

Procesamiento de Información Visual

Tema 1. Introducción

José Manuel Fuertes García
Manuel J. Lucena López

Departamento de Informática
Universidad de Jaén



4 de febrero de 2022

Índice

Conceptos básicos

Procesos de captura de información visual

Espacios transformados

Objetivos del Procesamiento de Información Visual

- ▶ Mejora de la información visual para ayudar en procesos de toma de decisión a intérpretes humanos.
- ▶ Procesamiento de la información visual de cara a su almacenamiento, transmisión y representación para el diseño de máquinas artificiales autónomas de visión.

¿Qué es el Procesamiento de Información Visual?

- ▶ Una imagen puede ser definida como una función bi-dimensional, $f(x, y)$, donde x e y son *coordenadas espaciales* y f es la amplitud de cualquier par de coordenadas (x, y) , también denominada *intensidad* o *nivel de gris* de la imagen para ese punto.
- ▶ Cuando los valores x e y son discretos y los de intensidad f están cuantizados, estamos en presencia de una *imagen digital*.
- ▶ El campo del *Procesamiento de Información Visual* hace referencia al procesamiento digital de dicha información realizado sobre un sistema artificial.
- ▶ Una imagen digital estará formada por un conjunto finito de elementos dispuestos en localizaciones espaciales específicas que almacenan un valor. A cada uno de esos elementos de la imagen se les denomina *píxel*.

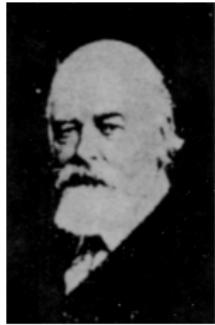
Los orígenes del Procesamiento Digital de Imagen

- ▶ Tal vez una de las primeras aplicaciones de las imágenes digitales fue en la industria periodística, cuando se iniciaron los envíos de imágenes a través de un cable submarino que unía Londres y Nueva York. El transporte sobre el Atlántico de las imágenes a través del cable suponía una demora de sólo 3 horas frente una semana. Equipos especializados de impresión codificaban las imágenes para su transmisión a través del cable y las reconstruían tras su recepción. La mala calidad visual de la información transmitida demandó técnicas que permitiesen su mejora de cara a los procesos de impresión finales.

Los orígenes del Procesamiento Digital de Imagen



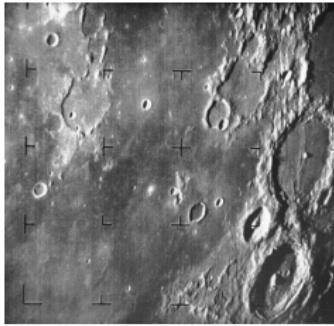
Ilustración digital de 1921 obtenida por una impresora telegráfica a partir de códigos de cinta



Ilustraciones digitales de 1922 y 1929 respectivamente, en las que ya se utilizan sistemas capaces de codificar 15 tonos de gris

Los orígenes del Procesamiento Digital de Imagen

- ▶ La introducción del concepto moderno de Computador Digital en 1940 por John von Neumann y los posteriores desarrollos del hardware y el software terminaron de hacerlos realidad en las dos décadas posteriores.
- ▶ A principios de 1960 ya existían ordenadores capaces de realizar tareas de procesamiento de las imágenes digitales que ya empezaban a ser una realidad.



Primera imagen de la luna obtenida por el *Ranger 7* el 31 de julio de 1964

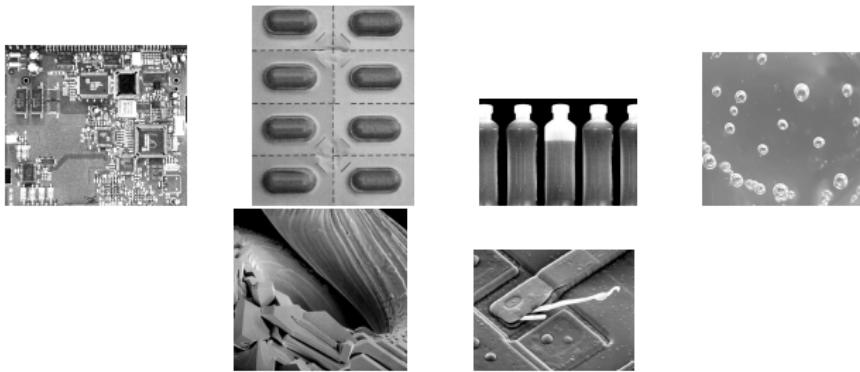
Campos de aplicación del Procesamiento Digital de Imagen

- ▶ Desde principio de los años 60 hasta nuestros días el campo de aplicación del procesamiento de imagen ha crecido vigorosamente.
- ▶ Campos como la medicina, biología, geología, astronomía, física, arqueología, defensa, procesos industriales..., son algunos ejemplos en los que dichas técnicas se están aplicando.
- ▶ Haciendo un esfuerzo de síntesis podríamos establecer dos grandes pilares de interés en el procesamiento de información visual:
 - ▶ Mejora de la información pictórica para ayudar en la toma de decisión de un experto humano.
 - ▶ Procesamiento de los datos de una escena para la percepción autónoma de la misma por parte de un sistema artificial.
- ▶ Mostraremos a continuación algunos de dichos campos.

Campos de aplicación del Procesamiento Digital de Imagen

Sistemas automáticos de inspección visual.

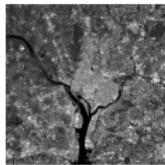
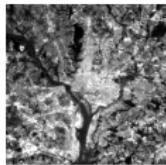
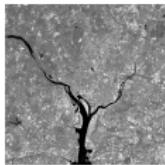
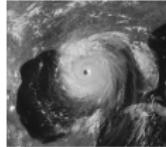
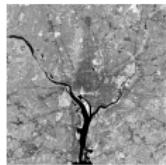
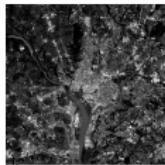
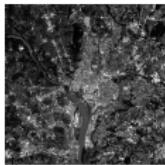
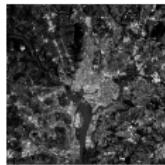
- ▶ Son sistemas básicos para el control de producción y calidad en la industria.
- ▶ Inspección automática de filamentos de lámparas incandescentes.
- ▶ Identificación de componentes defectuosos en sistemas electrónicos o electromecánicos.
- ▶ Sistemas de inspección automática de defectos en superficies.



Campos de aplicación del Procesamiento Digital de Imagen

Interpretación de escenas a partir de la teledetección.

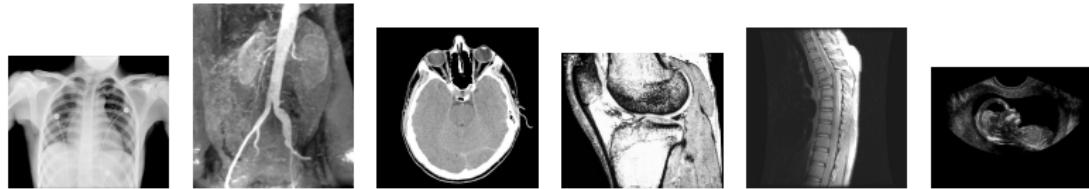
- ▶ Son sistemas básicos para análisis de recursos naturales en agricultura, hidrología, recursos geológicos, medio ambientales, etc
- ▶ La teledetección trabaja con imágenes multiespectrales capturadas de la tierra que son adquiridas por satélites o por dispositivos alojados en aviones.



Campos de aplicación del Procesamiento Digital de Imagen

Técnicas de análisis biomédico basadas en la imagen.

- ▶ En procesos de diagnóstico médico se emplean distintos tipo de imágenes: rayos X, imágenes de tomografía ayudada por ordenador (CT), ultrasonidos, etc.
- ▶ Sobre dichas imágenes se realizan procesos como: la localización de objetos, medidas reales sobre órganos, elementos de interés o tumores,etc.



Imágenes biomédicas de tipo rayos X, CT, MRI y ecografía.

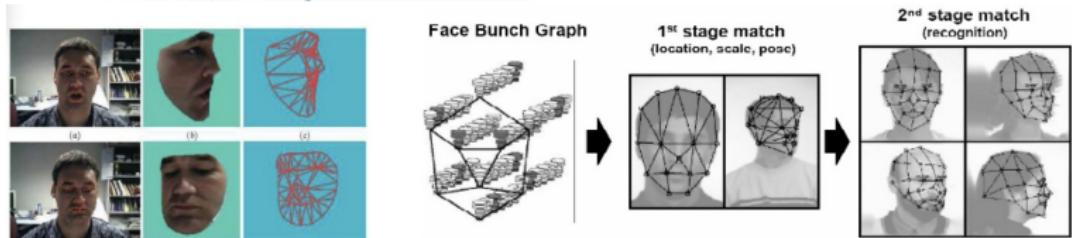
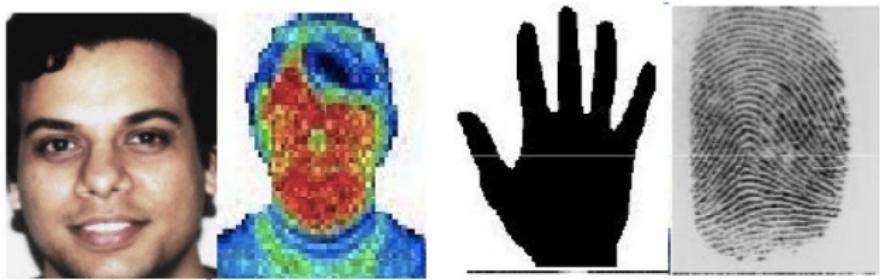
Campos de aplicación del Procesamiento Digital de Imagen

Sistemas de vigilancia inteligente.

- ▶ En los últimos años ha aumentado el interés en la protección de bienes y personas.
- ▶ Los sistemas de vigilancia tradicionales se basan en el uso de una vigilancia pasiva a partir de una monitorización pasiva.
- ▶ Los sistemas de vigilancia actuales requieren que esta ya no sea simple y reactiva, sino que sea inteligente y proactiva:
 - ▶ Control de accesos a áreas especiales.
 - ▶ Identificación de personas mediante características biométricas.
 - ▶ Identificación de acciones sospechosas.
 - ▶ Detección de situaciones de congestión de vehículos o personas.

Campos de aplicación del Procesamiento Digital de Imagen

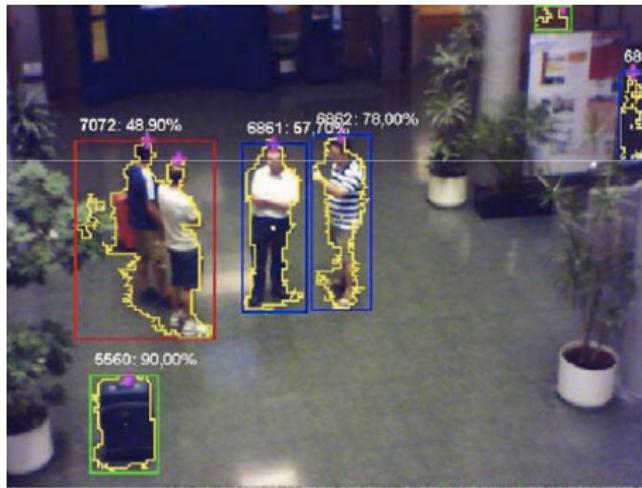
Sistemas de vigilancia inteligente.



Sistemas de identificación biométricos.

Campos de aplicación del Procesamiento Digital de Imagen

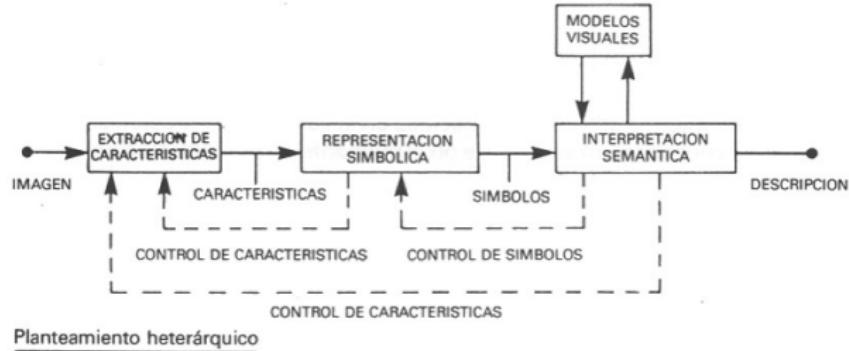
Sistemas de vigilancia inteligente.



Identificación de grupos, personas y objetos.

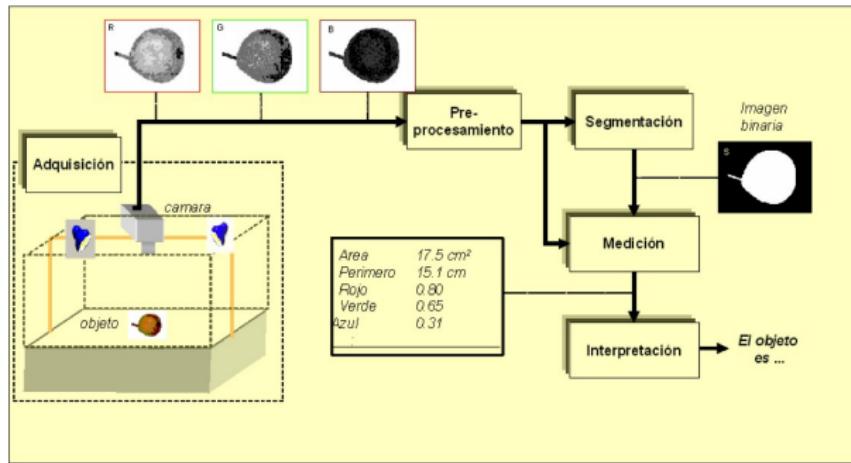
Componentes de un sistema de PDI

- ▶ Desde mediados de los años 80 se han propuesto numerosos modelos de sistemas para la realización de tareas de procesamiento digital de imagen.



Componentes de un sistema de PDI

- ▶ Desde mediados de los años 80 se han propuesto numerosos modelos de sistemas para la realización de tareas de procesamiento digital de imagen.



Esquema simplificado de bloques de un sistema general de interpretación de una escena.

