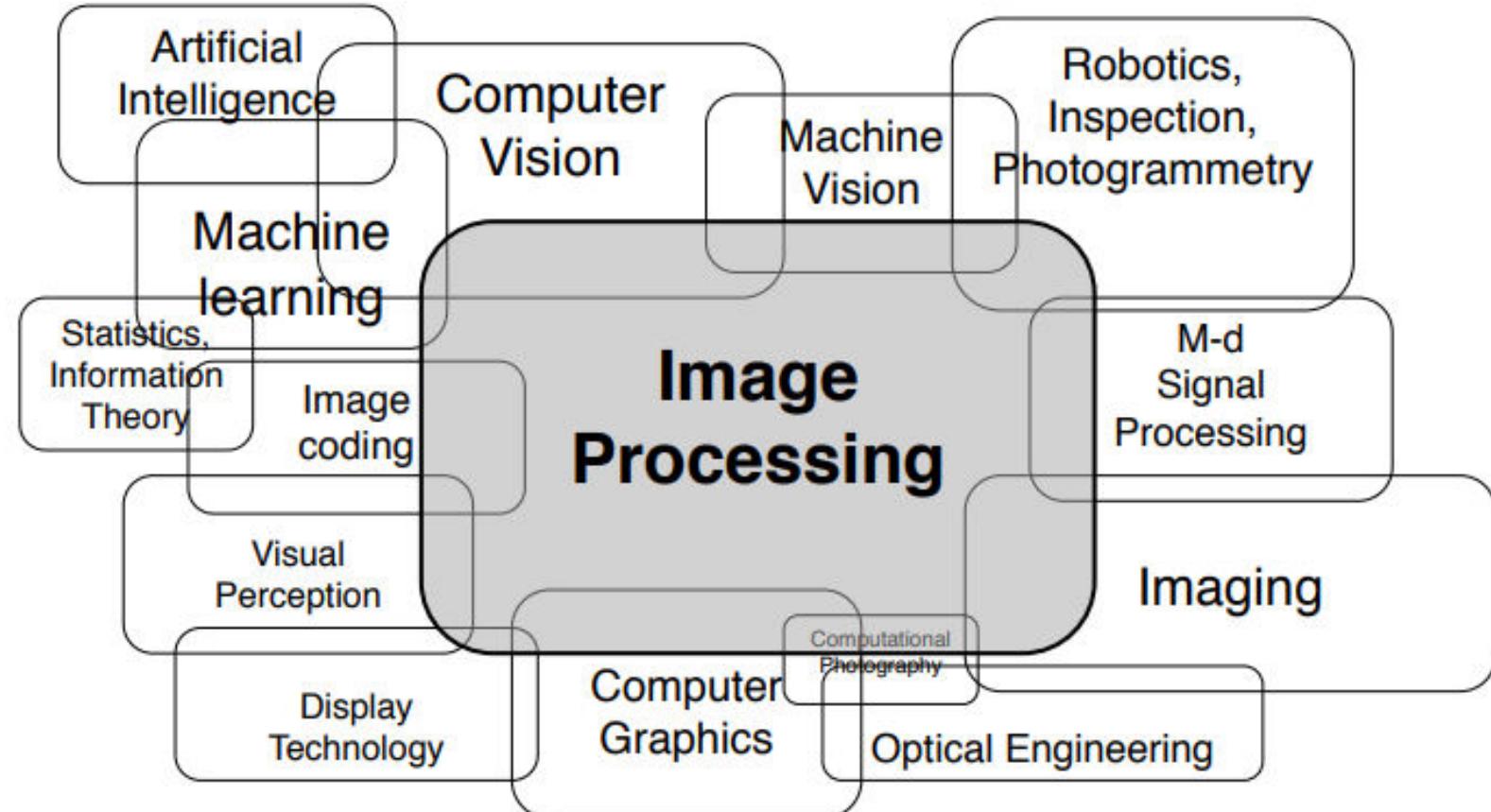
4.3. Historia (Cont.)

Tarea

- 1980 hasta hoy: las técnicas de procesamiento se utilizan para todo tipo de tareas en todo tipo de áreas:
 - Mejora de imagen / restauración,
 - Efectos artísticos,
 - Visualización médica,
 - Inspección industrial,
 - Interfaces de ordenador humano,
 - entre otros.



4.4. Áreas relacionadas (Cont.)

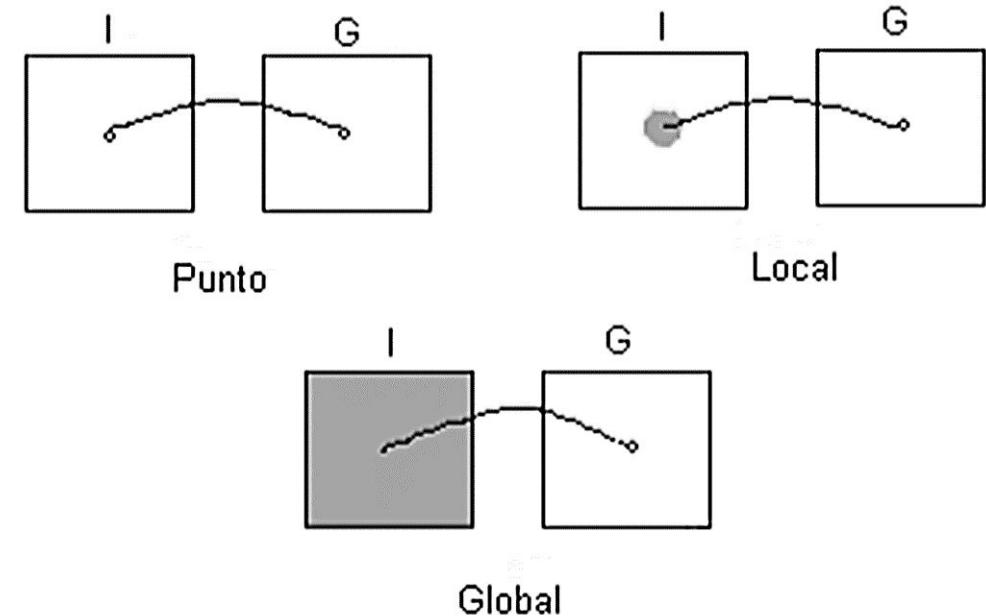


Tarea



4.5. Tipos de Operaciones

- Las operaciones sobre imágenes se pueden clasificar en tres tipos:
 - Punto
 - Local
 - Global





