



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE INGENIERÍA

DIVISIÓN DE INGENIERIA ELÉCTRICA ELECTRÓNICA

*“Textura multiespectral a partir de una variación
del algoritmo Local Binary Pattern”*

T E S I S

QUE PARA OBTENER EL GRADO DE:

INGENIERO
(INGENIERÍA EN COMPUTACIÓN)

PRESENTA:

BERENICE SÁNCHEZ SERRALDE

DIRIGIDA POR:

DR. JORGE ARTURO LIRA CHÁVEZ



Ciudad Universitaria, México, D.F.
2011



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



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.

Dedicatoria

A mis padres por haberme criado en una familia llena de amor y de valores.

A mi mamá Eloisa por su infinito amor y por enseñarme que las cosas se pueden lograr cuando uno quiere.

A mi papá Francisco por su paciencia, comprensión y amor.

A mi bebé Yaolitzin por ser la mayor bendición que me ha dado Dios.

A mi hermana Lorena y mis sobrinos Ámbar y Yael por estar siempre conmigo.

A mis abuelos Guadalupe, Odilón, Petra y José.

Agradecimientos

Gracias al Dr. Jorge Lira Chávez por ser mi asesor, por haberme permitido estar en el Laboratorio de Percepción Remota, por haberme facilitado una computadora y libros para poder realizar este trabajo, por haberme aceptado en la clase de P.R. en el postgrado de Geofísica, por haberme apoyado para poder terminar la tesis, por darme su ejemplo de profesionalismo. Gracias de todo corazón ya que sin su ayuda no habría sido posible la realización de este trabajo.

Gracias a mis sinodales por haber aceptado revisar este trabajo y estar presentes en el examen profesional.

Gracias al Mtro. Alejandro Martín Rodríguez Aguayo por su amistad y por todo su apoyo durante mi estancia en el Laboratorio, por toda su paciencia y ayuda tanto profesional como humana. A la Lic. Ericka G. García por su amistad y apoyo mientras estuve en el Laboratorio.

Gracias a la U.N.A.M y a la Facultad de Ingeniería por haberme permitido estudiar mi carrera. Gracias a todos mis profesores en la Facultad de Ingeniería ya que todos me transmitieron sus conocimientos.

Gracias al Instituto de Geofísica por haberme permitido realizar mi Servicio Social en el Departamento de Cómputo, también gracias a todo el personal administrativo que hizo muy acogedora mi estancia en el instituto.

Gracias al Comité de Becas del Instituto de Geofísica y al Dr. Jorge Lira por haberme brindado una beca para realizar mi tesis.

Gracias al Dr. Boris Escalante por haberme permitido asistir a la clase de Procesamiento Digital de Imágenes en el IMMAS. Gracias a la Maestra Clemencia Santos del Instituto de Geografía por darme asesorías y permitirme ir al Instituto.

Gracias al Lic. en Geografía Juan Nieto Hernández por brindarme asesoría para poder terminar este trabajo.

Gracias al Maestro Guillermo Sandoval por haberme dado trabajo desde el 2005 como profesora.

Gracias a mis padres Eloisa y Francisco por todo su amor, por su apoyo, por sus enseñanzas, por su comprensión, por su paciencia, por darme todos los recursos para estudiar una carrera, por transmitirme sus valores y costumbres.

Gracias a mis amigos de la Facultad de Ingeniería con los cuales compartí las aulas y soñamos con llegar a este momento.

Gracias a Dios por permitirme concluir este trabajo.

“POR MI RAZA HABLARÁ EL ESPÍRITU”

Cd. Universitaria, D.F., primavera 2011.

Berenice Sánchez Serralde

Introducción

La **Percepción Remota** es una ciencia moderna que ha cobrado importancia en las últimas décadas con el avance de la tecnología espacial; es un método de observación a distancia de un sistema físico.

La percepción remota nació como una de las técnicas relacionadas con las sondas espaciales Voyager enviadas a planetas cercanos al nuestro y tuvo un desarrollo acelerado con la prospección de los recursos naturales desde el espacio y ha culminado con una nueva rama en la ciencia (Lira, 1987).

