

Capítulo 2

2. Imagen Digital



Ilustración 1 Representación visual de una imagen digital a través de una escala de niveles de gris

2.1 Percepción visual

Debido a que la presente tesis se concentra en el área de imágenes digitales, es conveniente recordar que la mayoría de las imágenes digitales representan escenas naturales continuas (Ilustración 10). Las excepciones son imágenes digitales artificiales tales como pruebas de patrones que son creadas numéricamente en la computadora e imágenes construidas por sistemas tomográficos.

Es por tanto importante entender la física de la formación de una imagen por sensores y sistemas ópticos incluyendo la percepción de la visión humana, ya que en las aplicaciones del procesamiento de imágenes, uno de los objetivos es mejorar la calidad visual de las

mismas adaptándola lo mejor posible a la capacidad perceptiva del sistema de visión humana (Ilustración 11).



Ilustración 2 El ojo humano

Los humanos son criaturas principalmente visuales. No todos los animales dependen de sus ojos como nosotros lo hacemos para más del 99% de la información recibida acerca de nuestro entorno. Los murciélagos emplean sonidos de alta frecuencia, los gatos tienen una pobre visión pero su sentido del olfato está muy desarrollado, las serpientes localizan a sus presas por las emisiones de calor, los peces tienen órganos que detectan (y en algunos casos generan) campos eléctricos. Incluso los pájaros, quienes tienen un sentido de la vista altamente desarrollado, tienen una configuración visual diferente a la nuestra. Como otro ejemplo los búhos tienen sus ojos en los lados opuestos de su cabeza proporcionando así casi 360 grados de cobertura en la visión, ellos presentan de cuatro a cinco diferentes receptores de color, en cambio nosotros los seres humanos tenemos solo tres: el rojo, verde y azul. (González, 2002)

2.1.1 Elementos de la percepción visual

Percepción de la luz.

La luz es una energía radiante la cual, al incidir sobre los órganos de la visión les permite efectuar la función de la visión (Pratt, 2001).

Se conoce a la luz como una forma de radiación electromagnética, cuya longitud de onda pertenece a una porción del espectro electromagnético en la cual la retina del ojo humano exhibe una respuesta. La gama normal de longitud de onda de visión humana, oscila entre 390 nm y 780 nm aproximadamente. Debido a su dualidad onda-partícula la luz tiene simultáneamente propiedades tanto de onda como de partícula.

La luz entrante en el sistema de visión humano puede ser originado de un ente luminoso o bien de la luz reflejada sobre algún objeto o de la luz transmitida a través de un objeto translucido.

Si $E(\lambda)$ representa la distribución espectral de energía de la luz emitida y $t(\lambda)$ y $r(\lambda)$ denotan la longitud de onda transmitida y la reflejada de un objeto, respectivamente. Para el objeto que la transmite, la distribución espectral de la energía de luz observada es:

$$C(\lambda) = t(\lambda)E(\lambda) \quad (\text{II.1})$$

Mientras que para el objeto que la refleja la distribución espectral es:

$$C(\lambda) = r(\lambda)E(\lambda) \quad (\text{II.2})$$

Existen tres descriptores comunes de la sensación de luz: brillantez, color (matiz o tono) y saturación. Si observamos dos fuentes generadoras de luz, con la misma forma espectral, el recurso que hace aumentar la intensidad física para ser percibida es la brillantez.

El atributo de la luz que nos hace distinguir una luz roja de una luz verde o amarilla, es llamado el color o matiz o tono.

El tercer descriptor de la luz es la saturación, este atributo nos permite distinguir de una luz espectral (tono puro) a una luz pastel (tono pastel).

Los colores puros del espectro (tonos puros o luz espectral) están completamente saturados. En cambio los colores como el rosa (rojo + blanco) y el lavanda (azul + blanco) están menos saturados y corresponden a tonos pastel (luz pastel).

Los tonos pastel se encuentran menos saturados, siendo su grado de saturación inversamente proporcional a la cantidad de luz blanca añadida.

2.1.2 Formación de la imagen en el ojo

Estructura del ojo humano

El ojo humano puede considerarse como una disposición de doble lente positiva que forma una imagen real en una superficie fotosensible, al incidir la energía luminosa sobre el ojo humano, éste la transforma en señales que son recibidas e interpretadas por el cerebro. La Ilustración 12 muestra un corte simplificado del ojo humano.

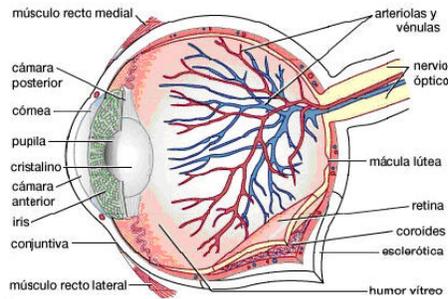


Ilustración 3 Corte horizontal del ojo humano

El ojo humano es una masa gelatinosa casi esférica contenida dentro de una estructura dura y flexible, la **esclerótica**.

El ojo es parecido a una esfera, con un diámetro promedio de 24 mm. Al nacer, el tamaño promedio del ojo es de 16 mm aproximadamente y crece hasta aproximadamente 24 mm de diámetro en la edad adulta.

La **esclerótica** es una membrana blanca y opaca la cual está conectada en forma contigua por la cornea en la parte frontal del ojo.

Sobresaliendo del cuerpo de la esfera está la **cornea** la cual contiene vasos por los que no circula sangre y es óptimamente clara, permitiendo que la luz entre a través de la parte frontal del ojo, la cornea sirve como primer y más fuerte elemento convexo del sistema de lentes.

La más interior de las membranas es la **retina**, la cual bordea el interior de la porción posterior de la pared. La retina convierte los estímulos luminosos en señales eléctricas que son transmitidas al cerebro a través del **nervio óptico**, el cual contiene aproximadamente un millón de fibras nerviosas.

Al interior de la envoltura esclerótica se halla una envoltura interna, la **coroides**, que es una capa oscura provista de una red de vasos sanguíneos que funcionan como la principal fuente de nutrición del ojo. La coroides se encuentra adyacente a la retina y conectada a la esclerótica. La coroides es muy pigmentada y por tanto ayuda a reducir la cantidad de luz innecesaria que entra al ojo.

En su extremo anterior, la coroides está dividida en el **cuerpo ciliar** y el **iris**. El segundo se contrae o expande para controlar la cantidad de luz que entra al ojo. La apertura central del iris (la pupila) varía en diámetro, desde 2 a 8 milímetros aproximadamente. El frente del iris contiene el pigmento visible del ojo, mientras que la parte posterior contiene un pigmento negro.

El lente está hecho de capas concéntricas de células fibrosas y está suspendido por fibras asociadas al cuerpo ciliar. Éste tiene una composición de 60 a 70 por ciento de agua, cerca de 6 por ciento de grasa y más proteína que cualquier otro tejido en el ojo. El lente es coloreado por una ligera pigmentación amarilla que se incrementa por el paso de los años. En casos extremos, el nublamiento excesivo del lente, causado por la aflicción comúnmente llamada como cataratas, puede provocar incapacidad para diferenciar los colores y la pérdida de una visión clara. El lente absorbe un 8 por ciento del espectro de luz visible y una relativamente más alta absorción en longitudes de onda más cortas. Tanto la luz infrarroja como la ultravioleta son absorbidas considerablemente por las proteínas contenidas en la estructura del lente, pero en cantidades excesivas pueden dañar al ojo.

Al frente del ojo se encuentra el **crystalino**, el cual se encuentra compuesto de aproximadamente 70% de agua y absorbe el 10% de la luz que incide sobre él. El cristalino forma imágenes sobre la retina de objetos que se encuentran de frente al ojo desde distancias que van desde 25 cm hasta el infinito.

Localizada detrás de la córnea se encuentra la **cámara anterior**, la cual contiene el **humor acuoso** que provee nutrientes a la córnea. En el centro del ojo, se encuentra el **humor vítreo**, un fluido claro que provee de nutrientes al ojo.

