

Tema 1

- Introducción. Percepción visual
- Conceptos básicos. Tipos de imágenes
- Representación. Espacios de color
- Etapas en el procesamiento de imágenes

PROCESAMIENTO DE IMÁGENES Y VIDEO

¿Qué es el procesamiento de imágenes?

Un conjunto de métodos, técnicas y herramientas destinadas a la manipulación y análisis de imágenes y video

Objetivo:

Mejorar el aspecto de las imágenes y hacer más evidentes en ellas ciertos detalles que se desean hacer notar.

Mejora, restauración y eliminación del ruido. Creación y manipulación del contenido. Compresión y transmisión. Extracción de información

La imagen puede haber sido generada de muchas maneras, por ejemplo, fotográficamente, electrónicamente, por medio de monitores de televisión.

El procesamiento de las imágenes se puede en general hacer por medio de métodos ópticos, o bien por medio de métodos digitales, en una computadora.

El teorema de Fourier es el eje central de los principios matemáticos implícitos en ambos métodos.

PROCESAMIENTO DE IMÁGENES Y VIDEO

Aplicaciones del procesamiento de imágenes

Es muy amplia y abarca muchos campos.

Biología y genética (análisis de huesos, tejidos y células,...)

Defensa/Inteligencia (interpretación automática de imágenes satélites en búsqueda de objetivos militares)

Procesamiento de documentos (detección y reconocimiento automático de caracteres dentro de un documento: cheques bancarios, correo postal,...)

Automatización industrial (inspección visual automática)

Forense/criminología (análisis de huellas dactilares, análisis de similitud de fotografías,...)

Análisis de materiales (clasificación de objetos, impurezas,...)

Fotografía y video (composición de escenas, adición de efectos especiales,...)

Imágenes satélites (análisis de paisajes: analizar recursos naturales, medir características de vegetación como contenido de agua, temperatura, características del cultivo,...)

Medicina (Imágenes obtenidas con fines de diagnóstico médico: Rayos X, resonancias magnéticas, sistemas de ayudas al diagnóstico; Imágenes: cerebrales 3D, cardíacas , vasculares,...)

Visión artificial (reconocimiento de formas, detección, seguimiento, video-vigilancia,...)

Capítulo 1. Imágenes digitales y sus propiedades

El sistema de visión humano nos suministra una percepción y comprensión del mundo que nos rodea

- ⇒ Más de un **70%** de la información total (imágenes, sonidos, sensaciones, etc.) que recibimos es visual.
- ⇒ Las escenas que percibimos suelen ser tridimensionales (**3D**) y cuando las capturamos mediante dispositivos (cámaras fotográficas o de vídeo, pantallas de rayos X, etc.) obtenemos imágenes bidimensionales (**2D**). Esta proyección a una dimensión inferior supone una gran pérdida de información.
- ⇒ Las escenas dinámicas donde aparecen objetos en movimiento complican más todavía la visión por ordenador.

Capítulo 1. Imágenes digitales y sus propiedades

Diferentes tipos de imágenes según el dispositivo utilizado:

⇒ Imágenes formadas mediante:

- **luz visible** (cámara)
- **rayos infrarrojos** (visión nocturna)
- **rayos ultravioleta** (espectrógrafo)
- **campos magnéticos** (campo magnético-electroimán)

⇒ Imágenes formadas a partir de:

- **ultrasonidos** (reflexión de ondas)
- **radiación de rayos X** (exploración continua con rayos)
- **impulsos eléctricos** (amplificación de impulsos)
- **isótopos radiactivos** (indicadores o trazadores con gammacámara)

Capítulo 1. Tipos de imágenes digitales y sus propiedades

- ⇒ Imágenes formadas por **rayos infrarrojos**, que permiten la visión nocturna.
- ⇒ Imágenes formadas por **rayos ultravioleta**, como **espectrogramas** (fotografía de un espectro luminoso obtenida mediante un espectrógrafo, sobre placa de cristal o película sensible).
- ⇒ Imágenes formadas a partir de **campos magnéticos**, como la **resonancia magnética** utiliza un potente campo magnético para obtener imágenes detalladas del corazón o del tórax. Se coloca a la persona dentro de un gran electroimán que causa una vibración de los núcleos de los átomos del organismo, produciendo unas señales características que son convertidas en imágenes.
- ⇒ Imágenes formadas a partir de **ultrasonidos**, como la **ecografía**, que forma una imagen por la reflexión de las ondas sonoras con determinadas partes del cuerpo. El **ecocardiograma**, que utiliza ondas ultrasonoras de alta frecuencia que chocan contra las estructuras del corazón y de los vasos sanguíneos y, al rebotar, producen una imagen móvil que aparece en una pantalla de vídeo.

Capítulo 1. Tipos de imágenes digitales y sus propiedades

⇒ Imágenes formadas a partir de una radiación de **rayos X**, como las **radiografías** de tórax o la **radioscopia** (fluoroscopia) que es una exploración continua con rayos X que muestra en una pantalla el movimiento del corazón y los pulmones. La tomografía computarizada que crea imágenes transversales de todo el tórax utilizando los rayos X y muestra la ubicación exacta de las anomalías. La **angiografía** muestra con detalle el aporte de sangre, por ejemplo, a los pulmones (se inyecta un líquido radiopaco dentro de un vaso sanguíneo que puede verse en las radiografías) o al cerebro. Detecta anomalías vasculares como la obstrucción de un vaso sanguíneo (ictus), aneurismas (bolsas en una arteria) o arteritis (inflamaciones).

