

S.E.P.

S.E.S.

D.G.E.S.T.

INSTITUTO TECNOLÓGICO DE TOLUCA

REPORTE FINAL DE AÑO SABATICO

OPCIÓN: Libro de Texto

TRATAMIENTO DIGITAL DE IMÁGENES

P R E S E N T A:

M. en C. ITZEL ABUNDEZ BARRERA

CONTENIDO

INTRODUCCIÓN	5
1.1 Un modelo simple de la imagen.	7
1.2 Muestreo y Cuantificación.	8
1.3 Relación entre píxeles.	9
Vecindad y Conectividad.	9
1.4 Medidas de distancia.	13
1.5 Áreas de Aplicación y Tareas Típicas.	13
GEOMETRÍA DE LA IMAGEN	16
2.1 Transformaciones Básicas.	16
Translación.	16
Transformación de Escala.	17
Rotación	17
Ángulos de las diversas rotaciones.	18
2.2 Transformaciones de perspectivas.	19
2.3 Modelo de Cámara.	23
2.4 Calibración de Cámara.	26
2.5 Imagen Estereoscópica.	36
MEJORAMIENTO DE LA IMAGEN DIGITAL	39
3.1 Introducción.	39
3.2 Operadores de punto	40
Ecuilización del Histograma	41
Procesamiento de Histograma	41
Mejoramiento de imagen.	42
Aumento del contraste de la imagen.	46
3.2 Operadores Locales en el Rango del Espacio.	48

Filtro espacial paso alto básico	48
3.3 Mejoramiento en el rango de frecuencia.	50
Método del Gradiente.....	51
Filtros estadísticos	52
Mejoramiento en el dominio espacial	53
filtros de paso bajo	53
Filtros de paso alto	54
Filtro de paso banda	55
EXTRACCIÓN DE CARACTERÍSTICAS DE LA IMAGEN DIGITAL	56
4.1 Detección de Discontinuidades.....	56
4.2 Detección de curvas y Líneas y transformada de Hough.....	58
4.3 Información de Profundidad por medio de la Geometría.....	58
4.4 Orientación de la superficie por modelos de Reflexión.¡Error! Marcador no definido.	
4.5 Flujo óptico.....	65
SEGMENTACIÓN DE IMÁGENES	67
5.1 Detección de Discontinuidades	68
5.2 Enlazado de Bordes y Detección de Límites.....	70
5.2.1. Método basado en las primeras derivadas.....	70
5.2.2. Método del Gradiente.....	71
5.2.3. Método Roberts.....	72
5.2.4. Método Prewitt	72
5.2.5. Método Sobel	73
5.2.6. Método Basado en la Segunda Derivada	73
5.2.7. Método Laplaciano Gaussiano	74
5.2.8. Método de Kirsch	75

5.3 Umbralización.	76
5.4 Segmentación Orientada a Regiones.	79
5.5 Utilización de Movimiento.	81
REPRESENTACIÓN Y DESCRIPCIÓN	85
6.1 Esquemas de Representación.	85
6.2 Descriptores de Contorno.	87
6.3 Descriptores de Región.....	87
Métodos Estadísticos	87
Métodos Estructurales.....	93
Métodos Espectrales.....	94
6.4 Descriptores Relacionales.	95
Descripción de la imagen	95

INTRODUCCIÓN

Una de las disciplinas de la Inteligencia Artificial (IA) es la Visión Artificial (VA). La VA desarrolla sistemas para dar a una computadora la capacidad de visualizar el mundo real e interactuar inteligentemente con él.

Los sistemas de VA tienen múltiples aplicaciones, algunas de ellas son:

- Reconocimiento automático de caracteres
- Reconocimiento automático de huellas digitales
- Mejora de imágenes de rayos X
- Reconocimiento de imágenes satelitales
- Análisis de imágenes medicas, entre otras.

Para aplicar la visión artificial es necesaria de un área como lo es el Tratamiento Digital de Imágenes (TDI), la cual deriva de dos áreas principales de aplicación: la mejora de la información pictórica para la interpretación humana y el procesamiento de los datos de una escena para la percepción autónoma por una máquina [González, 1996].

El Tratamiento Digital de Imágenes está formado por seis etapas que se encuentran interrelacionadas. En la figura 1 se muestra un diagrama representando las etapas y sus relaciones [González, 1996].

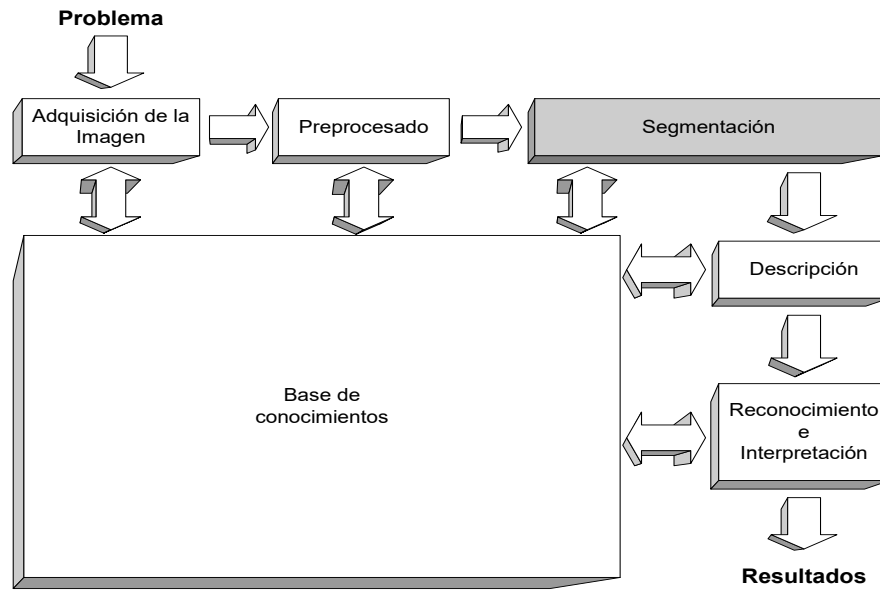


Figura 1.1 Etapas del Tratamiento Digital de Imágenes

Adquisición de la imagen digital: Se requiere de dos elementos, el primero es un dispositivo físico sensible a una determinada banda del espectro electromagnético (rayos X, ultravioleta, visible o infrarrojo), que produzca una señal de salida, el segundo, denominado digitalizador, es un dispositivo para convertir la señal de salida del sistema sensible a forma digital. Puede ser a través de una cámara de video, cámara fotográfica, etc.

Preprocesamiento de la imagen: Su función básica es la de mejorar la imagen, trata típicamente las técnicas de mejora de contraste, eliminar ruido y aislar regiones cuya textura indiquen información alfanumérica.

Segmentación: Consiste en dividir una imagen en sus partes constituyentes u objetos.

Representación y descripción: En la representación van a transformar los datos de píxel en bruto a una forma adecuada para ser posteriormente tratados por computadora. La descripción consiste en extraer rasgos con alguna información cuantitativa de interés o que sean fundamentales para diferenciar una clase de objetos de otra.

Reconocimiento e interpretación: Es el proceso en el que se asigna una etiqueta a un objeto basándose en la información proporcionada por sus descriptores. La interpretación es el proceso que asocia un significado a las etiquetas obtenidas en el reconocimiento.

El termino imagen se refiere a una función bidimensional de intensidad de luz $f(x, y)$, donde x y y , representan las coordenadas espaciales y el valor de f en un punto cualquiera (x, y) es proporcional al brillo de la imagen en ese punto [González, 1996].

1.1 UN MODELO SIMPLE DE LA IMAGEN.

Según González (1996), una imagen digital es una función bidimensional de intensidad de luz $f(x, y)$ digitalizada, donde x y y representan las coordenadas espaciales y f es el valor de intensidad (iluminación) de la imagen. Puesto que la luz es una forma de energía $f(x, y)$ debe ser mayor que 0 y finita, es decir:

$$0 < f(x, y) < \infty \quad (\text{Ec. 1})$$

Las imágenes que percibimos cotidianamente provienen normalmente de la luz reflejada por los objetos, pudiendo estar caracterizada por dos componentes como son la cantidad de luz incidente preecedente de la fuente sobre la escena contemplada y la cantidad de luz reflejada por los objetos de la escena. Para definir apropiadamente son componentes de iluminación y reflectancia indicada por $i(x, y)$ y $r(x, y)$.

En la figura 1.2 se observan tres figuras donde los óvalos cuentan con la misma intensidad, lo que hace el efecto de tener diferente intensidad es el fondo, entre más claro este sea el ovalo se percibirá más intenso.

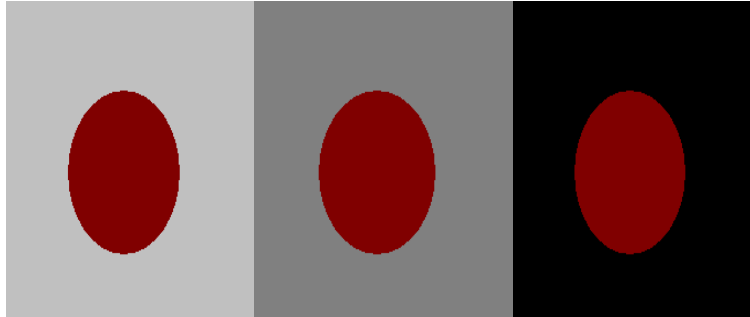


Figura 1.2 Ejemplo de contraste simultáneo

1.2 MUESTREO Y CUANTIFICACIÓN.

Una imagen digital puede ser representada por una matriz de dos dimensiones $N \times M$, donde sus elementos son llamados píxeles (Abreviatura de pictures elements), cuyos índices de fila y columna identifican un punto de la imagen que son obtenidas por sus coordenadas, y el valor del correspondiente píxel de la matriz indica una característica de dicho píxel, estas características pueden ser el nivel de gris, un color en particular o incluso el tipo de textura. Véase la figura 1.3.

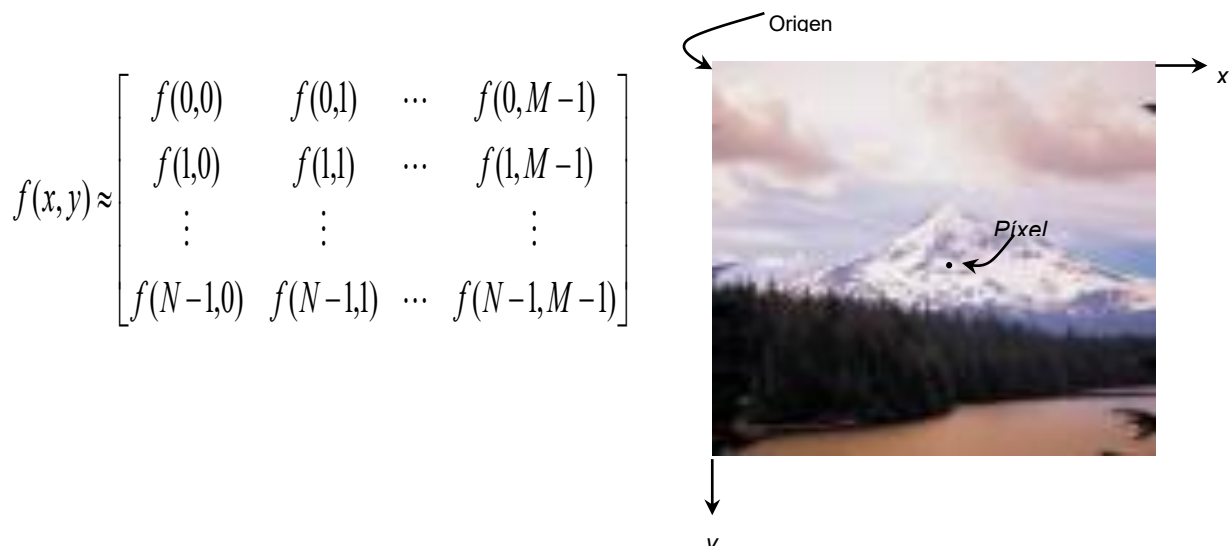


Figura 1.3 Representación Matricial de una imagen

1.3 RELACIÓN ENTRE PÍXELES.

Cuando se hace referencia a un pixel en particular, se emplearán las letras minúsculas, como p y q . Un subconjunto de pixeles de $f(x,y)$ se indican mediante S .

VECINDAD Y CONECTIVIDAD

En la descripción e interpretación de imágenes es necesario saber cuándo 2 píxeles son vecinos, o bien cuando un grupo de píxeles es conexo.

Para definir Vecindad y Conectividad se utilizan al menos tres tipos de concepto de vecinos:

- vecinos tipo 4.
- vecinos tipo 8.
- vecinos tipo 6.

Nota.- En la figura 1.4 se ilustran los diferentes tipos de vecindad.

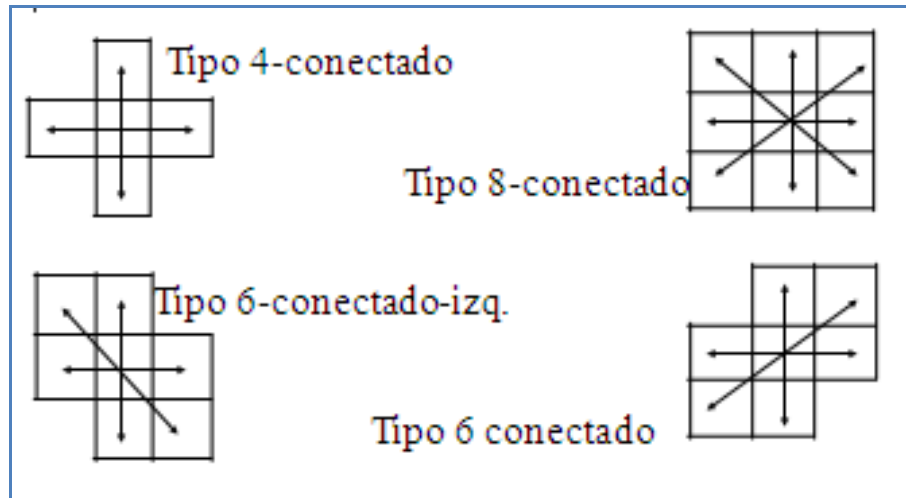


Figura 1.4 Tipos de Vecindad

La figura 1.5 muestra la relación entre el píxel que se está evaluando

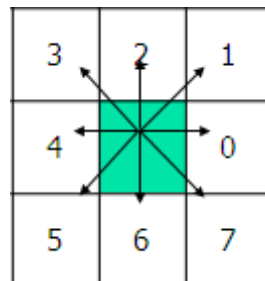


Figura 1.5 Vecindad entre píxeles

La vecindad tipo 4- conectado son:

- Cuando dos píxeles comparten un borde son vecinos del tipo 4- conectados.
- El píxel (x, y) tiene 4 vecinos de este tipo:

$$\checkmark (x-1, y), (x+1, y), (x, y-1), (x, y+1)$$

La figura 1.6 ilustra la conectividad tipo 4-connectado.

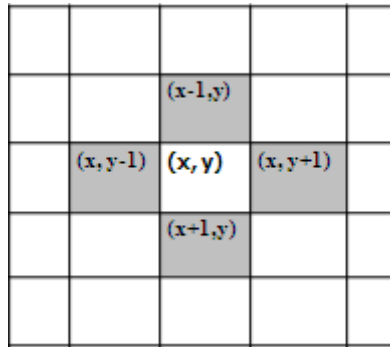


Figura 1.6 Conectividad tipo 4-conectado

La vecindad tipo 8 corresponde a los vecinos superior, inferior izquierda, derecha y los unidos por sus esquinas como lo ilustra la figura 1.7 donde el píxel a evaluar es el que se encuentra en el centro de tono claro.

- Cuando dos píxeles comparten un borde o una esquina son vecinos tipo 8-conectados.
- Cada píxel (x, y) tiene 8 vecinos de este tipo:

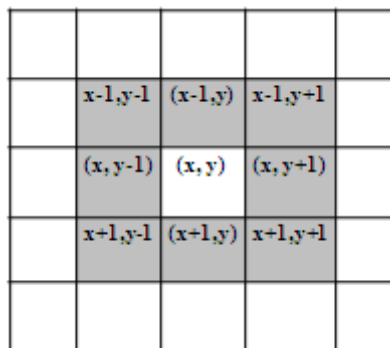
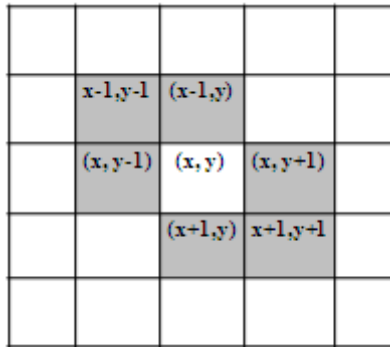


Figura 1.7 Conectividad tipo 8-conectado

La vecindad tipo 6 - conectado- izq. se da:

- Cuando dos píxeles comparten un borde o las esquinas en la dirección NO-SE, son vecinos tipo 6-conectados- izq.
- Cada píxel (x, y) tiene 6 vecinos de este tipo:

En la figura 1.8 se ilustra este tipo de conectividad.

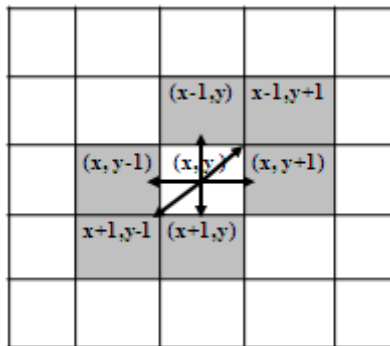


1.8 Conectividad tipo 6-conectado izq.

La vecindad tipo 6-conectado-derecha está dada:

- Cuando dos píxeles comparten un borde o las esquinas en la dirección NE-SO, son vecinos tipo 6-conectados-der.
- Cada píxel (x, y) tiene 6 vecinos de este tipo:

La figura 1.9 ilustra el tipo de conectividad.



1.9 Conectividad tipo 6-conectado der.

1.4 MEDIDAS DE DISTANCIA.

Para los píxeles p , q y z , de coordenadas (x, y) , (s, t) y (u, v) respectivamente D es la función de distancia o métrica si:

a) $D(p, q) \geq 0 (D(p, q) = 0 \text{ si } p = q)$,

b) $D(p, q) = D(q, p)$,

c) $D(p, z) \leq D(p, q) + D(q, z)$.

Distancia Euclidiana entre p y q

$$D_e(p, q) = \left[(x - s)^2 + (y - t)^2 \right]^{1/2}$$

1.5 ÁREAS DE APLICACIÓN Y TAREAS TÍPICAS.

Existen cuatro grandes áreas de aplicación de la VA a saber:

- Obtención de la distancia de los objetos en la escena tridimensional.
- Detección de objetos en movimiento.
- Reconocimiento de patrones y formas.
- Reconocimiento de objetos tridimensionales.

El objetivo final está en utilizar la visión para reconocer, manipular, predecir, navegar, etc.

El reconocimiento de formas y patrones, aún siendo un concepto general, puede concretarse en que los diferentes objetos van a estar caracterizados por una determinada forma de contorno, que se describe de alguna manera o bien por la región que delimitan y sus propiedades subyacentes [Pajares, 2002].

El tratamiento de imágenes expuesto anteriormente, requiere un elevado costo computacional. En efecto, en el tratamiento de imágenes esto constituye un verdadero problema. Muchos analistas han desarrollado técnicas para ser implementadas en sistemas basados en arquitecturas paralelas, a este respecto una buena referencia la constituye el libro de Hussain [Hussain, 1999].

En la actualidad es común encontrar diversas aplicaciones en el campo de la Industria automotriz, ensamblado, soldadura, etc. Ver figura 1.10.