La visión es posible en gran medida debido a la retina. La retina ocupa un área de aproximadamente 1000 mm², y está compuesta de aproximadamente 100 millones de sensores. Cada sensor convierte sólo una porción de la luz radiante de la imagen que incide en la retina en una señal eléctrica que es procesada por el cerebro.

Los sensores o receptores que contiene la retina son de dos tipos: los **conos** y los **bastones**.

Los sensores de tipo bastón proveen una visión burda de una imagen aunque sobre un campo de vista muy amplio, en cambio los conos proveen una buena resolución de la imagen dentro de un campo de vista estrecho.

El número de conos en cada uno de los ojos es de 6 a 7 millones. Se localizan ante todo en la porción central de la retina, llamada la fovea, y son altamente sensibles al color. Los humanos pueden resolver detalles finos en gran parte con los conos porque cada uno de éstos se encuentra conectado a su propia terminal nerviosa. Los músculos controlan la rotación del ojo hasta que la imagen de un objeto de interés cae en la fovea. La visión de cono es llamada fotópica o visión de luz brillante.

El número de bastones es mucho más grande; de 75 a 150 millones de éstos se hallan distribuidos sobre la superficie de la retina. El área de distribución más amplia y el hecho de que los bastones estén conectados a un solo extremo del nervio reduce la cantidad de detalles discernibles por estos receptores. Los bastones sirven para dar una imagen general, en conjunto, del campo de visión. Ellos no están involucrados en la visión a color y son sensibles a bajos niveles de iluminación. Por ejemplo, objetos que aparecen con viveza coloreados ante la luz del día cuando son vistos con luz de luna aparecen como formas sin color porque sólo los bastones son estimulados. Este fenómeno es conocido como escotópico o visión lumínica débil.

La Ilustración 13 muestra la densidad de bastones y conos de una sección transversal del ojo derecho pasando a través de la región de aparición del nervio óptico del ojo.

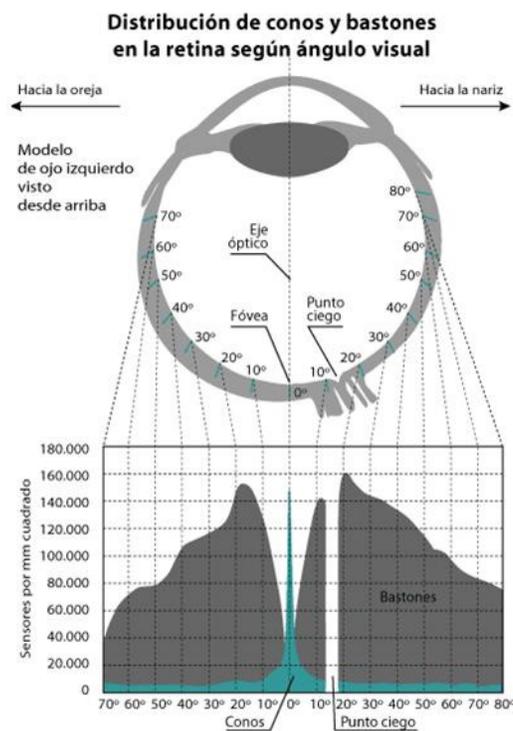


Ilustración 4 Densidad de conos y bastones en el ojo humano

En la Ilustración anterior podemos observar una ausencia de conos y bastones a aproximadamente 20° de la fovea con respecto al lugar donde se encuentra el nervio óptico. La ausencia de receptores en esta área se explica por la llamada mancha ciega.

La escasez de sensores ópticos en este punto produce un **punto ciego** en el campo de vista del ojo.

Para reducir el efecto del punto ciego, el ojo desarrolla micros rastreadores sobre una imagen a una tasa más rápida que la tasa de persistencia del ojo, produciendo una imagen que es percibida por el cerebro como si no contuviera un punto ciego.

La principal diferencia entre el lente del ojo y un lente óptico ordinario es que el primero es flexible. El radio de curvatura de la superficie anterior del lente es más grande que el radio de su superficie posterior.

Los objetos situados dentro del campo de vista del ojo son proyectados sobre la retina a través de la córnea y el cristalino.

La forma del lente es controlada por la tensión en las fibras del cuerpo ciliar. Para enfocar objetos distantes, los músculos controladores causan que el lente sea relativamente allanado. Asimismo, estos músculos permiten al lente volverse espeso de acuerdo al enfoque que se haga sobre objetos cercanos al ojo, el radio de curvatura del cristalino se incrementa automáticamente cuando se ajusta a objetos que están lejos o a objetos que se encuentren cerca.

La distancia entre el centro del lente y la retina (llamada longitud focal) varía de 17 a 14 mm aproximadamente, ya que la potencia refractiva de los lentes se incrementa a partir de su mínimo hasta su máximo. Cuando el ojo enfoca sobre un objeto más lejano que 3 metros, el lente muestra su potencia refractiva más baja. Cuando el ojo enfoca sobre un objeto cercano, el lente es altamente refractivo.

El rango dinámico del ojo es de aproximadamente 10 incrementos en magnitud, variando desde los niveles de luz escotópica a los niveles elevados de luz que producen un brillo deslumbrante. Sin embargo, el ojo no puede observar todo ese rango a la vez ya que se adapta a rangos pequeños de niveles de luz.

El punto de adaptación en el cual el ojo opera es determinado, tanto por el tamaño de apertura del iris, el cambio en la actividad neuronal, la pérdida de color y por la regeneración de los pigmentos receptores.

La inhibición de la actividad neuronal (la inhibición de las señales eléctricas recibidas e interpretadas por el cerebro) se efectúa cuando una porción de la retina es estimulada. En este punto, la región de la retina que ha sido estimulada y las regiones adyacentes son inhibidas, resultando en un cambio en la brillantez percibida. El resultado de un estímulo de luz constante sobre la retina ocasiona un cambio químico por el proceso de pérdida de color y

regeneración de conos y bastones. Este proceso toma aproximadamente un minuto y conlleva una menor sensibilidad de los receptores que están ubicados sobre la retina.

La brillantez que el ojo percibe es una función de la intensidad de luz incidente sobre el ojo. La variación de la brillantez percibida sobre un rango de luz varía desde la luz escotópica hasta niveles altos en los que la luz es deslumbrante.

Las imágenes digitales son un conjunto discreto de intensidades, por ello es importante conocer la capacidad del ojo para distinguir entre niveles de intensidad diferente.

La brillantez percibida es una función logarítmica de la intensidad de luz incidente sobre el ojo, por lo tanto la variación en la brillantez va desde la luz escotópica hasta niveles altos en los que la luz es extremadamente deslumbrante.

También es interesante la capacidad del ojo para discriminar entre cambios en intensidad de luz en cualquier adaptación específica.

2.2 Imagen

La palabra imagen proviene del latín *imago*, y se considera que es una representación de un objeto mediante técnicas diferentes de diseño, pintura, fotografía y video.

Su significado depende del campo científico que se trate, así:

- En matemáticas, una imagen de una función f es el conjunto de elementos y para los cuales existe un elemento x tal que $y = f(x)$.
- En física la imagen óptica de un objeto es la proyección de los puntos de la superficie visible del objeto sobre un plano.
- En informática puede tener dos significados:
 - una imagen puede ser un archivo codificado que, al abrirlo, muestra una representación visual de algo (ya sea fotografía, gráfica, dibujo, etc.)
 - también puede ser un archivo en el cual está toda la información de una unidad (disco duro, CD, DVD). Este archivo se puede usar como "molde" para configurar nuevos computadores sin tener que realizar todos los procesos de instalación de un sistema nuevo y a la vez sirve como copia de seguridad de los datos de ese ordenador antes de realizar alguna operación sobre él.

Para esta tesis se definirá como imagen a una señal bidimensional, es decir, una función de dos variables independientes $f(x, y)$, las cuales generalmente se encuentran representando una intensidad de radiación electromagnética. Las variables (x, y) representan las dos coordenadas y el valor o amplitud $f(x, y)$ proporciona la intensidad o brillo de la imagen en el punto (x, y) .