⇒ Imágenes formadas a partir de **impulsos eléctricos**, como el **electrocardiograma**, que amplifica los impulsos eléctricos del corazón y se registran en un papel en movimiento. El **electroencefalograma** es una imagen formada a partir de los impulsos eléctricos de 20 alambres (electrodos) colocados sobre el cuero cabelludo con el objeto de establecer el trazado y registro eléctrico de la actividad cerebral.

Capítulo 1. Tipos de imágenes digitales y sus propiedades

⇒ Imágenes formadas a partir de **isótopos radiactivos** (indicadores o trazadores). Los indicadores se reparten por todo el cuerpo y se detectan con una gammacámara (detección de coágulos de sangre en los pulmones). En la técnica de **tomografía computarizada** por emisión de fotones simples, distintos tipos de cámaras de registro de radiaciones pueden grabar una imagen simple o producir una serie de imágenes de secciones transversales amplificadas por el ordenador. En las imágenes formadas por *emisión de positrones*, como la **tomografía por emisión de positrones**, se inyecta una sustancia en la sangre que se desplaza hasta las estructuras cerebrales, permitiendo medir la actividad que desarrolla el cerebro. También permite analizar, por ejemplo, el funcionamiento del corazón.

Capítulo 1. Conceptos básicos

⇒ Cualquiera de estas imágenes se puede modelar por una **función continua** de dos o tres variables.

⇒ En el caso de imágenes estáticas los dos argumentos de la función son las coordenadas (x, y) del plano cartesiano, mientras que si las imágenes van cambiando en el tiempo (dinámicas) necesitamos una tercera coordenada t que nos especifica el tiempo.

$$\text{punto de la imagen} \equiv (x, y, t) \rightarrow f(x, y, t) \equiv \text{intensidad luminosa}$$

Dicha función suele ser una **función digital** (dominio discreto y rango discreto)

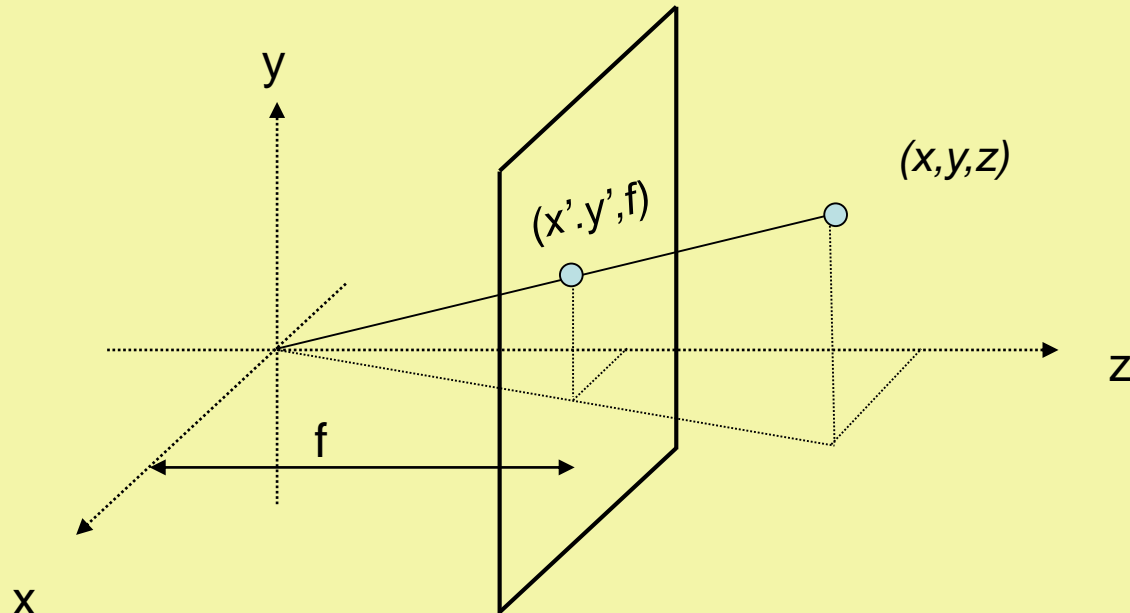
⇒ Los valores de la función corresponden a la luminosidad, brillo o intensidad de la radiación de los puntos de la imagen. También pueden representar temperatura, presión, distancia al observador, etc.

Capítulo 1. Conceptos básicos

Una imagen 2D es el resultado de una **proyección** de una escena 3D. Un punto de la escena 3D viene representado por su proyección en el plano imagen. Así, el punto $P=(x,y,z)$ se proyecta en el punto $P'=(x',y',f)$ cuyas coordenadas vienen dadas por las expresiones:

$$x' = \frac{xf}{z}, \quad y' = \frac{yf}{z}$$

que se obtienen directamente del teorema de **Tales**



Capítulo 1. Conceptos básicos

⇒ Cuenta la historia que un sacerdote egipcio le preguntó a **Tales de Mileto** (s. IV a. C) a cerca de la altura de la Pirámide de Keops, cuando ya las pirámides rondaban los 2.000 años de edad, y éste respondió con un método de lo más ingenioso para medir dicha altura.

