

Marco Teórico

A continuación se revisarán algunos conceptos teóricos utilizados en este trabajo.

4.1. Imágenes por Computadora

La importancia de las imágenes por computadora se debe a que el sentido más importante, al menos para el humano, es la visión y además que la información contenida en una imagen es muy rica; de ahí surge la expresión "una imagen habla más que mil palabras".

Las imágenes por computadora puede dividirse en dos grandes áreas:

- Visión por computadora (VC) La salida de un sistema basado en VC será usada por una máquina.
- Tratamiento de Imágenes (TI) La salida de un sistema basado en TI será usado por un ser humano.

Estas dos ramas aunque distintas comparten varias herramientas, conceptos, etc.

Históricamente, el TI surgió como una extensión del procesamiento (digital) de señales, mientras que las ciencias de la computación son responsables del desarrollo de la VC.

Uno de los tópicos más importantes de la VC es el análisis de imágenes (AI).

El AI envuelve el examen de la información contenida en una imagen para facilitar la solución de un problema dado, involucra dos grandes problemas:

- Extracción de rasgos La extracción de rasgos es el proceso de adquirir información de alto nivel de la imagen; por ejemplo: forma.
- Clasificación de patrones La clasificación de patrones toma esta información para identificar objetos en una imagen.

En cuanto al TI involucra al ser humano en el lazo visual. Estas técnicas usan para su diseño el conocimiento acerca del sistema visual humano para mejorar la apariencia de una imagen.

Los tópicos más importantes son:

- Mejoramiento de una imagen.
- Restauración de una imagen.
- Compresión de imágenes.
- Reconstrucción.

El mejoramiento de imágenes involucra en tomar una imagen como entrada, usando para esto las ventajas del sistema visual humano.

Los métodos resultantes son normalmente orientados por el problema. Esto quiere decir que el mismo método que funciona para cierto tipo de imágenes no funciona para otro tipo de imágenes.

La restauración de imágenes es el proceso que toma como entrada una imagen y usa un modelo conocido o estimado de la degradación (que afectó la apariencia de la imagen) tratando de restaurar la imagen a su estado natural.

Para lograr esto se requiere conocer algo acerca de la degradación a fin de obtener el modelo de la misma.

4.2. El modelo de imagen

El término imagen se refiere a una función de intensidad luminosa 2-dimensional $f(x, y)$ donde el valor de la amplitud de f en cada par (x, y) da la intensidad (brillantez) de la imagen en ese punto [González 1996⁸]. Como la luz es un tipo de energía, entonces:

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

Las imágenes percibidas consisten en la luz reflejada de los objetos componiendo una escena y las fuentes de iluminación asociadas. Las dos componentes son:

- La cantidad de luz incidente sobre los objetos: Componente de iluminación, $i(x, y)$
- La cantidad de luz reflejada por los mismos: Componente de reflexión, $r(x, y)$

El producto de estos dos componentes se forma $f(x, y)$, como se ve en la ecuación 4.1.

$$f(x, y) = i(x, y) r(x, y) \tag{4.1}$$

Donde

$$\begin{aligned} 0 < i(x, y) < \infty \\ 0 < r(x, y) < 1 \end{aligned}$$

Con respecto a la ecuación 4.1 vemos que $r(x, y)$ esta irradiada por abajo en 0 (absorción total) y 1 por arriba (reflexión total).

- ✓ La naturaleza de $i(x, y)$ está dada por las características de las fuentes de iluminación.
- ✓ La naturaleza de $r(x, y)$ está dada por las características de los objetos.

La intensidad f de una imagen en un punto dado, en las coordenadas (x, y) se llama el nivel de gris I en el punto

$$L_{\min} \leq I \leq L_{\max}$$

En la práctica se requiere que $L_{\min} \geq 0$ y $L_{\max} < \infty$. Por cuestiones computacionales $L_{\min} = 0$ y $L_{\max} = 127/256$ (1 byte). El término intervalo $[L_{\min}, L_{\max}]$ se llama escala de grises y en la práctica $[0, L]$, donde por conveniencia $L_{\min} = 0$ (negro) y $I = L_{\max}$ (blanco).

Todos los demás valores intermedios de niveles de gris varían de negro a blanco.

4.3. Componentes de un sistema de Imágenes por Computadora

Se compone de dos grandes cosas: Hardware y Software.

En cuanto al hardware, es un medio de:

- Adquisición: Cámaras, CCD, escáner, etc.
- Procesamiento: PC, DSP, etc.
- Despliegue: Monitor, impresora, película, vídeo grabadora, etc.
- Almacenamiento: Disco duro.

Y al software, que permite manipular y procesar imágenes, puede usarse para controlar también los procesos de adquisición y almacenamiento de la información.

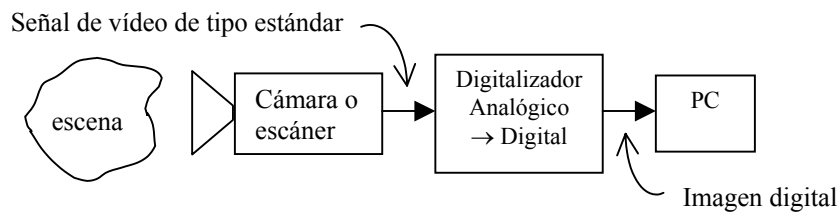


Fig. 4.1 Proceso de digitalización de una escena

El digitalizador o tarjeta de vídeo acepta una señal estándar (proveniente de un sensor) y entrega como salida a una imagen de manera que la computadora pueda almacenar, a esta forma se le llama imagen digital.

El proceso de transformar una señal de vídeo de tipo estándar a una imagen digital, recibe el nombre de digitalización.

Este proceso es necesario ya que la señal como sabemos, el vídeo estándar se encuentra en general en forma analógica (o continua), el procesador PC requiere de una versión digital de la misma.