La percepción remota involucra métodos de análisis de imágenes, diseño y construcción de arquitecturas computacionales espaciales, fabricación de sensores de alta eficiencia, y desarrollo de software computacional eficiente.

Una situación típica de la percepción remota involucra un detector adquiriendo una imagen de un objetivo el cual se encuentra a una gran distancia. La imagen es con frecuencia formada en un intervalo de longitudes de onda diferente al de la sensibilidad del sistema de visión humano.

Por ejemplo, un satélite con órbita geoestacionaria podría generar una imagen en infrarrojo de la costa noreste de México; como otro ejemplo, un radiotelescopio en el desierto del Sahara podría generar imágenes de micro-ondas de sitios distantes. Tales imágenes poseen un número interesante de problemas en la interpretación.

Las imágenes obtenidas por los satélites (imágenes multiespectrales) ofrecen una perspectiva única de la Tierra, sus recursos y el impacto que sobre ella ejercemos los seres humanos. La información extraída de las imágenes multiespectrales nos muestra una visión global de objetos y detalles de la superficie terrestre.

Los científicos dedicados al estudio del percepción remota, tales como geógrafos, geofísicos, vulcanólogos, etc., consientes de que la forma en la cual se interpreten los datos debe de ser la más adecuada dependiendo del estudio que se encuentren realizando se han dedicado al desarrollo de herramientas que faciliten la interpretación de las imágenes obtenidas.

Las observaciones o mediciones hechas a distancia con la ayuda de satélites llegan a formar una colección de datos en forma de imágenes que tienen que ser procesadas digitalmente por computadoras, por lo que no se puede pensar en la percepción remota sin el uso del análisis digital de imágenes. De acuerdo con Lira:

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

Un aspecto relevante a tomar en cuenta en el análisis de una imagen es la textura, ya que es una propiedad asociada con la superficie en estudio, como el suavizado, la rugosidad, la granularidad, la regularidad y la homogeneidad.

Además, en la textura podemos observar una repetición espacial de patrones en una superficie. Ejemplos de textura son: las puntadas de un suéter, las piedras de un camino empedrado y los granos de arena en la playa.

Estudiar las texturas en una imagen es importante ya que ellas están compuestas de varios subelementos estilizados que se repiten con cierto significado.

El operador **LBP** (**Local Binary Pattern**) es un método para analizar texturas y se ha convertido en una poderosa medida de la textura de una imagen, mostrando excelentes resultados en términos de precisión y complejidad computacional. El operador LBP puede ser visto como un acercamiento unificador hacia la estadística tradicionalmente divergente y los modelos estructurales del análisis de texturas.

Existen en la literatura numerosos métodos para caracterizar texturas; sin embargo su uso no es siempre satisfactorio al no considerar aspectos de las aplicaciones prácticas, como por ejemplo el uso de imágenes multiespectrales.

Con base en lo anterior es como surge el interés de realizar esta tesis: *“Textura multiespectral a partir de una variación del algoritmo LBP”* en la cual se propone tomar como referencia al operador LBP y su generalización para la cuantificación de textura multiespectral ya que es idealmente apropiado para aplicaciones que requieren de una rápida extracción de características de manera eficiente.

Esta tesis fue realizada en el **Laboratorio de Percepción Remota** del **Instituto de Geofísica, UNAM**, bajo la dirección del **Dr. Jorge Lira Chávez**.

Objetivo

Implementar una modificación del algoritmo LBP para el análisis textural de imágenes multiespectrales que de cómo resultado una correcta interpretación en el estudio de estas.

Objetivos particulares

- ◆ Interpretar la textura de una imagen multiespectral de una zona de estudio con ayuda del software desarrollado. La zona que se decidió estudiar es el Valle de México.
- ◆ Realizar una modificación del algoritmo LBP y combinarlo con el algoritmo de Divergencia para evaluar los resultados que arroja al analizar la textura y los bordes de una imagen multiespectral

Metodología

1.- Realizar modificaciones al algoritmo LBP para analizar textura de imágenes multiespectrales.

Debido a que los satélites proporcionan imágenes multiespectrales, habrá que realizar modificaciones al algoritmo LBP para adecuarlo a que pueda trabajar con imágenes multiespectrales.