La historia dice así:

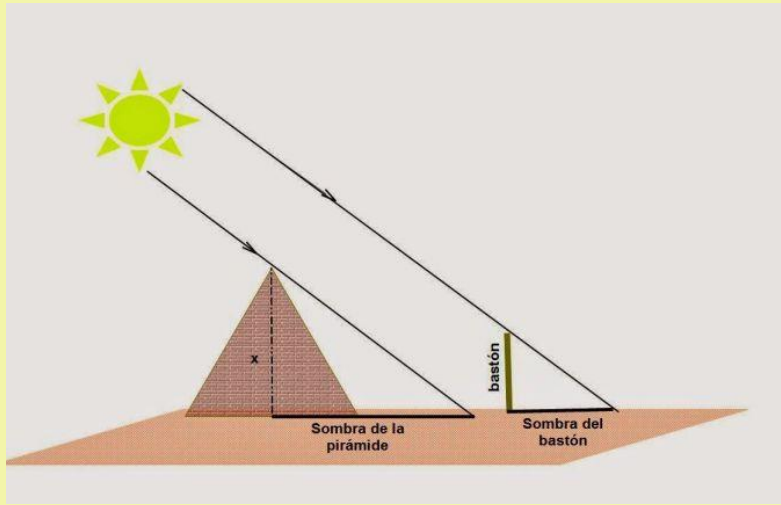
“Un sacerdote egipcio le pregunta sonriendo cuál puede ser la altura de la pirámide del rey Khufu (la pirámide de Keops). **Tales** reflexiona y a continuación contesta que no se conforma con calcularla a ojo, sino que la medirá sin ayuda de ningún instrumento. Se echa sobre la arena y determina la longitud de su propio cuerpo.

Los sacerdotes le preguntan qué es lo que está pensando, y Tales les explica: ‘Me pondré simplemente en un extremo de esta línea, que mide la longitud de mi cuerpo, y esperaré hasta que mi sombra sea igual de larga. En ese instante , la sombra de la pirámide de vuestro Khufu también ha de medir tantos pasos como la altura de la pirámide.’

El sacerdote, desorientado por la extrema sencillez de la solución, se pregunta si acaso no hay algún error, algún sofisma, **Tales** añade: ‘Pero si queréis que os mida esa altura, a cualquier hora, clavaré en la arena mi bastón.’

El método que utilizó **Tales de Mileto** para calcular la altura de la Pirámide de Keops es lo que conocemos como **Teorema de Tales**

Capítulo 1. Conceptos básicos



La sombra es la región donde no dan los rayos del sol. Se supone que los rayos que inciden en la pirámide y en el bastón son paralelos y el bastón está clavado perpendicularmente al suelo. De esta forma, los ángulos de los triángulos son iguales entre sí y, por tanto, **dichos triángulos son semejantes**.

$$\frac{\text{sombra pirámide}}{\text{sombra bastón}} = \frac{x}{\text{longitud bastón}}$$

Supongamos que las medidas se realizaron a una hora del día en la que:

sombra pirámide = 200 m

sombra bastón = 2.05 m

Y que el **bastón** medía **1.5 m**

Se tendría entonces:

$$\frac{200}{2.05} = \frac{x}{1.5}$$

De donde se obtiene:

$$x = 1,5 \cdot \frac{200}{2,05} \Rightarrow \underline{\underline{x = 146,34 \text{ m}}}$$

Valor aproximado que tenía la pirámide de Keops en la antigüedad (actualmente 137 m aproximadamente).

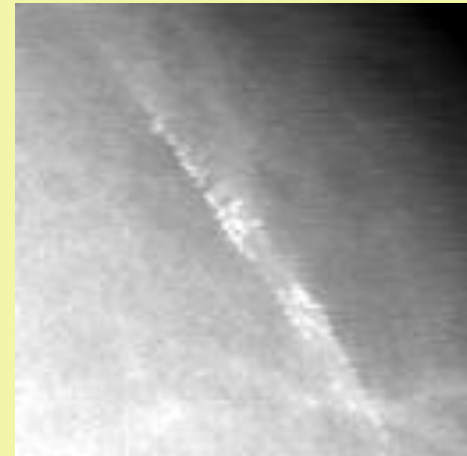
Capítulo 1. Conceptos básicos

⇒ En una **imagen monocromática** el rango de f tiene un valor mínimo (negro) y un valor máximo (blanco). Si se trata de una **imagen digital** el rango viene dado por un conjunto finito de valores, como puede ser el conjunto $\{0, 1, \dots, L-1, L\}$, donde el 0 corresponde al negro y el valor L a blanco. En este caso diremos que la imagen tiene $L+1$ niveles o **tonos de gris**. El dominio de la función digital es también una región acotada del plano, D , de la forma:

$$D = \{ (m, n): m=1,2,\dots,M, n=1,2,\dots,N \}.$$

Una **imagen monocromática** viene dada por una matriz

$$f = \begin{bmatrix} f(1,1) & f(1,2) & \dots & f(1,N) \\ f(2,1) & f(2,2) & \dots & f(2,N) \\ \dots & \dots & \dots & \dots \\ f(M,1) & f(M,2) & \dots & f(M,N) \end{bmatrix}$$