Una señal de vídeo típica contiene marcos de información de vídeo, cada marco corresponde a una fotografía o pantalla completa de información visual. Un marco puede descomponerse en dos campos y cada campo consiste en líneas de información.

Ambos campos forman lo que se conoce como un marco de información visual. Este modelo de dos campos por marco se conoce como vídeo entrelazado. Algunos tipos de señales de vídeo, llamada vídeo no entrelazado tiene un solo campo. Para el caso de las líneas de información, la señal eléctrica tiene la forma, por ejemplo:

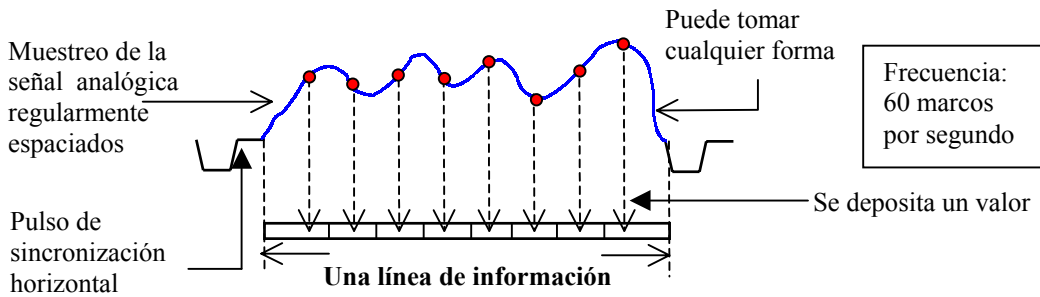


Fig. 4.2 Transformación de la señal analógica a digital

El pulso de sincronización horizontal indica al hardware de despliegue cuando empezar una nueva línea. Después de que un marco ha sido desplegado un pulso conocido corto, pulso de sincronía vertical indica al hardware de despliegue comenzar un nuevo marco.

La señal analógica es cambiada a señal digital (imagen digital) al hacer un muestreo de la señal continua, de manera continua.

En la figura 4.2, se puede ver como una línea de vídeo que esta haciendo un muestreo, es digitalizada al medir el nivel de voltaje a intervalos fijos de tiempo. El valor de voltaje en cada instante es convertido en un número que es almacenado en un buffer y corresponde al brillo de la imagen en ese punto.

Nótese que el brillo en ese punto depende tanto de las propiedades de los objetos en la escena, así como, de las condiciones de iluminación.

Cuando este proceso ha terminado, se tiene un marco de información de vídeo. En este sentido se dice que un marco de información de vídeo ha sido grabado y la computadora puede almacenar y/o procesar dicha imagen digital.

4.4. Muestreo y cuantización uniformes

Para que una imagen sea útil a la computadora, ésta debe ser digitalizada espacialmente (discretación, también conocida como muestreo) a través de un escáner y en amplitud (cuantización).

Si suponemos que una función continua $f(x, y)$ es digitalizada (hecha un muestreo y cuantizada), $f(x, y)$ puede ser vista como un arreglo de $n \times m$ elementos, donde cada elemento es una cantidad discreta.

Cada elemento en este arreglo se llama pixel (picture element) de una manera más formal, un pixel es una terna $(x, y, f(x, y))$, también los procesos de muestreo y cuantización (intensidad asociada) pueden verse como siguen:

Sean Z los conjuntos de los enteros y R los reales, el proceso de muestreo puede verse como un particionado del plano XY en una cuadrícula, siendo las coordenadas del centro de cada celda un par $Z \times Z = Z^2$ por tanto $f(x, y)$ es una imagen digital si (x, y) son enteros en $Z \times Z$ y f es una función que nos asigna un nivel de gris, un elemento de R a cada (x, y) [González 1996⁸].

Este proceso de asignación, es el proceso de cuantificación antes visto.

Si los niveles de gris son también enteros entonces, $R \rightarrow Z$ una imagen digital se convierte en una función 2D cuyas coordenadas y valores de amplitud son enteros.

La resolución o grado de detalle en una imagen depende fuertemente de la cantidad de muestras y los niveles de gris.

- Es claro que entre más grandes sean estos parámetros más cercana será la imagen digital a la original.
- Esto tiene que ver con el espacio requerido para almacenar tal información y por otro lado el tiempo de procesamiento.
- Una buena imagen es algo difícil de definir, depende sobretodo de la complejidad de la imagen y de los requerimientos de la aplicación.

A mayor nivel de gris y a mayor resolución es mejor [González 1996⁸].

El procesamiento de una imagen digital es un proceso jerárquico.

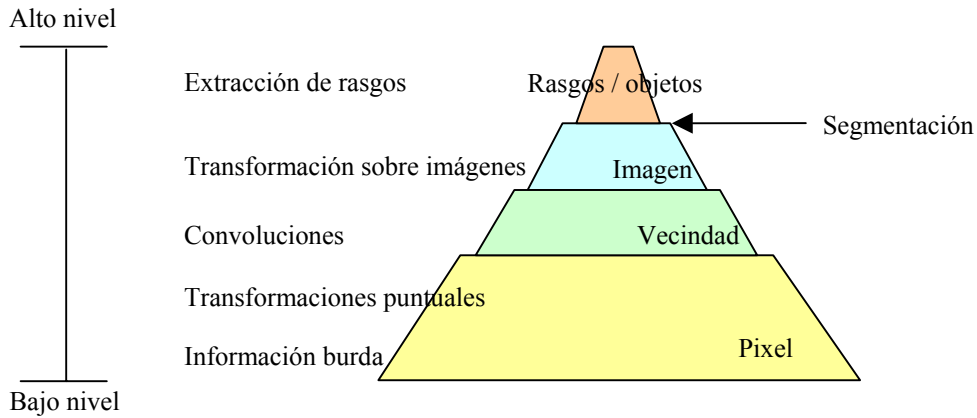


Fig. 4.3 Niveles de representación jerárquica en el procesamiento de una imagen digital

- En el bajo nivel se trata con pixeles individuales, donde se pueden realizar ciertos procesamientos.