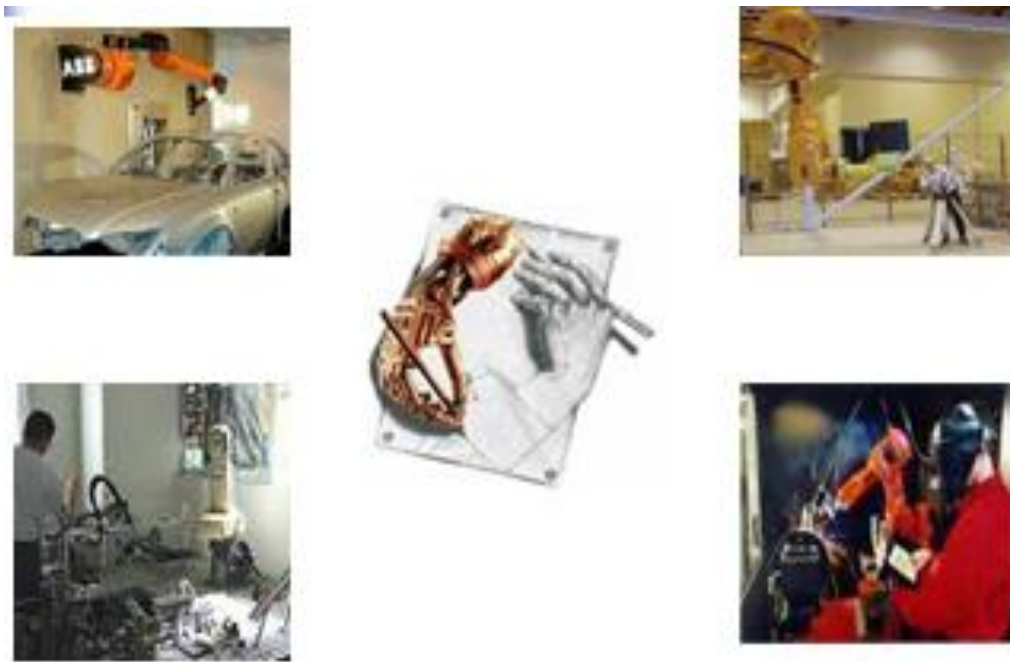


Figura 1.10 Aplicaciones en l industria

En el caso de la medicina la detección de zonas malignas como en el caso de imágenes colposcópicas, imágenes de mastografía o imágenes tomográficas, en el mejoramiento de la imagen para una mejor percepción como es el caso de las imágenes de rayos X y las imágenes de resonancia magnética. Ver figura 1.11.



Figura 1.11 Aplicaciones en la medicina

Existen otras aplicaciones como las imágenes multiespectrales, robótica en el espacio, satélites, etc. Ver figura 1.12.

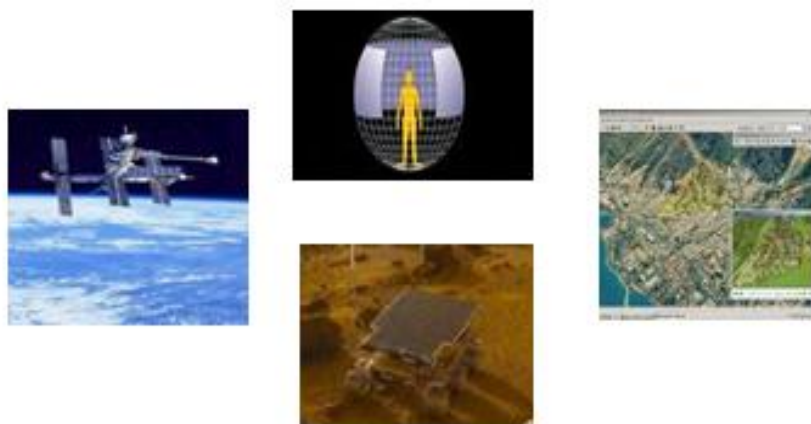


Figura 1.12 Otras aplicaciones

GEOMETRÍA DE LA IMAGEN

En este capítulo se mostrara las transformaciones básicas para la imagen, a través del desarrollo del modelo de cámara.

2.1 TRANSFORMACIONES BÁSICAS.

En los casos que indiquen imágenes bidimensionales (2-D), se mantendrá la convención anterior de representación en minúsculas (x, y) para indicar las coordenadas de un píxel.

TRANSLACIÓN.

Suponiendo que se desea trasladar un punto de coordenadas (X, Y, Z) a un nuevo emplazamiento, empleando los desplazamientos (X_0, Y_0, Z_0) . La traslación se realiza empleado las ecuaciones:

$$\begin{aligned}X' &= X + X_0 \\Y' &= Y + Y_0 \\Z' &= Z + Z_0\end{aligned}\tag{Ec.2.1.}$$

Esto puede escribirse en forma matricial como:

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & X_0 \\ 0 & 1 & 0 & Y_0 \\ 0 & 0 & 1 & Z_0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}\tag{Ec.2.2.}$$

Para normalizar la escritura y poder concatenar las diferentes transformaciones posibles, se introduce la cuarta coordenada y se trabaja en coordenadas homogéneas.

En este sistema de coordenadas, la translación está definida por la matriz de translación T :

$$T = \begin{bmatrix} 1 & 0 & 0 & X_0 \\ 0 & 1 & 0 & Y_0 \\ 0 & 0 & 1 & Z_0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (\text{Ec.2.3})$$

TRANSFORMACIÓN DE ESCALA

La matriz de transformación de escala está definida por:

$$S = \begin{bmatrix} a & 0 & 0 & 0 \\ 0 & b & 0 & 0 \\ 0 & 0 & c & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (\text{Ec.2.4.})$$

ROTACIÓN

Las diversas rotaciones alrededor de los ejes son efectuadas en el sentido contrario al de las agujas del reloj luego que se mira del origen a la parte positiva del eje. Ver figura 2.1.

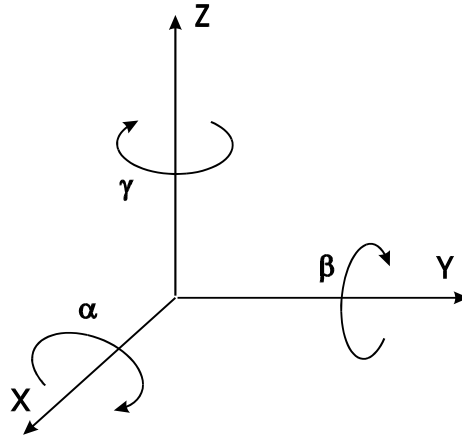


Figura 2.1 Rotación de un punto sobre cada uno de los ejes de coordenadas.

Ángulos de las diversas rotaciones.

La matriz de rotación de un ángulo γ alrededor del eje Z se escribe:

$$R_{\gamma} = \begin{bmatrix} \cos \gamma & \text{sen} \gamma & 0 & 0 \\ -\text{sen} \gamma & \cos \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (\text{Ec.2.5})$$

De la misma manera, las matrices de rotación de los ángulos α y β alrededor de OX y OY se escriben:

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

Una transformación compleja se escribe entonces por el producto de las matrices de transformaciones elementales correspondientes. Es necesario recordar que tales productos no cumplen la propiedad conmutativa.

2.2 TRANSFORMACIONES DE PERSPECTIVAS.

La imagen es el resultado de la proyección puntual de la escena sobre un plano: la transformación perspectiva que realiza, esto es una transformación no lineal a diferencia de las precedentes.

El sistema coordenado de referencia de la cámara (x, y, z) es el de la figura 2.2.

El plano imagen coincide con el plano (x, y) el eje óptico pasando por el centro de la lente coincide con el eje z . Ver figura 2.2.

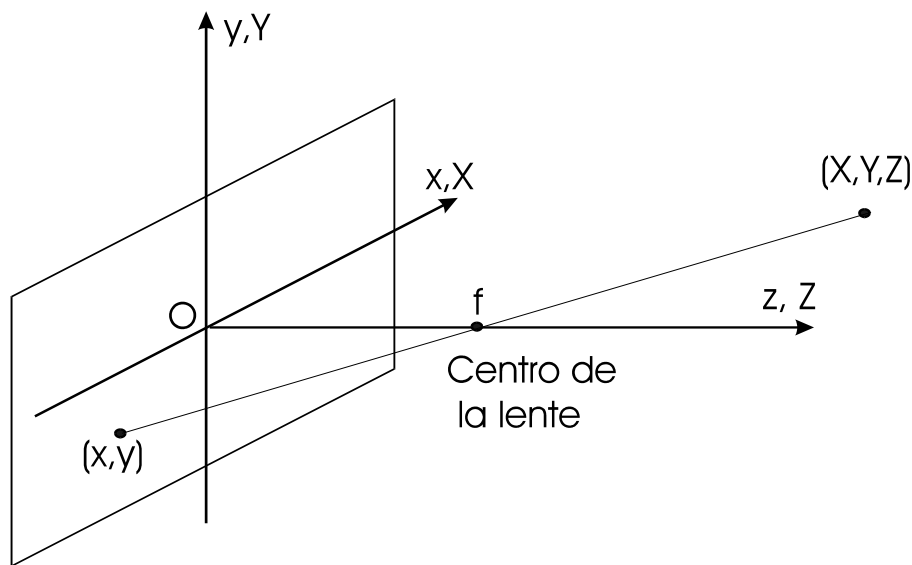


Figura 2.2 Modelo básico del proceso de la imagen

El centro del plano imagen está situado en el origen y el centro de la lente en el punto de coordenadas $(0, 0, f)$.

Si la calibración ha sido hecha, entonces f es la distancia focal de la lente.

Se supone, para la siguiente descripción, que el sistema de coordenadas de la cámara es idéntico al sistema del espacio (X, Y, Z) .

Sea (X, Y, Z) las coordenadas de un punto en el espacio con $Z > f$ y sea (x, y) su proyección sobre el plano imagen.

A continuación:

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

De donde se deducen las coordenadas dentro del plano imagen:

$$x = \frac{fX}{Z-f} \qquad y = \frac{fY}{Z-f}$$

Por homogeneidad con lo que precede, se expresa esta transformación no lineal bajo forma matricial recurriendo a las coordenadas homogéneas.

Las coordenadas homogéneas de un punto (X, Y, Z) son definidas por (kX, kY, kZ) donde k es una constante arbitraria no nula.

La matriz de transformación perspectiva está definida por:

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

Sea un punto w en coordenadas clásicas y sea w_h en coordenadas homogéneas k .

$$w = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \qquad w_h = \begin{bmatrix} kX \\ kY \\ kZ \\ k \end{bmatrix}$$

$$c_h = Pw_h$$

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

Las componentes de c_h son coordenadas homogéneas.

Se vuelve a las coordenadas clásicas dividiendo las tres primeras componentes por la cuarta.

$$c = \begin{bmatrix} \frac{kX}{f - Z} \\ \frac{kY}{f - Z} \\ \frac{kZ}{f - Z} \end{bmatrix}$$

Las dos primeras componentes son las coordenadas en el plano imagen, y la tercera no tiene interés.

La transformación perspectiva inversa está definida por la matriz P^{-1} .

$$w_h = P^{-1}c_h$$

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

Sea un punto c_h sobre el plano imagen:

$$c_h = \begin{bmatrix} kx \\ ky \\ 0 \\ k \end{bmatrix}$$

De esto se deduce el punto en coordenadas homogéneas w_h y el punto del espacio w :

$$w_h = \begin{bmatrix} kx \\ ky \\ 0 \\ k \end{bmatrix}$$

$$w = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Evidentemente, el punto no está en $Z=0$, pero se encuentra el hecho de que no se puede "reconstruir" a partir de una proyección.

El punto (x, y) corresponde al conjunto de puntos de R^3 que están situados sobre la recta pasando por $(x, y, 0)$ y $(0, 0, f)$.

Las ecuaciones de esta recta se escriben:

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

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

Reconstruir el punto exige entonces el conocimiento de Z .

A partir de esta observación, se puede formular la perspectiva inversa de forma diferente utilizando la componente Z como componente indeterminada:

$$c_h = \begin{bmatrix} kx \\ ky \\ kz \\ k \end{bmatrix}$$

$$w_h = \begin{bmatrix} kx \\ ky \\ kz \\ \frac{kz}{f} + k \end{bmatrix}$$

De donde, volviendo a las coordenadas cartesianas:

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

Expresando z en función de Z , se tiene:

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

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

2.3 MODELO DE CÁMARA.

A continuación se trabajara con un problema general donde los sistemas coordenados tanto el de la cámara como el real se encuentran en diferentes plano. En la figura 2.3 se muestra un sistemas de coordenadas (X, Y, Z) utilizadas para localizar los puntos de la camara como los 3D indicados por w , así como el sistema coordenado de la camara (x, y, z) y los puntos de la imagen indicados por c . Suponiendo que la camara esta montada sobre un soporte el cual le permite en el plano horizontal (*pan*) un ángulo θ e inclinarse (*tilt*) un ángulo α . El primer punto *pan* corresponde al ángulo entre los ejes y , mientras que *tilt* corresponde al angulo entre los ejes z y Z . La ceparación entre el centro del trípoide origen del sistema cooredenado global esta indicado por w_0 , y la separación del centro del

plano imagen con respecto al centro del pivote esta representado por el vector r , de componentes (r_1, r_2, r_3) .

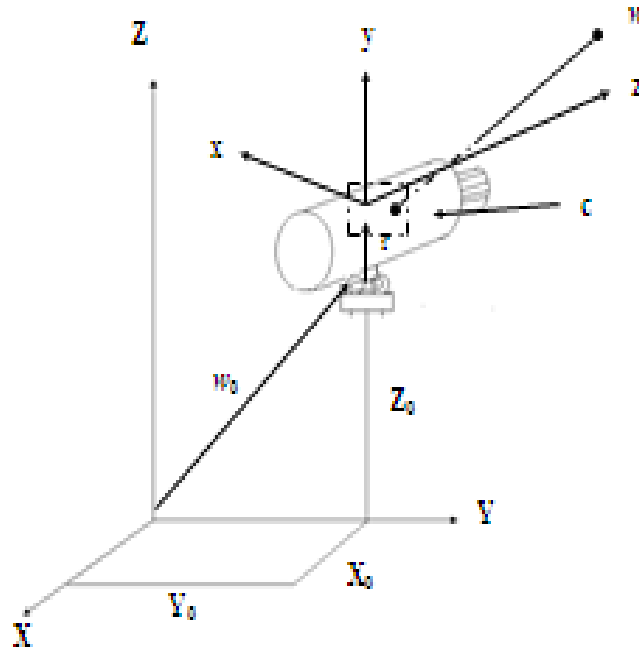


Figura 2.3 Geometría de la imagen con dos sistemas coordenadas

A continuación se muestra el ejemplo de un modelo de cámara montado sobre un robot utilizado en la tesis de Abundez [Abundez 2002], donde el origen del sistema coordenado del robot es coincidente con el origen del sistema coordenado de la cámara. Sin embargo, los ejes son tomados conforme lo descrito en la figura 2.4.



Figura 2.4. Geometría del sistema de navegación.

La figura 2.5 muestra como la medición de un punto del mundo real (3D) es representado en el sistema coordenado de la imagen. El punto espacial X_0 , se proyecta en el plano imagen pasando por el centro de la proyección X_c , originando el punto X_i , descrito por sus coordenadas en el plano imagen.

El punto localizado a la distancia f , frente al plano imagen es conocido como centro de proyección y la distancia f es conocida como distancia focal. El esquema que plantea el sistema coordenado de la cámara mostrado en la figura 2.5, es conocido como sensor de proyección posterior¹. El eje y , referido como eje óptico, indica la dirección del sensor al momento de capturar la escena.

¹ Proyección Posterior.- Se refiere a que el foco se presenta entre la lente y el objeto para una clara concepción de la geometría del sistema.

El campo de visión de la cámara es representado en una imagen de 360 x 240 píxeles, teniendo su origen al centro del plano imagen.

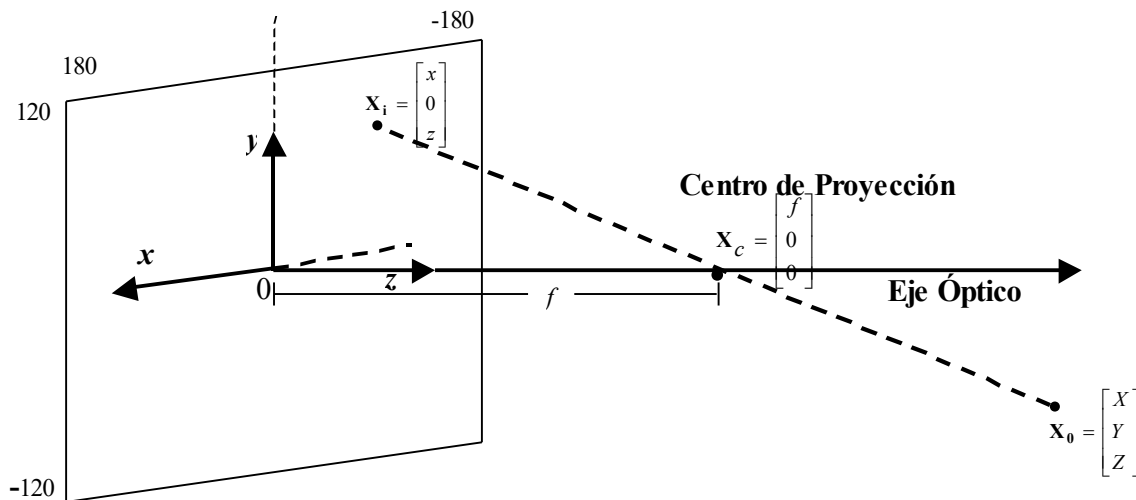


Figura 2.5. Sistema coordenado de la cámara

2.4 CALIBRACIÓN DE CÁMARA.

Existe un proceso ineludible al trabajar con VA para la determinación de distancias, el cuál es necesario efectuar antes de llevar a cabo la adquisición de la imagen. Este es la calibración de la cámara [Russ,1995],[Pratt,1991],[Trucco, 1999],[Schalkoff,1989], la cual permite conocer la equivalencia entre las coordenadas reales (X,Y,Z) de un punto en el espacio, y la representación del mismo punto en el plano imagen, donde sólo existen coordenadas x,z , como se muestra en el modelo de la cámara (figura 2.6). Puesto que la profundidad (y) se pierde en el proceso de digitalización, es necesario recuperarla posteriormente por medio de cálculos. Así, una forma de establecer esa equivalencia es realizar una medición física de varios puntos en el espacio obteniendo sus coordenadas (X,Y,Z) correspondientes, y luego tomar una imagen de esos mismos puntos, conociendo las coordenadas espaciales de la cámara. A este proceso se le llama

calibración. Si una cámara está calibrada, es posible establecer la relación entre las coordenadas 3D del mundo real de los objetos, vistos con sus correspondiente coordenadas 2D en el plano imagen y viceversa.

El proceso de calibración se describe en el diagrama de la figura 2.6. Para ejemplificar este proceso se mostrara un sistema donde la cámara está montada sobre un robot. El presente tema es extraído de la tesis Abundez [Abundez,2002].

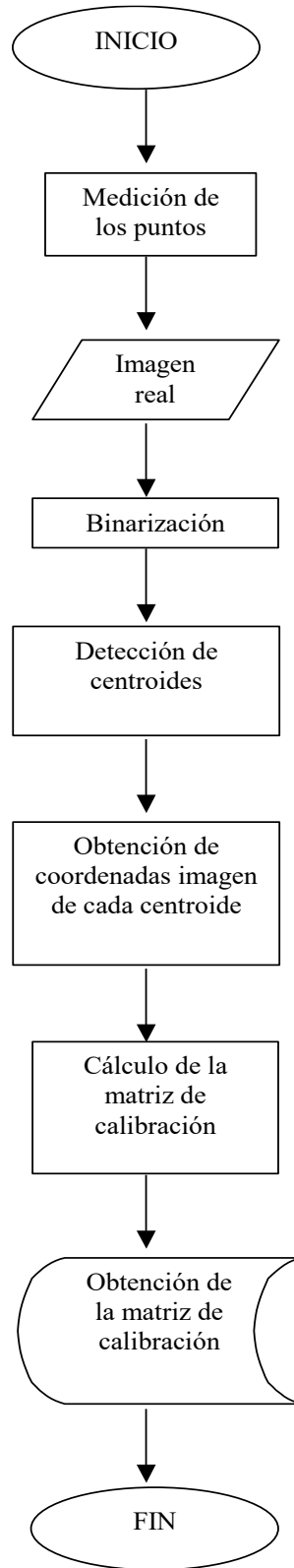


Figura 2.6. Calibración de Cámara.