Sus elementos se llaman **píxeles** y sus valores **tonos de gris**

Capítulo 1. Conceptos básicos

Un **píxel** de una imagen digitalizada es un punto del dominio de la función digital correspondiente que tiene asociado el valor de dicha función y cuya posición viene determinada por sus coordenadas (x, y) . En el análisis de imágenes juegan un papel importante los píxeles de un entorno (píxel próximos) de cada píxel. Por ello, es necesario definir una función distancia entre píxeles.

La **distancia Euclídea** entre los píxeles (i, j) y (h, k) viene dada por la expresión:

$$D_E[(i, j), (h, k)] = \sqrt{(h-i)^2 + (k-j)^2}$$

La **distancia rectangular** entre los píxeles (i, j) y (h, k) viene dada por la expresión:

$$D_4[(i, j), (h, k)] = |h-i| + |k-j|$$

Si se permiten movimientos en diagonal, como en el tablero de ajedrez, la distancia entre los píxeles (i, j) y (h, k) viene dada por la expresión:

$$D_8[(i, j), (h, k)] = \max\{|h-i|, |k-j|\}$$

y se llama **distancia de Tchebychev** o del tablero de ajedrez.

Capítulo 1. Conceptos básicos

La calidad de una imagen digital viene definida por dos características:

- **Profundidad de color.-** es el número de colores que tiene una imagen. Lo determina el número de bits que almacenan la información del color de un píxel. 1 bit = 2 posibles colores (blanco y negro, según su valor sea 0 o 1), 8 bits = 256 colores (256 son las combinaciones posibles de 8 bits que valen 0 o 1), 16 bits= 65.536 posibles colores, 24 bits = 16,7 millones de posibles colores.
- **Compresión de la información.-** es la forma en que se almacena la información de la imagen en relación al tamaño que ocupa. A mayor compresión, mayor pérdida de información y menor tamaño de archivo.

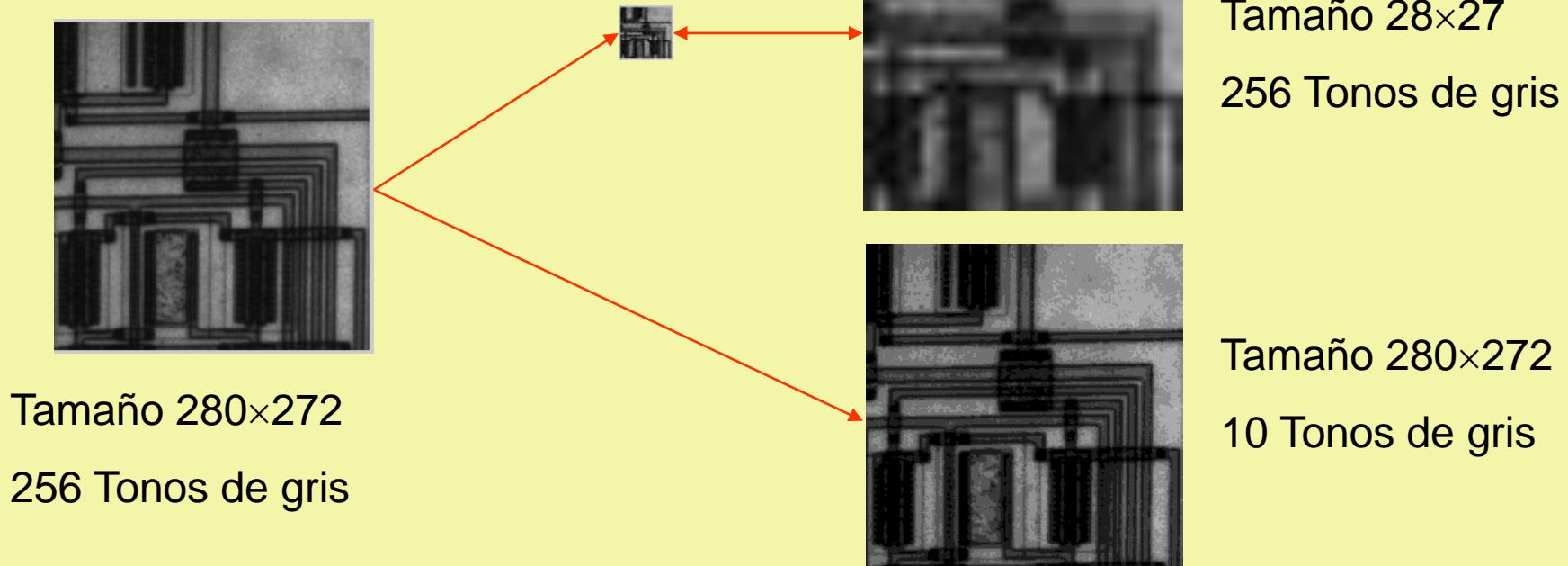
Capítulo 1. Conceptos básicos

La **resolución espacial** viene dada por la proximidad de la muestras de la imagen en el plano imagen. En el caso de imágenes digitalizadas es el número de puntos del dominio de la imagen, es decir, $M \times N$.