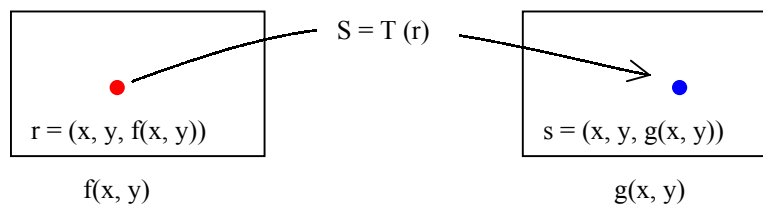


Fig. 4.4 Nivel de pixel

- En el siguiente nivel, se toma en consideración al pixel en cuestión y sus vecinos más próximos.

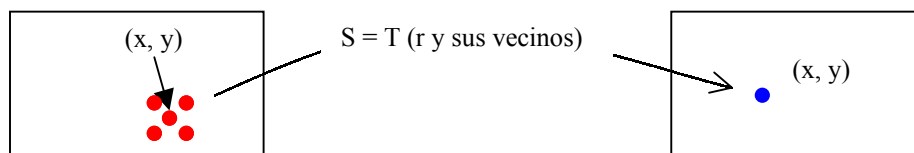


Fig. 4.5 Nivel de vecindad

- En el nivel de imagen, se toma toda la información contenida en la imagen a la vez para obtener una imagen transformada.
- Conforme subimos de nivel se obtienen niveles más altos de representación y por tanto una reducción en la cantidad de datos.

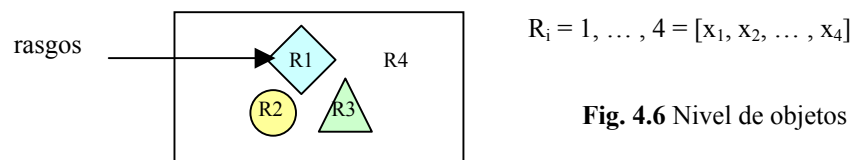


Fig. 4.6 Nivel de objetos

Para realizar el procesamiento digital de imágenes, es necesario que la imagen esté en formato digital. Existen básicamente tres maneras de obtener una imagen digital:

- 1) Adquirir la imagen de teledetección en formato analógico (por ejemplo, fotografía aérea) y posteriormente digitalizarla;
- 2) Adquirir la imagen de percepción remota ya en formato digital, tales como los datos en cintas CCT (Computer Compatible Tape), grabados por satélites como Landsat y SPOT.
- 3) Discretizar la imagen originalmente en papel a través de un escáner (por ejemplo, mapas temáticos).

4.5. Análisis de Imágenes (AI)

Como se vio en la sección 4.1, el AI involucra la manipulación de información contenida en una imagen para determinar los datos necesarios para ayudar a resolver un problema dado por computadora.

Sabemos que una imagen contiene mucha información, la mayoría de ésta no es necesaria para una aplicación dada, por lo tanto, una de las tareas primarias de AI es determinar la información mínima.

El AI puede usarse tanto en tratamiento de imágenes (TI) como en visión por computadora (VC).

- En VC el producto final consiste en información de alto nivel, por ejemplo:
 - Los parámetros de un objeto para controlar un robot.
 - Los rasgos de textura para distinguir automáticamente los tumores de piel.
- En TI el AI puede usarse para determinar el tipo de procesamiento requerido, así como los parámetros específicos para dicho procesamiento.
 - Poder determinar la función de degradación para restauración.
 - El desarrollo de una algoritmo de mejoramiento.
 - Cual es la información visualmente importante para un método de compresión.

De una manera muy general, el procesamiento de AI puede dividirse en tres etapas.

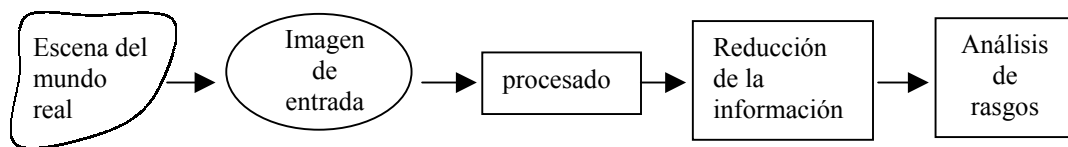


Fig. 4.7 Esquema de procesamiento con análisis de imágenes

1. Procesamiento: Remoción de ruido y eliminación de información innecesaria.

Ruido: Información no deseada, resultado del proceso de adquisición de la imagen.

2. Reducción de la información: Involucra reducir información, esto se puede hacer en el dominio frecuencial o en el dominio espacial para extraer características para el proceso de análisis.

3. Los rasgos extraídos son examinados y evaluados para su uso en la aplicación

- Después del procesamiento podemos segmentar la imagen en el dominio espacial o frecuencial.
- Después de esto podemos filtrar para reducir la información extrayendo rasgos útiles en el análisis, González [González 1996⁸] propone:

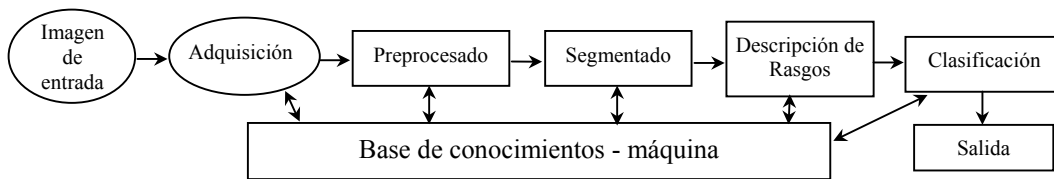


Fig. 4.8 Proceso completo con análisis de imágenes

4.5.1. Convolución

Se usa una "máscara" o "ventana" de convolución la cual es desplazada a lo largo de la imagen realizando sobre cada punto o pixel de la imagen extendida, la sigue una operación aritmética: El producto interno de la máscara con la región correspondiente, que se hace de la siguiente manera:

- Poner la máscara sobre una subimagen.
- Multiplicar términos coincidentes.
- Sumar productos y poner el resultado en la posición correspondiente del buffer.