2.- Realizar pruebas en MATLAB para comprobar la eficiencia del algoritmo LBP.

Debido a que MATLAB (abreviación de “MATriz LABoratory”) es un lenguaje de programación, que es una magnífica herramienta de alto nivel para desarrollar aplicaciones se programará primero en este entorno la modificación del algoritmo LBP para imágenes multiespectrales. Posteriormente se realizarán diversas pruebas para que el código sea el más eficiente posible en cuanto a extracción de características texturales como desempeño computacional.

3.- Realizar pruebas con imágenes multiespectrales

Tomar una zona de estudio con las imágenes multiespectrales correspondientes.

4.- Interpretar los resultados que arroje el algoritmo LBP

Analizar la imagen resultante la cual contendrá las características texturales de la zona en cuestión.

Contenido de la tesis

Capítulo I Introducción a la Percepción Remota

Con el objeto de que el lector tenga una mejor comprensión del objetivo de esta tesis se describen en este primer capítulo los temas principales y conceptos básicos referentes a la ciencia de la Percepción Remota, así como una introducción a las características de las imágenes que se ocuparan en el desarrollo del presente trabajo las cuales son imágenes multispectrales obtenidas del satélite Landsat TM.

Este capítulo tiene como objetivo que el lector se familiarice con ciertos términos y conceptos.

Capítulo II Imágenes

En este capítulo se toca el tema de la percepción visual, ya que en las aplicaciones del procesamiento de imágenes uno de los objetivos es mejorar la calidad visual de la imagen en un estudio adaptándola lo mejor posible a la capacidad perceptiva del sistema de visión humana, para ello se introducen conceptos básicos sobre la formación de las imágenes en el ojo humano.

Además se introduce una definición general del concepto de imagen y los tipos básicos de éstas.

Las imágenes obtenidas del sensado remoto ya sea desde un avión o una plataforma satelital están disponibles en formato digital, espacialmente los datos están compuestos por elementos discretos llamados pixels y radiométricamente están cuantificados dentro de niveles discretos de brillantez. La gran ventaja de tener los datos disponibles en formato digital es que pueden ser procesados por una computadora. Por esta razón este capítulo está también dedicado a las imágenes digitales para tener un panorama general del proceso de obtención y de las modificaciones que se le pueden realizar a las imágenes digitales.

Capítulo III Imagen Digital Multiespectral

En este capítulo se define lo que es una imagen digital multiespectral y algunos conceptos básicos sobre este tipo de imágenes, como por ejemplo: cuáles son los tipos de despliegue de estas imágenes, las relaciones de vecindad existentes en una imagen digital y los campos vectoriales asociados a las imágenes digitales multiespectrales

Capítulo IV Textura multiespectral

Este capítulo introduce varios conceptos para definir y analizar las texturas de imágenes complejas dentro de varios campos en el área científica, así como algunos métodos para analizar texturas, como por ejemplo la matriz de co-ocurrencia para imágenes monoespectrales además de la descomposición en componentes principales y del algoritmo de divergencia para cuantificar textura en imágenes multiespectrales

Capítulo V El Algoritmo LBP

Un descriptor de textura simple y relativamente nuevo es el Operador LBP (Local Binary Pattern, LBP). Este método está basado en el reconocimiento de ciertos patrones locales binarios, los cuales muestran propiedades fundamentales de la textura local de una imagen y la ocurrencia del histograma el cual provee una medida de las características texturales.

En este capítulo se detalla el funcionamiento del operador LBP además de algunas de sus aplicaciones en otros trabajos de investigación.

Como en principio el operador LBP no estaba pensado para imágenes multiespectrales, se realizó para esta tesis una modificación para adecuarlo al análisis de imágenes multiespectrales, en este capítulo se describen tales modificaciones y el proceso para llegar a éstas.

Capítulo VI Métodos y Materiales

En este capítulo se presentan algunos ejemplos de imágenes multiespectrales sobre los cuales se aplicó la modificación que se realizó al algoritmo LBP. Las imágenes de satélite que se emplearon pertenecen a la serie Landsat TM. Las escenas analizadas corresponden al centro de la República Mexicana.