4.6. Operaciones (Cont.)

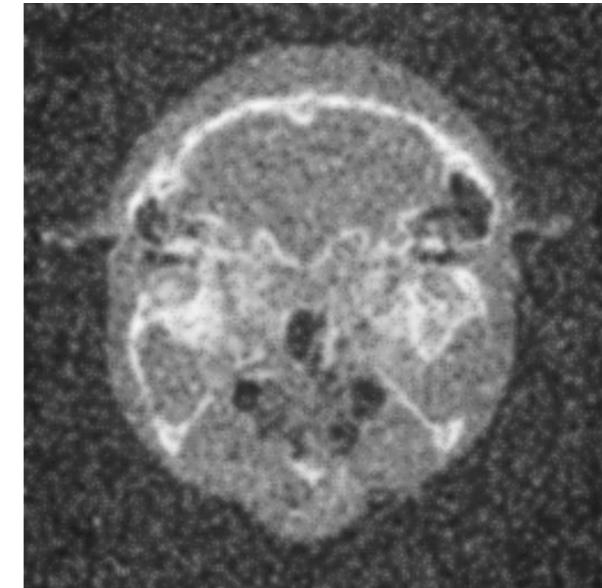
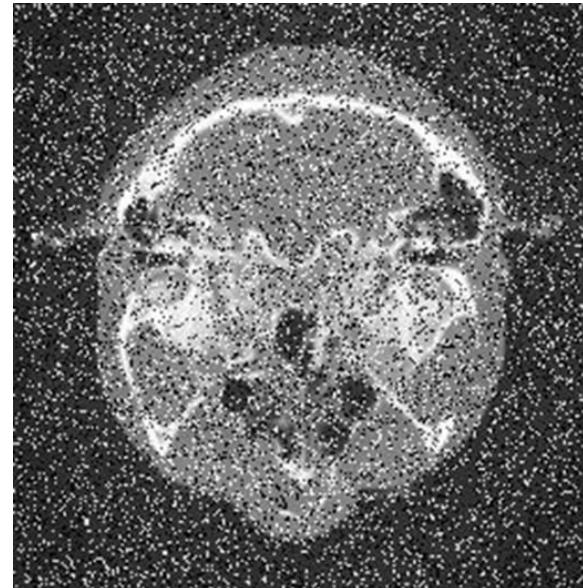
- Binarización





4.6. Operaciones (Cont.)

□ Filtrado





4.6. Operaciones (Cont.)

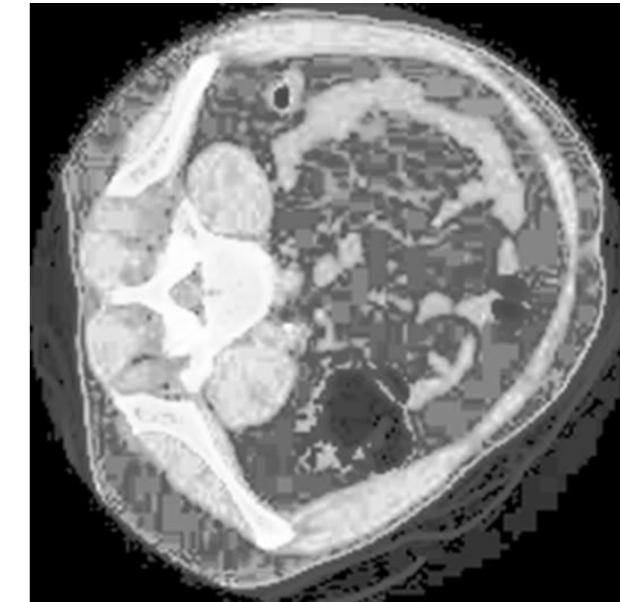
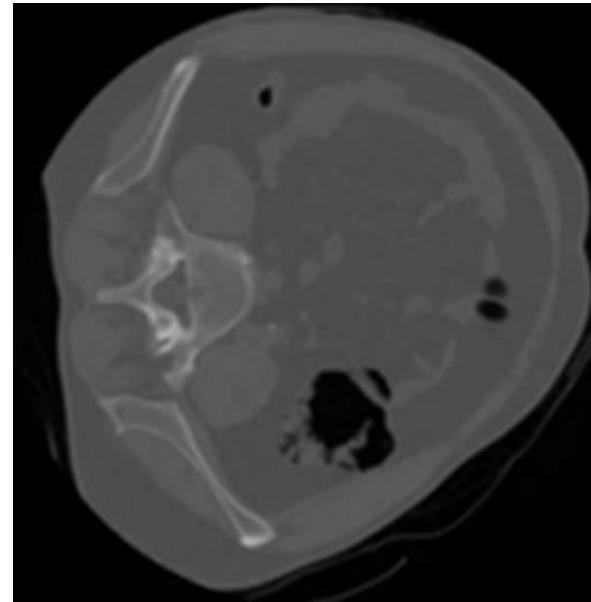
- Detección de bordes





4.6. Operaciones (Cont.)

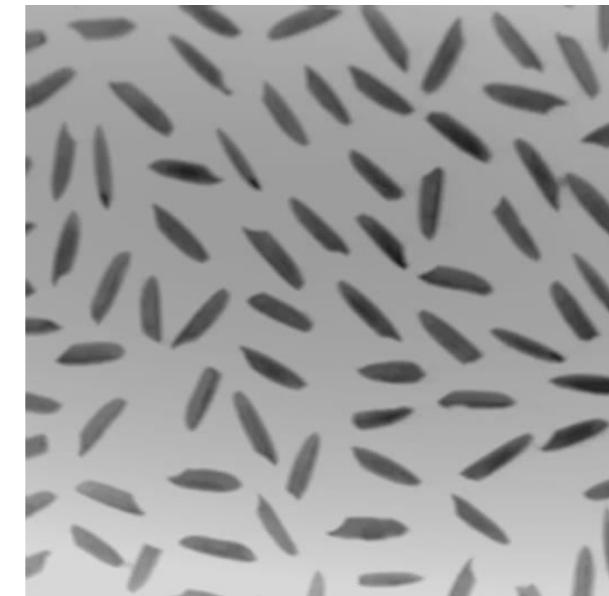
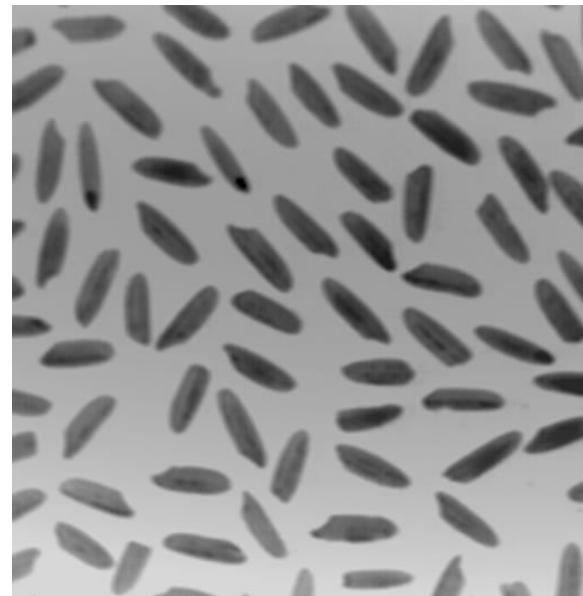
- Mejoramiento de contraste