Generalmente se referencia el punto (x, y) como el píxel (acrónimo del inglés: “Picture Element”) y lo evaluamos en $f(x, y)$ como el nivel de gris de una imagen f del píxel en la localidad (x, y) .

Las imágenes se pueden clasificar en dos categorías generales: continuas y discretas. Una imagen continua es considerada como una función de dos variables independientes, cuyo rango está sobre valores continuos. Por ejemplo, la intensidad de una imagen fotográfica grabada en un film es una función bidimensional $f(x, y)$ de dos variables independientes x y y de valores reales.

Una imagen discreta es considerada una función de dos variables independientes, cuyo intervalo se encuentra en valores enteros. Por ejemplo, la intensidad de una fotografía discretizada de $M \times N$ pixeles, es una función bidimensional $f(m, n)$ de dos valores discretos independientes m y n , donde m, n toman valores desde 1 hasta M, N respetivamente. En el siguiente capítulo se precisará el concepto de Imagen Discreta.

2.3 Tipos de imágenes

De acuerdo con John Goutsias, Profesor del centro de Ciencias de la Imagen y Departamento de Ingeniería eléctrica y en Computación de la universidad John Hopkings, existen tres tipos de imágenes dependiendo de los valores que toman sus pixeles al momento de realizar una representación visual (Goutias 2001):

Imágenes binarias

Son el tipo de imágenes más simples, ya que pueden tomar solo dos valores típicamente negro y blanco, ó 0 y 1.

Estas imágenes toman solo 1 bit por píxel para representar la imagen, así una imagen de tamaño 256×256 se encuentra representada por solo 65 536 bits de información

Estas imágenes son las más frecuentemente usadas en aplicaciones de visión computacional donde se requiere solo de datos específicos para representar información de forma y contorno o perfil. Las imágenes binarias (Ilustración 14) son a veces creadas de imágenes en escala de grises después de aplicar un proceso llamado umbralización, en el cual cada valor de gris sobre o igual al umbral toma el valor de 1 y aquellos valores que se encuentran bajo el umbral toman un valor de 0.

Ejemplos:



Ilustración 5 Ejemplos de imágenes binarias

Imágenes en escala de grises

Estas imágenes contienen solo información de brillo y no de color, típicamente son imágenes que están representadas por 256 niveles de gris. Estas imágenes toman 8 bits por píxel para representar una imagen, por ejemplo, una imagen de tamaño 256 x 256 es representada por 524 288 bits de información (Ilustración 15).

Esta representación provee mayor y más adecuada resolución de brillantez en términos de lo que el sistema de visión humana percibe.



Ilustración 6 Ejemplos de imágenes en escala de grises

Imágenes de color

Estas imágenes pueden ser modeladas como el conjunto de tres bandas de imágenes monocromáticas, donde cada banda de datos corresponde a un color fundamental diferente. Típicamente las imágenes de color son representadas con rojo, verde y azul, o bien RGB.

En esta representación se emplea un modelo en escala de grises de 8-bits por banda; así entonces, una imagen RGB es representada por $3 \times 8 = 24$ bits por píxel. Así por ejemplo una imagen en color de tamaño 256×256 es representada por 1 572 864 bits de información (Ilustración 16).

Cuando procesar una imagen a color no es apropiado a una aplicación específica, se usan algoritmos de procesamiento de imágenes en escala de gris, para procesar cada banda independientemente, debido a que existe una alta correlación entre las tres bandas.

Ejemplos:

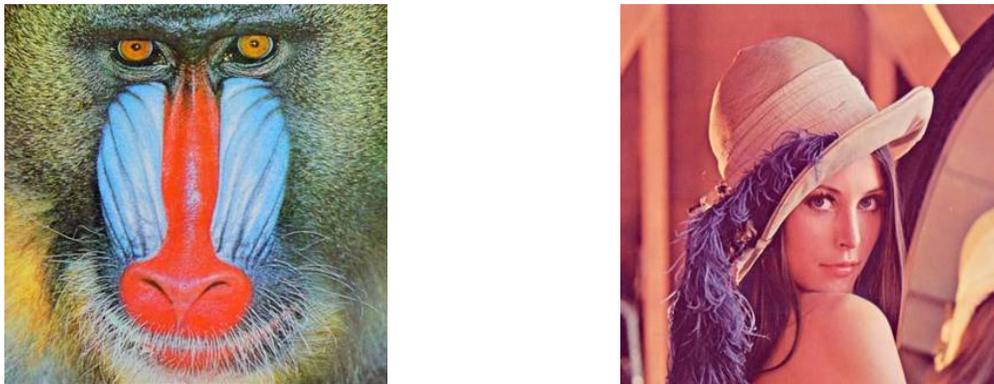


Ilustración 7 Ejemplos de imágenes en color

2.4 Imagen digital

Tal como vimos anteriormente (véase sección 2.2) el término “imagen monocromática” o simplemente “imagen”, se refiere a una función de intensidad de luz bidimensional $f(x, y)$, donde x, y indican las coordenadas espaciales y el valor de f en cualquier punto (x, y) es proporcional a la luminosidad (o nivel de gris) de la imagen en dicho punto.

Cuando una imagen es generada por un proceso físico, los valores son proporcionales a la energía radiada por el elemento físico, por ejemplo ondas electromagnéticas. Como consecuencia de lo anterior, $f(x, y)$ debe ser un número mayor de cero y finito, es decir

$$0 < f(x, y) < \infty$$

(II.3)

La función puede ser caracterizada por dos componentes:

- La cantidad de iluminación incidente sobre la escena que está siendo vista.
- La cantidad de iluminación reflejada por los objetos en la escena.

Una imagen digital es una imagen $f(x, y)$ que ha sido discretizada tanto en coordenadas espaciales como en luminosidad. Una imagen digital puede ser considerada como una matriz cuyos índices de renglón y columna identifican un punto en la imagen y el correspondiente valor de elemento de matriz identifica el nivel de gris en aquel punto. Los elementos de estos arreglos digitales son llamados elementos de imagen o píxeles.

Pero una imagen digital no debe verse sólo como una simple colección de números dispuestos en forma matricial, sino como la distribución espacial de la respuesta espectral de los elementos que componen la escena, donde la respuesta espectral corresponde a la variación de intensidad de energía radiante en función de su longitud de onda.

2.5 Muestreo y cuantización

Existen diversos caminos para adquirir una imagen, sin embargo, todos estos caminos tienen el mismo objetivo: generar una imagen digital de conjunto de datos. La salida para la mayoría de los sensores es un voltaje continuo en forma de señal, cuya amplitud y comportamiento espacial, están relacionados con el fenómeno físico del que fue sentido. Para crear una imagen digital, es requisito convertir los datos continuos sentidos en un formato digital. Esto involucra dos procesos: el muestreo y la cuantización.

La idea básica del muestreo y la cuantización es que, de una imagen continua $f(x, y)$, ésta se convierta a una con formato digital.

El muestreo es el proceso de convertir una señal (una función continua en el tiempo o en el espacio) en una secuencia numérica (una función discreta en el tiempo o en el espacio).

Una imagen puede ser continua con respecto a los ejes x , y , y puede serlo también en amplitud.

Para convertirla en un formato digital la función se tiene que muestrear en ambos ejes y también en la amplitud. Digitalizar los valores de las coordenadas se denomina muestreo, mientras que digitalizar los valores de amplitud se denomina cuantización.

El teorema de muestreo señala que la reconstrucción de una señal continua en el tiempo en banda base a partir de sus muestras es posible si la señal es limitada en banda y la frecuencia de muestreo es mayor que dos veces el ancho de banda de la señal.

El teorema de muestreo es comúnmente llamado teorema de muestreo de Shannon y también es conocido como teorema de muestreo de Nyquist-Shannon-Kotelnikov, Whittaker-Shannon-Kotelnikov o Whittaker-Nyquist- Kotelnikov-Shannon, WKS.