Definiciones

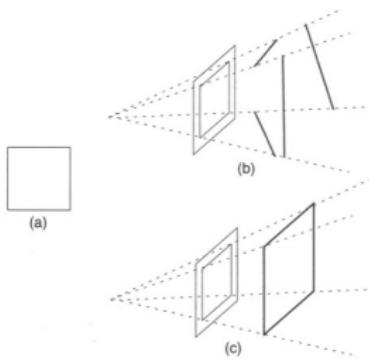
- ▶ La visión por ordenador puede definirse como la percepción por parte de un ordenador de información visual sensorial.
- ▶ *La visión es una tarea de proceso de información con bien definidas etapas de entrada y salida. La entrada consiste en una matriz de valores brillantes, que representan las proyecciones de una escena tridimensional registrada por una cámara o dispositivo similar de toma de imagen. Varias matrices de entrada proporcionan información en varias bandas espectrales (color) o desde múltiples puntos de vista (estéreo o secuencia temporal). La salida deseada es una descripción concisa de la escena tridimensional representada en la imagen, la naturaleza exacta de la cual depende según los objetivos y expectativas del observador. Generalmente implica una descripción de los objetos y sus interrelaciones, pero también puede incluir información tal como las estructuras tridimensionales de superficies, sus características físicas (forma, textura, color, materia y la ubicación de sombreados y fuentes de luz). (Barrow y Tenenbaum)*

Definiciones

- ▶ Se centra en la generación de descripciones de escenas tridimensionales a partir de la imágenes bidimensionales.
- ▶ *Ciencia que desarrolla las bases teóricas y algorítmicas para obtener información sobre el mundo real a partir de una o varias imágenes.* (HARALICK, 1992)
- ▶ *Disciplina que desarrolla sistemas capaces de interpretar el contenido de escenas naturales.* (CASTLEMAN, 1996)
- ▶ *Hacer que un computador vea. Esto, sin embargo, es todavía un problema no resuelto... Computer Vision ha emergido como una disciplina propia basada principalmente en las matemáticas y ciencias de la computación.* (FAUGERAS, 2001)
- ▶ *...la visión artificial o comprensión de imágenes, describe la deducción automática de la estructura y propiedades de un mundo tridimensional, posiblemente dinámico, bien a partir de una o varias imágenes bidimensionales de ese mundo.* (NALWA)

¿Hasta qué punto es difícil?

- ▶ El propio proceso de formación de la imagen no es más que la aplicación de muchas y diferentes escenas tridimensionales en una imagen bidimensional.
- ▶ **El problema inverso no tiene una solución única.**
- ▶ Barrow y Tenenbaum observan que al pasar de una escena a una imagen bidimensional, la imagen codifica mucha información sobre la escena, pero la información es confusa en el valor de brillo de cada uno de los puntos. Al proyectar sobre una imagen bidimensional, la información tridimensional se pierde.



(a) Imagen ambigua. (b) y (c) Dos escenas que darian lugar a la imagen

Índice

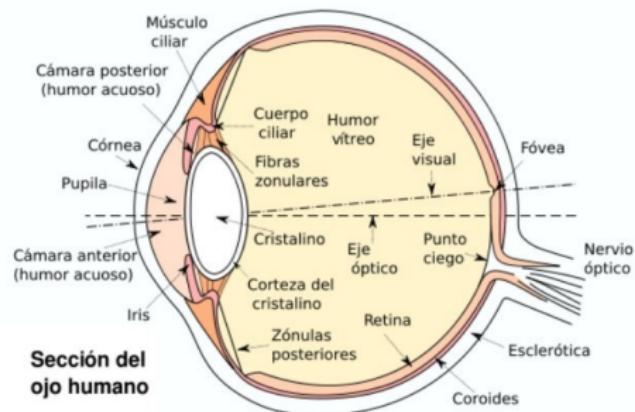
Conceptos básicos

Procesos de captura de información visual

Espacios transformados

Estructura del ojo humano

- ▶ Lo que se conoce del sistema visual humano más allá del propio ojo es principalmente disjunto, especulativo y escaso.
- ▶ La complejidad del sistema visual humano nos lleva a mostrar únicamente los mecanismos y parámetros que intervienen en el proceso de formación de la imagen en el humano frente a un sistema artificial.



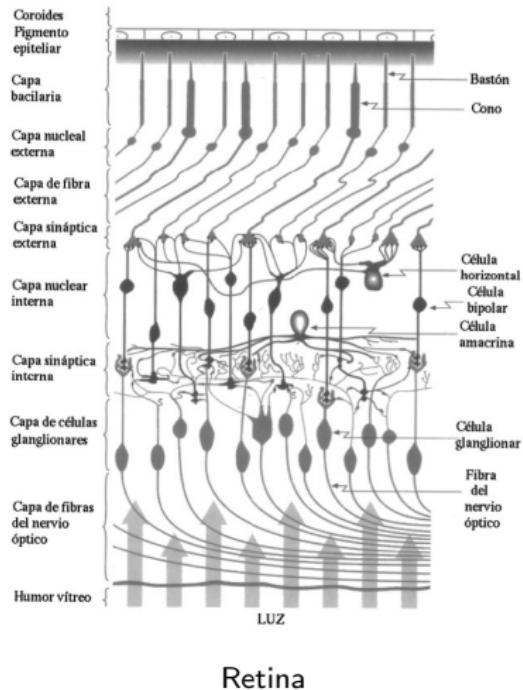
Esquema simplificado de una sección transversal del ojo humano.

Estructura del ojo humano

- ▶ Tiene un diámetro medio aproximado de 20mm.
- ▶ Está rodeado por tres membranas: la córnea y la esclerótica (cubierta exterior) y la coroides y la retina.
- ▶ La córnea es un tejido resistente y transparente que cubre la superficie anterior del ojo. Su prolongación, la esclerótica, encierra el resto del globo ocular.
- ▶ La coroides, situado entre la retina y la esclerótica, está muy pigmentada para reducir la cantidad de luz exterior que incide en el ojo.
- ▶ El iris se abre o cierra para controlar la cantidad de luz que entra en el ojo.
- ▶ El cristalino atenúa los niveles de radiación que inciden en el ojo, tanto del visible, como del ultravioleta e infrarrojos. Puede deformar su convexidad para formar la imagen en la superficie de la retina.

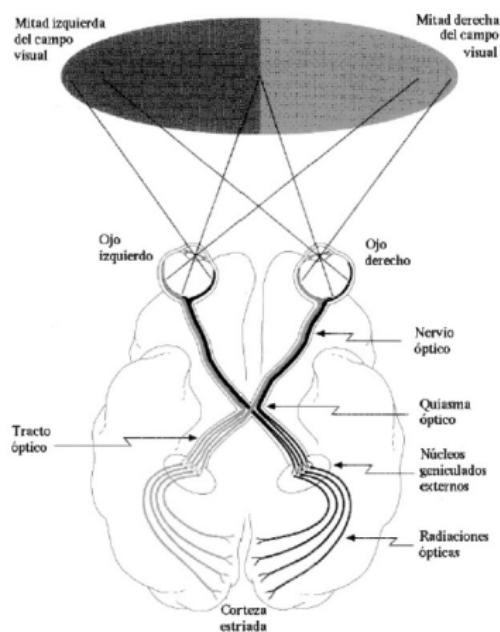
Estructura del ojo humano: La retina

- ▶ Es una membrana nerviosa compleja con un mosaico de fotorreceptores que cuando es estimulada por la luz, produce señales eléctricas nerviosas que evocan la experiencia de la visión.
- ▶ Existen dos clases de fotorreceptores:
 - ▶ Conos: Existen entre 6 y 7 millones en cada ojo. Localizados entorno a la fóvea son los responsables de la visión cromática, también llamada fotóptica.
 - ▶ Bastones: Existen entre 75 y 150 millones en cada ojo, distribuidos sobre la superficie retiniana. Son sensibles a niveles de iluminación bajos, siendo responsables de la visión escotóptica, por lo que no están implicados en la visión en color.



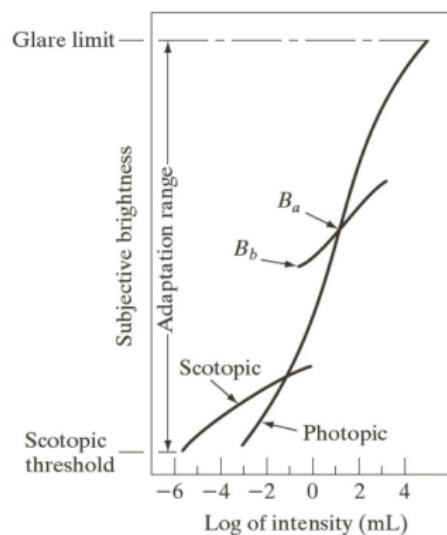
Del ojo al cerebro

- ▶ La conversión de las imágenes retinianas en conocimiento del mundo, conforma la auténtica barrera para nuestro entendimiento de la visión humana.
- ▶ Conocemos los caminos visuales del ojo al cerebro, pero sabemos poco de lo que ocurre en él.
- ▶ El campo visual izquierdo va al lado derecho del cerebro y el derecho al izquierdo.
- ▶ La disparidad entre las dos imágenes retinianas es la que hace posible la percepción de la profundidad estereoscópica dentro del campo de visión que comparten los dos ojos.



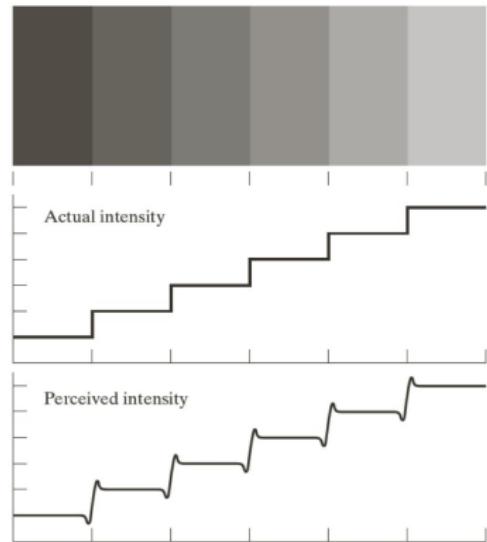
Adaptación a la iluminación

- ▶ El sistema visual humano es capaz de percibir un rango de niveles de intensidad de luz enorme (10^{10})
- ▶ La iluminación subjetiva, la percibida por el sistema visual humano, es una función logarítmica de la intensidad de luz incidente en el ojo.
- ▶ El sistema visual humano no puede percibir simultáneamente todo ese rango de intensidades que van desde el umbral escotópico hasta el límite de deslumbramiento.
- ▶ Cambios en la sensibilidad global permiten la adaptación a la iluminación.
- ▶ Un determinado nivel de adaptación a la iluminación permite percibir de forma correcta los distintos niveles de intensidad en ese rango (B_a y B_b).



Adaptación a la iluminación

- ▶ Hay fenómenos que ponen de manifiesto que la iluminación percibida no es función de la intensidad.
- ▶ El sistema visual tiende a sobrevalorar o infravalorar la intensidad cerca de los límites de dos regiones con intensidades diferentes.
- ▶ Aunque la intensidad de las franjas es constante, realmente se percibe un patrón de iluminación fuertemente escalonado, sobre todo cerca de los bordes entre franjas.
- ▶ Se denomina patrón de bandas de Mach.



Adaptación a la iluminación

- ▶ Existe otro fenómeno que se denomina contraste simultáneo.
- ▶ Está relacionado con el hecho de que la iluminación percibida de un área no depende únicamente de su intensidad.
- ▶ Todos los cuadrados tienen exactamente la misma intensidad; sin embargo aparecen a la vista como más oscuros conforme el fondo se hace más claro.

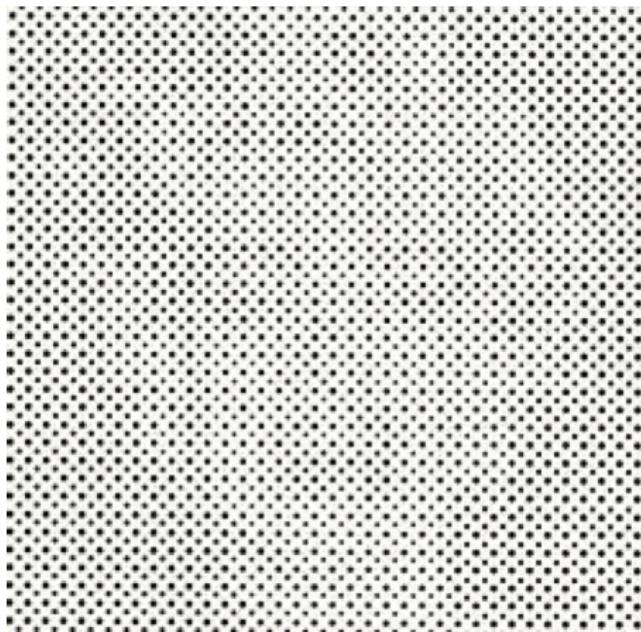


Ilusiones, Ambigüedades e Inconsistencias

- ▶ El sistema visual humano no es tan exacto y ajeno a errores como en principio podríamos pensar.
- ▶ Dichos errores podrían ser clasificados en dos grandes familias:
 - ▶ *Errores de captación de la información.* Debidos principalmente a problemas o limitaciones en el sistema de adquisición de las imágenes.
 - ▶ *Errores de interpretación de la información visual.* Debidos a una mala interpretación de la física o la geometría de los objetos o relaciones existentes entre los mismos:
 - ▶ Ilusiones.
 - ▶ Ambigüedades.
 - ▶ Inconsistencias.

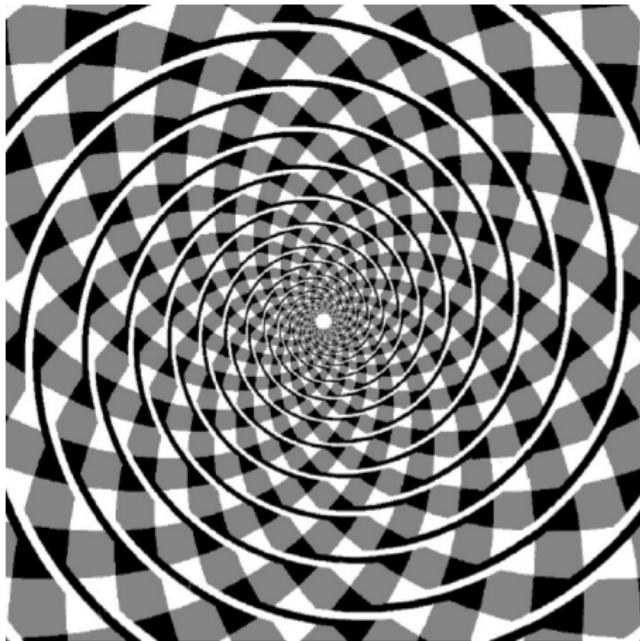
Ilusiones, Ambigüedades e Inconsistencias

Ilusiones.



