

3. Imágenes Digitales

1. Introducción

El campo de las imágenes digitales se encuentra en continua evolución. El análisis de la percepción humana y la interpretación de las imágenes por parte de máquinas, son dos de los campos de principal aplicación de estas técnicas. Una de las primeras aplicaciones se desarrolló en Londres, cuya meta era conseguir la transmisión de imágenes a través de un cable submarino entre dicha ciudad y Nueva York. Se consiguió por supuesto, reducir el tiempo de recepción de las imágenes, ya que inicialmente este trabajo se realizaba por transporte marítimo. El problema residía en la calidad de la imagen: en la transmisión se perdía o deterioraba gran cantidad de información.

Los métodos de recepción y envío de imágenes fueron avanzando poco a poco, beneficiándose de las primeras computadoras. En 1964 se da un paso fundamental: las imágenes de la Luna transmitidas a través del Ranger 7, fueron procesadas por un ordenador para la corrección de distorsiones. Dichas imágenes mejoradas eran mostradas por televisión con una calidad envidiable (teniendo en cuenta las limitaciones en cuanto a fiabilidad de las televisiones en esa época). Estas técnicas de corrección fueron la base de gran cantidad de estudios que condujeron a la explosión de la Visión Artificial. Su empleo es muy variado y raro es no encontrar un campo que no haga uso del análisis de las imágenes.

2. Imágenes Digitales

Una imagen se define como una función $f(x,y)$ que expresa la intensidad o nivel de gris de cada uno de los puntos o píxeles que la forman; 'x' e 'y' expresan las coordenadas espaciales de referencia. En ocasiones puede resultar interesante la representación de una imagen incluyendo una tercera coordenada como medición del nivel de gris. De esta forma se pueden determinar los “picos” o valores máximos de los píxeles, así como las zonas más constantes y los valores mínimos.

Las imágenes digitales son aquellas que han sido discretizadas tanto en las coordenadas espaciales como en sus niveles de gris. Se representa igualmente por $f(x,y)$, y puede considerarse como una matriz cuyas filas y columnas identifican un punto de la imagen con valor su intensidad. Aunque la talla de una imagen digital varía de una a otra, existen ciertas ventajas en las imágenes cuadradas, con iguales valores de alto y ancho, y con niveles de gris que son potencias de 2. Un ejemplo típico es el de una televisión de 512 x 512 con 128 niveles de gris. Este tipo de imágenes pueden ser en color o monocromas;

normalmente se trabaja con las segundas, salvo análisis muy específicos que necesiten irremediamente la información del color. Aún así, las operaciones son iguales para ambos casos.

2.1. Proceso de digitalización

Una imagen natural capturada con una cámara, un telescopio, un microscopio o cualquier otro tipo de instrumento óptico, presenta una variación de sombras y tonos continua. A este tipo de imágenes, se les conoce como *imágenes analógicas*.

Para que una imagen analógica, en blanco y negro, en escala de grises, o a color, pueda ser "manipulada" usando un ordenador, primero deben convertirse a un formato adecuado. Este formato es la *imagen digital* correspondiente.

A este proceso se le denomina *digitalización* y es el primer paso en cualquier aplicación de procesamiento de imágenes digitales. Este proceso consta de dos pasos: muestreo y cuantización.

Al pasar estas señales analógicas a información digital, sucede que se hacen discontinuas y ya no guardan similitud o analogía respecto del original, ya que son números. Son discontinuas porque sólo se digitaliza a la vez un fragmento o muestra de todo el conjunto de la imagen o de la información de que se trate. No guarda similitud, ya que los números son cantidades abstractas que pueden referirse a cualquier cosa, de ahí que la información numérica de una imagen pueda ser teóricamente idéntica a la de un sonido o a la de un texto. El resultado dependerá del código y programa con que se lea dicha información, que en sí misma no es ni una imagen ni un sonido ni una palabra escrita.