El proceso de muestreo sobre una señal continua que varía en el tiempo (o en el espacio como en una imagen u otra variable independiente en cualquier otra aplicación) es realizado midiendo simplemente los valores de la señal continua cada T unidades de tiempo (o espacio), llamado intervalo de muestreo. El resultado de este proceso es una secuencia de números, llamadas muestras, y son una representación de la imagen original. La frecuencia de muestreo f , cuyas unidades son Hz, es el recíproco del intervalo de muestreo

$$v = 1/T. \quad (\text{II.4})$$

Las condiciones que se deben tomar en cuenta en el proceso de muestreo son:

Limitar en banda a través de un filtro paso-bajas la señal a muestrear.

Siguiendo el criterio de Nyquist, si conocemos el ancho de banda de la señal, entonces la frecuencia de muestreo v para lograr una reconstrucción casi perfecta de la señal original deberá ser $v_N \geq 2WB$, donde WB es el ancho de banda de la señal original y la frecuencia de muestreo que sigue esta condición se le llama frecuencia de Nyquist.

Si las condiciones de muestro no se satisfacen, entonces las frecuencias se pueden llegar a traslapar; es decir, las frecuencias superiores a la mitad de la frecuencia de muestreo serán reconstruidas y aparentarán ser frecuencias por debajo de la frecuencia de muestreo. El resultado sería una distorsión llamada *aliasing*.

Aunque el teorema de muestreo está formulado para funciones de una sola variable, también puede ser extendido de la misma manera a funciones de varias variables arbitrarias. Por ejemplo, las imágenes en escala de grises son representadas frecuentemente como matrices de números reales representando las intensidades relativas de los pixeles localizados en las intersecciones de renglones y columnas. Como resultado, las imágenes necesitan dos

variables independientes o índices para especificar a cada pixel individualmente; una para los renglones y otra para las columnas.

Una imagen puede ser continua respecto al eje de coordenadas x , y , pero también puede ser continua en amplitud. Para convertir una imagen continua a su forma digital, como se vio anteriormente, a la digitalización de los valores de las coordenadas se le llama muestreo mientras que el proceso de digitalizar la amplitud es llamado cuantización.

Normalmente, el proceso de adquisición de la imagen se realiza usando una matriz de sensores. El número de sensores dentro de la matriz establece los límites del muestreo en ambas direcciones (Ilustración 17).

La digitalización de la amplitud o cuantización la realiza cada sensor asignando un valor discreto a ciertos intervalos de amplitudes continuas.



(a)



(b)

Ilustración 8 Muestreo y cuantización de imágenes

La Ilustración (a) muestra una imagen continua proyectada sobre un plano de arreglos y sensores. La Ilustración (b) muestra esa misma imagen después del proceso de muestreo y cuantización. Claramente se observa que la calidad de una imagen digital es determinada por el número de muestras y de los niveles de gris usados en el muestreo y la cuantización.

Las imágenes a color consisten regularmente de una composición de tres imágenes separadas en escala de grises, llamados modelos de color, cada uno representa los tres colores.

2.6 Modelos de color

Un **modelo de color** es la especificación de un sistema de coordenadas tridimensionales y se un subespacio de éste sistema en el que cada color queda representado por un único punto.

Existen diversos modelos de color, los más importantes son:

➤ Modelo de color RGB (Red, Green, Blue)

En este modelo cada color aparece con sus componentes espectrales primarias de rojo, verde y azul (Ilustración 18). Está basado en un sistema de coordenadas cartesianas y su subespacio de color está definido como un cubo unitario, de manera que los valores en cada punto de este espacio están normalizados de $[0,1]$.

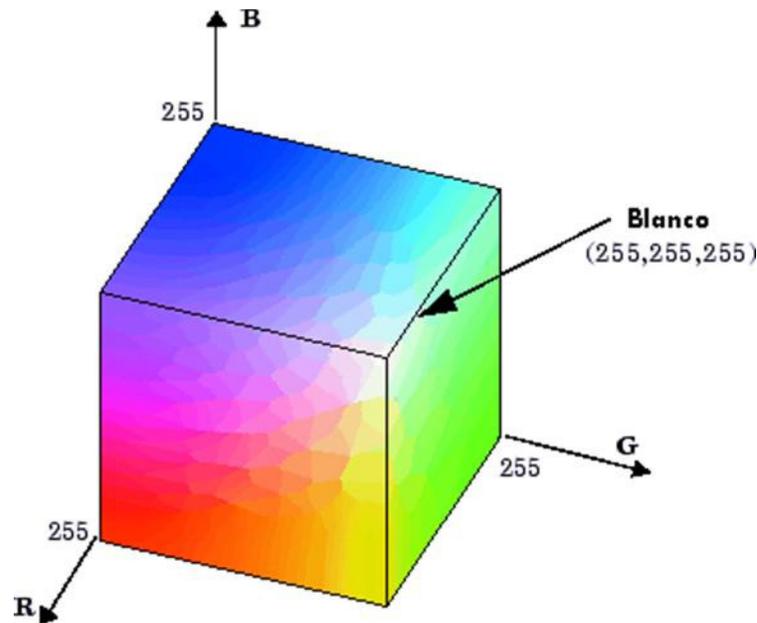


Ilustración 9 Modelo de color RGB

Las imágenes del modelo de color RGB consisten en tres planos de imágenes independientes, uno por cada color primario. Este tipo de modelo de color lo encontramos en monitores a color y cámaras de video.

➤ Modelo de color CMY (Cyan, Magenta, Yellow)

El cian, magenta y amarillo son los colores secundarios de luz, o bien los colores primarios de los pigmentos. Por ejemplo, cuando una superficie recubierta de un pigmento cian se ilumina con luz blanca, no se refleja la luz roja de ella. Es decir, el cian absorbe la luz roja presente en la luz blanca.

La mayoría de los dispositivos que depositan pigmentos coloreados sobre papel, tales como impresoras y fotocopiadoras en color necesitan una entrada CMY o bien una conversión interna de RGB a CMY.

La equivalencia entre los modelos RGB y CMY es la siguiente:

$$\begin{pmatrix} C \\ M \\ Y \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \begin{pmatrix} R \\ G \\ B \end{pmatrix} \quad (\text{II.5})$$

Normalizadas de [0,1]

➤ Modelo de color YIQ (Y=luminancia, I=fase, Q=cuadratura)

Este modelo se utiliza en emisiones comerciales de TV. Es de hecho una re modificación del modelo RGB utilizada por su eficacia en la transmisión y para mantener la compatibilidad de los estándares de TV blanco y negro. De hecho la componente Y (luminancia) del sistema proporciona toda la información de video que necesita un sistema de TV monocromática.

La equivalencia entre los modelos RGB y YIQ es la siguiente

$$\begin{pmatrix} Y \\ I \\ Q \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & 0.275 & 0.321 \\ 0.212 & 0.523 & 0.311 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} \quad (\text{II.6})$$

➤ Modelo de color HSI (Hue=tono, S=saturación, Q=cuadratura)

El tono es un atributo cromático que describe un color puro (amarillo puro, naranja puro, etc.), mientras la saturación proporciona una medida del grado en que un color puro está diluido en luz blanca. El modelo HSI es útil debido a que la componente de intensidad (I) está desacoplada de la información cromática, mientras que las componentes de saturación y tono están relacionadas con la forma en la que los humanos percibimos el color.

El modelo HSI es una herramienta ideal para el procesamiento de imágenes en color. Por ejemplo el modelo HSI se ocupa en el diseño de sistemas automáticos para determinar el grado de maduración de frutas y otros vegetales, hasta sistemas para comprar muestras de color o inspeccionar la calidad de productos coloreados.

