

71



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE INGENIERÍA

SISTEMAS DE EVALUACIÓN Y ESPECIFICACIÓN DE REQUERIMIENTOS EN IMÁGENES SATELITALES

T E S I S
PARA OBTENER EL TÍTULO DE:
INGENIERO EN COMPUTACIÓN
P R E S E N T A :
JUAN MIGUEL PANTOJA RAMOS

DIRECTOR DE TESIS: ING. ROBERTO REYES CHALICO

276825



CIUDAD UNIVERSITARIA

2000



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

AGRADECIMIENTOS

Doy gracias a Dios por haberme dado la fuerza para terminar la tesis.

Esta tesis esta dedicada a mi mamá Ernestina Ramos García, a mis hermanas Eva Sánchez Ramos y Minerva Sánchez Ramos. También esta dedicada a Karina Torres Reyes.

Por todo el apoyo que me han brindado de forma incondicional.

Gracias.

**SISTEMAS DE EVALUACIÓN Y ESPECIFICACIÓN
DE REQUERIMIENTOS EN IMÁGENES
SATELITALES**

ÍNDICE GENERAL

INTRODUCCIÓN

CAPÍTULO I

1 FORMACIÓN DE LA IMAGEN	3
1.1 Formación de una imagen	3
1.1.1 Fuente de iluminación	4
1.1.2 Paisaje	4
1.1.3 Escena	4
1.1.4 Sensor remoto	4
1.1.5 Plataforma	5
1.1.6 Sistema de procesamiento	5
1.2 Sistema sensor	5
Introducción	
1.2.1 Sistema Landsat	6
1.2.2 Sistema Spot	13
1.2.3 Escala de gris	17
1.2.4 Modelo RGB	21
1.2.5 Falso color	25
1.3 Atributos y elementos de análisis de una imagen	30
1.3.1 Atributos	30
1.3.1.1 Escala	30
1.3.1.2 Brillantez	30
1.3.1.3 Tono	30
1.3.1.4 Contraste	31
1.3.1.5 Resolución	31
1.3.1.6 Detectibilidad	31
1.3.1.7 Reconocibilidad	31
1.3.1.8 Firma	31
1.3.1.9 Textura	32
Método estadístico	33
Método estructural	35
Método espectral	36
1.3.2 Elementos de una imagen	38
1.3.2.1 Tonalidad o color	38
1.3.2.2 Comportamiento espectral	38
1.3.2.3 Tamaño	39
1.3.2.4 Forma	39
1.3.2.5 Altura	39
1.3.2.6 Sombra	39
1.3.2.7 Localización	39
1.3.2.8 Contexto	39

CAPÍTULO II

2 PROCESAMIENTO DE IMÁGENES	42
2.1 Modificación del histograma	42
2.2 Filtrado	56
Introducción	56
2.2.1 Filtrado espacial	58
2.2.1.1 Filtros lineales	59
Filtro paso bajas	60
Filtro paso bajas de media ponderada	61
Filtro paso altas	62
Filtro supresor de bandas	64
Filtro paso banda	64
2.2.1.2 Filtros no lineales	66
Filtro de mediana	66
Filtro de máxima y de mínima	67
Filtro de moda	67
Aplicación del gradiente(detección de bordes)	68
2.2.2 Filtrado frecuencial	71
2.2.2.1 Transformada de Fourier	73
2.2.2.2 Transformada rápida de Fourier	76
2.2.2.3 Propiedades de la transformada de Fourier	79
Traslación	79
Periodicidad y simetría	80
Rotación	81
Linealidad y escala	81
2.2.2.4 Filtro paso bajas	82
2.2.2.5 Filtro paso altas	85
2.2.2.6 Filtro rechazo banda	86
2.2.2.7 Filtro paso banda	87

CAPÍTULO III

3 CLASIFICACIÓN DE IMÁGENES	89
3.1 Clasificación de una imagen	91
Introducción	91
3.1.1 Conceptos básicos	92
Clases informacionales	92
Clases espectrales	92
Patrón	92
Patrón espectral	93
Patrón espacial	93
Pixel o CIV	93
Espacio de observaciones	93
Clase rechazo	93
Regla de clasificación	93
Clases discriminantes	94
Clasificador	94
Función discriminante	94
Regla de decisión	94
Superficie o superficie de decisión entre las clases	94
Distancia de Mahalanobis	95
Regla de decisión óptima de Bayes	96
Riesgo medio de Bayes	96
Regla de máxima verosimilitud	96
Conjunto de entrenamiento	97
Estimadores de error	97
3.1.2 Proceso de entrenamiento	97
3.2 Clasificación supervisada	99
3.2.1 Clasificación por el teorema de Bayes	100
3.2.2 Clasificación por máxima verosimilitud	102
3.2.3 Clasificación por mínimas distancias	103
3.2.4 Clasificación por paralelepípedos	104
3.2.5 Clasificación no paramétrica	105
3.2.5.1 Clasificación por medio de árboles	105
3.3 Clasificación no supervisada	108
3.3.1 Clasificación por migración de medias	108
3.3.2 Clasificación por agrupación de dos pasadas	109

CAPÍTULO IV

4 CASO PRÁCTICO	112
Introducción	114
4.1 Imagen de un muelle	115
4.1.1 Descripción	115
4.1.2 Delimitación de requerimientos	116
4.1.3 Selección de técnicas de procesamiento y clasificación	116
4.1.4 Análisis y discusión de resultados	123
4.2 Imagen de Washington DC	124
4.2.1 Descripción	124
4.2.2 Delimitación de requerimientos	125
4.2.3 Selección de técnicas de procesamiento y clasificación	125
4.2.4 Análisis y discusión de resultados	129
4.3 Imagen de un área urbana	130
4.3.1 Descripción	130
4.3.2 Delimitación de requerimientos	130
4.3.3 Selección de técnicas de procesamiento y clasificación	131
4.3.4 Análisis y discusión de resultados	134
Conclusiones	136
Bibliografía	138
Apéndice	141

INTRODUCCIÓN

Cada día los satélites representan un papel muy importante en las comunicaciones, pero su campo de desarrollo no es exclusivamente a éstas. Las imágenes obtenidas desde satélite han modificado el estudio de la superficie terrestre y la evaluación de los recursos naturales, y es aquí donde entra la aplicación de las técnicas de clasificación de imágenes con la finalidad de detectar zonas áridas, zonas boscosas dañadas por incendios o plagas, detectar agua, montañas, agricultura, maremotos, zonas rurales o urbanas. También se encarga de identificar cómo va cambiando la superficie terrestre con el tiempo y así poder predecir los cambios futuros que puede sufrir la área de estudio en determinado lapso de tiempo. Sin embargo las imágenes obtenidas por los satélites muchas veces no son aptas para la clasificación debido a que tienen mucho ruido por factores externos, no controlables como las nubes o por una mala recepción de la imagen mandada por satélite, es decir, que la imagen es alterada por interferencia durante la transmisión desde el satélite a la tierra, también porque la imagen tiene muy poco contraste por falta de una buena iluminación o demasiado contraste por causa de una saturación en los sensores remotos, otro de los problemas es la definición de los bordes debido a la distancia en la que se ubica el satélite y por su baja resolución espacial, en muchas ocasiones por el movimiento del satélite y de la tierra, provocan un movimiento uniforme al momento de la captura de la imagen obteniendo una imagen borrosa con muy poca nitidez. Son muchos los factores que puedes influir para no tener una imagen adecuada para la clasificación, por eso es necesario tener en cuenta el procesamiento digital de la imagen, ya que su función principal es la mejora de la imagen de forma que se aumenten las posibilidades de éxito en la clasificación de la imagen.

Se puede esquematizar a un sistema de procesamiento de imágenes digitales en tres bloques, conteniendo los siguientes elementos:

- ◆ Adquisición de imagen. Involucra la conversión de la escena a una representación digital que pueda ser procesada por una computadora. Dicha adquisición de imágenes puede hacerse por medio de un conjunto de sensores especialmente diseñados para observar una escena y proveer una representación digital de la misma.
- ◆ Procesamiento de la imagen. Este bloque provee el proceso digital de una o más imágenes para producir un resultado deseado. El procesamiento puede ir desde el mejoramiento de una imagen individual para un despliegue más detallado, a un proceso más complejo involucrando varias imágenes compuestas (por ejemplo en la clasificación multiespectral).
- ◆ Despliegue de la imagen. Provee un producto de salida que puede ser visto por un observador humano. El despliegue puede realizarse utilizando monitores o papel.

Por lo regular siempre se puede encontrar libros especializados en cómo se lleva a cabo la captura de una imagen, o libros que explican sobre los tipos de procesamiento digital de imágenes y otros sobre técnicas de clasificación.

El objetivo de esta tesis es integrar toda la información desde cómo se forma la imagen pasando por el procesamiento, la clasificación y la comprobación práctica utilizando tres imágenes mostrando al lector los criterios de evaluación y especificación de requerimientos en imágenes satelitales.

En el transcurso de los capítulos se hablará sólo de aquellos métodos más usados y que mejor resultados han dado para el procesamiento y clasificación en imágenes.

CAPÍTULO I

FORMACIÓN DE UNA IMAGEN

CONTENIDO

1 FORMACIÓN DE LA IMAGEN

1.1 Formación de una imagen

- 1.1.1 Fuente de iluminación
- 1.1.2 Paisaje
- 1.1.3 Escena
- 1.1.4 Sensor remoto
- 1.1.5 Plataforma
- 1.1.6 Sistema de procesamiento

1.2 Sistema sensor

- Introducción
- 1.2.1 Sistema Landsat
- 1.2.2 Sistema Spot
- 1.2.3 Escala de gris
- 1.2.4 Modelo RGB
- 1.2.5 Falso color

1.3 Atributos y elementos de análisis de una imagen

- 1.3.1 Atributos
 - 1.3.1.1 Escala
 - 1.3.1.2 Brillantez
 - 1.3.1.3 Tono
 - 1.3.1.4 Contraste
 - 1.3.1.5 Resolución
 - 1.3.1.6 Detectibilidad
 - 1.3.1.7 Reconocibilidad
 - 1.3.1.8 Firma
 - 1.3.1.9 Textura
 - Método estadístico
 - Método estructural
 - Método espectral
 - 1.3.2 Elementos
 - 1.3.2.1 Tonalidad o color
 - 1.3.2.2 Comportamiento espectral
 - 1.3.2.3 Tamaño
 - 1.3.2.4 Forma
 - 1.3.2.5 Sombra
 - 1.3.2.6 Altura
 - 1.3.2.7 Localización
 - 1.3.2.8 Contexto
-

CAPÍTULO I

FORMACIÓN DE UNA IMAGEN

1.1 FORMACIÓN DE UNA IMAGEN

INTRODUCCIÓN

Los satélites artificiales, los sensores remotos y el manejo de imágenes digitales por medio de computadora, han modificado el estudio de la superficie terrestre y la evaluación de los recursos naturales.

La ciencia que se encarga de esto es la **percepción remota** la cual se define como:

- ◆ La obtención de información acerca de una superficie o escena, utilizando luz visible e invisible, por medio del análisis automatizado de datos obtenidos a distancia por un sensor remoto.
- ◆ La ciencia de captar y coleccionar información de un objeto a través de medidas hechas a distancia, sin estar en contacto físico con él (Purdue University, 1976; Sabin, 1978).
- ◆ El conjunto de actividades que buscan la caracterización de las propiedades de los cuerpos naturales, a través de la detección, registro y análisis del flujo de energía electromagnética por ellos reflejados o emitidos (Steffen et al., 1981).

Con el moderno desarrollo de la ciencia de la computación electrónica digital se ha dado un auge importante a la percepción remota, pues la disponibilidad de estas herramientas permiten la evaluación cuantitativa de un gran volumen de datos. Así, en una simbiosis muy estrecha las ciencias de la computación han impulsado el desarrollo de la percepción remota, y está, a su vez, ha generado nuevos sistemas de procesamiento digital con propósitos específicos.

El desarrollo de la ciencia y de la tecnología ha permitido incorporar nuevos métodos y dispositivos de captura de datos a distancia, lo que ha hecho que la percepción remota extienda su campo de acción a fenómenos muy diversos, ya no sólo al mundo macroscópico sino también en el microscópico.

En la figura 1.1 observamos los elementos básicos que hoy componen la moderna ciencia de la percepción remota.

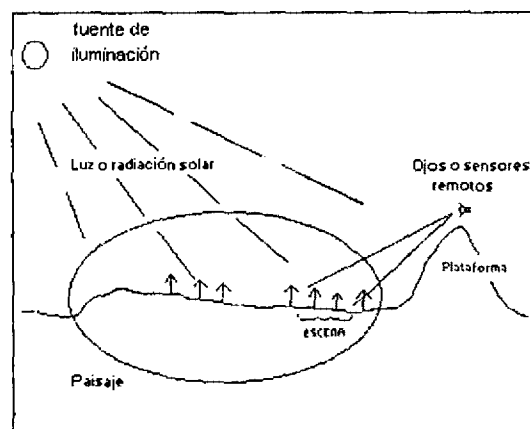


Figura 1.1 Elementos básicos que componen a la percepción remota.

1.1.1 FUENTE DE ILUMINACIÓN

Hay sistemas en percepción remota que utilizan una fuente de iluminación externa, sobre la cual no se puede tener control. Estos sistemas, llamados **sistemas pasivos**, pueden utilizar al sol o bien a una estrella lejana como fuente de iluminación.

Los **sistemas activos** cuentan con su propia fuente de iluminación de la que se tiene control, de aquí que la fuente de iluminación ya no emite luz visible para "iluminar" la escena. Existe una serie de posibilidades tales como: luz ultravioleta infrarroja, rayos gamma ó rayos X, o bien un haz de partículas como protones o neutrones, por lo que la percepción remota pueda aplicarse, no sólo al ámbito de la geofísica, sino también al de la medicina, la biología, la física nuclear y la industrial.

1.1.2 PAISAJE

El paisaje esta compuesto por todos los objetos presentes en el territorio explorado por el hombre, tales como ríos, colinas, vegetación o rocas.

Gracias a la nueva tecnología el territorio explorado por el hombre se extiende a:

- ◆ Un sistema microscópico (conjunto de átomos "iluminado" por un haz de electrones).
- ◆ Un sistema macroscópico (cultivo de bacteria "iluminado" por luz ultravioleta).
- ◆ El cuerpo humano ("iluminado" por rayos X con el propósito de obtener radiografías).

El paisaje es el elemento que presenta mayores retos al investigador porque:

- ◆ Generalmente es una parte muy compleja donde intervienen muchos factores, algunos de ellos ajenos al propósito de la investigación
- ◆ Con frecuencia el investigador no controla una parte o, en algunos casos, todos los aspectos que componen al paisaje.

Esto conduce a que una gran porción de errores que se cometen en el diseño y operación de un experimento de percepción remota provienen precisamente de la subestimación o sobresimplificación de la complejidad del paisaje.

1.1.3 LA ESCENA

Es la sección o superficie del paisaje donde se enfoca su interés empleando el manejo de conceptos lógicos, es decir, de entidades no físicas. Sin embargo, pueden obtenerse diferentes escenas del mismo paisaje (por ejemplo al variar el enfoque de la lente de una cámara fotográfica).

Sobre una fotografía resultante de una escena quedan grabadas permanente y visible el conjunto de objetos que se encontraban sobre la escena convirtiéndose en una representación bidimensional de los objetos enfocados, o lo que es lo mismo, en una imagen de la escena.

1.1.4 SENSOR REMOTO

El sensor remoto permite capturar la luz proveniente de la escena.

Hoy en día existen nuevos y modernos detectores que permiten capturar luz invisible como la luz ultravioleta o la infrarroja, y detectar radiaciones como los rayos X, gamma y partículas como electrones, protones y neutrones. Ejemplos de tales detectores son las placas fotográficas infrarroja, los detectores Geiger, los fotodetectores, los fotomultiplicadores y las cámaras de televisión. Algunos de éstos proporcionan imágenes en forma cuantitativa, es decir, en forma numérica, de tal manera que la información resultante puede ser almacenada en la computadora para proceder a un análisis matemático. Los sensores remotos están constituidos por dos subsistemas básicos: el de formación de imágenes y el sensor o registro de imágenes.

1.1.5 LA PLATAFORMA

La plataforma es el lugar donde se coloca o monta el sensor remoto para obtener una visión en conjunto de la escena.

A medida que una plataforma de observación se coloca a una mayor distancia de la escena, la visión de conjunto crece, pero no así el detalle o grado de distancia el que consecuentemente se hace menor. Distinguir objetos relativamente pequeños o muy próximos entre sí requiere de detectores de alta resolución, lo que hace que sean indispensables para el uso de plataformas, colocados a gran distancia de la escena (aplicaciones geofísicas). Existen plataformas que pueden estar a poca altura de la escena, que tienen otras aplicaciones como la observación de un cultivo de bacteria.

1.1.6 SISTEMA DE PROCESAMIENTO

Compuesto por el dispositivo para procesar cualitativa o cuantitativamente los datos proporcionados por el sensor remoto acerca de la escena.

Los sensores remotos actuales son capaces de producir imágenes de alta calidad en forma analógica o digital, es decir, pueden proporcionar una representación continua o discreta de la escena. Una imagen continua es aquella donde la variación de tonos de gris o color se presenta sin discontinuidades, sin líneas, o fronteras, aparte de las que pudiera tener la escena misma. Una imagen discreta, por su parte, es la que está compuesta por los elementos definidos y diferenciados como puntos o cuadrados. Es necesario aclarar que una escena es siempre continua, no así la imagen respectiva. Por tal motivo una imagen será catalogada continua o discreta dependiendo del grado de resolución que tenga el sensor y del detalle que desee distinguir. De echo, sólo las imágenes ópticas podrán considerarse siempre continuas.

1.2 SISTEMAS SENSORES

INTRODUCCIÓN

El objetivo y principio de estudio del análisis de percepción remota es precisamente el paisaje, sistema físico tridimensional, iluminado por una fuente de iluminación. Por medio del paisaje se encuentra la escena, superficie de carácter subjetivo virtual, sobre la cual se encuentra los objetos de interés. Sobre la escena se define un elemento con dimensiones físicas, denominado campo instantáneo de vista (CIV), cuyo tamaño depende de las dimensiones del paisaje, de la resolución del sensor remoto y de la distancia de éste al CIV. La energía luminosa reflejada o emitida por el CIV incide entonces sobre el sensor optoelectrónico, compuesto en primera instancia por un sistema óptico (generalmente un telescopio) y por un conjunto de dispositivos de conducción de luz que hacen llegar a ésta directamente hasta un juego de filtros de color. Esto es muy importante, pues cada objeto de la naturaleza tiene su muy particular manera de manifestarse cuando es visto bajo diferentes filtros de color.

Si se emplearan cuatro filtros, se obtendrán cuatro imágenes digitales o si utilizáramos siete filtros, se registrarán siete imágenes digitales. Resumiendo esto de manera lógica se puede decir que el conjunto de imágenes digitales (cuatro o siete) se les denomina "imagen multispectral" y al conjunto de números para CIV se les conoce como firma "espectral". De ahí que una imagen multispectral esté formada por un conjunto de imágenes digitales, cada una de ellas conteniendo aquellos aspectos de la escena que corresponde a un color determinado y que a su vez, la firma espectral contenga las características del conjunto de objetos que se encuentran en el CIV, para cada color empleado en la observación correspondiente de la escena.

Para registrar simultáneamente los diferentes aspectos de una escena se requiere no nada más diferentes filtros sino también varios sensores, al menos uno para cada color. En el proceso de observación de la escena, cada sensor genera una imagen digital llamada banda espectral o plano de color y el conjunto de sensores, es, por tanto, una imagen multispectral, la que a su vez queda registrada permanentemente en algún dispositivo de memoria, como por ejemplo una cinta magnética. Cabe agregar aquí que el sistema optoelectrónico formado por un conjunto de filtros y detectores se le conoce como barredor multispectral. Además, que el tiempo que le lleva al sensor remoto medir la energía del CIV se conoce como tiempo de residencia o tiempo de integración y que, en general, se desea que sea lo más corto posible.

Al realizar una medida sobre un CIV determinado, la energía luminosa para un color dado es transformada por el fotodetector correspondiente en un pulso de voltaje cuya intensidad es proporcional a la energía luminosa incidente; a su vez, es transformada a voltaje por medio de un dispositivo electrónico, llamado convertidor analógico/digital (A/D), dicho voltaje es transformado a un número entero en una escala que generalmente es de 0 a 255, con lo que un CIV oscuro emitirá poca energía luminosa y el voltaje correspondiente será de baja intensidad, con lo que se obtendrá un número cercano a cero. En el desarrollo de este proceso se va generando una cantidad enorme de números enteros que deben ser almacenados en algún dispositivo de memoria al mismo ritmo que son producidos. Como el tiempo de residencia es de una fracción muy pequeña de segundo, la velocidad a la que son generados estos números es muy grande y por lo tanto el sistema de almacenamiento debe de operar con una alta eficiencia. Con un sistema de almacenamiento masivo de memoria controlada por un microprocesador se puede realizar una especie de barrido sobre cada fotodetector, capturando el resultado de las medidas al mismo tiempo que se realiza el barrido sobre la escena; Dicho microprocesador no sólo capturará dicha medida, sino que la enviará al sistema de almacenamiento masivo con un código lógico de orden para que el pixel generado y almacenado guarde la misma posición relativa al CIV correspondiente, de otra manera se tendría una especie de rompecabezas en desorden y sería imposible reconstruir la escena.

1.2.1 SISTEMA LANDSAT

El sistema de satélite Landsat, originalmente llamado ERTS, es el primer satélite diseñado con el objeto de permitir la adquisición de datos espaciales, espectrales y temporales de la superficie de la tierra de una forma global y repetitiva.

Al principio, fueron lanzados a órbitas circulares casi polares en altitudes que varían entre 897 y 918 kilómetros. Llevaban consigo un barredor multispectral de cuatro canales, tres sensores de tipo Vidicom de rayo regresible y dos grabadoras de vídeo. El barredor multispectral operaba en los siguientes rangos espaciales:

- Banda 4: de 0.5 a 0.6 μm
- Banda 5: de 0.6 a 0.7 μm
- Banda 6: de 0.7 a 0.8 μm
- Banda 7: de 0.8 a 1.1 μm

Con estos dos sistemas los satélites eran capaces de captar escenas de la tierra 185 * 185 kilómetros con una resolución espacial de 80 metros. Sin embargo, esta resolución se mejoro con el lanzamiento del tercer satélite que contenía una banda más en el barredor multiespectral, específicamente en el infrarrojo térmico(10.4-12.6 µm). Además, los sensores de tipo Vidicom de rayo regresible fue también modificado significativamente empleando dos cámaras idénticas que cubrían el espectro de 0.53 a 0.75 µm, las cuales alineadas simultáneamente sobre el mismo punto obtenían escenas de 84*85 kilómetros en tierra aumentando su resolución espacial a 40 metros.

La serie Landsat consta hasta el momento de 6 satélites y se contempla extenderla hasta 7.

SATÉLITES	PERIODO	SENSOR
Landsat 1	23/07/1972 al 05/01/1978	MSS/RBV
Landsat 2	22/01/1975 al 27/07/1983	MSS/RBV
Landsat 3	05/03/1978 al 07/09/1983	MSS/RBV
Landsat 4	16/07/1982 operando	MSS/TM
Landsat 5	01/03/1984 operando	MSS/TM
Landsat 6	1990/1993 falló lanzamiento	ETM
Landsat 7	En planes	ETM/MLA

Tabla 1.1 Serie de satélites Landsat. MSS = Multiespectral scanner, RBV = Return beam vidicom, TM = Thematic mapper, MLA = Multilinear array.

Como se muestra en la tabla 1.1 la serie Landsat ha producido imágenes por medio de tres sensores diferentes MSS (imágenes en 4 bandas con un pixel rectangular de resolución 79m x 56m, el Landsat 3 contó con una quinta banda en la porción térmica del espectro que tenía una resolución de 120m por lado), RBV (sistema de televisión de 3 canales, el Landsat 3 tenía un solo canal) y TM (barredor multiespectral de 7 bandas y resolución de 30m). Sin embargo, se tiene contemplado que el último satélite que se tiene proyectado lanzar cuenta con 2 sensores nuevos que son el ETM, instrumento formador de imágenes semejantes al TM pero con 8 bandas y una resolución de 15m; y el MLA, barredor multiespectral de captura a lo largo de la dirección de desplazamiento del satélite y que contará con 8 bandas y una resolución máxima de 10 metros en la porción visible.

Los satélites Landsat se desplazan en una órbita casi polar y heliosíncrona, esto permitió a los Landsat 1, 2 y 3 obtener datos de una misma región de la tierra cada 18 días; y actualmente a los Landsat 4 y 5 cada 16; en condiciones semejantes de iluminación.

Los datos colectados son transmitidos vía microondas a las estaciones terrenas de recepción, donde son grabados y procesados. Los datos son ofrecidos a los usuarios como imágenes multiespectrales en formato de fotografías, en cintas magnéticas compatibles para computadora. En discos flexibles y actualmente en CD-ROM.

El MSS es un sistema formador de imagen cuyo barrido es perpendicular a la trayectoria y es obtenido a través de un espejo oscilante, con un campo de vista (ángulo de vista) de 11.6°. El sistema permite la obtención de líneas de barrido de aproximadamente 185Km.

La figura 1.2 muestra la estructura de captura de la imagen en el MSS, incluyendo el posicionamiento de los detectores con relación a la órbita del satélite y su área de captación en el suelo.

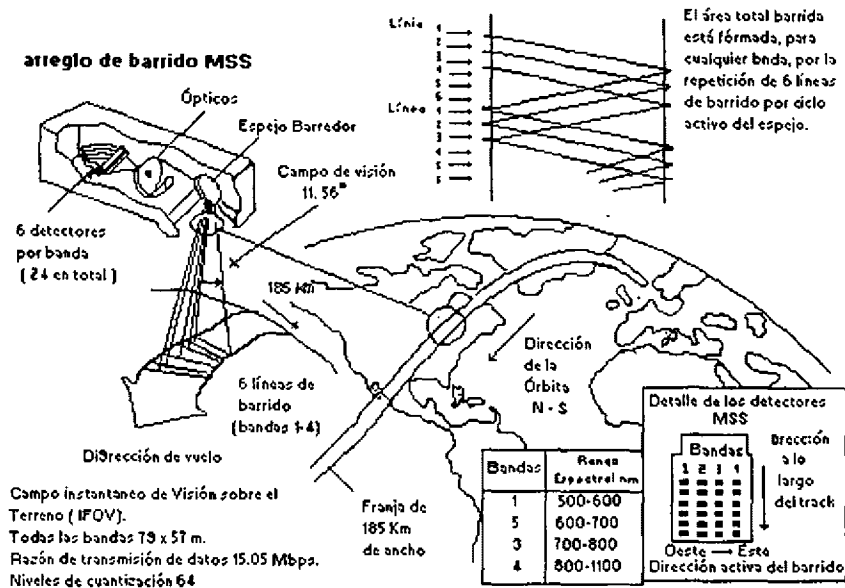


Figura 1.2 Estructura de captación de la imagen para MSS (tomada de Sorrenti, 1989).

La radiación proveniente de la escena, después de la reflexión en el espejo oscilante, es focalizado por el sistema óptico sobre una matriz de 24 terminales de fibra óptica de manera que cada terminal recibe el flujo de radiaciones en un área de 79m x 79m en la superficie.

Conducido por las fibras ópticas, el flujo de radiaciones es entonces descompuesto por medio de filtros en 4 conjunto, de las siguientes bandas:

- Banda 4: de 0.5 a 0.6 μm
- Banda 5: de 0.6 a 0.7 μm
- Banda 6: de 0.7 a 0.8 μm
- Banda 7: de 0.8 a 1.1 μm
- Banda 8: de 10.4 a 12.6 μm (únicamente Landsat 3).

Al alcanzar los detectores (6 por banda), la radiación es convertida en una señal eléctrica proporcional a la radiación de la superficie incluida en el elemento de resolución del terreno correspondiente. La señal que puede variar de 0 a 4Volts, es entonces cuantizada en 63 niveles posibles, siendo enseguida transmitida a la tierra o, eventualmente, grabadas para transmisiones posteriores.

El arreglo de la matriz de detectores (6 por banda) permite que, durante la oscilación del espejo 6 líneas sean barridas del oeste hacia el este una vez cada 9.95 μs . Como durante este intervalo el Landsat se desplaza apenas 56m en el sentido de barrido; a cada elemento de resolución de la imagen o pixel se debe asociar un área de 56m x 79m, en la superficie.

Generalmente una imagen MSS está formada por 2,983 líneas y 3,548 píxeles, cubriendo un área de 185Km x 185Km, aproximadamente, sobre la superficie.

El mapeador temático (Thematic mapper) es un instrumento sensor (captador de imágenes) proyectado para conseguir mejor resolución (30m. x 30m. máxima sobre la superficie), separación espectral (7 bandas) y fidelidad geométrica precisa, en comparación con el MSS. Es un sistema formador de imagen de barrido perpendicular a la trayectoria y la captación de la imagen es hecho a través del espejo oscilante.

Los datos del TM son adquiridos simultáneamente en 7 bandas espectrales que son:

- Banda 1: de 0.45 a 0.52 μm
- Banda 2: de 0.52 a 0.60 μm
- Banda 3: de 0.63 a 0.69 μm
- Banda 4: de 0.76 a 0.90 μm
- Banda 5: de 1.55 a 1.75 μm
- Banda 6: de 10.4 a 12.5 μm
- Banda 7: de 2.08 a 2.35 μm

La imagen está formada por bandas de 6,120 x 6,120 píxeles cubriendo un área de 185Km x 185Km aproximadamente. El tamaño del elemento de resolución del terreno es 30m para las bandas 1,2,3,4,5 y 7 y de 120m. para la banda 6.

Para las bandas 1 a 5 y la 7, los sensores están en forma de una estructura de línea y columnas de 16 elementos, mientras que la banda 6 cuenta únicamente con 4 detectores. La energía captada por el sensor es convertida en señales eléctricas de bajo nivel y registrada como uno de los 256 niveles de grises posibles. La figura 1.3 proporciona una vista de los detectores en el plano focal principal y la figura 1.4 da la estructura general de captación de una imagen.

Los satélites Landsat se desplazan de norte a sur en una órbita geocéntrica, circular, quasi-polar y heliosíncrona (el ángulo sol-tierra-satélite permanece constante), lo que garantiza condiciones semejantes de iluminación a lo largo del año en el área observada.

Banda	Longitud De onda, nm	Separación IFOV		ejes apagados		Fracción de 0.24°
		In	In	In	DEG	
6	10.4 a 12.5	7	0.14211	0.41760	0.024024	0.03648
		34.75	0.14211	0.38905	0.23213	0.36748
5	1.55 a 1.75	26	0.10606	0.24727	0.14756	0.61430
				0.14120		
7	2.08 a 2.35					
LED		45	0.1836			
EJES ÓPTICOS						
4	0.76 a 0.90	25	0.1020	0	0.019452	0
				0.04238	0.02530	
3	0.63 a 0.69	25	0.1020	0.14439	0.08617	0.35905
2	0.52 a 0.60	25	0.1020	0.24638	0.14705	0.51201
1	0.45 a 0.52			0.34838	0.20793	0.56636
				0.35552	0.21219	

Figura 1.3 Vista de los sensores del TM proyectados en el plano focal principal (tomada de Sorrenti, 1989).

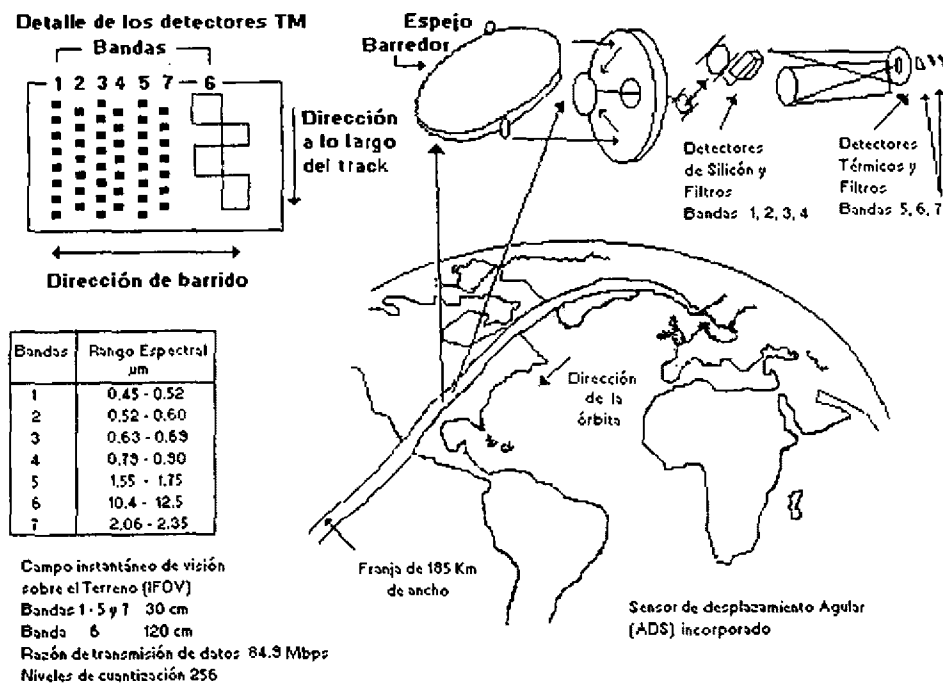


Figura 1.4 Visión general de la estructura de captación de una imagen para el sistema TM (tomada de Sorrenti, 1989).

Los satélites Landsat 1, 2, y 3 se localizan a una altitud de 920Km siguiendo la órbita ubicada a 99° con respecto al ecuador. Dichos satélites atraviesan el ecuador cada 103 minutos (9:30 de la mañana hora local), periodo durante el cual, debido al movimiento de rotación, el punto de cruce del satélite se desplaza 2752Km, que es la distancia entre las proyecciones de dos órbitas consecutivas. Al final de 24 horas, 14 órbitas son realizadas, las órbitas del día siguiente aparecerán desplazadas 160Km con relación a las correspondientes de ese día. Para el día 19 el desplazamiento con respecto al día 0 desaparece y comienza de nuevo el ciclo.

La órbita de los satélites Landsat 4 y 5 también es repetitiva y se encuentra a una altitud nominal de 705.3Km, formando un ángulo de 98.22°. Los satélites cruzan el ecuador de norte a sur a las 9:45 de la mañana en cada pasada. Cada órbita lleva aproximadamente 99 minutos para completarse, lo que permite un total de 14.5 órbitas por día, cubriendo la tierra entera en 16 días. Las figuras 1.5 y 1.6 muestran el paro de órbitas subsiguientes y órbitas adyacentes para los satélites Landsat 4 y 5.

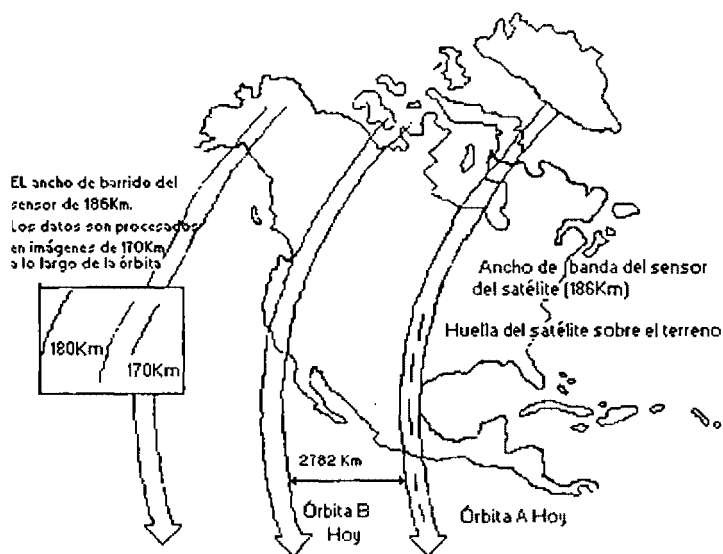


Figura 1.5 Órbita subsiguiente del Landsat 4 y 5 (tomada de Sorrenti, 1989).

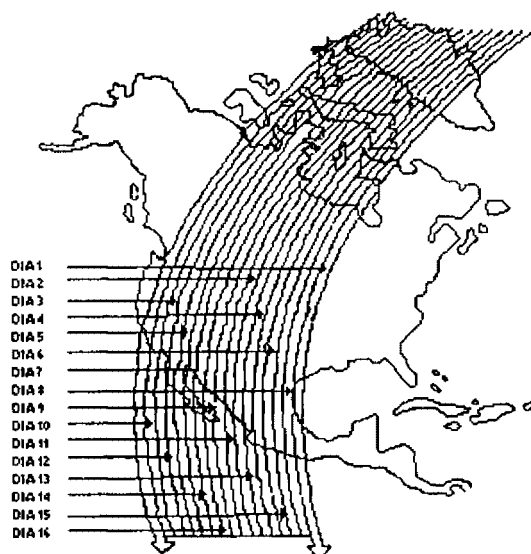


Figura 1.6 Órbitas adyacentes en 16 días consecutivos para Landsat 4 y 5 (tomada Sorrenti, 1989)

La sobreposición lateral de órbitas en el ecuador es de 7.3%, al moverse hacia el norte o hacia el sur, el porcentaje de sobre posición aumenta.

El sistema mundial de referencia (WRS, figura 1.7) es una llave de localización global para las imágenes Landsat. Está basado en un sistema de columnas y líneas cuyo cruce determinan la localización nominal del centro de cada escena. Este sistema proporciona a los usuarios la correlación con las latitudes y longitudes y la facilidad de identificar cualquier posición sobre la superficie terrestre.

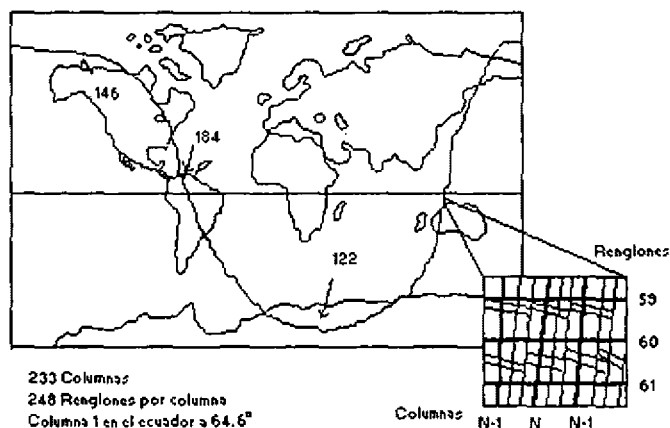


Figura 1.7 Sistema mundial de referencia (WRS) para el Landsat 4 y 5 (tomada de Sorrenti, 1989)

El esquema del WRS refleja la naturaleza repetitiva de las órbitas del Landsat y la manera como las imágenes son encuadradas después de la adquisición.

En el sistema WRS los números de las órbitas (columnas) están correlacionados con los centros nominales de cada órbita (figura 1.7). El número de líneas identifica la posición de la escena en el sentido vertical y son numerados del 1 al 124 en el sentido descendente del satélite y de 125 a 248 en el sentido ascendente (figura 1.7).

La hora solar media de cada línea individual en una órbita se mantiene fija, por lo tanto los centros para una línea de una determinada latitud tiene una misma hora solar. La figura 1.8 muestra las variaciones de las horas locales de los cruzamientos de las órbitas con las latitudes.

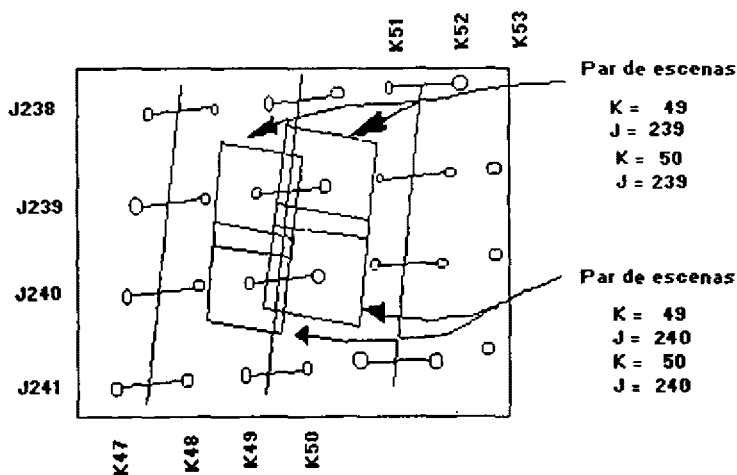


Figura 1.8 Variación de las horas locales con los cruzamientos de las órbitas con las latitudes (tomada de Sorrenti, 1989).

1.2.2 SISTEMA SPOT

SPOT es una fuente permanente de datos para información geográfica y ofrece características únicas en el campo de la percepción remota espacial: resolución terrestre de 10 a 20 metros por pixel, posibilidades flexibles de adquisición de datos para cualquier punto sobre la superficie de la Tierra, posibilidad de visión estereoscópica y una excelente aproximación geométrica.

El programa de satélite SPOT, (del francés: System Probatoire d'Observation de la Terre, Sistema de pruebas de Observación de la tierra) es un conjunto de adquisición y distribución de imágenes por satélite de vocación civil. Ha sido desarrollado bajo la responsabilidad del Centro Nacional de Estudios Espaciales CNES en Francia con la participación de Suecia y Bélgica. Consta de 3 satélites que poseen las mismas características. El primer satélite Spot fue puesto en órbita el 22 de febrero de 1986, el segundo en febrero de 1990 y el tercero fue lanzado en septiembre de 1993.

Los satélites Spot han sido colocados a una altitud media de 830Km, en una órbita circular quasi polar inclinada 8 grados con respecto al norte geográfico. Los satélites Spot dan una vuelta a la tierra en 1001 minutos, lo cual combinado con el movimiento de la rotación de la tierra hace que la traza en el suelo de 2 órbitas consecutivas aparezca desplazada en el ecuador 2,823Km (figura 1.9). Cada uno de los satélites Spot vuelve a pasar en la vertical de un mismo punto al cabo de 3,229 revoluciones, ósea, cada 26 días.

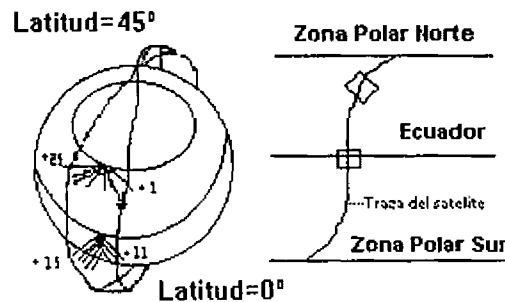


Figura 1.9 Órbita subsecuentes del satélite Spot

Los satélites Spot están dotados de 2 instrumentos idénticos, el HRV1 y HRV2 (Haute Résolution Visible) capaces de funcionar de forma independiente. Cada instrumento barre una banda de dimensión, en sentido este-oeste, de 60Km en mira vertical y puede alcanzar los 80Km en visión oblicua.

A lo largo de la traza del satélite se dividen las escenas cada 60Km. Así pues, la escena Spot cubre una superficie de 60Km x 80Km en visión oblicua (figura 1.10).

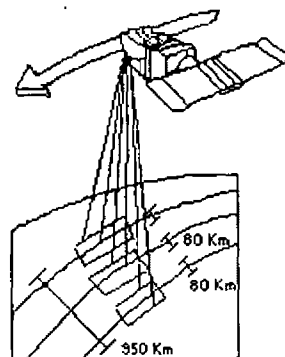


Figura 1.10 Cobertura de la mira vertical y oblicua del Spot (tomada de Spot image, 1989).

Ambos instrumentos cubren áreas distintas cuando realizan una mira vertical. Sin embargo, cuando los dos sensores funcionan al mismo tiempo en una toma de la mira vertical, las escenas adquiridas por ambos instrumentos sufren un traslape de 3Km pudiendo integrarse una sola franja de 117Km (figura 1.11). La orientación de las escenas con respecto al norte varía en función de la latitud y del ángulo de toma de vista.

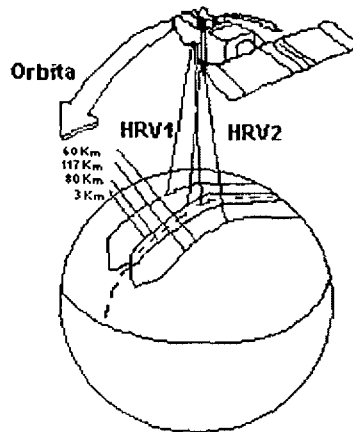


Figura 1.11 Vista vertical de los dos instrumentos HRV del Spot (tomada de CNES, 1968).