2.2. Procesos iniciales en el procesamiento de la imagen

Explicar y entender una aplicación de tratamiento de imágenes digitales resulta bastante sencillo, incluso sin tener conocimientos previos del tema.

El primer paso es la *adquisición* de la imagen. Requiere un dispositivo o sensor capaz de extraer digitalmente la señal captada. En la actualidad, los dispositivos más utilizados son los sensores CCD, elementos capaces de adquirir la intensidad lumínica y transformarla en señal eléctrica cuando la luz incide directamente sobre ellos. Están formados por gran cantidad de sensores distribuidos en forma de parrilla. Las características del sensor y la imagen que produce son elementos determinantes en la aplicación.

Una vez adquirida la imagen, el siguiente paso es el *preprocesado*. El objetivo fundamental es preparar la imagen para la obtención posterior de información de manera más eficaz y correcta. Las técnicas más importantes y las más utilizadas son realce del contraste, reducción de ruido y aislamiento de regiones con características semejantes.

El siguiente paso es la *segmentación*, a menudo definida como la partición de una imagen en sus componentes u objetos constituyentes. En general, la segmentación es una de las partes más complejas del procesado de una aplicación, pero si se realiza cuidadosamente, puede conducir a soluciones más que exitosas; también es cierto que una pequeña errata genera información falsa, muchas veces enmascarada con datos aparentemente válidos.

La salida de la operación anterior es una lista de datos de los píxeles que forman el contorno de un objeto, una región. En todos los casos, convertir la información extraída a datos fácilmente interpretables se hace imprescindible, no sólo para la comprensión por parte de la persona que realiza el desarrollo, sino también para completar correctamente la aplicación.

Con la *descripción o selección de características* se eligen los datos que se van a extraer, clave para diferenciar unos objetos de otros.

El paso más largo es el *reconocimiento*; trata de establecer “etiquetas” para cada una de los componentes extraídos basándose en las características establecidas.

La *interpretación* consiste en dar el significado correcto de cada una de las estructuras conseguidas en los pasos anteriores.

Pero más allá de lo expuesto hasta ahora, es necesario el análisis profundo de lo que se llama *conocimiento base*. No es más que aquellas particularidades conocidas previamente y que se suelen presentar en imágenes adquiridas en ambientes ya estudiados. Este conocimiento y las operaciones explicadas anteriormente deben interactuar entre sí. Como se observa en la figura 3.1, el centro de toda la aplicación es dicho *conocimiento base*. A partir de las relaciones que establece con el resto de pasos, se obtiene la respuesta final. Por supuesto, la obtención de los resultados no obliga a realizar todas las fases; en ocasiones, basta un pequeño procesado para obtener la información deseada de forma correcta.

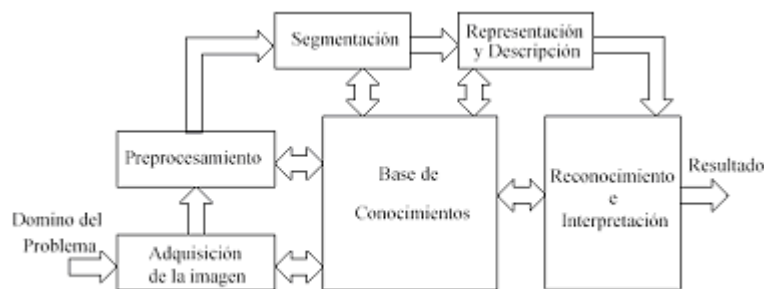


Figura 3.1: Pasos fundamentales en procesado de una imagen digital

2.3. Conceptos principales

2.3.1. Profundidad de bits

La profundidad de bits es determinada por la cantidad de bits utilizados para definir cada píxel [5]. Cuanto mayor sea la profundidad de bits, tanto mayor será la cantidad de tonos (escala de grises o color) que puedan ser representados, y por tanto, más exacta será la representación del color en la imagen digital. Las imágenes digitales se pueden producir en blanco y negro, a escala de grises o a color.