La conversión de RGB a HSI se realiza de acuerdo con las siguientes expresiones:

$$I = \frac{1}{3}(R + G + B) \quad (\text{II.7})$$

$$S = 1 - \frac{3}{(R + G + B)}[\min(R, G, B)] \quad (\text{II.8})$$

$$H = \frac{\theta}{360 - \theta} \quad (\text{II.9})$$

$$\theta = \cos^{-1} \frac{1/2[(R - G) + (R - B)]}{[(R - G)^2 + (R - B)(G - B)]^{1/2}} \quad (\text{II.10})$$

si $B \leq G$
si $360 - \theta \geq G$

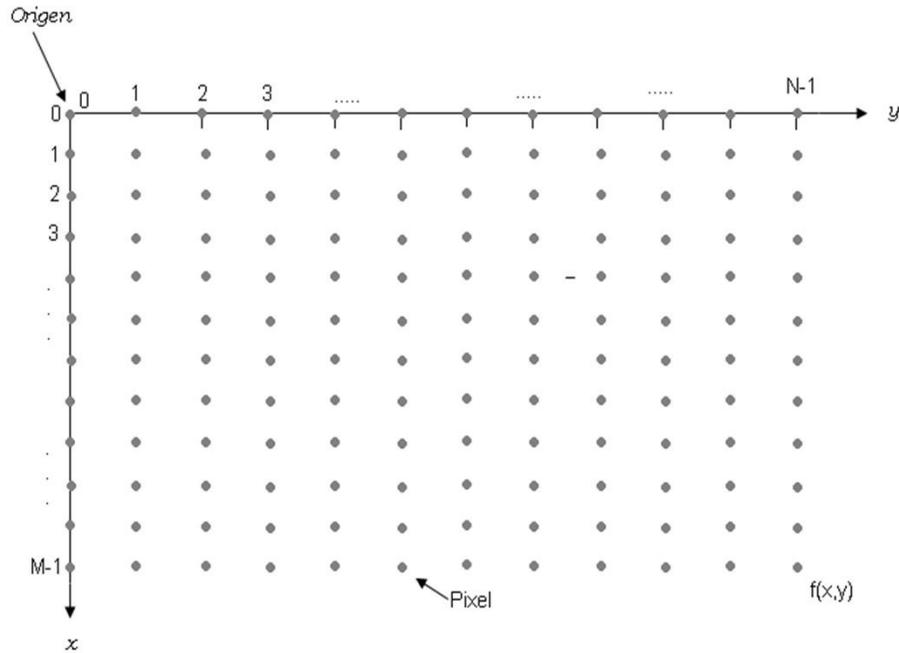
2.7 Representación de imágenes digitales

El resultado del muestro y la cuantización es una matriz de números reales. Si una imagen $f(x, y)$ es muestreada dando como resultado una imagen digital, esta última tendrá M filas y N columnas.

Los valores de las coordenadas (x, y) ahora se convertirán en cantidades discretas. Para una notación clara usamos valores enteros para las coordenadas discretas.

De esta manera los valores de las coordenadas del origen serán $(x, y) = (0, 0)$. Los valores de la siguiente coordenada a través de la primera fila de la imagen están representados por $(x, y) = (0, 1)$.

La notación anterior nos permite escribir la imagen digital completa M x N en una matriz de la siguiente forma:



Es posible escribir la imagen digital completa $M \times N$ en forma de matriz:

$$f(x,y) = \begin{pmatrix} f(0,0) & f(0,1) & \cdots & f(0,N-1) \\ f(1,0) & f(1,1) & \cdots & f(1,N-1) \\ \vdots & \vdots & \vdots & \vdots \\ f(M-1,0) & f(M-1,1) & \cdots & f(M-1,N-1) \end{pmatrix} \quad (\text{II.11})$$

El lado derecho de la ecuación es por definición una imagen digital. Cada elemento del arreglo de esta matriz es llamado un elemento de la imagen o píxel. En el caso de imágenes multiespectrales, cada matriz es una banda multiespectral, así cada banda de una imagen multiespectral está representado por una matriz.

En el capítulo VI se usa una notación de matriz más tradicional para denotar los elementos de una imagen digital:

$$A = \begin{pmatrix} a_{0,0} & a_{0,1} & \cdots & a_{0,N-1} \\ a_{1,0} & a_{1,1} & \cdots & a_{1,N-1} \\ \vdots & \vdots & \vdots & \vdots \\ a_{M-1,0} & a_{M-1,1} & \cdots & a_{M-1,N-1} \end{pmatrix} \quad (\text{II.12})$$

Donde $a_{ij} = f(x = i, y = j) = f(i, j)$

Si \mathbb{N} representa el conjunto de números enteros y \mathbb{R} el conjunto de números reales. El proceso de muestreo está denotado como una partición del plano (x, y) en una retícula, donde en el centro de cada cuadro se encuentra un par de elementos pertenecientes al producto cartesiano \mathbb{N}^2 , el cual es el conjunto de todos los pares ordenados de elementos (I_i, I_j) donde I_i y I_j son enteros pertenecientes a \mathbb{N} .

Por lo tanto $f(x, y)$ es una imagen digital si (x, y) son enteros pertenecientes a \mathbb{N} y f es la función que asigna los valores de niveles de gris, los cuales pertenecen al conjunto \mathbb{R} , para cada par de coordenadas (x, y) . Esta asignación es el proceso de cuantización.

Si los niveles de gris también son enteros \mathbb{Z} reemplaza a \mathbb{R} y la imagen digital se convierte en una función 2D cuyos valores de amplitud y coordenadas son enteros.

El proceso de digitalización requiere conocer los valores de M y N , y para el número L (el valor discreto de nivel de gris) para cada píxel. No se requiere de otros valores más que de los valores de M y N los cuales deben ser positivos (Ilustración 19). Sin embargo, el número de niveles de gris es típicamente un entero potencia de 2: $L = 2^k$.

Algunas veces el rango de valores espaciados por la escala de grises es llamado rango dinámico de la imagen, y se refiere a imágenes cuyos niveles de gris varían en una proporción significativa de una escala de grises y además que contienen un elevado rango dinámico.

Cuando un número significativo de píxeles presenta un elevado rango dinámico, la imagen muestra alto contraste. En cambio una imagen con poco rango dinámico tenderá a ser gris, sin contraste.

El número de bits b requeridos para guardar una imagen digitalizada es: $b = M \times N \times k$.

Cuando una imagen presenta 2^k niveles de gris, es común referirse a ella como una “imagen de k -bits”. Por ejemplo si una imagen presenta 256 posibles valores de niveles de gris, esta es llamada imagen de 8-bits, también se dice que la imagen está cuantizada a 8-bits.

El muestreo es el factor principal para determinar la resolución espacial de una imagen. Se entiende por resolución espacial al detalle más pequeño discernible en una imagen. La resolución de los niveles de gris es similar, y se refiere a qué tan discernible puede ser el cambio de un nivel de gris a otro.

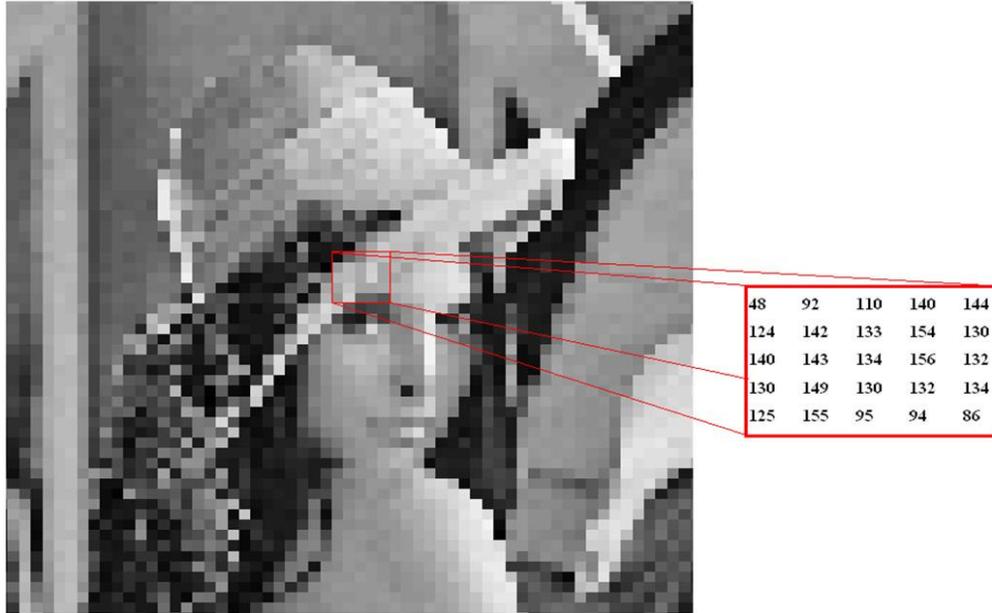


Ilustración 10 Representación visual de una imagen digital con los valores numéricos de los pixeles correspondientes

2.8 Procesamiento digital de imágenes

El término procesamiento digital de imágenes se refiere a la manipulación e interpretación de imágenes con ayuda de una computadora. Entre los intereses principales se encuentra la transformación de datos para poder relacionarlos con un marco de referencia adecuado y así poder tener su análisis e interpretación de la imagen. Otro interés básico es la extracción de ciertos rasgos o características de una imagen con el objeto de apoyar un análisis automático de tales datos.