Los instrumentos poseen dos arreglos de detectores, el primero de ellos consta de 3,000 sensores y se emplean para la captura de imágenes en el modo multiespectral, mientras el segundo tiene 6,000 sensores que permite la adquisición de imágenes pancromáticas. Por tanto, cada línea de barrido comprende 3,000 ERT's o píxeles para cada banda del modo multiespectral y 6,000 en el modo pancromático.

Una escena Spot consta de 3,000 líneas de 3,000 ERT's de 220m en modo multibanda y de 6,000 líneas de 6,000 píxeles de 10m en modo pancromático. En mira oblicua, hasta 27 grados hacia el este o hacia el oeste, la dimensión en el suelo de una línea varía entre 60Km (mira vertical) y 80Km (mira oblicua extrema), siendo constante el número de detectores. El paso de muestreo varía en consecuencia entre 10m y 13.5m o entre 20m y 27m según el modo espectral. A continuación las líneas se someten a un remuestreo a una distancia fija de 10m o 20m, con lo cual aumenta el número de píxeles en la línea. El descentrado observado entre cada línea se debe, entre otro, a la rotación de la tierra entre los registros de dos líneas consecutivas.

Dada su órbita quasi polar y heliosíncrona, Spot puede adquirir imágenes hasta una altitud de 84 grados norte-sur.

La elección de la órbita se hizo con el objeto de asegurar una sincronización perfecta con el sol (heliosíncrona).

Las condiciones de iluminación son por lo tanto idénticas para todas las escenas de una misma latitud en determinada época. Si a lo largo del año varían el ¹acimut y la elevación del sol, el ángulo entre la dirección solar, el centro de la tierra y el plano de la órbita permanece constante.

Los satélites Spot capturan la información mediante dos modos distintos de observación: el modo pancromático "P" y el modo multibanda o multiespectral "XS". Ambos modos funcionan indiferentemente en uno u otro de los instrumentos y de manera simultánea o individual. Sin embargo, debido a que el satélite posee únicamente dos canales para la transmisión de datos no pueden funcionar las 4 posibilidades a un mismo tiempo. Las combinaciones posibles son mostradas en la tabla 1.2.

HRV1	XS	*		*		*	
HVR1	P		*		*		*
HVR2	XS	*			*		*
HVR2	P		*	*			*

Tabla 1.2 Modos posibles de funcionamiento simultáneo de los instrumentos sensores HRV1 y HRV2.

Cuando un instrumento captura una escena, al mismo tiempo, tanto en modo pancromático como multiespectral, las dos imágenes generadas (XS y P) no capturan perfectamente debido a la disposición de los sensores dentro del instrumento.

Las longitudes de onda para los diferentes modos son:

- ◆ Banda pancromática: de 0.51 a 0.73 μm
- ◆ Banda XS1: de 0.50 a 0.59 μm
- ◆ Banda XS2: de 0.61 a 0.68 μm
- ◆ Banda XS3: de 0.79 a 0.89 μm

Como se nota, para el modo pancromático la observación se realiza en una banda espectral única, correspondiente a la parte visible del espectro sin el azul. Dicha toma de vista, efectuada en un solo canal, proporciona imágenes en blanco y negro. El tamaño del pixel es de 10m.

En el modo multibanda se obtiene imágenes para las porciones verde, rojo e infrarrojo cercano al espectro. La combinación de las 3 bandas posibilita la consecución de composiciones a color. El tamaño del pixel es de 20m.

Los dos instrumentos HRV del Spot están dotados de espejos orientables que permiten la visión oblicua, hacia el este o hacia el oeste de la traza, con un límite de más o menos 27° de la vertical. Para poder tener una mira oblicua, los espejos orientables se desplazan de 0.6° en 0.6°, lo que permite captar hasta 91 escenas o posiciones diferentes.

Tal inédita característica posibilita, por un lado, la visión estereoscópica, y por el otro, la observación repetida de una misma zona durante un ciclo de 26 días (figura 1.12 y 1.13).

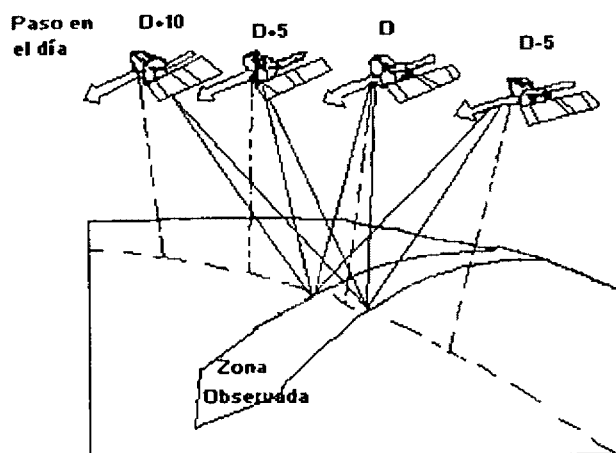


Figura 1.12 Repetitividad de las observaciones durante un ciclo de 26 días, para el satélite Spot (tomada de CNES, 1988)

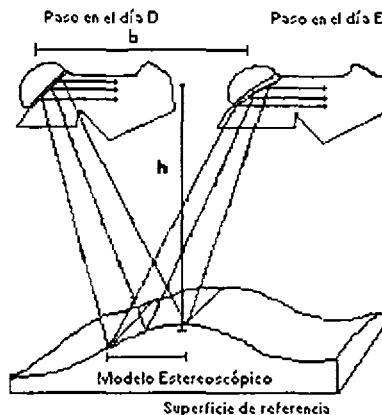


Figura 1.13 Observación estereoscópica debida al paralaje creada por dos miras con diferente ángulo (tomada de CNES, 1988)

Un desplazamiento del ángulo de vista de 24° al oeste y de 24° al este permite obtener una relación base/altura (b/h) de 1. Para un estereoscópico, compuesto por una vista vertical y otra realizada desde un ángulo de 27° , el valor de la relación b/h es de 0.5.

Gracias a la mira oblicua, resulta posible adquirir escenas dentro de una banda de 950Km, paralela a la dirección de desplazamiento del satélite y centrada en su traza como se mostró en la figura 1.10 Dicha técnica permite aumentar la frecuencia de observación de un mismo sitio durante un ciclo. Tal frecuencia varía en función de la latitud: en el ecuador se puede observar la misma región 7 veces durante los 26 días del ciclo orbital, lo que da una media de 3.7 días entre dos observaciones. Para una latitud de 45° , una misma región se puede observar 11 veces durante el ciclo orbital, dando una media de 2.4 días.

La serie Spot presenta también su propio sistema de referencia mundial para la ubicación geográfica de sus imágenes, denominados GRS (Grade Reference Spot, figura 1.14). La GRS es un sistema mediante el cual se atribuye a cada escena dos marcas de referencia K y J, correspondientes a columnas y líneas. Las columnas quedan orientadas paralelamente a la dirección de desplazamiento del satélite y las líneas son paralelas a las latitudes. La intersección de una columna con una línea determina un nodo. El centro de cada escena Spot está relacionada con un nodo GRS.

Una escena en la GRS cubre un área de 60Km x 60Km. Si una escena cubre un área mayor y su centro se encuentra desplazado con respecto a cualquier nodo de la GRS, debido a que fue tomada mediante una mira oblicua, se le incorpora al nodo más cercano.

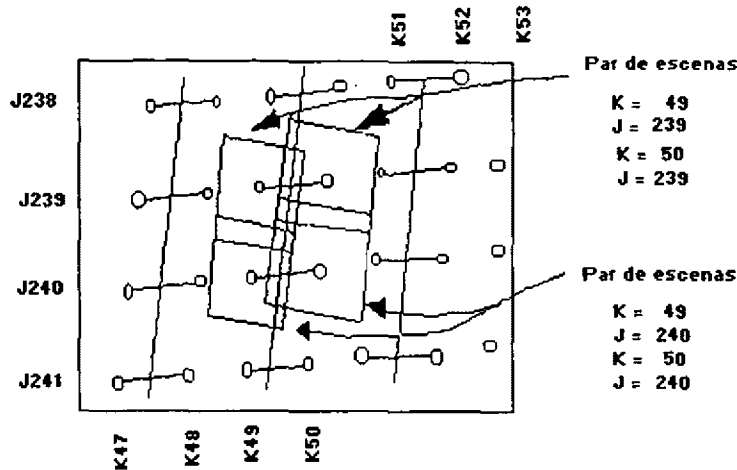


Figura 1.14 Sistema de referencia mundial (GRS) para el satélite Spot (tomada De Spot Image, 1986)

Una escena puede estar desplazada hasta 30Km (para la derecha o la izquierda de la K nominal), pero no existirá desplazamiento en el sentido norte-sur, esto es, el centro de la imagen siempre caerá en una J del sistema GRS.

1.2.3 ESCALA DE GRIS

Una escala de gris es una secuencia de tonos en la cual se corre en forma discreta el continuo que va desde el negro o ausencia de luz hasta el blanco o luz más intensa. En la práctica a cada tono se le asigna un número de acuerdo a una función.

El número de niveles de gris puede ser arbitrario dependiendo de la cantidad de tonos que desee distinguir. En un sensor siempre es un múltiplo de 2 dado que las computadoras trabajan con sistemas binarios.

Un sensor remoto tiene la ventaja sobre el ojo humano de que no sólo permite la digitización de un objeto en forma extremadamente rápida, sino en el contenido de información que puede discriminar, dado que el ojo humano sólo puede distinguir del orden de 64 tonos de gris y observar en la región

El procesamiento de imágenes es una rama de la teoría de graficación por computadora en la que típicamente manejan las imágenes como arreglos o series de elementos. El número de elementos en un arreglo, determina la resolución espacial de la imagen y el número de bits accesible a cualquier elemento del arreglo, determina el número de "colores" o escala de valores de gris que puede tener cada elemento.

La operación de un sistema de procesamiento digital de imágenes de propósito general puede dividirse en tres categorías principalmente: la digitización, el procesamiento y el despliegue. En este último concentraremos nuestra atención.

El proceso de despliegue permite que una imagen digital, pueda ser analizada e interpretada visualmente con propósitos de análisis digital. El sistema de despliegue produce una imagen en la cual la brillantez o densidad de cada pixel desplegado se controla directamente usando el nivel de gris correspondiente a ese pixel en la imagen digital.

Las características que determinan la calidad de un sistema de despliegue de imagen digital son las siguientes: manejo de la imagen, resolución fotométrica, resolución espacial, respuesta a bajas frecuencias y ruido.

En sistemas de despliegue, la resolución fotométrica se refiere a la precisión con la cual el sistema produce el valor de brillantez o densidad correcta en cada pixel. La resolución espacial es el número de niveles descritos de gris que el sistema puede generar. Esto depende, parcialmente, del número de bits utilizados para controlar la brillantez de cada pixel. Con 8 bits se generan 256 niveles que van desde el blanco hasta el negro, pasando por una escala de gris con una alta resolución. La resolución de una imagen depende en gran parte a los niveles de la escala de gris. El número efectivo de niveles de gris, nunca es más que el número de niveles de gris de los datos digitales.

Debe aclararse que el término imagen se refiere a una función de dos dimensiones luz-intensidad, denotadas por $f(x,y)$, en donde el valor o amplitud de f esta dado en coordenadas espaciales. (x,y) da la intensidad (brillo) de la imagen en un punto.

Esto nos lleva a introducir el termino de Luminancia: Desde que la luz es una forma de energía, $f(x,y)$ debe ser finita y no ser cero, esto es;

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

La naturaleza básica de $f(x,y)$ puede ser considerada como su esencia, caracterizada por dos componentes. Un componente es la suma del origen de la luz, incidida sobre la escena vista, mientras que el otro es la suma de la luz reflejada por el objeto en la escena. Estos componentes son llamados iluminación y reflejo de los componentes, y son denotados por:

$i(x,y)$ y $r(x,y)$, respectivamente.

Las funciones $i(x,y)$ y $r(x,y)$ combinadas como un producto forman $f(x,y)$:

$$f(x,y) = i(x,y) r(x,y) \tag{1.1}$$

donde $0 < i(x,y) < \infty$ (1.2)

y $0 < r(x,y) < 1$ (1.3)

La ecuación 1.3 indica el factor de reflejancia, el cual es limitado por 0 (total absorción) y por 1 (total reflejancia). La naturaleza de $i(x,y)$ es determinada por el origen de la luz, mientras que $r(x,y)$ es determinada por las características del objeto en la escena.

La intensidad de una imagen monocromática f en coordenadas (x,y) es llamada nivel de gris (I) en ese punto. Desde la ecuación 1.1 a la 1.3 es falso en el rango:

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

En teoría, se requiere que L_{min} sea positivo, y L_{max} sea finito. En la práctica, $L_{min} = (i_{min}) (r_{min})$ y $L_{max} = (i_{max}) (r_{max})$. El intervalo $[L_{min}, L_{max}]$ es llamado escala de gris. En la práctica es común variar estos intervalos numéricos a intervalos $[0,L]$, donde $I = 0$ es considerado como negro y $I = L$ es considerado como blanco en esta escala. Todos los valores intermedios son considerados matices de gris variando continuamente desde el negro al blanco.

Sin embargo, una imagen monocromática contiene escalas de gris equivalentes a una imagen de color, estas escalas de gris dependen de la señal de Luminancia (Y) cuya proporción se rige por la siguiente ecuación de porcentaje:

$$Y = 0.30R + 0.59V + 0.11A$$

donde R = rojo, V = verde y A=azul

Por ejemplo, para una señal de barras de color (patrón) los porcentajes son los siguientes:

$$\text{BLANCO} \quad Y = 0.30R + 0.59V + 0.11A = 1.00$$

Es decir, el color blanco contiene los tres colores primarios.

Para el caso de la señal referida la banda blanca se encuentra en el nivel más bajo de voltaje, esto es porque las señales de vídeo tienen la característica de transmisiones negativas, es decir, que el color negro se ubica en el bajo. En la figura 1.15 se muestran las 8 bandas de color para la señal y en porcentaje de la escala de gris. Los porcentajes para las bandas de color resultan:

BLANCO	Y=1.00
AMARILLO	Y=0.89
CYAN	Y=0.70
VERDE	Y=0.59
MAGENTA	Y=0.41
ROJO	Y=0.30
AZUL	Y=0.11
NEGRO	Y=0.00

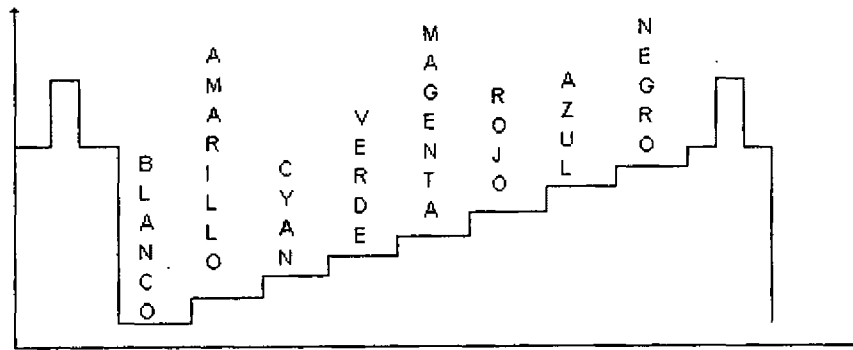


Figura 1.15 Pulso de una señal de vídeo (Y) monocromática para un patrón de barras de color.

Las imágenes monocromáticas tienen un espectro en frecuencia menos denso, es decir, se tienen un ancho de banda menor a los 64MHz, esto implica una frecuencia de muestreo más baja. El inconveniente de esto, es en la resolución de la imagen, puesto que solo contendrá la escala de gris desde el blanco hasta el negro, omitiéndose los detalles de color.

Este tipo de codificación es llamado codificación por componente, el cual muestra solamente la señal de luminancia y procesa sólo esa señal sin tomar en cuenta los detalles de color. Este tipo de codificación ahorra velocidad en el muestreador, también ahorra el número de bits por muestra puesto que en estas señales solo se codifica la escala de gris por el pixel, aumentando con esto el contenido del archivo de datos de imagen considerablemente de la señal Y.

En el caso de las señales monocromáticas se tendrá una gráfica de intensidad contra número de pixeles para el caso de escalas de gris en intensidad luminosa.

Si el ruido electrónico dentro del sistema de despliegue ocupa más de un nivel de gris, entonces el número efectivo de niveles de gris se reduce. Como regla general, el nivel de ruido RMS representa un límite suficientemente bajo para la resolución en la escala de gris. Se utiliza el nivel de ruido RMS ya que si el ruido tiene una distribución normal, entonces tendrá aproximadamente una desviación estándar de ± 1 el 68% del tiempo.

Otra característica importante de los sistemas de despliegue es la linealidad de la escala de gris. Es decir, el grado en el que la brillantez o densidad es proporcional al nivel de gris de entrada. Cualquier dispositivo de despliegue tiene un nivel de gris de entrada que produce una curva de transferencia de densidad o brillantez. Para su operación adecuada, esta curva debe ser razonablemente lineal y constante. No obstante, el usuario puede alterar la curva de transferencia para adecuar, a su gusto, la imagen desplegada. Sin embargo, el proceso debe de ser hecho en la computadora, y no en el dispositivo de despliegue, el cual debe de presentar los datos al operador sin ningún "realce".

Finalmente, la última característica importante en un sistema de despliegue es el ruido. El ruido electrónico en un sistema de despliegue produce variaciones tanto en la intensidad como en la posición del punto. El ruido aleatorio en el canal puede producir un efecto de sal y pimienta, particularmente visible en los campos planos. El nivel efectivo de cuantización es aproximadamente igual a la amplitud del ruido RMS. Si el ruido es periódico y de intensidad razonablemente alta, puede producir una muestra sobrepuesta en la imagen desplegada. Si el ruido es periódico y sincrónico con deflexión en la señal vertical y horizontal, puede producir un patrón de barras. La calidad general de un sistema de despliegue es adecuada si todo el ruido, aleatorio y periódico, permanece abajo de un nivel de gris de amplitud.

1.2.4 MODELO RGB

El ojo humano percibe el color a través de tres pigmentos visuales en los conos de la retina. Estos pigmentos tienen una sensibilidad pico en longitud de onda cercanas al 630nm (rojo), 530nm (verde) y 450nm (azul). Cuando dichos conos son sensibilizados, el cerebro mezcla las señales registradas por cada uno de los pigmentos y percibe el color de la luz.

Esta teoría de la visión de tres estímulos es la base de la fotografía a color y de la mayoría de los monitores de vídeo. Esta técnica hace uso de los 3 colores primarios rojo, verde y azul, por lo que se llama modelo RGB (debido a sus siglas en inglés: red, green, blue).

Este modelo se representa mediante el cubo unitario definido por los ejes R, G y B, como se muestra en la siguiente figura 1.16. El origen representa el negro y el vértice con coordenadas (1,1,1) es el blanco. Los vértices del cubo sobre los ejes determinan los colores primarios y los vértices restantes representan los colores complementarios de cada uno de ellos.

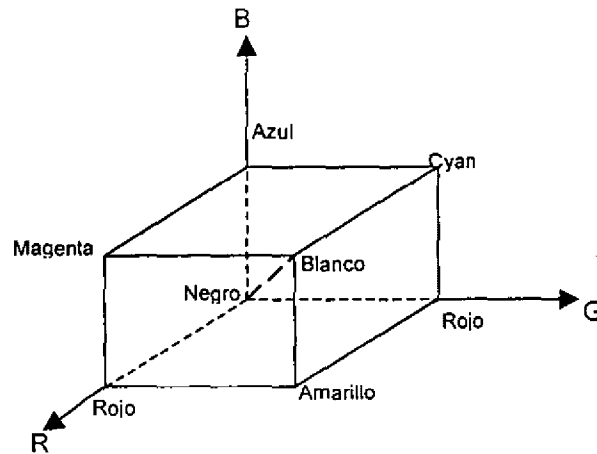


Figura 1.16 Cubo de colores del modelo RGB.

Este esquema de color es un modelo aditivo, esto es, las intensidades de los primarios se suman para producir otros colores.

Cada punto de color contenidos en los límites del cubo puede representarse como la triada (R, G, B) donde los valores R, G, B se asignan en el intervalo de 0 a 1. El vértice magenta se obtiene sumando el rojo y el azul para producir la triada $(1,0,1)$ y el blanco en $(1,1,1)$ es la suma de los vértices rojo, verde, azul. Las sombras de gris se representan a lo largo de la diagonal principal del cubo, del origen (negro) al vértice blanco. Cada punto situado en esta diagonal tiene una contribución igual de cada color primario, de manera que una sombra de gris en medio del negro y el blanco se representa como $(0.5,0.5,0.5)$

Los colores pueden ser obtenidos de dos maneras, por síntesis aditiva o por análisis substractivo (ver figura 1.17). La primera se basa en el principio de reconstrucción de la luz blanca por proyección en la obscuridad sobre una pantalla negra, de tres haces luminosos de los colores primarios.

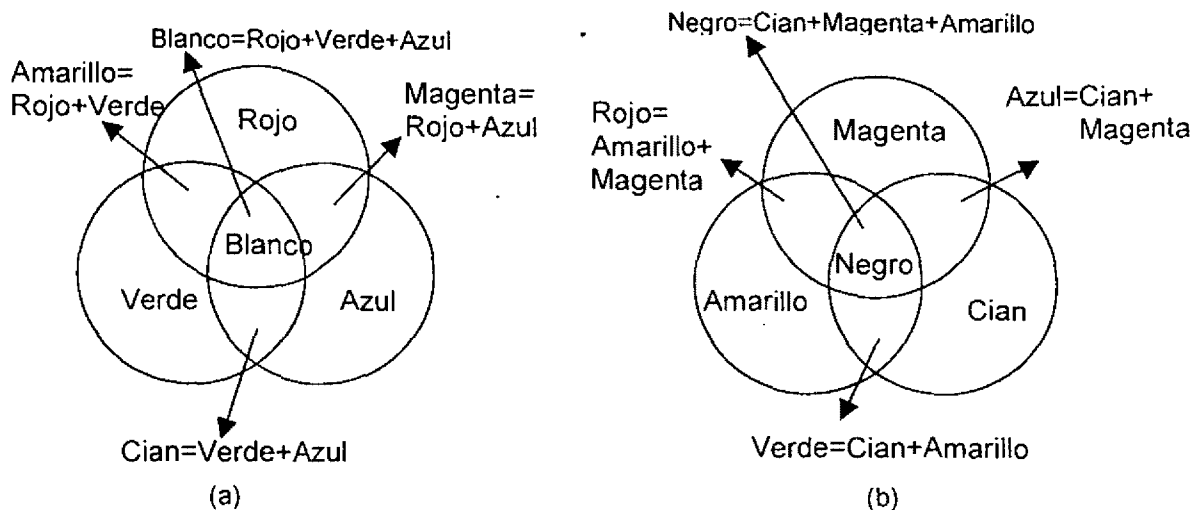


Figura 1.17 Sistema de combinación de color. (a) Aditivo. (b) Substractivo

Generalmente, un color cualquiera se conseguirá sobre la pantalla por la modulación de la intensidad de los tres colores. Este método es el que se utiliza en los televisores a color, en donde en un extremo existen tres "cañones", uno para cada color primario y así los colores en la pantalla se generan según la intensidad de cada cañón del cinescopio. El principio del análisis substractivo consiste en que el haz luminoso de una luz blanca cruza sucesivamente varios filtros coloreados en amarillo, cian y magenta quienes interceptan su color complementario, es decir, el filtro amarillo impide el paso del azul, rojo etc. Si los tres filtros son saturados, los tres componentes primarios de la luz blanca están totalmente absorbidos, quedando únicamente el negro a la salida de los tres filtros. Como en el caso de la síntesis aditiva, el nivel de saturación de los tres filtros permite la conservación de los componentes deseados de la luz blanca y obtener así a la salida la señal de un color cualquiera.

Para efectuar un análisis numérico del color de un objeto es necesario conocer las coordenadas de dicho color en un sistema de referencia, en este caso el azul, verde, rojo (A,V,R). Sea un color cualquiera C definido por sus coordenadas (a, v, r) . Se supone que estas tres cantidades pueden tomar valores enteros entre 0 y 255, cifrando 8 bits a cada valor. Por ejemplo, tómese el color $C(70,22,15)$. El análisis de C nos lleva a lo siguiente la proporción del color verde es la más importante, siendo esta 227/255. Entonces el tono primario dominante es el verde. Como la proporción de azul no es despreciable (156/25), este verde tendrá un poco de cian, entonces se puede definir a C como un verde cian.

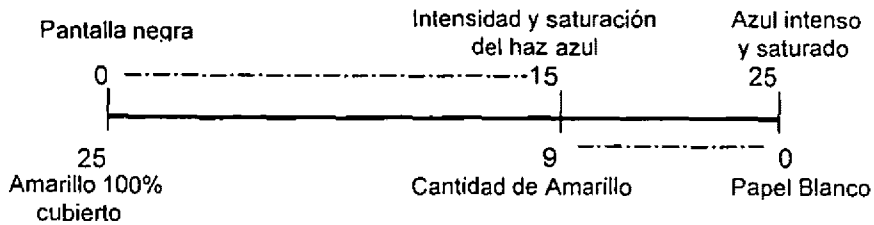
Por otro lado, los porcentajes de colores primarios son importantes, lo que nos permite decir que el color será bastante claro. Su descomposición numérica es la siguiente:

$$C = (156,227,70) = (70,70,70)+(86,157,0)$$

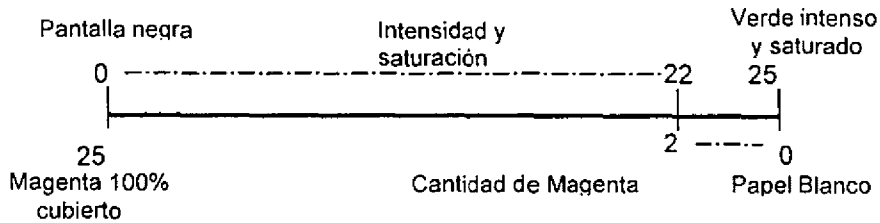
con $(70,70,70)$ que corresponde al gris o cantidad de blanco que aclara el color (colores primarios en igual intensidad dan gris),

$$C' = (86,157,0) = (86,86,0)+(0,71,0)$$

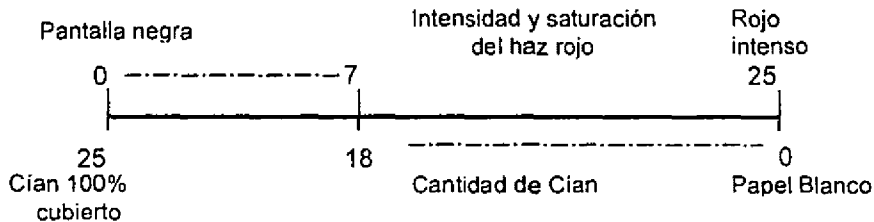
Donde $(86,86,0)$ es un cian puro (azul y verde en igual intensidad sin rojo) y se aprecia que la tonalidad dominante es verde en un $(71/255)*100\%$. Como los colores secundarios están perfectamente complementados por los primarios, un color de coordenadas (a,v,r) en el sistema (A,V,R), tendrá coordenadas $(255-a,255-v,255-r)$ en el sistema referencial (A,M,C) Amarillo, Magenta, Cian. En el ejemplo, el color C tendrá coordenadas $(99,28,185)$ en el sistema (A,M,C). Se puede representar la correspondencia entre la síntesis aditiva sobre una pantalla negra y el análisis substractivo sobre papel blanco de la siguiente forma gráfica y tomando como ejemplo al color C (Ver figuras 1.18, 1.19 y 1.20).



1.18 Relación entre síntesis aditiva y análisis sustractivo para el azul



1.19 Relación entre síntesis aditiva y análisis sustractivo para el verde



1.20 Relación entre síntesis aditiva y análisis sustractivo para el rojo

El procesamiento de imágenes en blanco y negro (B & N) utiliza el mismo principio que se utiliza en el de a color con la única diferencia que en esta se utiliza un color y no una intensidad de gris, a cada valor numérico.

Para obtener un procesamiento de este tipo es, ante todo, necesario disponer de una paleta ordenada de colores. La mayoría del software utiliza una tabla de colores llamada arco iris porque incluye todos los colores del espectro en orden creciente de longitud de onda: violeta, índigo, azul, verde, amarillo, naranja y rojo. Pero puede definirse una tabla de colores propia siguiendo cualquier orden con los colores deseados. El problema será entonces la interpretación; deberá mantenerse a la vista una escala de colores como referencia para ordenar y luego analizar la información.

Los tipos de procesamiento descritos anteriormente se aplican en la misma forma que para la visualización en B & N. La escala de gris utilizada en ella es evidente en su definición y en el orden adoptado. Se puede interpretar instantáneamente al ojo y no es necesaria una leyenda para comprenderla. Al contrario, la interpretación de un canal visualizado en una tabla de color es precisa cierta experiencia y a menudo la visualización de la escala de colores para poder interpretar algunos fenómenos. Existen, además, un paso obligatorio por colores más oscuros o más claros, de acuerdo a los matices, lo que no facilita la interpretación.

La visualización simultánea de varias bandas es posible en síntesis aditiva de colores, basta proceder como en el caso de la visualización en B&N escogiendo en primer lugar 3 bandas de la misma imagen y asociándolas a cada uno de los tres colores primarios: azul, verde y rojo. El número máximo de bandas visualizable es, desde luego, tres. En el caso de SPOT, las tres bandas están predefinidas. Falta por asignar los colores a ellas. Para elaborar una COMPOSICION EN COLOR (CC) de una imagen SPOT multiespectral, se asigna en general el azul a la banda XS1, el verde a la banda XS2 y el rojo a la banda XS3. Se visualizan las tres bandas en síntesis aditiva:

- ◆ La banda XS1 en niveles de azul (que van del negro al azul intenso pasando por una gama ordenada de azules).
- ◆ La banda XS2 en niveles de verde.
- ◆ La banda XS3 en niveles de rojo.

La visualización de las tres bandas en bruto en una síntesis aditiva en la pantalla no es evidentemente la óptima; debe adaptarse la dinámica de cada una de ellas. Desde luego que es posible efectuar cualquier tipo de adaptación dinámica en cada una de las bandas que integran la composición en colores.

En numerosos sensores remotos se encuentran las tres bandas que corresponden a los sectores de longitudes de onda. Una composición en colores clásica consiste entonces en "desplazar" estos sectores de longitudes de onda y asociarles los colores primarios en el orden de clasificación correspondiente a su longitud de onda (azul, verde, rojo). Se trata de hecho de un simple deslizamiento en un lugar del espectro de colores, respetando el orden lógico de las bandas:

SECTOR DE LONGITUD DE ONDA	COLOR ASIGNADO
verde	azul
rojo	verde
infrarrojo cercano	rojo

A este tipo de COMPOSICIÓN EN COLORES se le llama "COMPOSICIÓN EN FALSO COLOR"

Desde luego, es posible cualquier otra asignación de colores. En el caso de una imagen SPOT multiespectral, la selección de bandas es impuesta dado que hay solo tres, pero se pueden de todos modos permutar los colores, afectando por ejemplo XS1, XS2, XS3 a los colores respectivos, rojo, azul y verde. Falta todavía interpretar visualmente los colores de una composición de este tipo, es teóricamente posible construir $3! = 6$ composiciones en colores diferentes con una imagen SPOT XS. Para una imagen Landsat TM (de 6 bandas), existe 120 CC posibles ($6!/3!$), pero muchas de ellas son semejantes.

Dado que la composición en colores se efectúa mediante síntesis aditiva de los colores, es relativamente fácil descomponer sus diferentes elementos en "cantidades" relativas de colores básicos azul, verde y rojo, y remontar así hasta una estimación relativa de la respuesta de cada elemento en las diferentes bandas que forman la CC. Se llega así a recrear un símil de sello espectral y a proponer una hipótesis acerca del carácter del elemento estudiado.

En caso de una composición en falso color de una imagen SPOT XS, una parcela de color rojo vivo corresponde a una superficie con baja reflectancia en las bandas XS1 y XS2 (a las cuales han sido asignados los respectivos colores azul y verde) y una fuerte reflectancia en la banda XS3 (a la cual ha asociado el rojo). Conociendo el sello espectral típico de la vegetación, esta parcela debe corresponder a una cubierta vegetal con actividad clorofílica en la fecha de la toma de la imagen. Si, además, se conocen la región y el calendario agrícola, se puede deducir la posible ocupación del suelo.

Una imagen en falso color permite resaltar u observar rasgos de la superficie que a simple vista no pueden ser discriminados

1.2.5 FALSO COLOR

CONCEPTO DE COLOR

Cuando cada componente de frecuencia contenida en el intervalo del rojo a violeta contribuye más o menos en la misma cantidad de la energía total, el color se describe como blanco figura 1.21.

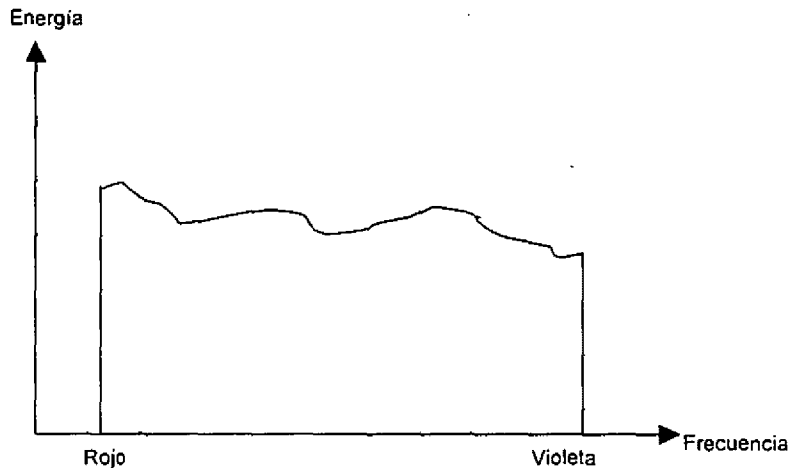


Figura 1.21 Distribución de energía para una fuente de luz blanca.

Cuando hay una frecuencia dominante figura 1.22 se dice que la luz tiene color correspondiente a dicha frecuencia. En la figura 1.22 la intensidad de la luz dominante es E_p y la contribución de las otras frecuencias produce luz blanca de intensidad E_w . La brillantez de la fuente se calcula como el área situada bajo la curva, lo que da la energía total emitida. La pureza depende de la diferencia entre E_p y E_w . Cuanto mayor sea la energía E_p de la frecuencia dominante en comparación con la componente de la luz blanca E_w , tanto más pura será la luz. Se tiene una pureza del 100% cuando $E_w = 0$ y una pureza del 0% cuando $E_w = E_p$.

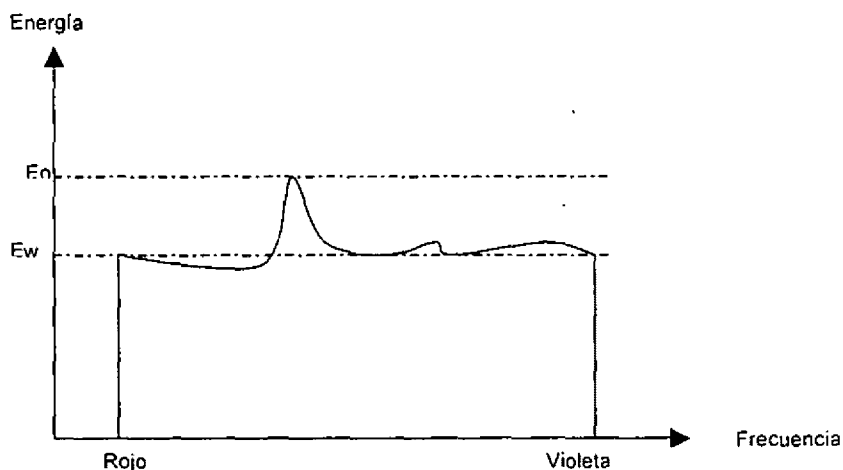


Figura 1.22 Distribución de energía de una fuente de luz con una frecuencia dominante.

Las características de la luz formada por la combinación de dos o más fuentes, son determinadas por ellas mismas. Dos fuentes de luz de diferente color, con intensidades elegidas adecuadamente pueden utilizarse para producir una gama de otros colores. Si las dos fuentes se combinan para producir luz blanca, éstos se conocen como colores complementarios.

Por lo general los modelos de color que se utilizan para describir combinaciones de luz hacen uso de tres colores para obtener una gama razonablemente amplia, llamada gama de colores de ese modelo. Los dos o tres colores para describir otros, en un modelo, se conocen como colores primarios. Desde esta forma la ecuación que nos permite generar un color a través de tres primarios es:

$$Color = IP_1 + IP_2 + IP_3$$

donde:

IP_1 = Intensidad del primer primario

IP_2 = Intensidad del segundo primario

IP_3 = Intensidad del tercer primario

No existe un conjunto único de colores primarios que describa todas las longitudes de ondas posibles contenidas en el espectro visible. Para ampliar la gama de colores deben agregarse más primarios, pero esto todavía requeriría un número infinito de ellos para incluir todos los colores posibles.

Tres primarios son suficientes para muchos fines y los colores que no están en la gama de un conjunto especificado, todavía puede describirse por medio de métodos extendidos. Por ejemplo, si un color no se obtiene de tres primarios dados, se pueden combinar uno o dos de estos con ese color para obtener un ajuste con los primarios restantes. En este sentido amplio, puede considerarse que un conjunto de colores primarios describe todos los colores (Hearn and Baker, 1988; Rogers, 1986).

Una norma internacional para los colores primarios se estableció en 1931 por la Comisión Internacional de iluminación ("Commission Internationale de l'ECLAIRAGE, también llamada CIE"). La finalidad de esta norma es la de permitir que todos los colores se definan como la suma de tres colores primarios.

$$Color = X + Y + Z$$

Puesto que no hay tres colores en el espectro visible que puedan lograr esto, los colores primarios estándar establecidos por la CIE no corresponden a ningún color real. El hecho de que estos primarios representen colores imaginarios no es una realidad importante, ya que cualquier color real se define como la cantidad de cada primario necesarias matemáticamente para producirlo. Cada primario de esta norma se define por sus curvas de distribución de energía (figura 1.23).

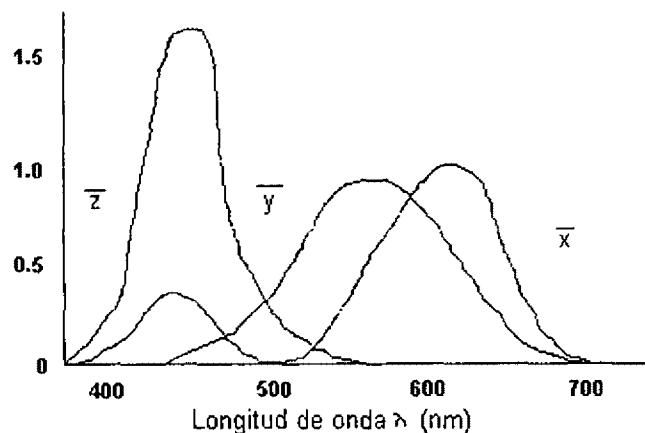


Figura 1.23 Colores estándares establecidos por la CIEB en 1931 (Rogers 1986)

La naturaleza tridimensional de color sugiere la graficación de los valores de cada componente a lo largo de ejes ortogonales (figura 1.24). El resultado es llamado espacio tristímulos. Cualquier color C es representado por el vector que va desde el origen hasta éste, y cuyas componentes son X, Y, Z requeridos para generar C . Los pesos relativos son llamados valores de cromaticidad y están dados por las coordenadas normalizadas:

$$X = \frac{X}{X+Y+Z}, \quad Y = \frac{Y}{X+Y+Z}, \quad Z = \frac{Z}{X+Y+Z}$$

donde

$$X + Y + Z = 1$$

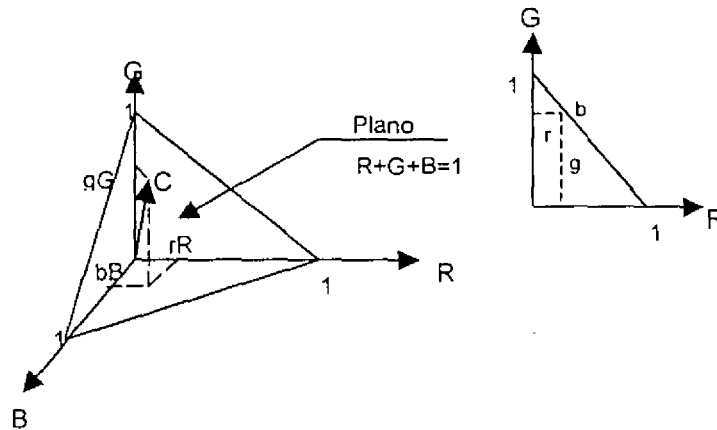


Figura 1.24 Espacio Tridimensional del color (espacio tristímulos)

(tomada de Rogers, 1986)

La proyección del plano unitario como se muestra en la figura 1.24 produce un diagrama de cromaticidad. El diagrama de cromaticidad directamente proporciona una relación entre dos colores e indirectamente con el tercero.

$$Z = 1 - X - Y$$

Esto hace posible que se represente todos los colores en un diagrama bidimensional.

Cuando se grafican valores (X, Y) para los colores del espectro visible, se obtiene la curva con forma de lengua (ver figura 1.25). Esta curva se denomina diagrama de cromatismo de CIE. Los puntos de color se rotulan a lo largo de la curva de acuerdo con la longitud de onda en nanómetros, del extremo rojo al violeta del espectro. El punto C del diagrama corresponde a la posición de luz blanca conocida como C iluminante. Que se utiliza como aproximación estándar de la luz natural media.

La curva de la figura 1.25 se denomina diagrama de cromatismo porque ofrece un medio para definir cuantitativamente la pureza y la longitud de la onda dominante.

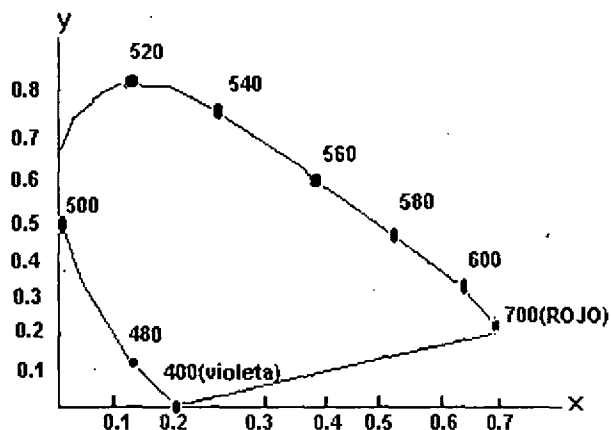


Figura 1.25 Diagrama de cromatismo de CIE (tomada de Heart and Baker, 1988)

Para cualquier punto de color, como C_1 de la figura 1.26, se define la pureza como la distancia relativa que existe desde el punto C de la luz blanca hasta C_1 , medida a lo largo de la recta que une C con la curva (que representa los colores del espectro visible) y pasa por C_1 . El color C_1 es aproximadamente 25% puro, ya que está situado cerca de una cuarta parte de la distancia total de C a C_2 . La longitud de onda dominante de cualquier color se define como el punto de cruce entre la curva espectral y la línea que une a C y ese punto de color. Para el punto de color C_1 la longitud de onda dominante está en C_2 .

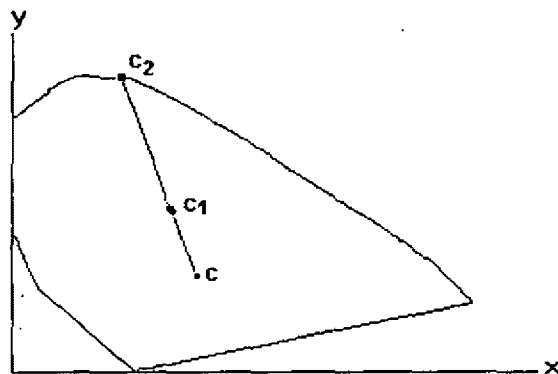


Figura 1.26 Definición de pureza y longitud de onda dominante (tomada de Heart and Baker, 1988)

Los colores complementarios se representan por los dos puntos extremos de una recta que cruza por el punto C , como en la figura 1.27. Cuando los dos colores C_1 y C_2 se combinan en proporciones adecuadas se produce la luz blanca.

La gama de colores se establece en el diagrama de cromatismo con líneas que unen los puntos de color que la definen (figura 1.28) Todos los colores situados en la línea que une los puntos C_1 y C_2 de la figura 1.28 pueden obtenerse combinando cantidades de los colores representados por los extremos. La gama de colores de tres puntos C_3, C_4 y C_5 es el triángulo que se forma. Estos tres colores pueden combinarse para generar cualquier color contenido en el triángulo, pero no podría producir puntos externos.