De aquí, se concluye que:

- ◆ Si la profundidad es de 1 bit, sólo existe la posibilidad de tener 2 colores.
- ◆ Si la profundidad es de 2 bits, es posible tener 4 colores (00 / 01 / 10 / 11).

- ◆ Si la profundidad es de 3 bits, es posible tener 8 colores, etc.

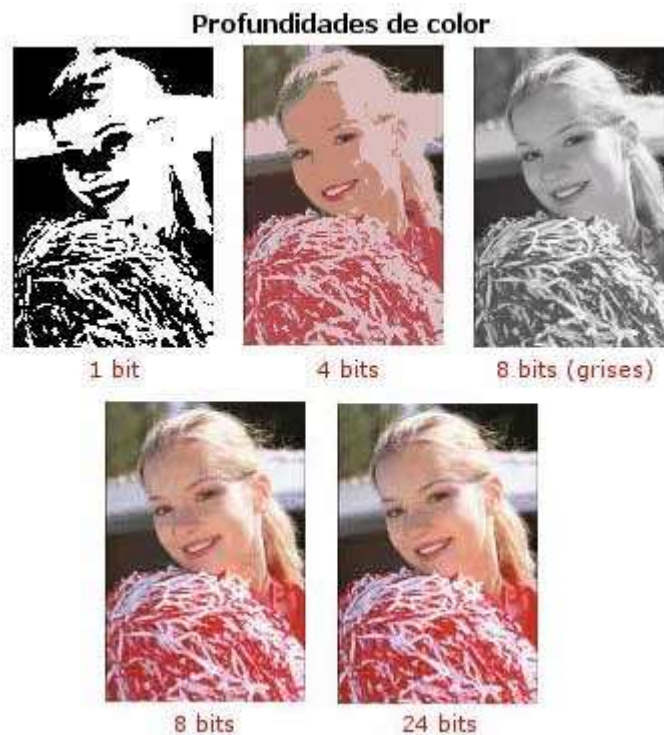


Figura 3.2: Imágenes con diferente profundidad

En artes gráficas se usan, principalmente, ficheros:

- ◆ De un **1 bit para imágenes de línea**, sólo blanco y negro sin grises medios.
- ◆ De **8 bits para imágenes de tono continuo** en B/N, con 256 matices de color.
- ◆ De **24 bits (8 por cada canal) para imágenes a todo color.**
- ◆ De **32 bits (8 por canal, cman) para ficheros de impresión.**

2.3.2. Resolución

El concepto de resolución está relacionado con el detalle de que una imagen puede ser impresa o visualizada por otros medios [5]. Se mide de la siguiente manera:

- En imágenes vectoriales, la resolución de la imagen depende del periférico que se use para imprimirla o verla.
- En imágenes de bits, es el número de píxeles que hay por unidad lineal, cm o pulgada.

Como la cantidad de píxeles de una imagen no crece cuando la estiramos mediante un programa de dibujo, a mayor tamaño, menor resolución de un mismo fichero.

Si el fichero se estira demasiado, los píxeles pueden llegar a hacerse visibles, lo cual sólo es aceptable cuando se trata de crear efectos gráficos especiales. Por ello, es necesario crear ficheros que tengan la resolución adecuada al tipo de impresión que se vaya a usar.

2.3.3. Rango dinámico

El rango dinámico es el rango de diferencia tonal entre la parte más clara y más oscura de la imagen [5]. Cuanto más alto sea el rango dinámico, más matices se podrán representar, a pesar de que el rango dinámico no se correlaciona en forma automática con la cantidad de tonos reproducidos.



Figura 3.3: Imágenes con diferente rango dinámico

En este ejemplo, se puede observar que la imagen de la izquierda posee un rango dinámico más amplio, pero una cantidad limitada de tonos representados (hay falta de detalles en las sombras, y en los toques de luz). En cambio, la imagen de la derecha, posee un rango dinámico más estrecho, pero una cantidad mayor de tonos representados.