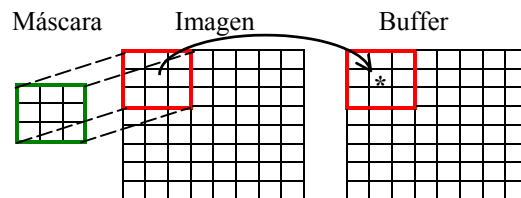


Fig. 4.9 Posición inicial de la convolución

El siguiente paso consiste en desplazar la ventana una posición a la derecha y repetir el producto interno, éste proceso continúa hasta terminar la fila correspondiente.

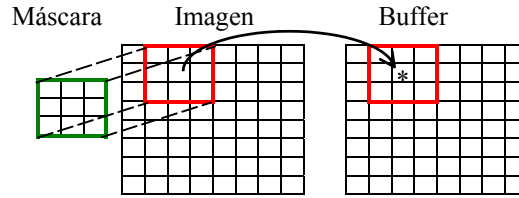


Fig. 4.10 Secuencia del procedimiento de convolución

Posteriormente se mueve la máscara hacia abajo una fila y repetir el proceso hasta que la imagen completa haya sido convolucionada.

Matemáticamente la operación que se está haciendo es la siguiente:

$$G(x, y) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} f(x - i, y - j) \cdot h(i, j)$$

En la práctica la máscara es recorrida sobre la imagen un pixel, antes y después, además, se prefiere usar el método de convolución, ya que se puede fácilmente hacerse en hardware.

4.6. Mejoramiento de imagen

El principal objetivo de las técnicas de mejoramiento de imagen, es procesar una imagen con el fin de enfatizar o realzar características para un mejor despliegue y un procesamiento posterior, depende por tanto del problema específico a resolver el que se emplee una u otra técnica. En general son desarrolladas en forma empírica

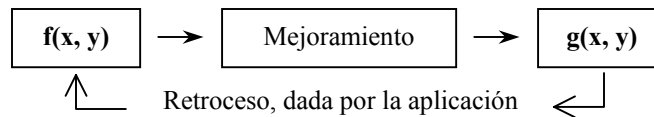


Fig. 4.11 Proceso del mejoramiento de la imagen

Los métodos de mejora de imagen se pueden dividir en dos campos diferentes:

- Métodos en el dominio espacial, se basan en manipulaciones directas sobre los pixeles de la imagen, caen en tres categorías:
 1. Operaciones puntuales.
 2. Operaciones basadas en vecindades.
 3. Operaciones globales.

- Métodos en el dominio frecuencial, se intenta modificar las componentes espectrales de la imagen, se basan en la transformada de Fourier.

4.6.1. Métodos que modifican el rango de niveles de gris

Estos métodos son esencialmente puntuales. Normalmente se usan funciones de recorrido o mapeo lineales para encontrar los nuevos valores de niveles de gris de la imagen, sin embargo se pueden usar también transformaciones no lineales.

Se aplican para el mejoramiento de contraste y realce de rasgos.

- Las operaciones básicas son usadas para expandir y/o comprimir el rango dinámico de una imagen.
- Otras veces se requiere expandir el rango dinámico y cortar valores abajo y arriba de ciertos límites.
- Hay otras técnicas como el rebanado de niveles de gris (gray-level slicing).

4.6.2. Técnicas de modificación del histograma

Estas técnicas van principalmente enfocadas a mejorar la visualización de una imagen. El histograma de una imagen es un gráfico que ofrece una descripción global de la apariencia de la misma, en el eje de abscisas se representa el rango de valores de píxeles, mientras que en el eje de ordenadas se representa el rango de valores que pueden tomar esos píxeles.

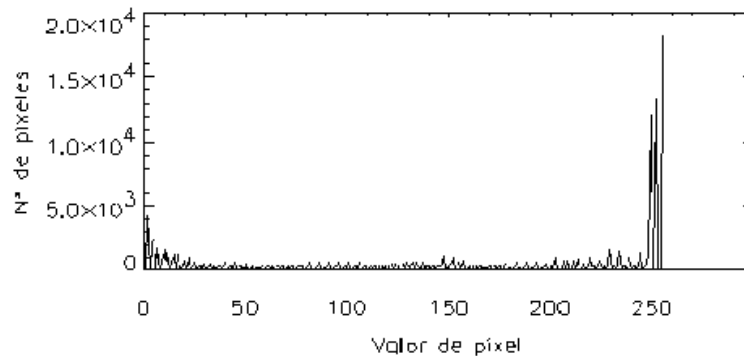


Fig. 4.12 Histograma de una imagen

El histograma de una imagen nos da la distribución de sus niveles de gris.

Si L es el número de niveles de gris de una imagen y $[0, L - 1]$ el rango de niveles de gris, entonces el histograma de la imagen es una función discreta:

$$p(\gamma_k) = n_k / n$$

- γ_k Es el k-esimo nivel de gris.
- n_k Es el número de pixeles del k-esimo nivel de gris.
- n Es el número de pixeles en la imagen.
- k 0, 1, 2, ..., $L - 1$

En general, un histograma:

- Con poca expansión involucra un bajo contraste.
- Con gran expansión involucra un alto contraste
- Con valores cargados hacia la izquierda involucra poco brillo (negra).
- Con valores cargados hacia la derecha involucra mucho brillo (clara).