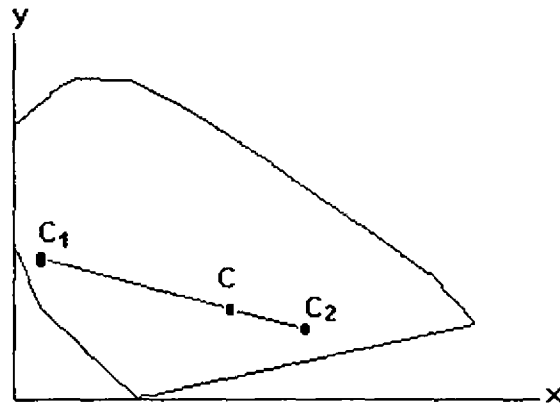


Figura 1.27 Representación de los colores complementarios (tomada de Hearn and Baker, 1988)

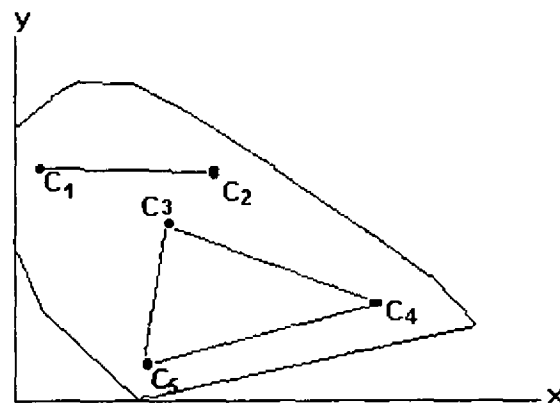
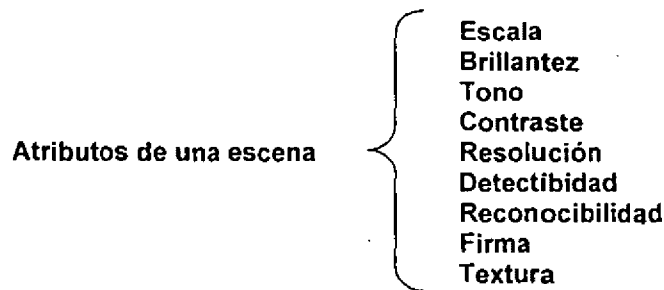


Figura 1.28 Gama de colores para los modelos de dos y tres colores primarios (tomada de Hearn and Baker, 1988)

1.3 ATRIBUTOS Y ELEMENTOS DE ANÁLISIS DE UNA IMAGEN

Todas las imágenes pueden ser descritas en términos de ciertos atributos y elementos fundamentales comunes que son: escala, brillantez, contraste, resolución, tono, textura etc.

1.3.2 ATRIBUTOS DE UNA IMAGEN



A continuación se describen cada una de ellas.

1.3.1.1 ESCALA

Es la razón de la distancia entre dos puntos en una imagen o mapa, a la correspondiente distancia sobre el terreno. La escala de una imagen es determinada por:

El campo de visión angular del dispositivo sensor.

- ◆ La altitud desde la cual la imagen es adquirida
- ◆ El factor de magnificación empleado en la reproducción de la imagen

La escala óptima de una imagen es determinada por la naturaleza del proyecto de interpretación.

1.3.1.2 BRILLANTEZ

Es la magnitud de la respuesta producida en el ojo por la luz, y es una sensación subjetiva que puede ser determinada únicamente en forma aproximada. La luminancia es una medición cuantitativa de la intensidad de luz a partir de una fuente, y es medido con un dispositivo llamado fotómetro. Las variaciones de brillantez pueden ser calibradas a partir de una escala de gris.

En imágenes generadas a partir de barredores térmicos la brillantez de un objeto es proporcional al calor irradiado desde el objeto. Para imágenes de radar la brillantez es proporcional al haz de energía siendo retroreflejado por el objeto y captado por la antena.

1.3.1.3 TONO

El término tono es usado para cada sombra distinguible desde el negro hasta el blanco.

En sistemas sensores como cámaras fotográficas y barredores multiespectrales, el tono de un objeto es primeramente determinado por la habilidad del objeto para reflejar la luz incidente, aunque los efectos atmosféricos y la sensibilidad del sistema sensor son también factores importantes que lo influyen.

1.3.1.4 CONTRASTE

La definición de contraste de una imagen es la razón entre las partes más brillantes y más oscuras de la imagen. Varias expresiones son comúnmente usadas para describir el contraste, puede haber confusión en su uso. La razón de contraste (Cr) es ampliamente usada y es definida como (Sabin, 1978):

$$Cr = \frac{B_{max}}{B_{min}}$$

Note que cuando $B_{min} = 0$, Cr no está definido: cuando $B_{min}=B_{max}$, $Cr=1$

Además de describir una escena completa la razón de contraste es también útil para describir la razón de brillantez entre un objeto y su fondo adyacente en la imagen. La razón de contraste es un factor vital para determinar la posibilidad de discriminar y detectar objetos.

1.3.1.5 RESOLUCIÓN

Existen cuatro características de una imagen de satélite con relación a su capacidad de resolución las cuales son:

- ◆ **Resolución espacial:** Es la capacidad de un sensor para distinguir dos o más elementos. A los cuales se les puede registrar como entidades separadas, donde la resolución es la distancia mínima entre ellos.
En una imagen, la resolución espacial está relacionada con el tamaño del área en la superficie, que representa un pixel.
- ◆ **Resolución espectral:** Se interpreta como la capacidad de un sensor para separar y detectar ciertos rangos de longitud de onda del espectro electromagnético.
Para una imagen, la resolución espectral está asociada con el número de bandas que posee, sus posiciones en el espectro electromagnético y el rango de frecuencias que cada una de ellas cubre.
- ◆ **Resolución radiométrica:** se define como la capacidad de un sensor para detectar la diferencia de intensidad, entre dos mediciones, cuando existe una variación en la energía que recibe el instrumento.
La resolución radiométrica en una imagen de satélite está dada por el número de niveles de gris que posee.
- ◆ **Resolución temporal (periodo de pasaje):** Es el tiempo mínimo en el que se puede obtener dos imágenes satelitales en forma consecutivas para la misma zona de la superficie terrestre.

1.3.1.6 DETECTIBILIDAD

Es la habilidad de un sistema formador de imágenes para registrar la presencia o ausencia de un objeto, aunque la identificación del objeto pueda ser desconocida. Un objeto puede ser detectado aunque éste sea más pequeño que la resolución espacial del sensor.

1.3.1.7 RECONOCIBILIDAD

Es la habilidad para identificar un objeto en una imagen. Los objetos pueden ser detectados y discriminados y aun así no ser reconocibles.

1.3.1.8 FIRMA

Una firma es la expresión de un objeto en una imagen que capacita al objeto para ser identificado. Las firmas son determinadas por las características de un objeto que determina su interacción con la energía electromagnética. Las firmas de los objetos son generalmente diferentes para cada longitud de onda.

1.3.1.9 TEXTURA

Es la menor superficie continua y homogénea distinguible en la imagen fotográfica y posible de representación. Esto implica una forma y dimensiones definidas para este elemento, y que la textura fotográfica sea un patrón de arreglos de elementos texturales.

Es posible por medio de la textura diferenciar dos regiones o dos áreas que presentan la misma tonalidad.

La distinción de los elementos es una función de la escala y de la resolución espacial de los productos utilizados y del contraste entre objetos o regiones de la superficie terrestre.

La textura sólo puede ser observada como "textura" en imágenes de gran escala. En imágenes de pequeña escala la textura es la combinación de los elementos de reconocimiento o interpretación, como drenaje, relieve, tonalidad, etc.

Un acercamiento importante para describir una región es la cuantificación de su textura. A pesar de que no existe una definición formal de textura, intuitivamente la describimos como la proporción de medida que nos caracterice su suavidad, tosquedad, regularidad, etc. de la imagen. La figura 1.29 muestra algunos ejemplos.

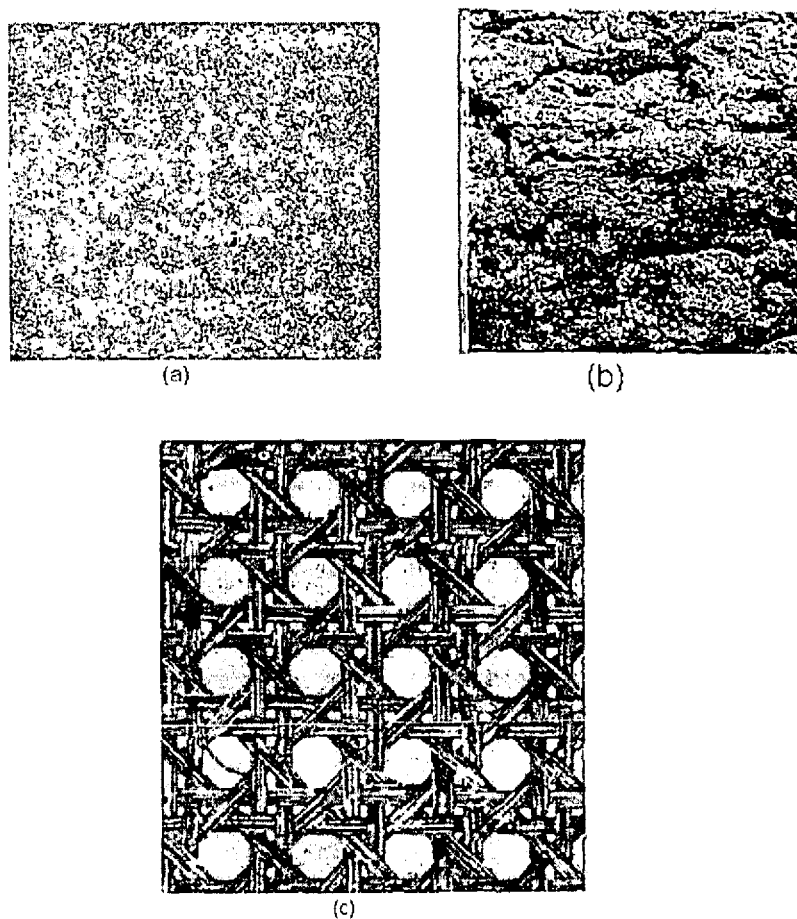


Figura 1.29 Ejemplos de textura: (a) lisa (b) tosca (c) textura regular.

Los tres principales métodos usados en el procesamiento de imágenes que describen la textura de una región son: Estadístico, Estructural y Espectral. El método estadístico caracteriza la textura como lisa, tosca, granosa etc. La técnica estructural trata con los argumentos primitivos de la imagen tal como la descripción de textura basada en la separación normal de líneas semejantes. La técnica espectral se basa en propiedades del espectro de Fourier y son usadas principalmente para detección periódica global en una imagen, identificar altos niveles de energía y picos estrechos en el espectro.

MÉTODO ESTADÍSTICO

Uno de los métodos más simples para describir la textura es usar los momentos, del histograma de niveles de gris, sea z una variable aleatoria que denotará la intensidad de la imagen discreta y $P(z)$ con $i=1,2,\dots,L$ será el histograma correspondiente, donde L es el número de los diferentes niveles de intensidad. El n -ésimo momento de z esta dado por:

$$\mu_n(z) = \sum_{i=1}^L (z_i - m)^n P(z_i)$$

donde m es el valor de z (la intensidad media)

$$m = \sum_{i=1}^L z_i P(z_i)$$

El segundo momento (llamado varianza y denotado $\sigma^2(z)$) es de particular importancia en la descripción de textura. Esta es una medida de contraste de nivel de gris será usada para constituir la descripción de suavidad. Por ejemplo

$$R = 1 - \frac{1}{1 + \sigma^2(z)}$$

R es 0 para áreas de intensidad constante ($\sigma^2(z) = 0$ si para toda z_i se tiene el mismo valor)

El tercer momento es una medida de asimetría del histograma mientras que el cuarto momento es una medida de lisura relativa. Del quinto momento en adelante, no son relativamente fáciles de obtener para la forma del histograma, pero proporcionan otras discriminaciones cuantitativas del contenido de textura.

Las medidas de textura calculadas por la computadora a través del histograma son limitadas porque no contienen información con respecto a la posición relativa de cada pixel con respecto a los otros pixeles. Un camino que brinde este tipo de información en el proceso de análisis-textura es no solo la distribución de intensidad sino la posición de los pixeles iguales y los más cercanos con el mismo nivel de intensidad.

Sea P un operador de posición, y sea A una matriz de $K \times K$ en la cual sus elementos a_{ij} son el número de veces que los puntos con nivel de gris z_i (en la posición especificada por P) relativos a los puntos con nivel de gris z_j con $1 \leq i, j \leq K$.

Por ejemplo; considere una imagen con 3 niveles de gris, $z_1 = 0$, $z_2 = 1$ y $z_3 = 2$ como sigue:

$$\begin{array}{ccccc}
 0 & 0 & 0 & 1 & 2 \\
 1 & 1 & 0 & 1 & 1 \\
 2 & 2 & 1 & 0 & 0 \\
 1 & 1 & 0 & 2 & 0 \\
 0 & 0 & 1 & 0 & 1
 \end{array}$$

Definimos el operador de posición P como "un pixel a la derecha y un pixel abajo" obteniendo la matriz A de 3×3 :

$$A = \begin{bmatrix} 4 & 2 & 1 \\ 2 & 3 & 2 \\ 1 & 2 & 0 \end{bmatrix}$$

Donde a_{11} (arriba a la izquierda) es el número de veces que un punto con nivel $z_1 = 0$ aparece localizado un pixel abajo y uno a la derecha con el mismo nivel de gris, y a_{13} (arriba a la derecha) es el número de veces que un punto con nivel $z_1 = 0$ aparece localizado uno abajo y a la derecha con un nivel de gris $z_3 = 2$. El tamaño de la matriz A es determinado estrictamente por el número de distintos niveles de gris que tiene la imagen de entrada. En la aplicación de los conceptos discutidos en esta sección, generalmente se requiere que las intensidades sean recuantizadas en 5 bandas en el orden de A .

Sea n el total de pares de puntos en la imagen que satisface P (en el ejemplo $n = 4 + 2 + 1 + 2 + 3 + 2 + 2 = 16$) si una matriz C es formada dividiendo cada elemento de A entre n y entonces c_{ij} es un estimador de probabilidad colectiva que un par de puntos satisfaga P tendrá valores (z_i, z_j) . La matriz C es llamada matriz de concurrencia de niveles de gris. Porque C depende de P , el presentar un parámetro de textura será detectado escogiendo el operador de posición adecuado. Para el ejemplo el operador usado es sensitivo a las bandas de intensidad constante sobre una línea a 45 grados (Note que el valor mas alto en A fue $a_{11} = 4$ partiendo de que tenemos una línea de puntos con intensidad igual a cero y a 45 grados). Generalmente, el problema es analizar una matriz C del orden que categorice la textura de la región sobre C que puede obtenerse computacionalmente. Un conjunto de descriptores para estos propósitos se incluye a continuación.

(1) probabilidad máxima

$$\max_{i,j} (c_{ij})$$

(2) diferencia elemental de momentos de orden K

$$\sum_i \sum_j (i-j)^k c_{ij}$$

(3) diferencia elemental inversa de momento de orden K

$$\sum_i \sum_j c_{ij} / (i-j)^k \quad i \neq j$$

(4) entropy

$$-\sum_i \sum_j c_{ij} \log c_{ij}$$

(5) uniformidad

$$\sum_i \sum_j c_{ij}^2$$

La idea básica es caracterizar el contenido de C vía su descripción. Por ejemplo para la primer propiedad da un indicador de las respuestas fuertes de P como en el presente ejemplo). La segunda descripción tiene un valor relativamente bajo cuando los valores altos de C se encuentran cerca de la diagonal principal, porque la diferencia $(i - j)$ son muy pequeñas. La tercera descripción tiene el efecto opuesto. La cuarta descripción es una medida aleatoria, alcanza su valor mas alto cuando todos los valores de C son iguales. La quinta descripción es la mas baja cuando todos los c_{ij} s son todos iguales.

MÉTODO ESTRUCTURAL

Como se menciona al inicio de esta sección, la segunda mejor categoría de descripción de textura esta basada en el concepto de su estructura. Supone que tenemos una regla de la forma:

$$S \rightarrow aS$$

La cual indica que el símbolo S será rescritó como aS (por ejemplo la aplicación de esta regla tres veces nos dará $aaaS$). Si a representa un círculo (figura 1.30) y el significado de "círculos a la derecha es asignado a la cadena de la forma $aaa.....$ entonces $S \rightarrow aS$ será generador del patrón de textura mostrado en la figura 1.30(b).

Supóngase que ahora aumentamos nuevas reglas a este esquema: $S \rightarrow bA$, $A \rightarrow cA$, $A \rightarrow c$, $A \rightarrow bS$, $S \rightarrow a$, donde la representación de b significa "círculos hacia abajo", y la presencia de una c "círculos a la izquierda". Podemos ahora generar una cadena de la forma $aaabccbaa$ que corresponde al círculo 3,3 de la matriz. El patrón de textura, tal como es mostrado en la figura 1.30c será generado fácilmente con el mismo camino (sin embargo, estas reglas pueden generar estructuras que no sean rectangulares).

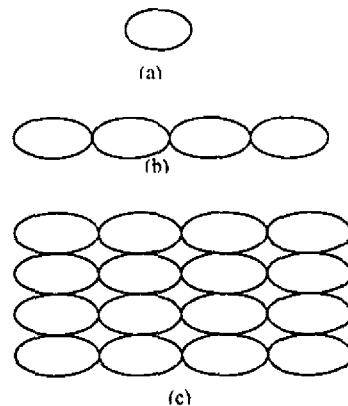


Figura 1.30 (a) textura primitiva (b) patrón generalizado para la regla $S \rightarrow aS$, (c) generación del patrón de textura en 2-DD para esta y otras reglas.

La idea básica en la discusión es que una "primitiva estructura" simple puede ser usada para formar patrones de textura complejos, pero también algunas reglas que limitan el número de arreglos posibles de las primitivas.

MÉTODDO ESPECTRAL

El espectro de Fourier es idealmente adecuado para describir la dirección del periodo o los patrones periódicos bidimensionales en una imagen. Estos patrones de textura global, aunque fácilmente distinguibles como concentración de altos niveles de energía en el espectro, generalmente son considerablemente difíciles de detectar con métodos espaciales por su naturaleza local de estos métodos.

En este trabajo consideramos tres características del espectro de Fourier que son usadas para describir textura:

- 1).- Puntas prominentes.- da la dirección principal de los patrones de textura.
- 2).- La localización de puntas en el plano frecuencial da el periodo espacial fundamental de los patrones.
- 3).- Eliminación de algunos componentes periódicos por filtrado dejando elementos no-periodicos de la imagen que podrán ser descritos por técnicas estáticas. Recordando que el espectro de una imagen real es simétrico al origen, por lo que solo la mitad del plano frecuencial será considerado. Para el propósito de análisis, cada patrón periódico es asociado con solo un pico o punta en el espectro frecuencial, antes que dos.

La detección e interpretación de las características espectrales mencionadas deben ser simplificadas por expresiones del espectro en coordenadas polares produciendo una función $S(r, \vartheta)$, donde S es la función espectral y r y ϑ son las variables en el sistema coordenado. Para cada dirección ϑ $S(r, \vartheta)$ será considerada una función unidimensional $S_{\vartheta}(r)$. Similarmente para cada frecuencia r , $S_r(\vartheta)$ es una función unidimensional. Analizando $S_{\vartheta}(r)$ para cada valor de ϑ produce el comportamiento del espectro (tal como la presencia de picos) por una dirección radial del origen. Mientras que analizamos $S_r(\vartheta)$ para un valor fijo de r produce un comportamiento circular centrado en el origen.

Una descripción más completa es obtenida por la integral (sumando las variables discretas) de estas funciones.

$$S(r) = \sum_{\vartheta=0}^n S_{\vartheta}(r)$$

y

$$S(\theta) = \sum_{r=1}^R S_r(\theta)$$

Donde R es el radio de un círculo con centro en el origen. Para un espectro de $N \times N$, R es generalmente escogida como $N/2$

El resultado de $S(r)$ y $S(\theta)$ constituye un par de valores $[S(r), S(\theta)]$ para cada par de coordenadas (r, θ) . Variando estas coordenadas podemos generar dos funciones unidimensionales, $S(r)$ y $S(\theta)$, que constituyen un descriptor de espectro-energía de textura para una imagen completa o una región a considerar. Además descriptores de estas funciones pueden ser computados en orden de caracterización cuantitativa. Descriptores típicamente usados para estos propósitos es la localización de los valores más altos. La media y varianza de ambos la amplitud y la variación axial, y la distancia entre la media y el valor más alto de la función.

Ejemplo.- La Figura 1.31 ilustra el uso de las ecuaciones $S(r)$ y $S(\theta)$ para descriptores de textura globales. La figura 1.31a muestra una imagen con textura periódica y la figura 1.31b muestra su espectro. La figura 1.31c y 1.31d muestran las gráficas de $S(r)$ y $S(\theta)$, respectivamente. La gráfica de $S(r)$ es una estructura típica, presentando altos niveles de energía cercanos al origen y progresivamente los valores bajan para frecuencias altas. La gráfica $S(\theta)$ picos prominentes en intervalos de 45 grados, correspondientes a lo blanco del periodo en el contenido de la textura de la imagen.

Como una ilustración de una gráfica puede ser usada para diferenciar entre dos patrones de textura en las direcciones vertical y horizontal. La figura 1.31f muestra la gráfica $S(\theta)$ para el espectro de su imagen (figura 1.31e). Esta gráfica muestra los picos altos en intervalos de 90 grados. Discriminar entre los dos patrones de textura analizando sus correspondientes $S(\theta)$.

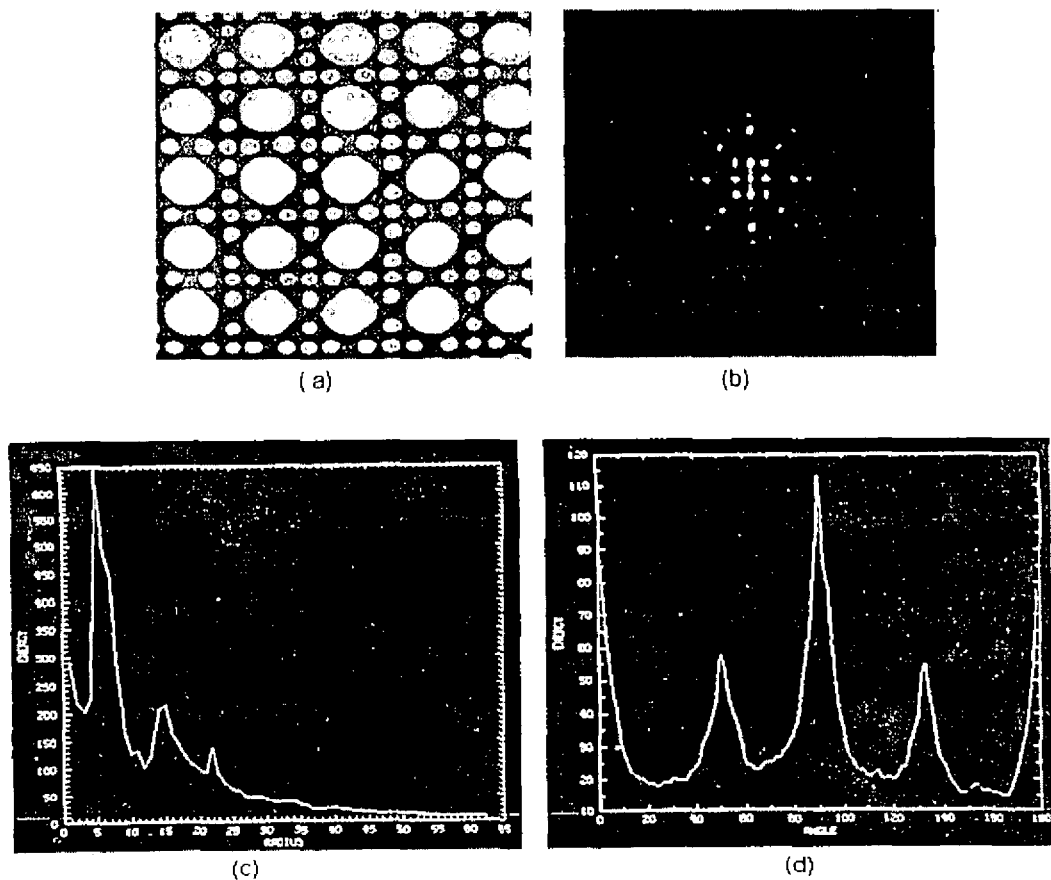
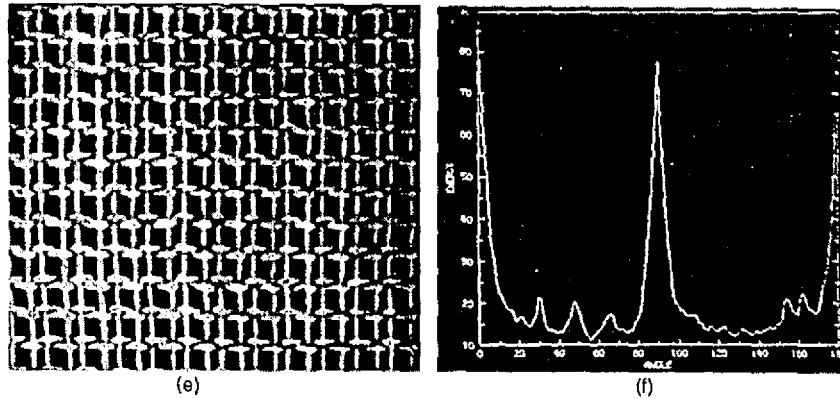
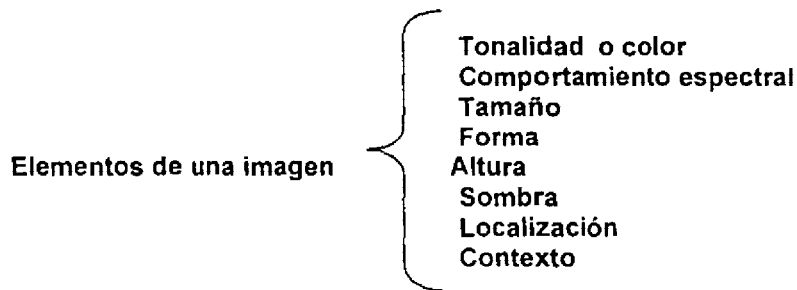


Figura 1.31 (a) muestra una imagen con textura periódica (b) su espectro (c) gráfica de $S(r)$ (d) gráfica de $S(\theta)$



Continúa Figura 1.31 (e) otra imagen con diferente tipo de textura periódica.

1.3.2 ELEMENTOS DE UNA IMAGEN



1.3.2.1 TONALIDAD O COLOR

La tonalidad representa el nivel de gris con que un objeto o grupo de objetos son registrados en una película fotográfica en blanco y negro. La tonalidad, para una imagen de satélite, está dada por la cantidad de energía reflejada y/o emitida por el objeto en una determinada banda espectral.

Las imágenes a color permiten que los objetos sean diferenciados a través de variaciones de color. La percepción de color es un elemento de nuestra conciencia sobre el medio ambiente.

El color y la tonalidad de los objetos (como representativos del comportamiento espectral) están sujetos a variaciones en función de las condiciones ambientales y de la adquisición de los datos.

1.3.2.2 COMPORTAMIENTO ESPECTRAL

En una imagen correspondiente al infrarrojo, un bosque tendrá tonalidades de gris claro, porque en esta región hay elevada reflexión de energía. En una imagen captada en la región del rojo, el bosque representara tonalidad oscura (negra), porque en esta porción del espectro la vegetación verde tiene elevada absorción de la energía debido a la clorofila.

1.3.2.3 TAMAÑO

El tamaño de un objeto es uno de los factores más útiles para su identificación. A través de la medida de un objeto desconocido de una fotografía aérea, el interprete puede eliminar varias consideraciones al respecto de todos los grupos de posibles identificaciones.

Para la evaluación de las dimensiones de un objeto como elemento de identificación, se debe de tomar en cuenta la escala de la imagen que está siendo analizada.

1.3.2.4 FORMA

Las formas indican la disposición espacial de los elementos texturales con propiedades comunes.

El valor de la forma para el fotointérprete es, que ella delimita la clase a la cual un objeto desconocido debe pertenecer. Frecuentemente, permite una identificación determinante y auxilia en el entendimiento del significado y función de objeto.

Es importante considerar que la forma de un objeto observado a partir de una perspectiva vertical es diferente con relación a la mira horizontal. La habilidad para comprender y discriminar un objeto a partir de su visión en un plano debe ser desarrollada.

1.3.2.5 ALTURA

La altura de los objetos ayuda también al interprete a diferenciar dos regiones que poseen tonalidades similares. Cuando se dispone de un par de estereoscópicos se puede tener una visión tridimensional del terreno y disponer de la altura de los objetos.

1.3.2.6 SOMBRA

Las sombras en la imagen son resultantes de la iluminación oblicua, por el sol, de la superficie del terreno, en el instante de la toma de la imagen.

En las fotografías aéreas las sombras algunas veces ayudan al interprete, proporcionándole representaciones del perfil de los objetos de interés. Ellas son particularmente útiles si los objetos son muy pequeños o hace falta contraste con respecto a los objetos vecinos.

1.3.2.7 LOCALIZACIÓN

La localización de los objetos puede no ser una llave para su identificación, pero puede informarnos su significado mediante la indicación de su uso. La localización de los objetos con relación a las características del terreno u otros objetos es frecuentemente una pista determinada para la identificación de su naturaleza.

1.3.2.8 CONTEXTO

Algunos objetos pueden ser identificados por la compañía que ellos tienen. Algunos objetos están asociados con otros y la presencia de uno, confirma la presencia de otros.

CAPÍTULO II

PROCESAMIENTO DE IMÁGENES

CONTENIDO

2 PROCESAMIENTO DE IMÁGENES

2.1 Modificación del histograma

- Métodos de modificación del histograma
 - Normalización de la imagen o "Contrast Stretching"
 - Ecuilización
 - Umbralización

2.2 Filtrado

- Introducción
 - 2.2.1 Filtrado espacial
 - 2.2.1.1 Filtros lineales
 - Filtro paso bajas
 - Filtro paso bajas de media ponderada
 - Filtro paso altas
 - Filtro supresor de bandas
 - Filtro paso banda
 - 2.2.1.2 Filtros no lineales
 - Filtro de mediana
 - Filtro de máxima y de mínima
 - Filtro de moda
 - Aplicación del gradiente(detección de bordes)
 - 2.2.2 Filtrado frecuencial
 - 2.2.2.1 Transformada de Fourier
 - 2.2.2.2 Transformada rápida de Fourier
 - 2.2.2.3 Propiedades de la transformada de Fourier
 - Traslación
 - Periodicidad y simetría
 - Rotación
 - Linealidad y escala
 - 2.2.2.4 Filtro paso bajas
 - 2.2.2.5 Filtro paso altas
 - 2.2.2.6 Filtro rechazo banda
 - 2.2.2.7 Filtro paso banda
-

CAPÍTULO II

PROCESAMIENTO DE IMÁGENES

2.1 MODIFICACIÓN DEL HISTOGRAMA

El histograma es una imagen digital con niveles de gris definidos en el rango $[0, L-1]$ es una función discreta expresada por la ecuación siguiente.

$$p(r_k) = n_k / n$$

Donde r_k es el k -ésimo nivel de gris, n_k es el número de píxeles en la imagen de cada uno de los niveles de gris, n es el número total de píxeles en la imagen y $k = 0, 1, 2, \dots, L-1$. El histograma se puede conceptualizar como un vector que contiene el mismo número de elementos que los niveles de cuantización, en cada elemento, es guardado el número de píxeles que tiene el valor de gris correspondiente al elemento del vector.

La función $p(r_k)$ nos da una estimación de la probabilidad de aparición del nivel de gris r_k . Una gráfica de esta función para todos los valores posibles de k nos da una descripción global de la apariencia de una imagen.

El histograma se ha convertido en una de las herramientas para la manipulación de las imágenes más utilizadas debido a que a pesar de su sencillez nos permite conocer a través de él una gran variedad de información sobre una imagen.

El histograma nos va a proporcionar la distribución de los niveles de gris de la imagen (aunque no nos localiza esos niveles de gris espacialmente). Así el histograma de niveles de gris se define como una curva representada sobre un eje cartesiano donde las abscisas representan los niveles de gris normalmente se toman 256 niveles de gris, de 0 a 255, medida muy aceptable ya que el ojo humano solo puede discernir 64 niveles de gris. Dentro de este rango el 0 representa el negro absoluto y el 255 el blanco absoluto (ver figura 2.1).

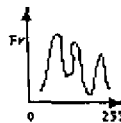


Figura 2.1 Modelo del histograma.

Hay otra manera de hallar el histograma. Tenemos una imagen que está definida por la función continua $D(x, y)$, la cual varía de forma que el nivel de gris más alto está en el centro y se va haciendo más débil conforme vamos avanzando hacia los bordes. Podemos seleccionar entonces un nivel de gris D y trazar una línea de contornos que conecte a todos los puntos que tengan ese mismo nivel; Conseguimos así una curva cerrada que rodea a las zonas con igual o mayor nivel de gris, a cuya área se denominará A . Ahora ya podríamos definir el histograma como.

$$\Pi(D) = \lim_{\Delta D \rightarrow 0} \frac{A(D) - A(D - \Delta D)}{\Delta D} = -\frac{d}{dD} A(D)$$

Por medio de un rápido vistazo al histograma correspondiente podremos saber si una imagen está muy oscura o muy clara (normalmente, si un histograma está muy desplazado hacia uno de los lados significa que estamos perdiendo niveles de gris), si tiene poco o mucho contraste e incluso en algunos casos el posible contenido de la imagen. Así por ejemplo podemos ver las siguientes fotografías de una actriz de Hollywood donde su imagen, dependiendo de su claridad u oscuridad, presenta un histograma inclinado hacia un extremo o hacia el otro.

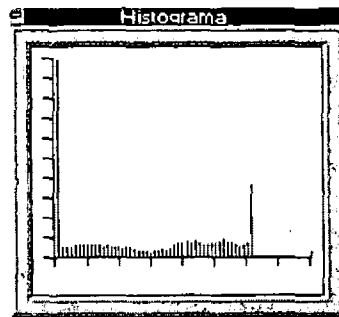


Figura 2.2 Fotografía original y su histograma

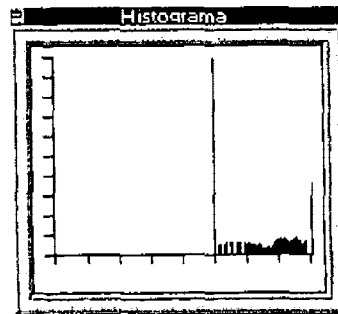


Figura 2.3 Fotografía clara y su histograma (desplazado hacia la derecha)

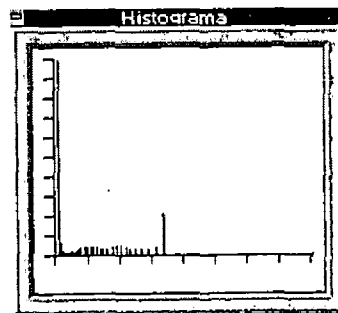
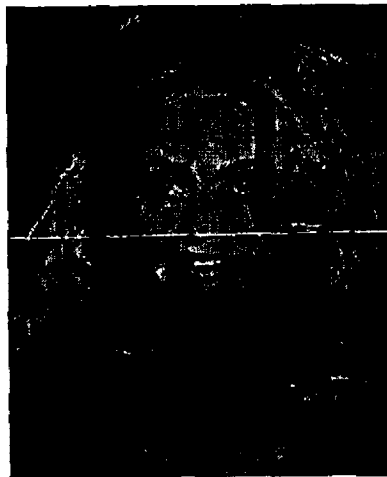


Figura 2.4 Fotografía oscura y su histograma (desplazado hacia la izquierda)

El histograma no caracteriza de forma unívoca a una imagen, solo muestra frecuencias de grises en la imagen pero no nos da información espacial. Es decir que no podemos reconstruir la imagen a través de su histograma, hecho que se tendría que cumplir para que el histograma fuera unívoco respecto a la imagen.

Como ya hemos comentado, el histograma no tiene información espacial, sino simplemente de frecuencia de escalas de gris. Por esto si nosotros tuvieramos dos imágenes diferentes (incluso de distinto tamaño) pero con la misma frecuencia de tonos de gris estas tendrían el mismo histograma.

Dos ejemplo buenos que confirmarían esta afirmación serían o bien una figura geométrica con el mismo fondo, pero situada en distintos puntos de cada imagen (ver figura 2.5)

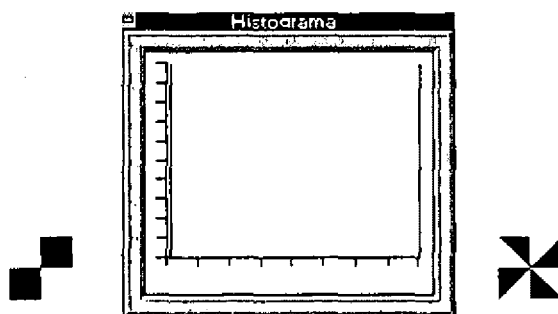


Figura 2.5 Imágenes distintas (a los lados) con un mismo histograma.

METODOS DE MODIFICACION DEL HISTOGRAMA

NORMALIZACIÓN DE LA IMAGEN O “CONTRAST STRETCHING”

Es una técnica de tratamiento de imágenes que consiste en mejorar el contraste de la imagen “alargando” el rango de valores de intensidad de gris de la gráfica del histograma o del rango de valores de intensidad de gris partiendo de la gráfica del histograma o de una parte de él. Difiere de la ecualización del histograma en que solo puede aplicar una función escalar lineal al valor de los pixeles de la imagen.



Figura 2.6 Simulación de normalización.

Para obtener la normalización que se muestra en la figura 2.6 es necesario especificar el límite superior e inferior de pixel antes de comenzar la normalización. Esto suele ser el mínimo y máximo valor de pixel de la imagen respectivamente.

El procedimiento más corto para realizar la normalización suele ser, una vez conocidos los límites inferior (a) y superior (b) de la imagen, escanear esta para encontrar el menor (c) y mayor (d) valor de pixel y aplicar a cada pixel (p) la siguiente formula:

$$p_{ent} = (p_{sal} - c) \left(\frac{b - a}{d - c} \right) + a$$



Figura 2.7 Imagen antes de normalización (izquierda) y la misma imagen una vez normalizada (derecha)

Además de este uso, la normalización también se utiliza cuando transformamos el tipo de imagen de valores de pixel en punto flotante a valores enteros de 8 bit. La normalización va a permitir en este caso una fácil conversión.

Un realce lineal de contraste es mejor aplicado para imágenes con histogramas de distribución gaussiana o semi-gaussiana, esto es, donde todos los valores de brillantez caen generalmente dentro de un simple, y relativamente angosto, rango del histograma y existe únicamente una moda. Desafortunadamente, este caso no siempre se da, especialmente para escenas con gran cantidad de cuerpos de agua.

Cuando en una imagen se toman el valor mínimo y máximo de brillantez y se expande a todo el rango posible, se dice que se tiene una elongación de contraste mínima-máxima. Si en lugar de estos valores el analista especifica un rango (min,max) que determine ciertos porcentajes en forma lineal. La figura 2.8 muestra estos dos tipos de contrastes.

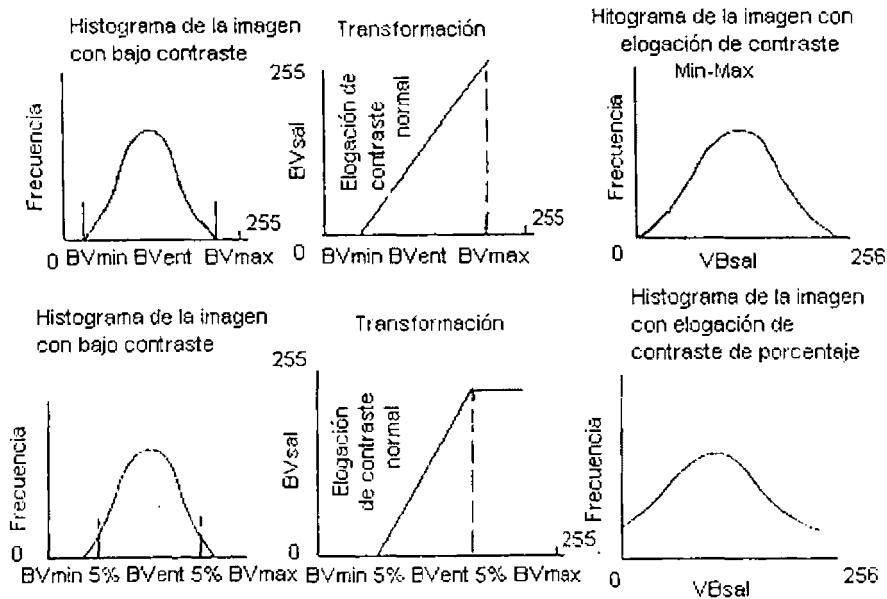


Figura 2.8 Elogación mínima-máxima y de porcentaje de un histograma.

Si el histograma de una imagen no es gaussiano (es decir es: bimodal, trimodal, etc.) es posible efectuar una elongación de contraste lineal por pedazos a la imagen, como es mostrado en la figura 2.9. Aquí el analista, identifica un número de pasos de realce lineal, que expande los rangos de brillantez solamente en las modas del histograma. Esto corresponde a determinar una serie de segmentos de valor mínimo de brillantez a expandir y valor máximo de brillantez a expandir y emplear la ecuación arriba mencionada, dentro de cada región del histograma seleccionada por el usuario.

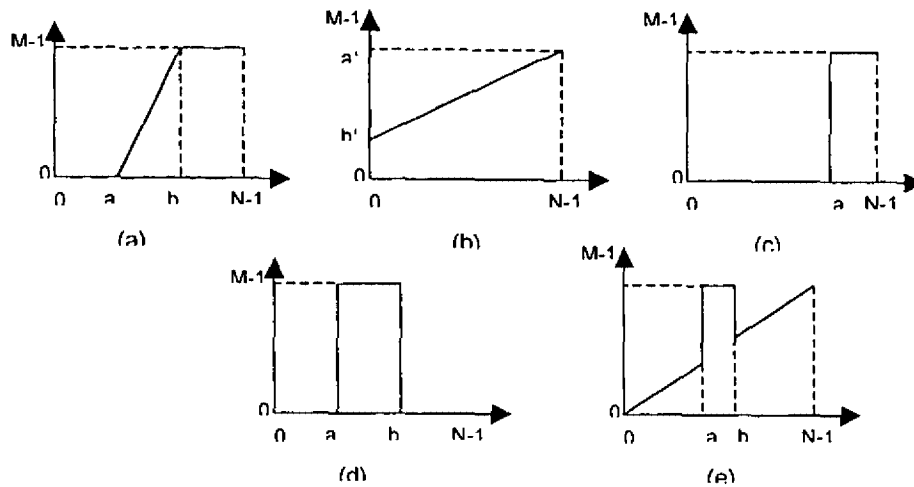


Figura 2.9 Elogación de contraste lineal por pedazos.

La función mostrada en la figura 2.9a realiza las variaciones dentro del intervalo (a,b), mientras que satura a 0 los valores menores que a, y a M-1, los valores mayores que B. En la figura 2.9b, los niveles de 0 a N-1 son comprimidos para el intervalo (a',b').

La definición de un valor de umbral es el efecto de la función mostrada en la figura 2.9c, la imagen resultante tiene solo dos valores 0 y M-1. El efecto de la función 2.9d es destacar los puntos de la imagen en el intervalo (a,b); la imagen resultante tiene, también, únicamente dos valores. La función de la figura 2.9e es semejante a la anterior, pero los niveles de gris fuera del intervalo (a,b) se mantiene.

La elogación de contraste lineal por pedazos también puede ser efectuada cuando ya se tiene un conocimiento previo de los valores de reflectancia de los objetos de interés, para resaltarlos del resto de la imagen

ECUALIZACIÓN

Es un sofisticado método para modificar el rango y contraste dinámico de una imagen. Este se altera aplicando un determinado modelo al histograma.

La ecualización de histogramas se realiza por medio de una función monótona, no lineal (al contrario de la herramienta anterior) que transforma los valores de pixel de una entrada por otros de salida que siguen una distribución uniforme de intensidades. Esta técnica es utilizada entre otras cosas para procesos de comparación de imágenes.