4.6. Operaciones (Cont.)

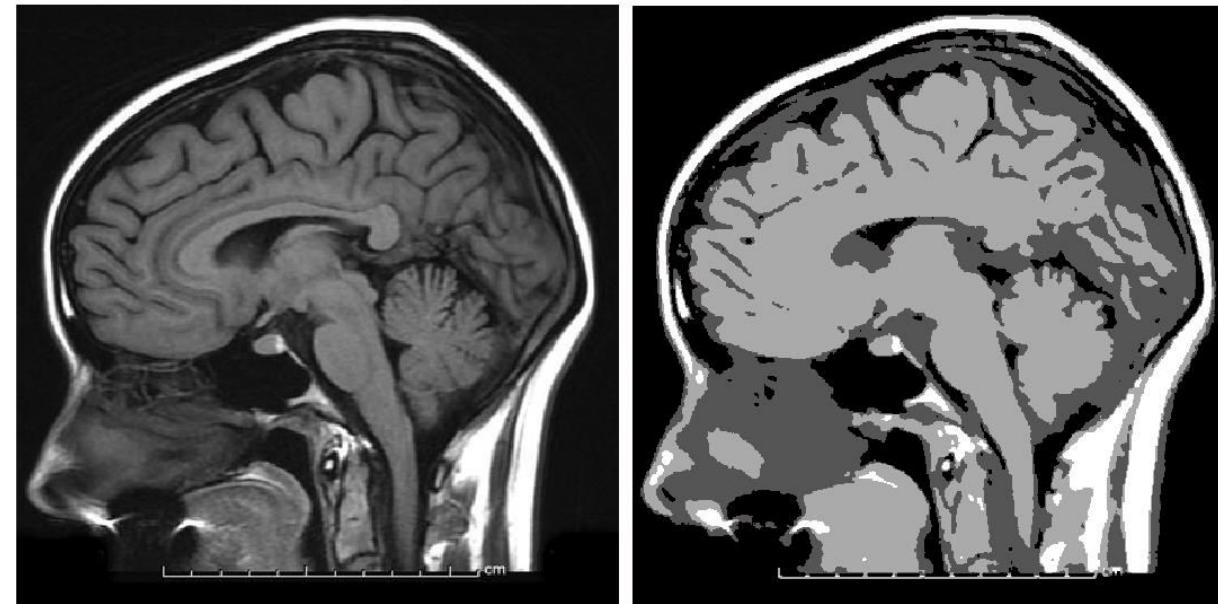
□ Morfología Matemática





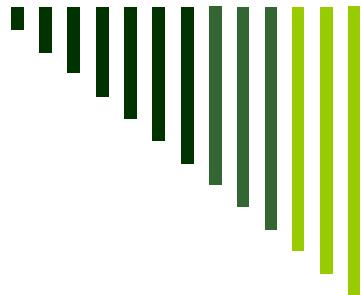
4.6. Operaciones (Cont.)

- Segmentación





5. BINARIZACIÓN



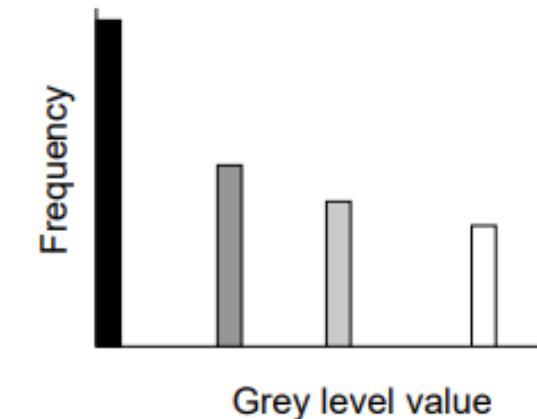
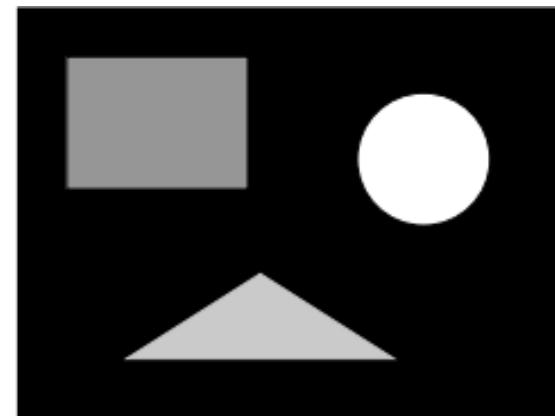
5.1. Binarización.

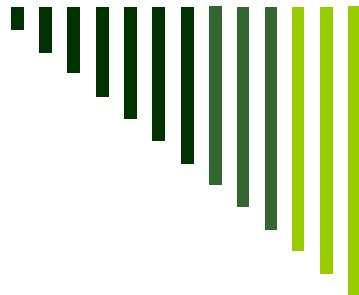
- Separar el objeto del fondo de la imagen.
- La entrada es una imagen en escala de gris.
- La salida es una imagen binaria.
- Una imagen binaria es una imagen que tiene solo dos tonos de gris (negro:0 y blanco:1).
- Utiliza un valor umbral, el cual es calculado empleando diversos métodos.
- Thresholding



5.1. Binarización (Cont.)

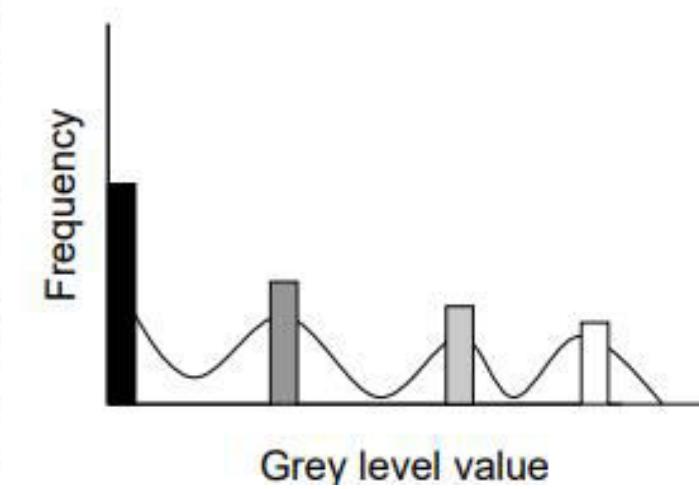
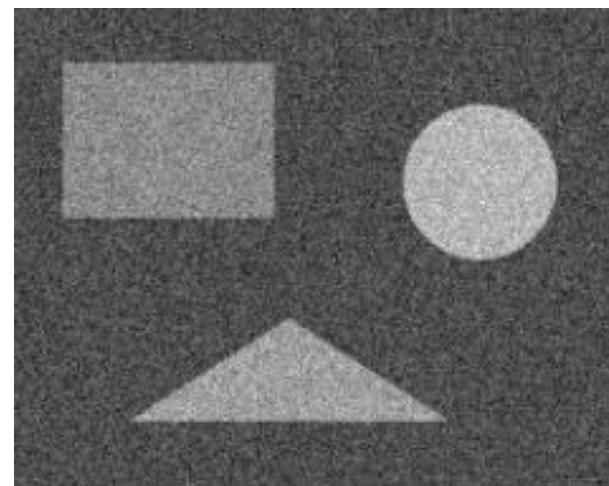
- El histograma de la imagen crea contenedores de intensidades y cuenta número de píxeles en cada nivel, también se puede emplear normalizado.