También se discuten los resultados obtenidos al aplicar la modificación al operador LBP divergencia sobre las imágenes descritas en el capítulo anterior.

La metodología utilizada se presenta en el capítulo 5. Se analiza la información de la textura obtenida a partir de objetos espectrales específicos, se discuten los resultados obtenidos sobre una imagen y se discute la aplicación de la modificación del operador LBP.

Capítulo VII Conclusiones Generales

En este capítulo se presentan las conclusiones del presente trabajo.

Anexo A Plataformas orbitales empleadas en percepción remota

En este anexo se presentan las fichas técnicas de algunas de las principales plataformas orbitales empleadas actualmente en la percepción remota para la generación de imágenes multiespectrales.

Anexo B El Sistema Landsat

En este anexo se proporciona la información más relevante del sistema Landsat TM
Aportación de la tesis

La percepción remota es una ciencia cada vez más importante en las ciencias del medio ambiente y en las ciencias planetarias, pues hace posible coleccionar datos de áreas inaccesibles o peligrosas de exploración directa. Algunas de las aplicaciones incluyen el monitoreo de deforestación en áreas tales como el Amazonas en Brasil, los efectos del cambio climático sobre los glaciares en las regiones del ártica y antártica, el análisis del crecimiento urbano en ciudades como el DF, el monitoreo de contaminación de ríos, lagos y océanos y la localización de minerales en el área geológica, entre otros. El análisis de textura en las imágenes satelitales resulta de gran ayuda al momento de realizar un estudio en específico.

Índice

Introducción.....	1
Índice	7
Índice de Ilustraciones	9
1. Introducción a la Percepción Remota.....	12
1.1 Conceptos básicos.....	12
1.2 Elementos de la percepción remota.....	14
1.3 Radiación electromagnética.....	17
1.3.1 Propiedades de las ondas electromagnéticas	18
1.4 El espectro electromagnético.....	19
1.5 Obtención de la información	21
1.6 Interacción radiación - materia.....	24
1.7 Imágenes obtenidas en percepción remota	24
2. Imagen Digital	26
2.1 Percepción visual.....	26
2.1.1 Elementos de la percepción visual.....	27
2.1.2 Formación de la imagen en el ojo.....	28
2.2 Imagen	33
2.3 Tipos de imágenes	34
2.4 Imagen digital	36
2.5 Muestreo y cuantización.....	37
2.6 Modelos de color	39
2.7 Representación de imágenes digitales	42
2.8 Procesamiento digital de imágenes.....	45
2.9 Imágenes digitales en percepción remota.....	52
2.9.1 Elementos de una imagen digital.....	52
2.9.2 Características de una imagen digital	56
3. Imagen Digital Multiespectral	57
3.1 Imagen digital multiespectral	57
3.2 Despliegue de imágenes	60
3.3 Patrones y clase de patrones	62
3.4 Ruido y artefactos	65
3.5 Relaciones de vecindad	65
3.6 Campos vectoriales en una imagen multiespectral.....	70
4. Textura Multiespectral.....	74

4.1 Textura.....	74
4.2 Primitivas de textura.....	76
4.3 Características texturales	76
4.4 Clasificación de textura	78
4.5 Matriz de co-ocurrencia.....	80
4.6 Textura multiespectral	82
4.6.1 Descomposición en componentes principales	83
4.6.2 Divergencia de una Imagen digital multiespectral	97
5. El Algoritmo Lbp	101
5.1 Análisis de texturas con patrones locales binarios	101
5.2 Formalismo teórico.....	105
5.3 Contraste y patrones de textura	109
5.4 Modificaciones al algoritmo LBP para imágenes multiespectrales.....	110
6.- Métodos y Materiales.....	112
6.1 Aeropuerto de la Ciudad de México.....	115
6.2 Sur del Distrito Federal Xochimilco, Tláhuac.....	121
6.3 Volcán Popocatepetl e Iztacíhuatl	127
7.- Conclusiones	133
Anexos	135
A. Plataformas Orbitales.....	135
A.1 Plataformas geoestacionarias.....	135
A.1.1 Goes 9	135
A.2. Plataformas polares.....	136
A.2.1 Aster.....	136
A.2.2 Landsat 7.....	137
A.2.3 Spot 5	138
A.2.4 Quickbird 2	139
A.2.5 Ikonos 2	140
B. El sistema Landsat	141
Bibliografía.....	144