Para llevar a cabo la calibración es necesario utilizar una rejilla de calibración [Pratt,1991],[Schalkoff,1999]. Usualmente, se utilizan puntos marcados en posiciones conocidas, sobre dos planos verticales unidos por una de sus aristas, en una configuración ortogonal. A continuación, se utiliza como rejilla de calibración un conjunto de figuras colocadas sobre una pared y parte del piso, esto con el fin de contar con una forma ortogonal que facilitará el procedimiento. Las figuras utilizadas fueron de dos tipos: Círculos, con un diámetro de 0.15 m, y elipses con ejes de 0.257 y 0.16 m a una distancia promedio de 0.50 m entre cada uno de ellos (figura 2.7). Se utilizaron elipses sobre el piso, ya que al utilizar círculos se apreciaban de menor tamaño por el ángulo de visión con que se observaban. La cámara se colocó a una altura sobre el nivel de piso de $Z = 0.60$ m, y a una distancia $Y = 3.60$ m con respecto al panel vertical de la rejilla.



Figura 2.7 Toma de la rejilla de calibración

Los objetos sobre el muro se colocaron sobre papel blanco con el fin de que existiera un buen contraste, al momento de la binarización.

Los índices visuales utilizados para determinar los valores experimentales para la calibración, son los centroides de cada objeto. Los centros de gravedad (x_c, y_c) están dados por las expresiones:

$$x_c = \frac{1}{N} \sum_{i=1}^N x_i \qquad y_c = \frac{1}{N} \sum_{i=1}^N y_i \qquad (\text{Ec. 2.11})$$

Donde N es el número de píxeles que componen al objeto, y (x_i, y_i) son las coordenadas de un píxel i .

La referencia universal utilizada para la rejilla de calibración, se muestra en la figura 2.8 donde el eje X es situado en la arista de la pared y el piso y el eje Z al centro de los objetos de la pared.

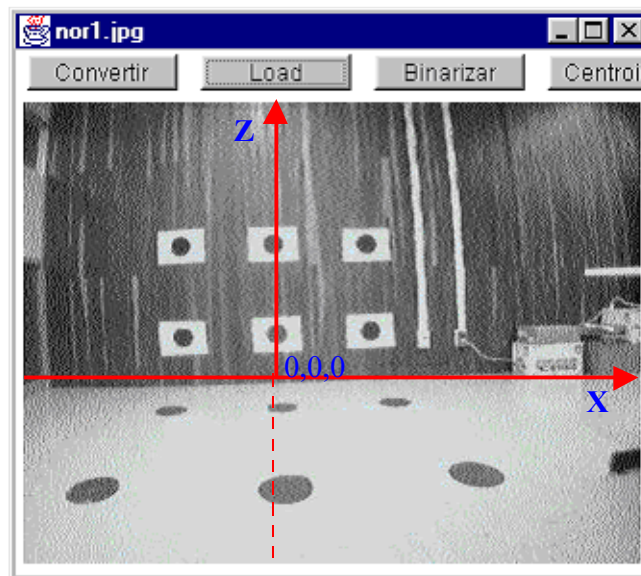


Figura 2.8 Referencia universal

Cabe mencionar que para la determinación de los centroides fue necesaria la detección de los objetos, lo cual se realizó a partir de la imagen previamente binarizada. El resultado del cálculo de los centros de gravedad, se muestra en la figura 2.9.

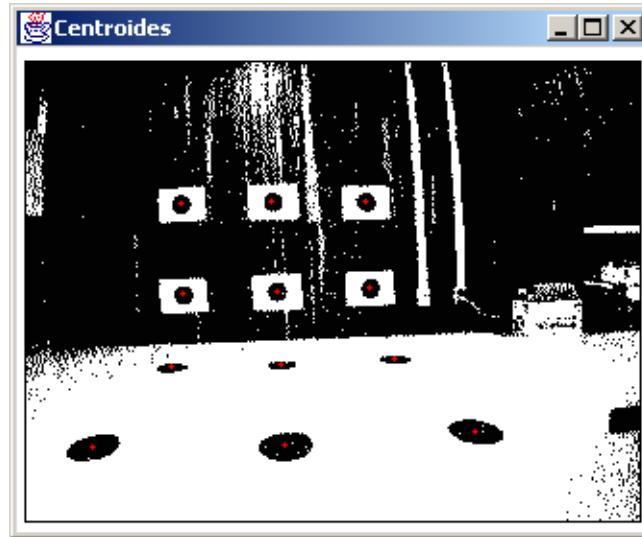


Figura 2.9. Detección de los centros de las marcas.

Las coordenadas (x_c, y_c) de cada objeto son almacenadas, con el fin de ser utilizadas en la búsqueda de la matriz de calibración. Para ello es necesario, conocer las mediciones en el mundo real con respecto a la cámara de cada par ordenado (x_c, y_c) . Las coordenadas del mundo real (X, Y, Z) , son mediciones realizadas de forma manual, considerando como referencia universal, la mostrada en la figura 2.9. En la tabla 2.1 se muestran las mediciones de cada objeto. Los puntos (Pa1 - Pa6) indican los índices de los objetos de la pared, y los puntos (Pi1-Pi6) representan los índices de los objetos del piso.

Tabla 2.1. Coordenadas resultantes de los objetos de referencia.

Puntos	Metros			Píxeles	
	X	Y	Z	xc	yc
Pa1	-0.650	3.600	0.453	100	31
Pa2	0.000	3.600	0.440	148	29
Pa3	0.650	3.600	0.425	197	29
Pa4	-0.650	3.600	-0.225	101	75
Pa5	0.000	3.600	-0.225	150	74
Pa6	0.650	3.600	-0.225	198	74
Pi1	0.530	2.771	-0.600	92	114
Pi2	0.000	2.771	-0.600	149	113
Pi3	0.530	2.771	-0.600	207	111
Pi4	-0.530	1.524	-0.600	44	168
Pi5	0.000	1.524	-0.600	151	169
Pi6	0.530	1.524	-0.600	258	163

El proceso para calcular las coordenadas reales, inicia almacenando en un vector las coordenadas del plano imagen ecuación (2.12). Este cálculo permitirá conocer sus correspondientes coordenadas físicas, con respecto a la referencia coordenada del robot.

$$\begin{bmatrix} w_i x_i \\ w_i y_i \\ w_i \end{bmatrix} = \begin{bmatrix} a_{11} a_{12} a_{13} a_{14} \\ a_{21} a_{22} a_{23} a_{24} \\ a_{31} a_{32} a_{33} a_{34} \end{bmatrix} \begin{bmatrix} x_o^w \\ y_o^w \\ z_o^w \\ 1 \end{bmatrix} \quad (\text{Ec. 2.12})$$

El vector del primer miembro contiene las coordenadas de los píxeles en el plano imagen (x_i, y_i) , de cada línea detectada y el valor de homogeneidad (w_i) . El segundo miembro de la ecuación contiene los parámetros de la Matriz de Calibración [Kvan,1999],[Pla,1997], la cual ha sido previamente calculada, el segundo vector corresponde a la coordenada del punto a desplazarse el robot (P3). El interés es encontrar estas coordenadas, para lo cual se resuelve un sistema de ecuaciones formulado de la ecuación (2.12), de donde mediante algunas manipulaciones algebraicas se llega a:

$$\begin{aligned} x_o(x_i a_{31} - a_{11}) + y_o(a_{32} x_i - a_{12}) &= a_{13} z_o + a_{14} - x_i a_{33} z_o - x_i a_{34} \\ x_o(y_i a_{31} - a_{21}) + y_o(a_{32} y_i - a_{22}) &= a_{23} z_o + a_{24} - y_i a_{33} z_o - y_i a_{34} \end{aligned} \quad (\text{Ec.2.13})$$

Donde se desea conocer x_o, y_o .

Los puntos de interés están localizados en el piso, por lo que z_o es considerado igual a -0.60 m, esto para poder tener una mayor visibilidad, ya que el robot cuenta con una altura de 0.22 m. Representando la ecuación (2.13) en forma matricial se tiene:

$$\begin{bmatrix} x_i a_{31} - a_{11} & a_{32} x_i - a_{12} \\ y_i a_{31} - a_{21} & a_{32} y_i - a_{22} \end{bmatrix} \begin{bmatrix} x_i \\ y_i \end{bmatrix} = \begin{bmatrix} a_{13} z_o + a_{14} - y_i a_{33} z_o + y_i a_{34} \\ a_{23} z_o + a_{24} - z_i a_{33} z_o + z_i a_{34} \end{bmatrix} \quad (\text{Ec. 2.14})$$

obteniendo una ecuación de la forma:

$$Ax = b \quad (\text{Ec. 2.15})$$

Posteriormente, premultiplicando toda la representación matricial del sistema de ecuaciones por A^{-1} :

$$[A^{-1}A]x = A^{-1}b$$

$$Ix = A^{-1}b$$

$$x = A^{-1}b \quad (\text{Ec.2.16})$$

Con la condición que A sea una matriz no singular, que en este caso se determina si $\det(A) \neq 0$

Los valores aquí determinados, corresponden a las coordenadas de los puntos extremos de las líneas de la carretera x_o^w, y_o^w . Estos valores son almacenados para ser utilizados en el proceso de generación de trayectoria.

Para la descripción del proceso de calibración, llamaremos x_o^w, y_o^w, z_o^w a las coordenadas del mundo real y a las coordenadas de los puntos, en el plano imagen x_i, y_i . Expresando en forma matricial, tenemos la ecuación 2.17 [Schalkoff,1989] que determina la matriz de calibración.

$$\begin{bmatrix} w_i x_i \\ w_i y_i \\ w_i \end{bmatrix} = \begin{bmatrix} a_{11} a_{12} a_{13} a_{14} \\ a_{21} a_{22} a_{23} a_{24} \\ a_{31} a_{32} a_{33} a_{34} \end{bmatrix} \begin{bmatrix} x_o^w \\ y_o^w \\ z_o^w \\ 1 \end{bmatrix} \quad (\text{Ec. 2.17})$$

Donde (w_i) , representa el valor de homogeneidad. La “matriz de calibración” (A) , representa las transformaciones (traslación, rotación y escalamiento), necesarias para igualar el sistema de coordenadas de la cámara con el origen de las coordenadas del mundo.

$$A = \begin{bmatrix} a_{11} a_{12} a_{13} a_{14} \\ a_{21} a_{22} a_{23} a_{24} \\ a_{31} a_{32} a_{33} a_{34} \end{bmatrix}$$

Aplicando las mediciones obtenidas (tabla 2.1) en la ecuación (2.17), obtenemos la matriz de calibración siguiente:

$$A = \begin{pmatrix} 208.4790, & 59.8034, & 94.1836, & 130.0002 \\ -5.0673, & 41.9076, & -97.5952, & 115.4334 \\ 0.0120, & 0.4597, & 0.8044, & 1.0000 \end{pmatrix} \quad (\text{Ec. 2.18})$$

En la Tabla 2.2 se muestran los cálculos obtenidos, tomando como entrada las coordenadas en píxeles de los puntos en el plano imagen (x_c, y_c) , y como

salida las coordenadas reales de cada uno de los puntos (X, Y, Z) , a partir de la matriz de calibración generada.

Tabla 2.2. Coordenadas del mundo real a partir de las coordenadas del plano imagen, mediante la matriz de calibración.

Puntos	Píxeles		Metros		
	x_c	y_c	X	Y	Z
Pa1	100	31	-0.655	3.600	0.453
Pa2	148	29	-0.001	3.600	0.440
Pa3	197	29	0.678	3.600	0.425
Pa4	101	75	-0.525	3.600	-0.225
Pa5	150	74	-0.001	3.600	-0.225
Pa6	198	74	0.514	3.600	-0.225
Pi1	92	114	0.531	2.771	-0.600
Pi2	149	113	-0.003	2.771	-0.600
Pi3	207	111	0.055	2.771	-0.600
Pi4	44	168	-0.530	1.524	-0.600
Pi5	151	169	-0.002	1.524	-0.600
Pi6	258	163	0.578	1.524	-0.600

Con el propósito de comprobar la precisión de la matriz de calibración obtenida, se aplicó un proceso inverso, dando como entrada, las coordenadas del mundo real y obteniendo las coordenadas del plano imagen. En la Tabla 2.3 se muestran los resultados obtenidos.

Tabla 2.3. Comprobación de la matriz de calibración teniendo como entrada, las coordenadas en el mundo real de los objetos de referencia.

Puntos	Metros			Píxeles	
	X	Y	Z	x_c	y_c
Pa1	-0.655	3.600	0.453	99	30
Pa2	0.000	3.600	0.440	147	28
Pa3	0.650	3.600	0.425	196	28
Pa4	-0.650	3.600	-0.225	100	75
Pa5	0.000	3.600	-0.225	148	74
Pa6	0.650	3.600	-0.225	196	74
Pi1	0.530	2.771	-0.600	92	114
Pi2	0.000	2.771	-0.600	149	113
Pi3	0.530	2.771	-0.600	207	110
Pi4	-0.530	1.524	-0.600	44	166
Pi5	0.000	1.524	-0.600	150	166
Pi6	0.530	1.524	-0.600	256	160

Haciendo un comparativo con los resultados reportados en las tablas 2.2 y 2.3, se valida la precisión de la calibración del sistema.

2.5 IMAGEN ESTEREOSCÓPICA

El interés de los métodos de tratamiento digital de imágenes deriva de dos áreas principales de aplicación: la mejora de la información pictórica para la interpretación humana y el procesamiento de los datos de la escena para la percepción autónoma por una máquina [González, 1996].

El término imagen como lo define González [González, 1996] se refiere a “una función bidimensional de intensidad de luz $f(x,y)$, donde x y y representan las coordenadas espaciales y el valor de f en un punto cualquiera (x,y) es proporcional al brillo (nivel de gris) de la imagen en ese punto”.

Ahora bien no solo existen imágenes como tales, para que el tratamiento digital de imágenes se lleve a cabo es necesario contar con imágenes digitales la cual “es una imagen $f(x,y)$ que se ha discretizado tanto en las coordenadas espaciales como en el brillo. Una imagen digital puede considerarse como una matriz cuyos índices de fila y columna identifican un punto de la imagen y el valor del correspondiente elemento de la matriz indica el nivel de gris en ese punto” [González, 2003].

Una imagen contiene un gran número de elementos que la componen llamados píxeles o pels, abreviatura inglesa de picture elements.

En la visión por computadora, la escena tridimensional es vista por una, dos o más cámaras para producir imágenes monocromáticas o en color del tipo 3D o imágenes estereoscópicas [Valero, 2002]. Las imágenes adquiridas pueden ser segmentadas para obtener de ellas características de interés tales como bordes o regiones. Posteriormente, de las características se obtienen las propiedades subyacentes mediante el correspondiente proceso de descripción. Tras lo cual se consigue la estructura de la imagen requerida por la aplicación de interés [Rowley, 1995].

De la idea anterior surgen las imágenes estereoscópicas las cuales proveen información acerca de la profundidad, dado que consta de la obtención de dos imágenes separadas del mismo objeto donde la distancia entre los centros de los lentes se llama línea base y el objetivo es hallar las coordenadas (X,Y,Z) del punto w que tiene las coordenadas (x_1,y_1) y (x_2,y_2) en las dos imágenes [González, 1996], esto puede ser representado en la figura 2.10.

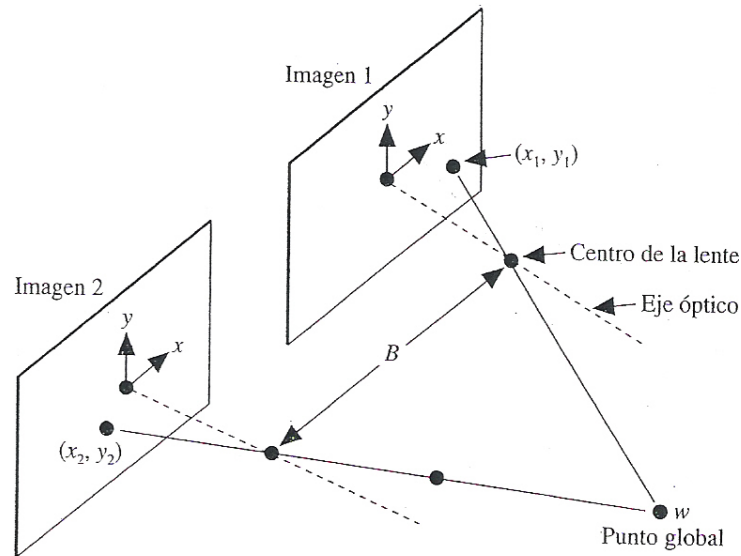


Figura 2.10 Modelo del proceso de imagen estereoscópica [De Fu, 1987]

Las imágenes de intensidad están íntimamente ligadas al concepto de luminosidad, dependiendo de la relación radiométrica que relaciona la cantidad de luz radiada por los puntos de las superficies en las imágenes y la cantidad de luz que emite [Rowley, 1995].

Una vez obtenida y definida la imagen esta será sometida al procesamiento digital de imágenes.

MEJORAMIENTO DE LA IMAGEN DIGITAL

3.1 INTRODUCCIÓN

El uso de las técnicas de mejoramiento de una imagen es el procesar las imágenes de manera que el resultado sea el más adecuado que la imagen de entrada para el uso de alguna aplicación específica. Para llevar a cabo este proceso es necesario el conocer la aplicación posterior esto con el fin de elegir el método adecuado para ser utilizado y garantizar el éxito en la etapa posterior.

El Filtrado es una técnica que permiten cambiar los niveles de gris en una imagen con objeto de realzar ciertas características geométricas o topológicas de los objetos de dichas imágenes. Elimina ruidos, realza discontinuidades entre regiones.

Operadores de bordes: Detectan y miden discontinuidades locales de intensidad luminosa o su gradiente.

Transformación de rango: Utiliza la geometría conocida acerca de las imágenes estereoscópicas para inferir la distancia que separa los puntos de interés en la escena de nuestro punto de vista.

Orientación de superficies: Cuando conocemos la fuente de iluminación así como las propiedades de reflectancia de las superficies observadas, es posible calcular la orientación de estas últimas. A este proceso se le conoce como obtención de la forma a partir del sombreado en la escena. Este proceso es particularmente sencillo cuando tenemos control sobre la iluminación. Con el podemos realizar las siguientes tareas como:

- Reducción de ruidos.
- Realzado de Características.

- Técnicas de Filtrado (procesamiento digital de imágenes):
 - ✓ Filtros Digitales (convolución con mascarar).
 - ✓ Filtros Estadísticos.
 - ✓ Filtros Morfológicos.

3.2 OPERADORES DE PUNTO

Los Procesos de punto a punto son las técnicas más simples para la mejora de la imagen.

En el procesamiento punto a punto, el valor de brillantez del píxel (x_p, y_p) en la imagen de salida depende tan solo de la brillantez del píxel de mismas coordenadas en la imagen de entrada con lo cual se realiza la “Ecualización” del histograma, razón por la que se le da el nombre de punto a punto.

La figura 3.2 se muestra gráficamente la convolución de una imagen teniendo una o varias imágenes de entrada y dando como resultado una imagen de salida.

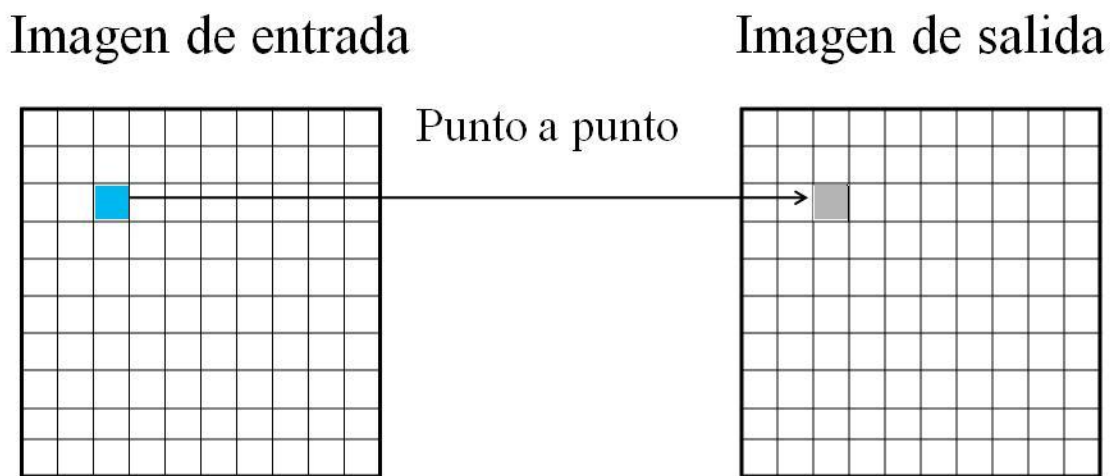


Figura 3.2 convolución de una imagen punto a punto

ECUALIZACIÓN DEL HISTOGRAMA

Un histograma es una función que proporciona la frecuencia de ocurrencia de cada nivel de gris.