Ilusiones, Ambigüedades e Inconsistencias

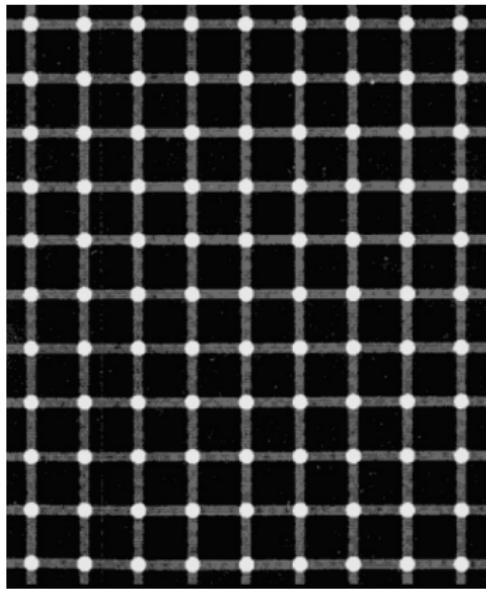
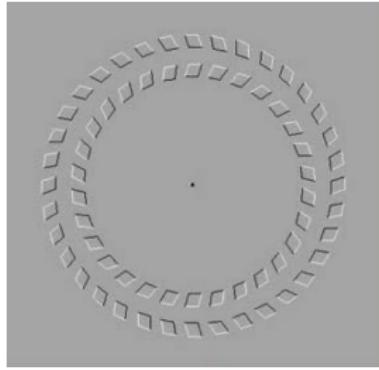
Ilusiones.



Espiral de Fraser.

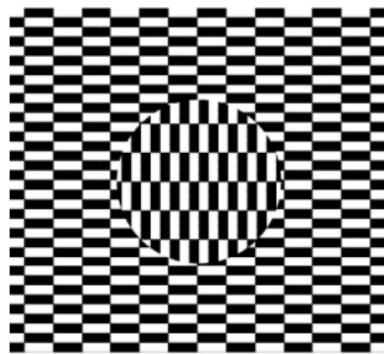
Ilusiones, Ambigüedades e Inconsistencias

Ilusiones.

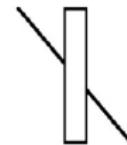


Ilusiones, Ambigüedades e Inconsistencias

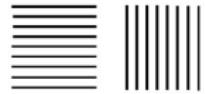
Ilusiones.



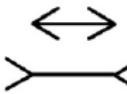
Ilusión de Zöllner (1860)



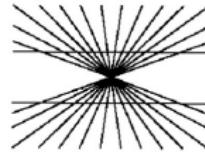
Ilusión de Poggendorff (1860)



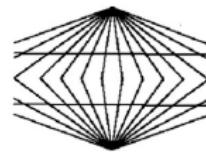
Cuadrados de Helmholtz (1866)



Ilusión de Müller-Lyer (1889)



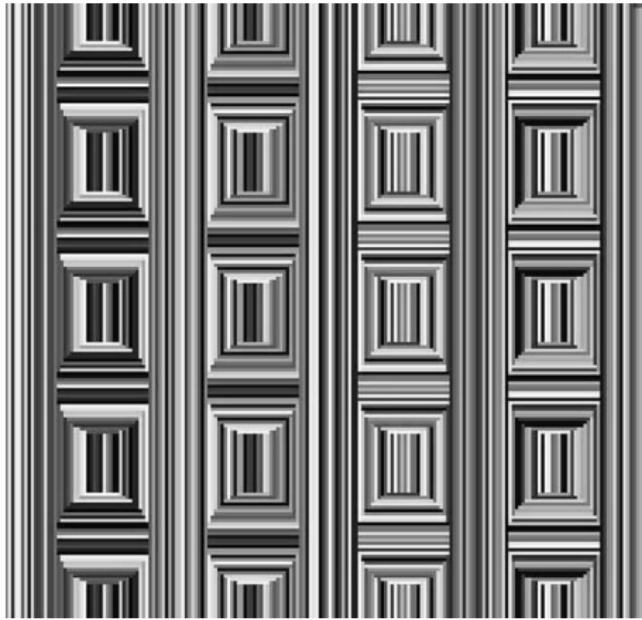
Ilusión de Hering (1861)



Ilusión de Wundt (1896)

Ilusiones, Ambigüedades e Inconsistencias

Ambigüedades.



Ilusiones, Ambigüedades e Inconsistencias

Ambigüedades.



Ilusiones, Ambigüedades e Inconsistencias

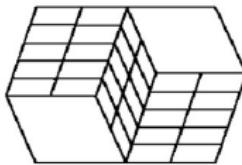
Ambigüedades.



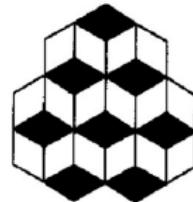
Mujer joven / Anciana



Jarrón / Caras



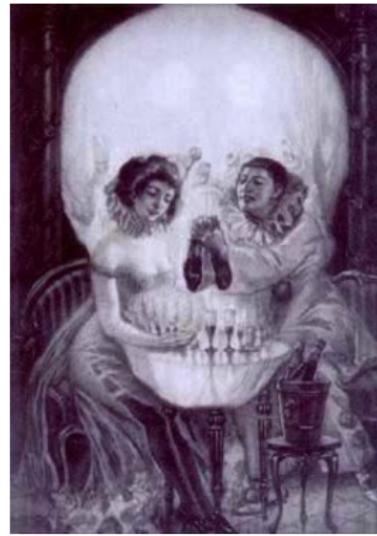
Hacia dentro / Hacia fuera



Seis cubos / Siete cubos

Ilusiones, Ambigüedades e Inconsistencias

Ambigüedades.



Ilusiones, Ambigüedades e Inconsistencias

Ambigüedades.



Ilusiones, Ambigüedades e Inconsistencias

Ambigüedades.



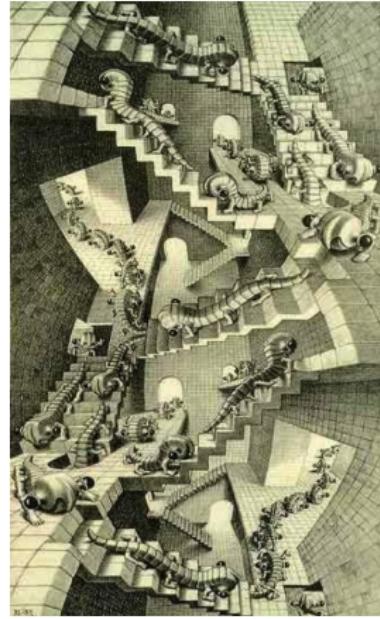
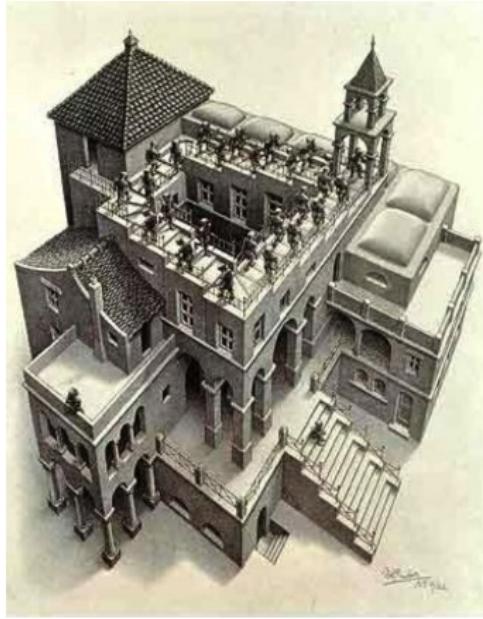
Ilusiones, Ambigüedades e Inconsistencias

Inconsistencias.



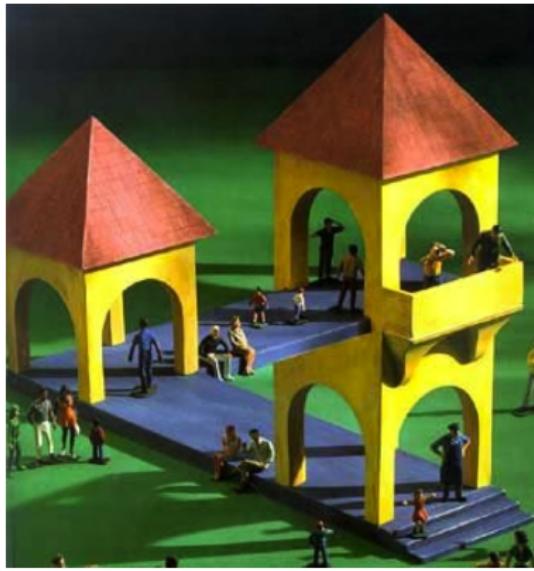
Ilusiones, Ambigüedades e Inconsistencias

Inconsistencias.



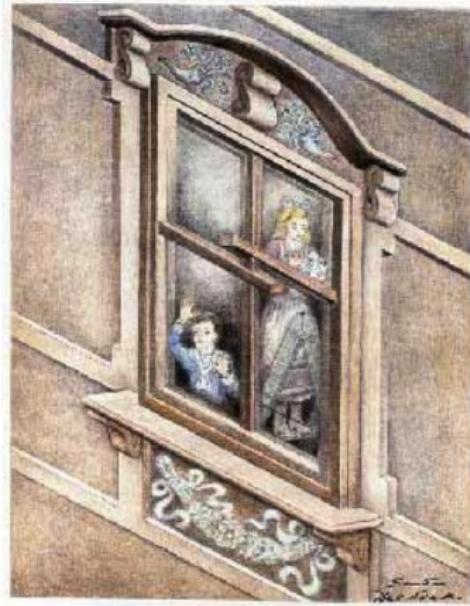
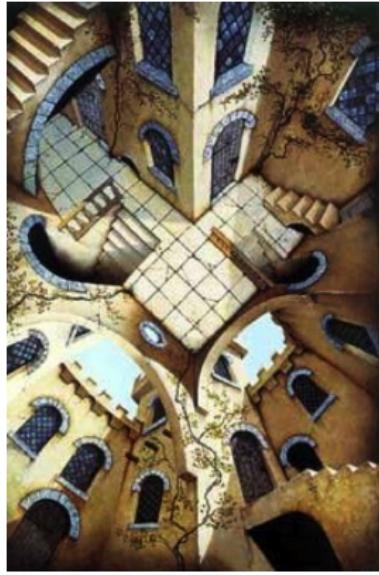
Ilusiones, Ambigüedades e Inconsistencias

Inconsistencias.



Ilusiones, Ambigüedades e Inconsistencias

Inconsistencias.

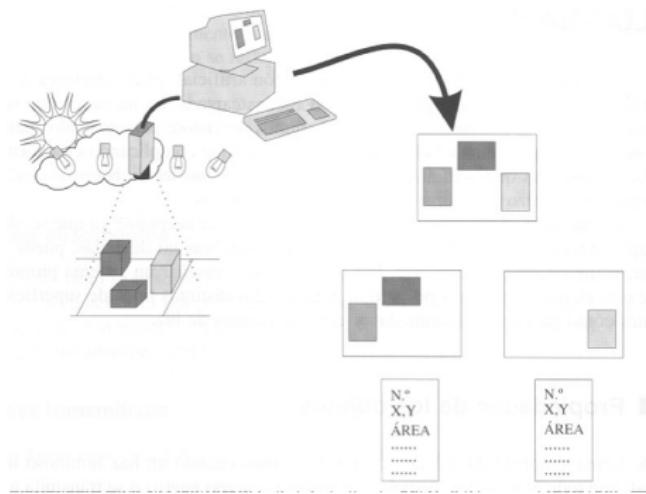


Ilusiones, Ambigüedades e Inconsistencias

- ▶ Las diferentes ilusiones, ambigüedades e inconsistencias nos llevan a preguntarnos si inferimos de nuestras imágenes en la retina más de lo que de verdad está soportado por la geometría y la física de la formación de imágenes.
- ▶ Helmholtz expresó en su *Handbook of Physiological Optics*:
 - ▶ Cada imagen es una imagen de algo sólo para aquel que conoce cómo leerla y que está capacitado, con la ayuda de la imagen, a formar una idea de la cosa.

Elementos de un sistema de percepción visual artificial

- ▶ El desarrollo de un sistema de percepción visual artificial, requiere de un amplio conjunto de elementos.

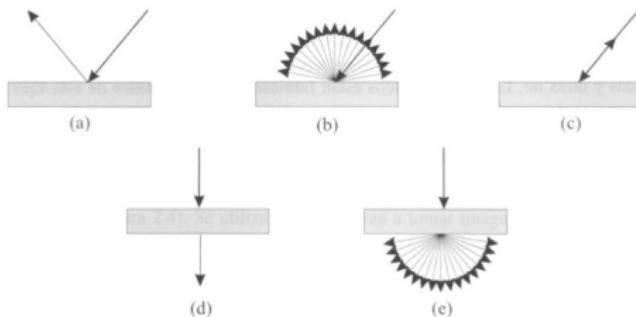


La iluminación

- ▶ La iluminación juega un papel vital al simplificar de manera considerable el análisis y posterior interpretación de la escena capturada.
- ▶ Afecta de forma considerable a la complejidad del algoritmo: Un problema se resuelve de formas más rápida cambiando la iluminación que complicando el algoritmo.
- ▶ La iluminación en un entorno dado no suele ser aceptable ya que genera imágenes con bajo contraste (poco o demasiado iluminadas), sombras no deseadas, etc.
- ▶ Un diseño propio de iluminación puede fijar y controlar todos esos fenómenos no deseados.
- ▶ Estudiar los distintos tipos de superficies nos puede dar las claves de cómo pueden ser iluminadas y con qué fuentes de luz.

La iluminación

- ▶ **Tipos de superficie:** Cuando un haz luminoso incide sobre un material puede ocurrir que:
 - ▶ Todo él se refleje (espejo).
 - ▶ Se absorba (cuerpo negro).
 - ▶ Se transmita a través de él.