El histograma puede ser usado para modificar los niveles de gris de una imagen y por lo tanto su apariencia, a través de:

- Expandir
- Comprimir
- Desplazar
- Cortar

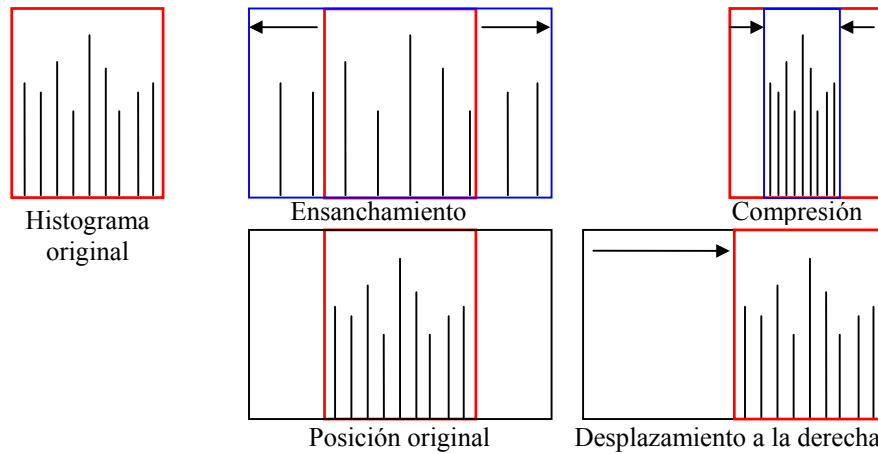


Fig. 4.13 Modificación de un histograma en forma gráfica

4.6.3. Modificación por histograma

Por expansión de la ecuación de recta, la ecuación de mapeo es:

$$g(x, y) = \frac{f(x, y) - \min}{\max - \min} (M - m) + m \quad (4.2)$$

max Valor de nivel de gris más grande en $f(x, y)$.

min Valor de nivel de gris más pequeño en $f(x, y)$.

M, m Son los valores \max' y \min' respectivamente a los cuales se quiere ensanchar el rango dinámico de la imagen, con $M \geq \max$ y $m \leq \min$ según sea el caso.

Algunas veces, si la mayoría de los valores caen dentro de un rango pequeño y unos cuantos en el complemento, este método no dará muy buenos resultados. En éste caso es útil cortar primero un porcentaje de los extremos inferiores y superiores del histograma y después expandir.

La operación contraria de expansión es la contracción, ésta disminuirá el contraste en la imagen al comprimir el rango dinámico.

$$g(x, y) = \frac{f(x, y) - \min'}{\max' - \min'} (M' - m') + m'$$

\max' Valor de nivel de gris más grande en $f(x, y)$.

\min' Valor de nivel de gris más pequeño en $f(x, y)$.

M', m' Son los valores \max y \min respectivamente a los cuales se quiere disminuir el rango dinámico de la imagen, con $M' \leq \max'$ y $m' \geq \min'$ según sea el caso.

La operación de desplazamiento del histograma, que puede usarse para oscurecer o aclarar una imagen, puede realizarse como sigue:

$$g(x, y) = f(x, y) \pm \text{offset}$$

offset es la cantidad de desplazamiento.

4.6.4. Filtros

Normalmente la generación de ruido se produce de forma aleatoria e inesperada y como causa de un mal control tanto en la fase de digitalización de la imagen como en la fase de tratamiento de la misma.

Aunque la terminología del filtrado es esencialmente del dominio de las frecuencias, el teorema de convolución nos permite hablar también de filtros en el dominio espacial aunque el sentido de dichos filtros no sea tan claro como en el dominio de las frecuencias.

Lo que viene a definir que la operación convolución de dos funciones en el dominio del espacio es igual que el producto de las mismas en el dominio de las frecuencias [González, 1996⁸], por tanto saber que formas adoptarán las funciones asociadas a determinados filtros espaciales nos dará el resultado de la operación.

Una característica común a todos los tipos de datos raster es la llamada frecuencia espacial, que define la magnitud de cambios de los datos por unidad de distancia en una determinada zona de la imagen:

- Áreas de la imagen con pequeños cambios o con transiciones graduales en los valores de los datos se denominan áreas de bajas frecuencias (como por ejemplo la superficie de una masa de agua en reposo).
- Áreas de grandes cambios o rápidas transiciones se conocen como áreas de altas frecuencias (por ejemplo suelo urbano con densas redes de carreteras).

Tanto los filtros espaciales como los de frecuencia tienen la misma base y su aplicación en uno u otro depende del lado de la igualdad del teorema de convolución en que estemos. De todas formas todos estos filtros los podemos clasificar en dos tipos:

- **Filtros de suavizado.**
Los cuales sirven para emborronar una imagen o para reducir su ruido. Los procesos de emborronamiento son fundamentales en muchos procesos de cálculo de procesamiento de una imagen.
- **Filtros de realce.**
Los cuales sirven para resaltar aquellas características de la imagen que por causas del mecanismo de captación, por error o por necesidad de resaltarla hayan quedado o estén emborronadas.

A continuación se describirán un conjunto de filtros muy utilizados que se pueden ubicar en cualquiera de las divisiones anteriores:

- **Mediana.**
El principal objetivo de este filtro es reducir el ruido sin emborronar las fronteras, por lo que podrá estar situado dentro de los filtros no lineales como un filtro de suavizado.
- **Enfatizado.**
Es un filtro tipo o conjunto de operaciones que van a realzar las altas frecuencias sin perder la contribución de las bajas, por lo que también se trata de un filtro no lineal, pero esta vez de realce.

- **Media.**
Por fin un filtro lineal de suavizado el cual produce una distribución normal de los píxeles dentro de la imagen, realizando un proceso de emborronamiento como es común a todos los filtros de suavizado.
- **Derivada o gradiente.**
También es un tipo genérico de filtros los cuales realizan el cálculo del gradiente de la imagen con diferentes estimadores, así tendremos los estimadores de Robert, Prewitt y Sobel los cuales se suelen utilizar comúnmente para la segmentación de imágenes y para la detección de bordes.

4.7. Clasificadores

Clasificar una imagen consiste en tomar decisiones, estas decisiones pueden ser del tipo asignar algún píxel a una determinada categoría A, siempre en función de un criterio de clasificación y atendiendo a alguna medición sobre el elemento a clasificar.