Uno de los realces de contraste más usados es la ecualización de histograma, el cual se efectúa mediante los siguientes pasos:

Primero, el histograma de la imagen es integrado (figura 2.10a) generándose su función de distribución (figura 2.10b), a partir de ésta se obtiene la función de distribución acumulativa (figura 2.10c), la cual actúa como función de transformación. Posteriormente, se comparan cada uno de los valores de probabilidad acumulada transformados contra los valores de brillantez originales, y se asignan al valor de brillantez más cercano a su valor. Aquellos valores de brillantez que según la función de transformación son asignados a un mismo nuevo valor son integrados aditivamente. La probabilidad de encontrar cada uno de los nuevos valores de brillantez, así como el número de pixeles que poseen son mostrados en las figuras 2.10d y 2.10e.

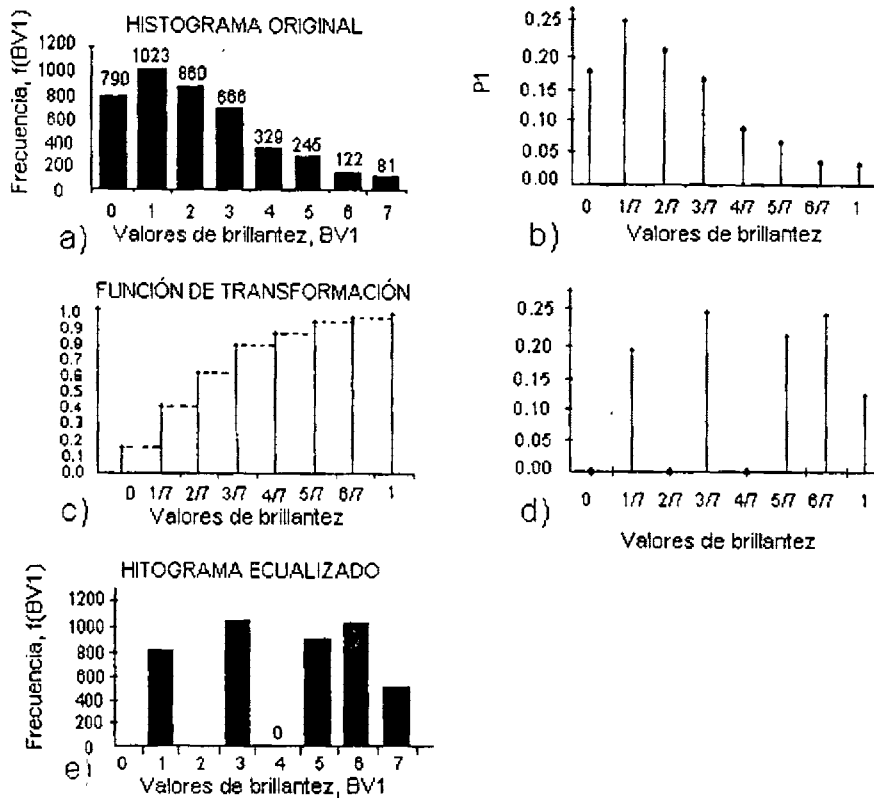


Figura 2.10 Proceso de ecualización de un histograma.

La ecualización de histograma aplica el mayor realce de contraste a los rangos más poblados de valores de brillantez en la imagen. Esto automáticamente reduce el contraste en las partes muy claras o muy oscuras, asociadas con los extremos del histograma(ver figura 2.11).



Figura 2.11 Imagen anterior una vez ecualizada

La ventaja que tiene este método es que se consigue un incremento radical del contraste, sin embargo, el resultado es más artificial que si se utiliza la normalización.

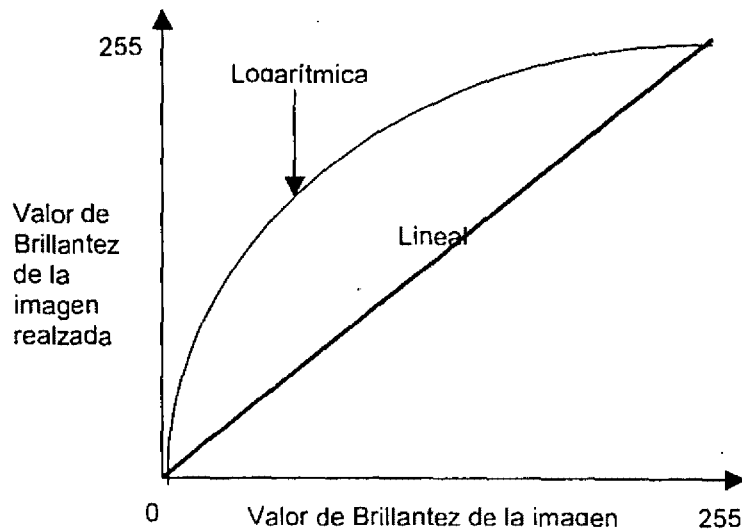


Figura 2.12 Elogación lineal y logarítmica de un histograma.

UMBRALIZACIÓN

La umbralización es la que separa las regiones en la imagen que corresponden a los objetos que nos interesan de su correspondiente fondo. Es también conocida como "Thresholding" esta técnica nos facilita un rápido camino para realizar esta segmentación en base a las diferentes intensidades o colores que tiene el fondo y los objetos contenidos en la imagen.

La más simple implementación es partir de la imagen original para devolver una imagen binaria que represente la segmentación, donde los píxeles negros corresponderán al fondo y los blancos al objeto (o viceversa). En esta implementación la segmentación está determinada por un parámetro conocido "Intensidad Umbral" (ver figura 2.13). En una simple pasada, cada píxel se compara con la intensidad umbral; si la intensidad del píxel es mayor que la del umbral el píxel se transforma, por ejemplo, a blanco y en caso contrario a negro (ver figura 2.14).

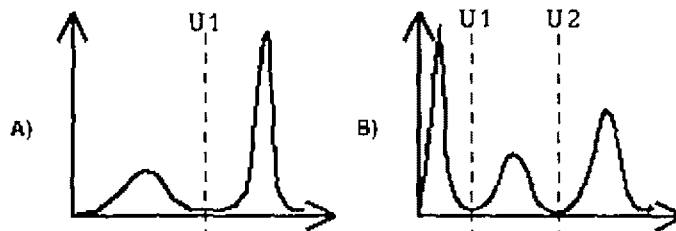


Figura 2.13 Dos ejemplos de cómo aplicar la umbralización.

La umbralización además del caso anteriormente expuesto tiene otra posible aplicación que es el ver que áreas de una imagen tiene sus píxeles dentro de un determinado rango o banda de intensidades (o colores). Estas dos aplicaciones hacen que esta herramienta sea ideal para exploraciones astrológicas, reconocimiento de formas en robótica, etc.



Figura 2.14 Ejemplo de umbralización

A continuación se presenta un ejemplo sencillo de Modificación del Histograma con el fin de visualizar el uso de esta técnica ya que el objetivo de este trabajo no es profundizar con esta técnica específicamente.

Para obtener el mapeo que transforme el histograma de la imagen y como consecuencia la apariencia de la imagen podemos emplear el siguiente algoritmo:

1. Primero obtenemos el histograma de la imagen de entrada el cual denotamos como H_f , supongase que el rango de niveles de gris es $[0,10]$ (ver figura 2.15).

Nivel de Gris	H_f	H^*f
0	0	0
1	0	0
2	2	2
3	4	6
4	8	14
5	2	16
6	2	18
7	2	20
8	0	20
9	0	20
10	0	20

Imagen Original

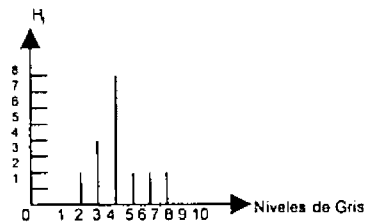


Tabla 2.1 Distribución frecuencial H_f y valores acumulados H^*f de los diferentes niveles de gris dentro de la imagen

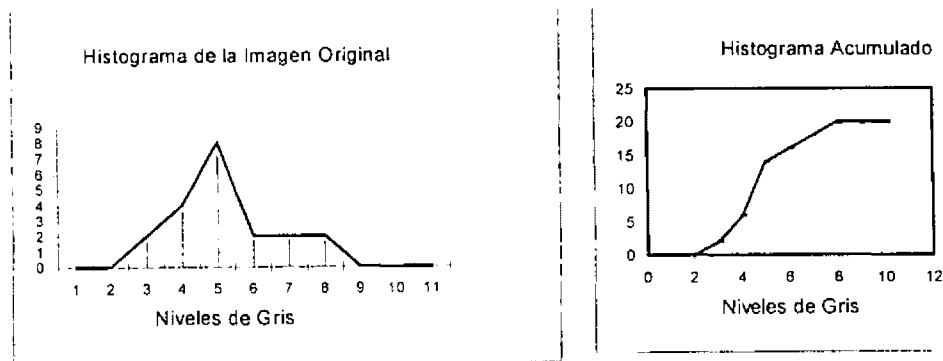


Figura 2.15 (a) Histograma de la imagen original. (b) Histograma acumulado

2. Definimos cual es la forma que deseamos darle al histograma, esto lo podemos obtener por medio de una función. Para el ejemplo que se proporciona en el presente trabajo se emplea una función de distribución normal, porque con ella podemos controlar el brillo y el contraste de la imagen, especificando la media y la varianza respectivamente, además, la función de distribución normal es simétrica con respecto a la media lo que permite crear una imagen con tonalidades de gris balanceados. A este histograma lo denotamos como H_d .

El efecto que deseamos darle a la imagen, es hacerla más brillante y con poco contraste. Para conseguir nuestro objetivo definimos un histograma con la forma de distribución normal con valores de la media $\mu = 8$ y la varianza $\sigma^2 = 1$. con lo que obtenemos la tabla 2.2 para valores enteros entre uno y diez.

n	Hd(n)
0	0.0000006349117
1	0.0000134985669
2	0.0000348132629
3	0.0005445710573
4	0.005166746339
5	0.02973257231
6	0.1037768744
7	0.2196956447
8	0.2820947918
9	0.2196956447
10	0.1037768744

Tabla 2.2 Distribución normal para 0 a 10 con $\mu = 8$ y $\sigma = 1$

Dado que requerimos que los histogramas tengan el mismo número de elementos para una comparación adecuada, multiplicamos estos resultados con el número total de pixeles que debe contemplar este histograma y redondeando los resultados obtenemos los datos que se muestran en la tabla 2.3.

n	Hd(n)	Hd(n)*20	Redondeando	H*d(n)
0	6.34912E-07	1.26982E-05	0	0
1	1.34986E-05	0.000269971	0	0
2	3.48133E-05	0.000696265	0	0
3	0.000544571	0.010891421	0	0
4	0.005166746	0.103334927	0	0
5	0.029732572	0.594651446	1	1
6	0.103776874	2.075537488	2	3
7	0.219695645	4.393912894	4	7
8	0.282094792	5.641895836	6	13
9	0.219695645	4.393912894	4	17
10	0.103776874	2.075537488	2	19

Tabla 2.3 Generación del histograma deseado de la imagen

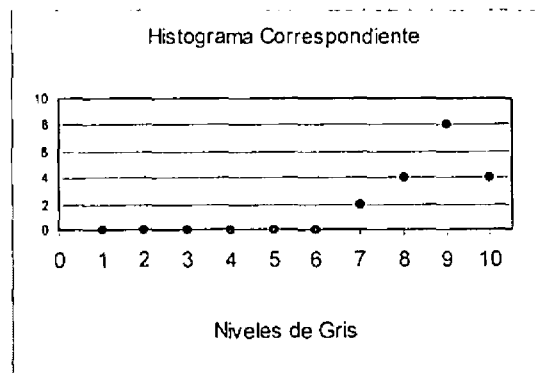


Figura 2.16 Gráfica del Histograma deseado

3. Calculamos los histogramas acumulados de la imagen de entrada y del histograma deseado. Definimos el histograma acumulado de H' como:

$$H'(0) = H(0)$$

$$H'(n+1) = H(n+1) + H'(n)$$

Pero al calcular el histograma acumulado observamos que el número total de elementos no es igual a veinte, el cual es un número de pixeles de la imagen de entrada. Para que nuestro algoritmo obtenga el mapeo correcto necesita que los histogramas acumulados contemplen el mismo número de pixeles. Para obtener nuestra tabla 2.4 ya ajustada podemos aplicar la expresión al ultimo elemento de la tabla.

$$H_d(n) = H_a(n) * 20 / H(19)$$

obteniendo así

N	fd(n)	fd(n) ajustado
0	0	0
1	0	0
2	0	0
3	0	0
4	0	0
5	1	1
6	3	3
7	7	7
8	13	13
9	17	17
10	19	20

Tabla 2.4 Histograma deseado ajustado.

4. El mapeo que aproxime el histograma de la imagen con el histograma deseado se define como:

$$m(r) = n$$

donde n es tal que $H'_d(x)$ sea el valor más próximo a $H'_f(x)$

La tabla 2.5 muestra el mapeo como se explica en el punto 4 del algoritmo.

n	Hf	Hd	m
0	0	0	0
1	0	0	0
2	2	0	6
3	6	0	7
4	14	0	8
5	16	1	9
6	18	3	10
7	20	7	10
8	20	13	10
9	20	17	10
10	20	20	10

Tabla 2.5 Tabla de Mapeo

Para visualizar la transformación graficamos el resultado del mapeo en la figura 2.17

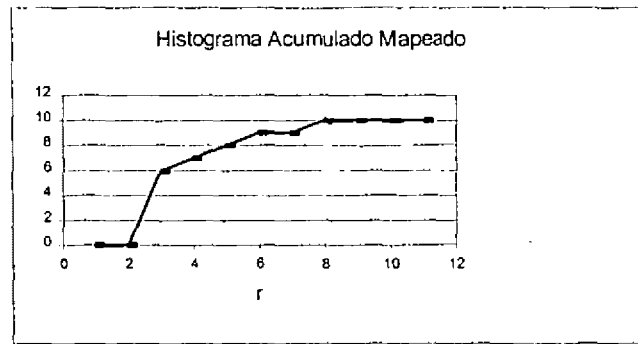


Figura 2.17 Histograma Modificado Acumulado

Una vez aplicando el mapeo encontrado en la imagen original obtenemos el histograma de salida (tabla 2.6).

n	H ⁱ	m=n	Fo(n)
0	0	0	0
1	0	0	0
2	2	6	2
3	6	7	4
4	14	8	8
5	16	9	2
6	18	9	2
7	20	10	2
8	20	10	0
9	20	10	0
10	20	10	0

Tabla 2.6a Aplicando el mapeo

Finalmente obtenemos como resultado

n	f ₀ (n)
0	0
1	0
2	0
3	0
4	0
5	0
6	2
7	4
8	8
9	4
10	2

Tabla 2.6b Resultado del mapeo

Y su histograma correspondiente se muestra en la figura 2.17

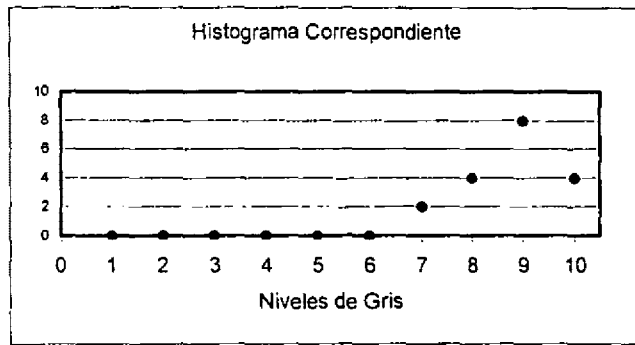


Figura 2.17 Histograma Deseado.

Cuando se aplica el algoritmo para obtener la función de distribución uniforme se dice que la imagen pasa por el proceso de ecualización del histograma. Por ejemplo para ecualizar el histograma del ejemplo anterior, primero obtenemos el histograma. Sabemos que nuestro histograma deseado debe de contemplar el mismo número de pixeles que el histograma de la imagen original entonces dividiendo $20/10$ nos da el valor de cada uno de los valores de gris. Por lo que obtenemos el siguiente histograma:

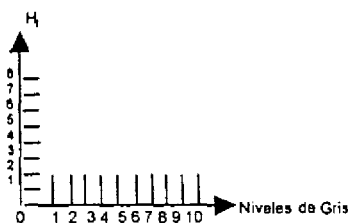


Figura 2.18 Histograma deseado de salida.

Acumulando ambos histogramas y aplicando el punto 4 de nuestro algoritmo obtenemos el mapeo.

N	H _i	H' _d	M
0	0	0	0
1	0	2	0
2	2	4	1
3	6	6	3
4	14	8	7
5	16	10	8
6	18	12	9
7	20	14	10
8	20	16	10
9	20	18	10
10	20	20	10

Tabla 2.7 Mapeo de los histogramas.

La gráfica del mapeo encontrado se muestra a continuación (figura 2.19).

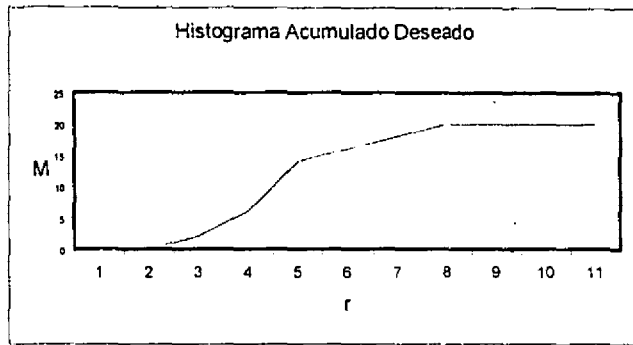


Figura 2.19 Histograma acumulado deseado

Aplicando el mapeo encontrado a la imagen original obtenemos la tabla 2.8.

n	$f_0(n)$
0	0
1	2
2	0
3	4
4	0
5	0
6	0
7	8
8	2
9	2
10	2

Tabla 2.8 Histograma final

Que da como resultado el siguiente histograma.

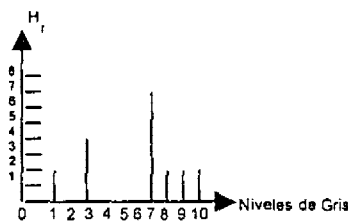


Figura 2.20 Histograma final modificado.

La figura 2.21 nos muestra una imagen y su histograma, a la cual se le aplica el algoritmo de especificación del histograma, cuyos resultados se pueden ver en las figuras 2.22 y 2.23

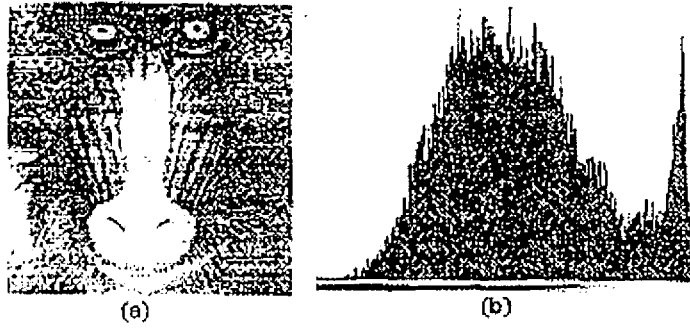


Figura 2.21 (a) Imagen Original; (b) Histograma de la imagen.

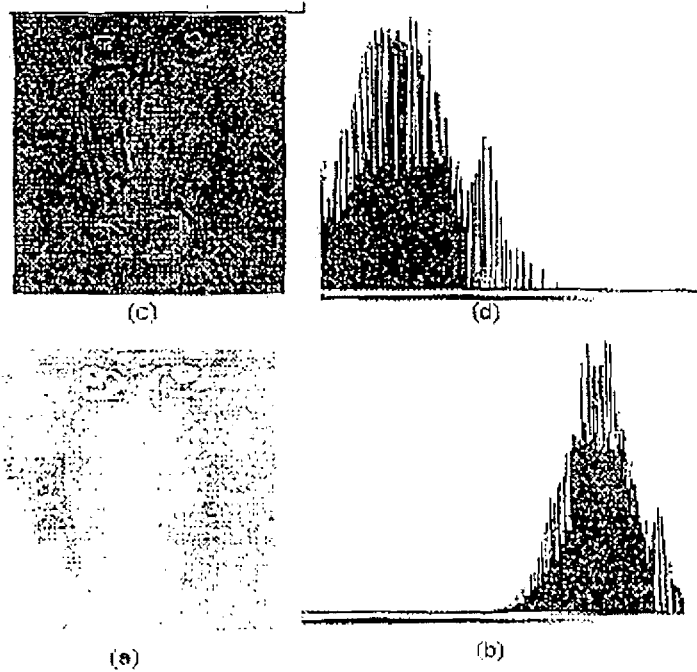


Figura 2.22 Imágenes después de aplicarles el algoritmo antes mencionado.

La figura 2.22a imagen después de la especificación del histograma con valores de $\mu = 200$ $\sigma = 500$ en la función de distribución normal, figura 2.22b histograma de la imagen; figura 2.22c imagen después de la especificación del histograma empleando también la función de distribución normal con valores de $\mu = 50$ $\sigma^2 = 1000$; figura 2.22d su histograma

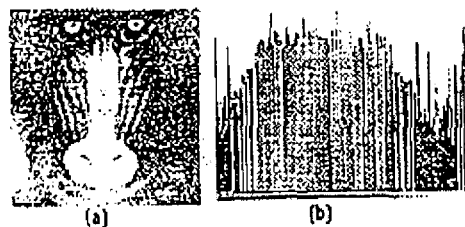


Figura 2.23 (a) imagen después de la equalización del histograma y (b) su histograma

La figura 2.23a muestra el resultado de la ecualización del histograma. La mejora comparada a la imagen original es muy evidente. La figura 2.23b muestra el histograma de la imagen ecualizada. Note que los niveles de gris de una imagen a la cual se le aplica la ecualización del histograma son extendidos y siempre tendiendo hacia el blanco. Este proceso incrementa el rango dinámico de niveles de gris y, consecuentemente, produce un incremento en el contraste de la imagen. En imágenes con histogramas estrechos y relativamente con pocos niveles de gris la ecualización del histograma mejora significativamente la apariencia visual de la imagen. Resultados similares de realce pueden haber sido conseguidos usando ajuste de contraste discontinuos. Claro que la ventaja de la ecualización del histograma sobre las técnicas manuales de manipulación de contraste es que el proceso es completamente automático.

2.2 FILTRADO

INTRODUCCIÓN

El filtrado tiene como objetivo, el realizar una restauración de la imagen o resaltar ciertas características. El filtrado puede ser efectuado en el dominio espacial o en el dominio de la frecuencia.

Las operaciones de filtrado pueden ser divididas en dos clases: filtrado lineal y filtrado no lineal. El filtrado lineal puede efectuarse mediante la operación de convolución, tanto para el dominio espacial como en el dominio de la frecuencia.

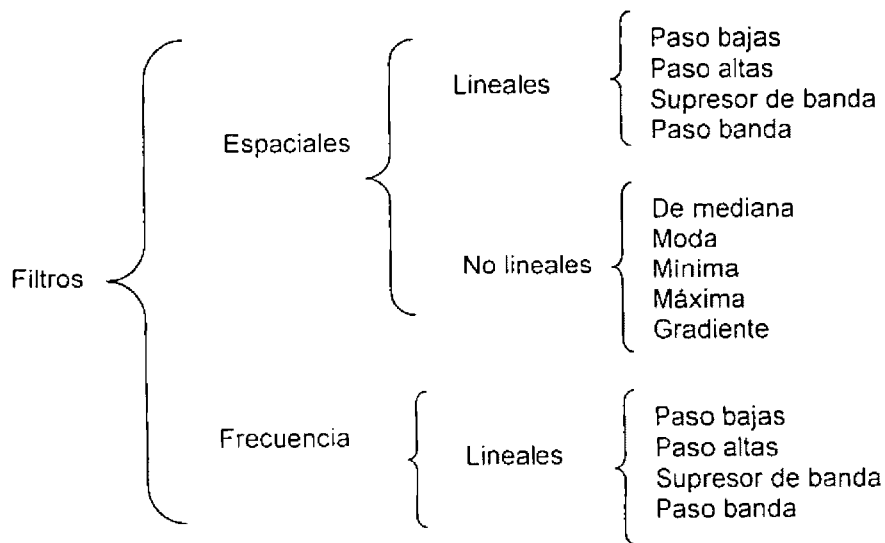
Las frecuencias en una imagen, también llamada frecuencia espacial, son análogas a las frecuencias en una señal en tiempo. Una señal senoidal con alta frecuencia alterna rápidamente mientras que una señal de baja frecuencia cambia lentamente con el tiempo. Similarmente, una imagen con alta frecuencia espacial, presenta cambios frecuentes en la brillantez al pasar de una posición a otra; mientras que una imagen con bajas frecuencias mantiene casi constante un mismo valor de brillantez en áreas extensas de la imagen.

Los filtros lineales más usados son: Los filtros paso bajas que atenúan o eliminan los componentes de alta frecuencia en el dominio de Fourier mientras las bajas frecuencias no son tocadas, esto es que el filtro deja pasar exclusivamente las frecuencias bajas. Los componentes de alta frecuencia caracterizan los bordes y otros detalles de la forma en una imagen, por lo tanto el efecto del filtrado paso bajas es crear imágenes borrosas. De manera similar los filtros paso altas atenúan o eliminan los componentes de bajas frecuencias. Debido a que estos componentes son responsables de pequeñas variaciones características de una imagen, como contraste total e intensidad promedio, el resultado del filtro paso altas es una reducción de estas características y una aparente aclarado de los bordes y otros detalles de la forma. El problema de detección de bordes es el de identificar un cambio súbito del nivel de gris entre dos regiones relativamente homogéneas. El problema de detección de bordes es un problema clásico en procesamiento digital de imágenes y varias formas de abordarlo han sido propuestas. En la mayoría de los casos se realiza un filtro paso altas, seguido del proceso de sacar un umbral, si la salida del filtro pasa el valor de umbral, un borde es detectado. Un tercer tipo de filtrado paso banda, remueve regiones de frecuencia seleccionadas entre bajas y altas frecuencias. Estos filtros son para restauración de imágenes y no son de interés en el realce de imágenes.

Aunque los filtros lineales proporcionan resultados aceptables, se cuenta también con la opción de aplicar a la imagen original un filtro no lineal el cual como su nombre lo indica, no cumple con las condiciones de linealidad (superposición), pero aplica criterios de localización estadística como son la mediana y la moda o criterios de máxima y mínima.

Cabe mencionar que en este trabajo se presentarán los filtros más usados y más conocidos sabiendo que existen otros que no son tan usados.

Clasificación de filtros



2.2.1 FILTRADO ESPACIAL

El contenido de una imagen puede ser revelado únicamente, cuando analizamos las relaciones espaciales de los niveles de gris.

En el filtrado espacial, el nivel de gris de un pixel P después de una transformación, depende del valor del nivel de gris original del pixel y de los pixeles vecinos. En general, los pixeles más próximos a P contribuyen más para el nuevo nivel de éste, a diferencia de los pixeles más apartados.

Las operaciones espaciales se refieren a procesamientos que operan ya no únicamente sobre pixeles, sino ahora sobre renglones de la imagen.

Las funciones de procesamiento de imágenes para operaciones espaciales pueden ser expresadas por la siguiente ecuación

$$g(x, y) = T[f(x, y)]$$

Donde $f(x, y)$ es la imagen de entrada, $g(x, y)$ es la imagen procesada, y T es un operador de, definido sobre alguna vecindad de (x, y) .

La **vecindad** de un pixel (x, y) es una subimagen de área cuadrada que tiene como centro a (x, y) . Por ejemplo los elementos de una vecindad de 3X3, son el pixel P colocado en las coordenadas (x, y) el cual tiene cuatro vecinos verticales y horizontales cuyas coordenadas son $(x+1, y)$, $(x-1, y)$, $(x, y+1)$, $(x, y-1)$ además de tener cuatro vecinos diagonales cuyas coordenadas son $(x+1, y+1)$, $(x+1, y-1)$, $(x-1, y+1)$, $(x-1, y-1)$. Este concepto es mostrado en la figura 2.24.

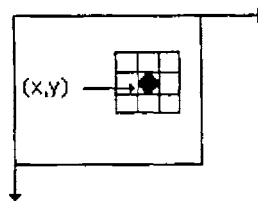


Figura 2.24 La vecindad de 3X3 de un pixel (x, y) en una imagen

Si el valor de gris en una vecindad pequeña no cambia, nos encontramos en un área de valores de gris constantes; esto puede significar que la vecindad está incluida en un objeto. Al contrario si el valor de gris cambia, podemos estar en el borde de un objeto. De esta forma podemos, reconocer áreas de valores de gris constantes y bordes, que tienen como objetivo el mejorar la calidad de la imagen.

Al realizar operaciones sobre vecindades, el centro de la subimagen es movida a través de todos los pixeles de la imagen, comenzando en la esquina superior izquierda, aplicando un operador en cada pixel (x, y) para obtener un nuevo valor g en ese pixel. Aunque existen otras formas de vecindades, como aproximaciones a círculos, los arreglos cuadrados son los más usados debido a su facilidad de implementación.

El uso de vecindades permite una gran variedad de aplicaciones que van más allá del realce de imágenes. No obstante la aplicación específica, la propuesta general es permitir que los valores de f en una vecindad predefinida de (x, y) determinen los valores de g en (x, y) . Uno de los enfoques principales dentro de esta formulación es basada en el uso de las llamadas máscaras, referidas también como temples, ventanas o filtros (ver figura 2.25). Básicamente, una máscara es un pequeño arreglo bidimensional, por ejemplo de 3×3 , en el cual los valores de los coeficientes determinan la naturaleza del proceso, como por el aclarado de imágenes.

El uso de máscaras espaciales en el procesamiento de imágenes se le conoce como **filtrado espacial** y las máscaras a su vez son llamadas filtros espaciales

2.2.1.1 FILTROS LINEALES (espaciales)

El filtrado lineal puede ser representado por la operación de convolución definida por la ecuación:

$$g(i, j) = \sum_{m=i-\frac{w}{2}}^{i+\frac{w}{2}} \sum_{n=j-\frac{w}{2}}^{j+\frac{w}{2}} f(m, n)h(i-m, j-n)$$

donde

$f(m, n)$ = imagen original

$h(m, n)$ = filtro

$g(i, j)$ = imagen filtrada

La aplicación de la operación de convolución bidimensional, en el dominio espacial, se efectúa trabajando con una máscara que se desplaza a través de toda la imagen como lo ilustra la siguiente figura 2.25.

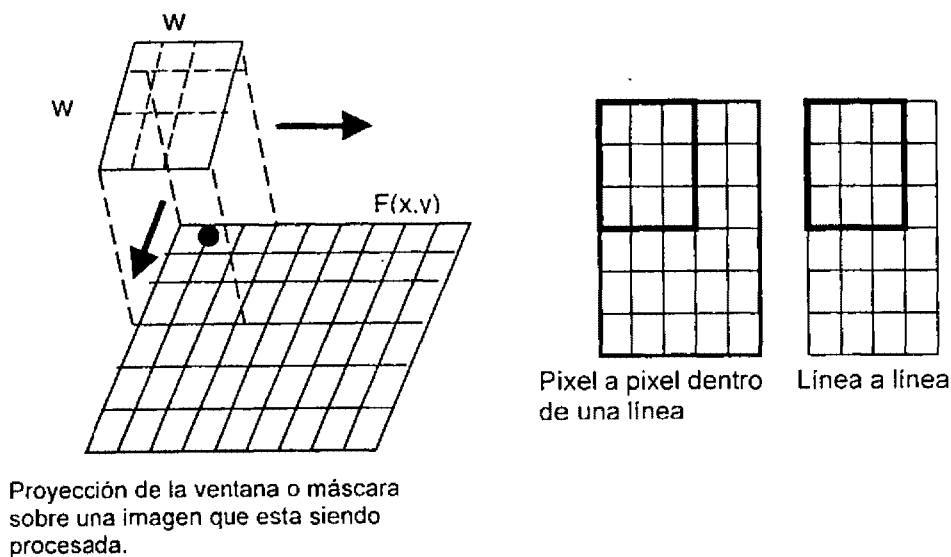


Figura 2.25 Interpretación de la operación de convolución.

Estas máscaras no son otra cosa que una matriz cuyos elementos son multiplicados por los valores de los píxeles de la imagen situados en la posición de la máscara y los resultados sumados. El resultado de la sumatoria se asigna al píxel, de la imagen filtrada, cuya posición coincide con el elemento central de la matriz. La respuesta de la máscara lineal es calculada por la ecuación 2.1

$$g = w_1 f_1 + w_2 f_2 + \dots + w_9 f_9 \quad (2.1)$$

w_1	w_2	w_3
w_4	w_5	w_6
w_7	w_8	w_9

Figura 2.26 Una máscara de 3x3 con coeficientes arbitrarios

En la operación hecha en el dominio del espacio, se deben calcular los puntos pertenecientes a los bordes de la imagen de una manera diferente de la de los demás, ya que éstos no poseen todos sus vecinos. Por cuestiones de simetría se usan, en la definición de los núcleos de los filtros, matrices de $n \times n$, donde n es un número impar. Por cuestiones de eficiencia computacional se prefieren valores pequeños para n (un máximo de 7).

FILTRO PASO BAJAS(Smoothing)

El efecto de un filtro paso bajas es de suavizar la imagen, es decir, las altas frecuencia que corresponden a las transiciones abruptas, son atenuadas. La suavización tiende también, por las mismas razones, a minimizar el efecto del ruido en las imágenes. El filtro paso bajas tiene, por otro lado el efecto de disminuir el contraste de la imagen y la pérdida en la definición de bordes.

La forma de la respuesta al impulso necesitada para implementar un filtro espacial paso bajas indica que el filtro debe tener todos los coeficientes positivos(ver la figura 2.27).

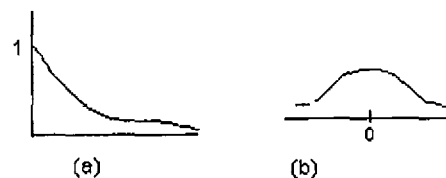


Figura 2.27 (a) Filtro paso bajas en el dominio de la frecuencia
(b) Sección transversal del filtro correspondiente al dominio espacial

Para un filtro paso baja espacial de $n \times n$, el arreglo más simple consiste en una máscara en la cual todos los coeficientes tengan el valor de uno. Sin embargo, la ecuación 2.1, la respuesta sería la suma de las multiplicaciones de los píxeles por nueve coeficientes, lo cual causaría que g sobrepase el rango de niveles de gris permitidos. La solución es escalar la suma dividiendo g entre 9. La figura 2.28a muestra

la máscara resultante. Máscaras más grandes siguen el mismo concepto, como lo muestra la figura 2.28b. Note que en todos los casos, la respuesta g es simplemente el promedio de todos los pixeles en el área de la máscara. Por esta razón, el uso de máscara de la forma mostrada en la figura 2.28 es referida como el promedio de la vecindad. A este tipo de máscaras se les conoce como **filtros paso bajas de media no ponderada** porque a ningún coeficiente de la máscara se multiplica por un factor de peso que depende de la importancia que se le quiera dar al coeficiente.

$$\frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad \frac{1}{25} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

(a) (b)

Figura 2.28 Filtros tipo media no ponderada (pasa bajas)

La figura 2.29 nos muestra un ejemplo de suavizado de una imagen por medio de la aplicación de máscara paso bajas de media no ponderada sucesivamente mayor. Note en particular la pérdida en la definición de bordes.

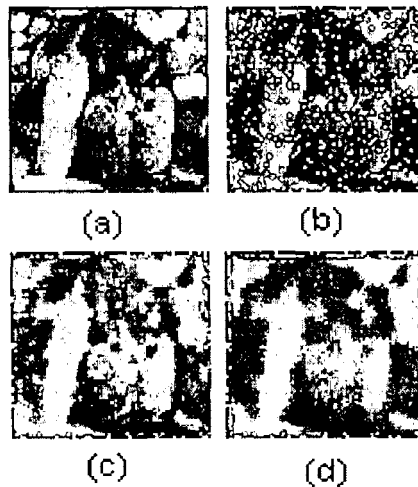


Figura 2.29 (a) Imagen original. (b) Imagen original con 10% de ruido. (c) filtro paso bajas de 3x3. (d) filtro paso bajas de 5x5

FILTRO PASO BAJAS DE MEDIA PONDERADA

A veces se asocia a los coeficientes w_1, w_2, \dots, w_k que pertenecen a la máscara ciertos factores o pesos z_1, z_2, \dots, z_k que dependen de la significación o importancia de cada uno de los coeficientes en este caso:

$$\bar{X} = \frac{z_1 w_1 + z_2 w_2 + \dots + z_k w_k}{z_1 + z_2 + \dots + z_k}$$

donde $k = n \times n$ (número de coeficientes de la máscara)

$$w_1 = 1, w_2 = 1, \dots, w_k = 1$$

Ejemplo de cómo obtener una máscara de media ponderada

$$g = \begin{bmatrix} z_1 w_1 & z_2 w_2 & z_3 w_3 \\ z_4 w_4 & z_5 w_5 & z_6 w_6 \\ z_7 w_7 & z_8 w_8 & z_9 w_9 \end{bmatrix} = \begin{bmatrix} 1*1 & 2*1 & 1*1 \\ 2*1 & 4*1 & 2*1 \\ 1*1 & 2*1 & 1*1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

Al igual que en los filtros paso bajas de media no ponderada. Para evitar que g sobrepase el rango de nivel de gris permitido. La solución es escalar la suma dividiendo g entre la suma de los coeficientes que pertenecen a la máscara resultante. Ver figura 2.30b

$$\frac{1}{10} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix} \qquad \frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

(a) (b)

Figura 2.30 Filtros paso bajas de tipo media ponderada

La figura 2.31 nos muestra un ejemplo de suavizado de una imagen por medio de la aplicación de la máscara paso bajas de media ponderada.

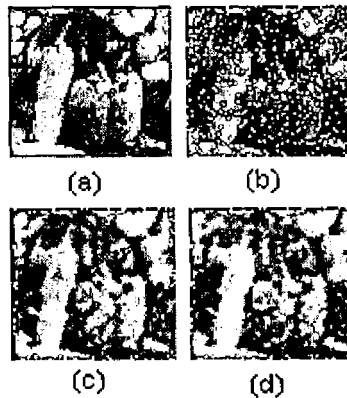


Figura 2.31 (a) Imagen original. (b) Imagen original con 10% de ruido.
 (c) filtro paso bajas ponderada de 3x3 (figura 2.30a).
 (d) filtro paso bajas ponderada de 3x3 (figura 2.30b)

En conclusión, si la suma de los pesos en un núcleo de una máscara es 1, el valor de la imagen resultante será igual a la media de la imagen original. Los filtros que tienen la característica de sacar la media a la imagen son filtros paso bajas.

FILTRO PASO ALTAS (Sharpening)

Los filtros paso altas espaciales tiene el efecto de obtener la agudización de los bordes de la imagen. Las transiciones entre regiones diferentes se toman más nítidas. El efecto indeseado es el enfatizar el ruido existente en la imagen.

La forma de la respuesta al impulso requerida para implementar un filtro espacial paso altas, indica que el filtro debe tener coeficientes positivos cerca del centro y coeficientes negativos en la periferia (ver figura 2.32b).

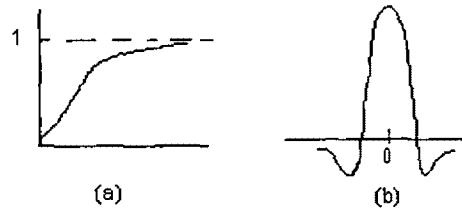


Figura 2.32 (a) Filtro paso altas en el dominio de la frecuencia
(b) Sección transversal del filtro correspondiente al dominio espacial

Debido a que sacar la media suaviza la imagen, es natural esperar que la resta de ésta en la imagen original tenga el efecto de agudizarla. Los filtros de diferenciación son caracterizados por poseer pesos negativos y entonces realizar las diferencias en la imagen y pueden ser representados por la máscaras mostradas en la figura 2.33.

$$\begin{matrix}
 1/4 \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix} &
 1/8 \begin{bmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{bmatrix} &
 1/8 \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix} \\
 \text{(a)} & \text{(b)} & \text{(c)}
 \end{matrix}$$

Figura 2.33 Filtros pasa altas

La figura 2.34 muestra el efecto que tiene los filtros paso altas sobre una imagen, aplicando el operador Laplaciano.

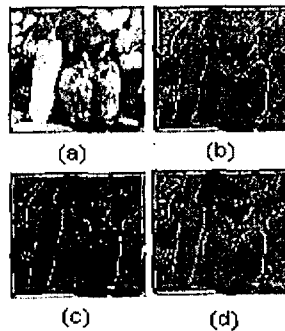


Figura 2.34 (a) Imagen original. (b) filtro paso altas de 3x3 (figura 2.33a)
(c) filtro paso altas de 3x3 (figura 2.33b).
(d) filtro paso altas de 3x3 (figura 2.33c)

FILTRO SUPRESOR DE BANDA

El filtro supresor de banda y paso banda son una extensión de los conceptos de filtro paso bajas y paso altas mencionados anteriormente.

Al aplicar un filtro rechazo banda a nuestra imagen en estudio, el efecto que tendrá sobre esta es: suavizar la imagen, disminuir el contraste y agudizar los bordes. Esto es posible aplicando a nuestra imagen original un filtro paso bajas y al resultado obtenido de este filtraje, un filtrado paso altas. La figura 2.35 muestra el efecto que tiene el filtro supresor de banda en una imagen.

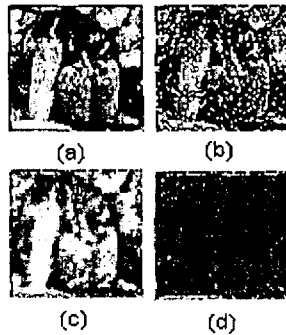


Figura 2.35 (a) Imagen original. (b) Imagen original con 10% de ruido.
(c) filtro paso bajas de 3x3 (máscara ocupada figura 2.30a).
(d) filtro paso altas de 3x3 (máscara ocupada figura 2.33b)

FILTRO PASO BANDA

Los filtros paso banda permiten el paso de una banda o región especificada, mientras atenúa, o suprime completamente todas las otras frecuencias. El filtro paso banda es lo opuesto al rechazo banda. El paso banda obtiene una nitidez entre las regiones de transición, es decir, se resaltan los bordes que posteriormente se suavizaran con un filtro paso bajas, el cual además disminuirá el ruido enfatizado por el filtro paso altas (Ver figura 2.36).

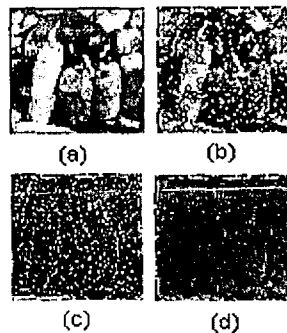


Figura 2.36 (a) Imagen original. (b) Imagen original con 10% de ruido.
(c) filtro paso altas de 3x3 (máscara ocupada figura 2.33b)
(d) filtro paso bajas de 3x3 (máscara ocupada figura. 2.30a).

Por tanto, podemos entender como filtro paso banda. Cuando a una imagen se le aplica un filtro paso altas y al resultado se le aplica un filtro paso bajas.

El teorema de convolución aplicado a los filtros, además, puede ser usado para realzar o detectar ciertas características en la imagen, tales como bordes, líneas, curvas o manchas. Idealmente la sección transversal de un borde presenta una forma de función escalón.

La figura 2.37 muestra máscaras de filtros para el realce de bordes en varios sentidos.

Norte	$\begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ -1 & -1 & -1 \end{bmatrix}$	Noreste	$\begin{bmatrix} 1 & 1 & 1 \\ -1 & -1 & 1 \\ -1 & -1 & 1 \end{bmatrix}$
Este	$\begin{bmatrix} -1 & 1 & 1 \\ -1 & -1 & 1 \\ -1 & 1 & 1 \end{bmatrix}$	Sureste	$\begin{bmatrix} -1 & -1 & 1 \\ -1 & -1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$
Sur	$\begin{bmatrix} -1 & -1 & -1 \\ 1 & -1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$	Suroeste	$\begin{bmatrix} 1 & -1 & -1 \\ 1 & -1 & -1 \\ 1 & 1 & 1 \end{bmatrix}$
Centro	$\begin{bmatrix} 1 & 1 & -1 \\ 1 & -1 & -1 \\ 1 & 1 & -1 \end{bmatrix}$	Noreste	$\begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & -1 \\ 1 & -1 & -1 \end{bmatrix}$

Figura 2.37 Máscaras de filtros de bordes direccionales.