La **resolución espectral** viene dada por el ancho de banda de las frecuencias de la radiación luminosa capturadas por el sensor.

La **resolución radiométrica** viene dada por el número de niveles o tonos de gris distinguibles.

La **resolución temporal** viene dada por el intervalo de tiempo entre muestras consecutivas en las que la imagen se captura.



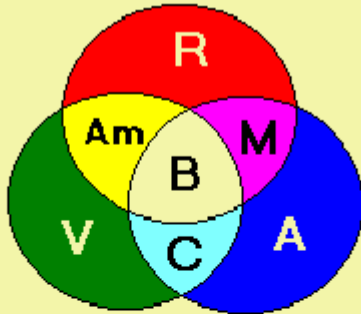
Capítulo 1. Conceptos básicos

El **color** es una propiedad de gran importancia en la percepción visual humana. El color está asociado con la capacidad de los objetos de reflejar ondas electromagnéticas de diferente longitud de onda.

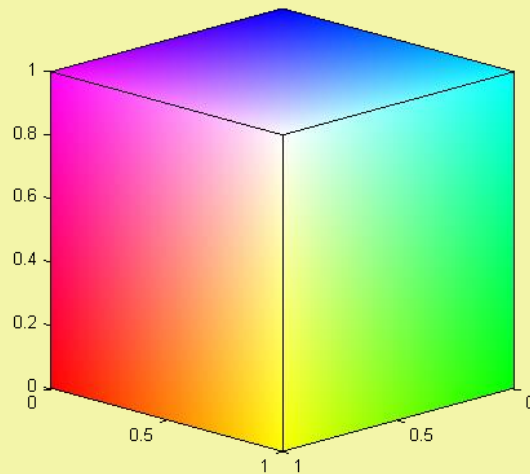
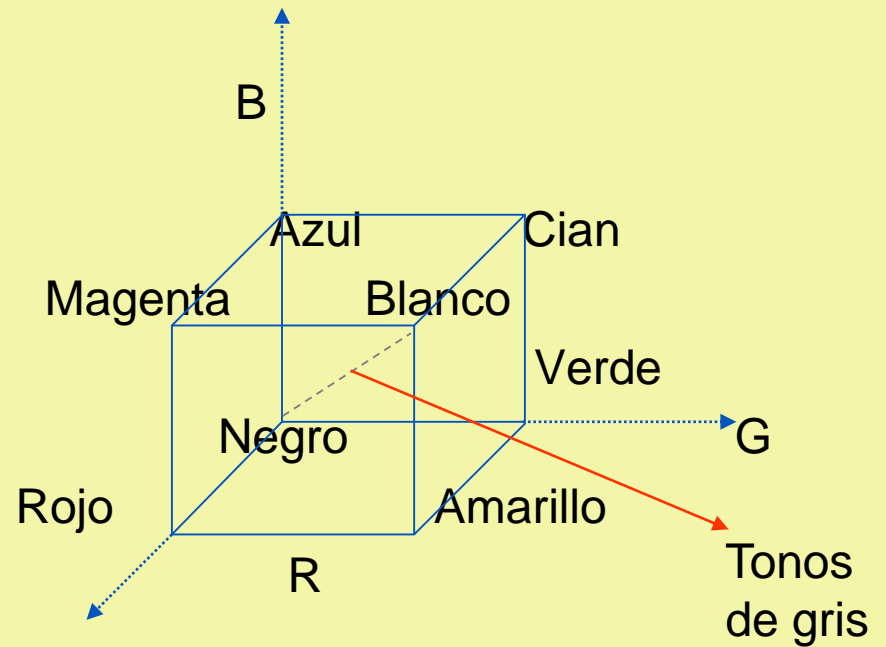
El ser humano detecta los colores como combinaciones de tres colores primarios, el rojo, el verde y el azul.

Las pantallas de nuestros ordenadores suelen utilizar el modelo **RGB** (red-green-blue), donde cada píxel tiene asociado un vector tridimensional (r,g,b) que nos conduce a un color determinado; el vector $(0,0,0)$ es el negro, (L,L,L) es el blanco, $(L,0,0)$ es el rojo “puro”, y así sucesivamente. Esto implica que tenemos L^3 colores diferentes que no todos serán albergados por el dispositivo, por lo que es común especificar un subconjunto de estos, llamado **paleta** de colores.