El objetivo de toda clasificación es el reconocimiento de clases o grupos cuyos miembros tengan ciertas características en común, el resultado ideal consiste en obtener clases que sean mutuamente excluyentes y exhaustivas.

Al conjunto de valores de un píxel podemos verlo desde el punto de vista de la clasificación como un patrón, y al conjunto de características o variables en las que se han tomado esos píxeles se les denominan rasgos, así, un patrón es un conjunto de medidas en una serie de rasgos, y la clasificación puede definirse como una forma de reconocimiento de patrones, donde los pasos a seguir serían:

1. Selección de los rasgos particulares que mejor describan el patrón.
2. Elección de un método adecuado para la comparación de los patrones.
3. Como paso opcional, la comprobación de los resultados; medición del grado de exactitud conseguido con esa clasificación.

Podemos diferenciar dos tipos principales de clasificación: supervisada, si partimos de un cierto conocimiento sobre la zona a clasificar y a partir del cual obtendremos criterios de clasificación, y no supervisada, si no tenemos ningún conocimiento previo y de lo que se trata es de discriminar píxeles en función de una magnitud cuantitativa de diferenciación entre unos y otros.

Parece lógico que la distancia euclidiana es una medida importante en la determinación de los parámetros entre clases, también es importante la escala en la cual se expresan los valores numéricos, si se cambia la escala de, por ejemplo el eje x, vemos que la interpretación de las relaciones entre puntos también cambian.

4.7.1. Clasificación supervisada

Previamente se han de seleccionar muestras de cada clase constitutiva, estas zonas, conocidas como zonas de entrenamiento, serán procesadas a fin de calcular sus estadísticas (media, rango, desviación típica, etc.) en todas las partes que intervengan en la clasificación.

4.7.2. Clasificación no supervisada

Supone la búsqueda automática de grupos de valores homogéneos, la intervención humana se centra más en la interpretación de resultados que en la consecución de los mismos.

Se proponen también algunos métodos mixtos, ya que el método supervisado puede resultar a veces demasiado subjetivo y el no supervisado puede carecer de significado físico real.

Independientemente del método utilizado, se hace necesario efectuar un análisis de las estadísticas, para comprobar que efectivamente se podrá efectuar una discriminación de los datos.

4.7.3. Fase de entrenamiento

Esta fase, que sólo se efectúa si se trata de una clasificación supervisada, consiste en seleccionar muestras, en forma de zonas de píxeles concretos, que se asignarán por el usuario a una clase concreta, basándose en el conocimiento previo que se tiene sobre la imagen, a partir de los puntos de las zonas de entrenamiento se extraerán ciertos parámetros estadísticos, la elección de los parámetros a calcular dependerá del método concreto de asignación que se vaya a emplear después.

Es fundamental la correcta selección de las muestras, ya que en caso contrario, las estadísticas no serán las correctas y los resultados pueden ser completamente erróneos. por otro lado, la validez de los resultados dependerá de dos factores:

1. El grado de fiabilidad de los posibles supuestos previos que se hayan hecho.
2. La representatividad de las muestras tomadas.

Las suposiciones que se toman como ciertas a priori sobre los datos, dependerán también del método de asignación empleado, aquí también, el método más complejo es en el que se realizan las suposiciones más rígidas.

4.7.4. Algoritmo Máximos y Mínimos

Es un algoritmo rápido marcadamente heurístico que emplea como único elemento formal la distancia euclídea, no requiere ninguna información *a priori* respecto al número de clases existentes.

Partiendo de un conjunto de muestras o vectores a agrupar, los pasos son los siguientes:

- Paso 1. Se escoge al azar un elemento individual de los P disponibles, digamos X_i , y se crea la clase α_i .
- Paso 2. Se calculan las distancias euclídeas de los $P-1$ vectores restantes no agrupados al elemento X_i y se toma la máxima distancia, de tal manera que el elemento implicado X_j produce la segunda clase α_2 .
- Paso 3. De nuevo se calculan las distancias de los $P-2$ elementos restantes no agrupados a los elementos X_i y X_j , de donde se toma la mínima del par de distancias calculadas, del total de distancias mínimas calculadas, se toma la máxima de ellas creando la tercer clase α_3 .
- Paso 4. El proceso es ahora similar, teniendo en cuenta que cada vez existe una clase más.
- Paso 5. Se continúa hasta que se cumpla la condición:

$$d_{max} > f * \text{distancia media entre clases}$$
 Donde, f es un umbral tal que, $0 < f < 1$, a partir del cual ya no es posible crear más clases.
- Paso 6. Los elementos restantes que quedan por agrupar se asignan a la clase más cercana.

Este algoritmo tiene un serio inconveniente, ya que es muy sensible al coeficiente f .

4.7.5. Algoritmo K-Medias

Este algoritmo ya esta consagrada en la literatura especializada y hace referencia a que existen K clases o patrones, por tanto es necesario conocer *a priori* el número de clases existentes.

Partimos de que se conocen con exactitud el número de clases, digamos K , de los P elementos a clasificar, los pasos son los siguientes:

- Paso 1. Se escogen K elementos, de forma que éstos van a constituir los centroides de las K clases.
- Paso 2. Se distribuyen los $P-K$ elementos entre las K clases, donde estos elementos se agregan a la clase más cercana.

- Paso 3. Una vez distribuidos los elementos entre las diferentes clases, se actualizan los centroides de las clases y se redistribuyen nuevamente.
- Paso 4. Se termina el algoritmo cuando ha alcanzado una situación estable, es decir, no existen cambios en la posición de los centroides, en otro caso se salta al paso 2.

Este algoritmo es sencillo y muy eficaz, pero se debe conocer con exactitud el parámetro K. Un valor superior al real dará lugar a clases ficticias, mientras que un K inferior dará menos clases a las reales.