5.1. Binarización (Cont.)

- El histograma de la imagen se puede ver afectado por el ruido.





5.1. Binarización (Cont.)

- La binarización puede ser vista como:

$$G(x, y) = \begin{cases} 0, & \text{si } I(x, y) \leq T \\ 1, & \text{si } I(x, y) > T \end{cases}$$

Para $I_{m \times n} \rightarrow G_{m \times n}$

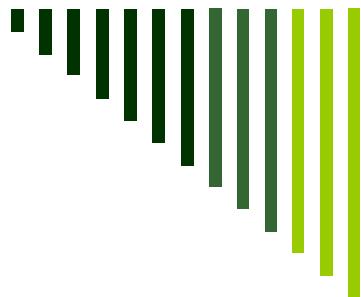


5.1. Binarización (Cont.)



$T = 150$





5.1. Binarización (Cont.)



$$T = 200$$



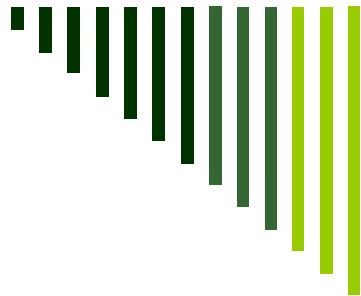


5.1. Binarización (Cont.)



$T = 250$

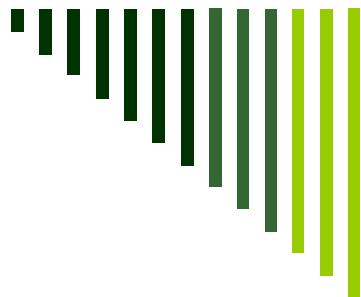




5.1. Binarización (Cont.)

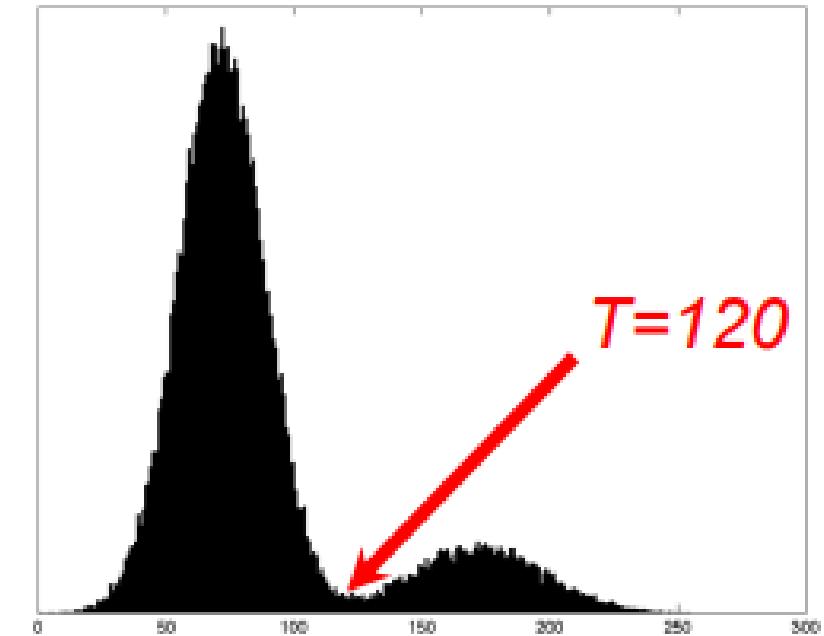
□ Problema: ¿Cómo encontrar el umbral adecuado?

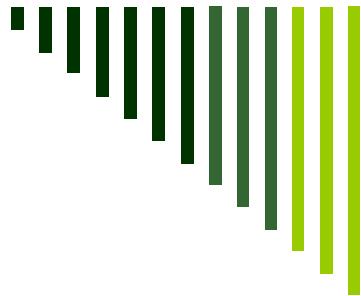
- Usando el histograma de la imagen
- Existen varios métodos.



5.1. Binarización (Cont.)

- Suposición: histograma bimodal
- Uso de frecuencias relativas
- Histograma normalizado.
- Valor umbral debe ubicarse entre 2 máximos