Índice de Ilustraciones

Ilustración 1 Imagen MODIS de la Tierra.....	12
Ilustración 2 Elementos Básicos de la Percepción Remota.....	14
Ilustración 3 Trayectorias de las plataformas Orbitales	16
Ilustración 4 Dispositivos de entrada y salida en el campo de la percepción remota	17
Ilustración 5 Espectro electromagnético.....	19
Ilustración 6 Geometría de observación en percepción remota.....	21
Ilustración 7 Componentes Radiométricos.....	23
Ilustración 8 Imagen óptica de la Ciudad de México (Composición a color RGB de las bandas 4,5 y3).....	25
Ilustración 9 Imagen de radar ASAR del norte del golfo de México, estado de Tamaulipas. Obtenida de: http://earth.esa.int	25
Ilustración 10 Representación visual de una imagen digital a través de una escala de niveles de gris	26
Ilustración 11 El ojo humano.....	27
Ilustración 12 Corte horizontal del ojo humano	29
Ilustración 13 Densidad de conos y bastones en el ojo humano	31
Ilustración 14 Ejemplos de imágenes binarias	35
Ilustración 15 Ejemplos de imágenes en escala de grises	35
Ilustración 16 Ejemplos de imágenes en color	36
Ilustración 17 Muestreo y cuantización de imágenes.....	39
Ilustración 18 Modelo de color RGB	40
Ilustración 19 Representación visual de una imagen digital con los valores numéricos de los pixels correspondientes.....	45
Ilustración 20 Realce de imágenes	47
Ilustración 21 Restauraciones de imágenes	47
Ilustración 22 Segmentación de imágenes.....	48
Ilustración 23 Región a segmentar delimitada por las líneas azul y roja	49
Ilustración 24 Segmentación de zonas.....	49
Ilustración 25 Segmentación espectral	50
Ilustración 26 Ejemplo de una segmentación de tipo contextual	50
Ilustración 27 Representación de una segmentación de tipo contextual	51
Ilustración 28 Esquema general que muestra la generación de una imagen digital multispectral a partir de una escena dada	52
Ilustración 29 Respuesta lineal de un sensor remoto frente a una cierta radiación incidente ..	54
Ilustración 30 Elementos de una Imagen digital	55
Ilustración 31 Imagen Digital Multiespectral.....	57
Ilustración 32 Imagen en tonos de gris	61
Ilustración 33 Pseudocoloración de una imagen	61
Ilustración 34 Composición RGB de la imagen multispectral (Bandas 7,4 y2).....	62
Ilustración 35 Imagen multispectral de cuatro bandas.....	63
Ilustración 36 Código de Freeman	66
Ilustración 37 Tipos de conectividad.....	67
Ilustración 38 Sistema de vecindad circular.....	69
Ilustración 39 Ejemplo de campo vectorial	71