Si el nivel de gris se codifica en q bits, se tienen 2^q niveles de cuantización: r_k está en $[0, 2^{q-1}]$.

PROCESAMIENTO DE HISTOGRAMA

El histograma de una imagen digital con niveles de gris en el rango $[0, L-1]$ es una función discreta $p(r_k) = n_k / n$, donde r_k es el k -ésimo nivel de gris, n_k es el número de píxeles de la imagen con este nivel de gris, n es el número total de píxeles de la imagen y $k = 0, 1, 2, \dots, L-1$.

Ejemplo: Supongamos que cada cuadro de imagen de la figura 3.2 a) cada cuadro representa un píxel, existen 4 niveles de color, el tamaño de la imagen es de 6x6 dando un total de 36 píxeles, b) muestra el diagrama del histograma, donde cada barra corresponde a niveles de color existentes y muestran la frecuencia existente de cada color ejemplo el color azul muestra que existen 17 píxeles de color azul y así con cada nivel de color, ver figura 3.2.

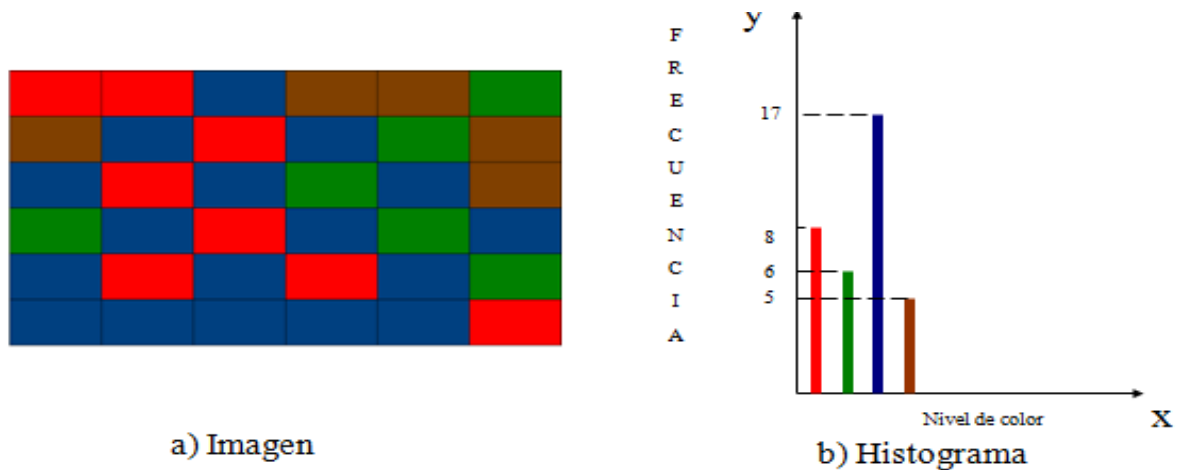


Figura 3.2 Ejemplo de Histograma

MEJORAMIENTO DE IMAGEN

En la figura 3.3 a) se muestra el histograma de una imagen oscura ya que los pixeles se encuentran concentrados a la izquierda del histograma donde se encuentran los tonos oscuros a diferencia del histograma de la figura 3.2 b) donde la concentración de colores se encuentra en los tonos blancos dando una imagen brillante. El histograma de la figura 3.3 c) muestra que la imagen a la que corresponde es una imagen de bajo contraste ya que los pixeles se encuentran concentrados en el centro de la imagen esto es no existen pixeles oscuros ni pixeles claros, para la figura 3.3 d) muestra el histograma que corresponde a una imagen con un alto contraste ya que los pixeles se encuentran distribuidos en todo el histograma.

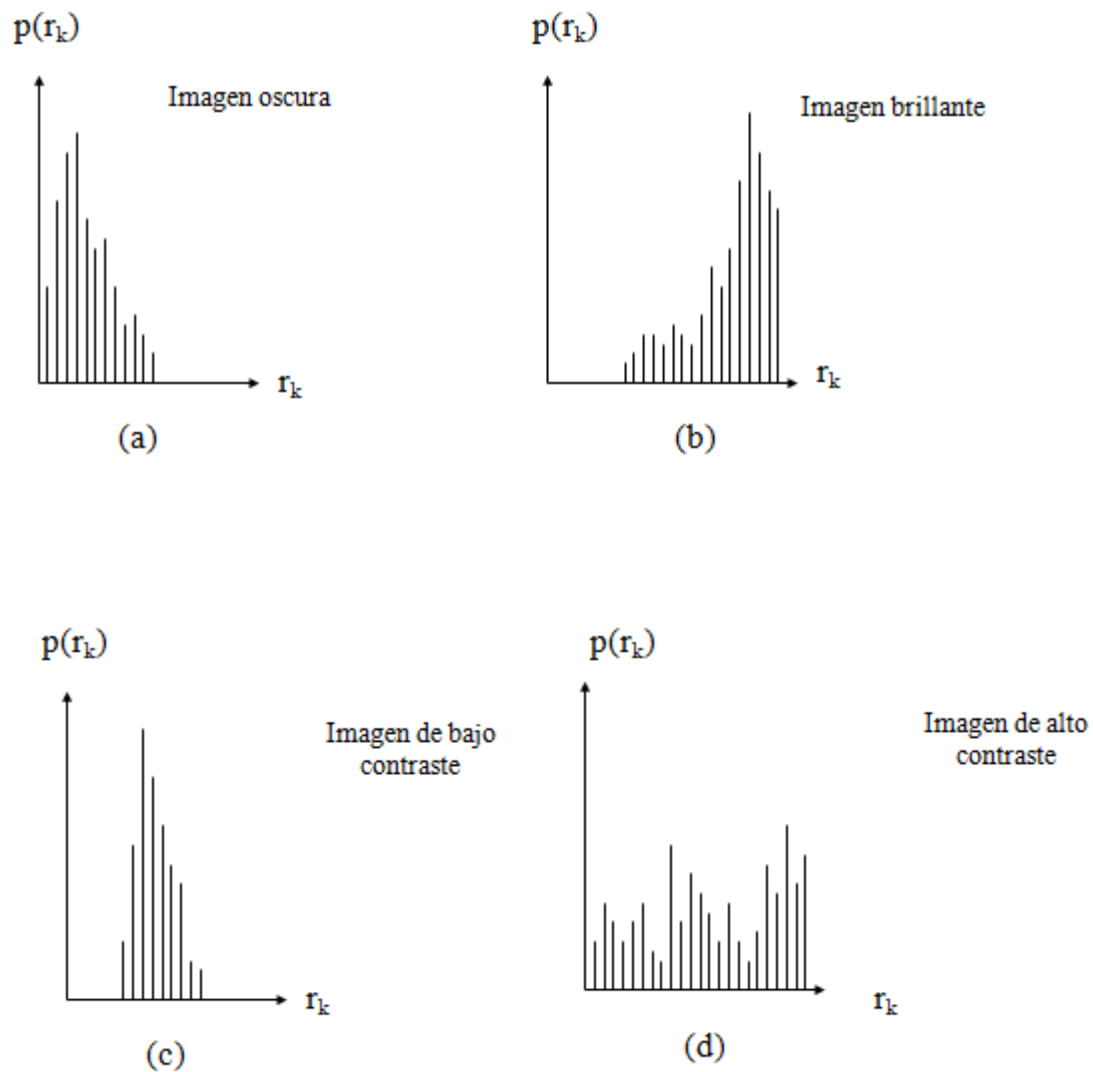


Figura 3.3 Histogramas correspondientes a cuatro tipos de imágenes básicas.

A continuación se muestra en la figura 3.4 una imagen con un bajo contraste (donde los píxeles se encuentran concentrados en el centro del histograma, esto es que no están bien distribuidos) y en la figura 3.5 se muestra el histograma de dicha imagen.



Figura 3.4 Imagen Original

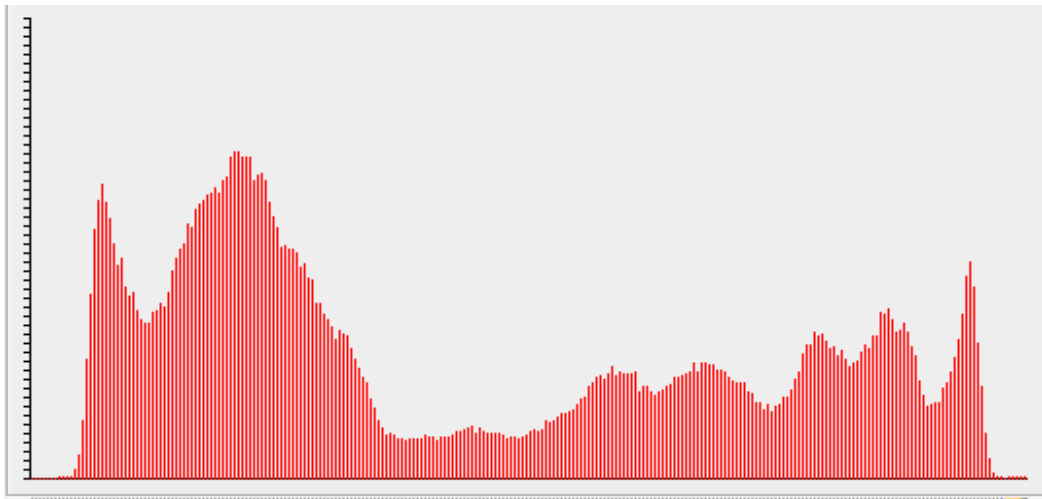


Figura 3.5 Histograma de la imagen 3.4

A continuación se muestra el resultado de ecualizar el histograma anterior dando como resultado el histograma ecualizado ver figura 3.6.

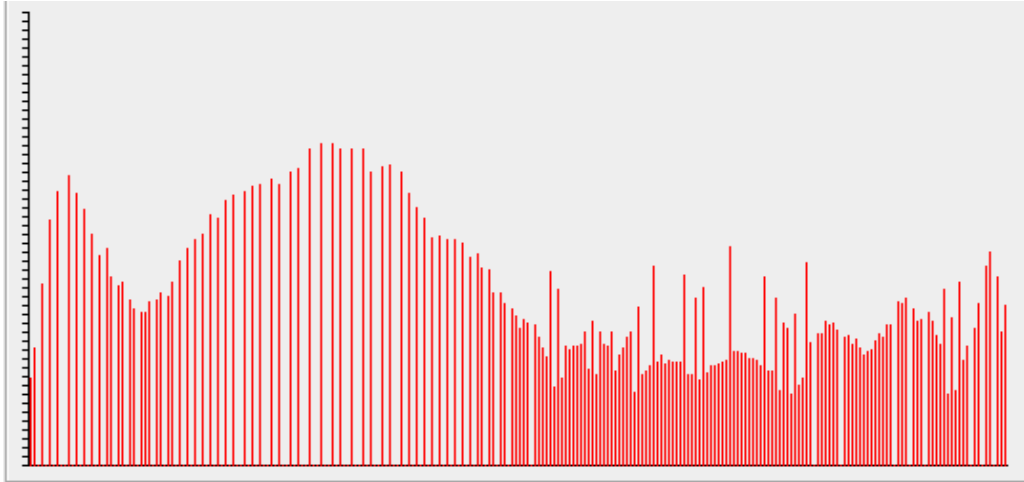


Figura 3.6 Histograma Ecualizado

Una vez ecualizado el histograma se procede a mostrar la imagen de acuerdo a los pixeles correspondientes al histograma ecualizado dando como resultado la imagen que se muestra en la figura 3.7 donde se puede observar esta imagen es clara con respecto a la original.



Figura 3.7 Imagen Ecualizada

Ecualización es una expansión del rango dinámico de la imagen aprovechando todos los niveles de gris.

AUMENTO DEL CONTRASTE DE LA IMAGEN.

La ecualización consiste en hacer igualmente probables todos los niveles de gris, es decir, obtener una *FDP* (*Función de Densidad de Probabilidad*) uniforme.

A continuación se muestra un ejemplo de la ecualización de un histograma tomado de [Pajares, 2002]:

Tomando el histograma de la figura 3.8 como entrada podemos observar que los píxeles se encuentran cargados hacia la derecha del histograma por lo tanto se puede deducir que el histograma corresponde a una imagen oscura.

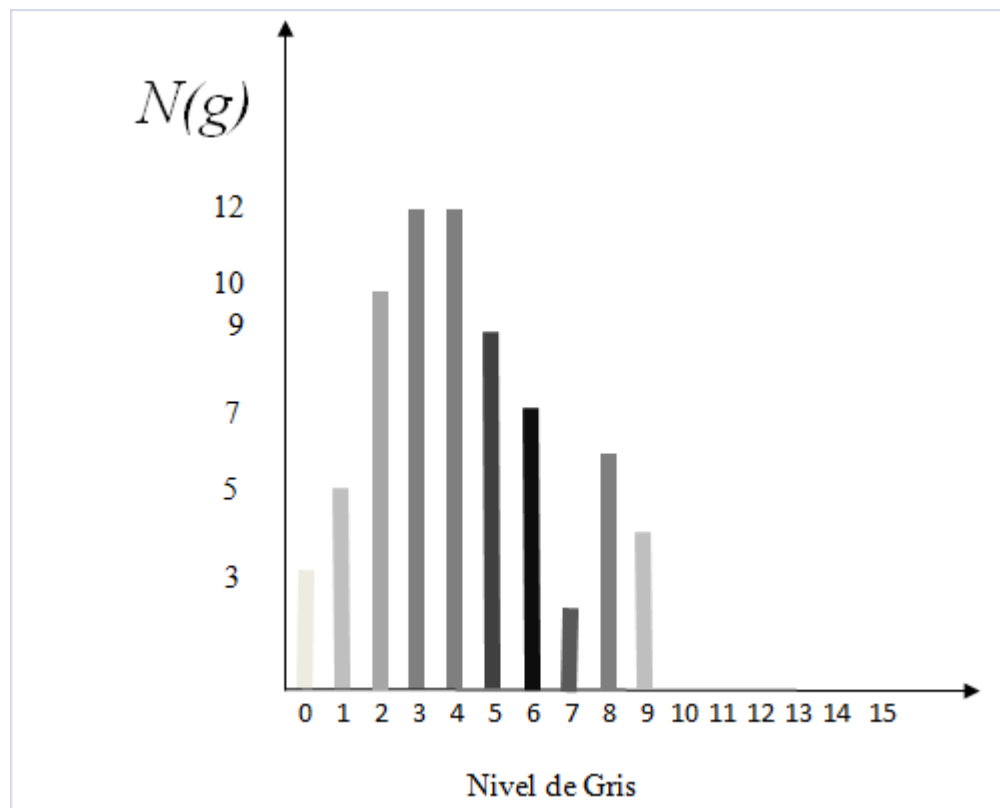


Figura 3.8 Imagen del histograma antes de la ecualización

Consideremos una imagen con 70 píxeles ($N \times M = 70$) y 16 niveles de gris ($N_g = 16$) con valores en el rango de $[0,15]$ g son los niveles de gris, f es la

frecuencia de aparición de dichos niveles de gris, F_a es la frecuencia acumulativa de los niveles de gris y $F(g)$ es el resultado de la transformación final

Utilizando la función analítica propuesta por Low (1991),

$$F(g) = \max \left\{ 0, \text{round} \left(\frac{N_g \times F_a}{N \times M} \right) - 1 \right\} \quad (\text{Ec. 3.1})$$

Tabla 3.1 Distribución de los niveles de Gris antes y despues de la palicación de la igualación

g	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
f	3	5	10	12	12	9	7	2	6	4	0	0	0	0	0	0
F_a	3	8	18	30	42	51	58	60	70	70	70	70	70	70	70	70
$F(g)$	0	1	3	6	9	11	12	13	14	15	15	15	15	15	15	15

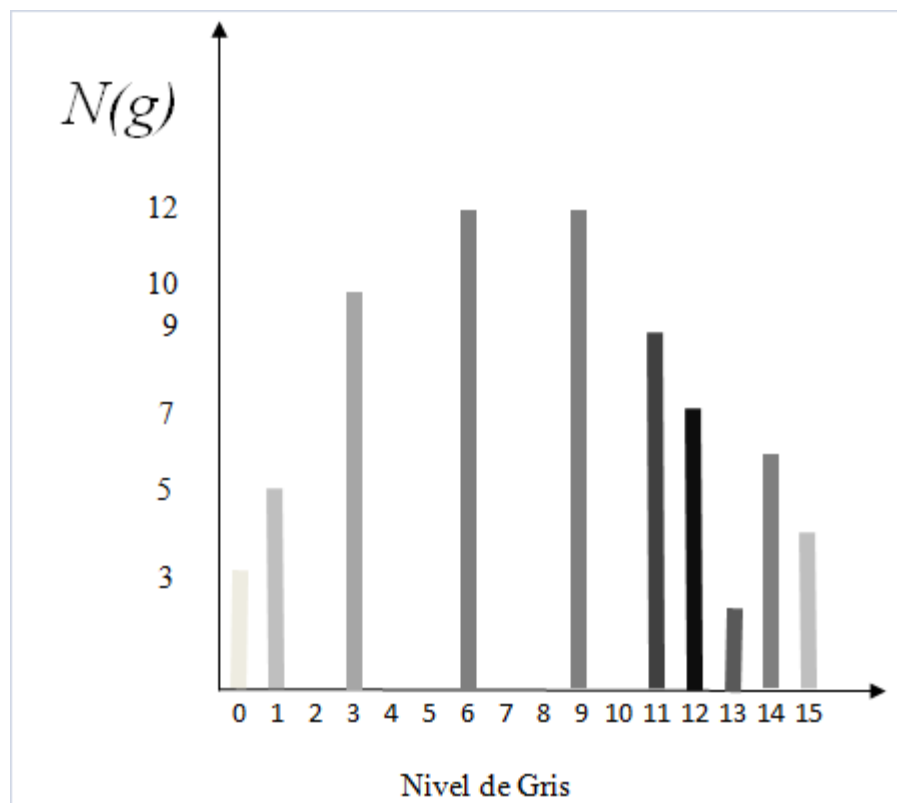


Figura 3.9 Imagen del histograma después de la ecualización

3.2 OPERADORES LOCALES EN EL RANGO DEL ESPACIO.

Los operadores en el rango *espacial* se refiere al conjunto de píxeles que componen una imagen y los métodos en el dominio espacial son procedimientos que operan directamente sobre los píxeles.

El empleo de máscaras espaciales para el procesamiento de las imágenes se denomina frecuentemente *filtrado espacial* (frente al *filtrado en el dominio de la frecuencia* empleando la transformada de Fourier), y las propias máscaras se les denomina *filtros espaciales*.

Los filtros denominados paso bajo atenúan o eliminan las componentes de alta frecuencia en el dominio de Fourier a la vez que dejan inalteradas las bajas frecuencias (es decir el filtro <<deja pasar>> las frecuencias bajas). Los componentes de alta frecuencia caracterizan los bordes y los restantes detalles muy marcados de la imagen, y por ello el efecto de introducir un filtrado paso bajo es el de hacer borrosa la imagen. De forma similar, los filtros de *paso alto* atenúan o eliminan las componentes de baja frecuencia. Siendo estas componentes las responsables de las características lentamente variables de la imagen, como el contraste global y la intensidad media, siendo el resultado de un filtrado paso alto es la reducción de bordes y de los restantes detalles finos. Un tercer tipo de filtrado, denominado *filtrado de paso banda*, elimina unas regiones elegidas de frecuencias intermedias, Estos filtros son empleados para la restauración de imágenes.