El procesamiento digital de imágenes es un campo de investigación abierto y en constante progreso. Se encuentra relacionado con diversas áreas de estudio como las matemáticas, la física, la ingeniería y la computación, lo cual se ve reflejado en el avance del estudio de imágenes en áreas tales como la medicina, la astronomía, geología, microscopía, urbanismo y percepción remota.

La percepción remota se basa en el procesamiento digital, el análisis de imágenes multiespectrales y el reconocimiento de patrones. Las imágenes obtenidas por satélites ayudan a estudiar los recursos y fenómenos que presenta la tierra, por ejemplo, un levantamiento geográfico, la predicción de cultivos agrícolas, el crecimiento urbano, la

predicción del clima, la clasificación y monitoreo de ambientes urbanos, agrícolas y marinos, la evaluación del medio ambiente y desastres naturales, la caracterización del tipo de suelos y texturas del terreno, entre otras.

El análisis de imágenes se descompone básicamente en dos grandes grupos de transformaciones: el procesamiento digital y el reconocimiento de patrones.

En algunas imágenes digitales los objetos de interés o patrones no siempre se encuentran en la representación o con la evidencia deseada, de tal forma que es necesario realzarlos antes de proceder a realizar una transformación en los mismos. Debido a esto, al grupo de procesamientos realizados a la imagen se les conoce como realce de patrones.

El procesamiento digital se divide a su vez en:

Realces

Los procesos de realce de imágenes consisten en una serie de técnicas cuyo objetivo es mejorar la apariencia visual de una imagen, ya sea en contraste, ruido, escala de grises, distorsiones, luminosidad, falta de nitidez y definición (Ilustración 20). Los métodos de realce de imágenes los podemos dividir en dos categorías:

➤ Los métodos de realce en el dominio espacial

Estos métodos consisten en la manipulación directa de los píxeles.

Dentro de los métodos de realce espacial podemos encontrar dos técnicas:

1. El *realce radiométrico* donde las operaciones son efectuadas directamente sobre un píxel sin importar o tomar en cuenta a los píxeles vecinos. Sirve para mejorar condiciones de bajo contraste, baja luminosidad o demasiada oscuridad. Ejemplo: ecualización del histograma.

2. El *realce con operaciones entre vecinos* en el que, a diferencia del realce radiométrico, las operaciones son efectuadas sobre un píxel pero tomando en cuenta a los píxeles que lo rodean. Sirve para eliminar ruido o para el mejoramiento de la nitidez. Ejemplo: kernel correspondiente a un filtro paso-bajas usando la convolución para realizar un filtrado espacial.

➤ Los métodos de realce en el dominio de la frecuencia.

Estos métodos corresponden a técnicas basadas en la representación de los píxeles, a través de una transformación hacia el dominio de la frecuencia y usa como operador de mapeo a la transformada de Fourier.



(a)



(b)

Ilustración 11 Realce de imágenes (a) imagen sin realce (b) imagen con realce por medio de la técnica normalización del histograma

a) Restauraciones

También son llamadas realces de Fidelidad, se supone que la imagen ha sido degradada por algún mecanismo de degradación conocido, a partir de cuyo modelo se construye una imagen ideal (Ilustración 21).



(a)



(b)

Ilustración 12 Restauraciones de imágenes (a) imagen con ruido (b) imagen restaurada

b) Segmentación

En el análisis de imágenes es útil separar patrones que son de interés, los cuales están relacionados directamente con los objetos de la escena, del resto de los demás en la imagen. Este proceso puede ser: real si los píxeles que hacen referencia a un patrón son asignados a una imagen intermedia en donde éstos ocupan sus mismas posiciones relativas, dando como resultado tantas imágenes como patrones deseamos separar. Puede ser virtual si los píxeles del patrón sólo se etiquetan como pertenecientes a una clase dada o si todos sus valores cambian a uno definido. Al resultado de particionar una imagen en regiones se le denomina segmentación (Ilustración 22).



Ilustración 13 Segmentación de imágenes (a) imagen original (b) imagen segmentada

En la ilustración 22 la segmentación se realizó empleando el algoritmo isodata, el cual es un algoritmo iterativo. A partir de la primera estimación del umbral, se van recalculando los parámetros y nuevos umbrales sucesivamente.

Una imagen digital contiene patrones relacionados con los objetos de una escena dada y habrá también patrones que no tengan ninguna relación con la escena. Para estudiar los patrones de interés es necesario separarlos del resto de la imagen.

El proceso de segmentación puede llevarse a cabo con base en una serie de modelos y criterios: por regiones, por bordes o líneas, por propiedades espectrales, por propiedades espaciales o por una combinación de estas dos últimas llamada contextual.

Segmentación espacial: Es posible que un objeto esté caracterizado por propiedades espaciales tales como textura y forma.

Para la banda 1 de una imagen Landsat en escala de grises, segmentaríamos físicamente con base en criterios espaciales en dos regiones, la definida por la línea roja que presenta diferentes escalas de gris y la zona definida por línea azul que presenta una textura más o menos uniforme (Ilustración 23 y 24).

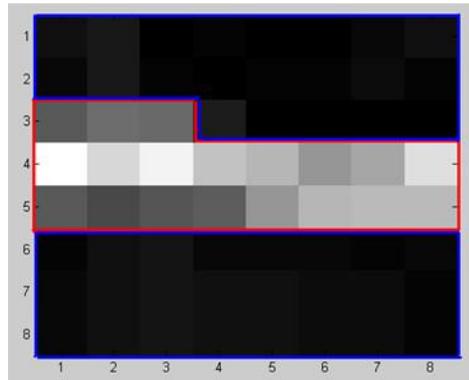


Ilustración 14 Región a segmentar delimitada por las líneas azul y roja

Podríamos delimitar dos zonas:

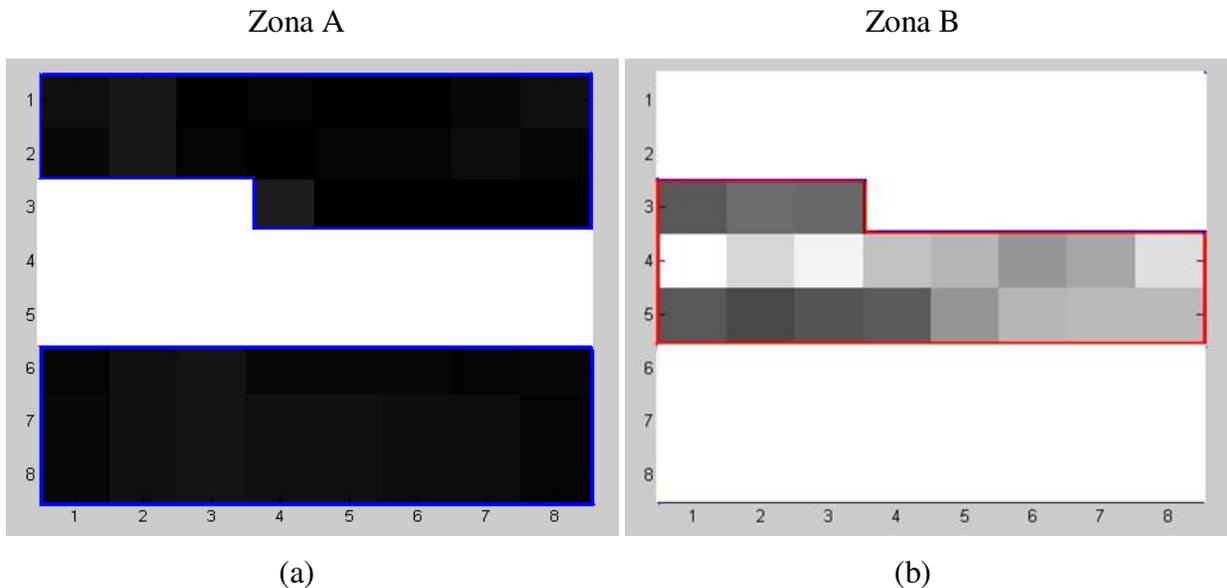


Ilustración 15 Segmentación de zonas

(a) Segmentación de la zona A delimitada por el contorno azul (b) Segmentación de la zona B limitada por el contorno rojo.