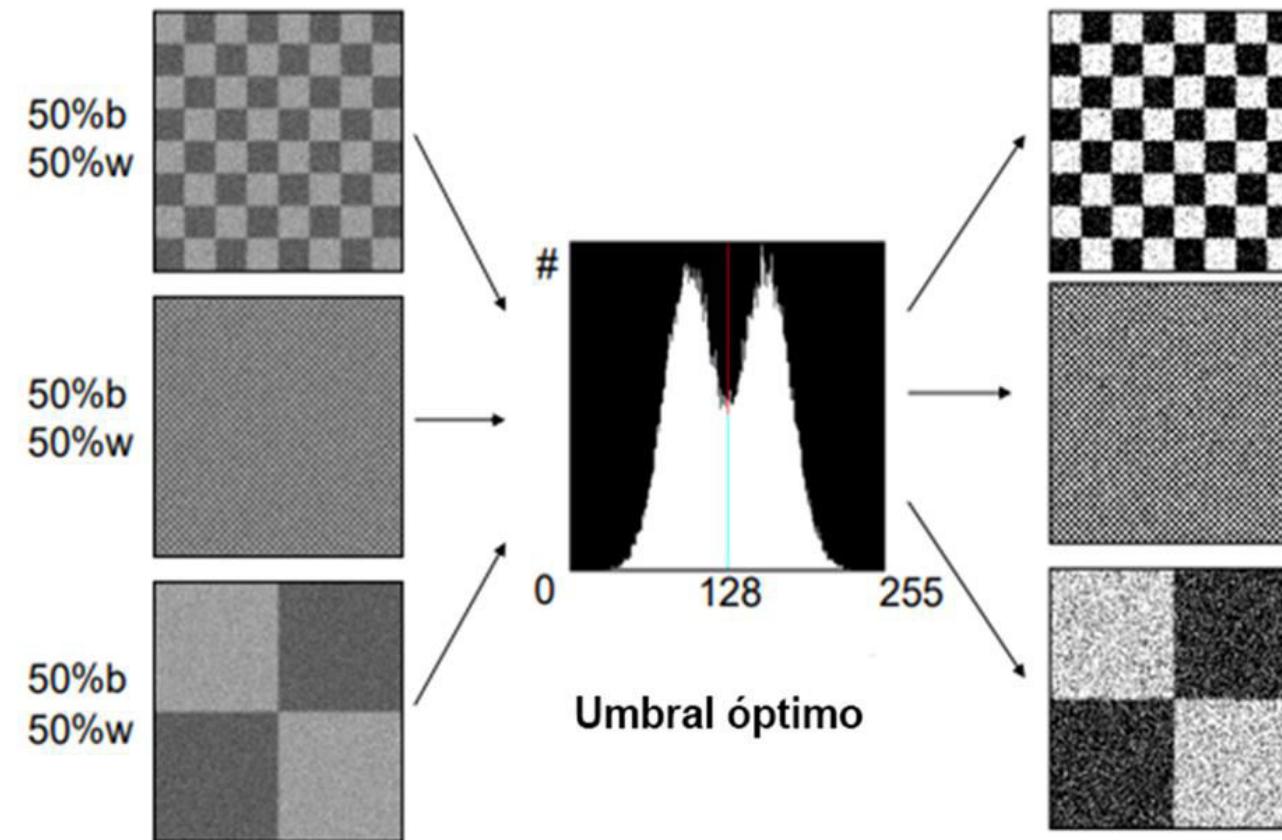




5.1.1. Método Otsu

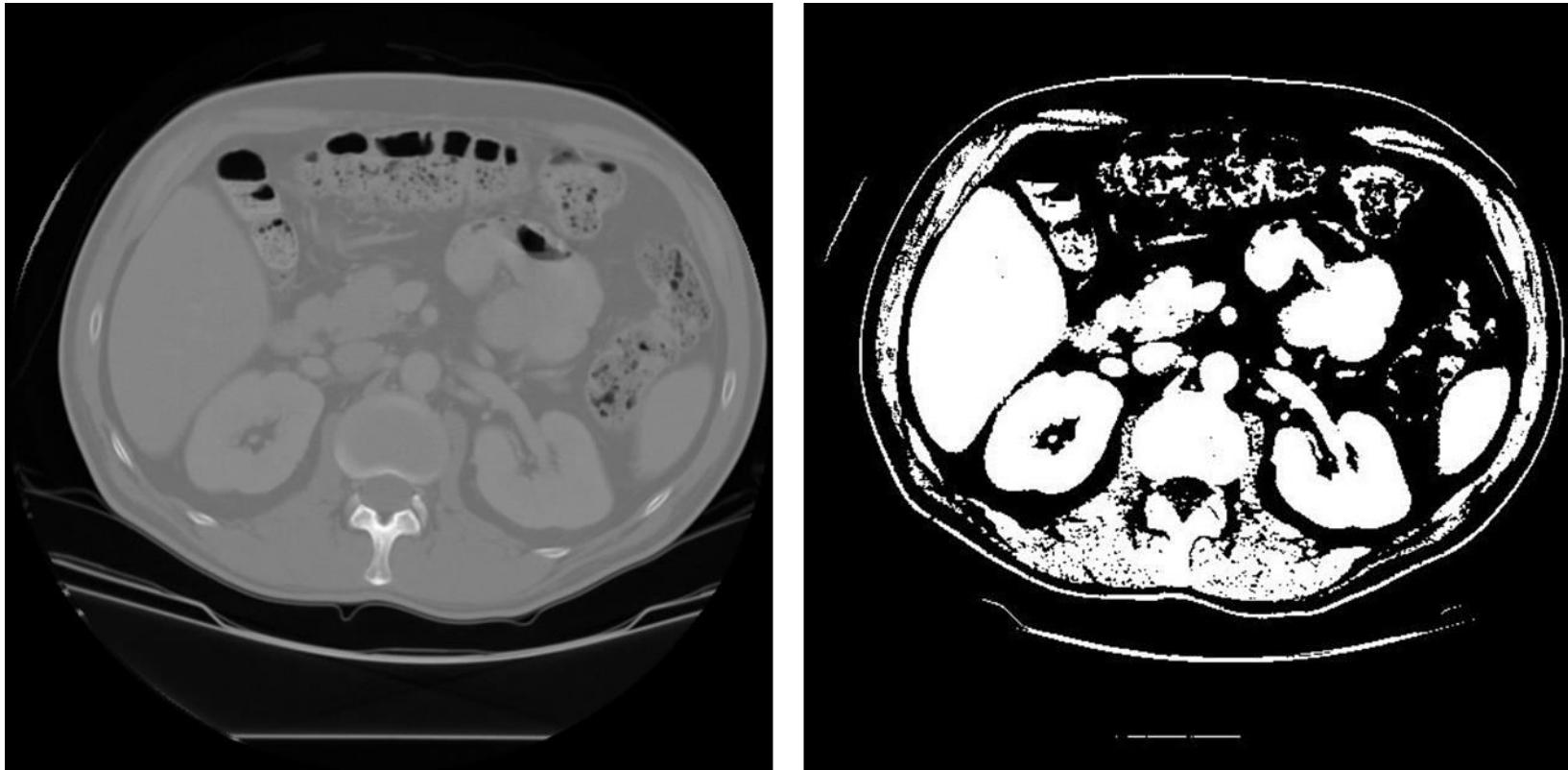
- Procedimiento no paramétrico.
- Considera histograma bimodal.
- Considera dos grupos o clases.
- Selecciona el umbral óptimo maximizando la varianza entre clases mediante una búsqueda exhaustiva y minimiza la suma ponderada de las variaciones dentro de cada.
- Tipo: _____

5.1.1. Método Otsu (Cont.)





5.1.1. Método Otsu (Cont.)





5.1.1. Método Otsu (Cont.)

□ Ventajas:

- Buena respuesta del método frente condiciones inadecuadas: ruido, sin máximos y mínimos diferenciados, mala iluminación, entre otros.
- No precisa de supervisión humana, preprocesamiento u otro tipo de información acerca de la misma.



(a) Two distinct modes



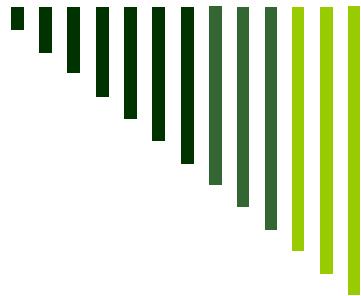
(b) Overlapped modes



5.1.1. Método Otsu (Cont.)

□ Desventajas:

- A mayor número de clases en la imagen aumenta, el método necesita mucho más tiempo para seleccionar un umbral multivariado adecuado.



5.1.1. Método Otsu (Cont.)

□ Funcionamiento:

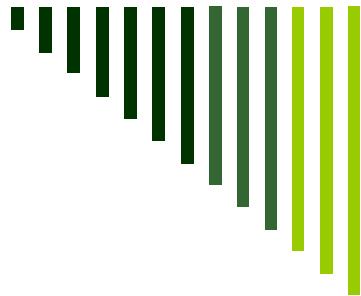
- Sea la imagen I en escala de gris con N píxeles y L posibles niveles diferentes.
- Probabilidad de ocurrencia del nivel de gris i en la imagen:

$$p_i = \frac{f_i}{N}$$

Donde:

f_i : Frecuencia de repetición del nivel de gris i -ésimo con $i = 1, 2, \dots, L$.