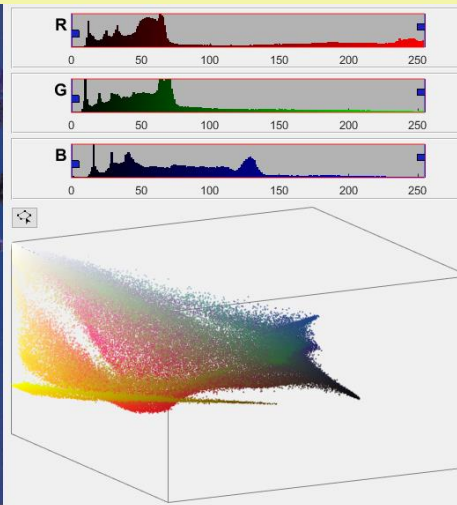
Capítulo 1. El espacio de color RGB



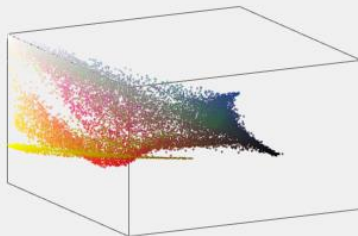
R= Rojo V= Verde A= Azul
Am= Amarillo M= Magenta C=Cyan
B= Blanco



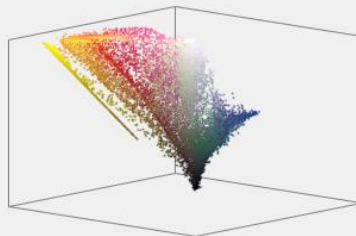
Capítulo 1. Diferentes espacios de color



RGB



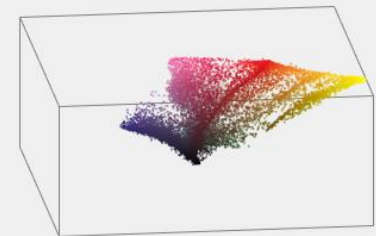
HSV



YCbCr



L*a*b*



Capítulo 1. Imagen indexada

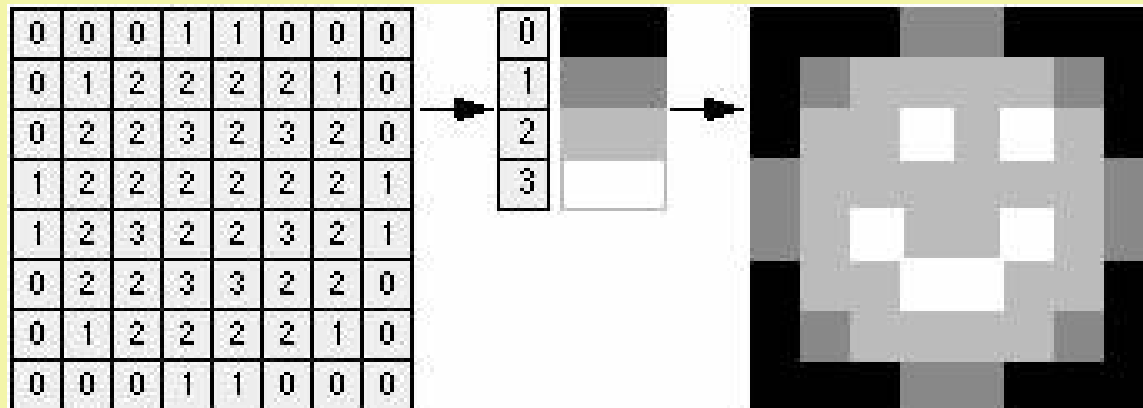
También se puede representar una imagen mediante dos componentes:

- una matriz X de tamaño $M \times N$ constituida por valores enteros del conjunto $\{1, 2, \dots, p\}$
- una matriz, llamada **mapa de colores**, de tamaño $p \times 3$, constituida por valores reales del intervalo $[0, 1]$.

El número p nos da el número de colores diferentes a utilizar.

Cada fila de la matriz mapa de colores especifica la componente roja, verde y azul de un color específico. Cada elemento de la matriz X especifica una fila de la matriz del mapa de colores, es decir, un color específico.

Una imagen así representada se dice que es una **imagen indexada**



Capítulo 1. Espacios de color

Un color se puede representar de dos maneras diferentes dependiendo del medio en que se reproduce:

a) Mediante un **sistema aditivo** que consiste en añadir colores al negro para crear nuevos colores.

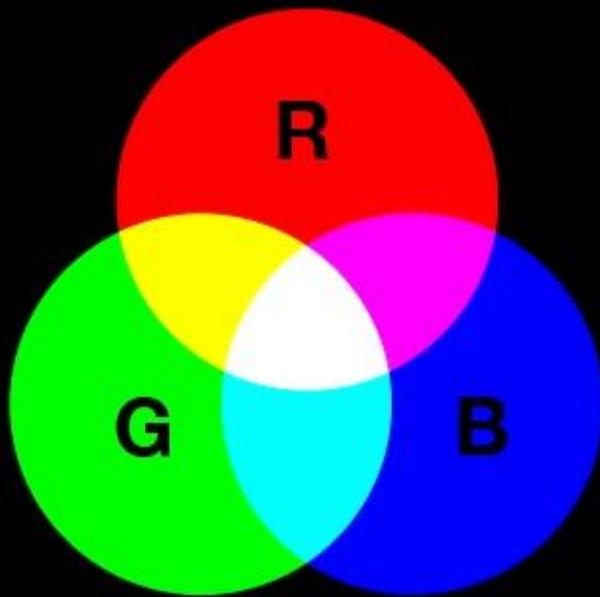
La presencia de todos los colores primarios es suficiente para crear un **blanco puro**, mientras que la ausencia de los colores primarios crea un **negro puro**. Los colores primarios de luz o colores primarios aditivos son el rojo, el verde y el azul. La suma de los tres colores primarios da lugar al blanco (la luz blanca está compuesta por luz roja, verde y azul a partes iguales). Los colores secundarios de la luz son el magenta (rojo + azul), el cian (verde + azul) y el amarillo (rojo + verde).

b) Mediante un **sistema sustractivo** en el que los colores primarios son sustraídos del blanco para formar nuevos colores.