Segmentación espectral: La segmentación se realiza empleando un modelo de la respuesta espectral de los píxeles involucrados.

Para la banda 5 de una imagen multispectral Landsat segmentaríamos virtualmente con base en criterios espectrales, por ejemplo, cuáles son los píxeles que tienen una respuesta espectral mayor o igual a 115 (Ilustración 25).

47	52	43	36	36	40	43	50
47	57	43	39	40	44	48	52
77	98	88	62	54	55	52	54
133	130	132	115	101	93	93	105
75	74	91	118	148	152	152	155
41	53	56	50	49	51	47	52
43	48	49	50	51	53	50	45
38	49	52	53	55	47	49	41

Ilustración 16 Segmentación espectral

Segmentación contextual: La segmentación se realiza en función de su respuesta espectral y del entorno espacial en el que se encuentra. Para la banda 3 se muestra en la Ilustración 26

18	21	13	13	13	13	15	17
17	24	12	13	16	14	18	19
36	50	48	28	23	23	22	23
70	67	75	62	57	51	52	57
42	36	44	59	81	85	86	87
15	21	23	19	19	20	20	21
14	19	18	18	20	20	20	15
18	20	21	19	23	18	18	15

Ilustración 17 Ejemplo de una segmentación de tipo contextual

Una forma de definir cada Zona sería la siguiente (Ilustración 27):

Zona **A** - Para valores de píxel de 12 a 23

Zona **B** - Para valores de píxel de 24 a 52

Zona **C** - Para valores de píxel de 53 a 67

Zona **D** - Para valores de píxel de 68 a 70

Zona **E** - Para valores de píxel de 71 a 81

Zona **F** - Para valores de píxel de 82 a 87

A	A	A	A	A	A	A	A
A	B	A	A	A	A	A	A
B	B	B	B	A	A	A	A
D	C	E	C	C	B	B	C
B	B	B	C	E	F	F	F
A	A	A	A	A	A	A	A
A	A	A	A	A	A	A	A
A	A	A	A	A	A	A	A

Ilustración 18 Representación de una segmentación de tipo contextual

La finalidad en el análisis de imágenes es la cuantificación de las propiedades de los objetos presentes en éstas, para lo cual es necesario entender los elementos que las componen, su origen y su naturaleza. Los métodos de análisis están estrechamente relacionados con estos aspectos. La toma de decisiones en muchas áreas del conocimiento se basan en el resultado de este análisis (Lira, 2002).

El conocimiento acerca de un objeto se adquiere únicamente a través de un proceso de formación de modelos, y la clasificación de un objeto presente en una imagen es la decisión por la cual se determina la categoría a la cual pertenece; el reconocimiento de un objeto implica su clasificación.

2.9 Imágenes digitales en percepción remota

2.9.1 Elementos de una imagen digital

Una imagen digital no es solo una colección de números dispuestos en forma matricial, es también la distribución espacial de la respuesta espectral de los elementos que componen la escena que se está observando (Ilustración 28). En percepción remota tenemos los siguientes elementos que se encuentran presentes al momento de obtener una imagen digital multispectral:

a) Fuente de radiación: La fuente de radiación puede ser de tipo ondulatoria (electromagnética o acústica) o de tipo corpuscular (neutrones, electrones, protones).

La radiación emitida por la fuente de radiación interacciona con un elemento físico de la escena, el cual es llamado Campo Instantáneo de Vista (CIV)

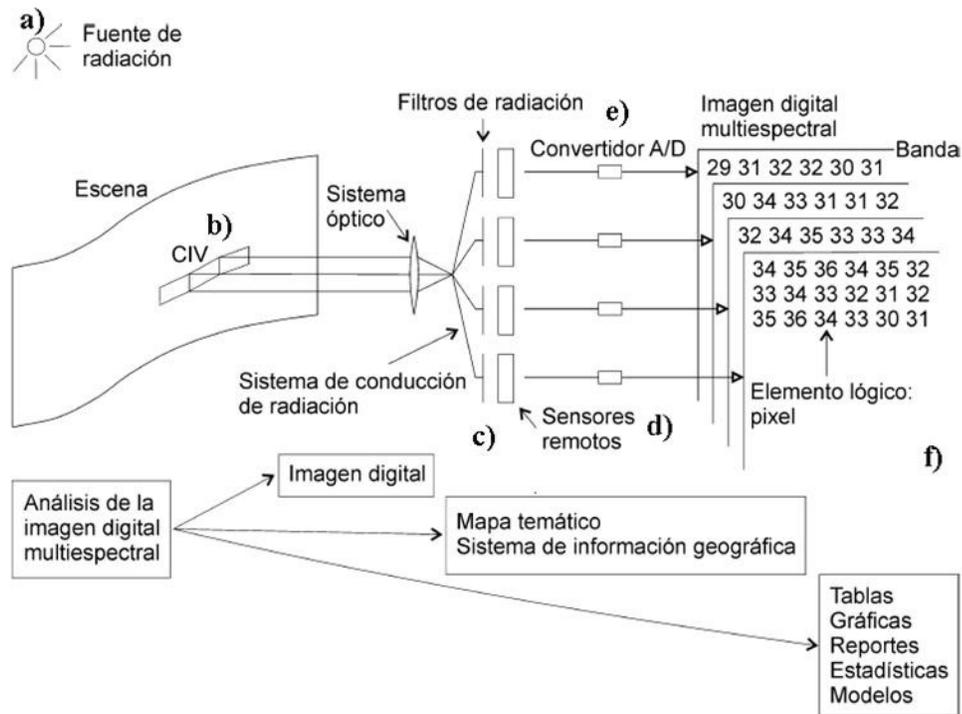


Ilustración 19 Esquema general que muestra la generación de una imagen digital multispectral a partir de una escena dada

(Adaptado de Lira, 2002)

a) Fuente de radiación: La fuente de radiación puede ser de tipo ondulatoria (electromagnética o acústica) o de tipo corpuscular (neutrones, electrones, protones).

La radiación emitida por la fuente de radiación interacciona con un elemento físico de la escena, el cual es llamado Campo Instantáneo de Vista (CIV).

b) Campo instantáneo de vista (CIV).- Es un elemento físico de la escena, se encuentra formando parte del sistema físico. El resultado de la interacción de la radiación con el CIV es también radiación, la energía o longitud de onda resultante no es necesariamente la misma que la que incide originalmente.

c) Sensor o detector.- Es el dispositivo que recibe la radiación que resulta al incidir la radiación proveniente de la Fuente de Iluminación y el CIV. De acuerdo con la geometría particular: Fuente – CIV – Detector.

Una parte de la radiación emitida, reflejada o transmitida por el CIV es captada, sobre un cierto ángulo sólido en una dirección predeterminada, por un arreglo de sensores (detectores).