Es similar a _____



5.1.1. Método Otsu (Cont.)

□ Funcionamiento:

- Los píxeles se dividen en dos clases C_1 y C_2 , con niveles de gris $[1,2,\dots,t]$ y $[t+1,t+2,\dots,L]$ respectivamente, donde las distribuciones de probabilidad de ambas clases son:

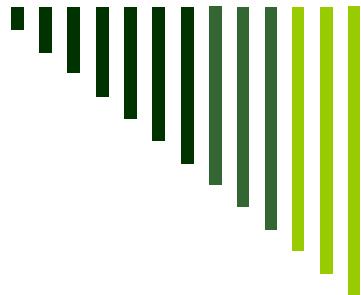
$$C_1 : \frac{p_1}{\omega_1(t)}, \dots, \frac{p_t}{\omega_1(t)}$$

Donde:

$$\omega_1(t) = \sum_{i=1}^t p_i$$

$$C_2 : \frac{p_{t+1}}{\omega_2(t)}, \frac{p_{t+2}}{\omega_2(t)}, \dots, \frac{p_L}{\omega_2(t)}$$

$$\omega_2(t) = \sum_{i=t+1}^L p_i$$



5.1.1. Método Otsu (Cont.)

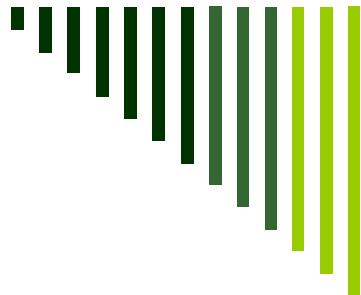
□ Funcionamiento:

- El promedio ponderado para cada una de las clases se definen como:

$$\mu_1 = \sum_{i=1}^t \frac{i \cdot p_i}{\omega_1(t)} \quad \mu_2 = \sum_{i=t+1}^L \frac{i \cdot p_i}{\omega_2(t)}$$

- La intensidad media total de la imagen es:

$$\omega_1 \cdot \mu_1 + \omega_2 \cdot \mu_2 = \mu_T \quad \omega_1 + \omega_2 = 1$$



5.1.1. Método Otsu (Cont.)

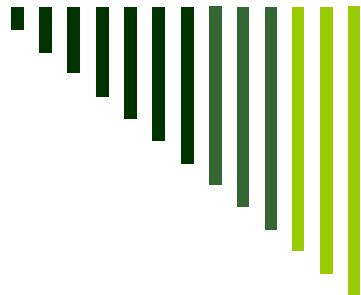
□ Funcionamiento:

- El promedio ponderado para cada una de las clases se definen como:

$$\mu_1 = \sum_{i=1}^t \frac{i \cdot p_i}{\omega_1(t)} \quad \mu_2 = \sum_{i=t+1}^L \frac{i \cdot p_i}{\omega_2(t)}$$

- La intensidad media total de la imagen es:

$$\omega_1 \cdot \mu_1 + \omega_2 \cdot \mu_2 = \mu_T \quad \omega_1 + \omega_2 = 1$$



5.1.1. Método Otsu (Cont.)

□ Funcionamiento:

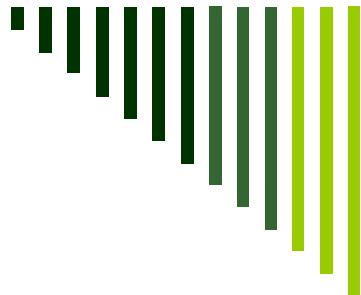
- Por medio del análisis discriminante, la varianza entre clases de una imagen umbralizada se define como:

$$\sigma_B^2 = \omega_1 \cdot (\mu_1 - \mu_T)^2 + \omega_2 \cdot (\mu_2 - \mu_T)^2$$

- Luego, se debe encontrar el umbral, t , que maximice la varianza (umbral óptimo):

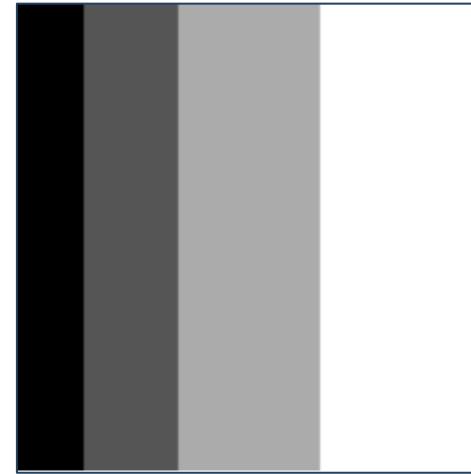
$$t^* = \underset{t}{\operatorname{Max}} \{\sigma_B^2(t)\}$$

Donde: $1 \leq t \leq L$

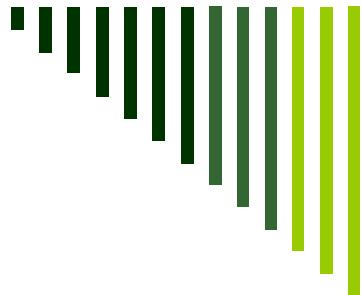


5.1.1. Método Otsu (Cont.)

- Ejemplo:
 - Sea la siguiente imagen con 4 tonalidades de gris.



$$\begin{aligned}L &= 4 \rightarrow [0,85,171,255] \\f_0 &= 10, \\f_{85} &= 20, \\f_{171} &= 30, \\f_{255} &= 40 \\N &= 10 \times 10 = 100\end{aligned}$$



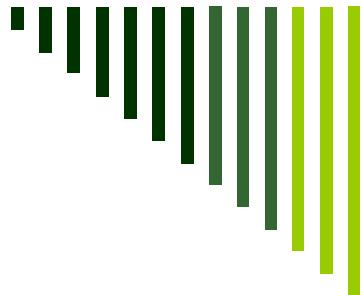
5.1.1. Método Otsu (Cont.)

□ Ejemplo:

- Calcular los valores w_1 y w_2 para $t = 85$

$$C_1 \leftrightarrow [0,85] \quad C_2 \leftrightarrow [171,255]$$

$$\omega_1(t=2) = \sum_{i=1}^t p_i = \frac{1}{10} + \frac{1}{5} = \frac{3}{10} \quad \omega_2(t=2) = \sum_{i=t+1}^L p_i = \frac{3}{10} + \frac{4}{10} = \frac{7}{10}$$



5.1.1. Método Otsu (Cont.)

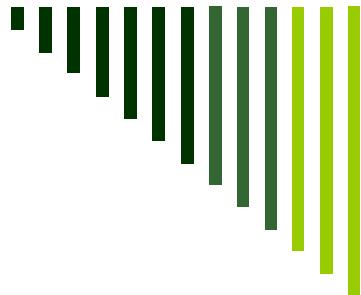
□ Ejemplo:

- Calcular la varianza entre clases para todo valor de umbral posible ($t = 85$):

$$C_1: \frac{p_1}{\omega_1(t=2)}, \frac{p_2}{\omega_1(t=2)} = \left\{ \frac{1}{3}, \frac{2}{3} \right\} \quad C_2: \frac{p_3}{\omega_2(t=2)}, \frac{p_4}{\omega_2(t=2)} = \left\{ \frac{3}{7}, \frac{4}{7} \right\}$$

$$\mu_1 = \sum_{i=1}^t \frac{ip_i}{\omega_1(t=2)} = \frac{1}{3} + 2 \times \frac{2}{3} = \frac{5}{3}$$

$$\mu_2 = \sum_{i=t+1}^L \frac{ip_i}{\omega_2(t=2)} = 3 \times \frac{3}{7} + 4 \times \frac{4}{7} = \frac{25}{7}$$