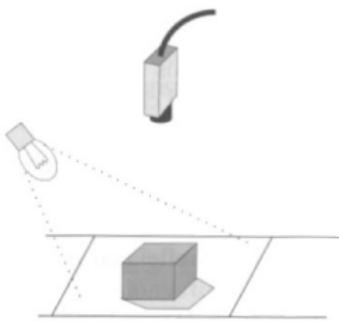


Diversas propiedades de los materiales: (a) especulares; (b) difusos; (c) reflectores;
(d) transparentes; (e) translúcidos.

La iluminación

► Tipos de iluminación: Direccional.

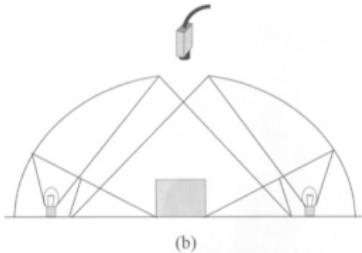
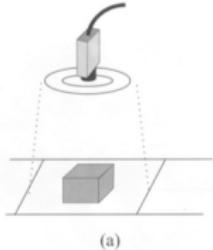
- Consiste en aplicar una iluminación orientada al objeto usando un haz altamente direccional.
- Se utiliza en aplicaciones como localización y reconocimiento de piezas o inspección de superficies de objetos.
- La forma del haz luminoso depende de la aplicación y del objeto a analizar.
- Se utiliza cuando es necesario iluminar un objeto para percibirlo o para aprovechar la formación de sombras que resaltan los defectos en los materiales.



La iluminación

► Tipos de iluminación: Difusa.

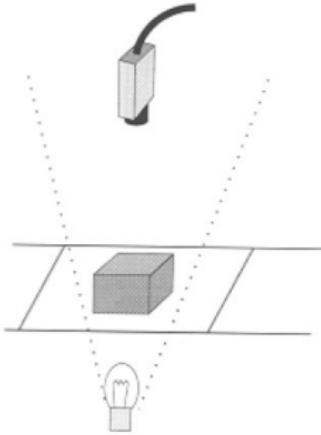
- ▶ Este tipo de iluminación intenta que los haces luminosos incidan sobre el objeto desde todas las direcciones.
- ▶ Se utiliza cuando se van a tomar imágenes a través de microscopios, los objetos presentan superficies suaves y regulares, o se necesitan imágenes desde diversos puntos de vista, por lo que no se pueden favorecer a unas zonas más que a otras.
- ▶ Hay que tener en cuenta que proporciona el mínimo contraste del objeto, debido a la ausencia de sombras.



La iluminación

► Tipos de iluminación: Contraluz.

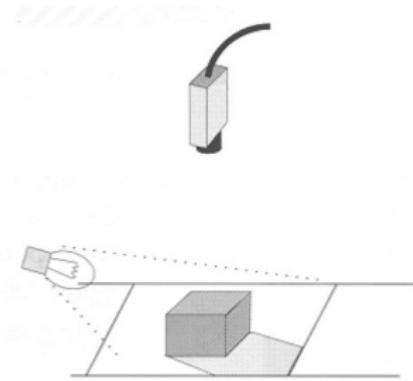
- ▶ Consiste en iluminar el objeto por detrás de forma que la fuente luminosa, el objeto y la cámara estén alineados.
- ▶ Produce imágenes prácticamente de sólo dos niveles de gris.



La iluminación

► **Tipos de iluminación:** Oblicua.

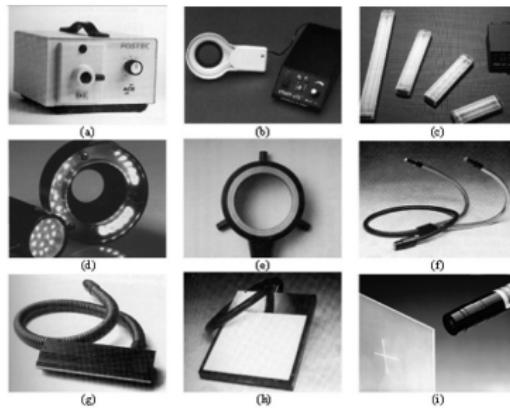
- Puede considerarse un caso particular de la iluminación direccional.
- Consiste en iluminar el objeto de tal forma que se generan sombras que aumentan el contraste de las partes tridimensionales.



La iluminación

► Fuentes de luz:

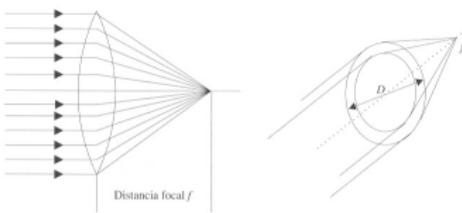
- Una vez elegida la forma en que se quiere que la fuente de luz llegue a los objetos, hay que definir qué fuente de luz va a producirla.



Incandescentes, fluorescentes circulares, lineales, led, fibra óptica y láser.

La óptica

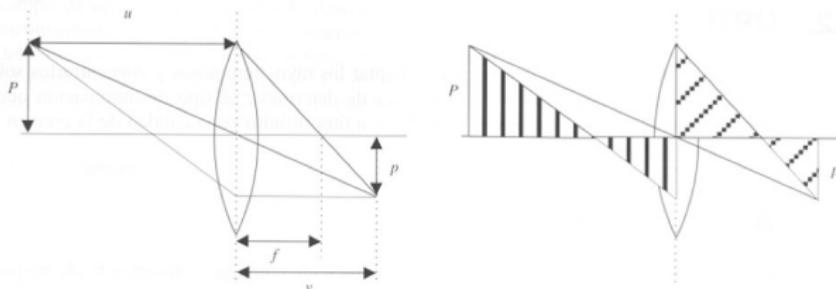
- ▶ La función de la óptica de una cámara es captar los rayos luminosos y concentrarlos sobre el sensor del sistema de adquisición.
- ▶ **Principales parámetros:**
 - ▶ Cuando los rayos paralelos pasan a través de una lente convexa, convergen hacia un punto que se denomina *punto focal*.
 - ▶ La distancia entre el eje de la lente y dicho punto se llama *distancia focal*.
 - ▶ La distancia focal es el parámetro principal a la hora de calcular la posición y el tamaño de los objetos en la imagen.
 - ▶ El número F es un parámetro que indica la relación entre la distancia focal y el diámetro del diafragma. $F = \frac{f}{D}$



La óptica

► Modelo de lente fina

- Un modelo de lente fina es aquel que establece que todo rayo que entra paralelo al eje óptico pasa por el foco posterior de la lente y todo rayo que pasa por el foco anterior sale de la lente paralelo al eje óptico.



- Mediante semejanza de triángulos obtenemos estas relaciones:

$$\frac{P}{u-f} = \frac{p}{f}$$

$$\frac{p}{v-f} = \frac{P}{f}$$

$$(u-f)(v-f) = f^2$$

$$uv - f(u+v) = 0$$

$$\frac{uv}{u+v} = f$$

$$\frac{uv}{u+v} = \frac{1}{f}$$

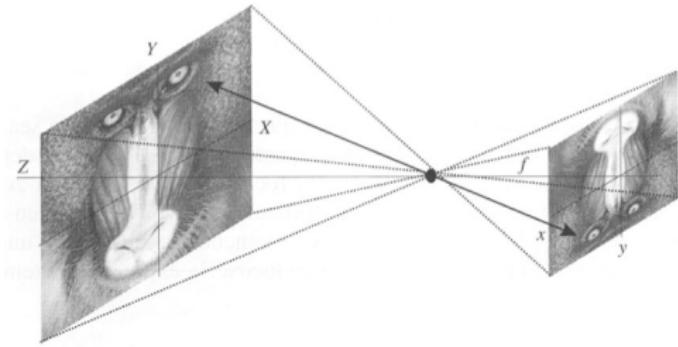
$$\frac{1}{u} + \frac{1}{v} = \frac{1}{f}$$

donde u es la distancia del objeto a la lente, v es la distancia del plano de enfoque a la lente y f es la distancia focal.

La óptica

► Modelo Pin-Hole

- ▶ Se trata de un modelo óptico muy simple.
- ▶ Este modelo reduce la óptica a un punto situado a la distancia focal de la imagen, por lo que de todos los rayos luminosos que refleja un punto perteneciente a un objeto, solamente es importante el que pasa directamente por la distancia focal.
- ▶ El modelo *pin-hole* supone que todos los puntos están enfocados y no tiene en cuenta ninguna de las aberraciones asociadas a la óptica.
- ▶ Mediante semejanza de triángulos obtenemos las relaciones entre un punto del espacio y su proyección en la imagen.



$$x = \frac{f}{Z} X$$
$$y = \frac{f}{Z} Y$$

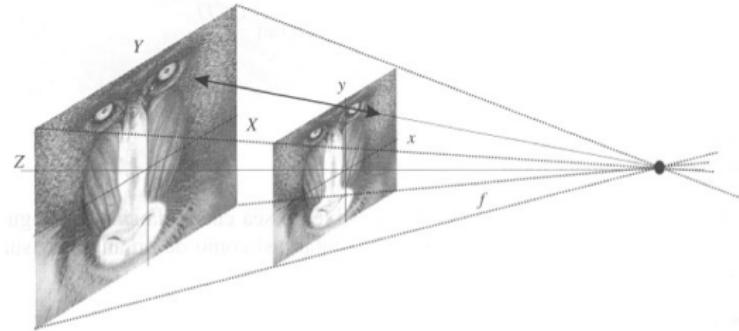
La óptica

► Modelo Pin-Hole

- ▶ El *ángulo visual* es un parámetro que relaciona la distancia focal y el tamaño del sensor.
- ▶ Se define como el ángulo que forman dos rayos que inciden sobre los bordes externos del sensor de la imagen cuando esta se encuentra enfocada a una distancia infinita.
- ▶ La expresión del *ángulo visual* está dado por:

$$w = 2 \cdot \arctan \left(\frac{CCD}{2f} \right)$$

donde *CCD* es el tamaño del sensor y *f* es la distancia focal.



Planos, Puntos y Líneas Rectas

- ▶ Un punto en un plano se representa como un par de coordenadas (x, y) en dicho plano \mathbb{R}^2
- ▶ Se puede considerar a \mathbb{R}^2 como un espacio vectorial en el que (x, y) es un vector.
- ▶ Podemos por tanto asociar un punto a un vector.
- ▶ Una línea recta se representa por la ecuación $ax + by + c = 0$
- ▶ Una línea recta puede ser representada por un vector $l = (a, b, c)^T$
- ▶ La correspondencia entre líneas y rectas no es uno a uno, ya que $(a, b, c)^T$ y $(ka, kb, kc)^T$ representan exactamente la misma línea recta para $k \neq 0$, sin embargo son vectores distintos.
- ▶ Estos vectores son considerados como equivalentes y se definen como *vectores homogéneos*.

Planos, Puntos y Líneas Rectas

- ▶ Un punto (x, y) se representa como por un vector $(x, y, 1)^T$
- ▶ Los vectores (kx, ky, k) , para $k \neq 0$, pueden considerarse como representantes del mismo punto (x, y) ya que satisfacen la ecuación $[a \ b \ c]^T [kx \ ky \ k] = k(ax + by + c) = 0$
- ▶ Se dice que los puntos en un plano bidimensional pueden ser representados por vectores homogéneos de tres dimensionales cuyos dos primeros elementos son las coordenadas del punto en el plano y el tercer elemento es 1.
- ▶ Si se tiene un vector homogéneo de tres dimensiones dado por $(x_1, x_2, x_3)^T$ que representa un punto en un plano, las coordenadas de ese punto en el plano están definidas como $(x, y) = (x_1/x_3, x_2/x_3)$.

Transformaciones básicas en un espacio 3D

- ▶ Las coordenadas homogéneas de un punto con coordenadas cartesianas $w = (X, Y, Z)^T$ se definen como $w_h = (kX, kY, k)^T$, donde k es una constante arbitraria distinta de 0.
- ▶ La conversión de coordenadas homogéneas a cartesianas se realiza fácilmente dividiendo las tres primeras coordenadas homogéneas por la cuarta.

Transformaciones básicas en un espacio 3D

► Traslación

- La traslación de un punto $w = (X, Y, Z)^T$, utilizando coordenadas homogéneas, viene dado por el producto matricial:

$$\begin{bmatrix} X' \\ Y' \\ Z' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & X_0 \\ 0 & 1 & 0 & Y_0 \\ 0 & 0 & 1 & Z_0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

donde $(X_0, Y_0, Z_0)^T$ es el vector de traslación.

Transformaciones básicas en un espacio 3D

► Escalado

- El escalado con los factores S_X, S_Y, S_Z en los ejes X, Y y Z , viene dado por la transformación:

$$\begin{bmatrix} X' \\ Y' \\ Z' \\ 1 \end{bmatrix} = \begin{bmatrix} S_X & 0 & 0 & 0 \\ 0 & S_Y & 0 & 0 \\ 0 & 0 & S_Z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

- Si se pretende realizar un escalado uniforme (los tres ejes se escalan por el mismo factor) se hace $S_X = S_Y = S_Z$ o bien se puede utilizar la matriz de transformación:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1/S \end{bmatrix}$$

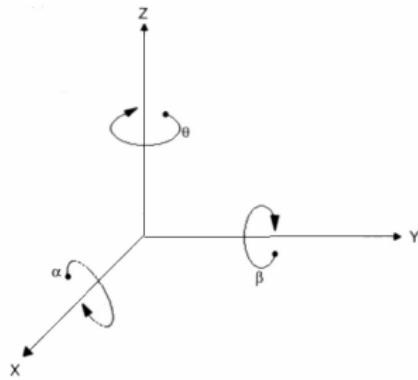
donde S es el factor de escala elegido.

Transformaciones básicas en un espacio 3D

► Rotación

- La rotación de un punto respecto a cada uno de los ejes de coordenadas viene dada por las matrices homogéneas:

$$R_\theta = \begin{bmatrix} \cos\theta & \sin\theta & 0 & 0 \\ -\sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad R_\alpha = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\alpha & \sin\alpha & 0 \\ 0 & -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad R_\beta = \begin{bmatrix} \cos\beta & 0 & \sin\beta & 0 \\ 0 & 1 & 0 & 0 \\ -\sin\beta & 0 & \cos\beta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$



Transformaciones básicas en un espacio 3D

► Concatenación

- ▶ La aplicación de varias transformaciones puede ser representada por una única matriz de transformación 4x4.
- ▶ Por ejemplo, la traslación, escalado y rotación con respecto del eje Z de un punto V viene dada por la composición matricial:

$$(R_\theta \cdot S \cdot T) \cdot V = A \cdot V$$

donde A es la matriz 4x4 obtenida mediante el producto $A = R_\theta \cdot S \cdot T$. Puesto que la multiplicación de matrices no es una operación conmutativa, el orden de las matrices es importante.