Ilustración 40 Campo vectorial formado por un subconjunto de valores de pixels de una imagen multiespectral empleando las bandas 1 y 7.....	73
Ilustración 41 Textura de una escena	74
Ilustración 42 Diferentes ejemplos de textura	75
Ilustración 43 Superficie generada por los valores de los pixels de la Banda 1, la cual muestra la textura que representa dicha banda.....	82
Ilustración 44 Primeras dos bandas de una imagen multiespectral Landsat TM.....	85
Ilustración 45 Representación de la descomposición en componentes principales.....	86
Ilustración 46 Imagen multiespectral del valle de México.....	93
Ilustración 47 Primeras tres componentes principales de la imagen multiespectral del valle de México	95
Ilustración 48 Campos vectoriales.....	98
Ilustración 49 Imagen obtenida al aplicar el algoritmo LBP a una imagen monoespectral ...	101
Ilustración 50 Relaciones del operador LBP con otros métodos de textura.....	104
Ilustración 51 Diferentes conjuntos de vecinos circulares simétricos con valores de P y R diferentes	105
Ilustración 52 Composición RGB (7, 4,2) del valle de México	112
Ilustración 53 Composición RGB (7, 4,2) del Aeropuerto de la Ciudad de México.....	115
Ilustración 54 Primera Componente Principal sin bordes del aeropuerto de la Ciudad de México	116
Ilustración 55 Primera Componente Principal con bordes del aeropuerto de la Ciudad de México	116
Ilustración 56 Divergencia de la Primera Componente Principal sin bordes del aeropuerto de la Ciudad de México.....	116
Ilustración 57 Divergencia de la Primera Componente Principal con bordes del aeropuerto de la Ciudad de México.....	116
Ilustración 58 Mapa del área de Texcoco.....	117
Ilustración 59 Primera componente principal del área de Texcoco sin bordes	117
Ilustración 60 Primera componente principal del área de Texcoco con bordes	117
Ilustración 61 Divergencia de la Primera componente principal del área de Texcoco sin bordes	118
Ilustración 62 Divergencia de la Primera componente principal del área de Texcoco con bordes	118
Ilustración 63 Mapa de Ciudad Netzahualcóyotl	119
Ilustración 64 Primera Componente Principal del área de Netzahualcóyotl sin bordes.....	119
Ilustración 65 Primera Componente Principal del área de Netzahualcóyotl con bordes.....	119
Ilustración 66 Divergencia de la Primera Componente Principal del área de Netzahualcóyotl sin bordes	120
Ilustración 67 Divergencia de la Primera Componente Principal del área de Netzahualcóyotl con bordes.....	120
Ilustración 68 Composición RGB (7, 4,2) de la zona de Xochimilco y Tláhuac	121
Ilustración 69 Primera componente del sur del Valle de México sin bordes	122
Ilustración 70 Primera componente del sur del Valle de México con bordes	122
Ilustración 71 Divergencia de la Primera componente del sur del Valle de México sin bordes	122
Ilustración 72 Divergencia de la Primera componente del sur del Valle de México con bordes	122

Ilustración 73 Mapa de la zona de Chalco.....	123
Ilustración 74 Primera Componente Principal de la zona de Chalco sin bordes.....	123
Ilustración 75 Primera Componente Principal de la zona de Chalco con bordes.....	123
Ilustración 76 Divergencia de la Primera Componente Principal de la zona de Chalco sin bordes.....	124
Ilustración 77 Divergencia de la Primera Componente Principal de la zona de Chalco con bordes.....	124
Ilustración 78 Mapa del área de Xochimilco.....	125
Ilustración 79 Primera componente principal del área de Xochimilco sin bordes.....	125
Ilustración 80 Primera componente principal del área de Xochimilco con bordes.....	125
Ilustración 81 Divergencia de la Primera componente principal del área de Xochimilco sin bordes.....	126
Ilustración 82 Divergencia de la Primera componente principal del área de Xochimilco con bordes.....	126
Ilustración 83 Composición RGB (7, 4,2) de los volcanes Popocatépetl e Iztacíhuatl.....	127
Ilustración 84 Primera componente principal de los volcanes Popocatépetl e Iztacíhuatl sin bordes.....	128
Ilustración 85 Primera componente principal de los volcanes Popocatépetl e Iztacíhuatl con bordes.....	128
Ilustración 86 Divergencia de la Primera componente principal de los volcanes Popocatépetl e Iztacíhuatl sin bordes.....	128
Ilustración 87 Divergencia de la Primera componente principal de los volcanes Popocatépetl e Iztacíhuatl con bordes.....	128
Ilustración 88 Imagen del volcán Popocatépetl.....	129
Ilustración 89 Primera componente principal del volcán Popocatépetl sin bordes.....	129
Ilustración 90 Primera componente principal del volcán Popocatépetl con bordes.....	129
Ilustración 91 Divergencia de la Primera componente principal del volcán Popocatépetl sin bordes.....	130
Ilustración 92 Divergencia de la Primera componente principal del volcán Popocatépetl con bordes.....	130
Ilustración 93 Imagen del volcán Iztacíhuatl.....	131
Ilustración 94 Primera componente principal del volcán Iztacíhuatl sin bordes.....	131
Ilustración 95 Primera componente principal del volcán Iztacíhuatl con bordes.....	131
Ilustración 96 Divergencia de la Primera componente principal del volcán Iztacíhuatl sin bordes.....	132
Ilustración 97 Divergencia de la Primera componente principal del volcán Iztacíhuatl con bordes.....	132
Ilustración 98 Imagen Goes.....	135
Ilustración 99 Imagen Aster.....	136
Ilustración 100 Imagen Landsat.....	137
Ilustración 101 Imagen Spot.....	138
Ilustración 102 Imagen Quickbird.....	139
Ilustración 103 Imagen Ikonos.....	140
Ilustración 104 Satélite Landsat.....	141