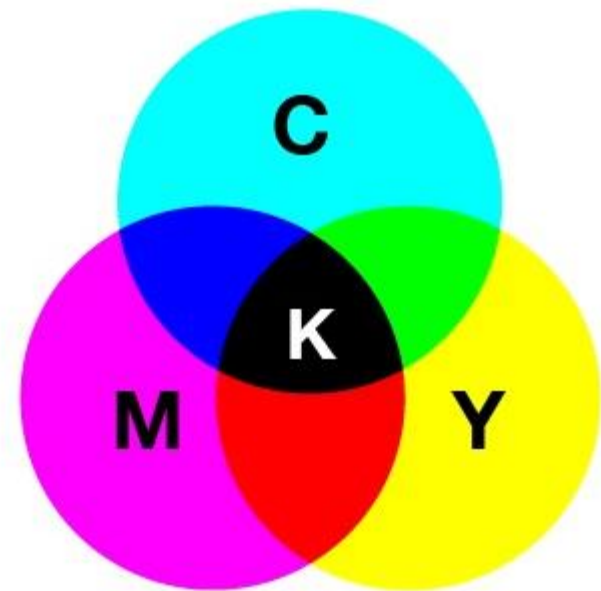
Cuantos más colores se mezclen se obtendrá algo más parecido al negro. Teóricamente, la presencia de todos los colores primarios sustractivos nos da el negro y su ausencia total el blanco. Un color primario de pigmento (sustractivo) es aquel que absorbe un color primario de luz y refleja los otros dos, por tanto, son el cian, el magenta y el amarillo. Cuando la luz incide sobre un color primario de pigmento lo que vemos es la combinación de los otros dos colores primarios de luz. Así, además del modelo RGB podemos tener otros modelos de color.

Capítulo 1. Espacios de color RGB - CMYK

Aditivo



Sustractivo



Capítulo 1. Espacio de color CMY

El espacio de color **CMY** (Cian, Magenta y Amarillo) se basa en los colores secundarios de la luz, es decir, en los colores primarios de pigmentos. La luz blanca pura se obtiene como diferencia entre el amarillo y el azul. la mayoría de los dispositivos que depositan pigmentos coloreados sobre el papel, como las impresoras o las fotocopadoras, utilizan entrada de datos en la configuración CMY o realizan una conversión de RGB a CMY mediante la transformación simple:

$$\begin{bmatrix} C \\ M \\ Y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

donde se supone que todos los colores han sido normalizados en el rango [0,1].

Teóricamente, con igual cantidad de pigmentos primarios, cian, magenta y amarillo, se producirá el color negro, pero en realidad, debido a las impurezas en las tintas, da lugar a un color café pardo; de ahí que para conseguir un color negro de calidad es conveniente añadir este color a las impresoras y fotocopadoras, pues es el que más utilizan, y se obtiene así el espacio de color **CMYK**.

Capítulo 1. Espacio de color YIQ

El sistema de color **NTSC** se usa para la televisión en Estados Unidos.

Una de las principales ventajas de este formato es que la información en tonos de gris se separa de los datos de color, y así la misma señal se puede utilizar para televisiones en blanco y negro y televisiones en color. En este sistema la imagen consta de tres componentes: La *luminancia* (Y), el *matiz* o *tonalidad* (I) y la *saturación* (Q).

- La luminancia, luminosidad o brillo es el flujo luminoso o intensidad luminosa emitido por un objeto con luz propia (emitida y no reflejada), tal como una bombilla o el Sol.
- El matiz o *tonalidad* es un atributo asociado con la longitud de onda dominante en una mezcla de ondas de luz y representa el color percibido por el observador.
- La saturación se refiere a cómo de puro es el color, es decir, la cantidad de luz blanca que se mezcla con él. Se parte del color blanco hasta llegar al color totalmente saturado



$$\begin{pmatrix} Y \\ I \\ Q \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.275 & -0.321 \\ 0.212 & -0.523 & 0.311 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

Capítulo 1. Espacio de color YUV

- El Sistema de color **PAL** (Phase Alternation Line), o modelo **YUV**, es el utilizado para la televisión en Europa. Fue desarrollado por los Laboratorios de Telefunken en Hannover (Alemania).
- En el año 1967, la República Federal Alemana y Reino Unido comenzaron a usarlo y posteriormente lo implantaron la mayoría de los países europeos que comenzaron a transmitir televisión en color, y algunos de Sudamérica.
- Las coordenadas que transmite son la *luminancia* Y , y dos coordenadas denominadas U y V , derivadas de las señales diferencia de color $R-Y$ y $B-Y$. Se obtiene del modelo RGB mediante la transformación lineal:

$$\begin{aligned}Y &= 0.30 R + 0.59 G + 0.11 B \\U &= 0.493 (B - Y) \\V &= 0.877 (R - Y)\end{aligned}$$