- ▶ La siguiente expresión representa la estructura de una transformación, resultado de la concatenación de matrices de transformación.

$$A = \begin{bmatrix} R & T \\ 0 & S \end{bmatrix}$$

donde R es de dimensión 3x3 y contiene los elementos que realizan la rotación, T es de dimensión 3x1 y contiene los elementos que realizan la traslación y S, de dimensión 1x1, realiza el escalado uniforme del vector al que se le aplica A.

Transformaciones Proyectivas 2D

- ▶ La geometría proyectiva 2D es el estudio de las propiedades del plano proyectivo \mathbb{P}^2 que son invariantes bajo un grupo de transformaciones conocidas como proyectividades.
- ▶ Una proyectividad es una transformación invertible dada por $h : \mathbb{P}^2 \rightarrow \mathbb{P}^2$ de manera tal que una línea recta es transformada como una línea recta. La proyectividad está definida como:

$$h(m) = m' = Hm$$

donde H es una matriz 3×3 no singular.

- ▶ Se dice que m' es la transformación lineal H de m .
- ▶ Esta transformación es biunívoca entre dos planos 2D, cuyos puntos son representados homogéneamente por m y m' . Es decir, un punto en un plano 2D tiene una única correspondencia en un punto de otro plano 2D, y cada punto en un plano tiene un sólo punto correspondiente en el otro plano.

Transformaciones Proyectivas 2D

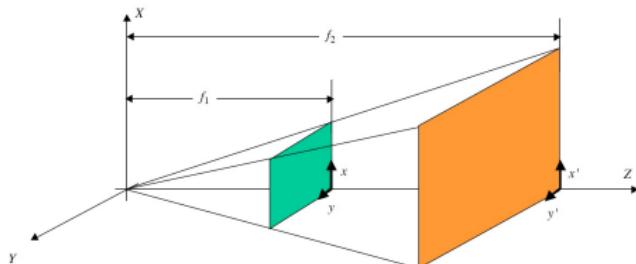
- ▶ La ecuación puede escribirse de manera explícita como:

$$\begin{bmatrix} x'_1 \\ x'_2 \\ x'_3 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

- ▶ Se observa que los resultados en la transformación de coordenadas no son afectados si se cambia H por kH , para $k \neq 0$, lo cual quiere decir que H es una matriz homogénea.

Transformaciones Proyectivas 2D

► Cálculo de la matriz proyectiva entre planos paralelos



- Utilizando el Teorema de Thales, se obtienen las siguientes relaciones:

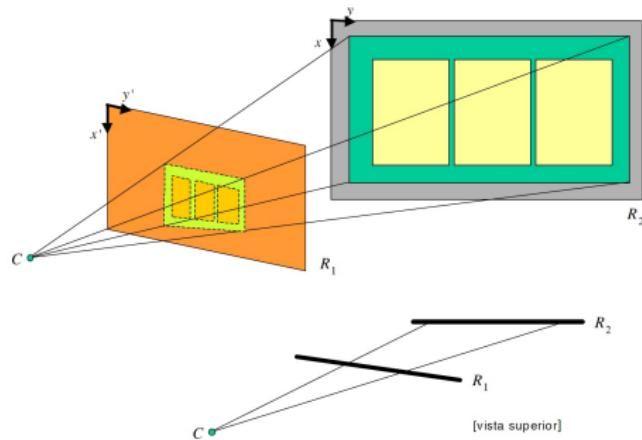
$$\frac{x'}{f_2} = \frac{x}{f_1} \quad ; \quad \frac{y'}{f_2} = \frac{y}{f_1}$$

- Estas ecuaciones se pueden escribir usando la forma matricial $m' = Hm$, con $m = [x \ y \ 1]^T$, $m' = [x' \ y' \ 1]^T$ y

$$H = \begin{bmatrix} f_2/f_1 & 0 & 0 \\ 0 & f_2/f_1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Transformaciones Proyectivas 2D

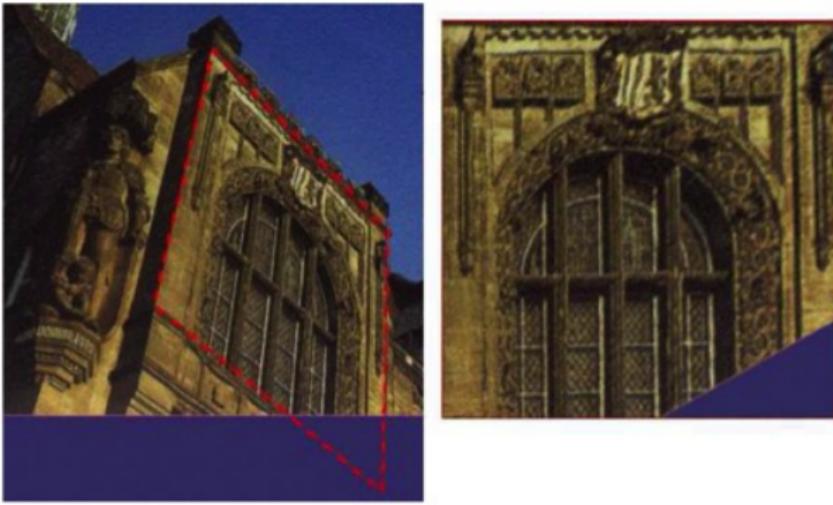
► Cálculo de la matriz proyectiva entre planos no paralelos



- En este caso se busca la matriz H que logre transformar una imagen que presente una distorsión proyectiva, en una imagen libre de distorsión.

Transformaciones Proyectivas 2D

- ▶ **Cálculo de la matriz proyectiva entre planos no paralelos**
 - ▶ Ejemplo de rectificación de la distorsión proyectiva presente en la fotografía de la fachada de una iglesia.



Transformaciones Proyectivas 2D

► Cálculo de la matriz proyectiva entre planos no paralelos

- ▶ Como datos de entrada para el problema de rectificación tomaremos las coordenadas de n puntos (x', y') en la imagen original y las coordenadas deseadas de esos puntos (x, y) en la imagen rectificada.
- ▶ Se obtiene para cada punto (x', y') y su correspondiente (x, y) las ecuaciones:

$$x' = \frac{x'_1}{x'_3} = \frac{h_{11}x + h_{12}y + h_{13}}{h_{31}x + h_{32}y + h_{33}}$$

$$y' = \frac{x'_2}{x'_3} = \frac{h_{21}x + h_{22}y + h_{23}}{h_{31}x + h_{32}y + h_{33}}$$

Siendo (x'_1, x'_2, x'_3) las coordenadas homogéneas para el punto (x', y') .

- ▶ En el problema interesa encontrar los 9 elementos de H . Sin embargo, como la matriz H es homogénea, kH también sería solución al problema.
- ▶ De esa manera es posible dividir cada elemento de H por h_{33} para obtener una matriz H con sólo 8 elementos desconocidos, ya que el último elemento sería igual a uno.

Transformaciones Proyectivas 2D

► Cálculo de la matriz proyectiva entre planos no paralelos

- Con $h_{33} = 1$ las ecuaciones anteriores pueden ser escritas de manera matricial $Ah = b$, como:

$$\begin{bmatrix} x & y & 1 & 0 & 0 & 0 & -x'x & -x'y \\ 0 & 0 & 0 & x & y & 1 & -y'x & -y'y \end{bmatrix} \begin{bmatrix} h_{11} \\ h_{12} \\ h_{13} \\ h_{21} \\ h_{22} \\ h_{23} \\ h_{31} \\ h_{32} \end{bmatrix} = \begin{bmatrix} x' \\ y' \end{bmatrix}$$

Se observa que para cada correspondencia de puntos se obtienen dos ecuaciones.

Transformaciones Proyectivas 2D

► Cálculo de la matriz proyectiva entre planos no paralelos

- Suponiendo n pares de puntos correspondientes se puede establecer el siguiente sistema de $2n$ ecuaciones y 8 incógnitas:

$$\begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{bmatrix} h = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

- El sistema de ecuaciones puede ser expresado com $Ah = b$.
- Para $n = 4$ existe una solución directa dada por $h = A^{-1}b$.
- Si $n > 4$ el sistema queda sobredimensionado. En ese caso se utiliza el método de los mínimos cuadrados en el que se encuentra un vector h que minimice $\|Ah - b\|$.
- La solución entonces estaría dada por $h = [A^T A]^{-1} A^T b$.

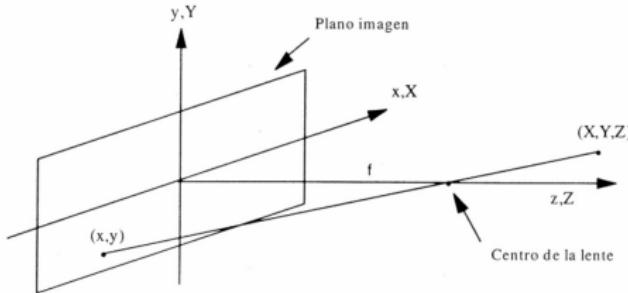
Transformaciones Proyectivas 2D

| | | | | | |
|---|--|---|---|---|--|
|  | |  |  |  |  |
| original | <i>Euclídea</i> | <i>similitud</i> | <i>afín</i> | <i>general</i> | |
| Transformación | Matriz \mathbf{H} | Invariantes | | | |
| Euclídea | $\begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix}$ | longitud entre puntos. | | | |
| Similitud | $\begin{bmatrix} s\mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix}$ | ángulos entre rectas, razón entre dos distancias. | | | |
| Afín | $\begin{bmatrix} \mathbf{A} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix}$ | líneas paralelas, razón entre dos áreas. | | | |
| General | $\begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}$ | razón de cruz. | | | |

Transformaciones Proyectivas 3D

► Sistema de coordenadas coincidente

- Una transformación perspectiva proyecta puntos del espacio tridimensionales en el plano.
- Este modelo de proyección no distorsiona los objetos en la proyección, sino que éstos se proyectan invertidos y escalados de acuerdo con un factor dado por la distancia focal.



$$\frac{x}{f} = -\frac{X}{Z-f} \quad x = \frac{fX}{f-Z}$$
$$\frac{y}{f} = -\frac{Y}{Z-f} \quad y = \frac{fY}{f-Z}$$

Transformaciones Proyectivas 3D

► Sistema de coordenadas coincidente

- Una transformación perspectiva se puede expresar como:

$$c_h = Pw_h = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{f} & 1 \end{bmatrix} \begin{bmatrix} kX \\ kY \\ kZ \\ k \end{bmatrix} = \begin{bmatrix} kX \\ kY \\ kZ \\ -\frac{kZ}{f} + k \end{bmatrix}$$

donde las coordenadas de c_h son las del punto proyectado en su forma homogénea.

- La matriz P se denomina *matriz de transformación de perspectiva*.

Transformaciones Proyectivas 3D

► Sistema de coordenadas coincidente

- Para obtener el valor c de coordenadas cartesianas se dividen las tres primeras por la cuarta:

$$c = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \frac{fx}{f-z} \\ \frac{fy}{f-z} \\ \frac{fz}{f-z} \end{bmatrix}$$

- Las dos primeras componentes de c son las coordenadas (x, y) en el plano imagen del punto tridimensional (X, Y, Z) proyectado.
- La tercera componente carece de significado, por lo que en muchas aplicaciones se elimina la tercera fila de la matriz de transformación de perspectiva P .

Transformaciones Proyectivas 3D

► Sistema de coordenadas coincidente

- ▶ Puesto que la transformación definida es de *varios a uno*, la transformación inversa no está totalmente definida, es decir, dado un punto proyectado es imposible conocer el punto tridimensional que lo generó, siendo tan sólo posible conocer la recta de proyección.
- ▶ Un punto del plano imagen (x_0, y_0) puede ser la proyección de cualquier punto tridimensional que se encuentre en la recta que pasa por $(x_0, y_0, 0)$ y $(0, 0, f)$, y que viene dada por la intersección de los planos:

$$X = \frac{x_0}{f}(f - Z)$$

$$Y = \frac{y_0}{f}(f - Z)$$

Transformaciones Proyectivas 3D

► Sistema de coordenadas coincidente

- Las anteriores expresiones también pueden obtenerse a partir de la inversa de la matriz de transformación de perspectiva:

$$P^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1/f & 1 \end{bmatrix}$$

Suponiendo:

$$c_h = \begin{bmatrix} kx_0 \\ ky_0 \\ kz \\ k \end{bmatrix}$$

donde (x_0, y_0) son las coordenadas del punto proyectado y z una variable libre.

Transformaciones Proyectivas 3D

► Sistema de coordenadas coincidente

- Obtenemos:

$$w_h = P^{-1}c_h = \begin{bmatrix} kx_0 \\ ky_0 \\ kz \\ \frac{kz}{f} + k \end{bmatrix}$$

que al pasar a coordenadas cartesianas proporciona:

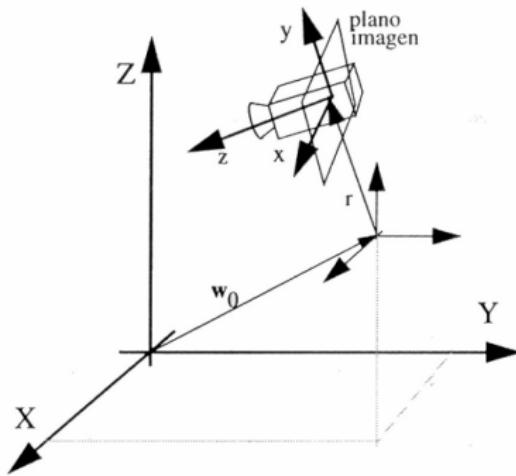
$$w = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} \frac{fx_0}{f+z} \\ \frac{fy_0}{f+z} \\ \frac{fz}{f+z} \end{bmatrix}$$

Despejando z en función de Z y sustituyendo se obtienen las expresiones de las intersecciones de los planos.