2.4. Tipos de imágenes digitales

Según el modo en el que se forma una imagen, y cómo se guarda esa información, existen básicamente dos sistemas [6]:

- ◆ **Imágenes vectoriales.** El aspecto de las imágenes creadas con elementos vectoriales es el de formas, regulares o irregulares, que contienen color plano, color en degradados definidos matemáticamente, o texturas. Los elementos constituyentes del vector, en una imagen vectorial, son las curvas de Bézier.

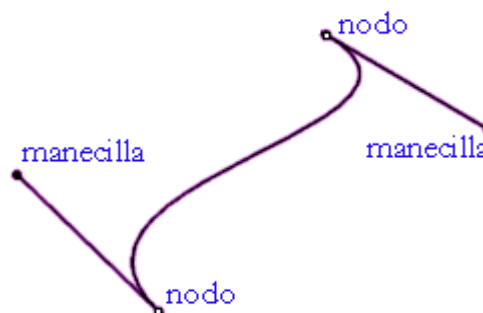


Figura 3.4: Curva de Bézier

Una curva Bézier se define por cuatro puntos. Los puntos inicial y final de la curva -nodos o puntos de anclaje- y dos puntos de control –manecillas o manejadores-, estos últimos como sus nombres lo indican, sirven para definir la forma de la curva y no aparecerán en la imagen final. Así, para modificar la curva sólo se tiene que mover alguno de los nodos. Estas curvas son fáciles de manejar, muy estilizadas y versátiles, ya que adoptan tanto curvaturas muy suaves (casi líneas rectas), como muy fuertes (curvas complejas). Además, pueden adaptarse a infinitud de formas. El diseño de iconos y logotipos es uno de los usos más frecuentes que se les dan. Pero, también, son muy útiles en el diseño y manipulación de fuentes de texto.

- ◆ **Mapa de bits.** Las imágenes de mapa de bits están formadas por una rejilla de celdas. A cada una de las celdas, se le asigna un valor de color y luminancia propios. Por esto, cuando vemos el conjunto de celdas, tenemos la ilusión de una imagen de tono continuo. Una forma importante de clasificar las imágenes de mapa de bits es según la cantidad y tipo de información que se asigne a cada píxel.

Un fichero de mapas de bits consta de las siguientes partes: una *cabecera* y los datos como mapa de bits. También es posible que se almacene una paleta de colores, para el caso de imágenes indexadas, que suele estar después de la cabecera y si el fichero contiene varias imágenes, puede existir un mapa de bits que contenga las direcciones, en forma de desplazamientos, donde están los mapas de bits de cada imagen. La cabecera contiene información sobre la imagen, ofreciendo datos como el tamaño de la imagen (número de filas y número de columnas), formato de almacenamiento (bmp, jpeg, gif, tiff...), etc. y en el mapa de bits está la imagen en sí, almacenada por filas o columnas o por planos de bits.

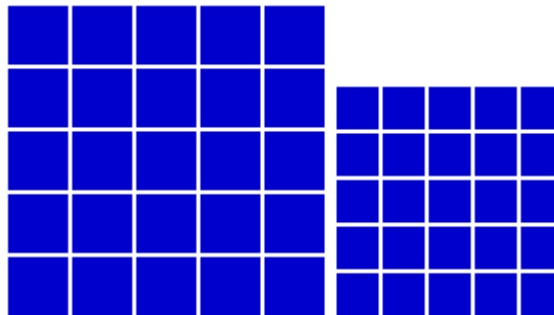


Figura 3.5 : Rejillas de 5 píxel

Después de apreciar las diferencias básicas entre los tipos de codificación, se podría plantear qué tipo de imagen es más conveniente. A pesar de las diferencias, no hay una codificación de imagen mejor una que otra. Cada una es diseñada para adaptarse a un tipo diferente de imagen. Los mapas de bits ofrecen sus mejores prestaciones con imágenes fotográficas y los gráficos vectoriales con imágenes de gestión (presentaciones, estadísticas, etc.). Las imágenes fotográficas normalmente tienen multitud de colores y cambios de tono irregulares. Utilizando mapas de bits, estas características no tienen ningún problema, pero si para los gráficos vectoriales debido a la enorme dificultad de identificar y delimitar los objetos que forman una imagen. En la mayoría de los casos es casi imposible conseguir un resultado con la calidad parecida a la de un mapa de bits.