4.7.6. Evaluación de las clasificaciones

Una vez que la clasificación se ha efectuado, podemos determinar el grado de error cometido, en términos de número de píxeles incorrectamente clasificados, los resultados de las clasificaciones tanto supervisadas como no supervisadas se pueden evaluar con la ayuda de las áreas de comprobación, el diseño de muestreo de dichas áreas y su localización debe realizarse previamente a la clasificación para evitar posibles sesgos, estas áreas de comprobación serán perfectamente conocidas y tendremos tabuladas la superficie que posee cada una de las clases, en estas mismas áreas se evalúan las superficies de cada clase que se han obtenido a través de las clasificaciones realizadas con la computadora y se comparan con los anteriores resultados, la utilización de áreas de entrenamiento para este propósito es claramente inadecuado.

Siempre se puede realizar una clasificación que se adapte a lo que marcamos, pero lo que intentamos medir es el grado de aproximación con la realidad, no con las muestras que hemos introducido, este es por tanto el inconveniente del método ya que es difícil conseguir esa información de referencia.

En el capítulo 6 se adoptará este método para evaluar a los clasificadores utilizados en la metodología propuesta.

4.8. Operadores Morfológicos

El lenguaje de la Morfología matemática binaria (tratado de formas) es el de la teoría de conjuntos, es un método no lineal para procesar imágenes digitales basándose en la forma, su principal objetivo es la cuantificación de estructuras geométricas, es decir, describir las regiones de una imagen como conjuntos, aquí los filtros también vienen definidos por su máscara, pero no es una máscara de convolución sino un elemento estructurante [Díaz-De-León 2000³⁰].

4.8.1. Dilatación

La dilatación binaria fue usada primero por Minkowski, y en la literatura matemática recibe el nombre de “Suma de Minkowski”, viene definida por la siguiente expresión:

$$A \oplus B = \{c \mid c = a + b \text{ para algún } a \in A \text{ y } b \in B\} = \bigcup_{b \in B} A_b$$

Donde A_b representa la traslación de A por b . intuitivamente, para cada elemento no-cero b_{ij} de B , A es trasladado i, j y sumado con el operador “or”, por ejemplo en la figura 4.14.

En la práctica los conjuntos A y B no son simétricos, el primer elemento de la dilatación, A , está asociado con la imagen que se está procesando y el segundo recibe el nombre de elemento estructural, la forma que actúa sobre A en la dilatación para producir $A \oplus B$.

La dilatación usado con imágenes binarias, donde cada píxel es 1 ó 0, la dilatación es similar a la convolución, sobre cada píxel de la imagen se superpone el origen del elemento estructurante, si el píxel de la imagen no es cero, cada píxel que cae en la estructura es añadido al resultado aplicando el operador “or”.

Cuando se realiza una dilatación con un elemento estructural que contiene el cero, lo que realizamos es la expansión de una imagen y es fácil pensar en una implementación paralela.

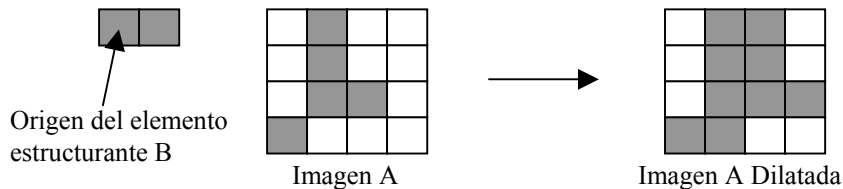


Fig. 4.14 Proceso de dilatación

En el ejemplo de la figura 4.14, el origen del elemento estructurante B está en la posición $(0,0)$.

Usado con imágenes en escala de grises, la dilatación se efectúa tomando el máximo de una serie de sumas, puede ser usado para implementar el operador de “máxima vecindad” con la forma de la vecindad dada en el elemento estructurante.

4.8.2. Erosión

La erosión (denominada por algunos autores substracción de Minkowsky) viene definida por la siguiente expresión.

$$A \ominus B = \{x \mid x + b \in A \text{ para cada } b \in B\} = \bigcap_{b \in B} A_{-b}$$

Donde $A_{\cdot b}$ representa la traslación de a por b ; una interpretación cualitativa de esta operación puede ser la siguiente: Se copia B en cada píxel de A y se marcan los píxeles de A en los cuales la copia de B este totalmente contenida en A . La erosión no es conmutativa ni asociativa.

La erosión es casi lo opuesto a la dilatación; realiza con el fondo lo que la dilatación al primer plano, también en este caso, existe un elemento estructurante que se utiliza para operar con la imagen, los efectos son de “encogimiento”, “contracción”, o “reducción”, puede ser utilizado para eliminar islas menores en tamaño que el elemento estructurante.

Sobre cada píxel de la imagen se superpone el origen del elemento estructurante, si cada elemento no-cero de dicho elemento está contenido en la imagen, entonces el píxel de salida es puesto a 1, haciendo a $\ominus b$ como representación de la erosión de una imagen a por el elemento estructurante b , por ejemplo en la figura 4.15.

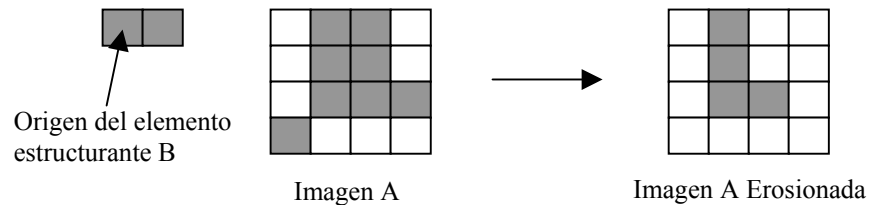


Fig. 4.15 Proceso de erosión

En el ejemplo de la figura 4.15, el origen del elemento estructurante B está en la posición $(0,0)$.

Usado en imágenes en escala de grises, la erosión se efectúa tomando el mínimo de una serie de diferencias, puede ser usado para implementar el operador de “mínima vecindad” con la forma de la “vecindad” dada por el elemento estructurante.