Transformaciones Proyectivas 3D

► Sistema de coordenadas no coincidente

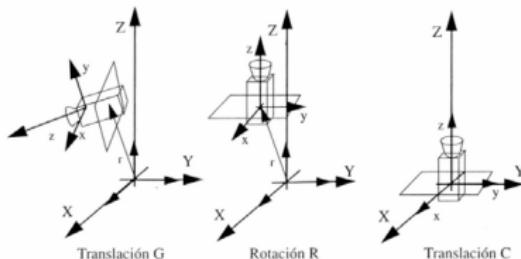
- Consideremos ahora un caso más general en el que los dos sistemas están separados y son, por tanto, no coincidentes:



Transformaciones Proyectivas 3D

► Sistema de coordenadas no coincidente

- Lo primero que tendremos que hacer es que los sistemas de cámara y mundo sean coincidentes: para ello concatenaremos una traslación, una rotación y otra traslación.
- La transformación de perspectiva se deberá multiplicar por la concatenación de transformaciones definidas.



- Un punto w del espacio se proyecta en el punto del plano imagen con coordenadas homogéneas:

$$c_h = PCRGw_h = Aw_h$$

donde A es la matriz resultante de concatenar las diversas transformaciones efectuadas.

Transformaciones Proyectivas 3D

► Sistema de coordenadas no coincidente

- Las coordenadas cartesianas (x, y) del punto del plano imagen se determinan dividiendo la primera y la segunda componente de c_h por la cuarta:

$$x = \frac{c_{h1}}{c_{h4}} = \frac{a_{11}X + a_{12}Y + a_{13}Z + a_{14}}{a_{41}X + a_{42}Y + a_{43}Z + a_{44}}$$

$$y = \frac{c_{h2}}{c_{h4}} = \frac{a_{21}X + a_{22}Y + a_{23}Z + a_{24}}{a_{41}X + a_{42}Y + a_{43}Z + a_{44}}$$

donde los a_{ij} son los elementos de la matriz A que se obtienen, en la práctica, mediante la calibración de la cámara.

- Reescribiendo las expresiones para ponerlas como función implícita de las variables tridimensionales (X, Y, Z) se obtiene la ecuación de los planos que definen la recta de proyección, en coordenadas del mundo, para cada punto del plano imagen:

$$(a_{11} - a_{41}x)X + (a_{12} - a_{42}x)Y + (a_{13} - a_{43}x)Z + (a_{14} - a_{44}x) = 0$$

$$(a_{21} - a_{41}y)X + (a_{22} - a_{42}y)Y + (a_{23} - a_{43}y)Z + (a_{24} - a_{44}y) = 0$$

Transformaciones Proyectivas 3D

|  | |  |  |  |  |
|---|--|---|---|---|--|
| <i>original</i> | | <i>Euclídea</i> | <i>similitud</i> | <i>afín</i> | <i>general</i> |
| Transformación | Matriz \mathbf{H} | Invariantes | | | |
| Euclídea | $\begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix}$ | longitud entre puntos, volumen. | | | |
| Similitud | $\begin{bmatrix} s\mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix}$ | ángulos entre planos, forma. | | | |
| Afín | $\begin{bmatrix} a_{11} & a_{12} & a_{13} & t_x \\ a_{21} & a_{22} & a_{23} & t_y \\ a_{31} & a_{32} & a_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$ | líneas paralelas. | | | |
| General | $\begin{bmatrix} h_{11} & h_{12} & h_{13} & h_{14} \\ h_{21} & h_{22} & h_{23} & h_{24} \\ h_{31} & h_{32} & h_{33} & h_{34} \\ h_{41} & h_{42} & h_{43} & h_{44} \end{bmatrix}$ | intersección y tangentes de superficies en contacto. | | | |

Calibración de la cámara

- ▶ En un sentido amplio, la calibración de la cámara hace referencia a la obtención de todos aquellos parámetros que inciden en la formación de la imagen, tanto los denominados geométricos, involucrados en el proceso geométrico de formación de la imagen, como los radiométricos, que tienen que ver con el brillo de los objetos proyectados.
- ▶ Abordaremos exclusivamente la calibración geométrica, que es con diferencia la de más interés práctico para los problemas de visión.
- ▶ A ella nos referiremos, en lo sucesivo, simplemente con el término *calibración*.

Calibración de la cámara

- ▶ La calibración de la cámara consiste en determinar los parámetros de transformación entre puntos 3D de la escena y puntos 2D de la imagen, que no es más que la propia transformación de perspectiva.
- ▶ Básicamente existen dos métodos de calibración:
 - ▶ **Auto-Calibración.**
 - ▶ **Calibración Fotogramétrica.**

Calibración de la cámara

► Auto-Calibración.

- ▶ En este tipo de calibración se parte de un conjunto de imágenes de la misma escena.
- ▶ Mediante la correspondencia entre puntos de distintas imágenes se pueden encontrar los mejores parámetros del modelo que ajusten dicha correspondencia.
- ▶ La reconstrucción 3D realizada con el modelo encontrado está afectada sin embargo por un factor de escala ya que en este método no se puede saber cuál es el tamaño real de los objetos captados por las cámaras.
- ▶ Un objeto pequeño cerca del centro óptico puede tener la misma imagen que el mismo objeto agrandado más cerca del plano de imagen.

Calibración de la cámara

► Calibración Fotogramétrica.

- ▶ Se utiliza cuando lo que se busca es una reconstrucción 3D precisa.
- ▶ Esta calibración utiliza un objeto 3D de referencia cuya geometría es conocida a la perfección.
- ▶ N puntos de interés son escogidos del objeto de referencia, obteniendo así las coordenadas $M'_i = (X'_i \ Y'_i \ Z'_i \ 1)^T$, para $i = 1, \dots, N$.
- ▶ El objeto es a continuación captado por la cámara y sus puntos de interés son vistos como puntos 2D con coordenadas $w_i = [u_i \ v_i \ 1]^T$.
- ▶ Teniendo un modelo de la proyección es posible obtener una estimación teórica de los puntos 3D.

Calibración de la cámara

- ▶ La transformación entre puntos 3D de la escena y puntos 2D de la imagen no es más que una función no lineal que depende de un vector de 11 parámetros, los parámetros el modelo.
- ▶ El problema de calibración se transforma en un problema de optimización mediante el cual una función objetivo que mide el error entre la proyección estimada y la proyección medida debe ser minimizada.
- ▶ Existen dos tipos de parámetros que definen esta transformación:
 - ▶ **Parámetros intrínsecos:** Son aquellos que caracterizan las propiedades inherentes de la cámara y de la óptica, es decir, los parámetros involucrados en la transformación de puntos 3D en el sistema de referencia de la cámara a puntos 2D del plano imagen.
 - ▶ Se suelen considerar 5 parámetros:
 - ▶ Distancia focal: f
 - ▶ Desplazamiento del centro de la imagen: c_x, c_y
 - ▶ Coeficientes de distorsión: k_1, k_2

Calibración de la cámara

- ▶ En un modelo *pin hole* la imagen se forma sin distorsión.
- ▶ En un modelo óptico más real la imagen se desplaza a las coordenadas (c_x, c_y) y se distorsiona según un *modelo radial*:

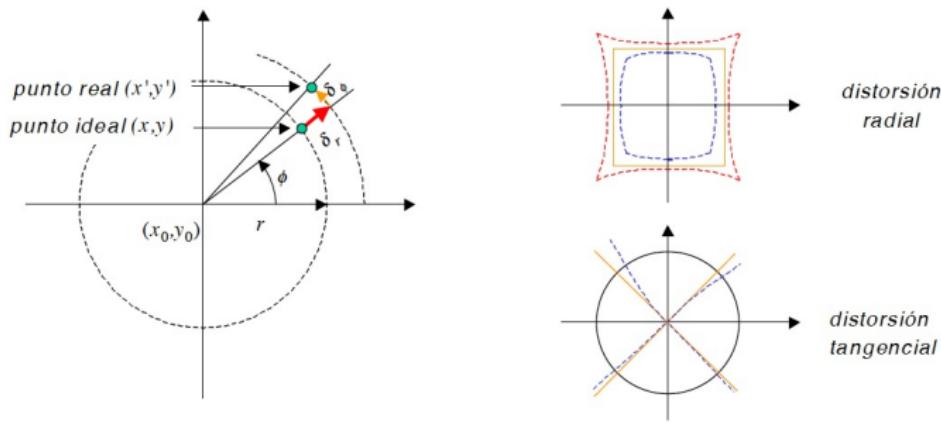
$$\begin{aligned}x_d + d_x &= x_u & \text{con} & \quad d_x = x_d(k_1 r^2 + k_2 r^4) \\y_d + d_y &= y_u & & \quad d_y = y_d(k_1 r^2 + k_2 r^4)\end{aligned}$$

$y \quad r = (x_d^2 + y_d^2)^{1/2}$

donde (x_d, y_d) son las coordenadas del píxel distorsionado y (x_u, y_u) las del píxel sin distorsionar.

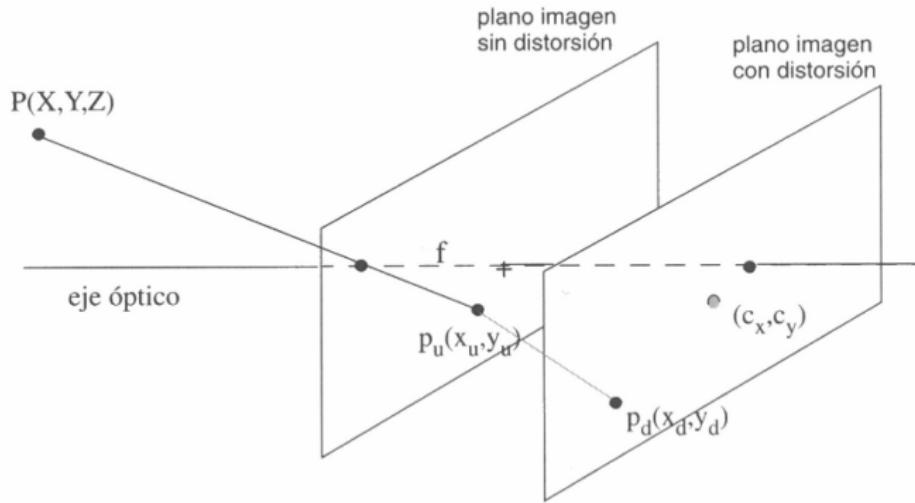
- ▶ Añadida a la anterior o de forma independiente nos podemos encontrar con una *distorsión tangencial* que a diferencia de la anterior que depende del radio, ésta sólo depende del ángulo.
- ▶ Las coordenadas distorsionadas son conocidas y las otras, las reales, no.

Calibración de la cámara



Modelo de distorsión de lente en componente radial y tangencial

Calibración de la cámara



Parámetros intrínsecos involucrados en un modelo de proyección con distorsión

Calibración de la cámara

- ▶ **Parámetros extrínsecos:** Son los seis parámetros que definen la posición y orientación de la cámara con respecto al sistema de referencia absoluto:
 - ▶ Traslación: T_x, T_y, T_z
 - ▶ Rotación: ángulos α, β, γ
- ▶ Sólo si una cámara está calibrada es posible establecer la relación entre las coordenadas 3D de los objetos y sus correspondientes proyecciones 2D y al revés.
- ▶ Su importancia es vital en todas aquellas tareas de procesamiento de información visual en donde se involucren las posiciones o geometrías de los objetos.

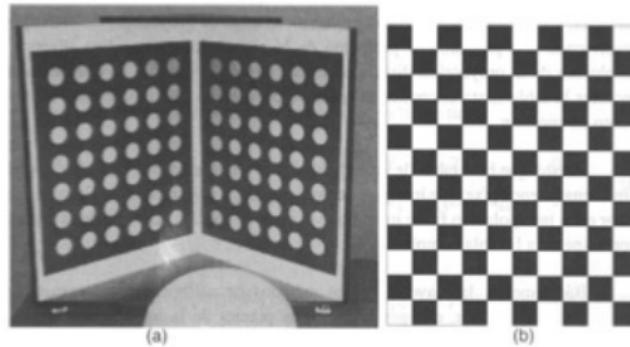
Calibración de la cámara

► Procedimiento general

- Aunque existen distintos métodos para la calibración de una cámara, el procedimiento básico es el mismo en todos ellos:
 1. Determinar con precisión un conjunto de puntos tridimensionales.
 2. Determinar sus correspondientes proyecciones en la imagen.
 3. Obtener los parámetros que mejor resuelven la correspondencia entre unos y otros.
- Los dos primeros pasos requieren conocer una serie de puntos 3D y sus correspondientes proyecciones en la imagen. Se denominan *puntos de calibración*.
- Dependiendo del tipo de método empleado, estos puntos pueden o no ser coplanarios.

Calibración de la cámara

- ▶ Los puntos de calibración se deben tomar sobre una plantilla que permita el posicionamiento exacto de éstos en el espacio y su localización automática y precisa en la imagen.
- ▶ Para facilitar el posicionamiento de los puntos y siempre que sea posible, la plantilla se coloca paralela al plano $X_w Y_w$, con lo cual todos los puntos tendrán la misma cota Z_w .



a) Plantilla de puntos no coplanarios, b) Plantilla de puntos coplanarios.