5.1.1. Método Otsu (Cont.)

□ Ejemplo:

- Calcular la varianza entre clases para todo valor de umbral posible ($t = 85$):

$$\mu_T = \omega_1(t = 2)\mu_1 + \omega_2(t = 2)\mu_2 = 3$$

$$\sigma_B^2 = \omega_1(t = 2)(\mu_1 - \mu_T)^2 + \omega_2(t = 2)(\mu_2 - \mu_T)^2 \approx 0.7619$$

$$\sigma_B^2(t = 1) = 0.4444$$

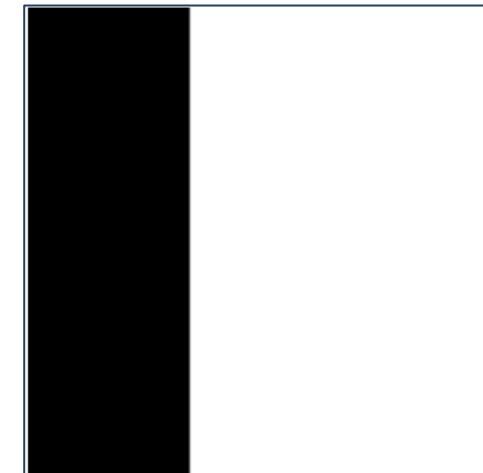
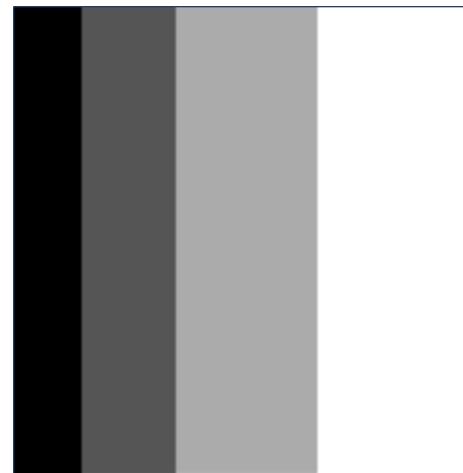
$$\sigma_B^2(t = 3) = 0.6667$$

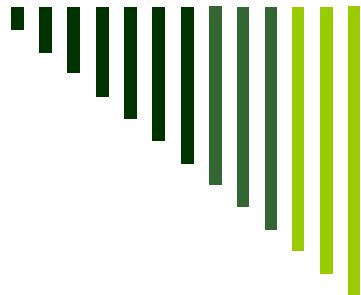


5.1.1. Método Otsu (Cont.)

□ Ejemplo:

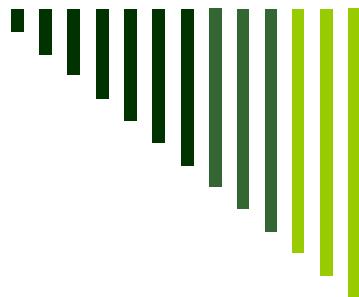
- El mejor umbral para Otsu es: $t = 85$



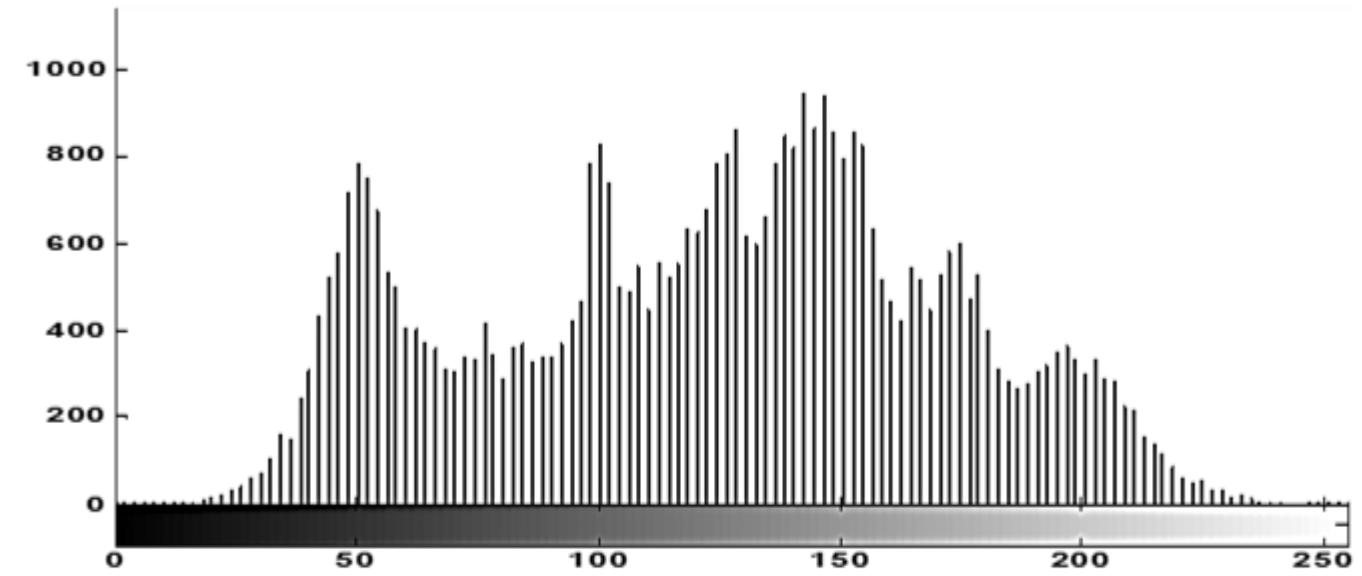


5.1.2. Máxima Entropía

- La entropía de la imagen especifica la incertidumbre en los valores de la imagen.
- Mide la cantidad promedio de información requerida para codificar los valores de la imagen.
- Eventos poco frecuentes proveen más información que los eventos frecuentes.
- La entropía es una medida de dispersión del histograma.