Los formatos de de almacenamiento más usados son:

- **BMP** - Formato de fichero gráfico BitMap muy popular en ordenadores Windows. La representación de la imagen consiste en filas y columnas de puntos. El valor de cada punto se almacena mediante uno a más bits de datos. A más bits utilizados

para representar un punto, más colores podrán ser mostrados. La densidad de los puntos (Resolución), determina la nitidez de la imagen que se expresa normalmente en puntos por pulgada (DPI), o simplemente por el número de filas y columnas (ej. 640x480). Es un formato de fichero **no comprimido**.

- **JPEG** (o extensión JPG) - (“Join Photographic Experts Group”) - Es un formato de compresión gráfica de imágenes mas usado. En él se definen tres niveles de procesamiento: Codificación básica, extendida y sin pérdidas. El formato JPEG comprime la imagen segmentándola en pequeños bloques de forma que se obtenga una óptima relación entre calidad y dimensión del fichero, haciéndolo altamente aconsejable para la definición de imágenes fotográficas.
- **GIF** - Siglas de “Graphics Interchange Format”. GIF es un formato de fichero gráfico que utiliza un esquema de compresión desarrollado por *CompuServe*. Es un formato de imagen de 8 bits por píxel de profundidad de color, lo que significa que el número máximo de colores es de 256, aunque podemos reducir este número. Fue el primero que se estandarizó para su uso en Internet y es el más utilizado en la World Wideeb. Además este tipo de ficheros goza de dos importantes variantes gráficas tales como la transparencia de uno de los colores y la animación de la imagen. Es aconsejable para todas las imágenes excluyendo las fotográficas (por su elevado numero de colores).
- **TIFF** (o extensión TIF) - “Tagged Image File Format” o Formato de archivo de imagen etiquetada - Formato gráfico de representación de imágenes sin pérdidas desarrollado por Aldus como formato estándar internacional, compatible con múltiples plataformas, y de excelente relación entre ocupación de espacio y calidad.

Además de estos formatos, existen multitud de tipos de ficheros gráficos, algunos genéricos y otros usados por una única (y normalmente muy conocida) aplicación. Algunos de los más extendidos son los formatos PNG, TGA, PSD, DWG, PCX, PIC, EPS, ó VMF.

2.4.1. Representación de imágenes digitales

1. Imágenes binarias

Son las imágenes más simples. Se caracterizan porque cada píxel sólo tiene dos valores posibles, 0 (negro) y 1 (blanco). También se les conoce como imágenes de 1bit/píxel.

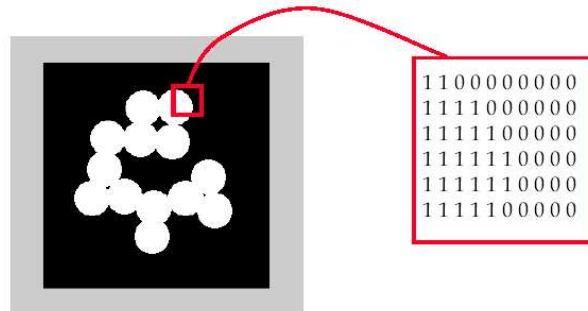


Figura 3.6: Imagen binaria

Los tonos intermedios (grises) se crean ordenando los píxeles blancos y negros, para simular gris, es lo que se llama técnica de difuminado. Se usan donde se requiere forma o contorno. Estas imágenes se obtienen a partir de imágenes en escala de grises, mediante umbralización.