Las líneas en las imágenes pueden ser detectadas a través de los filtros de convolución dados por las máscaras de la figura 2.38. La figura 2.38a muestra un filtro de 3×3 para líneas verticales mientras que el filtro de la figura 2.38b realza líneas horizontales.

De modo análogo lo mostrado en las figuras 2.38a y 2.38b pueden ser construidas máscaras para líneas inclinadas a 45° . Los filtros de la figura 2.38 deben ser usados con cautela pues ellos pueden responder a características que no son líneas, tales como bordes o picos. Una solución a este problema es usar operadores no lineales.

$\begin{bmatrix} -0.5 & 1 & -0.5 \\ -0.5 & 1 & -0.5 \\ -0.5 & 1 & -0.5 \end{bmatrix}$	$\begin{bmatrix} -0.5 & -0.5 & -0.5 \\ 1 & 1 & 1 \\ -0.5 & -0.5 & -0.5 \end{bmatrix}$
(a)	(b)

Figura 2.38 Máscaras para realce de líneas.

2.2.1.2 FILTROS ESPACIALES NO LINEALES

Los filtros espaciales no lineales operan también en vecindades. Sin embargo, las operaciones están directamente en los valores de los píxeles en la vecindad bajo consideración y no usan explícitamente coeficientes.

La detección de características tales como bordes, líneas, curvas y manchas pueden ser hecha también con filtros no lineales.

FILTRO DE MEDIANA

Si nuestro objetivo es reducir ruido sin crear imágenes borrosas, una alternativa es el uso de filtros de mediana. Esto es, los grados de nivel de cada píxel es reemplazado por el valor central de los niveles de gris en una vecindad del píxel, en lugar del promedio. Este método es particularmente efectivo cuando se quiere reducir el ruido, preservando la forma de los bordes.

La mediana de m entre un conjunto de valores es aquel al cual la mitad del conjunto de valores es menor que m y la mitad son mayor que m . Para llevar a cabo el filtrado de mediana en la vecindad de un píxel, primero ordenaremos los valores de los píxeles y sus vecinos, determinamos la mediana, y asignamos este valor al píxel. Por ejemplo, en unas vecindades de 3×3 la mediana es el 5to valor mayor, es una vecindad de 5×5 el 13vo valor mayor y así sucesivamente.

Cuando varios valores en una vecindad son iguales, todos esos valores serán agrupados. Por ejemplo, suponga que una vecindad de 3×3 tiene los valores (10,20,20,20,15,20,20,25,100). Estos valores son ordenados como (10,15,20,20,20,20,20,25,100)., de lo cual nos da como resultado una mediana de 20.

La función principal del filtrado de mediana es forzar los puntos con diferentes intensidades para que se parezcan a sus vecinos, realmente elimina los picos aislados de intensidad que aparecen en el área de la máscara.

El desempeño del filtro de mediana es deficiente cuando el número de píxeles que contienen ruido en la vecindad es mayor que la mitad del número de píxeles en la vecindad.

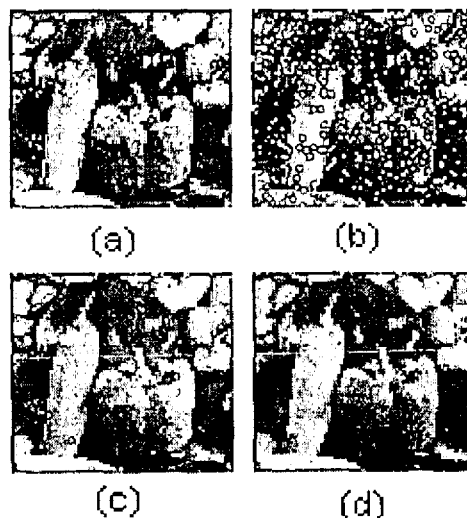


Figura 2.39 (a) Imagen original. (b) Imagen original con 10% de ruido.
(c) Imagen mejorada con un filtro de mediana de 3×3 .
(d) Imagen mejorada con un filtro de mediana de 5×5

FILTRADO DE MÁXIMA Y DE MÍNIMA

Podemos calcular otro tipo de filtrado, en base a la ordenación de los elementos en una vecindad, como tomar los valores mínimos o máximos en la vecindad. El objetivo del filtrado de máxima (con una respuesta $H = \max\{H_k | k = 1, 2, 3, K, 9\}$), es para buscar los puntos más brillantes en una imagen, y el filtro de mínima (con una respuesta $H = \min\{H_k | k = 1, 2, 3, K, 9\}$), en el cual es usado para propósito opuesto.

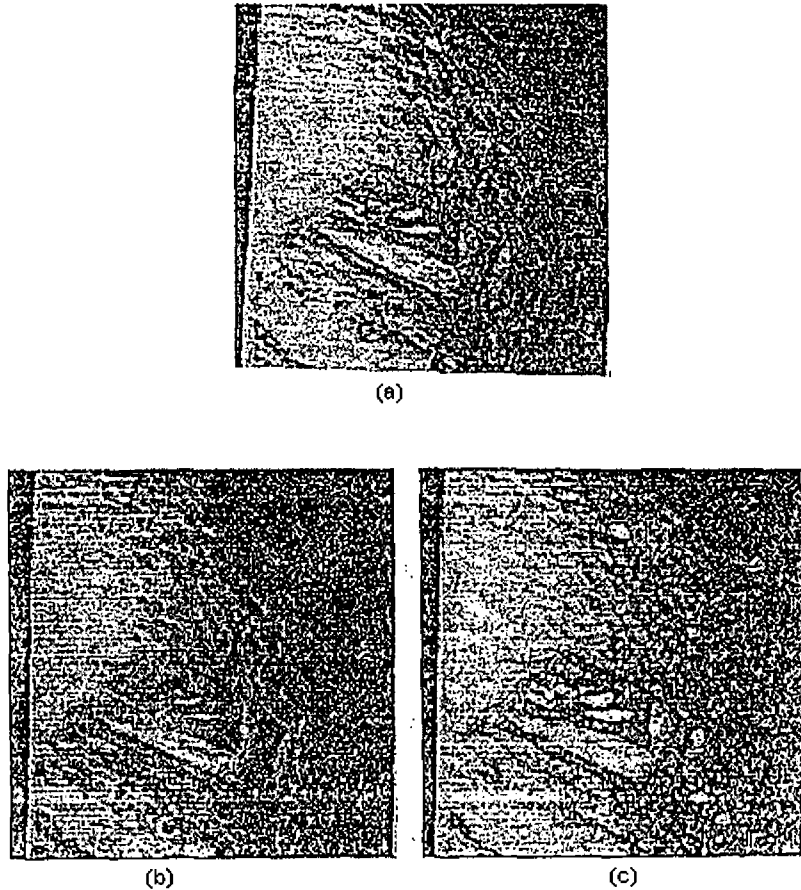


Figura 2.40 (a) Imagen original (b) Aplicación del filtro de mínima.
(c) Aplicación del filtro máxima.

En caso de que la imagen sea binaria, escoger el valor mínimo puede ser obtenido con el operador AND lógico de los valores de la vecindad y, escoger el valor máximo correspondiente se puede hacer mediante el operador OR lógico.

FILTRADO DE MODA

Una alternativa que produce resultados interesantes es tomar el valor más frecuente de una vecindad. Por ejemplo, en unas vecindades de 3x3 que esta constituida por los siguientes valores (10,15,20,29,21,23,20,28,40) nos da como resultado una moda=20.

APLICANDO EL GRADIENTE (Para detección de bordes)

Promediar los píxeles sobre una región tiende a hacer borroso los detalles en una imagen. Como promediar es análogo a la integración, se puede esperar que la diferenciación tenga el efecto opuesto y agudice una imagen.

El método más común de diferenciación en una aplicación de procesamiento de imágenes es el **gradiente**. Para una función $f(x, y)$, el gradiente de f en las coordenadas (x, y) está definido como el vector de la expresión 2.2.

$$\Delta f = \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{pmatrix} \quad (2.2)$$

De análisis vectorial es conocido que el vector gradiente apunta en la dirección del nivel máximo de cambio de f en (x, y) . En la detección de bordes una cantidad importante es la magnitud de este vector, que se calcula con expresión 2.3.

$$\Delta f = \text{mag}(\Delta f) = \left[\left(\frac{\partial f}{\partial x} \right)^2 + \left(\frac{\partial f}{\partial y} \right)^2 \right]^{1/2} \quad (2.3)$$

La ecuación 2.3 es la base para varias aproximaciones a la diferenciación de imágenes. Note de las ecuaciones 2.2 y 2.3, el cálculo del gradiente de una imagen está basado en la obtención de las derivadas parciales $\partial f / \partial x$ y $\partial f / \partial y$ en cada píxel de la imagen. Considere la región de una imagen mostrada en la figura 2.41(a), donde las letras w denotan los valores de los niveles de gris. La ecuación 2.2 puede ser aproximada en el punto w_5 de varias formas. La más simple es usar la diferencia $(w_5 - w_6)$ en la dirección y combinadas como se muestra en la expresión 2.4a

$$\Delta f \approx \left[(w_5 - w_8)^2 + (w_5 - w_6)^2 \right]^{1/2} \quad (2.4a)$$

en lugar de usar cuadrados y raíces cuadradas, podemos obtener resultados similares usando valores absolutos, como en la expresión 2.4b

$$\Delta f \approx |(w_5 - w_8)| + |(w_5 - w_6)| \quad (2.4b)$$

Esta aproximación es muy importante debido a que el cálculo de dos diferencias elevadas al cuadrado y una raíz cuadrada hace muy lenta la obtención de la imagen realzada. Por esta razón se elige la aproximación por medio de la suma de los valores absolutos de las diferencias.

Se puede verificar el error de nuestra aproximación con la expresión 2.4c.

$$\frac{(|w_5 - w_8| + |w_5 - w_6|) - [(w_5 - w_8)^2 (w_5 - w_6)^2]^{1/2}}{[(w_5 - w_8)^2 (w_5 - w_6)^2]^{1/2}} \quad (2.4c)$$

Calculando los errores porcentuales (es decir, el resultado de la expresión 2.4c multiplicado por 100) para valores de w en el intervalo $[0,255]$, que es precisamente el intervalo que comprende la escala de grises, se puede comprobar que los valores de error son variables pero no rebasan el 41.2% de error.

Si se toman los pixeles indicados en la expresión 2.4a para calcular el gradiente en una vecindad de 2×2 se corre el riesgo de no detectar un borde, debido a que puede existir una significativa diferencia con respecto al pixel w_9 , el cual forma parte de la vecindad. Para evitar esta situación podemos obtener mejores resultados al obtener los bordes usando las diferencias cruzadas, como se aprecia en la expresión 2.5a

$$\Delta f \approx [(w_5 - w_9)^2 + (w_6 - w_8)^2]^{1/2} \quad (2.5a)$$

o usando valores absolutos como en la expresión 2.5b

(2.5b)

$$\Delta f \approx |(w_5 - w_9)^2 + (w_6 - w_8)^2|$$

La ecuación 2.5 puede ser implementada tomando los valores absolutos de la respuesta de las dos máscaras mostradas en la figura 2.42b y sumar los resultados. Estas máscaras son llamadas los operadores **gradiente cruzado de Roberts**.

Es difícil implementar máscaras de tamaño mayor. Una aproximación a la ecuación 2.3, en el punto w_5 , pero ahora usando una vecindad de 3×3 , es calculada con la expresión 2.6.

$$\Delta f \approx |(w_7 + w_8 + w_9) - (w_1 + w_2 + w_3)| + |(w_3 + w_6 + w_9) + (w_1 + w_4 + w_7)| \quad (2.6)$$

La diferencia entre el tercer y el primer renglón de 3×3 aproxima la derivada en la dirección x , y la diferencia entre el tercer y la primera columna aproxima el derivativo en la dirección y . La máscara mostrada en la figura 2.42c, llamada **operadores Prewitt**, pueden usarse para implementar la ecuación 2.6. Finalmente, la figura 2.42d muestra otra pareja de máscaras, llamadas operadores de **Sobel** para aproximar la magnitud del gradiente.

$$\begin{pmatrix} w_1 & w_2 & w_3 \\ w_4 & w_5 & w_6 \\ w_7 & w_8 & w_9 \end{pmatrix}$$

(a)

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

(b) Robert. Cuando se aplica una máscara de 2x2

$$\begin{pmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix} \quad \begin{pmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix}$$

(c) Prewitt

$$\begin{pmatrix} -1 & 2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{pmatrix} \quad \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix}$$

(d) Sobel

Figura 2.42 Una región de 3x3 de una imagen (las letras w son los valores de nivel de gris) y varias máscaras usadas para calcular el derivativo en el punto de la etiqueta w_5 . Note que todos los coeficientes de las máscaras suman cero, indicando una respuesta de cero en áreas constantes, como es esperado de un operador derivativo.

Una de las desventajas del operador de Roberts es su anisotropía, o sea, su asimetría. Dependiendo de la dirección, ciertos bordes son más realzados que otros, aunque tengan igual magnitud.

El operador de Sobel tiene la ventaja de diferenciar y dar un efecto de suavizado, lo cual es atractivo porque los derivativos realzan el ruido. Para calcular el gradiente con los operadores mostrados en la figura 2.42 debemos convolucionar por separado ambas máscaras en cada vecindad de la imagen, para después obtener los valores absolutos de las dos convoluciones y finalmente sumarlos, lo que nos da el gradiente del pixel.

2.2.2 FILTRADO EN FRECUENCIA

INTRODUCCIÓN (Transformada de Fourier)

En esta sección vamos a estudiar las consecuencias y adaptaciones necesarias para poder usar la transformada de Fourier sobre una imagen digital en lugar de sobre una función continua. El primer punto a tener en cuenta es la relación existente entre una función continua y una imagen digital. Una imagen digital se obtiene a partir de una señal continua por un proceso de muestreo. Así pues, habrá que estudiar el efecto que este proceso de muestreo tiene sobre el cálculo de la transformada de Fourier.

Una imagen digital puede mirarse como el resultado de efectuar sobre una señal continua un proceso de muestreo en dos direcciones perpendiculares con un paso o distancia entre muestras que se fija para cada dirección. La siguiente gráfica de la figura 2.43 muestra este proceso de muestreo sobre una función de una variable

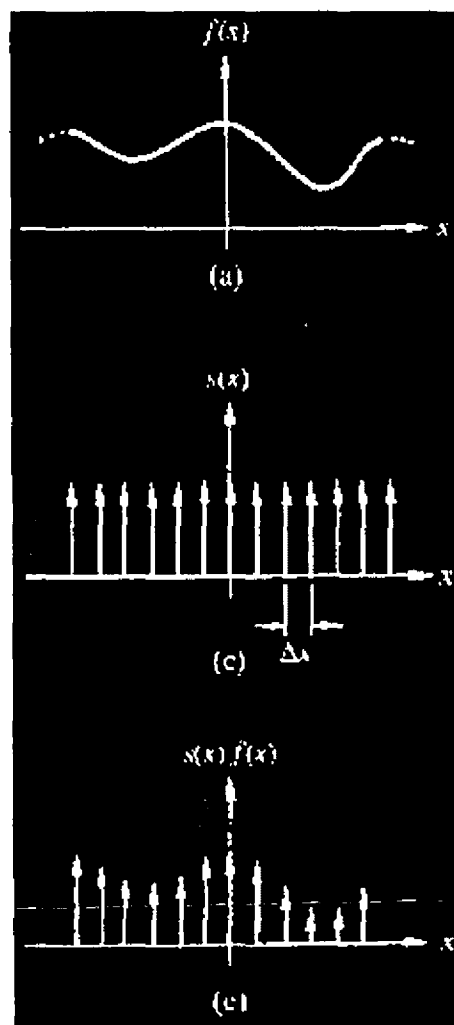


Figura 2.43 a) Señal continua unidimensional
b) Muestreador con un incremento Δx c) Señal Muestreada.

Una representación intuitiva de este efecto viene dado si consideramos a la imagen digital como el resultado de superponer una malla rectangular sobre la función y considerar que los valores de los pixeles de la imagen son una función de los valores de la señal en el interior de cada rectángulo definido por la malla. Funciones típicas del caso bidimensional, son aquellas que asignan a cada pixel la media de los valores de la señal en el interior de cada rectángulo o aquellas que asignan como valor representativo de cada rectángulo el valor en su centro. Con cualesquiera de los criterios anteriores es evidente que se producirá una pérdida de información ya que resumimos en un punto el comportamiento de la señal en una región del espacio.

Ya que nuestro estudio de la transformada de Fourier (TF) esta orientada al estudio de imágenes digitales, el resultado de aplicar esta transformación sobre una imagen debe ser otra imagen de igual dimensión. Pero el rango de valores de los pixeles habrá cambiado de forma muy notable. En la imagen original el valor de un pixel será un número entero entre $[0 - 255]$ representando el grado de brillo, pero en la imagen que representa el espectro de Fourier los valores de los pixeles son números reales no negativos con un rango de valores mucho mayor que 255. Entonces para poder representar adecuadamente estos valores introduciremos la siguiente transformación sobre los valores del espectro de Fourier

$$D(u, v) = k * \log(1 + |F(u, v)|)$$

Donde k es una constante de proporcionalidad que se fija en función de los valores del logaritmo. Es evidente que el uso del logaritmo sobre los valores de $|F(u, v)|$ reduce adecuadamente su rango dinámico.

La ecuación anterior facilita mucho la interpretación visual del espectro, de la transformada de Fourier. Un ejemplo de la obtención del espectro se observa en la figura 2.44, donde muestra una imagen del planeta Saturno, su espectro normal, y el espectro cuando se le suma uno a $|f(u, v)|$.

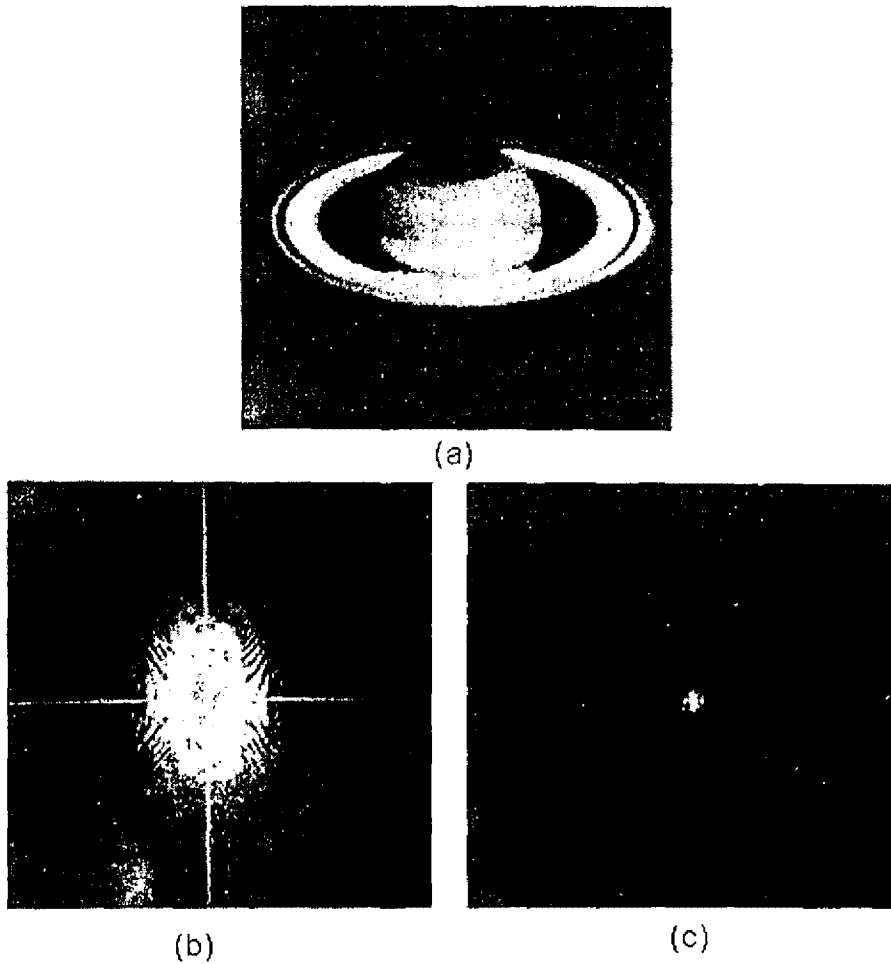


Figura 2.44. (a) Muestra una imagen del planeta Saturno.
 (b) Imagen resultante después de aplicar $|F(u, v)|$.
 (c) Imagen resultante después de aplicar $\log(1 + |F(u, v)|)$

2.2.2.1 TRANSFORMADA DE FOURIER (Caso Bidimensional)

La operación que nos permite pasar del dominio espacial al de las frecuencias es la transformada de Fourier, mientras que la transformada inversa de Fourier pasa del dominio de la frecuencia al dominio espacial.

Una función puede ser vista como la función discreta $f(x, y)$ del campo de radiación electromagnética de una región de la tierra. Dicha función se encuentra definida sobre una malla rectangular equiespaciada de muestra dada por

$$m(x_0 + x\Delta x, y_0 + y\Delta y)$$

Para $x = 0, 1, 2, \dots, M-1$, $y = 0, 1, 2, \dots, N-1$, siendo (x_0, y_0) el punto inicial de la discretización, Δx el desplazamiento entre muestras en la dirección x y Δy el desplazamiento entre muestras en la dirección y , como $x_0, y_0, \Delta x, \Delta y$ son constantes se puede representar la imagen como.

$$f(x, y) = m(x_0 + x\Delta x, y_0 + y\Delta y)$$

Considerando esto, se tiene que la transformada discreta de Fourier bidimensional que se aplica a funciones muestreadas está definida por la ecuación

$$F(u, v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \exp \left[-j2\pi \left(\frac{ux}{M} + \frac{vy}{N} \right) \right]$$

para $u = 0, 1, 2, \dots, M-1$, $v = 0, 1, 2, \dots, N-1$

La ecuación que define la transformada discreta inversa de Fourier bidimensional es

$$f(x, y) = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) \exp \left[j2\pi \left(\frac{ux}{M} + \frac{vy}{N} \right) \right]$$

para $x = 0, 1, 2, \dots, M-1$, $y = 0, 1, 2, \dots, N-1$

Los valores $u = 0, 1, 2, \dots, M-1$, y $v = 0, 1, 2, \dots, N-1$ en la transformada discreta de Fourier corresponde a muestras de la transformada continua dentro de una malla equiespaciada dada por

$$f(x, y) = F(u\Delta u, v\Delta v)$$

Donde Δu , Δv son los desplazamientos entre muestras y se mantienen constantes.

Debe notarse que las muestras de $F(u, v)$ comienzan en el origen del eje de las frecuencias. Los incrementos de muestreo en el dominio espacial y de frecuencias están relacionados por

$$\Delta u = \frac{1}{M\Delta x}, \quad \Delta v = \frac{1}{N\Delta y}$$

Cuando las imágenes son muestreadas en un arreglo cuadrado, entonces $M = N$ y

$$F(u, v) = \frac{1}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \exp \left[\frac{-2j\pi(ux + vy)}{N} \right]$$

para $u, v = 0, 1, 2, \dots, N-1$, y

$$f(x, y) = \frac{1}{N} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} F(u, v) \exp \left[\frac{2j\pi(ux + vy)}{N} \right]$$

para $x, y = 0, 1, 2, \dots, N-1$

Nótese que ahora se ha incluido el término $1/N$ en ambas expresiones.

Debido a que $F(u, v)$ y $f(x, y)$ son un par de ecuaciones que siempre se emplean juntas, la agrupación de estas constantes multiplicativas ($1/N$) es arbitraria (Gonzalez and Wintz, 1985). En la práctica normalmente se trata de emplear imágenes cuadradas para hacer uso de la transformada de Fourier.

Ahora bien, la transformada de Fourier de una función real es generalmente compleja, por lo que

$$F(u, v) = R(u, v) + jI(u, v)$$

Donde $R(u, v)$ y $I(u, v)$ son, respectivamente, las componentes real e imaginaria de $F(u, v)$. Todo número complejo puede expresarse en forma exponencial por lo que

$$F(u, v) = |F(u, v)| e^{j\Phi(u, v)}$$

donde

$$|F(u, v)| = [R^2(u, v) + I^2(u, v)]^{1/2}$$

$$\Phi(u, v) = \arctan \left[\frac{I(u, v)}{R(u, v)} \right]$$

La función de magnitud $|F(u, v)|$ es llamada el espectro de Fourier de $f(x, y)$, y $\Phi(u, v)$ es el ángulo de fase. El cuadrado del espectro es comúnmente referido como el espectro de potencia o de densidad espectral y se denota como, $P(u, v)$. Las variables (u, v) son llamadas variables de frecuencia. Este nombre surgió del hecho de usar la formula de Euler, el termino exponencial, $\exp[-j2\pi uvx]$, debe ser expresado en la forma: $\exp[-j2\pi uvx] = \cos 2\pi uvx - j \text{sen } 2\pi uvx$

$$P(u, v) = |F(u, v)|^2$$

$$P(u, v) = R^2(u, v) + I^2(u, v)$$

Es práctica común desplegar el espectro de Fourier $|F(u, v)|$ en forma de imagen, como una función de intensidad para visualizar la función transformada. Desgraciadamente, muchas veces el espectro decrece muy rápidamente y entonces, las altas frecuencias tienen tendencias de permanecer obscuras cuando son desplegadas en forma de imagen. Una técnica que compensa esta dificultad es efectuar el despliegue mediante la siguiente función

$$D(u, v) = \log(1 + |F(u, v)|)$$

2.2.2.2 TRANSFORMADA RAPIDA DE FOURIER (Caso Unidimensional)

La transformada unidimensional es efectuada a través del algoritmo denominado transformada rápida de Fourier (Fast Fourier Transform, FFT). La Transformada unidimensional de Fourier es:

$$f(x) = \frac{1}{N} \sum_{u=0}^{N-1} F(u) \exp\left[\frac{-2j\pi ux}{N}\right]$$

El número de multiplicaciones complejas y sumas requeridas para implementar la transformada unidimensional de Fourier es N^2 . Esto puede ser visto fácilmente a través de notar que, para cada uno de los N valores de u , la expresión de la sumatoria requiere N multiplicaciones complejas de $f(x)$ por $\exp[-2j\pi ux/N]$ y $(N-1)$ sumas de los resultados. El término $\exp[-2j\pi ux/N]$ puede ser calculado una vez y almacenado en una tabla para todas las subsecuentes aplicaciones. Por esta razón, la multiplicación de u por x es generalmente no considerada una parte directa de la implementación.

El algoritmo de FFT está basado en el método de "dobles sucesivos" (Successive doubling). Para llevar a cabo el desarrollo del método es conveniente expresar la transformada unidimensional en la forma

$$F(u) = \frac{1}{N} \sum_{x=0}^{N-1} f(x) W_N^{ux}$$

donde

$$W_N = \exp\left[\frac{-2j\pi}{N}\right]$$

Asumiendo que N es igual a una potencia entera de 2, $N = 2^n$, donde n es un entero positivo. Reescribiendo la función de la transformada

$$F(u) = \frac{1}{2M} \sum_{x=0}^{2M-1} f(x) W_{2M}^{ux} = \frac{1}{2} \left[\frac{1}{M} \sum_{x=0}^{M-1} f(2x) W_{2M}^{u(2x)} + \frac{1}{M} \sum_{x=0}^{M-1} f(2x+1) W_{2M}^{u(2x+1)} \right]$$

De la definición de W_N se tiene

$$W_{2M}^{2ux} = W_M^{ux}$$

por lo que la expresión toma la forma

$$F(u) = \frac{1}{2} \left[\frac{1}{M} \sum_{x=0}^{M-1} f(2x) W_M^{ux} + \frac{1}{M} \sum_{x=0}^{M-1} f(2x+1) W_M^{ux} W_{2M}^u \right]$$

si se define

$$F_{par}(u) = \frac{1}{M} \sum_{x=0}^{M-1} f(2x) W_M^{ux}$$

para $u = 0, 1, 2, K, M - 1$, y

$$F_{non}(u) = \frac{1}{M} \sum_{x=0}^{M-1} f(2x+1)W_M^{ux}$$

para $u = 0, 1, 2, K, M - 1$,

Debe notarse que las funciones $F_{par}(u)$ y $F_{non}(u)$ presentan la misma forma de la transformada de Fourier $F(u)$, pero que en la primera únicamente se consideran los puntos pares del juego de las muestras y en la segunda los puntos nones. Por esta razón se tiene que $F_{par}(u)$ es la transformada de Fourier de un subjuego de muestras que considera únicamente los puntos pares y $F_{non}(u)$ es la transformada de Fourier de los puntos nones. De las definiciones anteriores la transformada de Fourier se puede escribir como:

$$F(u) = \frac{1}{2} [F_{par}(u) + F_{non}(u)W_{2M}^u]$$

Esto es, la transformada de Fourier $F(u)$ puede ser obtenida a través de la transformada de Fourier de sus muestras pares $F_{par}(u)$ y nones, $F_{non}(u)$.

Evaluando la transformada en la posición $u + M$, de la definición de W_N se tiene que

$$W_M^{u+M} = W_M^u \text{ y } W_{2M}^{u+M} = -W_{2M}^u$$

La transformada en la posición $u + M$ es

$$F(u + M) = \frac{1}{2} [F_{par}(u) - F_{non}(u)W_{2M}^u]$$

De las expresiones obtenidas se ve que la transformada $F(u)$ para una función muestreada de N puntos puede ser calculada dividiendo la ecuación original en dos partes, una transformada par $F_{par}(u)$ y otra non $F_{non}(u)$. El cálculo de la primera mitad de $F(u)$ requiere la evaluación de las transformadas $F_{par}(u)$ y $F_{non}(u)$, las cuales emplean únicamente $N/2$ puntos. Los valores de $F_{par}(u)$ y $F_{non}(u)$ son sustituidos en la ecuación que define $F(u)$ para $u = 0, 2, K, (N/2 - 1)$. La otra mitad es obtenida directamente de la ecuación que da $F(u + M)$ sin hacer cálculos adicionales a los realizados para la primera mitad.

La implementación de las ecuaciones para $F(u)$, $F(u + M)$, $F_{par}(u)$ y $F_{non}(u)$ constituyen el algoritmo de FFT por dobleces sucesivos. El nombre viene del hecho de que la transformada de una función de dos puntos es calculada a partir de dos transformadas de un punto, la transformada de una función de 4 puntos es generada a partir de dos transformadas de 2 puntos, así sucesivamente, para cualquier N que es una potencia entera de 2.

El principal punto es tener en mente en la implementación del algoritmo de la FFT, es que los datos de entrada deben estar arreglados en el orden requerido para una sucesiva aplicación de las ecuaciones de $F_{par}(u)$ y $F_{non}(u)$ (Gonzalez Wintz, 1987).

El procedimiento puede ser ilustrado mediante el siguiente ejemplo. Supóngase que se desea obtener la transformada de una función de 8 puntos $\{f(0), f(1), f(2), f(3), f(4), f(5), f(6), f(7)\}$. La transformada par $F_{par}(u)$ usa las muestras con los argumentos pares, $\{f(0), f(2), f(4), f(6)\}$, y $F_{non}(u)$ usa las muestras con los argumentos nones, $\{f(1), f(3), f(5), f(7)\}$. Ahora bien, para obtener la transformada de cada subjuego de 4 muestras se requiere también hacer uso de las ecuaciones de $F_{par}(u)$ y $F_{non}(u)$. Así para el primer subjuego sus muestras pares son $\{f(0), f(4)\}$ y nones $\{f(2), f(6)\}$; y para el segundo subjuego las pares son $\{f(1), f(5)\}$ y nones $\{f(3), f(7)\}$. Ninguna nueva subdivisión es requerida pues cada subjuego de 2 muestras posee únicamente una muestra par y una non.

Combinando los resultados anteriores el arreglo de entrada debe ser $\{f(0), f(4), f(2), f(6), f(1), f(5), f(3), f(7)\}$. El algoritmo de dobles sucesivos opera sobre este arreglo considerando inicialmente 4 transformadas de 2 puntos que son $\{f(0), f(4)\}$, $\{f(2), f(6)\}$, $\{f(1), f(5)\}$ y $\{f(3), f(7)\}$. Posteriormente utiliza estos resultados para calcular las transformadas de los subjuegos de 4 puntos y finalmente emplea los resultados de éstas para producir la transformada deseada.

El calculo de la transformada inversa se hace mediante la siguiente consideración. Si la transformada discreta inversa de Fourier unidimensional esta definida por

$$f(x) = \sum_{u=0}^{N-1} F(u) \exp \left[\frac{j2\pi ux}{N} \right]$$

Tomando el complejo conjugado y dividiendo ambos lados por N

$$\frac{1}{N} f^*(x) = \frac{1}{N} \sum_{u=0}^{N-1} F^*(u) \exp \left[\frac{-j2\pi ux}{N} \right]$$

Comparando esta ecuación con la definición de la transformada de Fourier directa que se va de lado derecho presenta la misma forma. Si se mete el complejo conjugado $F^*(u)$ dentro de un algoritmo diseñado para calcular la transformada directa, el resultado será $f^*(x)/N$. Tomando el complejo conjugado y multiplicando por N se obtendrá la transformada inversa de $f(x)$. Ahora bien, si $f(x) = f^*(x)$.

Para el caso bidimensional el complejo conjugado es:
100

$$f^*(x, y) = \frac{1}{N} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} F^*(u, v) \exp \left[\frac{j2\pi(ux + vy)}{N} \right]$$

2.2.2.3 PROPIEDADES DE LA TRANSFORMADA DE FOURIER

TRASLACIÓN

Estudiaremos ahora la relación que existe entre la TFD (Transformada Directa de Fourier) de una sucesión de valores y la TFD de una traslación de dicha sucesión.

Si en el par de ecuaciones que define la TFD tomamos para el caso directo dos valores correctos de las variables u y v que notamos por u_0 y v_0 tenemos que

$$\frac{1}{N^2} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \exp \left[\frac{-i2\pi(u_0 x + v_0 y)}{N} \right] = F(u - u_0, v - v_0)$$

de igual manera podemos ver

$$\sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(u, v) \exp \left[\frac{-i2\pi(u_0 x + v_0 y)}{N} \right] = f(x - u_0, y - v_0)$$

De las expresiones anteriores se deduce que fijando los correspondientes valores de la exponencial se produce una traslación de valores de la transformación correspondiente. En muchas ocasiones y con objeto de realizar el cálculo de la TDF (Transformada Directa de Fourier) de forma más eficiente nos interesará el caso de $u_0 = v_0 = N/2$. En este caso lo que hacemos es considerar el origen de la TDF en el centro del cuadrado que define la TDF de la función $f(x, y)$. La figura 2.45, muestra los efectos de la traslación del espectro de Fourier al centro del cuadrado. Es importante hacer notar que una traslación en los valores de la función $f(x, y)$ no afecta las magnitudes de su TFD.

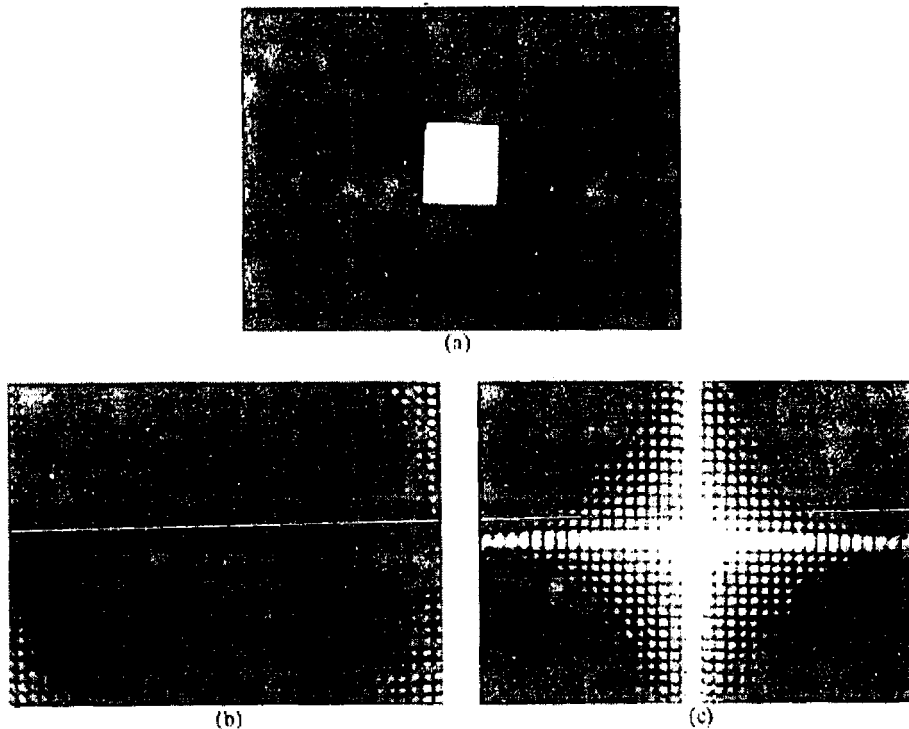


Figura 2.45 Efectos de la traslación del espectro de Fourier al centro del cuadrado

ESTA TESIS NO DEBE SALIR DE LA BIBLIOTECA

PERIODICIDAD Y SIMETRÍA

Del par de ecuaciones que definen la TFD (Transformada Directa de Fourier) puede verse fácilmente que la TFD de una función cualesquiera es una función periódica con periodo N , es decir

$$F(u, v) = F(u + N, v) = F(u, v + N) = F(u + N, v + N)$$

De esta propiedad se deduce fácilmente que podemos caracterizar una función en el dominio de frecuencia especificando solamente los N^2 valores que define su periodo. Entonces si consideramos los valores de un periodo de $F(u, v)$ y le aplicamos la TFD inversa veremos que podemos obtener la misma propiedad de periodicidad en el dominio espacial sobre la función. La principal conclusión que obtenemos de esta propiedad es que la TFD considera que la función es periódica y la muestra de N^2 valores representa un periodo completo. En el caso de señales con valores reales como son las imágenes puede demostrarse que

$$|F(u, v)| = |F(-u, v)| = |F(u, -v)| = |F(-u, -v)|$$

Lo que implica una simetría de los valores del espectro de Fourier de la función respecto del origen de coordenadas. Si, además, tenemos en cuenta la propiedad de periodicidad establecida antes nos encontramos que el espectro de Fourier de una función real además de ser periódico es una función que tiene tan solo $N/2$ valores distintos en cada dimensión. La figura 2.46 muestra los efectos señalados de periodicidad.

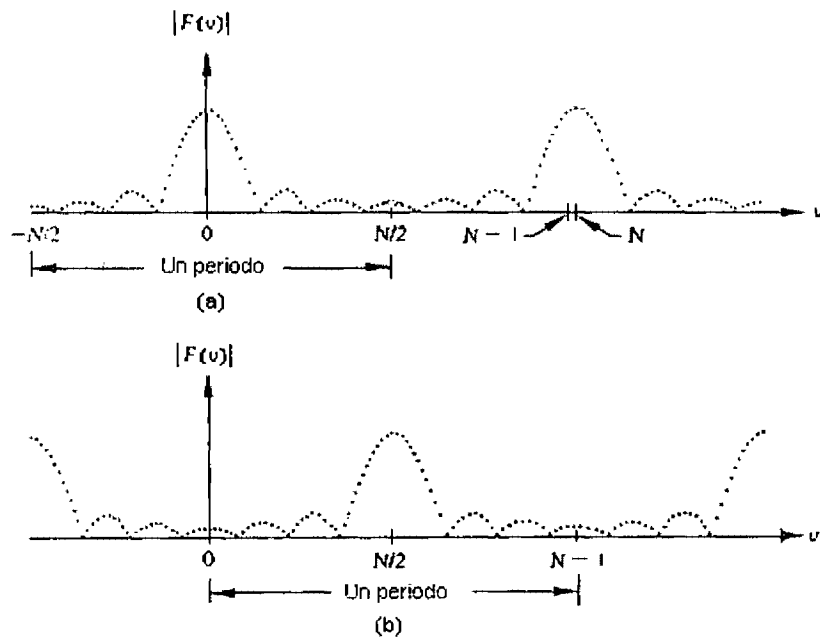


Figura 2.46 Propiedad de Periodicidad y Simetría

ROTACIÓN

Podemos demostrar que si rotamos una función $f(x, y)$ con un ángulo Φ_0 la transformada de Fourier de dicha función se rota por el mismo ángulo. Este resultado es también verdad si comenzamos rotando la función $F(u, v)$. La figura 2.47 muestra los resultados de la rotación

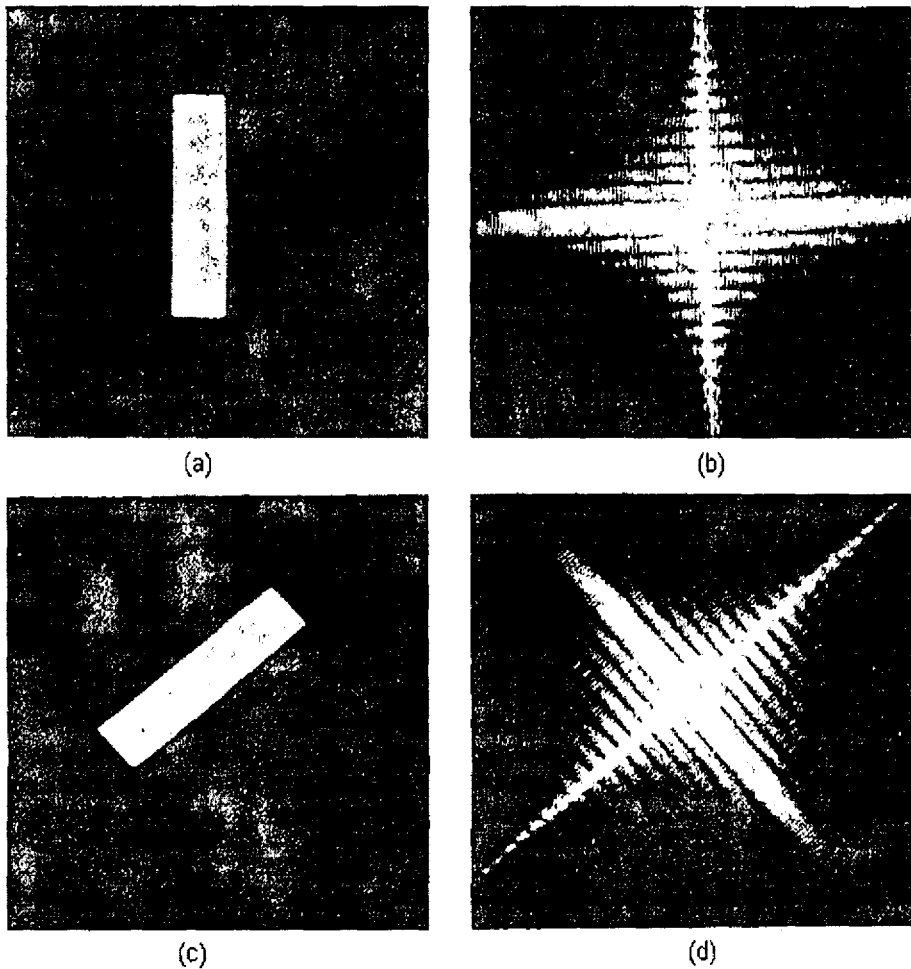


Figura 2.47 Propiedad de Fourier de Rotación

LINEALIDAD Y ESCALA

La transformada de Fourier considerada como una transformación de funciones $F(f)$ es una transformación lineal, es decir

$$F(af_1 + bf_2) = a * F(f_1) + b * F(f_2)$$

Donde a y b son constantes dadas. Puede asimismo demostrarse que la transformada de Fourier de la

función $f(ax, by)$ es igual a $\frac{F(u/a, v/b)}{|ab|}$

INTRODUCCIÓN A FILTROS EN EL DOMINIO DE LA FRECUENCIA

En todos los sistemas hay una función de entrada (o función de excitación) y una función de salida (o función de respuesta). Un sistema está completamente caracterizado si se conoce la naturaleza de la dependencia de la salida sobre la entrada.

Si se supone que la respuesta de un sistema a la excitación $f_i(t)$ es la función $f_o(t)$, y si la respuesta de ese sistema a la excitación $f_i(t) = a_1 f_{i1}(t) + a_2 f_{i2}(t)$ es $f_o(t) = a_1 f_{o1}(t) + a_2 f_{o2}(t)$ se dice que es un sistema lineal.