Las ecuaciones de conversión son las siguientes:

$$\begin{cases} Y' &= 0,299R + 0,587G + 0,114B \\ U &= 0,493(B - Y) \\ &= -0,147R - 0,289G + 0,436B \\ V &= 0,877(R - Y) \\ &= 0,615R - 0,515G - 0,100B \end{cases}$$

Esto equivale a escribir, en forma matricial:

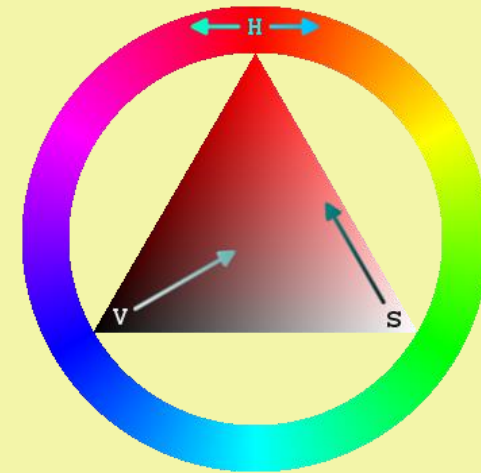
$$\begin{bmatrix} Y' \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0,299 & 0,587 & 0,114 \\ -0,147 & -0,289 & 0,436 \\ 0,615 & -0,515 & -0,100 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

Se asume que R , G y B están en el rango 0 a 1, con 0 representando la intensidad mínima y 1 la máxima. Y está en el rango 0 a 1, U está en el rango -0,436 a 0,436 y V está en el rango -0,615 a 0,615.

Capítulo 1. Espacio de color HSI-HSV

El espacio de color alternativo de más relevancia para el procesamiento de imágenes es el modelo **HSI** (**Hue, Saturation, Intensity**) o bien su equivalente **HSV** (**Hue, Saturation, Value**).

Espacio de color HSV como una rueda de color.



Los espacios de color hasta ahora estudiados no son demasiado adecuados para describir los colores en términos de su interpretación humana. Por ejemplo, nadie describe el color de su coche en términos de la composición de sus colores primarios.

El ser humano suele describir un objeto de color por su brillo, su tonalidad y su saturación. Sin embargo, el brillo es una descripción subjetiva que es difícil de medir. Por ello, en este modelo se separa el brillo de la información sobre el color (el matiz y la saturación corresponden a la percepción humana).

Este modelo es una herramienta muy útil para el desarrollo de algoritmos en procesamiento de imágenes. Existe una transformación no lineal para pasar de RGB a HSI.

Capítulo 1. Etapas

Etapas en el procesamiento de imágenes

1. Adquisición de la imagen. Puede ser: binaria, escala de grises, médica, microscópica, satélite, sintética,...
2. Mejorar la imagen de forma que aumenten las probabilidades de éxito en los procesos posteriores. Puede tratarse de : Eliminar ruidos, suavizado, realce, mejora del contraste,...
3. Dividir la imagen en sus partes sustituyentes u objetos. Segmentación.
4. Representación: Convertir los datos de los píxeles (de un contorno o una región), obtenidos tras la segmentación, en una forma adecuada para el procesamiento por ordenador.
5. Descripción: Extraer rasgos con alguna información cuantitativa de interés o que sean fundamentales para diferenciar una clase de objetos de otra. Descriptores estadísticos, geométricos y topológicos.
6. Reconocimiento: Asignar una etiqueta a un objeto basándose en la información proporcionada por los descriptores obtenidos en el apartado anterior. Interpretación: Asignar significado a un conjunto de objetos reconocidos.
7. Compresión: Reducción de la cantidad de datos necesarios para representar una imagen digital. Eliminación de datos redundantes.

Capítulo 1. Conclusiones

- El **Procesamiento Digital de Imágenes** es una materia en constante evolución en su vertiente tecnológica, tanto en capacidad de procesamiento y transmisión con crecimiento exponencial, pero también por el gran avance que han experimentado las técnicas de procesamiento digital en áreas como el filtrado, compresión y análisis de imágenes.
- Las **aplicaciones actuales del Tratamiento Digital** de Imágenes son prácticamente innumerables y prometen darnos grandes sorpresas en un futuro próximo. La televisión digital, los videojuegos, la telefonía móvil, el cine de animación, son solo algunos ejemplos del impacto que tiene actualmente en la sociedad de consumo. En el ámbito industrial y de servicios destacan las aplicaciones de visión artificial en inspección y robótica, así como el desarrollo espectacular que ha producido en la generación y tratamiento de imágenes médicas.
- Desde una **perspectiva histórica**, el Procesamiento Digital de Imágenes representa una generalización multidimensional de las técnicas de tratamiento digital de la señal unidimensional. Los orígenes del procesamiento digital de la señal se remontan al siglo XIX, aunque su desarrollo práctico no se produce hasta la década de los años 60, cuando Cooley y Tukey propusieron un algoritmo eficiente para el cálculo de la transformada de Fourier. Después, con la irrupción de los microprocesadores, se diseñaron procesadores adaptados para su cálculo. Dos circunstancias que fueron