FILTRO ESPACIAL PASO ALTO BÁSICO

Un filtro de paso espacial de paso alto (realce) indica que el filtro debe de tener coeficientes positivos cerca de su centro y coeficientes negativos en la periferia. Para una máscara de 3 X 3 esta condición se cumple escogiendo un valor positivo

en el centro y tomando coeficientes negativos en el resto. Observe la figura 3.8 que la suma de los coeficientes es 0.

$$\frac{1}{9} * \begin{array}{|c|c|c|} \hline -1 & -1 & -1 \\ \hline -1 & 8 & -1 \\ \hline -1 & -1 & -1 \\ \hline \end{array}$$

Figura 3.8 Mascara de un filtro espacial de paso alto básico

A continuación la figura 3.9 y 3.10 muestran otros ejemplos de mascaras

z_1	z_2	z_3
z_4	z_5	z_6
z_7	z_8	z_9

(a)

1	0
0	-1

0	-1
-1	0

(b) Roberts

Figura 3.9 Mascaras de Roberts

-1	-1	-1		-1	0	1
0	0	0		-1	0	1
1	1	1		-1	0	1

(c) Prewitt

-1	-2	-1		-1	0	1
0	0	0		-2	0	2
1	2	1		-1	0	1

(d) Sobel

Figura 3.10 c) Mascara de Prewitt horizontal y vertical, d) Mascara de Sobel horizontal y vertical respectivamente.

Obsérvese que la suma de todos los coeficientes de una máscara es siempre igual a 0, indicando una respuesta de 0 en las áreas constantes, como sería de esperar tratándose de un operador diferencial.

3.3 MEJORAMIENTO EN EL RANGO DE FRECUENCIA.

La base de estas técnicas en el dominio de la frecuencia es el teorema de convolución. Sea $g(x, y)$ una imagen formada por la convolución de una imagen $f(x, y)$ y un operador lineal invariante de posición $h(x, y)$ ¹, es decir

Se trata de calcular la transformada de Fourier de la imagen a transferir, multiplicar el resultado por la función de transferencia de un filtro y, finalmente, tomar la transformada de Fourier inversa para llegar a una imagen mejorada.

Filtro paso bajo.- Dejar pasar las bajas frecuencias. Algunos filtros son:

- Filtro ideal
- Filtro Butterworth

Filtro paso alto.- Dejar pasar las altas frecuencias. Algunos filtros son:

- Filtro ideal pasa alto
- Filtro Butterworth pasa alto

MÉTODO DEL GRADIENTE

Para una función $f(x, y)$, el gradiente de f en el punto de coordenadas (x, y) se define a continuación ver ecuación 3.1.

$$\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} \quad (\text{Ec. 3.1})$$

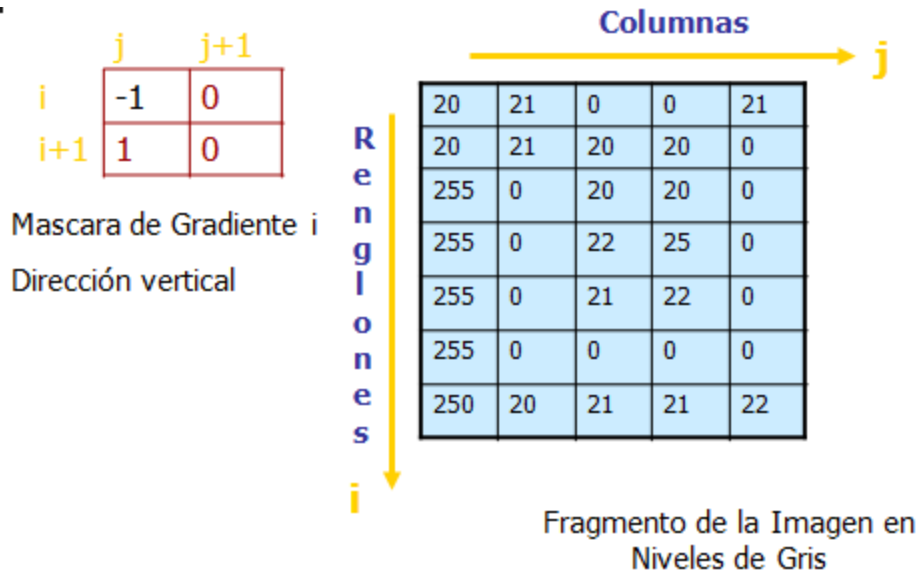


Figura 3.8 Gradiente

FILTROS ESTADÍSTICOS

Como filtro estadísticos podemos encontrar

- ✓ Promedio
- ✓ Media
- ✓ Mediana
- ✓ Gradiente estadístico

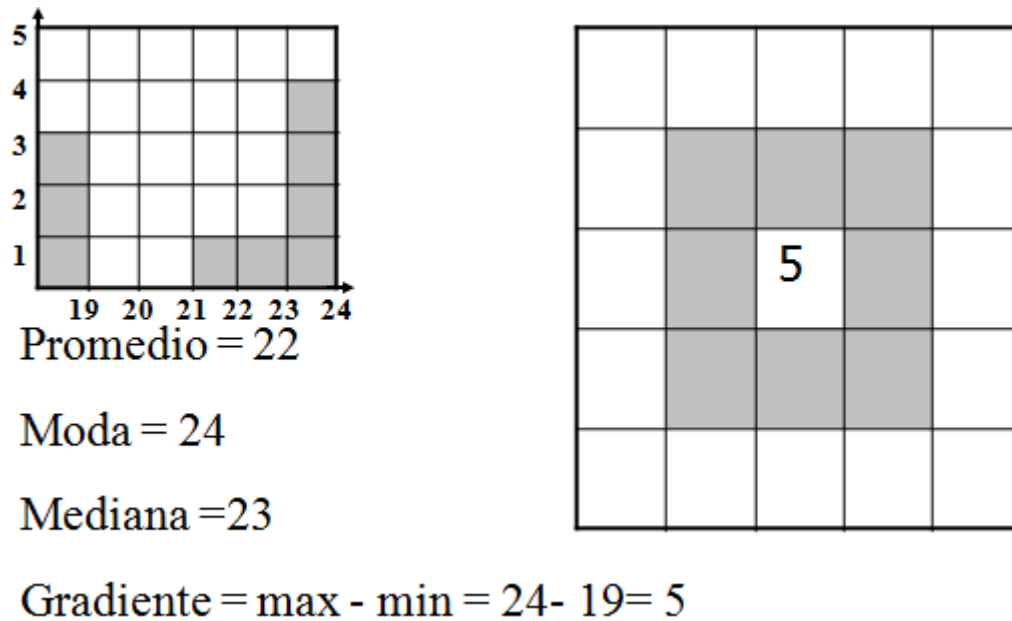


Figura 3.9 Filtros estadísticos

MEJORAMIENTO EN EL DOMINIO ESPACIAL

El empleo de máscaras espaciales para el procesamiento de las imágenes se denomina frecuentemente filtrado espacial, y las propias máscaras se denominan filtros espaciales.

FILTROS DE PASO BAJO

Atenúan o eliminan las componentes de altas frecuencias en dominio de Fourier a la vez que se dejan inalteradas las bajas frecuencias (es decir, el filtro <deja pasar> las frecuencias bajas.)

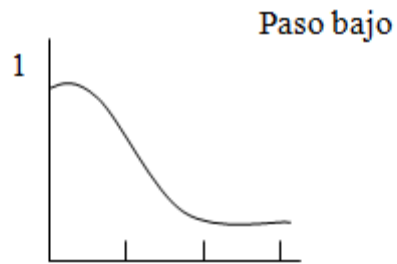


Figura 3.6 Filtro paso bajo

FILTROS DE PASO ALTO

Atenúan o eliminan los componentes de baja frecuencia. Como estas componentes son responsables de las características lentamente variables de la imagen, como el contraste global y la intensidad media, el resultado neto de un filtrado de paso alto es la reducción de estas características y, en correspondencia.

Una aparente agudización de estas características y, en correspondencia, una aparente agudización de los borde y de los restantes detalles finos.

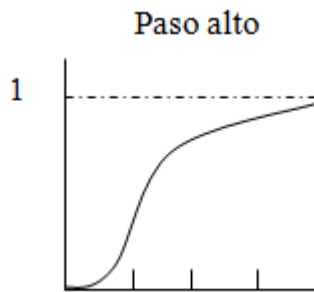


Figura 3.7 Filtro paso alto

FILTRO DE PASO BANDA

Elimina unas regiones elegidas de frecuencias intermedias. Estos filtros se emplean para la restauración de imágenes.

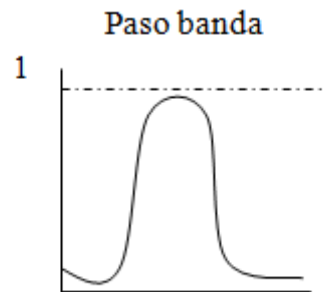


Figura 3.8 Filtro paso banda

EXTRACCIÓN DE CARACTERÍSTICAS DE LA IMAGEN DIGITAL

4.1 DETECCIÓN DE DISCONTINUIDADES.

Discontinuidad de una imagen digital: puntos, líneas y bordes. En la práctica, la forma más común de ver las discontinuidades es pasar una máscara a través de la imagen de la forma descrita.

Para una máscara 3x3 como la que se muestra en la figura 4.1, este procedimiento implica calcular la suma de los productos de los coeficientes por los niveles de gris contenidos en la región encerrada por las máscaras.

W_1	W_2	W_3
W_4	W_5	W_6
W_7	W_8	W_9

Figura 4.1 Usando una máscara general de 3x3

Considerando las mascarar de la figura 4.2. Si la primer mascara que se muestra es trasladada por toda la imagen, podría responder con más fuerza a líneas (un píxel continuo) orientados horizontalmente. Con un fondo constante, la respuesta máxima resultará cuando la línea pase por la fila central de la máscara. La segunda máscara responde mejor a líneas orientadas a 45°, la tercer máscara a a las líneas verticales y la cuarta en dirección de -45°.

-1	-1	-1
2	2	2
-1	-1	-1

-1	-1	2
-1	2	-1
2	-1	-1

-1	2	-1
-1	2	-1
-1	2	-1

2	-1	-1
-1	2	-1
-1	-1	2

Figura 4.2 Máscara de línea

La Figura 4.3(a) muestra que la primera derivada del perfil de nivel de gris es positiva en el borde de entrada de la transición, negativa en el borde de salida y cero en las zonas de nivel de gris constante. La segunda derivada es positiva en la parte de la transición asociada con el lado oscuro del borde, negativa en la parte de transición asociada con el lado claro y cero en las zonas de nivel de gris constante. Por lo tanto, el módulo de la primera derivada se puede utilizar para detectar la presencia de un borde en la imagen, y el signo de la segunda derivada se puede utilizar para determinar si un pixel del borde está situado en el lado oscuro o claro del mismo [González, 1996]. En la Figura 4.3 se muestra que la segunda derivada tiene un paso por cero en el punto medio de una transición de nivel de gris.

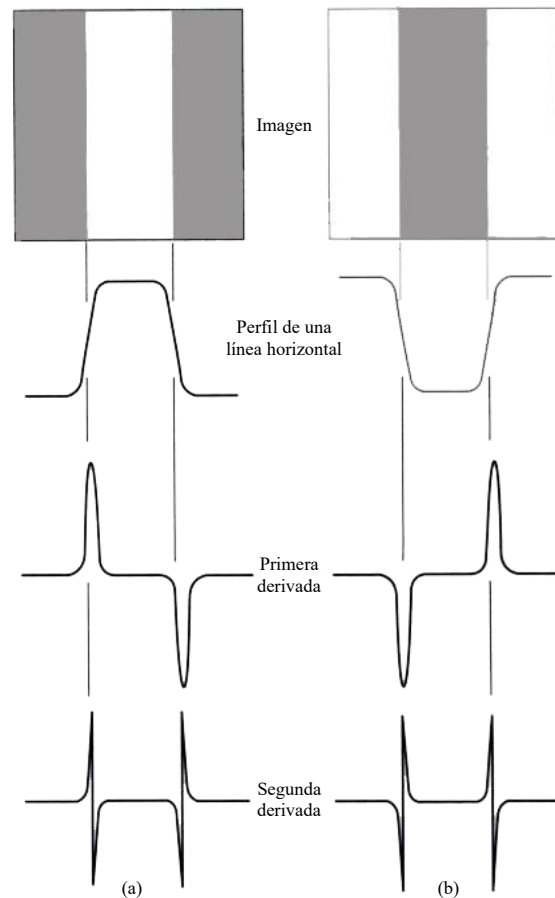


Figura 4.3 Detección de bordes por operadores de derivación: a) Banda clara sobre fondo oscuro; b) banda oscura sobre fondo claro. Obsérvese que segunda derivada tiene un paso por cero en la posición de cada borde.

4.2 DETECCIÓN DE CURVAS Y LÍNEAS Y TRANSFORMADA DE HOUGH.

En algunas ocasiones es necesario que una vez detectados los contornos se realice el enlazado de los mismos, esto con el fin de reunir píxeles que tengan alguna característica en común y que fueron detectados como segmentos discontinuos. Para ello existen diferentes técnicas, una de ellas es la transformada

de *Hough* la cual se explica a detalle a continuación debido a que fue utilizada en el desarrollo del presente trabajo.

Esta técnica se aplica para encontrar todos aquellos píxeles detectados por algún extractor de bordes, que pertenezca al contorno de un objeto cuya forma puede ser descrita por una curva paramétrica (recta o cónica). La transformada de Hough es intrínseca robusta a ruidos materializados como píxeles faltantes en el contorno o borde espurios pues se trata fundamentalmente de un método de interpolación (y linealización en el caso de rectas) [Pratt,1991].

Se considere un punto (x_i, y_i) y la ecuación general de una recta en forma explícita (punto-pendiente).

$$y_i = ax_i + b \quad (\text{Ec. 4.1})$$

A través de (x_i, y_i) pasan un número infinito de rectas, pero todas ellas satisfacen la ecuación (4.1) para diferentes valores de a y b . Sin embargo escribiendo la ecuación (4.2)

$$b = -ax_i + y_i \quad (\text{Ec. 4.2})$$

considerando el plano ab (también denominado espacio parámetro) se obtiene la ecuación de una única línea para un par determinado (x_i, y_i) . Además, un segundo punto (x_j, y_j) también tiene una línea en el espacio parámetro asociado con él, y esta línea corta a la línea asociada con (x_i, y_i) en (a, b) (a', b'), donde a' es la pendiente y b' la ordenada en el origen de la línea que contiene a (x_i, y_i) y (x_j, y_j) en el plano xy . De hecho todos los puntos contenidos en esta línea tienen líneas en el espacio parámetro que corta a (a', b') . La figura 4.4 ilustra los conceptos

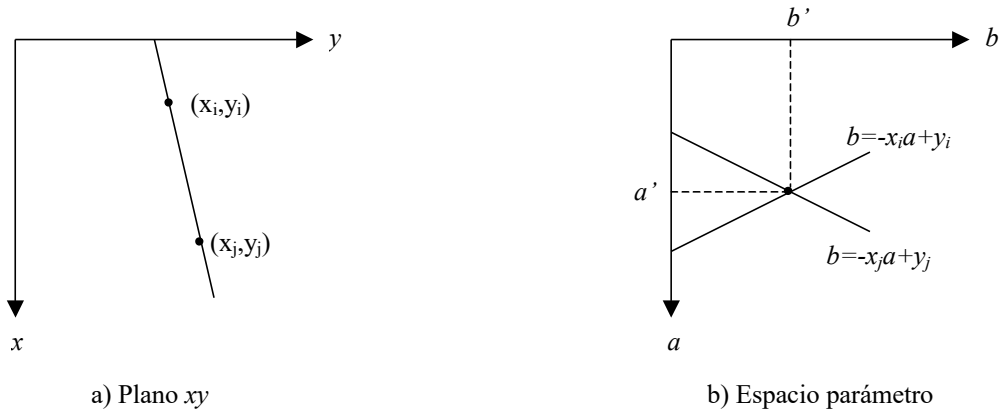


Figura 4.4 Representación de la Transformada de Hough

El cálculo de la *transformada de Hough* subdivide el espacio parámetro en las denominadas *células acumulador*, como se muestra en la figura 4.5, donde $a(a_{\min} - a_{\max})$ y sus columnas por el rango esperado para la ordenada en el origen $b(b_{\min} - b_{\max})$ son los valores de la pendiente y ordenada. La célula de coordenadas (i, j) , con valor acumulador $A(i, j)$, corresponde al cuadrado asociado con las coordenadas del espacio parámetro. Inicialmente estas células están inicializadas en cero. Después de cada punto (x_k, y_k) del plano imagen, se fija el parámetro a igual a cada uno de los valores permitidos de subdivisión sobre el eje x y se resuelve para el b correspondiente utilizando la ecuación 4.2. Las b resultantes se redondean después al valor más próximo permitido del eje x . Si una elección de a_p resulta ser la solución b_q , se fija $A(p, q) = A(p, q) + 1$. Al final de este procedimiento, un valor de M en $a(i, j)$ corresponde a M puntos del plano (xy) situados en la línea $y = a_i x + b_j$.

Al subdividir el eje x en K incrementos dados, para cada punto (x_k, y_k) K valores de b corresponderán a los K posibles valores de a . Con n puntos imagen, este método implica nK cálculos. Por ello el procedimiento es *lineal* en

n , y el producto nK no se acerca al número de cálculos expuestos al principio de esta sección a menos que j se aproxime o supere a n .

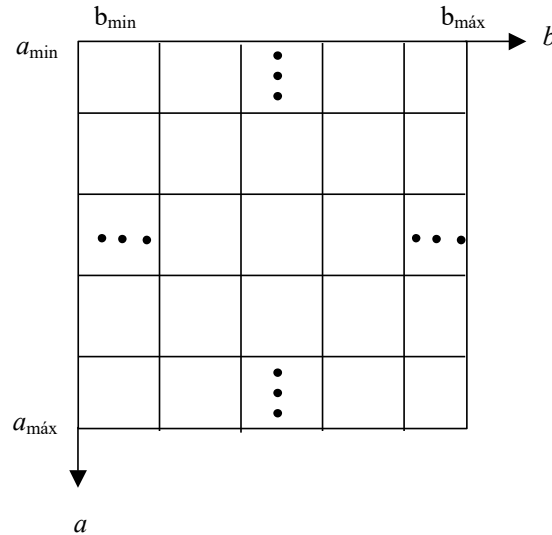


Figura 4.5. Representación de la matriz de acumuladores.

Un problema identificado, es que, cuando se utiliza $y = ax + b$ para representar una línea es que tanto la pendiente como la ordenada en el origen tienden a infinito cuando la línea se acerca a la vertical. Una forma de evitar esta dificultad es utilizar la representación normal de una recta:

$$x \cos \theta + y \sin \theta = \rho \quad (\text{Ec. 4.3})$$

La figura 4.3 muestra el significado de los parámetros utilizados en la ecuación (4.3). El emplear esta representación es la construcción de una tabla de acumuladores idéntico al expuesto para la representación punto-pendiente. Sin embargo en lugar de líneas rectas, los lugares geométricos son curvas sinusoidales en el plano $\theta\rho$. Como antes, M puntos colineales sobre una línea $x \cos \theta_j + y \sin \theta_j = \rho_i$ producen M curvas sinusoidales que cortan $(\theta_i \rho_j)$ en el espacio parámetro. Al incrementar θ y resolver para él ρ correspondiente da M

entradas en el acumulador $A(i, j)$ asociado con la calcula determinada por $(\rho_i \theta_j)$ ver figura 4.4. Ilustra la división del espacio parámetro.