Los sensores poseen un subsistema de enfoque y conducción de radiación, los cuales pueden ser arreglos de lentes (telescopio), fibras de conducción (cuando se emplea radiación electromagnética).

d) La radiación posteriormente es conducida hasta un grupo de detectores, frente a los cuales se encuentra un conjunto de filtros cuya función es la de seleccionar un conjunto de intervalos de energía $\{\Delta E_1, \Delta E_2, \dots, \Delta E_\gamma\}$ o de longitudes de onda $\{\Delta \lambda_1, \Delta \lambda_2, \dots, \Delta \lambda_\gamma\}$ para los que cada detector está diseñado. Es decir, el detector i -ésimo está adecuado para detectar la radiación del intervalo ΔE_i .

Es posible tener uno o varios detectores para cada intervalo dependiendo de si se mide la radiación de un CIV o un grupo de ellos al mismo tiempo.

Si se tienen γ intervalos de energía y se miden n CIV para cada uno de éstos, entonces se requieren de $\gamma \times n$ detectores para operar el sistema sensor.

Para cada intervalo de energía los detectores miden la cantidad de radiación por unidad de ángulo sólido por unidad de tiempo; las unidades son: mW/esterad/s.

Los detectores generan un voltaje con intensidad proporcional al resultado de haber integrado la radiación (Ilustración 29).

e) Convertidor analógico-digital. La señal de voltaje generada por los detectores, a medida que observan un CIV tras otro, se introduce a un convertidor analógico-digital que la cuantiza empleando una escala que usualmente va del 0 al 255 (256 posibles niveles); la razón de utilizar 256 niveles es que este número es almacenable en un byte de 8 bits de una computadora digital. Los detectores están calibrados de tal manera que se genera un 255 para el CIV más brillante de la escena y un 0 para el más oscuro.

La respuesta del detector, en la mayoría de los casos es lineal, como se puede observar en la Ilustración 29.

Siempre hay una señal de salida o respuesta para cualquier radiación incidente incluso si ésta es nula, esto se denomina polarización del detector (Ilustración 29).

La ganancia es la relación constante entre la radiación incidente y la señal de salida producida por éste.

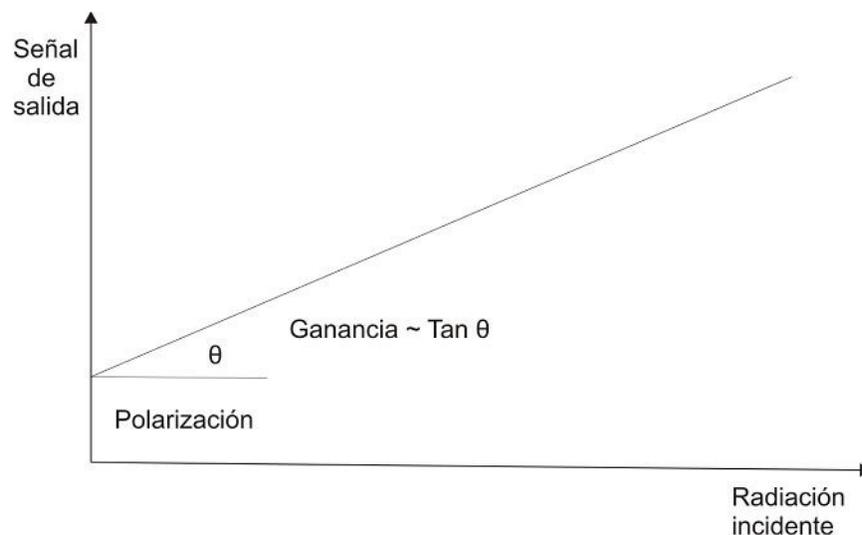


Ilustración 20 Respuesta lineal de un sensor remoto frente a una cierta radiación incidente
(Adaptada de Lira, 2002)

En imágenes de satélite la diferencia de ganancia de los detectores del sistema sensor, produce un efecto conocido como bandeo y cuya eliminación se realiza por medio de un filtraje.

f) Elemento lógico o píxel. Para cada intervalo de energía y para cada CIV se genera un número, cuyo valor se guarda en un medio de almacenamiento en una posición lógica que va de acuerdo con el lugar relativo que tiene el CIV correspondiente en la escena.

Para cada CIV se producen γ números que conjuntamente describen su respuesta espectral, los que arreglados en forma vectorial se escriben como: $p_{m,n} = \{b_1, b_2, \dots, b_\gamma\}_{m,n}$ b_i es el valor numérico (entero) obtenido para el intervalo de energía ΔE_i .

La colección de estos números, arreglados secuencialmente (en cuando a la energía o longitud de onda de la radiación empleada) forman lo que se conoce como píxel el cual es un elemento lógico de la imagen digital. La Ilustración 30 muestra los elementos de una Imagen Digital

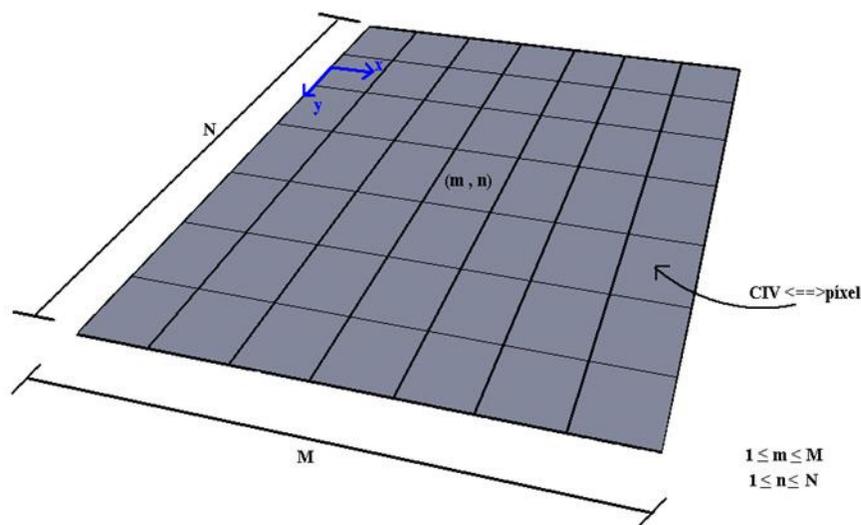


Ilustración 21 Elementos de una Imagen digital

Los enteros (m, n) representan las coordenadas del píxel en su posición dentro de la imagen digital. Las características espectrales de un CIV quedan representadas por el píxel correspondiente. M y N son números enteros que representan el número de CIV en los cuales se dividió la escena en las direcciones x , y respectivamente.

Al terminar de medir todos los CIV pertenecientes a una escena, se tiene una colección de matrices de números, para cada uno de los intervalos de energía empleados.

Se obtiene una matriz de $M \times N$ por cada banda por cada intervalo de energía ΔE_i

El tamaño de las matrices es de $M \times N$ y cada una de ellas forma una imagen digital conocida como **banda**, ya que representa la distribución espacial de valores de radiación provenientes de la escena para una banda o intervalo de energía.

El conjunto de estas bandas para una escena dada se conoce como **Imagen Digital Multiespectral**, ya que caracteriza la distribución espacial y espectral de la radiación proveniente de la Imagen.

La Imagen Multiespectral forma un campo vectorial donde los píxeles que la componen son vectores de tantos elementos como bandas se hayan generado. El píxel es el elemento digital básico de una Imagen Multiespectral.

2.9.2 Características de una imagen digital

Para generar una Imagen Digital se llevan a cabo procesos de discretización, estos pueden ser de dos tipos, espacial y radiométrica.

- Discretización espacial.- Se realiza una división de la escena en CIV empleando una rejilla virtual. Digitización o Digitalización: produce o define la medición del arreglo matricial de cada banda.
- Discretización radiométrica. Conversión a un número de la Intensidad de flujo radiante integrada por el sensor remoto. Cuantización: define el número de niveles o tonos de la imagen.

En cuanto a los elementos básicos que determinan el grado de detalle observable en una imagen de percepción remota se encuentran:

1. Digitalización
2. Cuantización
3. Contenido de ruido
4. Enfoque o nitidez de la Imagen