4.8.3. Apertura y Cierre

La apertura de una imagen A por un elemento estructurante B , se define como:

$$A \circ B = (A \ominus B) \oplus B$$

El cierre de la imagen A por elemento estructurante B , se define como:

$$A \bullet B = (A \oplus B) \ominus B$$

El resultado de aplicar iterativamente dilataciones y erosiones es la eliminación del detalle específico en la imagen menor que el elemento estructurante, sin la distorsión geométrica global de características no suprimidas, por ejemplo:

- Abrir una imagen con una estructura en disco, suaviza los contornos, rompe istmos y elimina pequeñas islas, picos y cabos, ver la figura 4.16, y
- Cerrar una imagen con un elemento estructurante en forma de disco, elimina pequeños agujeros y rellena brechas en los contornos, ver figura 4.17.



Fig. 4.16 Apertura Morfológica



Fig. 4.17 Cerradura Morfológica

4.8.4. Hit or Miss y adelgazamiento

Se puede combinar la erosión y la dilatación para producir un operador "hit or miss" que opere de la siguiente forma: Unos elementos realizan "hit" (acertar) y otros elementos realizan "miss" (fallar); este operador puede definirse como sigue:

$$A \otimes (J, K) = (A \ominus J) \cap (A^c \oplus K)$$

El operador "hit or miss" puede interpretarse de la siguiente forma: El elemento estructural J busca el mejor contorno interno y el elemento estructural K busca el mejor contorno externo.

Se puede utilizar el operador "hit or miss" para añadir o eliminar puntos de un objeto; si se eliminan se realiza un adelgazamiento, éste es el resultado de quitar a la imagen original el contorno encontrado por medio de "hit or miss" y el operador viene definido por la siguiente expresión [Díaz-De-León 2000³⁰]:

$$A - (A \otimes (J, K))$$

4.9. Tipos de imágenes

Como vimos en la sección 4.2, la función de imagen digital $f(x,y)$ representa un valor de brillo en ese punto, pero cuando se quiere representar una imagen en color, esa interpretación puede cambiar, por ejemplo, en función de la división de colores en un cubo RGB (Red-Green-Blue), podemos descomponer toda la imagen en tres bandas, cada una de ellas representando el brillo de rojo, de verde y de azul respectivamente.

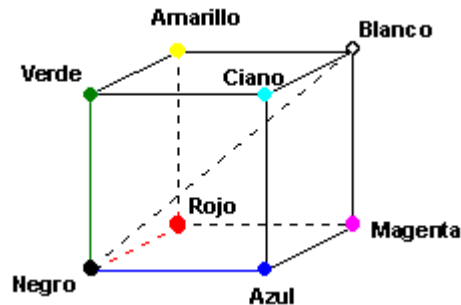


Fig. 4.18 Cubo RGB

El color visualizado en un punto concreto será el resultante de combinar los valores de ese punto en las tres bandas, en la proporción indicada por esos valores o bien se puede tener una tabla de colores (una paleta) y hacer que los valores de la imagen estén “apuntando” a los colores de esa paleta, de cualquiera de las formas podemos distinguir: blanco y negro, escala de grises, color indexado 256 y color real, los tipos de datos se diferencian por la resolución en bits (número de bits de información de la imagen por píxel) y por el número de canales que comprende la imagen, la resolución en bits determina cuántos colores o niveles de gris puede representar cada píxel.

El tipo de imágenes con las que nos enfrentamos pueden caer en algunas de las siguientes clases:

- Imágenes binarias
- Imágenes en niveles de gris (monocromáticas)
- Imágenes en color

4.9.1. Imágenes binarias

Estas son las más simples con las que podemos tratar, cada píxel de una imagen en blanco y negro es normalmente blanco o negro.

$$\text{Negro} = 0 \quad \text{y} \quad \text{Blanco} = 1.$$

Los tonos intermedios (grises) se crean ordenando los píxeles blancos y negros para simular gris, es lo que se llama técnica de difuminado (dithering).

Una imagen binaria es referida también como una imagen 1 bit/píxel, se usan sobretodo donde se requiere forma o contorno.

Son creados normalmente a partir de imágenes en niveles de gris por umbralización U , donde cada píxel p :

- Arriba de un umbral, $p \rightarrow 1$.
- Abajo e igual del mismo umbral, $p \rightarrow 0$.

El problema aquí es determinar el mejor nivel de umbralización U .

4.9.2. Imágenes en niveles de gris

Son conocidas también como monocromáticas o imágenes a un solo color, contienen solo información de brillo (no de color), cada píxel de una imagen en escala de grises puede ser uno de los 256 valores distintos de gris, del negro (0) al blanco (255), este tipo de datos muestra suaves cambios de tonalidad utilizando valores intermedios de gris.

La resolución de una imagen en escala de grises determina el tamaño de los píxeles y en consecuencia el número de píxeles de una imagen, cuanto mayor sea la resolución los cambios de tono de gris serán más suaves y por tanto más exacta la representación de la imagen (las imágenes de alta resolución también utilizan más memoria).

Una imagen típica contiene 8 bits/píxel \rightarrow 256 niveles de grises \rightarrow (0, 255). Sin embargo en aplicaciones como astronomía o medicina, 12 o incluso 16 bits/píxel son normales.

4.9.3. Imágenes en color

En una imagen en color real RGB, cada píxel puede ser uno de los más de 16,7 millones de colores disponibles, el tipo de datos de color real RGB es actualmente el mejor tipo de datos para editar imágenes en color debido a que retiene mayor cantidad de información de la imagen.

Se pueden modelar como 3 imágenes monocromáticas, cada una relacionada con su banda espectral. En general se usan 24 bits/píxel, 8 bits para cada color.

La relación que existe entre sus componentes para pasarlo a una sola dimensión es:

$$f(x, y) = R \cdot 2^0 + G \cdot 2^8 + B \cdot 2^{16}$$

En muchas aplicaciones la información RGB es transformada a otro espacio matemático y permite desacoplar la información de brillo de la del color.