2. Imágenes en escala de grises

En este tipo de imágenes, cada píxel puede tener uno de los 256 valores diferentes, desde el 0 (Negro) hasta el 255 (Blanco) (se utilizan 8bits/píxel). Esto es así porque contienen sólo información de brillo (no de color).

Este tipo de datos muestra suaves cambios de tonalidad usando valores intermedios de gris.

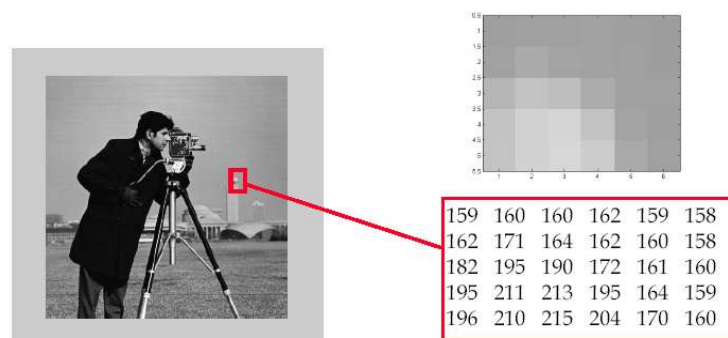


Figura 3.7: Imagen en escala de grises

3. Imágenes en color

Este tipo de imágenes, se modelan como tres imágenes monocromáticas. La representación de estas imágenes es a través de tres matrices, cada una conteniendo la componente roja (R), verde (G), y azul (B) respectivamente, como si fueran tres imágenes de intensidad independientes.

El color para un determinado píxel en la posición (x,y) será: $[R(x,y) \ G(x,y) \ B(x,y)]$.

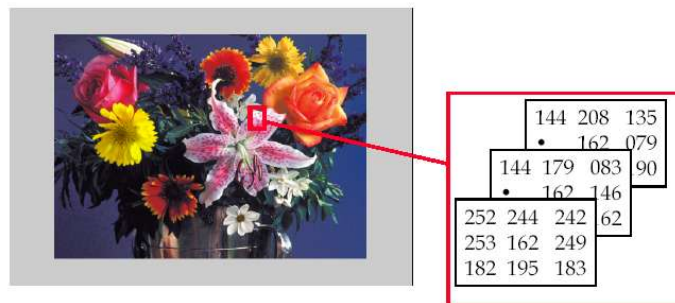


Figura 3.8: Imagen en color

Normalmente bastan 8 bits para definir el grado de intensidad de cada color en cada uno de los píxeles. Por lo tanto la información del color relativa a cada píxel se almacenaría en 3 bytes (uno por cada color, se dice codificación en 3×8 bits = **24 bits**), por lo que una imagen de 640×480 píxeles requeriría $640 \times 480 \times 3$ bytes = 921600 bytes = 900 KBs de memoria. Esta codificación resulta excesiva ya que gran mayoría de dispositivos de adquisición de imágenes no tiene la sensibilidad suficiente para aprovechar esta capacidad de información, por lo que para cada píxel, los 2 ó 3 bits menos significativos de cada componente de color son ruido. Por tanto sólo conservaríamos 5 bits para cada color, 15 bits en total, codificables por 2 bytes (de ahí el nombre: color de 2×8 bits = **16 bits**).

Este formato es soportado por la mayoría de las aplicaciones de tratamiento gráfico actuales, pero es inadecuado para el tratamiento digital de la imagen, porque se disocia la relación unívoca byte/componente de color (se tiene tres componentes de color y sólo 2 bytes). Por ello las aplicaciones suelen convertir el formato de 16 a 24 bits para trabajar con este tipo de imágenes. Con estos 24 bits, tenemos tres imágenes de intensidad monocromas (rojo, verde y azul), cada una codificada por 8 bits. A pesar de este sistema de 24 bits para el tratamiento de imágenes, existen casos en los que es insuficiente para medir pequeñas variaciones en algunas imágenes, por ello existen formatos de hasta **36 bits**.