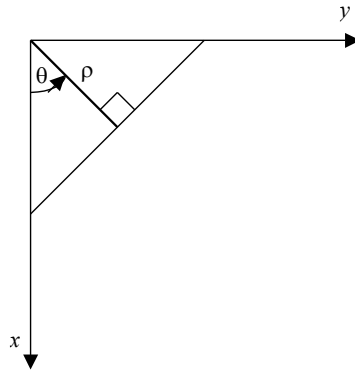


Figura 4.3.- Representación de la recta en el plano xy

El rango del ángulo θ es $\pm 90^\circ$, medidos con respecto al eje x . Por ello, con referencia a la figura 4.3 una línea horizontal tiene $\theta = 0^\circ$, siendo igual a la x positiva. De forma similar una línea vertical tiene $\theta = 90^\circ$, siendo ρ igual a la y positiva, o $\theta = -90^\circ$, siendo ρ igual a la y negativa

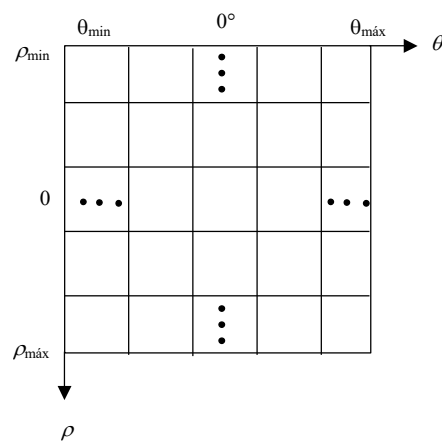


Figura 4.4 Cuantificación del plano $\theta\rho$ en la celdas

Ejemplo 1: Representación sinodal de cuatro puntos con los siguientes valores:

$x_1=5$ $y_1=15$ $x_2=10$ $y_2=10$ $x_3=15$ $y_3=15$ $x_4=20$ $y_4=0$

Como se puede observar el punto 1, 2, 4 forman una línea, ya que ambas curvas sinodales se intersecan (figura 4.5).

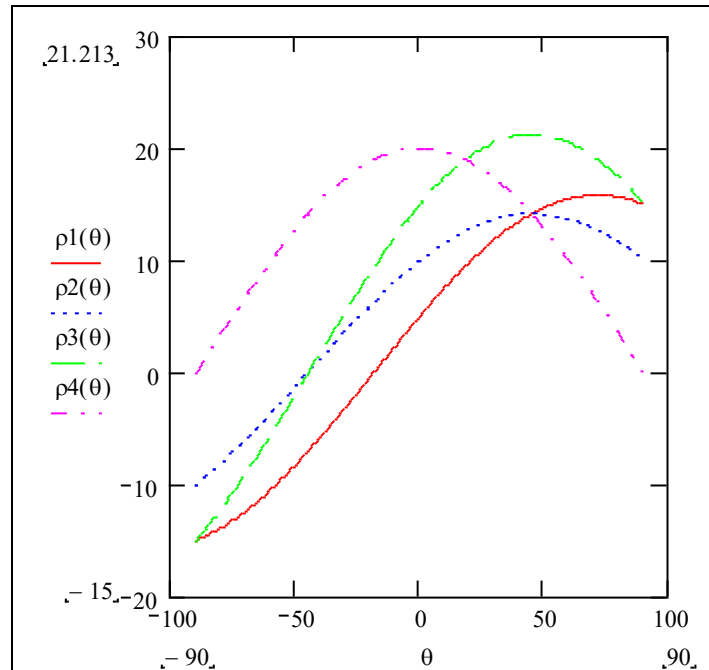


Figura 4.5 Representación sinoidal de 4 puntos

El rango del ángulo siempre varía entre $\pm 90^\circ$ con respecto al eje de las X , de modo que una línea horizontal tiene $\theta = 0^\circ$, siendo ρ igual a la x positiva. Una línea de $+90^\circ$ hará que ρ sea igual a la y positiva y $\theta = -90^\circ$ grados hará que ρ sea igual a y negativa (sustitúyanse los valores anteriores en Ec. 4.3). Es decir, el nuevo espacio-paramétrico se muestra en la figura 4.6.

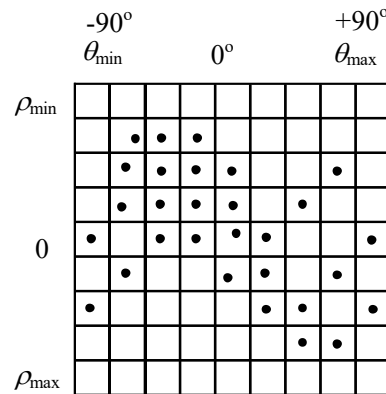


Figura 4.6 Nuevo espacio-parámetro

El rango ésta dado por:

$$R\rho = \pm\sqrt{A^2 + B^2} \quad (\text{Ec. 4.2.4})$$

donde A y B son el ancho y alto de la imagen respectivamente.

La *transformada de Hough* es muy útil para la determinación de líneas rectas, sin embargo se complica más para la determinación de las curvas. Este método podría emplearse para eliminar los puntos aislados que no pertenecen a ninguna figura, de modo que funcione como mecanismo de limpieza de imagen.

Algoritmo de la *Transformada de Hough*:

- Calcular el gradiente de una imagen
- Subdividir específicamente el plano $\rho\theta$.
- Examinar el contenido de las células del acumulador con altas concentraciones de píxel.
- Examinar la relación (principalmente de continuidad) entre los píxeles de la celda elegida. En este caso el concepto de continuidad se basa normalmente en el cálculo de la distancia entre los píxeles desconectados, que se han identificado durante el recorrido del conjunto de píxeles desconectados que se han identificado durante el recorrido del conjunto de píxeles correspondientes

a una celda acumulador dada. Un vacío en cualquier punto es significativo si la distancia entre ese punto y el vecino más próximo exceda de un determinado umbral.

4.5 FLUJO ÓPTICO.

El flujo óptico puede surgir del movimiento relativo de los objetos y el observador (aunque no siempre es así).

Puede dar importante información sobre la **localización espacial** de los objetos vistos y la **velocidad de cambio** de esta localización.

Sin embargo, la **relación** entre el **flujo óptico** en el plano imagen y la **velocidad de los objetos** en el mundo tridimensional **no** es necesariamente **obvia**.

Un objeto en movimiento puede producir un patrón de brillo constante en la imagen. Por ejemplo:

- esfera rotando iluminada por una fuente puntual
- estática. No hay cambio en los niveles de
- intensidad: flujo óptico nulo.

Un objeto detenido puede dar un patrón de brillo variable en la imagen. Por ejemplo:

- esfera estática e iluminada por una fuente puntual en movimiento. Hay cambio en los niveles de intensidad de la esfera: flujo óptico no nulo.

La forma simple de estimar el movimiento en la escena sin tener que relacionar cada pixel de la imagen con un punto 3D es:

- Medir los desplazamientos aparentes del patrón de intensidad de la imagen.
- Este movimiento aparente se llama flujo óptico ver figura 4.7.

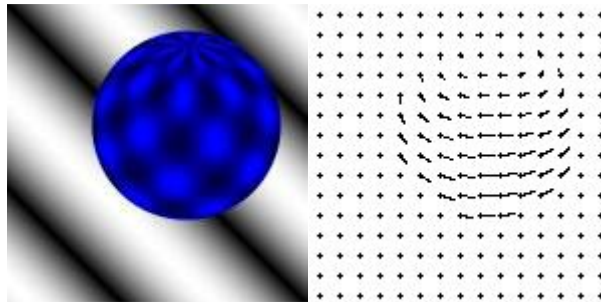


Figura 4.7 Flujo Óptico de una esfera

SEGMENTACIÓN DE IMÁGENES

En la etapa de segmentación se subdivide una imagen en sus partes constituyentes u objetos. El nivel al que se lleva a cabo esta subdivisión depende del problema a resolver. Esto es, la segmentación deberá detenerse cuando los objetos de interés de una aplicación hayan sido aislados.

En general, la segmentación autónoma es una de las tareas más difíciles del procesamiento de imágenes. Esta etapa del proceso determina el eventual éxito o fracaso del análisis. De hecho, la segmentación rara vez llega a alcanzar una solución satisfactoria. Por esa razón, se debe poner un considerable cuidado en aumentar la probabilidad de tener una segmentación robusta.

Algunas técnicas para realizar este proceso son la detección de puntos, detección de líneas, detección de bordes (Figura 3 (b)), detección combinada, umbralización, segmentación por medio del color o los niveles de gris (Figura 3(c)), segmentación por textura, entre otras [González, 1996].

Ejemplo:

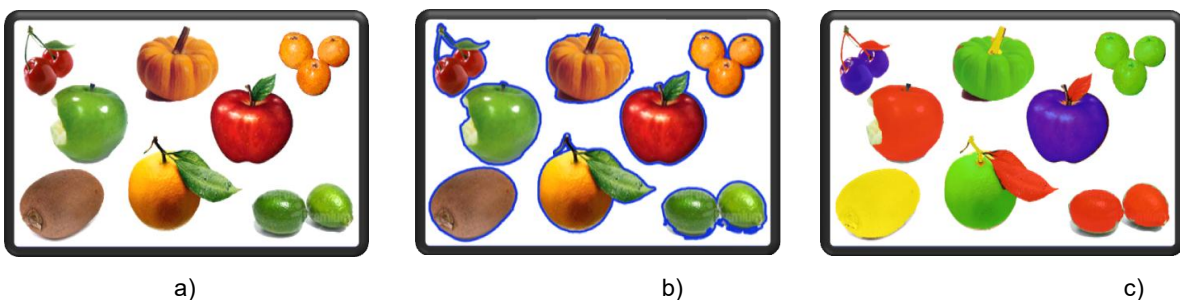


Figura 5.1. Técnicas para la segmentación, (a) Imagen original; (b) Detección de bordes; (c) Segmentación por medio del color

5.1 DETECCIÓN DE DISCONTINUIDADES

La forma más común de ver las discontinuidades para los tres tipos básicos de discontinuidades de una imagen digital (puntos, líneas y bordes), es pasar una máscara a través de la imagen. Utilizando por ejemplo una máscara 3 x 3, como se muestra en la figura 5.2, el procedimiento implica calcular la suma de los productos de los coeficientes por los niveles de gris contenidos en la región encerrada por la máscara.

W_1	W_2	W_3
W_4	W_5	W_6
W_7	W_8	W_9

Figura 5.2 Una máscara general de 3 x 3

Esto es, la respuesta de la máscara en un punto cualquiera de las imágenes:

$$R = w_1z_1 + w_2z_2 + \dots + w_9z_9 = \sum_{i=1}^9 w_i z_i$$

Donde z_i es el nivel de gris asociado con el coeficiente de la máscara w_i [González, 1996].

Extracción de bordes.- La extracción de bordes es el método más común para detectar discontinuidades significativas en el nivel de gris.

Un borde es la frontera entre dos regiones con propiedades de nivel de gris relativamente distintas, ver figura 5.3.

Se supone que las regiones en cuestión son lo suficientemente homogéneas para que la transición entre dos de ellas se pueda determinar sobre la base de las discontinuidades del nivel de gris solamente [González, 1996].

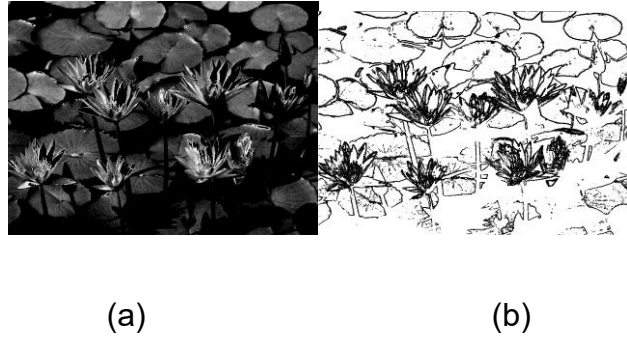


Figura 5.3 (a) Imagen original; (b) Imagen con bordes detectados

La conectividad.- Conceptualmente, las fronteras de los objetos se pueden encontrar uniendo los puntos de borde en una imagen, que han sido encontrados. Un pixel puede ser considerado “conectado a cuatro” o “conectado a ocho” (figura 5.4) cuando tiene las mismas propiedades que alguno de los cuatro u ocho vecinos más próximos.

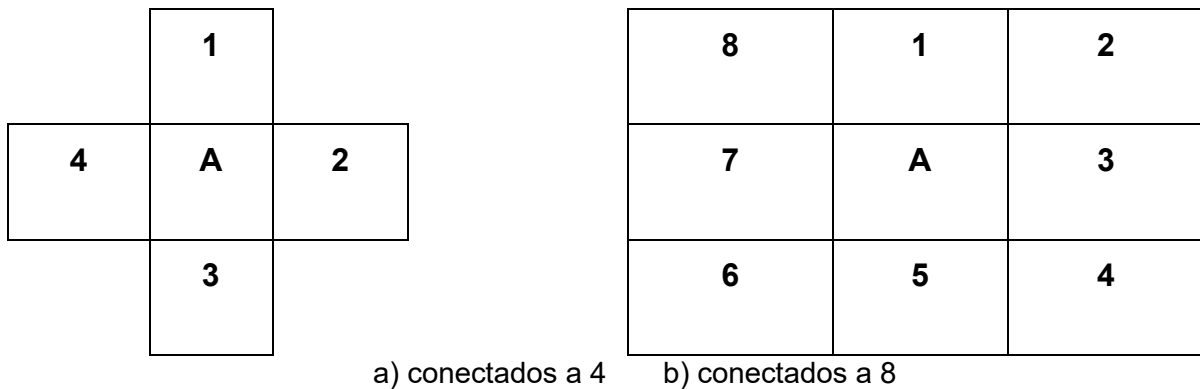


Figura 5.4 Conexión a 4 y conexión a 8

Llegando a una paradoja: Si suponemos “conexión a 4”, vemos que los segmentos 1, 2, 3 y 4 serían clasificados como disjuntos. Porque hay un hueco entre los pixeles vecinos. Esto no sucede si aplicamos “conexión a 8”. Una solución a esto

	j	j+1
i	-1	0
i+1	1	0

	j	j+1
i	-1	0
i+1	1	0

	j	j+1
i	-	
i+1		

suele ser considerar “conexión a 8” para el objeto (borde) y “conexión a 4” para el fondo.

5.2 ENLAZADO DE BORDES Y DETECCIÓN DE LÍMITES.

Una imagen digitalizada puede representarse matemáticamente por una función f cuyo valor en el punto de coordenadas (x, y) , representa la iluminación existente en ese punto, por lo tanto un valor del píxel va en una escala del 0 al 255. La imagen digital muestra un cambio de iluminación de un punto que está representado gráficamente por la pendiente que representa a $f(x, y)$, y normalmente resulta diferente según sea la dirección considerada, la dirección en la que el cambio es más intenso está dada por el gradiente. Un cambio brusco en la intensidad de un píxel a otro nos da la localización de un contorno [Arellano, 2005].

5.2.1. MÉTODO BASADO EN LAS PRIMERAS DERIVADAS.

En este método primero se miden los cambios de intensidad en todos los puntos de la imagen, haciendo uso de la primera derivada, después se seleccionan como puntos de bordes aquellos puntos en los que el cambio de intensidad rebasa algún umbral preestablecido.

Una imagen puede representarse matemáticamente por una función f cuyo valor en el punto (x, y) representa la iluminación existente en ese punto, el cambio de iluminación de un punto en la imagen está representado gráficamente por la pendiente que allí tiene la superficie que representa a f , y normalmente resulta diferente según sea la dirección considerada, la dirección en la que el cambio es más intenso está dada por el vector gradiente en ese punto [Arellano, 2005].

$$\begin{aligned}\frac{\partial f(x,y)}{\partial x} &= \Delta_x = \frac{f(x+d_x,y) - f(x,y)}{dx} \\ \frac{\partial f(x,y)}{\partial y} &= \Delta_y = \frac{f(x,d_y+y) - f(x,y)}{dy}\end{aligned}\quad (\text{Ec.5.1.})$$

El gradiente de una imagen $f(x,y)$, se define como un vector bidimensional dado por la ecuación 5.1, siendo un vector perpendicular al borde.

5.2.2. MÉTODO DEL GRADIENTE

Como resumen podemos indicar que estos métodos se fijan en los cambios bruscos del nivel de gris alrededor de un pixel de la imagen, estos cambios pueden detectarse con la primera derivada o gradiente. Para empezar podemos definir un borde como la zona que separa dos regiones de distinta iluminación, por lo tanto, podemos aplicar a un pixel y su alrededor una matriz (máscara) de forma que se evalúen las diferencias y poder asegurar si hay o no un borde. Mientras más diferentes sean los pixeles opuestos, con más seguridad existe un borde en el central.

$$\begin{array}{cc} & J & j+1 \\ i & \begin{array}{|c|c|} \hline -1 & 0 \\ \hline \end{array} \\ i+1 & \begin{array}{|c|c|} \hline 1 & 0 \\ \hline \end{array} \end{array}$$

$$v_i = f(i+1, j) - f(i, j) \quad (\text{Ec. 5.2})$$

$$\begin{array}{cc} & J & j+1 \\ i & \begin{array}{|c|c|} \hline -1 & 1 \\ \hline \end{array} \\ i+1 & \begin{array}{|c|c|} \hline 0 & 0 \\ \hline \end{array} \end{array}$$

$$v_i = f(i, j+1) - f(i, j) \quad (\text{Ec. 5.3})$$

5.2.3. MÉTODO ROBERTS

Es equivalente al anterior, si se rotan los ejes 45 grados, las mascarar serán entonces:

	V_i			V_j	
	j		j		
i	-1	0	i	0	1
	0	1		-1	0

Figura 5.5 Mascara de Roberts

Una gran desventaja del operador de Roberts es su sensibilidad al ruido, ya que se usan pocos pixeles para aproximar al gradiente.

5.2.4. MÉTODO PREWITT

Para aproximar la derivada sobre el pixel (x, y) , en la dirección horizontal, se toma en consideración no sólo el cambio de intensidad hacia la derecha de ese punto, sino que también el que aparece hacia la izquierda, además la promedia con los valores que obtiene en la misma posición de las filas contiguas encima y debajo; de este modo se llega a una estimación de la derivada horizontal que es muy resistente al ruido presente en las imágenes. Para estimar la derivada en dirección vertical se usa un procedimiento análogo [Arellano, 2005]. Las mascarar que se utilizan son:

V_i

	-1	-1	-1	
i	0	0	0	i

V_j

	1	0	1	
i	-1	0	1	

Figura 5.5 Mascara de Prewitt

5.2.5. MÉTODO SOBEL

Sigue la misma lógica que el anterior, sólo que da más peso a la variación de intensidad que aparece en la línea que contiene al píxel sobre el que se está estimando la derivada. Eso se implementa duplicando los coeficientes de esa línea. Las mascarar utilizadas son:

V_i		V_j																		
j		j																		
<table border="1"><tr><td>-1</td><td>-2</td><td>-1</td></tr><tr><td>0</td><td>0</td><td>0</td></tr><tr><td>1</td><td>2</td><td>1</td></tr></table>	-1	-2	-1	0	0	0	1	2	1		<table border="1"><tr><td>-1</td><td>0</td><td>1</td></tr><tr><td>-2</td><td>0</td><td>2</td></tr><tr><td>-1</td><td>0</td><td>1</td></tr></table>	-1	0	1	-2	0	2	-1	0	1
-1	-2	-1																		
0	0	0																		
1	2	1																		
-1	0	1																		
-2	0	2																		
-1	0	1																		

Figura 5.6 Mascara de Sobel

5.2.6. MÉTODO BASADO EN LA SEGUNDA DERIVADA

Consiste en determinar directamente en que punto de la imagen aparecen los picos de los cambios de iluminación utilizando las segundas derivadas. En los puntos en que los cambios de intensidad tienen un máximo. Para esto se puede cambiar el cálculo de las derivadas por el Laplaciano.

$$V^2 f(x, y) = \frac{\partial^2}{\partial x^2} f(x, y) + \frac{\partial^2}{\partial y^2} f(x, y) \quad (\text{Ec. 5.4})$$

Este operador tiene la ventaja de que sólo exige calcular dos segundas derivadas, para calcularlo en imágenes digitales se puede aproximar utilizando la máscara siguiente:

j

0	1	0
1	-4	1
0	1	0

Ventajas del método: producir puntos de contorno que al ser agrupador forman trazos finos, por lo que no se requieren otros procesos de adelgazamiento. Entre los inconvenientes está no ofrecer directamente información sobre la dirección ni el sentido en que tiene lugar la transición de claro a oscuro en cada punto de contorno ni la magnitud de dicha transición.

5.2.7. MÉTODO LAPLACIANO GAUSSIANO

Se basa en los cruces por cero de la segunda derivada, explora el hecho de que un salto de borde corresponde a un cambio de la función de la imagen. La primera derivada de la función de la imagen debe tener un extremo en la posición correspondiente a un borde en la imagen, y la segunda derivada debe ser cero en la misma posición.