Por lo tanto, un sistema lineal se puede definir como un sistema al cual se le puede aplicar el principio de superposición que establece: la respuesta de un sistema lineal producida por varias excitaciones actuando simultáneamente es igual a la suma de las respuestas que produce en el sistema cada excitación actuando por separado.

Los filtros lineales están basados en el concepto que establece que la respuesta al impulso es igual a la transformada inversa de Fourier de la función de transferencia del sistema.

A continuación se presentan las descripciones de los siguientes filtros en el orden de aparición:

- ◆ Paso Bajas
- ◆ Paso Altas
- ◆ Supresor de Banda
- ◆ Paso Banda

2.2.2.4 FILTRO PASO BAJAS

Los bordes y otros tipos de transiciones; tales como ruido, líneas y fronteras; en los niveles de gris de una imagen contribuyen fuertemente al contenido de altas frecuencias de una transformada de Fourier. Un desvanecimiento puede ser efectuado por medio del dominio de la frecuencia, atenuando un rango específico de componentes de alta frecuencia en la transformada de una imagen, a través de la transformada dada por:

$$G(u, v) = H(u, v)F(u, v)$$

Donde $F(u, v)$ es la transformada de la imagen que se desea suavizar y $H(u, v)$ es llamada función de transferencia o filtro. El problema es seleccionar una función $H(u, v)$ que produzca $G(u, v)$ a través de atenuar la componente de alta frecuencia de $F(u, v)$. La transformada inversa de $G(u, v)$ producirá entonces la imagen suavizada deseada $g(x, y)$. Las componentes de alta frecuencia son filtradas y la información en el rango de bajas frecuencias es "pasada" sin atenuación, este método es comúnmente referido como filtrado paso bajas.

Un filtro paso bajas ideal bidimensional (ILPF) es aquel que cada función de transferencia satisface la relación

$$H(u, v) = \begin{cases} 1, & \text{si } D(u, v) \leq D_0 \\ 0, & \text{si } D(u, v) > D_0 \end{cases}$$

Donde D_0 es una distancia umbral no negativa, y $D(u, v)$ es una distancia del punto (u, v) al origen del plano de frecuencias; esto es

$$D(u^2 + v^2)^{1/2}$$

La figura 2.48 muestra el filtro esquemáticamente. El nombre del filtro ideal viene del hecho de que todas las frecuencias dentro de un círculo de radio D_0 son pasadas sin ninguna atenuación, mientras que todas las frecuencias fuera de este son completamente atenuadas.

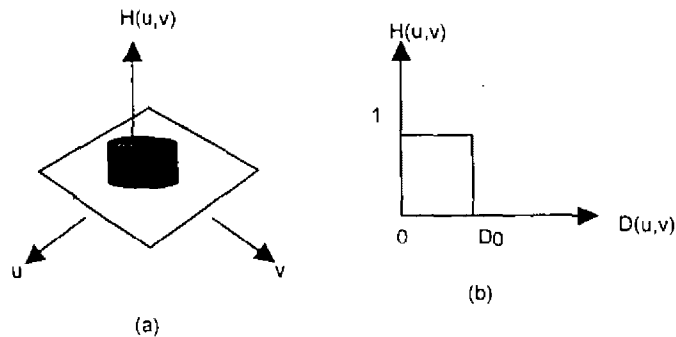


Figura 2.48 Filtro paso bajas ideal, ILPF. (A) Gráfica en perspectiva. (b) Sección transversal (tomada de Gonzalez and Wintz 1987).

El ILPF (filtro paso bajas ideal) es simétrico radialmente a partir del origen. Esta simetría radial está basada en la suposición de que el origen de la transformada de Fourier haya sido centrado en el cuadrado aprovechando las propiedades de la periodicidad y simetría de la transformada de Fourier vista anteriormente.

Para un filtro paso bajas ideal el punto de transición entre $H(u, v) = 1$ y $H(u, v) = 0$ es llamada la **frecuencia de corte**.

Un filtro ampliamente usado para suavizar imágenes es el filtro paso bajas Butterworth (BLPF). La función de transferencia del BLPF de orden n es

$$H(u, v) = \frac{1}{1 + \left[\frac{D(u, v)}{D_0} \right]^{2n}}$$

Donde $D(u, v) = (u^2 + v^2)^{1/2}$ es la distancia del punto (u, v) al origen y D_0 es la distancia a la que se encuentra localizada la frecuencia de corte. La figura 2.49 muestra un ejemplo de la aplicación de un filtro paso bajas butterworth sobre una imagen.

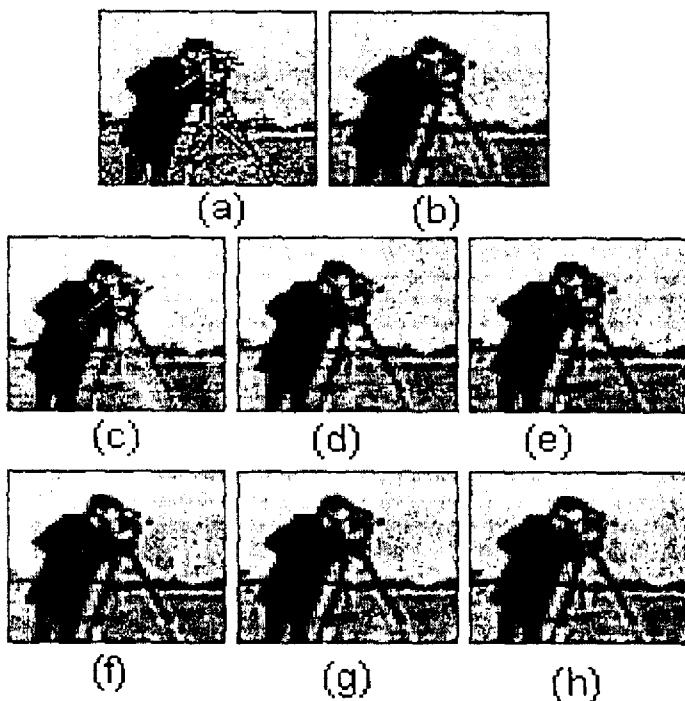
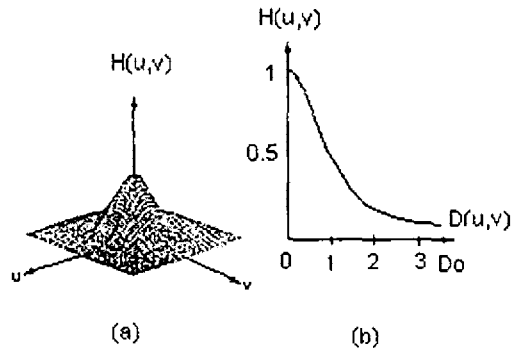


Figura 2.49 (a) imagen original. (b) Imagen filtrada con un filtro simétrico circular "ideal"
 (c) Filtro butterworth primer orden, (d) Filtro butterworth segundo orden
 (e) Filtro butterworth tercer orden, (f) Filtro butterworth cuarto orden
 (g) Filtro butterworth quinto orden, (h) Filtro butterworth sexto orden

A diferencia del ILPF (Filtro paso bajas ideal), la función de transferencia del BLPF (Filtro paso bajas butterworth) no tiene una fuerte discontinuidad que establezca claramente el corte entre las frecuencias que son pasadas y aquellas que se filtran. Para este tipo de filtros es costumbre definir una frecuencia de corte localizada en el punto para el cual $H(u, v)$ desciende hasta una cierta fracción de su máximo valor. Como se ve en la figura 2.50 el filtro BLPF (Filtro paso bajas butterworth) de orden $n = 1$ desciende un 50% de su máximo, esto es, $H(u, v) = 0.5$, cuando $H(u, v) = D_0$. Otro valor comúnmente usado es $1/2^{1/2}$ del valor máximo de $H(u, v)$.



2.50 Filtro paso bajas Butterworth, BLPF. a) Gráfica en perspectiva
b) Sección transversal.

2.2.2.5 FILTRO PASO ALTAS

Debido a que los bordes y otros cambios abruptos en los niveles de gris están asociados con los componentes de altas frecuencias, la agudización de una imagen puede ser llevada a cabo en el dominio de las frecuencias a través de un filtrado paso altas, el cual atenúa las componentes de bajas frecuencias sin alterar la información de alta frecuencia en la transformada de Fourier.

El IHPF (filtro paso altas ideal) está definido como

$$H(u,v) = \begin{cases} 0, & \text{si } D(u,v) \leq D_0 \\ 1, & \text{si } D(u,v) > D_0 \end{cases}$$

Donde D_0 es la distancia que se encuentra localizada en la frecuencia de corte y $D(u,v)$ es la distancia del punto (u,v) al origen del plano de frecuencias. La figura 2.51 da una gráfica de la función de transferencia del IHPF (filtro paso alta ideal). Nótese que este filtro es exactamente opuesto al filtro ideal paso bajas, por lo que este atenúa completamente todas las frecuencias dentro del círculo de radio D_0 , mientras deja pasar sin atenuación todas las frecuencias fuera del círculo.

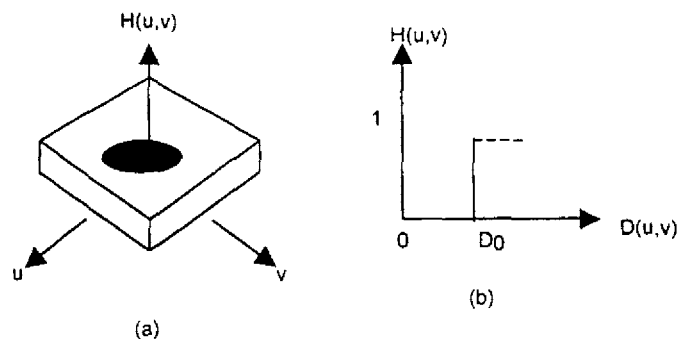


Figura 2.51 Filtro paso altas ideal IHPF. (a) Gráfica en perspectiva. (b) Sección transversal (tomada de Gonzalez and Wintz, 1987)

La función de transferencia del BHPF (filtro paso altas Butterworth) de orden n y frecuencia de corte localizada a una distancia D_0 del origen está definida por la relación

$$H(u, v) = \frac{1}{1 + \left[\frac{D_0}{D(u, v)} \right]^{2n}}$$

donde $D(u, v)$ es la distancia del punto (u, v) al origen del plano. La gráfica de la función de transferencia del BHPF es mostrada por la figura 2.52

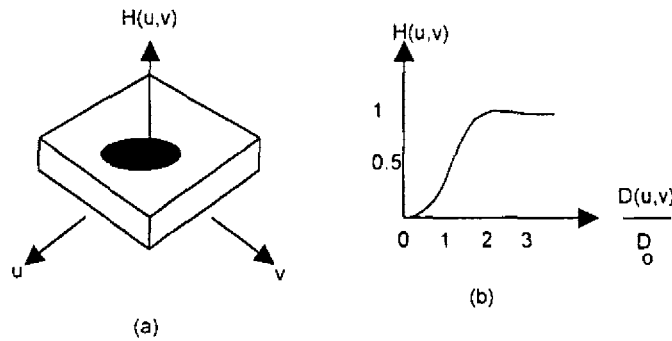


Figura 2.52 Filtro paso altas Butterworth, BHPF. (a) Gráfica perspectiva. (b) Sección transversal (tomada de Gonzalez and Wintz, 1987)

Note cuando $D(u, v) = D_0$, $H(u, v)$ es un 50% de su valor máximo. Como en el caso del filtro paso bajas Butterworth, es práctica común seleccionar la posición de la frecuencia de corte en el punto para el cual $H(u, v) = 1/2^{1/2}$ de su valor máximo.

2.2.2.6 FILTRO RECHAZO BANDA

Los filtros paso banda y rechazo banda (band reject) son una extensión de los conceptos de filtro paso bajas y paso altas mencionados.

Una manera simple de generar filtros que atenúan las frecuencias de una vecindad circular al punto u_0, v_0 es ejecutando una traslación de coordenadas del filtro paso altas ya visto.

Un IBRF (filtro rechazo banda ideal), el cual suprime todas las frecuencias en una vecindad de radio D_0 al punto (u_0, v_0) está dado por la relación

$$H(u, v) = \begin{cases} 0, & \text{si } D_1(u, v) \leq D_0 \quad \text{ó} \quad D_2(u, v) \leq D_0 \\ 1, & \text{Cualquier otro caso.} \end{cases}$$

Donde

$$D_1(u, v) = \left[(u - u_0)^2 + (v - v_0)^2 \right]^{1/2}$$

$$D_2(u, v) = \left[(u - u_0)^2 + (v - v_0)^2 \right]^{1/2}$$

Debido a la simetría de la transformada de Fourier, el filtro rechazo banda que no se encuentra en el origen debe de ser efectuado en pares de simetría para obtener resultados adecuados. El procedimiento puede ser extendido de una manera similar a cuatro o más regiones (La figura 2.53 muestra este tipo de filtro).

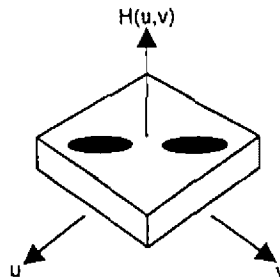


Figura 2.53 Filtro rechazo banda ideal (IBRF)
(tomada de Gonzalez and Wintz, 1987)

Si lo que desea es remover una banda de frecuencia centrada en el origen, se puede considerar filtros simétricos. Un filtro rechazo banda ideal radialmente simétrico, el cual remueve una banda de frecuencias a partir del origen está dado por

$$H(u, v) \begin{cases} 1, & \text{si } D(u, v) < D_0 - W/2 \\ 0, & \text{si } D_0 - W/2 \leq D(u, v) \leq D_0 + W/2 \\ 1, & \text{si } D(u, v) > D_0 + W/2 \end{cases}$$

donde W es el ancho de la banda y D_0 es la distancia a su centro a partir de su origen.

Un IBRF (filtro rechazo banda ideal) de orden n tiene la función de transferencia

$$H(u, v) = \frac{1}{1 + \left[\frac{D(u, v)W}{D^2(u, v) - D_0^2} \right]^{2n}}$$

donde W es el ancho de banda definido y D_0 es su centro

2.2.2.7 FILTRO PASO BANDA

Los filtros paso banda permiten el paso de una banda o región especificada, mientras atenúan, o suprimen completamente todas las otras frecuencias.

Un filtro pasa banda ideal, el cual suprime todas las frecuencias fuera de una vecindad de radio D_0 al punto (u_0, v_0) está dado por la relación.

$$H(u, v) \begin{cases} 1, & \text{si } D_1(u, v) \leq D_0 \quad \text{ó} \quad D_2(u, v) \leq D_0 \\ 0, & \text{Cualquier otro caso.} \end{cases}$$

Donde

$$D_1(u, v) = \left[(u - u_0)^2 + (v - v_0)^2 \right]^{1/2}$$

$$D_2(u, v) = \left[(u - u_0)^2 + (v - v_0)^2 \right]^{1/2}$$

Un filtro ideal paso banda puede representarse como se muestra en la figura 2.54

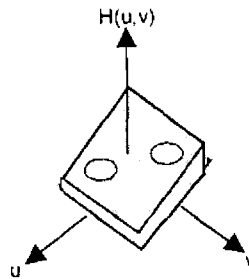


Figura 2.54 Filtro paso banda ideal
(tomada de Gonzalez and Wintz, 1987)

Estos filtros son exactamente lo opuesto de los filtros rechazo banda, la correspondiente función puede ser obtenida por la relación

$$H(u, v) = [H_R(u, v) - 1]$$

obteniendo como resultado

$$H(u, v) = \frac{1}{1 + \left[\frac{D^2(u, v) - D_0^2}{D(u, v)W} \right]^{2n}}$$

donde W es el ancho de banda definido y D_0 es su centro

CAPÍTULO III

CLASIFICACIÓN DE IMÁGENES

CONTENIDO

3 CLASIFICACIÓN DE IMÁGENES

3.1 Clasificación de una imagen

Introducción

3.1.1 Conceptos básicos

Clases informacionales

Clases espectrales

Patrón

Patrón espectral

Patrón espacial

Pixel o CIV

Espacio de observaciones

Clase rechazo

Regla de clasificación

Clases discriminantes

Clasificador

Función discriminante

Regla de decisión

Superficie o superficie de decisión entre las clases

Distancia de Mahalanobis

Regla de decisión óptima de Bayes

Riesgo medio de Bayes

Regla de máxima verosimilitud

Conjunto de entrenamiento

Estimadores de error

3.1.2 Proceso de entrenamiento

3.2 Clasificación supervisada

3.2.1 Clasificación Paramétrica

3.2.1.1 Clasificación por el teorema de Bayes

3.2.1.2 Clasificación por máxima verosimilitud

3.2.1.3 Clasificación por mínimas distancias

3.2.1.4 Clasificación por paralelepípedos

3.2.2 Clasificación no Paramétrica

3.2.2.1 Clasificación por medio de árboles

3.2 Clasificación no supervisada

3.3.1 Clasificación por migración de medias

3.3.2 Clasificación por agrupación de dos pasadas

CAPÍTULO III

CLASIFICACIÓN DE IMÁGENES

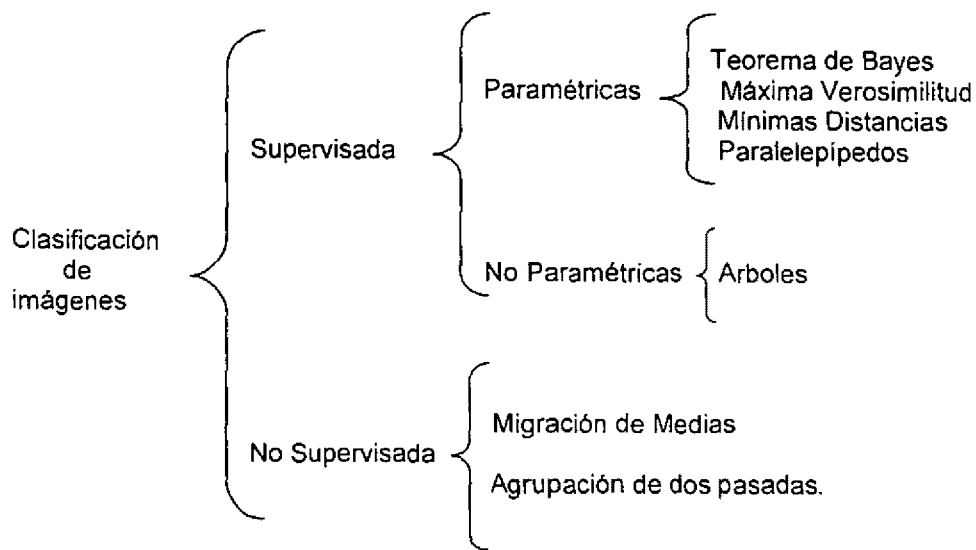
3.1 CLASIFICACIÓN DE UNA IMAGEN

INTRODUCCIÓN

Desde el lanzamiento del primer satélite dedicado a la observación de los recursos terrestres (LANDSAT-1,1972) se ha abierto una gran vía de investigación multidisciplinaria que ha posibilitado un conocimiento profundo de la Tierra. El desarrollo tecnológico permite hoy día el planteamiento de numerosos proyectos de investigación para desarrollar sensores de muy alta resolución, tanto espacial como espectral, que permitan obtener una gran cantidad de información sobre la superficie terrestre y abordar proyectos sobre recursos naturales con menor margen de error. En la próxima década se espera el lanzamiento de numerosos satélites con estas características. Pero hoy en día se cuenta con ciertos algoritmos computacionales que nos ayudan a interpretar la información proporcionada por los sensores. Uno de los métodos más usados para la extracción de información a partir de una imagen es la clasificación multispectral automática. Este método supone que, en una zona que es captada en diferentes regiones del espectro electromagnético, cada rasgo o cobertura sobre la superficie refleja energía dentro de una o más porciones específicas del espacio espectral. De esta forma el método consiste en determinar la región o regiones en las que una cobertura refleja energía. En otras palabras el objetivo final de las técnicas de clasificación es categorizar o asignar una etiqueta mediante un procedimiento automático a todos los píxeles de una imagen para obtener un mapa temático. Estas etiquetas son nombres asignados a diferentes cubiertas terrestres que se esperan encontrar en la escena bajo estudio. Para que la técnica de clasificación cumpla su objetivo hacen uso de información sobre la superficie en estudio y de funciones de probabilidad, generalmente, asociadas a los patrones de las clases. Sin embargo, las funciones de probabilidad son desconocidas usualmente, por lo que deben estimarse a partir de un conjunto de entrenamiento. En algunos casos se conoce la "forma" de la función de probabilidad, por lo que tan sólo deben estimarse algunos parámetros: se habla entonces de métodos paramétricos. Si no se conoce la forma de la función de probabilidad se habla de métodos no paramétricos. Los métodos paramétricos son generalmente más fáciles de implementar pero requieren un alto grado de conocimiento a priori acerca de la naturaleza de los patrones. Los métodos no paramétricos son potencialmente mejor considerando el grado de ajuste del valor estimado con el verdadero valor de probabilidad, aunque generalmente tienen asociado un alto coste computacional y requieren un elevado número de muestras de entrenamiento.

Los métodos paramétricos y no paramétricos son catalogados como **clasificadores supervisados**. Sin embargo, cuando se desconoce su función probabilística y no se tiene ninguna referencia de la superficie en estudio, se toma como parámetros la distancia espectral entre los píxeles para esto nos ayudamos de medios computacionales quedando la responsabilidad de etiquetar los grupos al analista. Estos métodos son catalogados como **clasificadores no supervisados**.

En este capítulo presentamos la clasificación de imágenes con la siguiente secuencia; inicialmente se presentan definiciones necesarias para el tema, profundizando un poco más en la clase rechazo y en el proceso de entrenamiento por considerarlos de mayor importancia para una exitosa clasificación. Posteriormente se presentan los métodos de clasificación, cabe aclarar que no son los únicos métodos existentes, sin embargo, son algunos de los más usados. Se han estructurado de la siguiente manera.



A continuación se presentan algunos conceptos y definiciones básicas que nos ayudarán a entender mejor los métodos de clasificación.

3.1.1 CONCEPTOS BÁSICOS

Clases informativas

Se refieren a los objetos que se desean discriminar con la clasificación de una imagen (en otras palabras, a las leyendas de un mapa temático) y son definidas por el investigador.

Clases espectrales

Están relacionadas con el espacio de observaciones: son los agrupamientos o nubes de datos que podemos encontrar en el espacio de observaciones. Es decir, las clases espectrales son inherentes a los datos de percepción remota y deben ser identificados y etiquetados por el analista.

Nota:

Idealmente podemos esperar una correspondencia directa entre clases espectrales y clases informativas; desafortunadamente este no es un caso muy común. Un analista debe entender bien las características espectrales del sistema sensor y ser capaz de relacionar estos parámetros a los tipos y proporciones de materiales encontrados dentro de la escena de observación o CIV del píxel. Si estos parámetros son entendidos, las clases espectrales pueden ser reetiquetadas como clases de información cuando sean necesarios.

Patrón

Mediante este término entendemos a un modelo que caracteriza una clase.

Patrón Espectral

La familia de clasificadores basados únicamente en la información individual de cada pixel se engloban bajo la categoría del reconocimiento espectral de patrones.

Patrón Espacial

En el proceso de clasificación existen clasificadores que usan información espacial (textura, proximidad espacial, contexto o direccionalidad) para el reconocimiento espacial de patrones.

Pixel o CIV (Campo Instantáneo de Vista)

Definimos una *observación* o un *pixel* o *CIV*, como vector aleatorio con d componentes,

$$X = [X^1, X^2, \dots, X^d]^T$$

donde $X^i \in G_i$,

X^i es una variable aleatoria que representa el valor de la observación X en la banda i . (Se suele utilizar el término característica para referenciar a una variable de este tipo). Cada variable aleatoria X_i toma valores en un conjunto G_i . En imágenes de Satélites X^i es el valor muestreado y digitalizado de reflectividad en la banda i . Valores típicos de reflectividad (función de la reflectividad espectral) están en el conjunto $G = \{0, 1, 2, \dots, 255\}$. Estos valores pueden representarse "visualmente" como niveles de gris, escalados desde 0 (negro) a 255 (blanco). En este caso $G_i = G$ para $i = 1, 2, \dots, d$

Espacio de observaciones

Definimos el espacio de observaciones Φ como el conjunto de todos los valores posibles que puede tomar una observación X , esto es,

$$\Phi = \bigotimes_{i=1}^d G_i$$

donde \otimes denota el producto cartesiano.

Clases informacionales

Si suponemos que todos los objetos presentes en una imagen son elementos potenciales de J clases distintas numeradas ω_j , $j = 1, 2, \dots, j$, llamaremos $\Omega = \{\omega_1, \omega_2, \dots, \omega_j\}$ al conjunto de las *clases informacionales*.

Clase rechazo

El conjunto Ω puede extenderse con la incorporación de una nueva clase, la clase rechazo, ω_0 , que es asignada a todas las observaciones para las que no se tiene una certeza aceptable de ser clasificados correctamente en alguna de las clases de Ω . Así, el conjunto extendido de clases informacionales se notará $\Omega^* = \{\omega_1, \omega_2, \dots, \omega_j, \omega_0\}$

Regla de clasificación

La regla de clasificación es una función $d : \Phi \rightarrow \Omega^*$ que esta definida sobre las observaciones X tal que para toda observación X , $d(X) \in \Omega^*$.

Clases discriminantes

Cuando se representan observaciones de diferentes clases, los agrupamientos presentes son generalmente distinguibles entre sí y es posible, en estos casos, asociar regiones disjuntas en el espacio de observaciones a cada una de las clases representadas. En este caso se dice que las clases son discriminantes.

Clasificador

Un clasificador es una partición de Φ en $J+1$ subconjuntos disjuntos $R_0, R_1, R_2, \dots, R_J$ donde $\Phi = \bigcup_i R_i$ tal que para todo $X \in R_i$ se verifica que $d(X) = \omega_i$ para $i = 0, 1, 2, \dots, J$.

Cada una de las regiones R_i se conocen como regiones de decisión.

El objetivo de un clasificador es dividir Φ en regiones disjuntas, una por cada una de las clases discriminantes, o lo que es equivalente, construir una regla que identifique cualquier observación X como perteneciente a la clase c si la región en la que está situada es R_c .

Función discriminante

Una función discriminante para la clase i , $i = 0, 1, 2, \dots, J$, a la que denotaremos por $g_i(X)$ tiene la propiedad que $g_i(X)$ alcanza un valor mayor que cualquier otra función discriminante $g_j(X)$, $i \neq j$, si X es una observación que pertenece a la i -ésima región de decisión.

Regla de decisión

La regla de decisión basada en funciones discriminantes. Si ω_i denota la i -ésima clase,

$$d(X) = \omega_i \quad \text{si y solo si} \quad g_i(X) \geq g_j(X) \quad \forall j = 1, 2, \dots, J$$

Si se presenta el caso de $g_i(X) = g_j(X)$ donde $i \neq j$, algún criterio adicional debe añadirse a la regla de decisión.

Superficie o frontera de decisión entre las clases

La superficie o frontera de decisión entre las clases ω_i y ω_j esta formada por aquellos puntos X para los que se verifican que $g_i(X) = g_j(X)$, o lo que es igual, es la región del espacio que separa dos regiones de decisión adyacentes.

Si se conocen las superficies de decisión entre todas las clases, la decisión acerca de la clase de una observación cualquiera puede hacerse en base a su posición relativa, respecto al conjunto total de superficies de decisión.

La estrategia básica a seguir para la clasificación óptima de un conjunto de observaciones es la de minimizar la pérdida media, lo que es equivalente a minimizar la probabilidad media de error en la clasificación. Gracias a la propiedad de la distribución normal, se define una métrica sobre el espacio de observaciones, la distancia de Mahalanobis.

Distancia de Mahalanobis

Se define la distancia de Mahalanobis de una observación X a una clase i , y la notamos por M_X^i , como

$$M_X^i = (X - \mu_i)^T \Sigma_i^{-1} (X - \mu_i)$$

donde Σ_i es la matriz de covarianza de la clase i , Σ_i^{-1} es la matriz inversa de Σ_i , μ_i es el valor medio de la clase i y $(X - \mu_i)^T$ es el vector traspuesto de $(X - \mu_i)$

Los contornos de densidad constante son hiperelipses que verifican que para todo X situado en ellas la distancia de Mahalanobis M_X^i es constante. Al estar ponderada por la inversa de la matriz de covarianza, esta métrica considera la distinta variabilidad de las variables en el espacio.

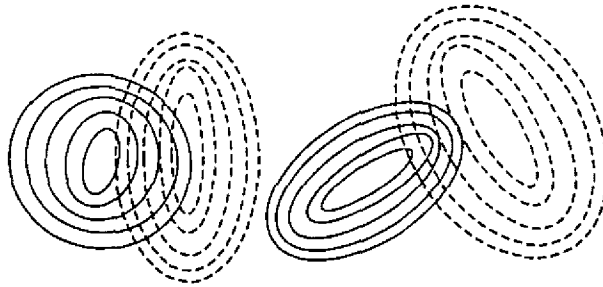


Figura 3.1 Ejemplo de fronteras de decisión cuadráticas.

En la figura 3.1 se muestra un ejemplo de fronteras de decisión cuadráticas. Cada serie de elipses concéntricas representa una clase. Las elipses indican zonas del espacio de igual probabilidad, esto es. Los puntos que se encuentran a igual distancia de Mahalanobis del centro de la clase.

En la práctica μ_i y Σ_i son desconocidos y deben estimarse a partir de las muestras de entrenamiento de la clase i . Sean $\hat{\mu}_i$ y $\hat{\Sigma}_i$ estimadores no sesgados de μ_i y Σ_i . Entonces, $\hat{\mu}_i$ y $\hat{\Sigma}_i$ se calculan por máxima verosimilitud como sigue:

$$\hat{\mu}_i^j = \frac{1}{n_i} \sum_{l=1}^{n_i} X_{l=i}^j \quad \text{para } j=1,2,\dots,d$$

$$\text{donde } \hat{\Sigma}_{ijk} = \frac{1}{n_i - 1} \sum_{l=1}^{n_i} (X_{l=i}^j - \hat{\mu}_i^j)(X_{l=i}^k - \hat{\mu}_i^k) \quad \text{para } j=1,2,\dots,d \text{ y } k=1,2,\dots,d$$

Donde n_i es el número de muestras de entrenamiento de la clase i y todas las observaciones X usadas para el cálculo de $\hat{\mu}_i$ y $\hat{\Sigma}_i$ siguen la misma distribución de probabilidad y pertenecen al conjunto de aprendizaje S .

El problema más serio relacionado con la estimación es encontrar el valor adecuado de n_i que garantice que los estimadores obtenidos sean estables.

Regla de decisión óptima de Bayes

Dado una observación X ,

$$d(X) = \omega_c \quad \text{si} \quad R(\omega_c|X) = \min_{i=1,\dots,J} \{R(\omega_i|X)\}$$

Riesgo medio de Bayes

Dado una observación X , la pérdida media o riesgo medio de Bayes resultante al clasificarlo como de clase ω_i está dada por

$$R(\omega_c|X) = \sum_{j=1}^J \lambda(i|j) p(\omega_j|X)$$

Donde $\lambda(i|j)$ es la pérdida o coste.

Para calcular las probabilidades a posteriori $p(\omega_i|X)$ utilizaremos la Regla de Bayes y dado que el denominador es constante e independiente de la clase, podemos expresarla como

$$p(\omega_c|X) \propto p(X|\omega_i) \pi_i$$

por lo que la pérdida media puede escribirse como

$$R(\omega_c|X) = cte \sum_{j=1}^J \lambda(i|j) p(X|\omega_j) \pi_j$$

$\lambda(i|j)$ es la pérdida o coste asociado a la decisión de clasificar una observación a la clase i cuando realmente debería clasificarse como de clase j .

tomaremos:

$$\lambda(i|j) = \begin{cases} 0 & \text{si } i = j \\ 1 & \text{si } i \neq j \end{cases}$$

en las que no hay coste es una decisión aceptada representada por 0 y donde hay coste es una unidad de decisión fallida representada por 1, de ahí que se conozca como función de coste cero-uno. Puede demostrarse que la regla de decisión óptima de Bayes (mínima pérdida media) con la función de coste cero-uno es equivalente a escoger la clase que maximiza la probabilidad a posteriori.

Regla de máxima verosimilitud

Dado una observación X ,

$$d(X) = \omega_c \quad \text{si} \quad L_X^c = \max_{i=1,2,\dots,J} \{L_X^i\}$$

donde $L_X^i = p(X|\omega_i)$

De todo lo expuesto se deduce que el conocimiento del valor de $p(X|\omega_i)$ es necesario para aplicar la regla de clasificación dada por la regla de máxima verosimilitud. Dado que el valor exacto de éste es generalmente imposible de conocer se suele estimar usando un modelo paramétrico o no paramétrico, de acuerdo con el conocimiento que se disponga a priori sobre la naturaleza del patrón de las clases.

Por la forma en que se ha definido la regla de decisión de máxima verosimilitud, ésta clasifica siempre cualquier observación como perteneciente a una de las clases de Ω (clase rechazo). Sin embargo, es posible que una observación no pertenezca a alguna de las clases, porque su clase cierta no se haya considerada al definir Ω porque presenta valores "extraños", muy diferentes a los demás debidos a problemas en la captación (llamados outliers). En cualquier caso, estas observaciones deberían asignarse a la clase rechazo.

La asignación de una observación a la clase rechazo se llevara a cabo por una técnica de umbralización, en la cual, el valor de máxima probabilidad a posteriori es comparado con un valor umbral. Si el valor de probabilidad es menor que el umbral, entonces la observación es asignada a la clase de rechazo. La elección de un valor umbral depende de la regla de clasificación utilizada y de la escena a clasificar, por lo que depende de los valores de las observaciones presentes en la imagen.

Una regla de decisión debe construirse en base a un conjunto de experiencias pasadas o casos conocidos para los que se sabe qué decisión tomar: el clasificador está basado en el conocimiento a priori. Este hecho está reflejado en la existencia de un conjunto de aprendizaje C formado por una serie de observaciones para cada una de las J clases a discriminar junto con la clase a la que se sabe con certeza a la que pertenecen (su "clase cierta"). A partir de S queremos inferir a qué clase pertenece una observación X , independiente de las observaciones de aprendizaje. La existencia de un buen conjunto de aprendizaje es fundamental ya que usamos S para construir y evaluar el clasificador.

Conjunto de entrenamiento

Un conjunto de aprendizaje o conjunto de entrenamiento S , $S \in \Phi \otimes \Omega$, consiste en un conjunto de parejas (X_i, c_i) donde $X_i \in \Phi$ y $c_i \in \Omega$ a las que llamamos muestras de aprendizaje.

$$S = \{(X_1, c_1), (X_2, c_2), \dots, (X_N, c_N)\}$$

donde N es el número de muestras de aprendizaje y c_i es la "clase cierta" de la observación

Estimadores de error

Dado un clasificador d , denotado mediante $R^*(d)$ a su *razón cierta de error*, entendida como una estimación de la proporción de puntos que clasificará correctamente.

El error real cometido en una clasificación es difícilmente calculable, salvo en situaciones muy concretas y en la práctica resulta incalculable. Por esta razón se proporcionan estimadores de error cometido. Los tres estimadores usados generalmente son: el estimador por resustitución, el estimador mediante un conjunto de pruebas y el estimador por validación cruzada. El error real está acotado inferiormente por la medida de resustitución y superiormente por la medida de validación cruzada.

3.1.2 PROCESO DE ENTRENAMIENTO

El primer paso en el proceso de clasificación supervisada, es el entrenamiento. Durante el proceso de entrenamiento se seleccionan sobre la imagen unas áreas que se consideran representativas de las clases de Ω . Estas áreas se denominan áreas de entrenamiento, y se usan para calcular el patrón que define cada clase. Estos patrones se utilizarán como referencia en el proceso de clasificación. Para esta selección se debe disponer de información adicional que nos permita conocer la clase cierta de los píxeles seleccionados. Esta información puede obtenerse a través de mapas, de fotografías aéreas, conocimiento de la zona, o aún mediante la fotointerpretación de compuestos en falso color generados a partir de las bandas de la misma imagen.

Respecto a la obtención de información basada en citas de campo, ésta se debe hacer en un período de tiempo cercano al de la captura de la imagen, de otra manera pueden existir errores debidos a variaciones temporales.

Identificando los patrones de cada clase se entrena al clasificador (en definitiva, se construye S) de forma que en el proceso de clasificación se puedan identificar pixeles similares como pertenecientes a la misma clase. La selección de las áreas de entrenamiento debe ser el resultado de un muestreo serio y responder a criterios de representatividad y validez estadística, además al ambiente del cual fueron obtenidos es relativamente homogéneo. Los datos de entrenamiento serán de valor si el ambiente del cual fueron obtenidos es relativamente homogéneo. Si esto no ocurre, se tendrán problemas de extensión de firma, esto es, no será posible extender nuestros datos de entrenamiento a través del espacio de coordenadas

En cuanto a la representatividad de las áreas de entrenamiento, la selección de áreas de entrenamiento será incorrecta si alguna se sitúa en zonas de transición entre clases. Al tratarse de una zona heterogénea, el patrón extraído será un promedio de los distintos componentes: será una clase híbrida y en consecuencia, puede confundirse con otras clases. En el otro extremo, también sería incorrecto seleccionar únicamente áreas muy homogéneas, donde no exista apenas diferencia entre los niveles digitales de los pixeles, obviando la variabilidad que existe en la Naturaleza. Se trata, por lo tanto, de encontrar un equilibrio en la homogeneidad de los pixeles que forman un área de entrenamiento.

La manera más fácil de remediar esta situación es, recopilar todos los factores ambientales que puedan contribuir a los problemas de extensión de firmas, como diferencias en los tipos de suelo, turbidez del agua, especies de cultivos, condiciones de humedad inusual en el suelo, etc. Estas condiciones ambientales deben ser cuidadosamente anotadas en la imagen y la selección de los sitios de entrenamiento hecha basada en la estratificación geográfica de estos datos.

En cuanto a la validez estadística, se deben cumplir ciertos requisitos sobre el muestreo, referentes al tamaño y a la distribución de la muestra de entrenamiento o aprendizaje. Sobre el tamaño de la muestra, se requiere un mínimo de $d + 1$ pixeles por clase. Sin embargo, diversos autores recomiendan superar con creces este número y fijarlo entre $10 \times d$ y $100 \times d$ pixeles por clase, esta condición permitirá encontrar el inverso de la matriz de covariancia para cada clase a ser calculada, lo cual es importante en algunos algoritmos de clasificación, sin embargo, no sería estadísticamente representativos. Los clasificadores entrenados con muy pocas muestras suelen ser muy inestables, estando muy influenciados por la inclusión o no de una determinada muestra.

Se deben usar los datos de entrenamiento para estimar los parámetros del algoritmo de clasificación particular a ser usado; estos parámetros son las propiedades del modelo de probabilidad empleado o las ecuaciones que definen las particiones en el espacio espectral. El juego de parámetros, para una clase dada, es llamado **la firma de la clase**.

En el reconocimiento de patrones se usan dos aproximaciones diferentes: clasificación supervisada y clasificación no supervisada.

- En la **clasificación supervisada** los clasificadores se construyen en base a clases informacionales. El analista "supervisa" el proceso de clasificación especificando al clasificador de las diferentes clases informacionales de interés: se intenta construir un modelo o patrón para cada clase. Para ello se usan muestras representativas de las diferentes cubiertas a partir de áreas de entrenamiento de forma que los patrones que caracterizan las diferentes clases se obtienen de estas áreas.
- En la **clasificación no supervisada** no se utilizan datos de entrenamiento como base para la clasificación: los clasificadores se construyen a partir de las agrupaciones espectrales existentes en la imagen. La idea básica que subyace en estos métodos (llamados de agrupamiento o clustering) es que los pixeles de una misma clase deben tener valores espectrales similares, mientras que pixeles de diferentes clases deberían tener valores diferentes y distinguibles.

3.2 CLASIFICACIÓN SUPERVISADA

A continuación se presentan los pasos a seguir para efectuar una clasificación:

1. Como primer paso de una clasificación supervisada se debe definir las categorías o tipos de cobertura a ser identificados dentro de la imagen. Estas son las clases de información y pueden ser, por ejemplo, agua, zonas urbanas, bosques, etc. Las categorías de interés deben ser cuidadosamente seleccionadas para llevar a cabo una clasificación exitosa.
2. Se debe establecer un esquema apropiado de clasificación para las coberturas. Si ya existen sistemas de clasificación reconocidos en la literatura internacional, es mejor adoptar o modificar éstos que inventar uno nuevo que sea usado exclusivamente por uno mismo. Esto permite mostrar la importancia de los resultados de nuestra clasificación a la luz de otros estudios y facilitar la mezcla de datos.
3. Debe ser notado que hay una relación entre el nivel de detalle en un esquema de clasificación y la resolución espacial. Esto sugiere que el nivel de detalle en el sistema de clasificación deseado, dicta el sistema de percepción remota que debería ser usado para proporcionar información. La resolución espectral es también una consideración importante. Sin embargo, no es un parámetro tan crítico como los datos de resolución espacial, pues la mayoría de los sistemas sensores (Landsat, Spot, etc.) registran la energía en aproximadamente las mismas porciones del visible y el infrarrojo medio.
4. Una vez que los factores de extensión de firma han sido evaluados y el analista ha seleccionado los sitios de entrenamiento representativos para cada una de las clases, se deben extraer las estadísticas espectrales de los sitios de entrenamiento. Las estadísticas de entrenamiento son analizadas para seleccionar las bandas que son más efectivas en discriminar las clases, esto es, aquellas que son más apropiadas para ser usadas en el proceso de clasificación.
5. Posteriormente se obtiene la discriminación de bandas, comúnmente llamada selección de rasgos, en donde el objetivo es ignorar, a partir de un análisis, aquellas bandas que proporcionan únicamente información espectral redundante. De esta manera la dimensionalidad (es decir, el número de bandas a ser procesadas) del juego de datos puede ser reducida. Este proceso minimiza el costo de la clasificación digital de la imagen, pero no la exactitud. La selección de rasgos puede involucrar tanto métodos gráficos, como análisis estadístico para determinar el grado de separación entre las clases. Combinaciones de n bandas son normalmente hechas, de acuerdo a su habilidad potencial para discriminar cada clase de las otras.
6. Una vez definidas las bandas que permiten una mejor discriminación de los rasgos, se lleva a cabo la definición del algoritmo a emplear. La selección de una regla de clasificación o decisión depende de la naturaleza de los datos de entrada y de la salida deseada. Algoritmos de clasificación paramétricos asumen que las clases consideradas son gaussianas en naturaleza (es decir, normalmente distribuidas). Algoritmos de clasificación no paramétrica no hacen tal suposición. Los algoritmos de clasificación más frecuentemente usados son: paralelepípedos, mínimas distancias, máxima verosimilitud y bayesiano.
7. Las firmas obtenidas a partir de los campos de entrenamiento, se emplean como entrada del algoritmo de clasificación seleccionado para etiquetar cada uno de los píxeles en la imagen, dentro de uno de los tipos de cobertura deseados (clases de información). En este paso el total de la imagen es clasificada mientras que en la fase de entrenamiento únicamente un porcentaje pequeño (entre 1% y 5%) de píxeles es identificado por otros medios.

La clasificación de la imagen se efectúa con n bandas que permiten una mejor discriminación de las clases.

Dos tipos de salida son disponibles a partir de una clasificación.

- El mapa temático, (o de clase) como se mencionó anteriormente, en el cual los píxeles son etiquetados y representados por un color o símbolo para identificar que ellos pertenecen a una clase.
- La otra es una tabla que resume el número de píxeles en la imagen que pertenece a cada una de las clases y el área que éstos cubren.

En una clasificación supervisada, la identidad y posición en la imagen de algunos de los tipos de cobertura tales como: áreas urbanas, agrícolas, bosques, etc. son conocidas a priori a través de una combinación de trabajo de campo, análisis de fotografías aéreas, mapas y experiencia personal. El analista intenta localizar sitios en la imagen que representen ejemplos homogéneos de esos tipos de cobertura conocidas. Estas áreas son llamadas sitios de entrenamiento porque sus características espectrales son usadas para entrenar al algoritmo de clasificación para el eventual mapeo de coberturas del resto de la imagen. Los parámetros estadísticos (media, desviación estándar, matriz de covariancia, matrices de correlación, etc.) son calculados par cada sitio de entrenamiento.