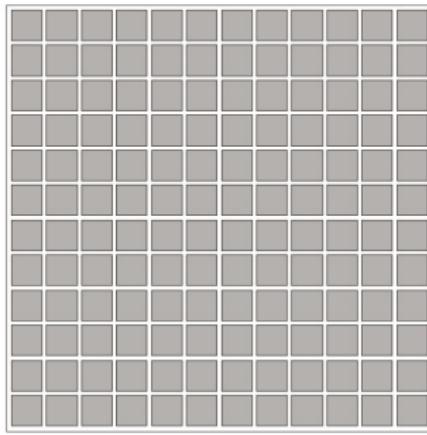
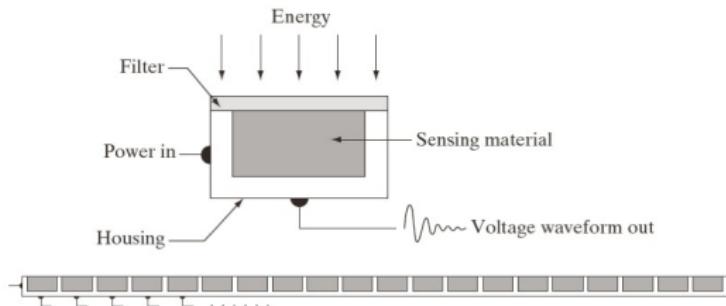
Concepto de imagen digital

- ▶ El concepto de imagen digital está asociado a una función bidimensional $f(x, y)$ cuya amplitud o valor será el grado de iluminación (intensidad de la luz) en el espacio de coordenadas (x, y) de la imagen para cada punto.
- ▶ El valor de esta función depende de la cantidad de luz que incide sobre la escena vista, así como de la parte que sea reflejada por los objetos que componen dicha escena: componentes denominados *iluminación* $i(x, y)$ y *reflexión* $r(x, y)$
- ▶ El producto de ambas funciones proporciona la función $f(x, y)$:

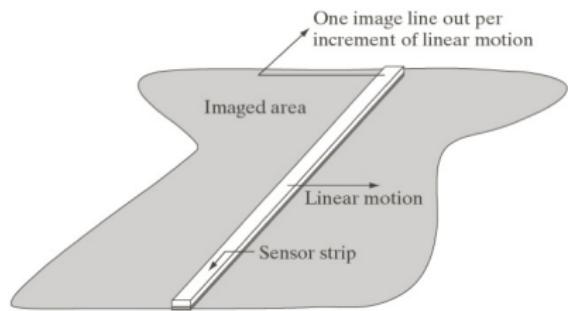
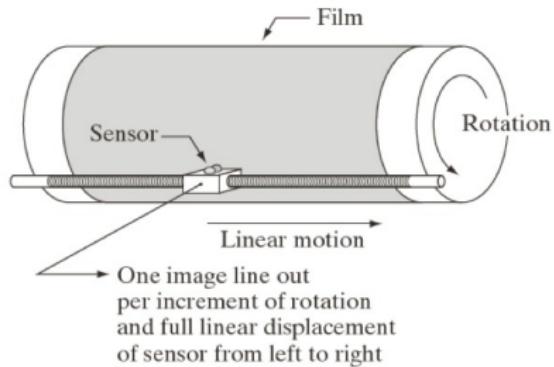
$$f(x, y) = i(x, y)r(x, y)$$

- ▶ La naturaleza de la iluminación viene determinada por la fuente de luz, mientras que las reflexión depende de las características del objeto en la escena.

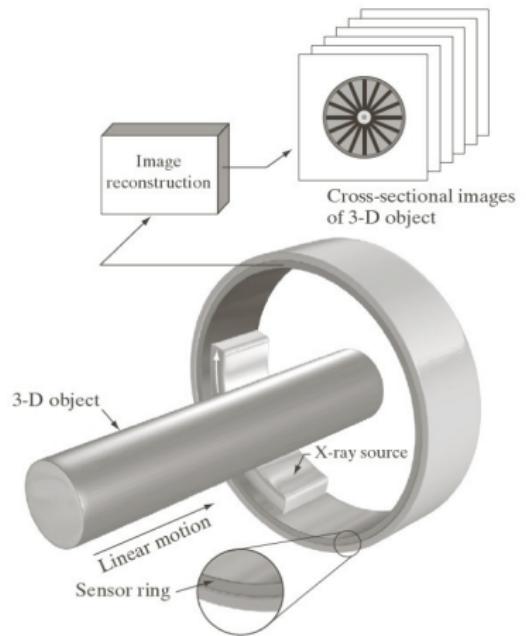
Sensores



Sensores

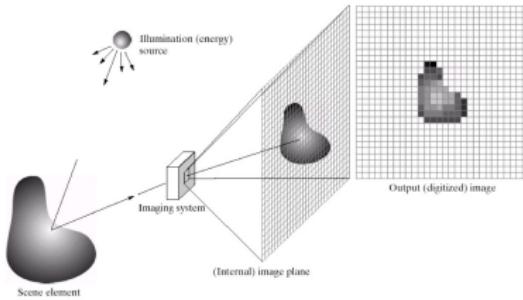


Sensores

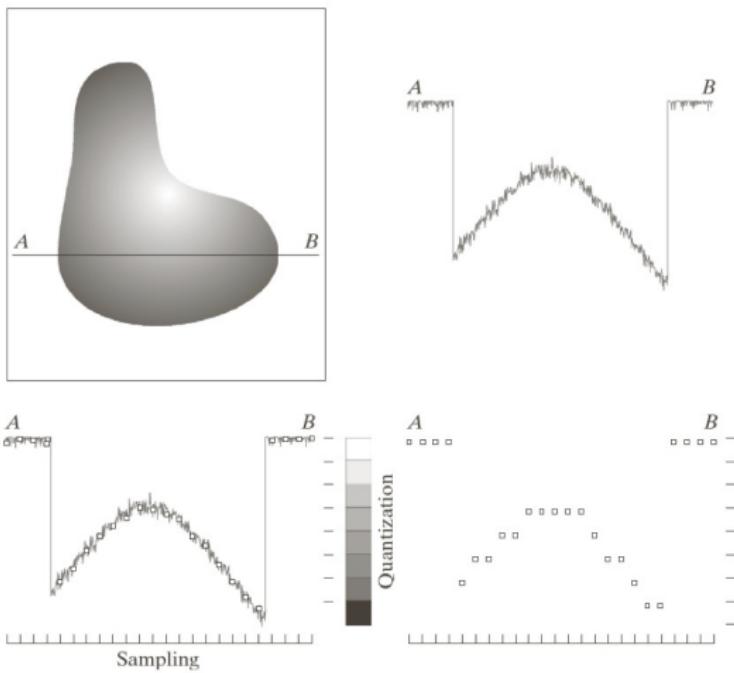


Muestreo y Cuantización

- ▶ Una señal es una función que depende de alguna variable que tiene un significado físico.
- ▶ Dependiendo del número de variables puede ser de distintas dimensiones, en el caso de las imágenes es bidimensional.
- ▶ Las funciones tienen un dominio y un rango.
- ▶ Si el dominio y el rango son continuos, la señal es continua o analógica; si el dominio es discreto pero el rango no, la señal también será discreta, y si el dominio y el rango son discretos, caso de las imágenes, la señal es digital.

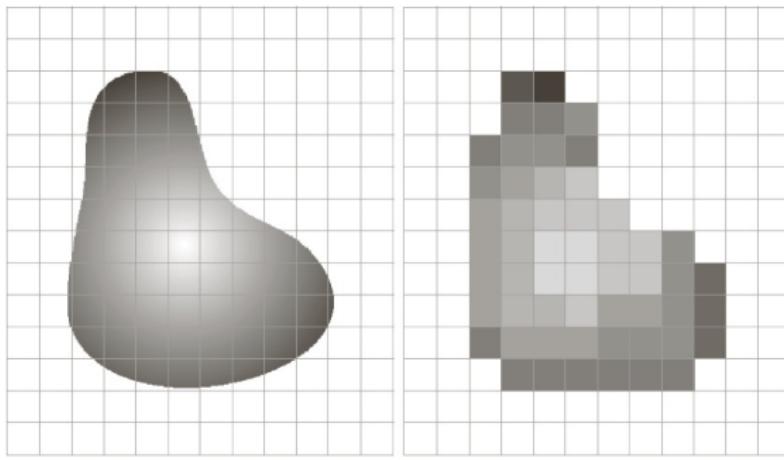


Muestreo y Cuantización



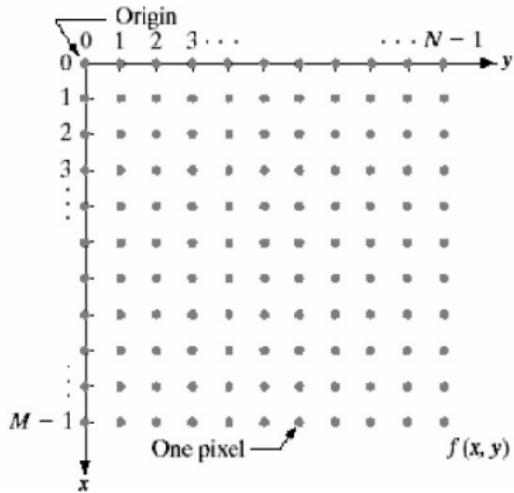
Muestreo y Cuantización

- ▶ La digitalización de las coordenadas espaciales, su dominio, está asociada al concepto de *muestreo*, mientras que la digitalización de su rango está asociado al de *cuantización* de los niveles de gris.
- ▶ El muestreo es la conversión que sufren las dos dimensiones espaciales de la señal analógica y que genera la noción de píxel.
- ▶ La cuantización es la conversión que sufre el rango de la señal analógica y genera el concepto de nivel de gris o intensidad.



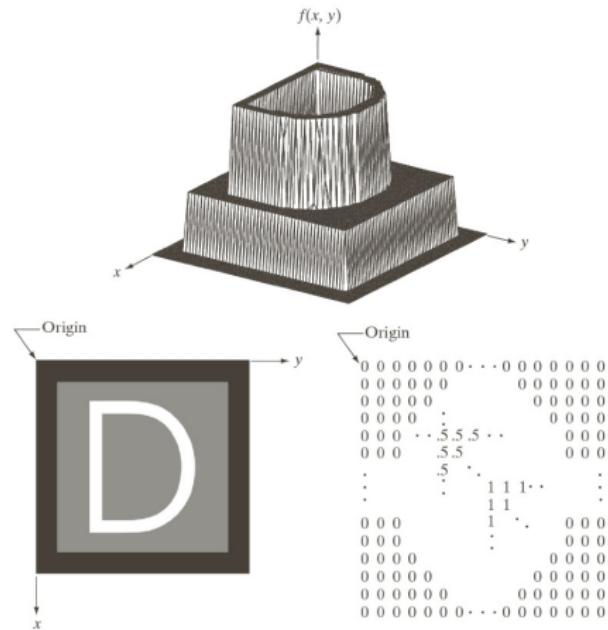
Muestreo y Cuantización

- ▶ Una imagen $f(x, y)$ está soportada por una matriz de $M \times N$ elementos.



$$f(x, y) = \begin{bmatrix} f(0, 0) & f(1, 0) & \cdots & f(N - 2, 0) & f(N - 1, 0) \\ f(0, 1) & f(1, 1) & \cdots & f(N - 2, 1) & f(N - 1, 1) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ f(0, M - 2) & f(1, M - 2) & \cdots & f(N - 2, M - 2) & f(N - 1, M - 2) \\ f(0, M - 1) & f(1, M - 1) & \cdots & f(N - 2, M - 1) & f(N - 1, M - 1) \end{bmatrix}$$

Muestreo y Cuantización

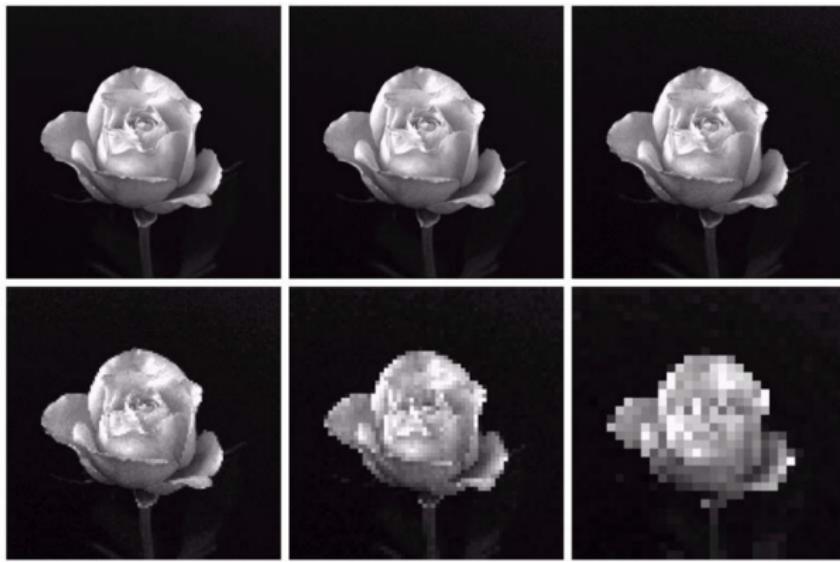


Muestreo y Cuantización



A 1024×1024 , 8-bit image subsampled down to size 32×32 pixels. The number of allowable gray levels was kept at 256.

Muestreo y Cuantización



| | | |
|---|---|---|
| a | b | c |
| d | e | f |

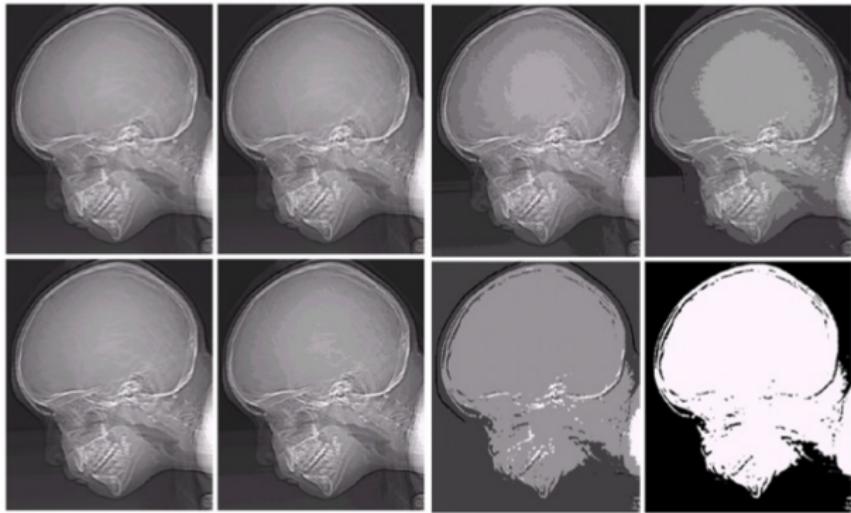
(a) 1024 × 1024, 8-bit image. (b) 512 × 512 image resampled into 1024 × 1024 pixels by row and column duplication. (c) through (f) 256 × 256, 128 × 128, 64 × 64, and 32 × 32 images resampled into 1024 × 1024 pixels.

Muestreo y Cuantización



Efecto de la reducción de la resolución espacial. Desde 1250 dpi, 300 dpi, 150 dpi y 72 dpi.