El punto crucial es calcular la segunda derivada de una forma robusta. Reducir el ruido y después calcular la segunda derivada es una posibilidad. Aunque estos dos criterios pueden entrar en conflicto, la distribución que optimiza simultáneamente ambos es la distribución gaussiana definida a continuación:

$$LoG(x, y) = -1 / \pi \sigma^4 \left[1 - x^2 + y^2 / 2\sigma^2 \right] e^{x^2 + y^2 / 2\sigma^2} \quad (\text{Ec. 5.5})$$

Se pueden usar aproximaciones discretas del filtro Laplaciano gaussiano para ser utilizadas como mascarar de convolución. La máscara de convolución utilizada para implementar el filtro Laplaciano gaussiano con $\sigma = 2.5$ es :

Tabla 2.1.-Mascara de convolución Laplaciano Gaussiano

.00032	.00024	.00016	.00012	.00016	.00024	.00032
.00024	.00008	.00010	.00017	.00010	.00008	.00024
.00016	.00010	.00035	.00046	.00035	.00010	.00016
.00012	.00017	.00046	.00058	.00046	.00017	.00012
.00016	.00010	.00035	.00046	.00035	.00010	.00016
.00024	.00008	.00010	.00017	.00010	.00008	.00024
.00032	.00024	.00016	.00012	.00016	.00024	.00032

5.2.8. MÉTODO DE KIRSCH

Como resumen podemos indicar que los métodos basados en gradiente se fijan en los cambios bruscos del nivel de gris alrededor de un pixel de la imagen, estos cambios pueden detectarse con la primera derivada o gradiente:

El algoritmo se reduce multiplicando elemento a elemento y sumarlo todo. Esto nos dará un valor que habrá que escalar y que pondremos en lugar del pixel evaluado. Si es un valor suficientemente grande, será un borde.

Lo que hace Kirsch es aplicar ocho matrices (una para cada dirección, como se muestra en la figura 5.14) al pixel en cuestión. De los resultados nos quedamos con el de mayor módulo y ese es el que tomamos. Con esto tenemos dos ventajas: la primera es que directamente tenemos la dirección en la que según el borde si es que existe (la dirección del valor que nos quedamos), ahorrándonos tener que calcular el arco tangente para saberlo.

5	5	5	-3	5	5f	-3	-3	5	-3	-3	-3
-3	0	-3	-3	0	5	-3	0	5	-3	0	5
-3	-3	-3	-3	-3	-3	-3	-3	5	-3	5	5

-3	-3	-3	-3	-3	-3	5	-3	-3	5	5	-3
-3	0	-3	5	0	-3	5	0	-3	5	0	-3
5	5	5	5	5	-3	5	-3	-3	-3	-3	-3

Figura 5.14.- Mascara de convolución del método Kirsch

Los resultados posibles son los siguientes:

- Valor obtenido cercano a cero: *no es un borde*
- Valor obtenido parecido en varias direcciones: *no hay información de dirección*
- Valor obtenido mayor al resto de las direcciones: *tenemos un borde e información de dirección.*

Las matrices de Kirsch, como las de los demás operadores gradiente, tienen la sumatoria de sus elementos igual a cero para no influir en la energía del pixel. Como resumen podemos decir que el operador de Kirsch es uno de los que mejor resultado ofrecen, ya que mira en un número suficiente de direcciones (si mirase en más direcciones aumentaría la matriz y por tanto la complejidad) sin elevar en exceso la complejidad.

5.3 UMBRALIZACIÓN.

El método busca segmentar la imagen mediante la definición de uno o varios umbrales con respecto a tonos de gris de pixeles que permiten generar agrupamiento de pixeles que definan los objetos presentes en la imagen.

La umbralización de una imagen en escala de grises, nos permite dado un umbral definido por algún método, separar lo que es fondo de los que es objeto, siempre y cuando el fondo y el objeto tengan sus niveles de gris agrupados en dos modos dominantes.

Si fijamos un umbral de nivel de gris T , digamos 150, y clasificamos a los píxeles con un valor menor a T como negros y los píxeles mayores a T como blancos generamos una imagen binaria. En esta imagen binaria los píxeles negros definirán al objeto y los blancos el fondo, logrando la segmentación.

La operación queda definida como:

$$l_n(x, y) = \begin{cases} 0 & l(x, y) < T \\ 1 & l(x, y) \geq T \end{cases}$$

A continuación se realiza un ejemplo de Umbralización utilizando una imagen en niveles de gris. Ver figura 5.16.

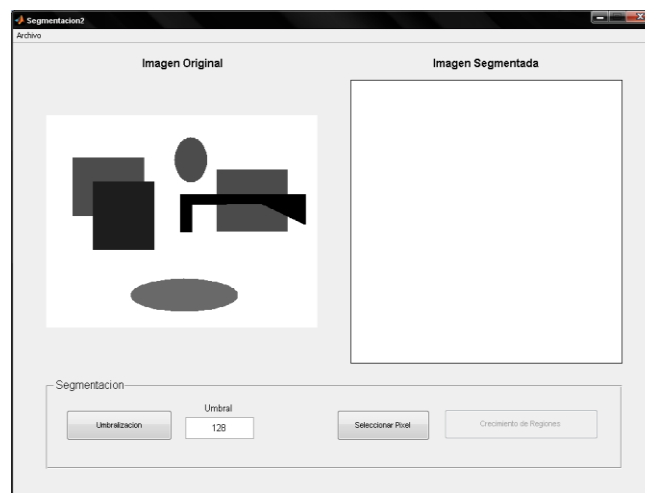


Figura 5.16 Segmentación por Umbralización.

El siguiente paso es seleccionar el valor del umbral a manera de ejemplo: $T=50$ se obtendrá un resultado donde los valores menores a el umbral establecido

tomaran como negros y los pixeles mayores a T como blancos generamos una imagen binaria. Ver figura 5.17.

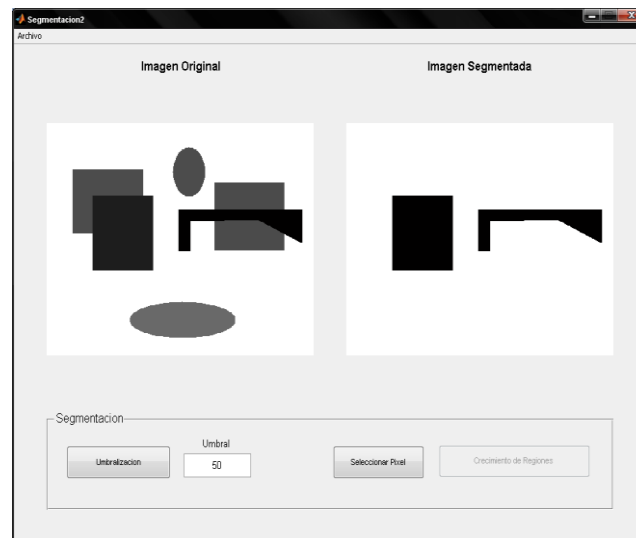


Figura 5.17 Resultado de Segmentación por Umbralización.

Enseguida realizar otro ejemplo modificando el valor de $T=106$ y se podrá observar que ahora es posible que todos los elementos contenidos en la imagen se visualicen en un solo grupo. Ver figura 5.18.

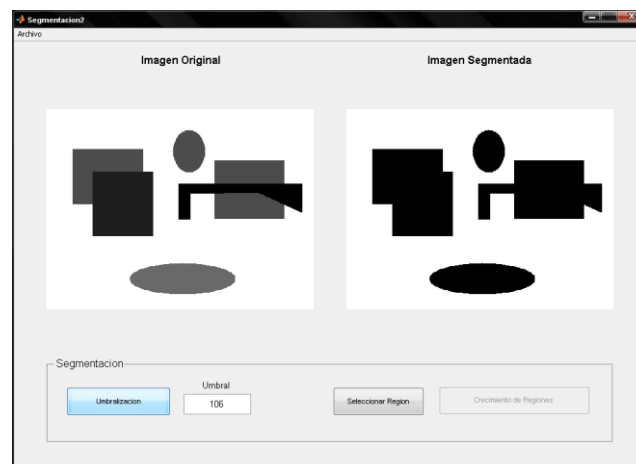


Figura 5.18 Segmentación por Umbralización $T=106$.

Con este material solo se da una pequeña introducción al Tratamiento Digital de Imágenes para lo cual existe una gran variedad de técnicas para el desarrollo de esta área.

5.4 SEGMENTACIÓN ORIENTADA A REGIONES.

Uno de los métodos de segmentación orientada a regiones de tiene la segmentación por Crecimiento de Regiones que como su nombre lo dice, el crecimiento de regiones es un procedimiento que agrupa píxeles o sub-regiones dentro de regiones más grandes. La sencillez de este método radica en la agregación de píxeles, que comienza con un conjunto de puntos generadores a partir de los que van creciendo las regiones al agregar a cada uno de estos puntos los píxeles próximos que tienen propiedades similares (como nivel de gris, textura, color). En este método, las regiones crecen mediante agregación de píxeles similares en valor respecto a la propiedad P que se utilice para realizar la segmentación.

Este tipo de algoritmos necesita que el usuario seleccione un conjunto de puntos semilla en la imagen. Estos puntos semillas servirán como puntos de comienzo del proceso de crecimiento de las regiones, de forma que a partir de cada semilla crecerán regiones añadiendo píxeles a dicha semilla de entre aquellos píxeles vecinos que tengan propiedades similares (como pueden ser el color, la intensidad, la textura).

El resultado de la segmentación dará lugar como mucho a tantas regiones como semillas haya, puede darse el caso de que dos de esas semillas correspondan a píxeles de la misma región, y en este caso el crecimiento desde una de las semillas absorberá a la otra uniendo las regiones para formar una sola.

Para realizar la agregación de píxeles similares será necesario definir el concepto de similaridad, que no tiene porque ser el mismo para todo tipo de aplicaciones. Una propiedad P que se puede utilizar para incluir un píxel en una de las regiones sería que la diferencia absoluta entre el nivel de gris de un píxel N_i y el generador N_g sea menor que un umbral T . Además, será necesario utilizar un criterio para decidir en caso de empate.

Por ejemplo, en la figura 6 se toma como puntos semilla los puntos de coordenadas (3,2) y (3,4). Para la propiedad $P = |N_i - N_g| < T$.

Si la diferencia entre dos píxeles, T es menor o igual (caso de empate) a 3 se tienen dos regiones, en cambio si $T=8$ se tendrá una sola región. Ver figura 5.19.

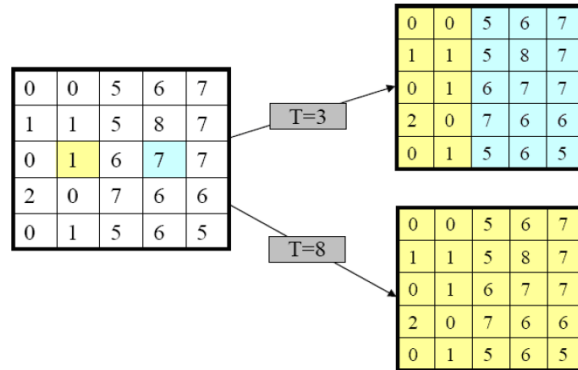


Figura 5.19. Proceso de crecimiento de regiones

En la figura 5.20 se ve la secuencia de crecimiento, en la segmentación de un hueso en una imagen obtenida por resonancia magnética. La región va creciendo hasta alcanzar el borde, donde se detiene el crecimiento:



Figura 5.20 Secuencia del crecimiento de regiones

5.5 UTILIZACIÓN DE MOVIMIENTO.

El movimiento es una poderosa herramienta utilizada por los seres humanos y los animales para extraer objetos de interés de un fondo de detalles irrelevantes. En las aplicaciones de imágenes, el movimiento se deriva de un desplazamiento relativo entre el sistema de sensores y la escena que se está presentando.

El desarrollo de métodos eficientes para calcular el movimiento y la estructura a partir de secuencias de imágenes es una de las áreas de mayor investigación en el campo de la visión artificial. La estimación del movimiento en imágenes tiene una amplia gama de aplicaciones, que pueden ser clasificadas dentro de tres grandes grupos:

- Recuperación de la estructura a partir de secuencias de imágenes (Estructura obtenida del movimiento, SFM), como son la detección de obstáculos para robots de navegación autónoma, la modelización del entorno, la adquisición automática de modelos para diseño asistido por computador (CAD), etc.
- Compresión y reconstrucción de secuencias de imágenes, como son los casos de codificación y decodificación MPEG, y la reconstrucción de imágenes afectadas por ruido.
- Seguimiento y caracterización dinámica de objetos en movimiento para casos como la determinación de parámetros de tráfico de automóviles, sistemas de seguridad por visión artificial y previsión meteorológica. La recuperación del campo de movimiento parece ser la tarea esencial de cualquier sistema de visión artificial que extraiga información a partir de una secuencia de imágenes. Sin embargo, el único dato disponible es la variación espacial y temporal del patrón de brillo de la imagen. De ellas es posible obtener una aproximación del campo de movimiento denominado flujo óptico. El campo de movimiento y el flujo óptico son iguales sólo en el

caso de que las variaciones espaciales del patrón de brillo correspondan a características estructurales de las superficies.

La segmentación de las imágenes consiste en identificar una serie de características locales (objetos, bloques, líneas, etc.) en una imagen, tratar de obtener las correspondencias con la otra imagen y, por último, determinar los movimientos. Tiene el inconveniente de que su complejidad crece con el número de particularidades con el que cuenta la imagen. En la actualidad se comienza a utilizar un método híbrido que utiliza el método de los gradientes para la extracción de contornos y luego se determinan los movimientos de tales contornos con técnicas propias de algoritmos de segmentación. La determinación del movimiento por segmentación o características locales tiene mayor interés en aplicaciones donde se quiere conocer el movimiento global de objetos contenidos en las imágenes, tales como seguimiento de objetos, la determinación de la estructura tridimensional del entorno, etc.

Uno de los métodos más sencillos para detectar cambios entre dos fotogramas $f(x, y, t_i)$ y $f(x, y, t_j)$, tomados en los instantes t_i y t_j respectivamente, es comparar las dos imágenes píxel a píxel.

Un procedimiento para hacerlo es formar una imagen diferencia.

Supóngase que se tiene una imagen de referencia que contiene solamente componentes estacionarios. Al comparar esta imagen con una imagen subsiguiente que tenga el mismo entorno, pero que incluya objetos en movimiento, se obtiene la diferencia de las dos imágenes, que elimina los componentes estacionarios, dejando solamente las entradas distintas de cero, que corresponden a los componentes no estacionarios de la imagen.

Una imagen diferencia de otras dos tomadas en los tiempos t_i y t_j se puede definir como:

$$d_{i,j}(x,y) = \begin{cases} 1 & \text{si } |f(x,y,t_i) - f(x,y,t_j)| > \theta \text{ siendo un umbral} \\ 0 & \text{en caso contrario} \end{cases}$$

Este mecanismo es posible utilizarlo con imágenes grabadas y con la iluminación relativamente constante.

Obsérvese que $d_{i,j}(x,y)$ tiene un 1 como coordenadas espaciales (x,y) solamente si la diferencia de nivel de gris de dos imágenes es apreciablemente diferente de estas coordenadas, según el valor determinado por el umbral θ .

En el análisis dinámico de imágenes, todos los píxel de $d_{i,j}(x,y)$ con valor 1 se consideran como el resultado del movimiento del objeto. Esta aproximación es aplicable solamente si las dos imágenes están grabadas y la iluminación es relativamente constante entre los límites establecidos por θ .

En la práctica, las entradas de $d_{i,j}(x,y)$ de valor 1 aparecen a menudo a causa del ruido. Normalmente, estas entradas son puntos aislados de la imagen diferencia, y un método sencillo para eliminarlas consiste en formar regiones 4 u 8 conexas de valores 1 en $d_{i,j}(x,y)$ e ignorar después cualquier región que tenga menos de un número predeterminado de entradas.

Si bien de esto puede resultar que se ignoren los objetos pequeños y/o de movimiento lento, este método aumenta la probabilidad de que el resto de las entradas de la imagen diferencia sean realmente el resultado del movimiento.

La figura siguiente ilustra estos conceptos. La figura 5.21(a) muestra un fotograma de la imagen de referencia tomado en el instante t_i que contiene un objeto sencillo de

intensidad constante que se está moviendo con velocidad uniforme sobre la superficie del fondo, también de intensidad constante. La figura 5.21(b) muestra un fotograma tomado en el instante t_j , y la figura 5.21(c), la diferencia entre ambas imágenes.

Obsérvese que el proceso de diferenciación genera dos regiones disjuntas: una es el resultado del borde anterior y la otra del borde posterior del objeto en movimiento.

000000111110000000	000000001111100000	000000110001100000
000001111111000000	000000011111110000	000001100000110000
000011111111100000	000000111111111000	000011000000011000
000111111111110000	000001111111111100	000110000000001100
000111111111110000	000001111111111100	000110000000001100
000011111111100000	000000111111111000	000011000000011000
000001111111000000	000000011111100000	000001100000110000
000000111110000000	000000001111100000	000000110001100000

Figura 5.21 (a) Imagen tomada en el instante t_i , (b) Imagen tomada en el instante t_j , (c) Imagen diferencia.

REPRESENTACIÓN Y DESCRIPCIÓN

Una vez segmentada la imagen en regiones, el conjunto resultante de píxeles segmentados, es necesario que se representen y describan adecuadamente para una posterior identificación.

Para representar una región se cuentan con dos posibilidades la de hacerlo en términos de sus características externas (su contorno) o de acuerdo a sus características internas (los píxeles que componen la región). Elegir el esquema de representación es una tarea de hacer útiles los datos para una posterior etapa.

6.1 ESQUEMAS DE REPRESENTACIÓN.

Las técnicas de segmentación producen datos en bruto de forma de píxeles de un contorno o de una región por lo cual es necesario aplicar esquemas de representación.

CÓDIGO DE CADENA

Los códigos de cadena son utilizados para representar un contorno mediante una sucesión conexa de segmentos de longitud y dirección especificadas. Esta representación se basa en segmentos de conectividad 4 u 8. La dirección de cada segmento es codificada utilizando un esquema de numeración como el que se muestra en la figura 6.1.

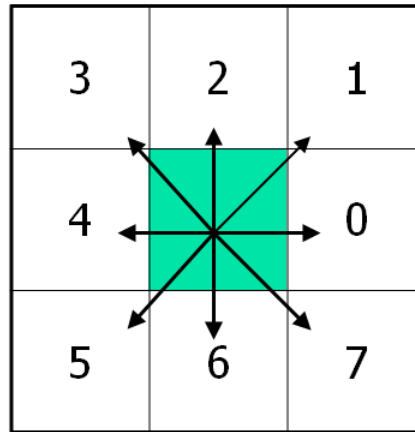


Figura 6.1 Código de cadena de 8 direcciones

Para la obtención del código es necesario establecer un punto inicial y recorrer el contorno de acuerdo al código de cadena a utilizar. Este código de cadena dependerá del punto de partida. Ver figura 6.2, observe las direcciones que sigue de acuerdo al código establecido de 8 direcciones.

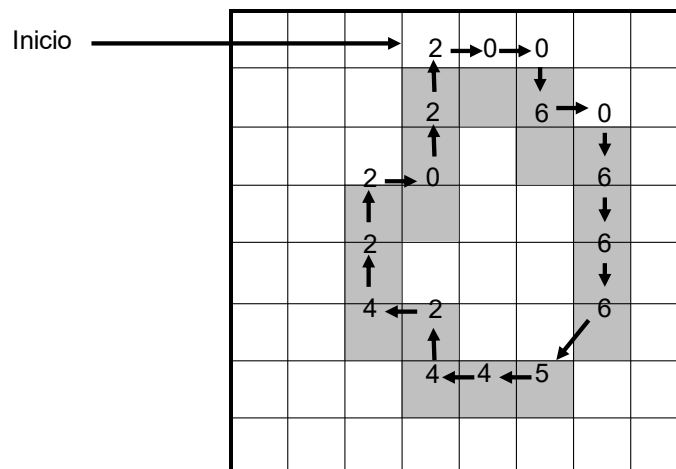


Figura 6.2 Código de cadena usando 8 direcciones

Para el ejemplo de la figura 6.2 el código obtenido es:
0,0,6,0,6,6,6,5,4,4,2,4,2,2,0,2,2

6.2 DESCRIPTORES DE CONTORNO.

Como su nombre lo indican estos descriptores sólo utilizan la información del contorno de la figura. Uno de los descriptores más simples es la longitud de un contorno. Contando sencillamente el número de píxeles a lo largo del contorno se obtiene una aproximación somera de su longitud.