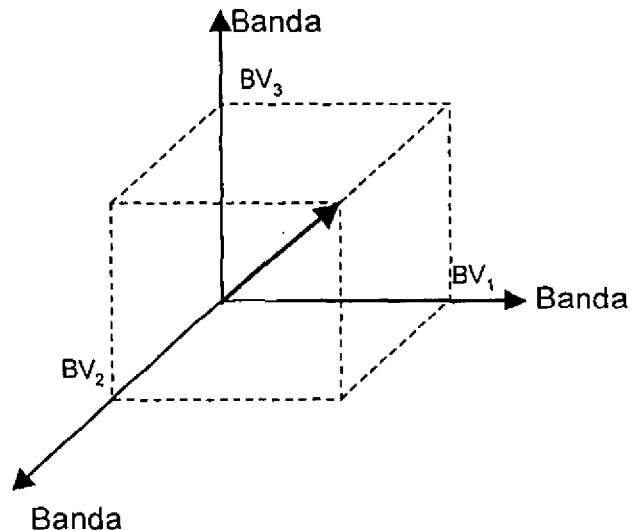


Figura 3.2 Representación de un píxel en el espacio espectral

3.2.1 CLASIFICACIÓN POR EL TEOREMA DE BAYES

Esta regla de decisión asigna cada píxel que tiene un vector de posición X , en el espacio espectral, a la clase que sea más probable de haberle dado origen.

El algoritmo asume que los datos estadísticos de entrenamiento para cada clase están normalmente distribuidos. Los datos de entrenamiento con histogramas bi o trimodales no son ideales. En tales casos, las modas individuales probablemente representan clases que deberían ser tratadas por separado y etiquetadas como clases aparte. Esto produciría clases de entrenamiento unimodales que cumplirían con el requerimiento de distribución normal.

Las clases espectrales son representadas por:

c_i , donde $i=1,2,3,\dots,m$; y m es el número total de clases.

Para tratar de determinar la clase a la cual un pixel en la posición X pertenece, el método se basa estrictamente en la probabilidad condicional.

$$p(c_i / X) \quad i=1,2,3,\dots,m$$

Esto es, $p(c_i / X)$ da la probabilidad de que la clase sea c_i dado que se está examinando un pixel en la posición X , o, la probabilidad de que la clase correcta sea c_i , para un pixel en la posición X .

La regla de clasificación es ejecutada de acuerdo a:

$$X \in c_i \text{ si } p(c_i / X) > p(c_j / X) \quad \forall j \neq i$$

Para clasificar el vector de mediciones X de un pixel desconocido dentro de una clase, la regla de decisión calcula el valor $p(c_i / X)$ para cada clase. Ésta asigna a la clase que tiene el valor mayor o máximo.

Los pixeles de entrenamiento pueden ser usados para estimar la distribución de probabilidades de cada tipo de cobertura, las cuales describen la posibilidad de encontrar un pixel de la clase c_i y se representa por $p(X / c_i)$, la cual se interpreta como: la probabilidad de encontrar un pixel en la posición X dado que estamos en la clase c_i .

Probabilidad Conjunta es la probabilidad de que un pixel en la posición espectral sea un pixel de la clase c_i y está dada por el producto:

$$p(X, c_i) = p(X / c_i)p(c_i)$$

Donde $p(c_i)$ es la probabilidad de que la clase c_i ocurra en la imagen.

Como la probabilidad conjunta expresa la posibilidad de que X y c_i están juntas, ésta se puede escribir también como.

$$p(c_i, X) = p(c_i / X)p(X)$$

Donde $p(X)$ es la probabilidad de encontrar un pixel en la posición X a partir de cualquier clase.

De las definiciones de probabilidad conjunta se establece que:

$$p(c_i / X) = p(X / c_i)p(c_i) / p(X)$$

El cual es conocido como **teorema de Bayes**; $p(c_i)$ es la probabilidad a priori porque es la probabilidad con la cual cualquier pixel puede ser adivinado antes de la clasificación. De esta misma manera $p(c_i / X)$ es la probabilidad a posteriori.

Empleando el teorema de Bayes en la regla de clasificación se tiene que:

$$X \in c_i \text{ si } p(X/c_i)p(c_i) > p(X/c_j)p(c_j) \quad \forall j \neq i$$

Donde $p(X)$ ha sido removido por ser factor común.

Resulta conveniente matemáticamente definir la función.

$$P_i(X) = \ln\{p(X/c_i)p(c_i)\}$$

$$P_i(X) = \ln\{p(X/c_i)\} + \ln\{p(c_i)\}$$

Introduciendo esta función la regla de clasificación queda como:

$$X \in c_i \text{ si } P_i(X) > P_j(X) \quad \forall j \neq i$$

La función $P_i(X)$ es llamada **función discriminante**.

Como se mencionó anteriormente se asume que la distribución de probabilidad para cada clase tiene la forma de un modelo normal, un modelo multivariado. De esta forma, considerando n bandas se tiene que:

$$p(X/c_i) = (2\pi)^{-n/2} \det(V_i)^{-1/2} \exp\{-\frac{1}{2}(X - M_i)' V_i^{-1} (X - M_i)\}$$

Donde M_i y V_i son el vector de medias y la matriz de covariancia de los datos de entrenamiento para la clase c_i . Sustituyendo el modelo de distribución normal multivariado en la definición de función discriminante $P_i(X)$ y considerando la regla de clasificación se tiene:

$$P_i(X) = \ln\{p(c_i)\} - \frac{1}{2} \ln\{\det(V_i)\} - \frac{1}{2} (X - M_i)' V_i^{-1} (X - M_i)$$

El término resultante $-\frac{n}{2} \ln(2\pi)$ es común a todos los $P_i(X)$ por lo que al ser sustituido en la regla de clasificación se eliminaría, además de no ayudar a la discriminación, motivo por el cual es ignorado.

3.2.2 CLASIFICACIÓN POR MÁXIMA VEROSIMILITUD

La regla de decisión de máxima verosimilitud es idéntica a la de Bayes excepto que ésta asume que cada clase es igualmente probable.

La ecuación que define a $P_i(X)$ asume que todas las clases tienen una probabilidad a priori, $p(c_i)$, de ocurrencia en el terreno. El sentido común nos recuerda que en la mayoría de las aplicaciones de sensores remotos hay una alta probabilidad de encontrar algunas clases más frecuentemente que las otras, por lo que a través de este término es posible incluir esta probabilidad a priori en la decisión de clasificación. Sin embargo, frecuentemente el analista no tiene información útil o suficiente sobre esta probabilidad $p(c_i)$ en cuyo caso se supone una probabilidad a priori igual. Esta consideración hace que:

$$p(c_i) = p(c_j) \quad \forall j \neq i$$

Por lo que el término $\ln\{p(c_i)\}$ es una constante a todos los $P_i(X)$ que al ser sustituido en la regla de clasificación se eliminará, motivo por el cual puede ser ignorado. De la misma forma el factor común $\frac{1}{2}$ puede ser removido.

El algoritmo de máxima verosimilitud hace uso del vector de medias, M_{j_i} , para cada clase y de la matriz de covariancia de la V_{j_i} . La regla de decisión aplicada a vectores de medición desconocidos, X , es:

$$X \in c_i \text{ si } P_i(X) > P_j(X) \quad \forall j \neq i$$

donde $j, i = 1, 2, \dots, m$ posibles clases, y

$$P_i(X) = -\ln\{\det(V_{j_i})\} - (X - M_{j_i})' V_{j_i}^{-1} (X - M_{j_i})$$

Donde $P_i(X)$ es la probabilidad del vector X de pertenecer a la clase c_i y $\det(V_{j_i})$ es el determinante de la matriz de covariancia V_{j_i} .

Para clasificar el vector de mediciones X de un pixel desconocido dentro de una clase, la regla de decisión calcula el valor $P_i(X)$ para cada clase y asigna el pixel a la clase que tiene el valor mayor o máximo.

3.2.3 CLASIFICACIÓN POR MÍNIMAS DISTANCIAS

Cuando el número de muestras de entrenamiento por clase es limitado puede ser más efectivo emplear un clasificador que no haga uso de la información de covariancia, sino que dependa únicamente de la posición de las medias de las clases espectrales.

El clasificador de mínimas distancias, o clasificador de la distancia mínima a la media de las clases hace uso únicamente de los vectores de medias de las clases espectrales, M_{j_i} , y se asigna el pixel a la clase cuya media se encuentra más cercana. De esta manera la regla de clasificación es:

$$X \in c_i \text{ si } D(X, M_{j_i}) < D(X, M_{j_j}) \quad \forall j \neq i$$

donde $j, i = 1, 2, \dots, m$ posibles clases.

Las dos distancias más usadas son la distancia euclidiana, basada en el teorema de Pitagoras y la distancia de alrededor del bloque.

La distancia euclidiana en el espacio de n-dimenciones se define como:

$$D(X, M_k) = \left[\left(\sum_{k=1}^n (\mu_{ik} - BV_{j_{ik}})^2 \right) \right]^{1/2}$$

La distancia de un bloque en el espacio n-dimensional se define como:

$$D(X, M_i) = \sum_{k=1}^n |\mu_{ik} - BV_{j_{ik}}|$$

donde μ_{ik} es el valor de la clase c_i en la banda k y BV_{ijk} es el valor del pixel en la posición (j,i) de la imagen en la banda k .

Es obvio que cualquier pixel será definitivamente asignado a una de las clases, es decir, no habrá pixeles sin clasificar.

Muchos algoritmos de mínimas distancias permiten al analista especificar una distancia o umbral de pertenencia a partir del valor medio de la clase. De esta forma, un pixel no será asignado a una categoría, aunque ésta sea el valor medio más cercano, si la distancia a ella sobrepasa este umbral.

Debido a que los datos de covariancia no son usados en mínimas distancias los modelos de las clases son simétricos en el dominio espectral. Clases elongadas no serán bien modeladas. Varias clases espectrales pueden ser necesarias con este algoritmo donde sólo una es encontrada por máxima verosimilitud.

Esta regla de decisión es computacionalmente simple y puede resultar en clasificaciones de exactitud comparable a otros algoritmos computacionales más intensos, tales como el de máxima verosimilitud.

En una clasificación por máxima verosimilitud, cada clase es modelada por una distribución de probabilidad normal multivariada, que toma en cuenta la dispersión de los datos en cualquier dirección particular del espectro a través de la matriz de covariancia. Sin embargo, cuando se considera la matriz de covariancia, V_i , de todas las clases, a ser diagonal e igual y la variancia en cada componente a ser idéntica

$$V_i = \sigma^2 I \quad \forall i$$

Los modelos de las clases son simétricos y las clasificaciones por mínimas distancias y máxima verosimilitud son equivalentes.

Existe una métrica que tiene la propiedad de considerar la variabilidad de las variables en las distintas direcciones del espacio que es la distancia de Mahalanobis y para el método de mínimas distancias se define como sigue.

Regla de clasificación de mínimas distancia de Mahalanobis

$$d(X) = c_i \quad \text{si} \quad M_X^i(X) = \min_{i=1, \dots, J} \{M_X^i(X)\}$$

Donde $M_X^i(X)$ está definida según la distancia de Mahalanobis antes ya definida.

3.2.4 CLASIFICACIÓN POR PARALELEPÍEDOS

Este algoritmo está basado en una simple lógica booleana "y/o" los valores de brillantez para cada pixel de entrenamiento, en la imagen multiespectral, son usados para producir vectores de medias n-dimencional, $M_i = (\mu_{i1}, \mu_{i2}, \mu_{i3}, \dots, \mu_{in})$, con μ_{ik} siendo el valor medio de los datos de entrenamiento para la clase c_i en la banda k de las m posibles clases previamente definidas, σ_{ik} es la desviación estándar de los datos de entrenamiento de la clase c_i en la banda k de las m posibles.

El método emplea dos fronteras de decisión una inferior y otro superior sobre cada eje del espacio espectral, definidas como:

$$Inf_{ik} = \mu_{ik} - x\sigma_{ik}$$

$$Sup_{ik} = \mu_{ik} + x\sigma_{ik}$$

Donde x es el número de veces que se tomará la desviación estándar (σ_{ik}) de la clase c_i en la banda k para establecer el valor umbral de las fronteras a partir del valor medio μ_{ik} .

De esta manera un algoritmo de paralelepípedo decide si BV_{jik} está en la clase c_i si y sólo si:

$$Inf_{jik} \leq BV_{jik} \leq Sup_{jik}$$

Donde $i=1,2,\dots,m$ número de clases.
 $k=1,2,\dots,n$ número de bandas.

Estas fronteras de decisión forman un paralelepípedo n -dimensional en el espacio de las bandas. Si el valor del pixel cae sobre el valor umbral inferior y bajo el valor umbral superior para todas las n bandas evaluadas, el pixel es asignado a esta clase. Cuando el valor de un pixel desconocido no satisface ninguno de los criterios de lógica booleana, éste es asignado a la categoría de desconocido.

Incrementando el tamaño del umbral, de 2 o 3 veces la desviación estándar, se incrementa el tamaño del paralelepípedo y se reduce el número de pixeles asignados a la clase desconocida, pero esta acción puede también introducir una cantidad de traslape significativo entre paralelepípedo, resultando en una clasificación errónea.

El algoritmo de paralelepípedos es un método computacionalmente eficiente para clasificar. Desafortunadamente, debido a que algunos paralelepípedos se traslapan, es posible que un pixel satisfaga los criterios para más de una clase. En tales casos es usual asignarlo a la primera clase para la cual éste cumple todos los criterios. Otra solución es tomar en criterio de asignar el pixel a aquella clase cuyo centro se encuentre a la mínima distancia.

Los métodos hasta aquí citados son clasificadores estadísticos paramétricos porque la regla de clasificación supervisada paramétrica utiliza parámetros que fueron obtenidos en el proceso de aprendizaje, por las clases ideales, previo a la clasificación: pero existen otros clasificadores que no utilizan parámetros y a continuación se presentan.

3.2.5 CLASIFICACIÓN NO PARAMÉTRICA

Bajo la denominación de clasificación no paramétrica se encuentran aquellos clasificadores en los que la única información disponible es un conjunto de muestras de entrenamiento y no se realizan hipótesis acerca de la distribución de probabilidad subyacente.

Existen distintas técnicas no paramétricas de clasificación dentro del campo de reconocimiento de patrones. Pero en este trabajo se presenta únicamente los árboles de clasificación por ser una de las técnicas más usadas que mejor resultado dan.

3.2.5.1 CLASIFICACIÓN POR MEDIO DE ARBOLES

Los árboles de clasificación representan el conocimiento jerárquicamente estructurado en forma de árbol. La idea en estos métodos es realizar una partición recursiva del espacio de observaciones basada en un conjunto de muestras de entrenamiento de acuerdo a algún criterio de bondad.

Podemos esquematizar, de forma sencilla, el funcionamiento de un árbol de decisión como sigue. Cada nodo interno del árbol representa una decisión acerca de un atributo o variable. Si la evaluación de la condición sobre la observación actual se satisface, se pasa a una de las ramas de ese nodo; si no se satisface, se pasa a otra rama. La clasificación de una observación consiste en repetir este proceso desde el nodo raíz del árbol hasta alcanzar una hoja. Los nodos hoja están etiquetados y por lo tanto representan una decisión, asignando la etiqueta de ese nodo a la observación en curso.

Los árboles de clasificación son una poderosa alternativa a los métodos clásicos, tanto paramétricos como no paramétricos empleados para clasificación y regresión. En esta sección introduciremos las bases de esta metodología aplicada a la clasificación.

Los árboles de clasificación se encuentran dentro de las técnicas de clasificación supervisada. El objetivo final es construir una regla de decisión basada en el conjunto de aprendizaje S . La asignación de una clase a una observación cualquiera X se realiza mediante la evaluación de una serie de condiciones acerca del valor de las variables X^1, X^2, \dots, X^d .

Un clasificador supervisado puede entenderse como un proceso que divide el espacio de observación Φ en regiones disjuntas, R_1, R_2, \dots, R_j de forma que

$$\Phi = \bigcup_i R_i, \quad R_i \cap R_j = \emptyset \text{ para } i \neq j$$

Los conjuntos R_i forman una partición de Φ , de forma que para todo $X \in R_i$ se verifica que:

$$d(x) = \omega_i$$

El fundamento de los árboles de clasificación es realizar predicciones mediante la construcción de un árbol binario a partir de un análisis sistemático de los datos de entrenamiento. Cuando el árbol está construido se procede a la clasificación. Estos pasos pueden resumirse como sigue.

1.- Construcción del árbol. El proceso de construcción del árbol de clasificación se realiza mediante una partición recursiva Φ en base a las muestras de entrenamiento disponibles. Se parte del nodo raíz, el cual tiene, asignado todas las muestras. Se realiza una partición de este conjunto evaluando una regla lógica (simple o compleja) sobre el/los valor(es) de alguna(s) variable(s). Las muestras que verifican la condición se asigna a uno de los nodos hijo (normalmente el izquierdo) y las restantes se asignan al otro.

Este proceso se repite recursivamente para cada uno de los nodos creados hasta que se verifica una condición de parada; en este momento se decide la partición y se continúa con otro nodo. Las muestras que están asociadas a un nodo constituyen un agrupamiento homogéneo, por lo que al nodo se le asigna una etiqueta, que normalmente la de la clase más representada. Estos nodos se conocen como **nodos de hoja o terminales**, en contraposición con el resto, que se conocen como **nodos intermedios** (Ver figura 3.3).

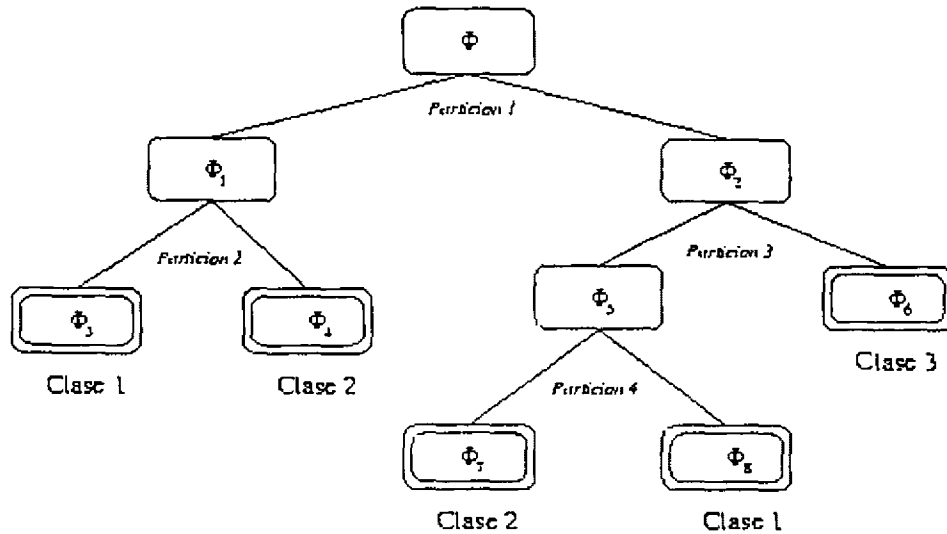


Figura 3.3 Un clasificador estructurado en forma de árbol binario

Este proceso de esquematización en la figura 3.3 para todo el árbol se verifica que los conjuntos de muestras asociados a sus hijos son disjuntos. Los nodos terminales, representados con un doble semirectángulo, cumplen que:

$$\Phi_3 \cup \Phi_4 \cup \Phi_7 \cup \Phi_8 \cup \Phi_6 = \Phi$$

de esta forma, las regiones de decisión quedan establecidas por las siguientes ecuaciones:

$$R_1 = \Phi_1 \cup \Phi_8, \quad R_2 = \Phi_4 \cup \Phi_7, \quad R_3 = \Phi_6$$

2. Clasificación. La clasificación de una observación X , en base a un árbol de clasificación, se realizaría como sigue: se evalúa la condición del nodo raíz. Si es cierta, se procede de la misma forma con el árbol izquierdo y si es falsa, con el árbol derecho. Se continúa esta cadena de evaluación hasta alcanzar un nodo hoja, asignándose la clase asociada a este nodo.

Para la construcción de un árbol de clasificación hay que especificar cuatro elementos básicos:

1. Un conjunto Q de preguntas binarias de la forma $\{¿Es X \in R ?\}$, $R \subset \Phi$ que permitan la partición de un nodo del árbol.
2. Una función que mida la bondad del criterio de partición $\Psi(s, t)$ que permita evaluar cualquier partición s de cualquier nodo t , en definitiva, algún criterio para escoger la mejor partición.
3. Una regla de parada, detener el proceso de partición.
4. Una regla para asignar cada nodo terminal a una clase y una estimación del error.

3.3 CLASIFICACIÓN NO SUPERVISADA

En una clasificación no supervisada, las identidades de los tipos de cobertura a ser especificadas como clases dentro de una escena, generalmente no son conocidas a priori, porque se carece de un trabajo de campo o los rangos de la superficie en la imagen no están bien definidos. La computadora es requerida para agrupar los pixeles dentro de diferentes clases espectrales, de acuerdo a algún criterio estadístico determinado. Es, por tanto, responsabilidad del analista etiquetar estos grupos.

Los pixeles que pertenecen a una agrupación particular, son espectralmente similares. Muchos criterios de similitud han sido propuestos, pero aquellos empleados comúnmente en procedimientos de agrupación, son generalmente la simple medición de distancias en el espacio multiespectral. Las distancias más frecuentemente empleadas son la euclidiana y la de alrededor de bloque.

3.3.1 CLASIFICACIÓN POR MIGRACIÓN DE MEDIAS

Es conocida también como el procedimiento de optimización interactiva de agrupamiento; y está basado en asignar los pixeles de la imagen dentro de unos candidatos o grupos y posteriormente moverlos de un grupo a otro, de manera que la suma del error cuadrático se reduzca.

A continuación se presenta el algoritmo de esta técnica:

1.- Se inicializa el procedimiento mediante el establecimiento de C puntos en el espacio multiespectral que sirven como candidatos de agrupación llamados:

$$M_i \quad i=1,2,\dots,C$$

La selección de los M_i es arbitraria, con la excepción de que no puede haber dos centros con el mismo valor.

En la práctica el número de agrupaciones real u óptimo es desconocido. Es frecuente seleccionar un número alto de grupos, teniendo en cuenta que los grupos pueden ser unidos después que el proceso ha terminado o durante uno de los pasos de interacción.

Cuando no se tiene un conocimiento a priori de los centros de los grupos que permita llevar a cabo la inicialización, éstos generalmente son seleccionados uniformemente espaciados a lo largo de la diagonal que une el origen del espacio espectral con el punto correspondiente a los máximos valores de brillantez. Esta selección puede ser redefinida, si el usuario ha calculado el histograma, en tal caso los centros de los grupos serán inicializados uniformemente a lo largo de la diagonal que une los extremos de los datos multiespectrales reales.

2.- Se barre todos y cada uno de los pixeles en la imagen, para establecer su localización X en el espacio espectral y medir su distancia con respecto a los candidatos a centro de grupo, $D(X, M_i)$. Las distancias de un pixel a los diferentes centros son examinadas y el pixel es asignado al candidato a centro de grupo más cercano (criterio de mínima distancia). Esta asignación será hecha con base en la medición de distancia euclidiana o de alrededor de bloque o cualquier otra que halla sido seleccionada.

3.- Para cada M_i se calcula un valor medio a partir de los pixeles que haya sido asignados a encontrarse alrededor de él. De esta forma se calcula un nuevo juego de medias denotado por:

$$m_i \quad i=1,2,\dots,c$$

Se compara el juego de candidatos a centro de grupos, M_i , con el juego de valores medios m_i , de tal forma que

Si $m_i = M_i$ para toda i , el procedimiento es terminado.

Si $m_i \neq M_i$ entonces $M_i = m_i$ y el control es retornado a 2.

En este paso se permite examinar los agrupamientos para ver si alguno de los grupos contiene tan pocos elementos que pueda ser despreciado, si dos o más centros de grupo son tan cercanos que pueden ser unidos o si un agrupamiento demasiado alejados, debe ser separado en dos nuevos grupos.

La mayor limitación de la técnica de migración de medias es la necesidad de especificar el número de centros de agrupamiento. Si esta especificación es demasiado grande, entonces una unión de grupos deberá ser hecha posteriormente. Si es demasiado pequeña se tendrán algunas clases multimodales.

Sin importar si la especificación inicial de centros de grupo es grande o pequeña, el algoritmo es computacionalmente caro debido al gran número de iteraciones y operaciones que se realizan.

Al terminar un agrupamiento, a todos aquellos pixeles dentro de un grupo se les asigna un símbolo para indicar que todos ellos pertenecen al mismo grupo. Usando estos símbolos se puede producir un mapa de grupos o clases espectrales.

3.3.2 CLASIFICACIÓN POR AGRUPACIÓN DE DOS PASADAS

Este algoritmo recibe su nombre debido a que barre la imagen dos veces. En la primera pasada, el programa construye las agrupaciones y obtiene el vector de media asociado con cada una de ellas. En la segunda pasada, un algoritmo de clasificación por mínima distancia al valor medio es aplicado a toda la imagen, donde cada pixel es asignado a uno de los grupos creados en la pasada uno.

Inicialmente el analista proporciona:

- ◆ Un radio en el espacio espectral (R), usado para determinar cuándo una nueva agrupación deberá ser formada.
- ◆ Un parámetro de distancia espectral (C), usado cuando se unen dos agrupaciones.
- ◆ El número de pixeles (N), a ser evaluado antes de hacer una unión de agrupaciones o entre uniones de agrupaciones.
- ◆ El número máximo de agrupaciones C_{max} , que serán identificadas por el algoritmo.

Comenzando por el origen de la imagen (línea 1, columna 1), los pixeles son evaluados secuencialmente de izquierda a derecha. Después que una línea es procesada la siguiente línea de datos es evaluada.

Primero, se determina que el pixel 1 en la imagen, X_1 , representa el vector de medias para la agrupación uno (M_1 , la dimensión n del vector depende del número de bandas usadas en la clasificación no supervisada).

$$M_1 = X_1, n_1 = 1 \quad \text{y} \quad cont = 1$$

Donde n_1 es el número de elementos en la agrupación 1 y $cont$ es el número de agrupaciones definidas.

Se considera el pixel 2, X_2 , y se calcula la distancia espectral de este pixel a la agrupación 1, $D(X_2, M_1)$.

Si $D(X_2, M_1) > R$ y $cont < C_{max}$ entonces $M_2 = X_2, n_2 = 1$ y $cont = 2$

Debido a que todavía no se han identificado todas las agrupaciones (C_{max}).

Si $D(X_2, M_1) < R$ entonces $n_1 = 2$ y $M_1 = \frac{1}{2}(X_1 + X_2)$

Se considera el siguiente pixel X_i y se calcula la distancia espectral de éste a cada una de las agrupaciones, $D(X_i, M_i)$

Si $D(X_i, M_i) > R$ y $cont < C_{max}$, entonces $M_j = X_i, n_j = 1$ y $cont = cont + 1$

donde M_j representa una nueva clase definida y $j < C_{max}$. Debido a que todavía no se han identificado todas las agrupaciones (C_{max}).

Si $D(X_i, M_i) < R$, entonces $n_i = n_i + 1$ y $M_i = \left(\frac{1}{n_i}\right)\{n_i(M_i) + x_i\}$

donde n_i es el número de elementos de la clase M_j antes de ser incrementado en uno.

La acumulación de agrupaciones continúa hasta que el número de pixeles evaluados es mayor que N o $cont$ es mayor o igual que C_{max} . En este punto, el programa checa si existen agrupaciones que se unan. Para lograr esto se calcula la distancia entre los vectores medios de las agrupaciones $D(M_i, M_j)$ y verifica si existen dos agrupaciones separadas una distancia espectral menor que C .

Si $D(M_i, M_j) < C$, entonces $n_i = n_i + n_j, M_i = \left(\frac{1}{n_i}\right)\{n_i(M_i) + n_j(X_j)\}$ y $cont = cont + 1$

donde n_i es el número de elementos de la clase M_j antes de ser incrementado en n_j .

Este proceso se ejecuta hasta que no hay más grupos separados una distancia menor que C . Entonces el siguiente pixel es considerado. El proceso continúa iteractuando hasta que la imagen entera es examinada.

Algunos programas no evalúan cada línea y cada columna de la imagen para calcular los vectores medios de las agrupaciones. En vez de esto, ellos realizan un muestreo cada i -ésimo renglón y cada j -ésima columna para identificar el C_{max} de agrupaciones. La decisión de muestreo en la imagen es dejada al analista.

Mientras más puntos son agregados a una agrupación, el desplazamiento de la media es menos drástico debido a que la nueva media calculada es pesada por el número de pixeles actuales en la agrupación. El punto final es la posición espectral del vector de medias que es usado como una firma en el clasificador de mínimas distancias aplicado en el paso 2.

Los datos finales del vector de medias son usados dentro de un algoritmo de mínimas distancias para clasificar todos los píxeles en la imagen dentro de alguna de las agrupaciones calculadas.

En contraste de la clasificación supervisada, la clasificación no supervisada requiere de una cantidad mínima de entradas por parte del analista. Es un proceso por medio del cual ciertas operaciones numéricas son ejecutadas para buscar la agrupación natural de las propiedades espectrales de los píxeles, siendo examinados dentro del espacio multispectral de los rasgos.

El usuario permite que la computadora seleccione las medias de las clases y las matrices de covariancia a ser usadas en la clasificación. Una vez que los datos son clasificados, el analista intenta, a posteriori, asignar estas clases naturales o espectrales a las clases de información de interés.

Para llevar a cabo la correlación entre las clases espectrales y las clases de información es necesario evaluar la localización de las agrupaciones en la imagen y en diagramas de dispersión bi o tridimensionales, etiquetarlas si es posible, y ver si cualquiera de ellas se puede combinar.

La etiquetación de las agrupaciones, es generalmente ejecutada, por medio de desplegar interactivamente todos los píxeles asignados a una agrupación individual sobre la pantalla. De esta manera es posible identificar, su localización y asociaciones espaciales. Este análisis visual interactivo, en conjunto con la información proporcionada por los dispersogramas, permite al analista asociar las agrupaciones dentro de clases de información.

Generalmente hay más información presente en una clasificación no supervisada que en una supervisada, esto se debe a que muchas veces no se muestrean todas las clases durante el proceso de entrenamiento.

CAPÍTULO IV
CASO PRÁCTICO

CONTENIDO

4 CASO PRÁCTICO

Introducción

4.1 Imagen de un muelle

- 4.1.1 Descripción
- 4.1.2 Delimitación de requerimientos
- 4.1.3 Selección de técnicas de procesamiento y clasificación
- 4.1.4 Análisis y discusión de resultados

4.2 Imagen de Washington DC

- 4.2.1 Descripción
- 4.2.2 Delimitación de requerimientos
- 4.2.3 Selección de técnicas de procesamiento y clasificación
- 4.2.4 Análisis y discusión de resultados

4.3 Imagen de un área urbana

- 4.3.1 Descripción
 - 4.3.2 Delimitación de requerimientos
 - 4.3.3 Selección de técnicas de procesamiento y clasificación
 - 4.3.4 Análisis y discusión de resultados
-

CAPÍTULO IV

CASO PRÁCTICO

INTRODUCCIÓN

El objetivo de este capítulo es de mostrar la aplicación de las técnicas de procesamiento y clasificación de imágenes según lo requieran las tres imágenes que se van a analizar.

Se desarrollo un programa ocupando Visual Basic 5, porque era el lenguaje de programación que más conocía y porque permite crear una interfaz gráfica para el manejo del programa de forma sencilla y rápida, sin embargo, Visual Basic 5 no es capaz de descomponer el nivel de intensidad de un pixel en RGB, tampoco puede leer el valor del pixel, etc. estos problemas se resolvieron al crear una dll en Visual C ++ 5 que posteriormente se manda a llamar desde Visual Basic.

El programa consta de los siguientes módulos:

Archivos. Se ocupa para abrir el archivo donde se encuentra la imagen a analizar.

Procesamiento. Solo se consideraron los módulos de programación de filtro paso bajas, filtro paso altas, modificación del histograma y moda, debido a las características que presentan las imágenes era necesario mejorar el contraste, crear una superficie más uniforme, realzar bordes y eliminar pixeles aislados por medio de la moda. También se necesito el falso color para resaltar u observar rasgos de la superficie que a simple vista no pueden ser discriminados.

Clasificación. Se programo el método de minimas distancias.

Tanto el procesamiento como la clasificación se realizó en el dominio espacial.

Para la manipulación del histograma se ocupo Corel 5 como software de apoyo.

4.1 IMAGEN DEL MUELLE

4.1.1 DESCRIPCIÓN

La Imagen 4.1 muestra un muelle con varios botes anclados, fue tomada desde un helicóptero por una cámara fotográfica. A pesar de que no fue capturada desde un satélite, se decidió tomarla en cuenta, debido a su complejidad, que consiste en la diversidad de tonalidades que maneja cada objeto (muelle, mar, botes), ocasionando que la tonalidad de un pixel puede encontrarse en los tres objetos al mismo tiempo.

Aplicando un preanálisis visual a la imagen 4.1, si se deseara clasificar al mismo tiempo el mar, el muelle y los botes, el principal problema es que probablemente el tono de un pixel puede pertenecer tanto al mar o al muelle o a un bote, es aquí donde radica la importancia de esta imagen para aplicar alguna técnica de clasificación.

Y también porque dependiendo de las características de la imagen como: contraste, ruido, textura, definición de bordes, nos permitirá llevar a la práctica los métodos de procesamiento digital de imágenes investigados, como son: modificación del histograma, filtro paso bajas, filtro de media etc., según lo requiera la imagen.

La imagen tiene un tamaño de 500 x 333 pixeles.
Cada pixel debe representar aproximadamente 40 centímetros cuadrados.

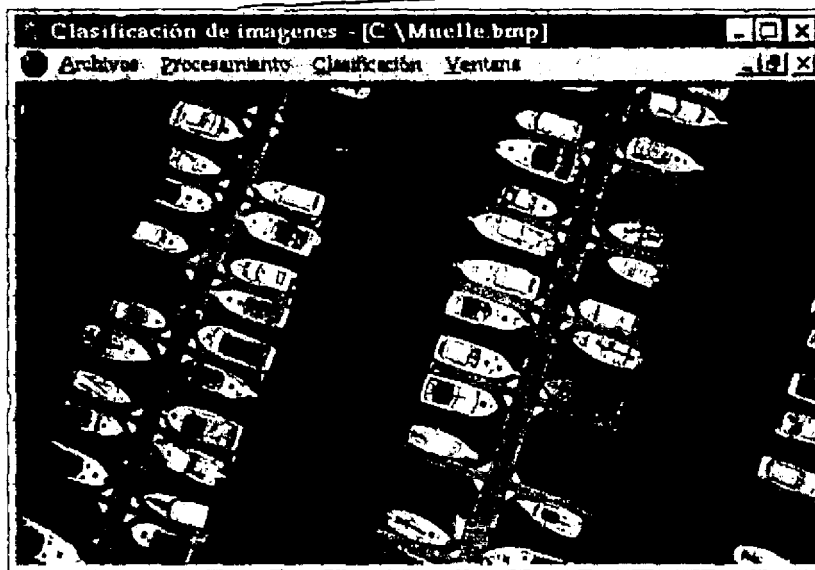


Imagen 4.1 Puerto

4.1.2 DELIMITACIÓN DE REQUERIMIENTOS

El objetivo de esta imagen es clasificar el mar, el muelle y los botes. También se desea eliminar de la imagen el mar y el muelle, dejando solo los botes con su tonalidad original.

Los bordes que delimitan un objeto de otro esta bien definidos, por lo que no tendremos problemas por esa parte, tampoco existe mucho ruido al grado de que distorsione algún bote, muelle o el mar, pero no se puede empezar la clasificación, debido a que presenta algunas partes oscuras y esto puede ocasionar que se pierda información que existe, y que por medio de un análisis visual no es posible percibir. Para lograr obtener toda la información será necesario mejorar el contraste de la imagen.

La mejora del contraste se logrará con la manipulación del histograma, ya que si un histograma esta muy sesgado hacia uno de los dos lados nos puede indicar si tiene poco o mucho contraste.

A lo que se desea llegar, es que el histograma de la imagen 4.1 este sesgado a la derecha para obtener un mejor contraste.

4.1.3 SELECCIÓN DE TÉCNICAS DE PROCESAMIENTO Y CLASIFICACIÓN

De acuerdo con las características de la imagen antes mencionada, el problema que presenta es que si se clasifica tal como esta hay riesgo de perder información por el bajo contraste, por lo tanto se analizará el histograma de la imagen 4.1.

Como se puede observar en la figura 4.2 la imagen no posee un histograma con una distribución normal y las altas frecuencias se encuentran sesgadas al lado izquierdo de la media, indicando que dentro de la imagen predominan las tonalidades oscuras.

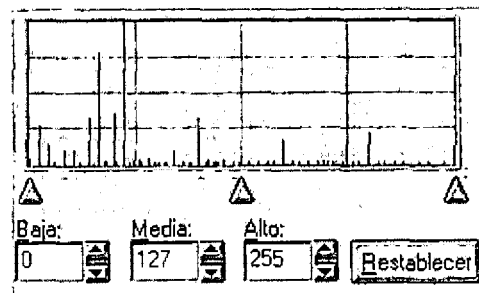


Figura 4.2 Histograma de la imagen 4.1

No es recomendable reducir el intervalo de las tonalidades debido a que después de la inspección visual lo más probable es que un pixel pueda pertenecer a dos o a los tres objetos a identificar y por lo tanto si reducimos el intervalo de tonalidades, la probabilidad de que un pixel pertenezca a más de una clase que identifica a los objetos a clasificar será mayor. Lo que haremos es mover la media hacia el lado izquierdo, para que las altas frecuencias se carguen hacia la derecha, esto nos ayudara a tener un mejor contraste en la imagen (ver figura 4.3).

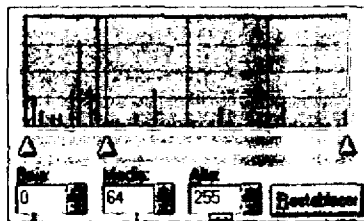
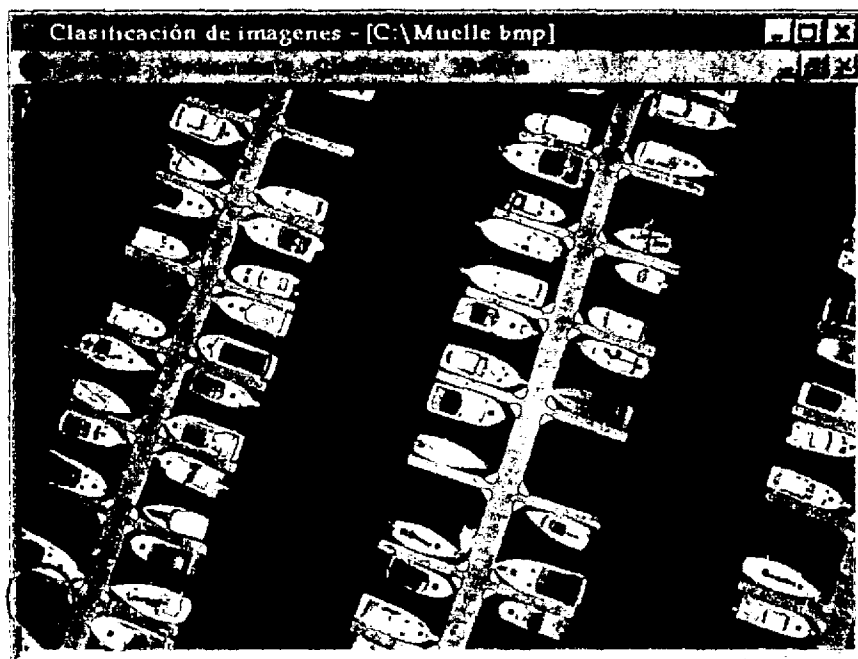
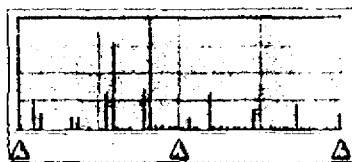


Figura 4.3 Modificación de la media del histograma de la imagen 4.1

La imagen 4.4a presenta el resultado obtenido al modificar el histograma.



(a)



(b)

Imagen 4.4 (a) Resultado después de modificar su histograma
(b) Histograma de la imagen 4.4a

El círculo de color rojo en la imagen 4.4a es una prueba de los resultados que se pueden obtener con la modificación del histograma. El círculo rojo nos indica la información que existe dentro de la imagen pero debido a las características del histograma de la imagen 4.1 no se podía identificar a simple vista.

Una vez que se tiene la imagen lista para ser clasificada, se seleccionan las clases de entrenamiento de cada objeto a clasificar.

En nuestro caso fueron necesario más de dos clases para cada objeto a clasificar.

¿Cómo seleccionar las muestras para que posteriormente se formen las clases?. Aquí se opto por generar una rutina, que consiste en lo siguiente:

Con el mouse se selecciona un punto dentro de la imagen, una vez que se tiene el punto se da un click en el botón derecho del mouse, donde como resultado muestra una pequeña pantalla donde nos indica el valor RGB del pixel seleccionado (Ver imagen 4.5), esto se hace tantas veces como sea necesario para encontrar las muestras más representativas para la clase.

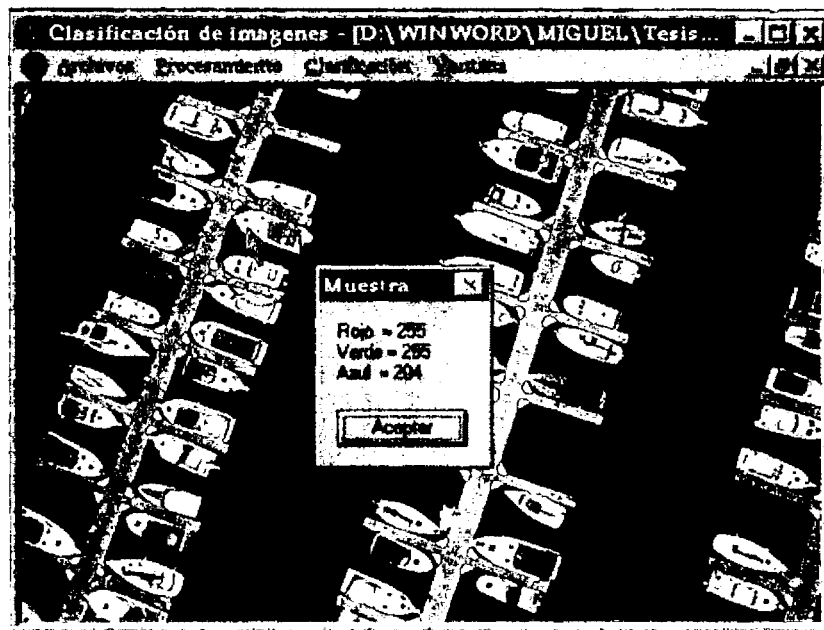


Imagen 4.5 Muestra como se obtienen los valores RGB de un pixel

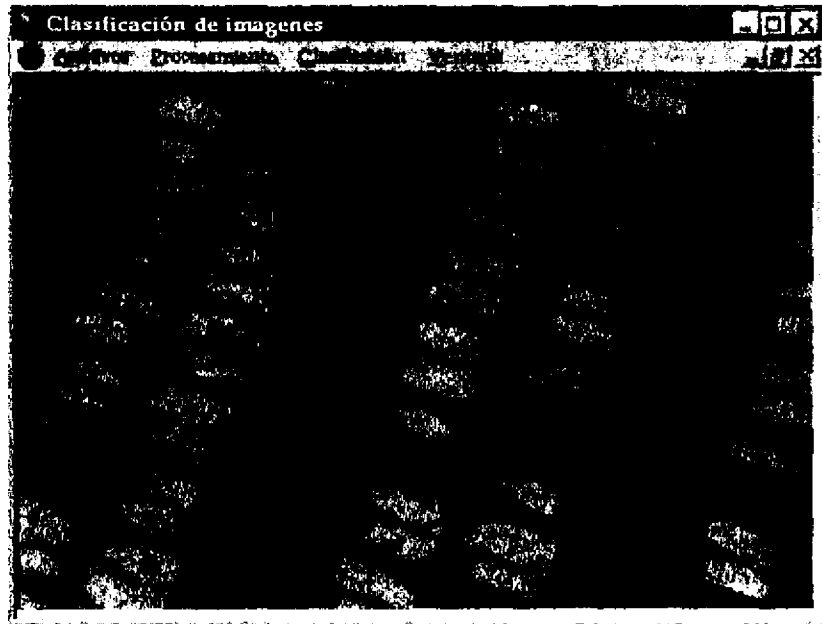
Ya que se tienen las clases se selecciona el método de clasificación.