Muestreo y Cuantización



El color

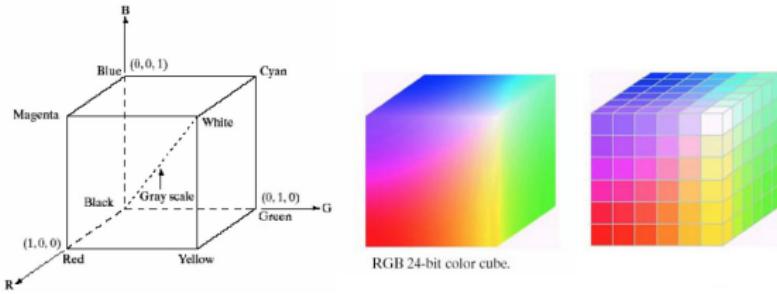
- ▶ Aunque el color es una de las características más importantes que definen a los objetos, hasta hace poco tiempo se le prestaba poca atención debido al coste computacional, a la memoria para procesar dichas imágenes y a la no existencia de resultados robustos que mostraran la ventaja de utilizar imágenes en color frente a sus correspondientes versiones en niveles de gris.
- ▶ El espacio de color como soporte de las sensaciones de color:
 - ▶ Brillo: sensación que indica si un área está más o menos iluminada.
 - ▶ Tono: sensación que indica si un área es similar al rojo, amarillo, verde o azul o a una proporción de dos de ellos.
 - ▶ Coloración: sensación por la que un área tiene un mayor o menor tono.
 - ▶ Luminosidad: brillo de una zona frente a otra blanca en la imagen.
 - ▶ Croma: coloridad de un área respecto al brillo de un blanco de referencia.
 - ▶ Saturación: relación entre coloridad y brillo.
- ▶ Los parámetros psíquicos de la percepción del color son la luminosidad, el tono y la saturación.

El color

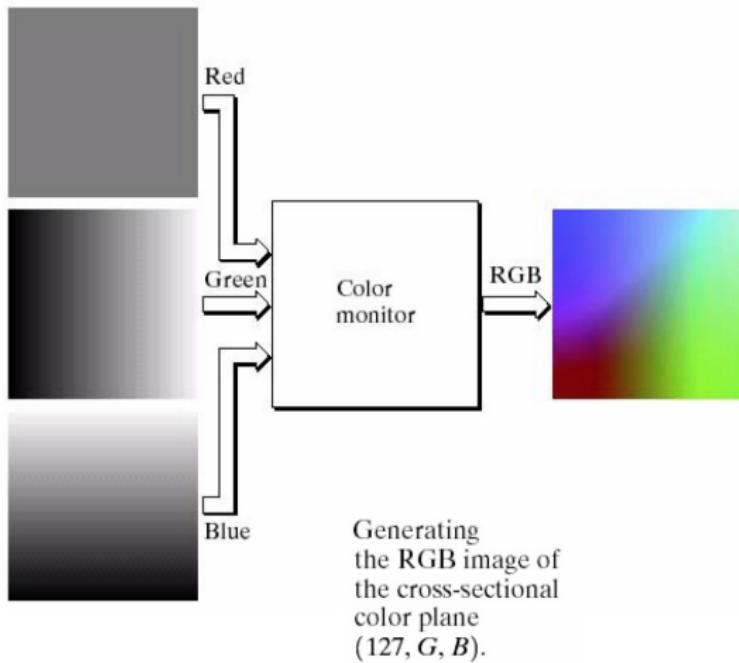
- ▶ Un espacio de representación del color es un método por el que se puede especificar, crear o visualizar cualquier color.
- ▶ **RGB**
 - ▶ El espacio RGB se basa en la combinación de tres señales de luminancia cromática distinta: el rojo, el verde y el azul.
 - ▶ La manera más sencilla e intuitiva de conseguir un color concreto es determinar la cantidad de color rojo, verde y azul que se necesita combinar. Para ello se realiza la suma aritmética de las componentes:

$$X = R + G + B$$

- ▶ Representación geométrica:



El color



El color



El color

► HSI

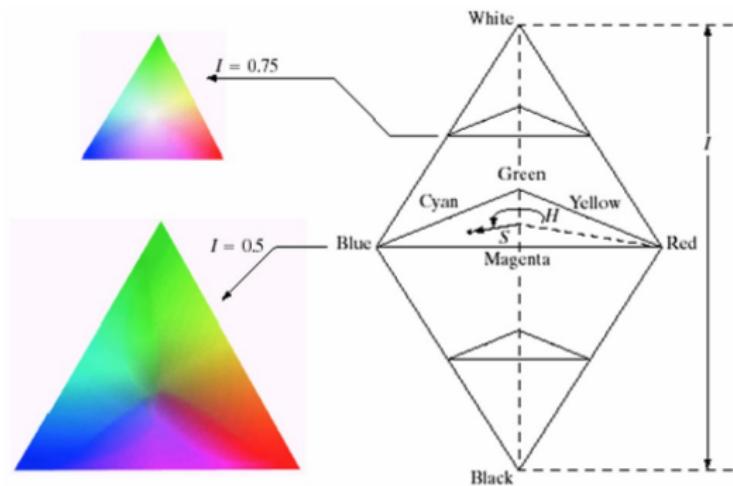
- ▶ El espacio de color *HSI* se basa en el modo de percibir los colores que tenemos los humanos.
- ▶ Dicho sistema caracteriza el color en términos de tono o tinte (Hue), saturación o cromatismo (Saturation) y brillo (Intensity).
- ▶ Transformaciones matemáticas para proyectar colores en RGB a HSI:

$$I = \frac{R + G + B}{3}$$

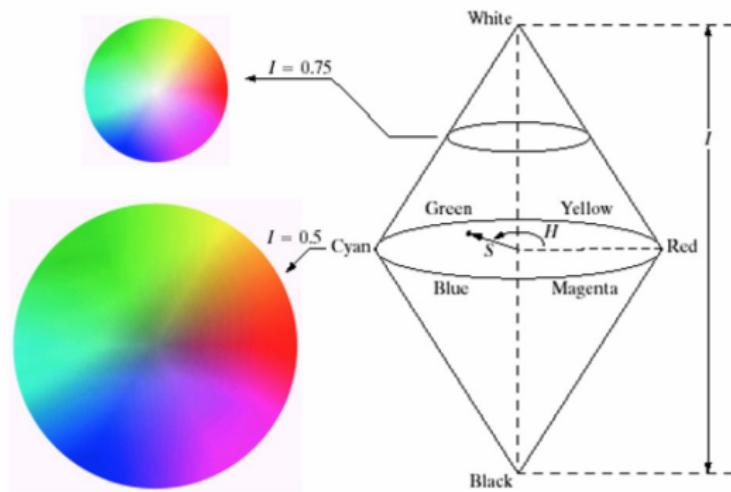
$$H = \arctan \left(\frac{\sqrt{3}(G - B)}{(R - G) + (R - B)} \right)$$

$$S = 1 - \frac{\min(R, G, B)}{I}$$

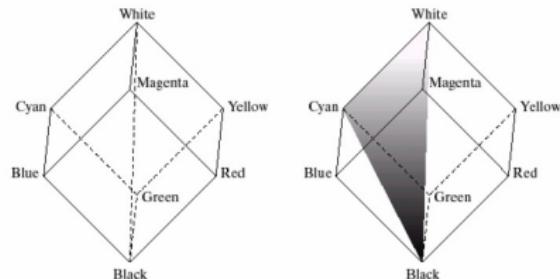
El color



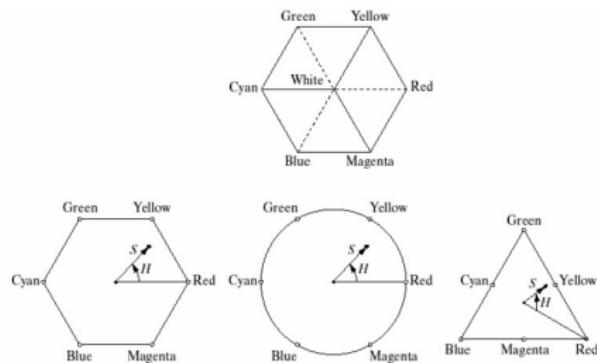
El color



El color



Conceptual relationships between the RGB and HSI color models.



Hue and saturation in the HSI color model. The dot is an arbitrary color point. The angle from the red axis gives the hue, and the length of the vector is the saturation. The intensity of all colors in any of these planes is given by the position of the plane on the vertical intensity axis.

El color

► HSI

- La proyección inversa que permite recuperar de nuevo la información de la imagen en RGB, viene dada por las siguientes relaciones en función de los valores del tono:

$$0^\circ < H \leq 120^\circ$$

$$B = \frac{1}{3}(1 - S)$$

$$R = \frac{1}{3} \frac{(1 + S \cos H)}{\cos(60 - H)}$$

$$G = 1 - (B + R)$$

$$120^\circ < H \leq 240^\circ$$

$$H = H - 120^\circ$$

$$R = \frac{1}{3}(1 - S)$$

$$G = \frac{1}{3} \frac{(1 + S \cos H)}{\cos(60 - H)}$$

$$B = 1 - (R + G)$$

$$240^\circ < H \leq 360^\circ$$

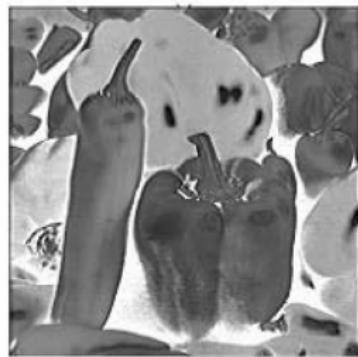
$$H = H - 240^\circ$$

$$G = \frac{1}{3}(1 - S)$$

$$B = \frac{1}{3} \frac{(1 + S \cos H)}{\cos(60 - H)}$$

$$R = 1 - (G + B)$$

El color



El color

- ▶ **XYZ**
- ▶ **CIE Luv**
- ▶ **CIE Lab**
- ▶ **YIQ**
- ▶ **YUV**
- ▶ ...

Índice

Conceptos básicos

Procesos de captura de información visual

Espacios transformados

- ▶ Todas las aproximaciones al procesamiento de imagen hasta este momento trabajan directamente sobre el píxel de la imagen de entrada sobre su *dominio espacial*.
- ▶ En muchas ocasiones las técnicas de procesamiento de imagen aplicadas requieren de una serie de propiedades que no están presentes en este dominio espacial.
- ▶ Ello genera la necesidad de buscar *espacios transformados* en los que dichas propiedades estén presentes, lo que implica dotarnos tanto de mecanismos de proyección de nuestros datos en dichos espacios, como de procesos de transformación inversa que nos permitan devolver a dichos datos a su espacio de representación original.
- ▶ En la literatura nos podemos encontrar infinidad de espacios de representación con propiedades y características bien distintas.

- ▶ Una clase importante de transformaciones lineales bidimensionales, notadas como $T(u, v)$, pueden expresarse de forma general como:

$$T(u, v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) r(x, y, u, v)$$

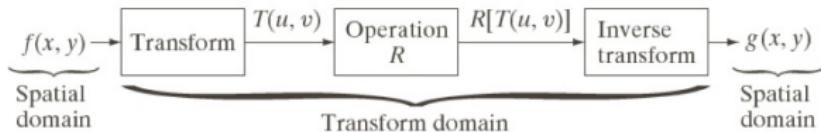
donde $f(x, y)$ es la imagen de entrada, $r(x, y, u, v)$ es el núcleo de la transformada directa, $u = 0, 1, 2, \dots, M - 1$ y $v = 0, 1, 2, \dots, N - 1$, x e y son variables espaciales, mientras que M y N son las filas y columnas de F . Las variables u y v son las *variables de transformación*.

- ▶ A la función $T(u, v)$ se le denomina *transformada directa* de $f(x, y)$.
- ▶ La *transformada inversa* de $T(u, v)$ está dada por:

$$f(x, y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} T(u, v) s(x, y, u, v)$$

donde $x = 0, 1, 2, \dots, M - 1$, $y = 0, 1, 2, \dots, N - 1$ y $s(x, y, u, v)$ es el denominado *núcleo de transformación inverso*.

- ▶ Esquema general de procesamiento en un dominio transformado.



- ▶ Un núcleo de transformación directo será separable si:

$$r(x, y, u, v) = r_1(x, u)r_2(y, v)$$

- ▶ Así mismo, será simétrico si $r_1(x, y)$ es funcionalmente igual a $r_2(x, y)$:

$$r(x, y, u, v) = r_1(x, u)r_1(y, v)$$

- ▶ Mismos comentarios para el núcleo de transformación inverso cambiando r por s en las anteriores expresiones.

- ▶ La transformada de Fourier bidimensional discreta tiene como núcleos de transformación directos e inversos los siguientes:

$$r(x, y, u, v) = e^{-j2\pi(ux/M+vy/N)}$$

$$s(x, y, u, v) = \frac{1}{MN} e^{j2\pi(ux/M+vy/N)}$$

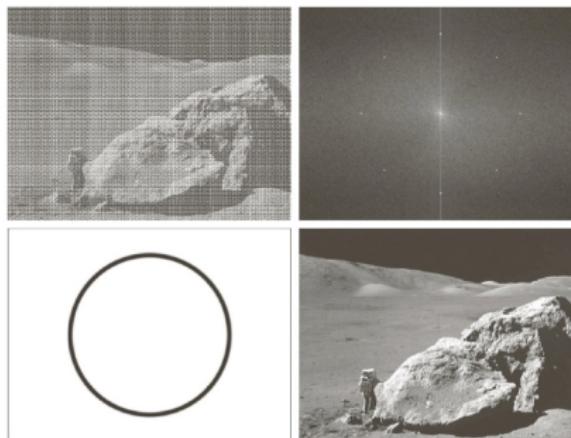
donde $j = \sqrt{-1}$, por lo que los núcleos son complejos.

- ▶ Dando lugar al siguiente par de transformadas directa e inversa:

$$T(u, v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi(ux/M+vy/N)}$$

$$f(x, y) = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} T(u, v) e^{j2\pi(ux/M+vy/N)}$$

- ▶ Los núcleos de la transformada de Fourier bidimensional son separables y simétricos, lo que permite que transformaciones 2D puedan ser computadas como transformadas 1D.
- ▶ La transformadas de Fourier no es el único espacio transformado en procesamiento de la información visual: Walsh, Hadamard, coseno discreta, Haar, etc.



- a) Imagen corrompida por interferencias sinusoidales, b) Espectro de Fourier en el que se pueden apreciar los patrones de energía responsables de las interferencias, c) Máscara usada para filtrar dicha imagen, d) Resultado obtenido.