Dentro de los descriptores de contorno podemos encontrar el código de de Freeman o código de cadena visto anteriormente.

6.3 DESCRIPTORES DE REGIÓN.

Aunque no existe una definición formal de la textura [Gonzalez,96], éste descriptor proporciona intuitivamente medidas de propiedades tales como suavizado, rugosidad y regularidad. Los tres métodos principales más utilizados en el procesamiento de imágenes para describir la textura de una región son los estadísticos, estructurales y espectrales. Las soluciones estadísticas proporcionan características de textura tales como suavidad, rugosidad, granulosidad y otras similares. Las técnicas estructurales consisten en la composición de primitivas de imágenes, tales como la descripción de texturas basadas en líneas paralelas regularmente espaciadas. Las técnicas espectrales se basan en las propiedades del espectro de Fourier y se utilizan primordialmente para detectar la periodicidad global de una imagen mediante la identificación de picos estrechos de alta energía del espectro [Dougherty,1994].

MÉTODOS ESTADÍSTICOS

Uno de los métodos más simples para describir la textura consiste en utilizar momentos del histograma de nivel de gris de una imagen o región. Sea z una variable aleatoria que indica la intensidad de una imagen discreta y sea $p(z_i), i = 1, 2, \dots, L$ el correspondiente histograma, donde L es el número de niveles

de intensidad diferentes. El momento n-ésimo de z respecto a la media se muestra en la Ecuación 6.3.1 [Dougherty,1994] [Gonzalez,2008].

$$\mu_n(z) = \sum_{i=1}^L (z_i - m)^n p(z_i) \quad (\text{Ec. 6.3.1})$$

Siendo m el valor medio de z (la intensidad media):

$$m = \sum_{i=1}^L z_i p(z_i) \quad (\text{Ec. 6.3.2})$$

Se observa de la Ecuación 6.3.1 que $\mu_0 = 1$ y $\mu_1 = 0$. El segundo momento (denominado también la varianza y representado por $\sigma^2(z)$) es de particular importancia en la descripción de textura. Siendo una medida del contraste de nivel de gris que se puede utilizar para establecer descriptores de suavidad relativa. Por ejemplo, la medida de la ecuación 6.3.3.

$$1 - \frac{1}{1 + \sigma^2(z)} \quad (\text{Ec. 6.3.3})$$

Donde, el resultado es 0 para áreas de intensidad constante ($\sigma^2(z) = 0$ si todas las z_i tienen el mismo valor) y se aproxima a 1 para valores grandes de $\sigma^2(z)$. El tercer momento es una medida de la desviación del histograma, mientras que el cuarto es la monotonía relativa. El quinto momento así como los siguientes no son tan fácilmente relacionados con la forma del histograma, pero proporcionan una mayor discriminación cuantitativa del contenido de la textura.

Las medidas de la textura calculadas utilizando solamente histogramas presentan la limitación de no contener información referente a la posición relativa de cada pixel con respecto a los otros. Una forma de introducir este tipo de información en

el proceso del análisis de la textura consiste en tomar en cuenta la distribución de intensidades y también las posiciones de pixeles que tienen valores iguales o similares de intensidad.

Sea P un operador de posición y A una matriz $k \times k$ cuyo elemento a_{ij} es el número de veces que aparecen pixeles con nivel de gris z_i (en la posición especificada por P) en relación con los pixeles con nivel de gris z_j con $1 \leq k$. Por ejemplo:

Considérese una imagen de 10×10 pixeles con 256 niveles de gris (ver Figura 6.3.1 (a)). Al necesitar la matriz de co-ocurrencia con niveles de gris similares, el primer paso es convertir la imagen a 8 niveles de gris de acuerdo con las escalas obtenidas de la división de 256 entre los niveles deseados (para este caso son 8 niveles), el resultado de la operación dará como resultado el tamaño de los rangos a realizar como se muestra en la figura 6.3.1(c). Al obtener los rangos, el siguiente paso es convertir el nivel de 256 a su equivalente de acuerdo a los rangos obtenidos, en el caso del nivel de 255 de acuerdo al rango se convierte al nivel 8 y así con cada uno de los valores obteniendo la matriz de la Figura 6.3.1 (b).

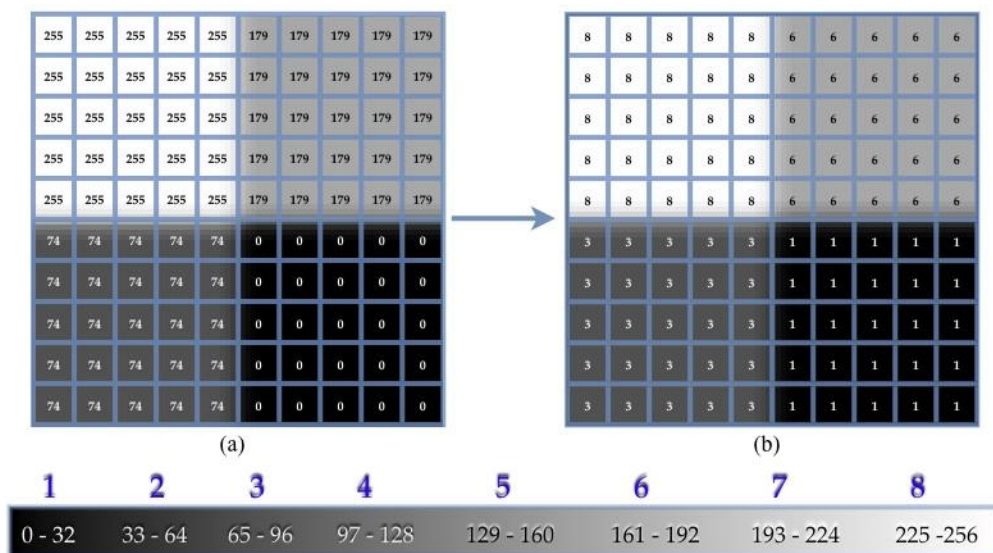


Figura 6.3.1. Representación matricial de una imagen de 10×10 pixeles. (a) Imagen con 256 niveles de gris; (b) Imagen con 8 niveles de gris; (c) Escala de valores para la conversión de 256 a 8 niveles de gris.

Obtenida la imagen a 8 niveles de gris, se procede a obtener la matriz de co-ocurrencia definiendo el operador de posición P como un pixel a la derecha (0°). Entonces se busca la ocurrencia de cada pixel con el que tenga a su derecha para generar la matriz de co-ocurrencia de 8×8 ver Figura 6.3.2.

20	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
5	0	20	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	20	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	5	0	20

Figura 6.3.2. Matriz de co-ocurrencia resultante de la Figura 6.3.1.

El tamaño de la matriz de co-ocurrencia está determinado estrictamente por el número de niveles de gris diferentes de la imagen de entrada. Por ello la aplicación de los conceptos presentados normalmente requiere que se vuelvan a cuantificar las intensidades en unas pocas bandas de nivel de gris para mantener el tamaño de la matriz razonable.

Sea n el número total de pares de pixeles de la imagen que satisfacen P (en el ejemplo anterior $n = 90$). Si se forma una matriz C dividiendo cada elemento de A por n , entonces c_{ij} es una estimación de la probabilidad conjunta de que un par de pixeles que satisfacen P tengan valores (z_i, z_j) . La matriz C se denomina matriz de co-ocurrencia de nivel de gris (Figura 6.3.3). Como C depende de P , se puede

detectar la presencia de un patrón de textura dado eligiendo un operador de posición apropiado. Por ejemplo, el operador utilizado en el ejemplo anterior es sensible a bandas de intensidad constante ejecutando a 0°. Más generalmente, el problema consiste en analizar una matriz dada C con el fin de establecer la categoría de la textura de la región sobre la que se calculó C . A continuación se muestran un conjunto de descriptores útiles para este fin. Ver Tabla 6.3.1 y 6.3.2.

Tabla 6.3.1. Medidas estadísticas de primer orden

Medidas Estadísticas de Primer Orden		
Descriptor	Definición	Formula
Media	Es la esperanza matemática de los niveles de intensidad de la imagen.	$m = \sum_{i=1}^l c(i)$
Kurtosis	Es la medida de lo plana que es una imagen respecto a una distribución normal.	$S = \frac{\sum_{i=0}^{l-1} (c_i - \mu)^4}{(N-1)\sigma^4}$
Entropía	Es la medida de aleatoriedad de los elementos del histograma.	$e = -\sum_{i=0}^{l-1} c_i \log_2 c_i$
Momento 2 (Varianza)	Es la medida del ancho de la función de densidad en una imagen.	$\sigma^2 = \sum_{i=0}^{l-1} (i - m)^2 c_i$
Momento 3 (Desplazamiento)	Es la medida de la desviación estandar del histograma en una imagen.	$\sigma^3 = \sum_{i=0}^{l-1} (i - m)^3 c_i$

Tabla 6.3.1. Medidas estadísticas de primer orden (continuación...).

Momento 4	Es la medida de la monotonía relativa de una imagen.	$\sigma^4 = \sum_{i=0}^{l-1} (i - m)^4 c_i$
Energía	Es la medida de la uniformidad de una imagen, su valor resultante es 1 cuando la uniformidad es constante.	$E = \sum_{i=0}^{l-1} c_i ^2$
Skewness	Es la medida que permite identificar si los datos se distribuyen de forma uniforme alrededor del punto central (Media aritmética).	$S = \frac{\sum_{i=0}^{l-1} (c_i - \mu)^3}{(N - 1)\sigma^3}$

Tabla 6.3.2. Medidas estadísticas de segundo orden

Medidas Estadísticas de Segundo Orden		
Descriptor	Explicación	Formula
Máxima Probabilidad	Es la medida de la ocurrencia más predominante en la imagen.	$\max_{i,j} (c_{ij})$
Contraste	Es la medida de intensidad de contraste entre un pixel y su vecino. El valor resultante es 0 cuando el contraste en una imagen es constante.	$\sum_i \sum_j (i - j)^k c_{ij}$
Homogeneidad	Es la medida de proximidad espacial de los valores de la matriz de co-ocurrencia en su diagonal.	$\sum_{i=1}^k \sum_{j=1}^k \frac{c_{ij}}{1 + i - j }$

Entropía	Es la medida de aleatoriedad de los elementos de la matriz de co-ocurrencia.	$-\sum_{i=1}^l \sum_{j=1}^l c_{ij} \log c_{ij}$
Energía	Es la medida de la uniformidad de una imagen, su valor resultante es 1 cuando la uniformidad es constante.	$\sum_{i=1}^k \sum_{j=1}^l c_{ij} ^2$
Correlación	Es la medida que verifica que tan relacionado está un pixel con respecto a su vecino.	$\frac{\sum_{i=1}^l \sum_{j=1}^l (i - \mu_x)(j - \mu_y) c_{ij}}{\sigma_x \sigma_y}$

La idea básica es caracterizar el contenido de c por medio de estos descriptores. Un método para utilizar estos descriptores es preparar a un sistema con los valores de los descriptores representativos para un conjunto de texturas diferentes. Por consiguiente, la textura de una región desconocida se determina por la exactitud con la que sus descriptores concuerdan con los almacenados en la memoria del sistema [González, 2008].

MÉTODOS ESTRUCTURALES

Como se mencionó anteriormente, una segunda categoría importante de descripción de la textura se basa en conceptos estructurales. Supóngase que se tiene una regla de la forma $S \rightarrow aS$, esto indica que el símbolo S se puede volver a escribir como aS (por ejemplo, tres aplicaciones de esta regla producirán la cadena $aaaS$). Si a representa un círculo y se asigna el significado de círculos a la derecha a una cadena de la forma $aaa\dots$, la regla $S \rightarrow aS$ permite generar el patrón de la textura que se muestra en la Figura 6.3.3.

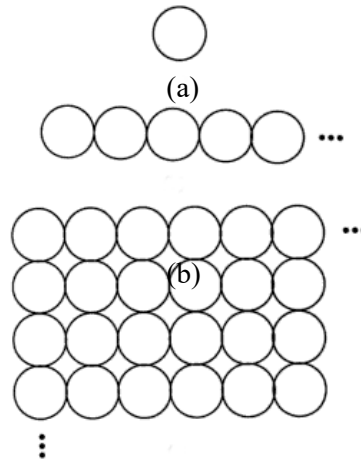


Figura 6.3.3. Patrón de textura generada por la regla $S \rightarrow aS$; (a) Textura primitiva; (b) Patrón generado por la regla $S \rightarrow aS$; (c) Patrón de textura bidimensional generado por ésta y otras reglas.

Supóngase a continuación que se añaden algunas nuevas reglas a este esquema: $S \rightarrow bA$, $A \rightarrow cA$, $A \rightarrow c$, $A \rightarrow bS$, $S \rightarrow a$, donde la presencia de una b significa círculo bajo y la presencia de una c significa círculo a la izquierda. Podemos ahora generar una cadena de la forma $aaabccbaa$ que corresponde a una matriz de círculos de 3×3 . De igual forma se pueden generar fácilmente patrones de textura mayores, tales como el que se muestra en la Figura 6.3.3(c).

La idea básica de la exposición anterior es una simple textura primitiva se pueda utilizar para formar patrones de estructuras más complejas por medio de algunas reglas que limitan el número de posibles composiciones de la(s) primitiva(s) [Chacon,07] [Gonzalez,08] .

MÉTODOS ESPECTRALES

El espectro de Fourier está idealmente indicado para describir la direccionalidad de patrones bi-dimensionales periódicos de una imagen. Estos patrones de textura global, aunque son fácilmente distinguibles como concentraciones de alta energía

del espectro, generalmente son bastante difíciles de detectar con métodos espaciales a causa de la naturaleza local de estas técnicas. Se consideran 3 características del espectro de Fourier que son útiles para la descripción de la textura: 1) Picos prominentes del espectro que dan la dirección principal de los patrones de textura; 2) Localización de los picos en el plano de frecuencia que dan el período espacial fundamental de los patrones, y 3) Eliminación de los componentes periódicos mediante el filtrado que deja elementos de la imagen no periódicos, que se pueden describir por técnicas estadísticas. Se recuerda que el espectro de una imagen real es simétrico con respecto al origen, así sólo se necesita considerar la mitad del plano de frecuencia. Por lo tanto, a efectos del análisis, cada patrón periódico está asociado con un solo pico del espectro, en lugar de dos. La figura 6.3.4 ilustra el empleo de las funciones espectrales para una descripción global de la textura [Gonzalez, 1996] [Dougherty, 1994].

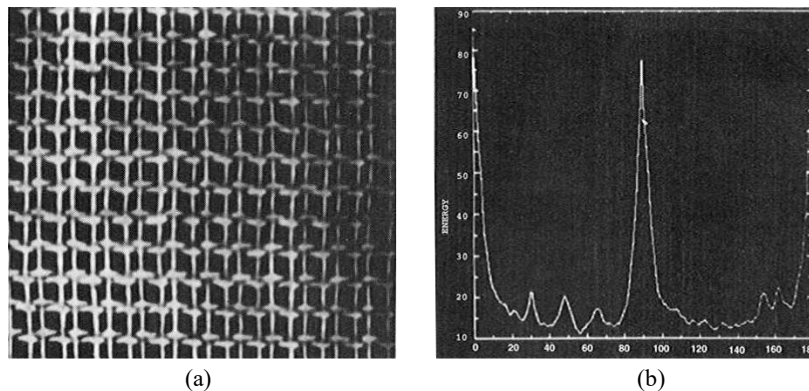


Figura 6.3.4. Descripción global de textura empleando funciones espectrales; (a) Imagen con textura periódica; (b) Diagrama de la energía espectral de la textura.

6.4 DESCRIPTORES RELACIONALES.

DESCRIPCIÓN DE LA IMAGEN

Modelo que representa las características de la imagen de relevancia para los fines del Sistema de Visión Artificial específico. Los modelos que describen una imagen pueden ser estructurales o sintácticos, se obtienen a partir de las imágenes previamente segmentadas.

A partir de una imagen segmentada debemos obtener las primitivas de los objetos presentes en la imagen, las cuales deberán utilizarse en la construcción del modelo matemático de forma que representara a dicho objeto.

Las primitivas de un objeto son:

Primitivas Globales: presentan información sobre la totalidad del objeto, por tanto nos permiten localizar la posición y orientación de un objeto en el plano. Por ejemplo: los momentos de inercia, el baricentro, el perímetro, el área, entre otras.

Primitivas Locales: informan sobre un segmento del objeto (por ejemplo, una parte de su contorno). Encontramos formas geométricas elementales: rectas, segmentos de curva, ángulos.

Descripción

Representación de la imagen segmentada por modelos (matemáticos, geométricos).

- Polígonos formados por rectas.
- Modelo del contorno: continuo o secuencial.
- Ocupación espacial del objeto 2D considerado.
- Objetos 2D formados por un conjunto de píxeles.

La descripción ya no es una imagen.

- Descripción por contornos
 - ✓ Transformada de Hough.
 - ✓ Firma.
 - ✓ Código de Freeman.
- Descripción por regiones.
 - ✓ Quad tree.
 - ✓ Etiquetado.
 - ✓ Crecimiento.

Bibliografía:

- [Abundez, 2002] Abundez Barrera Itzel María "Desplazamiento de un Robot Móvil para el Seguimiento de Camino en Línea", Tesis de Grado, Instituto Tecnológico de Toluca. 2002.
- [Chacon, 2007] Chacon Mario, "Procesamiento Digital de Imágenes". Trillas, 2007.
- [Dougherty, 1992] Dougherty Edward R., "Mathematical Morphology in Image Processing". Ilustrada Edición., CRC Press, 1992.
- [Dougherty, 1994] Dougherty Edward R., "Digital Image Processing Methods". Optical Engineering, 1994.
- [Escalera, 2001] De la Escalera Hueso Arturo. Visión por Computador Fundamentos y Métodos. Editorial Prentice Hall. Año 2001.
- [González, 1996] R.C. González, and Woods. Tratamiento Digital de Imágenes. Editorial Addison-Wesley / Díaz de Santos. Año 1996.
- [Gonzalez, 2004] Gonzalez Rafael C., "Digital Image Processing Using MATLAB". Illustrated Edición., Pearson Prentice Hall, 2004.
- [Gonzalez, 2008] Gonzalez Rafael C., "Digital Image Processing". Tercera Edición., Prentice Hall, 2008.

- [Kvan, 1999] Kvan C., Phipps Gary, "Autonomous Robotic Vehicle Road Following", IEEE Transaction On Robotics And Automation , 1999, pp. 648 – 659.
- [Pajares, 2002] Gonzalo Pajares. Visión por Computador. Editorial Alfaomega RA-MA. Año 2002.
- [Pla, 1997] Pla F, Sanchiz J .M, "Building Perspective Models To Guide A Row Crop Navigation Vehicle", Image And Vision Computing , 1997 , pp.465 – 473.
- [Pratt, 1991] Pratt W.K, "Digital Image Processing", Second Edition. E.U.A, Wiley Interscience, 1991.
- [Rowley, 1995] Henry A. Rowley, Shummet Baluja, Takeo Kanade, 2002, "Human Face Detection in Visual Scenes"
- [Russ, 1995] Russ K. C. "The Image Processing Handbook", Second Edition. E.U.A. CRC Press, 1995, pp. 98-103.
- [Schalkoff, 1989] Schalkoff R.J, "Digital Image Processing and Computer Vision." , Wiley, 1989.
- [Trucco, 1999] Trucco E, Verri Alessandro, "Introductory Techniques For 3-D Computer Vision", Prentice Hall, 1999, pp. 123 – 138.
- [Valero, 2002] Valero Ruiz C., González Gil J. A., 2002, "La imagen estereoscópica en formato digital. Un nuevo medio de expresión". XIV Congreso Internacional de ingeniería grafica Santander España.
- [Arellano, 2005] Arellano, Díaz Fernando, Odin Salgado Tinoco, Un Algoritmo de Correspondencia entre Imágenes Estereo a Múltiples Resoluciones, (Agosto, 2005).