A pesar de que fueron varias muestras por cada objeto a clasificar, las muestras son muy pequeñas y no servirían para aplicar métodos de clasificación que involucran probabilidad y estadística, cuando se tienen muestras pequeñas, es recomendable ocupar el método de mínimas distancias porque como su nombre lo dice, el criterio que toma para determinar si un pixel pertenece a una clase o no, es la distancia mínima que existe a la media de cada clase.

Para obtener los resultados deseados fue necesario agregar un criterio adicional al método de mínimas distancia, que consiste en lo siguiente:

Alrededor de cada pixel se analizaron los pixeles que se encontraban a su alrededor de un tamaño de una máscara de 3x3 pixeles. Si el pixel central cumplía con la mínima distancia, el segundo criterio es determinar si la mayor parte de los pixeles que lo rodeaban tenían la misma tonalidad que el pixel central, si así era el caso, se podía asegurar que formaba parte del muelle, mar o de los botes.

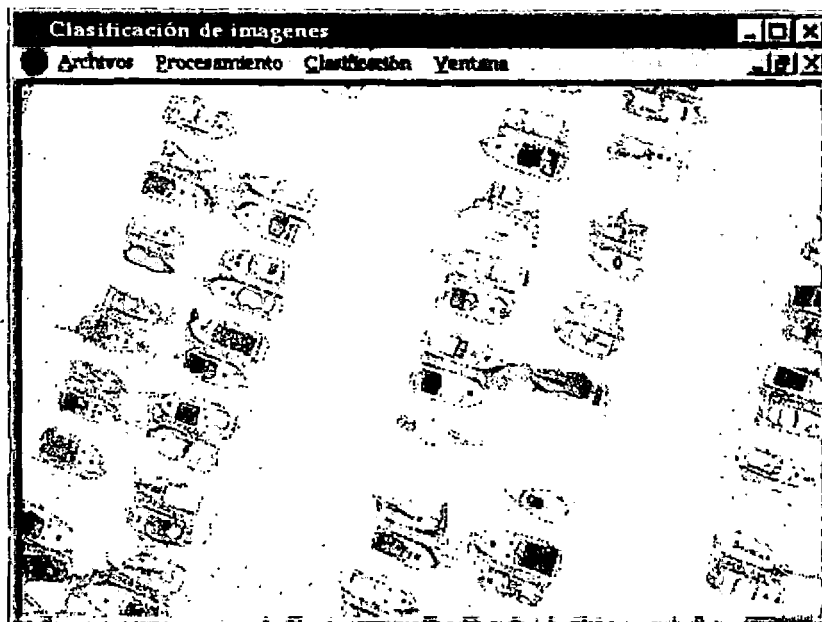
A continuación se muestran los resultados de la clasificación del mar, muelle y de los botes.



4.6 Resultado de la clasificación de mínimas distancias.

Los resultados obtenidos en la imagen 4.6 son muy buenos porque la imagen por si sola es capaz de indicar que estamos clasificando mar, muelle y botes, debido a que los bordes de cada objeto a clasificar no se perdieron durante la clasificación aunque existen algunas partes en la imagen 4.5 donde el grado de complejidad para clasificar era mayor, debido a que tanto el muelle y los botes principalmente cumplan con características muy similares en el ámbito de tonalidades.

Si se desea eliminar de la imagen 4.4a el muelle y el mar, entonces una vez que se tiene el resultado de la clasificación en la imagen 4.6, se barre la imagen 4.6 y todo lo que esta en color azul (mar) y color verde (muelle) se elimina de la imagen 4.4a, obteniendo como resultado la imagen 4.7.



4.7 Muestra el resultado de la clasificación de los botes

Como se puede observar en la imagen 4.7, se eliminó casi en su totalidad el muelle y el mar, pero se pierde información con respecto a los botes, lo que aparentemente se consideraba una buena clasificación resulta que aun le falta aplicarle algo para obtener los resultados deseados.

Para recuperar más información de los botes, lo que haremos, es aplicar un filtro de moda a la imagen 4.6. El filtro de moda tiene la finalidad de determinar si el píxel central de un área de 3x3 píxeles pertenece al objeto (botes) o no, esto se logra analizando los píxeles que se encuentran alrededor del píxel central, si las tonalidades que predominan alrededor de éste píxel, en este caso muy en particular son de color rojo (botes) entonces se puede asegurar que el píxel central forma parte de los botes aunque tenga tonalidad verde (muelle) o azul (mar) y, por consiguiente, el píxel no se pondrá en blanco dentro de la imagen 4.4a.

El resultado después de aplicar el filtro de moda sobre los botes se muestra en la imagen 4.8

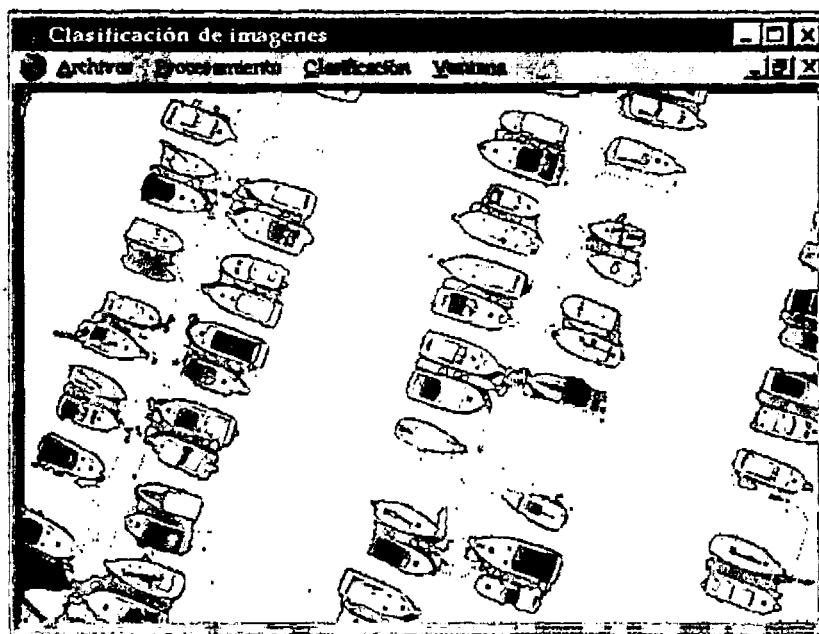


Imagen 4.8 Resultados obtenidos después de aplicar el filtro de moda la imagen 4.6, con la finalidad de mejorar la información con respecto a los botes

A pesar de que se ganó una mejor definición en los botes, también apareció información no deseada dentro de la clasificación, por lo tanto se aplicara un filtro de moda a la imagen 4.6, pero en esta ocasión será para definir bien el muelle y después eliminarlo de la imagen 4.8, aunque en la imagen 4.8 el muelle en su mayoría ya no existe, pero con esto se lograra eliminar en su totalidad los pixeles que no pertenecen a los botes, ver imagen 4.9.

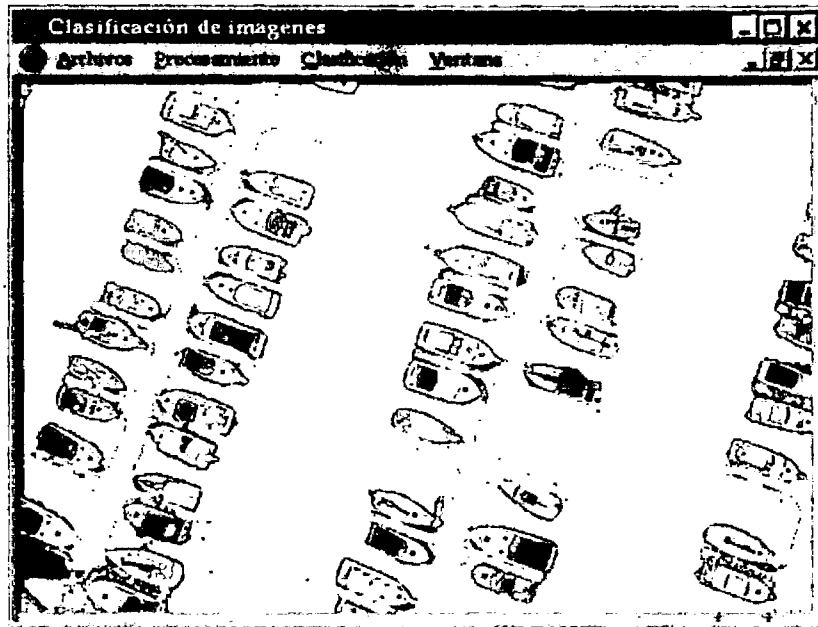


Imagen 4.9 Resultados obtenidos después de aplicar el filtro de moda la imagen 4.6, con la finalidad de mejorar la información con respecto al muelle y que posteriormente fue eliminada de la imagen 4.8

Como lo muestra la imagen 4.9 se logra obtener un mejor resultado, y esto se considera así porque se logra clasificar a los botes, recuperando aproximadamente más del 90% de la información y casi se logra eliminar los píxeles que no pertenecen a los botes.

Para obtener la imagen 4.9 a partir de la imagen 4.6, se consideró un solo proceso que consiste en barrer la imagen 4.6 dos veces. En el primer barrido a la imagen 4.6 se aplica el filtro de moda para definir mejor los botes, para que posteriormente nos permita identificarlos en la imagen 4.4a y poner en tonalidad blanca todo aquello que no pertenezca a los botes. El segundo barrido es un filtro de moda para definir el muelle y después poner en blanco los píxeles en la imagen 4.4a que pertenecen al muelle.

Los resultados intermedios se manejan en memoria.

4.1.4 ANÁLISIS Y DISCUSIÓN DE RESULTADOS

Como se pudo demostrar, el procesamiento es una parte muy importante antes o después de la clasificación de una imagen, porque si se aplica antes nos permite eliminar ruido, mejorar el contraste, definir bordes, mejorar la textura etc., todo con la finalidad de obtener una imagen que permita una clasificación independiente del método que se utilice, en la imagen 4.1 presenta un bajo contraste, motivo por el cual no es posible detectar visualmente la información restante. Para mejorar el contraste de a imagen 4.1 se ocupo la manipulación del histograma.

A pesar de que fueron varias clases las que se ocuparon para identificar cada objeto. El tamaño de las muestras para cada clase es muy reducida y esto no permite obtener buenos resultados con métodos de clasificación que manejen probabilidad y estadística como son: el teorema de Bayes o por máxima verosimilitud.

Cuando el número de muestras de entrenamiento por clase es limitado, resulta más efectivo emplear un clasificador que no haga uso de la información de la covarianza sino que dependa únicamente de la posición de las medias de las clases. Esta fue la razón por la cual se decidió ocupar el método de mínimas distancias y no otro.

La clasificación también se puede ocupar para eliminar objetos, y dejar sólo aquellos que nos interesan, en este caso en particular se elimino el mar y el muelle de la imagen 4.4a y se dejo todos los botes sin alterar su tonalidad.

Por la variedad de tonalidades que tienen los botes y muchas de ellas coinciden con el mar y con el muelle, fue necesario agregar un criterio adicional al método de mínimas distancias que consiste en que una vez que es clasificado el pixel se analizan los que se encuentran a su alrededor con una máscara de 3x3 y si existe más del 50% de los pixeles que pertenecen a la misma clase del pixel central, se considera que el pixel se ubica dentro del área espacial que pertenece al objeto a clasificar y si sucede lo contrario se descarta que el pixel pertenezca al objeto.

Una vez obtenida la clasificación del muelle, mar y botes (ver imagen 4.6) se ocupa esta imagen para lograr el segundo objetivo, que es eliminar el muelle y el mar realizando un mapeo de la imagen ya clasificada (ver imagen 4.6) sobre la imagen procesada (ver imagen 4.4a), los pixeles que representan mar y muelle (imagen 4.6) son eliminados de la imagen procesada y solo se deja los botes sin afectar sus tonalidades obteniendo una imagen solo con botes (ver imagen 4.7), pero se pierde información de los botes y esto se puede deber a:

- a) El método de clasificación no era el apropiado, pero esto se descarta, debido a que las muestras son muy pequeñas y el método de clasificación que da mejores resultados cuando las muestras son pequeñas es el de mínimas distancias.
- b) El procesamiento que se aplico a la imagen no era el correcto o le falto aplicar uno adicional. En la imagen 4.1 en particular solo era necesario mejorar su contraste para mostrar la información restante que no era posible ver a simple vista.

Para recuperar la información se manejo un procesamiento adicional después de obtener la clasificación. Primero se aplico un filtro de moda a la imagen ya clasificada (imagen 4.6) con la finalidad de definir mejor las áreas que representan los botes, la imagen resultante se manejo en memoria para que posteriormente se realizara un mapeo que consiste en eliminar todos los pixeles de la imagen 4.4a que no representan los botes, dando como resultado la imagen 4.8, donde se logra recuperar información de los botes pero trae información no deseada, la cual se elimina aplicando un segundo filtro de moda sobre la imagen 4.6 para recuperar la información restante del muelle y del mar guardando el resultado en memoria para que posteriormente se aplique un mapeo ocupando esta ultima imagen que esta en memoria y la imagen 4.8, se elimina todo lo que sea muelle o mar y da como resultado la imagen 4.9 donde se muestra los botes. Con esto se comprueba que el procesamiento puede ir antes o después de la clasificación.

4.2 IMAGEN DE WASHINGTON DC

4.2.1 DESCRIPCIÓN DE LA IMAGEN

La imagen 4.10 muestra una parte de la ciudad de Washington, DC, tomada por el satélite Landsat TM el 23 de Agosto de 1994, esta imagen fue formada con las bandas 3,2,1 (RGB) y tiene una resolución espacial de 5x5 metros.

Las tonalidades blancas que se presentan dentro de la imagen 4.10 o en cualquier otra imagen que se analice son generadas cuando la intensidad de luminosidad que presenta esa área satura al sensor remoto.



Imagen 4.10 Ciudad de Washington, DC, tomada por el satélite Landsat TM el 23 de Agosto de 1994

4.2.2 DELIMITACION DE REQUERIMIENTOS

El objetivo es identificar el área verde dentro de la imagen 4.10.

Con un análisis previsual se puede observar, que donde se encuentra los edificios agrupados se carece de área verde y muy probablemente una gran parte del área verde esta dañada, por lo tanto, va a ser necesario mejorar el contraste de la imagen con la finalidad de recuperar áreas que no estas dañadas pero que por el bajo contraste, no se puede asegurar en la situación actual en la que se encuentran, esto se lograra utilizando la manipulación del histograma, ya que por medio de un vistazo rápido al histograma correspondiente a la imagen 4.10 podremos saber si una imagen esta muy oscura o muy clara (normalmente, si un histograma esta muy desplazado hacia uno de los lados significa que estamos perdiendo niveles de gris), si tiene poco o mucho.

4.2.3 SELECCIÓN DE TÉCNICAS DE PROCESAMIENTO Y CLASIFICACIÓN

Como se puede observar en el punto anterior esta imagen en particular necesita ser procesada antes de ser clasificada. Cuando se menciona que la imagen necesita ser procesada, esto no implica que se tenga que aplicar todos los métodos de procesamiento existentes, en este caso, lo que se espera después del procesamiento, es mejorar el contraste de la imagen y esto se lograra con la manipulación de histograma.

La mejora del contraste se lograra moviendo la media del histograma hacia la izquierda

La figura 4.11 muestra el comportamiento del histograma de la imagen 4.10

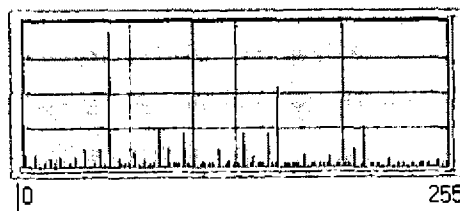
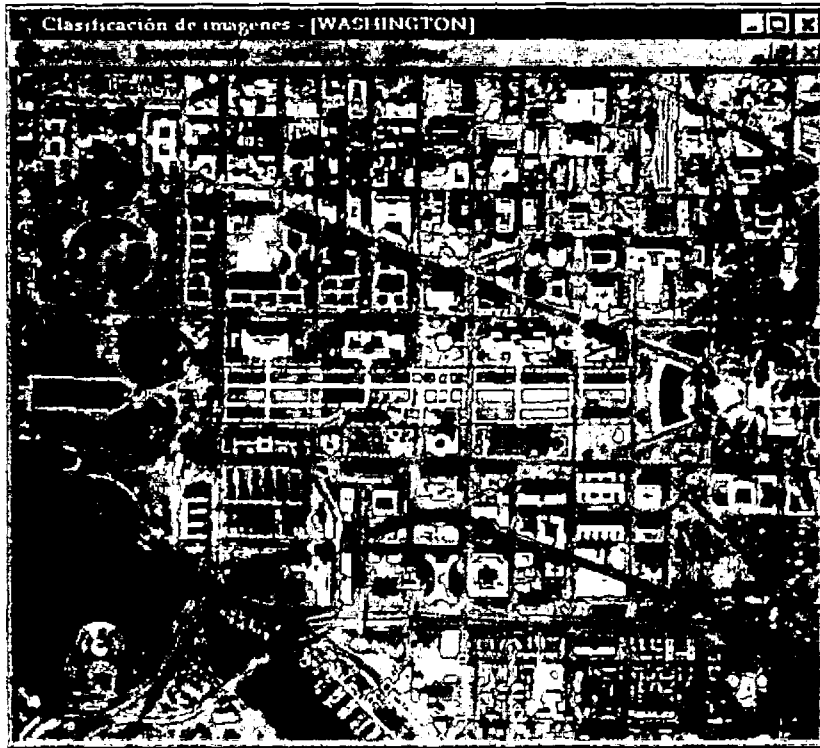
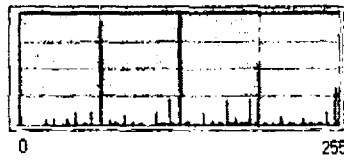


Figura 4.11 Histograma de la imagen 4.10

Los resultados de la manipulación del histograma se muestran en la imagen 4.12.



(a)



(b)

Imagen 4.12 (a) Resultado después de modificar el histograma
(b) Histograma actual de la imagen 4.10

Como se puede observar en la imagen 4.12 las tonalidades oscuras tienden a ser más claras permitiéndonos obtener más información de la imagen.

Una vez que la imagen queda lista para ser clasificada se empieza por la selección de las muestras para obtener las clases, esto se obtiene de la misma forma como se realizó en la imagen del muelle que se clasificó antes que esta (imagen 4.1).

Para identificar el área verde se definieron cinco clases, cuatro de ellas representan el área verde y la quinta es la clase rechazo.

Para clasificar el área verde se ocupará el método de mínimas distancias debido a que las muestras son muy pequeñas.

El resultado de la clasificación se puede observar en la imagen 4.13

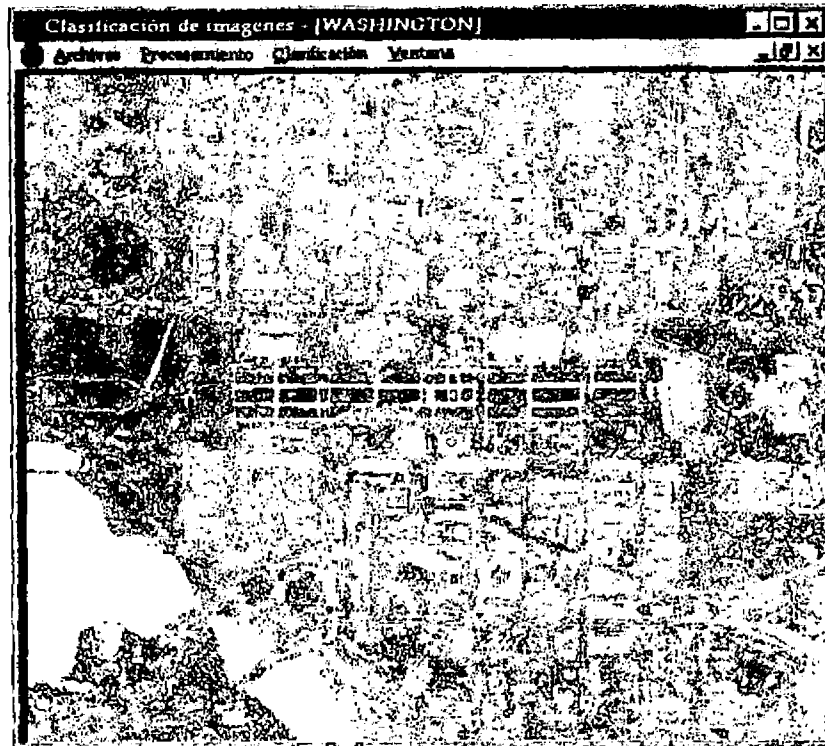


Imagen 4.13 Resultado después de aplicar el método de clasificación de mínimas distancia a la imagen 4.10e

Como se puede observar en la imagen 4.13, después de la clasificación aun quedaron píxeles que cumplan con las características de lo que se desea clasificar dentro de la imagen, por lo tanto, se decidió aplicar a la imagen 4.13 un filtro de moda con una máscara de 3x3 (ver imagen 4.14).

Se decidió el filtro de moda porque nos permite analizar los píxeles vecinos que se encuentra a su alrededor y si el píxel central se encuentra muy aislado se puede interpretar esto como que no pertenece o no forma parte del área verde a pesar de que pertenezca a alguna.



Imagen 4.14 Resultado después de aplicarle un filtro de moda con una máscara de 3x3 a la imagen 4.11

Los resultados obtenidos en la imagen 4.11 mejoran la interpretación de la clasificación.

Para tener una mejor opinión sobre los resultados obtenidos, la imagen 4.15, muestra los resultados finales de la identificación del área verde en color rojo sobre la imagen 4.12a.

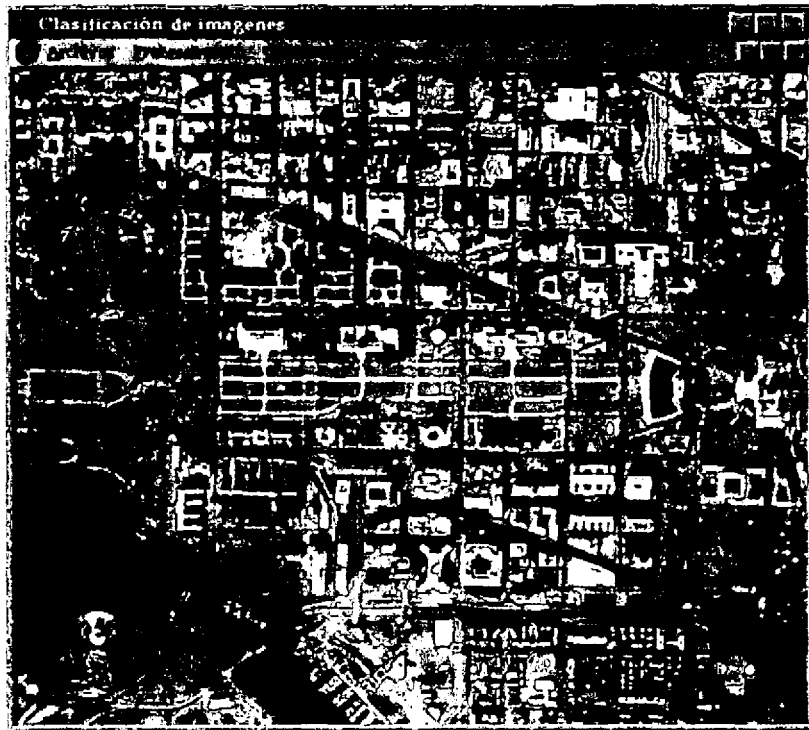


Imagen 4.15 Muestra el resultado de la clasificación en color rojo sobre la imagen 4.12a.

4.2.4 ANÁLISIS Y DISCUSIÓN DE RESULTADOS

El resultado de la clasificación es el esperado ya que se comprueba que las zonas urbanizadas tienden a carecer de áreas verdes y si llegara a existir alrededor de las áreas urbanizadas, como es el caso de esta imagen, las áreas verdes tienen mayor probabilidad de estar dañadas parcialmente como se puede observar en la imagen 4.15 que muestra el resultado de la clasificación en color rojo.

Se recurre al método de mínimas distancias para la clasificación, porque las muestras son pequeñas, pero trae información no deseada, sin embargo no se disminuyen las clases porque se perdería más información a comparación de los píxeles adicionales que trae la clasificación, se opta por aplicar un filtro de moda a la imagen ya clasificada con la finalidad de agrupar las áreas verdes y desaparecer los píxeles que se encuentran muy aislados.

4.3 IMAGEN DE UN ÁREA URBANA

4.3.1 DESCRIPCIÓN DE LA IMAGEN (área urbana)

La siguiente imagen muestra una vista panorámica de una área urbana



Imagen 4.16

4.3.2 DELIMITACIÓN DE REQUERIMIENTOS

Dentro de las técnicas de procesamiento de imágenes digitales, el realce de bordes tienen como objetivo que la imagen pueda ser visualizada a través de una pantalla de computadora de tal forma que se pueda distinguir las fronteras de los objetos para que el sistema visual humano pueda hacer la mejor interpretación posible de lo que en ella aparece.

Esta sección tiene como objetivo la aplicación de la técnica de realce de bordes aplicando filtros espaciales paso alta.

Diversos factores pueden influir en la ausencia de borde en una imagen, el más relevante puede ser, sensores con un rango de valores de iluminación bajo o cualquier otro factor equivalente. La idea subyace en este tipo de transformación es modificar los valores de los píxeles de manera que se produzca un realce en las fronteras de los objetos, esto se logra aplicando un filtro paso alta a la imagen.

Debido a la resolución que tiene la imagen es un poco complejo distinguir donde empieza y termina un edificio, también el problema que presenta la imagen es de poseer un bajo contraste que provoca que no se pueda detectar más información de forma visual.

4.3.3 SELECCIÓN DE TÉCNICAS DE PROCESAMIENTO Y CLASIFICACIÓN

Se aplicará un filtro espacial paso bajas para suavizar la imagen, es decir, las altas frecuencias que corresponden a las transiciones abruptas, son atenuadas y con esto se logra eliminar el ruido existente en la imagen

El filtro espacial paso bajas que se ocupara es el siguiente:

$$\begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{pmatrix} \frac{1}{16}$$

El resultado obtenido a partir de aplicar el filtro paso bajas se muestra en la imagen 4.17.



Imagen 4.17 Aspecto de la imagen original después de aplicarle un filtro paso bajas

La desventaja de aplicar un filtro paso bajas es la disminución del contraste y la pérdida de bordes como se puede observar en la imagen 4.17, pero este efecto se puede contrarrestar si se aplica un filtro espacial paso alta que tiene el efecto de resaltar el realce de los bordes. Las transiciones entre regiones diferentes se toman más nítidas.

El filtro paso alta que se ocupara es el siguiente.

$$\begin{pmatrix} -1 & -1 & -1 \\ -1 & 9 & -1 \\ -1 & -1 & -1 \end{pmatrix} \begin{matrix} 1 \\ 1 \\ 1 \end{matrix}$$

El resultado obtenido a partir de aplicar el filtro paso alta se muestra en la imagen 4.18.



Imagen 4.18 Resultado después de aplicar el filtro espacial paso alta.

Con el filtro espacial paso alta se logra tener una mejor definición de los bordes como se muestra en la imagen 4.18 ya que disminuye considerablemente los efectos colaterales de bajo contraste y la pérdida de bordes que muestra la imagen 4.71.

4.3.4 ANÁLISIS Y DISCUSIÓN DE RESULTADOS

Como se pudo observar solo se muestran las máscaras de los dos filtros espaciales que se ocuparon para llegar a los resultados deseado. La razón de ello es porque no existe una formula para definir que valores debe tener la máscara todo en su mayor parte se basa en prueba y error, la otra parte restante depende de la experiencia del comportamiento de la imagen con determinadas características con respecto a algunas máscaras y del conocimiento de la formula de aplicación del filtro.

El filtro espacial paso bajas tiene el efecto de suavizar la imagen, es decir, las altas frecuencias que corresponden a las transiciones abruptas, son atenuadas y con esto se logra eliminar el ruido existente en la imagen.

La desventaja de aplicar un filtro paso bajas es la disminución del contraste y la pérdida de bordes como se puede observar en la imagen 4.17, pero este efecto se puede contrarrestar si se aplica un filtro espacial paso alta que tiene el efecto de resaltar los bordes. Las transiciones entre regiones diferentes se toman más nitidas y esto se puede observar en la imagen 4.18

Enseguida se demuestra como el tamaño de la máscara y los valores que contiene influyen en éxito o fracaso de la aplicación del filtro.

A la imagen original se le aplicó el siguiente filtro.

$$\begin{matrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{matrix} \cdot \frac{1}{16}$$

Los resultados se muestran en la imagen 4.20



Imagen 4.20. Aplicación de un filtro paso bajas con una mascara de 5x5

Como se puede observar en la imagen 4.20 existe una pérdida de bordes y una disminución del contraste que impide que se clasifique a la imagen.

CONCLUSIONES

En el presente trabajo se ha realizado una investigación desde los elementos que conforman la captura de una imagen, pasando por los tipos de procesamiento más usados y llegando hasta los métodos de clasificación más conocidos y se concluye con la aplicación práctica sobre algunas imágenes con objetivos muy específicos.

Los resultados y conclusiones más importantes que pueden obtenerse tras la realización de este trabajo pueden resumirse en los siguientes puntos.

- ♦ Para obtener los resultados deseados o esperados cuando se clasifica una imagen, ayuda mucho la experiencia de haber trabajado con otras imágenes ya que esto nos permitirá detectar anomalías que se presentan dentro de la imagen que impidan que el método de clasificación que se aplique tenga el mayor éxito posible. Aunque no es indispensable la experiencia pero es de gran ayuda para tomar el camino más corto para obtener los resultados deseados.

- ♦ Las imágenes analizadas en el capítulo IV por contener muestras muy pequeñas por clase se empleó el método de mínimas distancias llegando a los siguientes resultados:

Cuando el número de muestras de entrenamiento por clase es limitado puede ser más efectivo emplear un clasificador que no haga uso de la información de covariancia, sino que dependa únicamente de la posición de las medias de las clases espectrales (método de mínimas distancias).

La utilización de un conjunto reducido y representativo de muestras seleccionadas del conjunto de entrenamiento original reduce el error cometido en la clasificación ocupando el método de mínimas distancias, aunque se ha puesto de manifiesto que en determinadas circunstancias algunos píxeles que pertenecen al objeto de clasificación pueden llegar a desaparecer o tender a ser más de los que no deberían ser, ya que muchos de ellos pertenecen a otros objetos. Sin embargo este efecto puede contrarrestarse con la aplicación de la moda que es un método de procesamiento que permite decidir si un píxel pertenece o no, a un objeto por medio del análisis de sus píxeles vecinos.

- ♦ A pesar que no se ocupó otro método de clasificación debido a las características de las imágenes analizadas en el capítulo IV se puede concluir que:

Los métodos de clasificación supervisada requieren el conocimiento de los valores de algunos parámetros tales como las matrices de covarianza, los vectores de medias, las probabilidades a priori, las probabilidades condicionadas, etc. Estos parámetros expresan algún tipo de conocimiento preliminar acerca del problema y deben estimarse a partir de áreas de entrenamiento o utilizando otros métodos de estimación en los que interviene, además de la imagen a clasificar, la clasificación previa que se desea corregir incorporando la información contextual.

Cuando se conoce la "forma" de la función de probabilidad y por lo tanto sólo se deben estimarse algunos parámetros: se habla entonces de métodos paramétricos. Si no se conoce la forma de la función de probabilidad se habla de métodos no paramétricos. Los métodos paramétricos son generalmente más fáciles de implementar pero requieren un alto grado de conocimiento a priori acerca de la naturaleza de los patrones. Los métodos no paramétricos son potencialmente mejores considerando el grado de ajuste del valor estimado con el verdadero valor de probabilidad, aunque generalmente tienen asociado un alto coste computacional y requieren un elevado número de muestras de entrenamiento.

Los métodos de clasificación no supervisada, las identidades de los tipos de cobertura a ser especificadas como clases dentro de una escena, generalmente no son conocidas a priori, porque se carece de un trabajo de campo o los rangos de la superficie en la imagen no están bien definidos. La computadora es requerida para agrupar los píxeles dentro de diferentes clases espectrales, de acuerdo a algún criterio estadístico determinado. Es por tanto responsabilidad del analista etiquetar estos grupos.

Como conclusión final.

Existen varios métodos de procesamiento de imágenes tanto para el dominio espacial o frecuencial, pero esto no implica que sea necesario aplicar todos a una imagen y en un determinado orden, siempre dependerá de las características que tenga la imagen y esto también aplica para los métodos de clasificación.

También no existe un orden en particular al aplicar el procesamiento y la clasificación, o viceversa, o bien que no exista una etapa de procesamiento.

Indicar que métodos de clasificación dan mejores resultados no sería adecuado porque la selección de una regla de clasificación o decisión depende de la naturaleza y características de los datos de entrada y de que se desee como salida, a partir de una imagen, lo mejor es establecer un esquema apropiado de clasificación para las coberturas.

Si ya existen sistemas de clasificación reconocidos en la literatura internacional, es mejor adoptar o modificar éstos que inventar uno nuevo que sea usado exclusivamente por uno mismo. Esto permite mostrar la importancia de los resultados de nuestra clasificación a la luz de otros estudios y facilitar la mezcla de datos.

BIBLIOGRAFÍA

- ♦ **La percepción remota: nuestros ojos desde el espacio**
Jorge Lira
Editorial Fondo de Cultura Económica, S.A. de C.V.
México D.F. 1995
- ♦ **Satellite Communication Systems Design**
Tirró
Editorial Sebastiano Tirrò
U.S.A. 1993
- ♦ **Digital Image Processing**
Rafael C. González, Richard E. Wood
Editorial Addison-Wesley Iberoamericana, S.A.
U.S.A. 1992
- ♦ **Signal System and Transforms**
Leand B. Jackson
Editorial Addison – Wesley Publishing Company
U.S.A. 1991
- ♦ **Digital Image Processing and Analysis**
Rama Chellappa and Alexander A. Sawchuk
Editorial IEEE Computer Society Press
U.S.A. 1985
- ♦ **Fundamentals of Digital Image Processing**
Anil K. Jain
Editorial Prentice Hall
New Jersey 1989
- ♦ **Digital Image Processing**
William K. Pratt
Editorial Awiley – Interscience Publication
U.S.A. 1991
- ♦ **Digital Image Processing and Computer Vision**
Robert J. Schalkoff
Editorial Wiley John Wiley & Sons
U.S.A 1989
- ♦ **Remote Sensing and Image Interpretation**
T.M. Lillesand y R. Kiefer
Editorial John Wiley & Sons
1987
- ♦ **Classification of Remotely Sensed Image**
Thomas. V.M. Benning, y N.P. Ching
Editorial Adam Illger
1987

-
- ◆ The WWW Virtual Library:
<http://vl.bwh.harvard.edu/cgi-bin/htsearch>

 - ◆ Remote Sensing Introduction - An Overview of Remote Sensing Remote Sensing is an extensive science, drawing from many areas for support and development. It depends greatly on the support of governments and private industries worldwide. Satellite and digital imagery play an..
<http://www.ucalgary.ca/UofC/faculties/SS/GEOG/Virtual/rem...>

 - ◆ Lycos Environment Guide: Remote Sensing - Help Feedback Search for: Lycos Network: Tripod.com | Angelfire.com | MailCity.com | WhoWhere.com | HotBot.com All Sites.. Click here to visit site Lycos Home> Education> Areas of Study> Sciences> Earth science> Environment> Geology> Top 5..
http://www-prodigy.lycos.com/wguide/wire/wire_484377_5066...

 - ◆ Thesis on Image Texture Analysis using Zero Crossings Information - Doctoral thesis in Image Texture Processin
<http://www.cssip.elec.uq.edu.au/~guy/thesis/home.html>

 - ◆ Color Image Processing Resources - color theory and spaces, and image processing resources (links to hardware, software, and publications).
<http://eleceng.ukc.ac.uk/~rls3/>

 - ◆ The Fast Fourier Transform - Next: Background Up: Overview Previous: Overview 3. The Fast Fourier Transform Background Introduction to the Fourier Transform Sampling and the Sampling Theorem The Discrete Fourier Transform The Fast Fourier Transform -- FFT Structure of the..
<http://www.vislab.usyd.edu.au/vislab/CP3/Four3/node1.html>

 - ◆ The Inverse Transform - The Inverse Transfor
<http://www.engr.trinity.edu/~paul/fourier/fourier/node4.html>

 - ◆ Image classification - Image segmentation and classification Image classification After the segmentation process which subdivides an image into its constituent parts or objects, classification is the following procedure to extract further information from the image..
<http://kazan.zeam.fu-berlin.de/euromet/courses/english/sa...>

 - ◆ Land Use/Land Cover Classification - Land Use/Land Cover Classification Classification of remotely sensed imagery is an effective method for rapid mapping of land use and land cover information. In supervised classification, which was used for this work, the analyst decides categories
http://www.ead.anl.gov/~kuiper/apg_modeling/lulc.htm

 - ◆ Supervised Classification - Supervised Classification 10/12/98 Click here to start Table of Contents Supervised Classification Review Classification Schemes PPT Slide Two Types of Classification Supervised Classification Training Site Selection Training Site Selection..
<http://www.clas.ufl.edu/users/mbinford/geo5145/lecture15r...>

 - ◆ Image pixel classification - Image pixel classification Corresponding pixels in multispectral images may be classified through their relative densities in the different spectral bands. Scene features such as grassland, water, bare earth and built-up areas have characteristic..
http://www.phy.hw.ac.uk/~peckham/envphy_4/land_surface%21...

 - ◆ End-to-End Remote Sensing Tutorial Page 1-18 - Minimum Distance Classification We initiate our survey of supervised classification by producing one using the minimum distance routine. The IDRISI program acts on DN's in multidimensional band space to organize the pixels into the classes we chose.
http://eerst.gsfc.nasa.gov/Sect1/Sect1_18.html
-

-
- ◆ Publications (Unsupervised Classification) - Unsupervised Classification Our research addresses the issue of comparison of non-hierarchical unsupervised classification algorithms. We discuss various techniques for comparing classification methods and conclude that the most promising approach.

<http://www.cs.ualberta.ca/~upal/cluster/>

APÉNDICE

Banda espectral

A cada una de las imágenes digitales que forman una imagen multispectral se le conoce como banda o plano de color.

Brillantez

Es la magnitud de la respuesta producida en el ojo por la luz, y es una sensación subjetiva que puede ser determinada únicamente en forma aproximada. La luminancia es una medición cuantitativa de la intensidad de luz a partir de una fuente, y es medida con un dispositivo llamado fotómetro. Las variaciones de brillantez pueden ser calibradas a partir de una escala de gris.

En imágenes generadas a partir de barredores térmicos, la brillantez de un objeto es proporcional al calor irradiado desde el objeto. Para imágenes de radar la brillantez es proporcional al haz de energía siendo retroreflejado por el objeto y captado por la antena.

Civ (Campo Instantáneo de vista)

Sobre la escena se define un elemento con dimensiones físicas, denominado campo instantáneo de vista, cuyo tamaño depende de las dimensiones del paisaje, de la resolución del sensor remoto y de la distancia de éste al CIV.

El tamaño del campo instantáneo de vista define la resolución espacial, es decir da el grado de detalle que se puede discernir de la escena, que es a lo que se llama digitización.

Color

Se puede definir al color como la componente de radiación de longitud de onda delta/alfa que es reflejada (o emitida) por un objeto que recibe luz blanca, es decir, el objeto recibe luz con todas las longitudes de onda de la banda visible del espectro, pero refleja preferentemente las longitudes de onda delta de la banda delta/alfa que corresponde a un color dado (dependiendo de la longitud de onda que se trate).

Contraste

La definición de contraste de una imagen es la razón entre las partes más brillantes y más oscuras de la imagen. Varias expresiones son comúnmente usadas para describir el contraste, puede haber confusión en su uso. La razón de contraste (C_r) es ampliamente usada y es definida como (Sabin, 1978):

$$C_r = \frac{B_{max}}{B_{min}}$$

Note que cuando $B_{min} = 0$, C_r no está definido: cuando $B_{min} = B_{max}$, $C_r = 1$

Además de describir una escena completa la razón de contraste es también útil para describir la razón de brillantez entre un objeto y su fondo adyacente en la imagen. La razón de contraste es un factor vital para determinar la posibilidad de discriminar y detectar objetos.

Detectibilidad

Es la habilidad de un sistema formador de imágenes para registrar la presencia o ausencia de un objeto, aunque la identificación del objeto pueda ser desconocida. Un objeto puede ser detectado aunque éste sea más pequeño que la resolución espacial del censor.

Firma

Una firma es la expresión de un objeto en una imagen que capacita al objeto para ser identificado.

Las firmas son determinadas por las características de un objeto que determina su interacción con la energía electromagnética. Las firmas de los objetos son generalmente diferentes para cada longitud de onda.

Imagen Digital Multiespectral

- ◆ Una Imagen Multiespectral es aquella que esta formada por un conjunto de imágenes digitales, cada una de ellas conteniendo aquellos aspectos de la escena que corresponden a un color determinado.
- ◆ Es la variación espacial y espectral de la "energía luminosa" proveniente de la escena.

Pixel

Es la representación numérica o lógica del campo instantáneo de vista y es la medida de la energía promedio que proviene de dicho campo; a cada pixel le corresponde solo un campo.

Patrón

- ◆ Es definido como un arreglo espacial de los objetos. El patrón es una importante llave para descubrir el origen o función, de un objeto.

Algunos patrones son básicamente culturales, otros naturales, dando origen a muchos patrones que resultan de la interacción de factores naturales y culturales.

Regiones culturales (tales como ciudades, carreteras, reforestación, áreas agrícolas, etc.) son claramente visibles ya que estas se conforman por líneas rectas u otras configuraciones regulares. Muchas de las actividades del hombre dejan marcas en la tierra que perduran por un largo tiempo, aun después de que las actividades hayan cesado (sitios arqueológicos).

- ◆ Es un arreglo de entes sin tomar en cuenta la naturaleza, de éstos. Puede afirmarse que el universo está formado por patrones. Por su parte, un ente es cualquier objeto físico o lógico: un número, una figura geométrica, o una ave son ejemplos de entes. Entonces la sucesión de números 4,17,19,16, es un patrón, diferente al patrón 17,4,16,19.

En la representación numérica de una imagen digital sólo podemos hablar de patrones lógicos, no así en la representación visual de la misma imagen, donde éstos tienen una representación física.

Patrón Espectral

Es la sucesión de valores sobre los diferentes planos que componen la imagen digital multiespectral.

La sucesión de valores que presenta un CIV determinado cuando es visto bajo diferentes filtros, esta sucesión es única para cada objeto de la naturaleza, de tal forma que dicho objeto tiene un patrón espectral único.

En una imagen multiespectral digital los patrones espectrales son siempre de tipo lógico y los patrones espaciales son lógicos en la representación numérica y física en la representación visual.

Reconocibilidad

Es la habilidad para identificar un objeto en una imagen. Los objetos pueden ser detectados y discriminados y aun así no ser reconocibles.

Resolución

Existen cuatro características de una imagen de satélite con relación a su capacidad de resolución las cuales son:

- ◆ Resolución espacial: Es la capacidad de un sensor para distinguir dos o más elementos, a los cuales se les puede registrar como entidades separadas, donde la resolución es la distancia mínima entre ellos.

En una imagen, la resolución espacial está relacionada con el tamaño del área en la superficie, que representa un pixel.

-
- ◆ **Resolución espectral:** Se interpreta como la capacidad de un sensor para separar y detectar ciertos rangos de longitud de onda del espectro electromagnético.

Para una imagen, la resolución espectral está asociada con el número de bandas que posee, sus posiciones en el espectro electromagnético y el rango de frecuencias que cada una de ellas cubre.

- ◆ **Resolución radiométrica,** se define como la capacidad de un sensor para detectar la diferencia de intensidad, entre dos mediciones, cuando existe una variación en la energía que recibe el instrumento.

La resolución radiométrica en una imagen de satélite está dada por el número de niveles de gris que posee.

- ◆ **Resolución temporal (período de pasaje).** Es el tiempo mínimo en el que se puede obtener dos imágenes satelitales en forma consecutivas para la misma zona de la superficie terrestre.

Textura

Es una manifestación espacial de la frecuencia de cambios y arreglos de los tonos de una imagen. Textura fina media y gruesa son algunos de los términos usados para describirla.

Tono

El término tono es usado para cada sombra distinguible desde el negro hasta el blanco.

En sistemas sensores como cámaras fotográficas y barredores multispectrales, el tono de un objeto es primeramente determinado por la habilidad del objeto para reflejar la luz incidente, aunque los efectos atmosféricos y la sensibilidad del sistema sensor son también factores importantes que lo influncian.