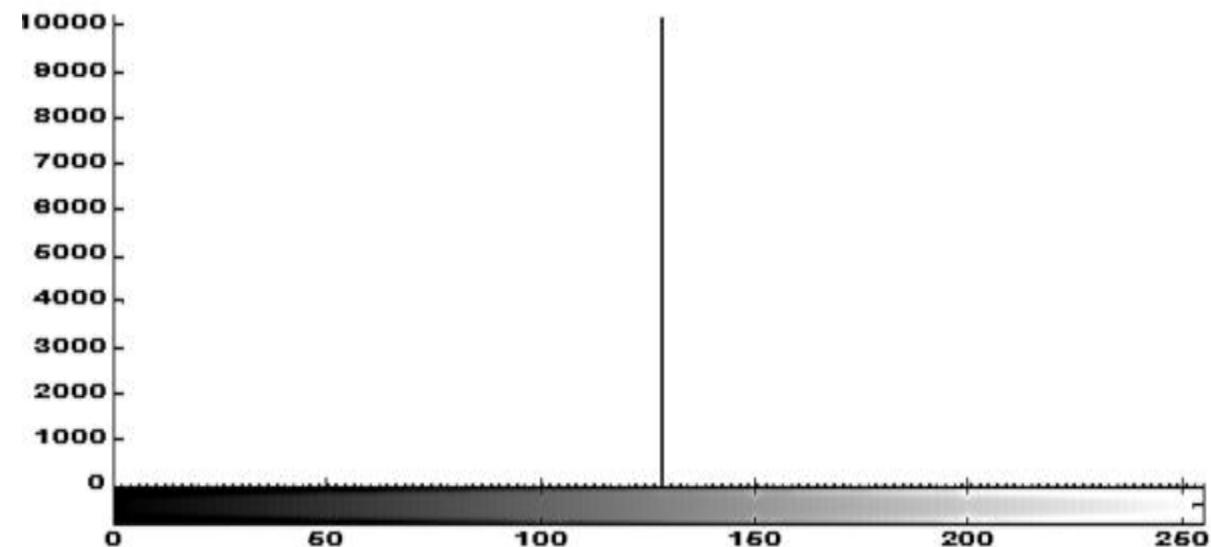
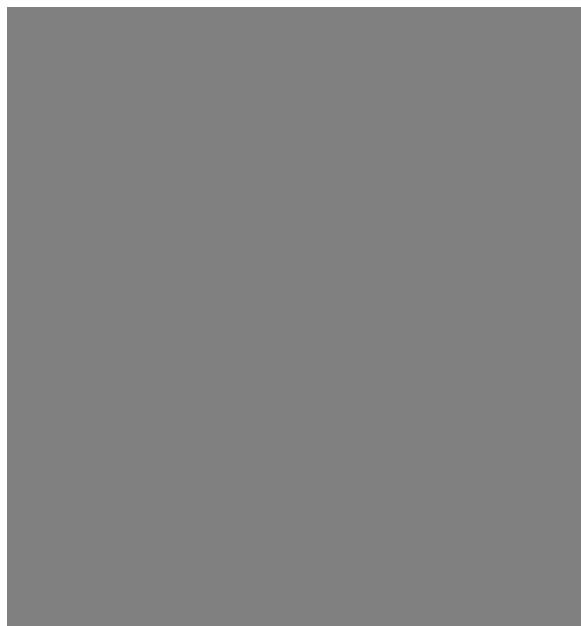


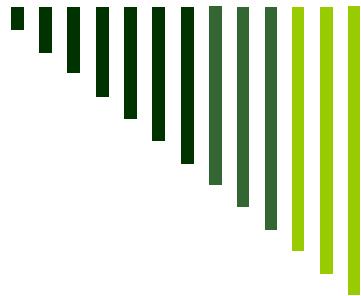
5.1.2. Máxima Entropía (Cont.)





5.1.2. Máxima Entropía (Cont.)





5.1.2. Máxima Entropía (Cont.)

- El método de la máxima entropía es un método de binarización donde el valor óptimo del umbral puede ser encontrado maximizando la entropía entre las clases resultantes (fondo y objeto).
- Es uno de los métodos de selección de umbral basado en entropía considera una imagen como una fuente de información.
- Supongamos que $h(i)$ es el histograma normalizado, el cual será usado por el método.
- Método Kapur.



5.1.2. Máxima Entropía (Cont.)

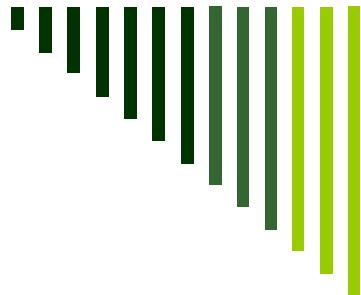
□ Funcionamiento:

- Calcular la entropía de los pixeles que serían de color negro para cada posible valor umbral

$$entropia_w = - \sum_{i=1}^t \left(\frac{P(i)}{\sum_{j=1}^t P(j)} \log_2 \frac{P(i)}{\sum_{j=1}^t P(j)} \right)$$

- Calcular la entropía de los pixeles que serían de color blanco de la imagen para cada posible valor umbral.

$$entropia_b = - \sum_{i=t+1}^L \left(\frac{P(i)}{1 - \sum_{j=1}^t P(j)} \log_2 \frac{P(i)}{1 - \sum_{j=1}^t P(j)} \right)$$

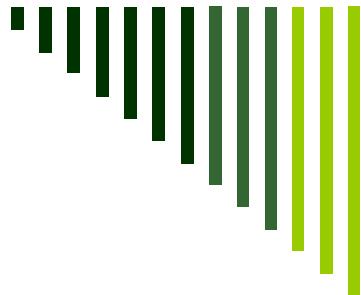


5.1.2. Máxima Entropía (Cont.)

□ Funcionamiento:

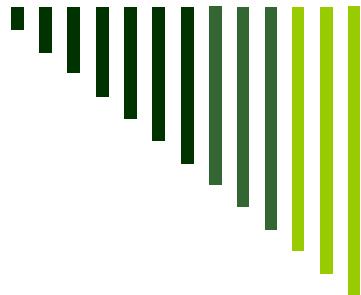
- El umbral óptimo se obtiene maximizando la suma de la entropía de ambos niveles de grises.

$$t = \operatorname{argmax}_t [entropia_b(t) + entropia_w(t)]$$



5.1.3. ISODATA

- Iterative SelfOrganizing Data Analysis Technique y Ay.
- Método de clasificación no supervisado.
- No es necesario saber el número de agrupaciones
- El algoritmo divide y fusiona agrupaciones
- El usuario define valores de umbral para parámetros
- La computadora ejecuta el algoritmo a través de muchos variaciones iteraciones hasta alcanzar el umbral



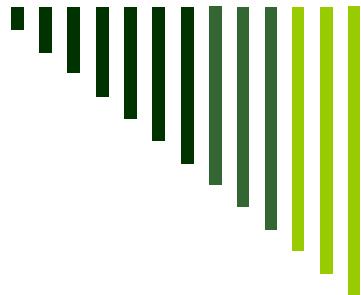
5.1.3. ISODATA (Cont.)

□ Funcionamiento:

- Calcular el histograma normalizado de la imagen (h_N)
- Multiplicar los valores del histograma normalizado por su peso (nivel de intensidad) se obtiene histograma ponderado (h_p).
- Calcular

$$u_j = \frac{\frac{\sum_{i=0}^t h_p(i)}{\sum_{i=0}^t h_N(i)} + \frac{\sum_{i=t+1}^L h_P(i)}{\sum_{i=t+1}^L h_N(i)}}{2}$$

Hasta que $u_j = u_{j-1}$, cuando esto sucede t será *el umbral óptimo* (*).



5.1.4. Método Local

- Usa una mascara de vecindad, preferentemente vecindades grandes.
- Ventaja: umbral adecuado a cada zona de la imagen.
- Desventaja: elevado tiempo de procesamiento.
- El proceso se realiza a todos los pixeles de la imagen
- Se puede aplicar con:
 - Usar el promedio es el paso más común
 - Aplicar otro método de binarización



GRACIAS..