Capítulo 1

1. Introducción a la Percepción Remota



Ilustración 1 Imagen MODIS de la Tierra

Esta imagen es la más detallada en color de la Tierra, se obtuvo usando una colección de observaciones basadas en satélite; científicos y visualizadores trabajaron juntos por meses haciendo observaciones de la superficie terrestre, océanos, hielos, nubes, etc. La mayor parte de la información contenida en esta imagen proviene del sensado remoto, del dispositivo de la NASA llamado Moderate Resolution Imaging Spectroradiometer (MODIS).

Imagen Obtenida de: <http://visibleearth.nasa.gov/>

1.1 Conceptos básicos

En la década de los 60's del siglo pasado, nuestra vista de la Tierra y el universo estaba restringida a las observaciones y fotografías usando la luz visible. Vistas distantes de algún lugar eran obtenidas solo de aviones y telescopios (Ilustración 1).

Hoy en día la ciencia de la Percepción remota provee instrumentos que ven nuestro universo en longitudes de onda diferentes que aquellas de la luz visible.

Tales instrumentos son desplegados en satélites y aviones para registrar imágenes de la Tierra y el sistema solar que pueden ser analizadas digitalmente con nuestras computadoras personales para proveer información sobre diversos temas. Algunos de éstos incluyen el medio ambiente, el uso de suelo, los recursos renovables y no renovables, desastres naturales, geología y urbanismo.

La percepción remota es la ciencia de la adquisición, proceso e interpretación de imágenes y datos relacionados, obtenidos de aviones y satélites que graban la interacción entre la materia y la radiación electromagnética.

Algunas definiciones de esta ciencia son:

“La percepción remota es la ciencia y el arte de obtener información sobre un objeto, área o fenómeno a través del análisis de datos adquiridos por un mecanismo que no está en contacto con el objeto, área o fenómeno de investigación.” (Lillesand, 2000).

“La percepción remota es la obtención de información acerca de una escena empleando radiación electromagnética (EM) en la región óptica y de microondas por medio del análisis automatizado de datos obtenidos a distancia por una sistema formado por un grupo de sensores remotos” (Lira, 2007).

De estas definiciones se puede considerar que esta ciencia involucra cuatro elementos fundamentales, los cuales son complementarios y no podríamos entender el uno sin el otro:

- a) Obtención de la información: Las imágenes obtenidas por un sensor remoto contienen un gran número de datos, que al ser analizados arrojan información proveniente de la escena.
- b) Radiación electromagnética: La interacción de la radiación con la escena genera información la cual es transportada por la señal electromagnética hacia los sensores remotos.
- c) Análisis automatizado de Imágenes: Es el conjunto de algoritmos computacionales programados organizados en un software de procesamiento de imágenes que permiten extraer información del comportamiento de la escena.
- d) Datos obtenidos a distancia: De las observaciones que efectúan los sensores remotos se obtienen los datos a través de medir la cantidad de radiación proveniente de la escena de estudio.

1.2 Elementos de la percepción remota

Con el desarrollo moderno de la ciencia de la computación y la electrónica digital se ha dado un auge importante a la percepción remota, pues 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 ésta, a su vez, ha generado nuevos sistemas de procesamiento digital con propósitos específicos.

A su vez, el desarrollo de la ciencia y la tecnología ha permitido incorporar nuevos métodos y dispositivos de captura a distancia, lo que ha hecho que la percepción remota extienda su campo de acción a fenómenos muy diversos.

Dentro de la percepción remota se encuentran los siguientes elementos básicos (Ilustración 2):

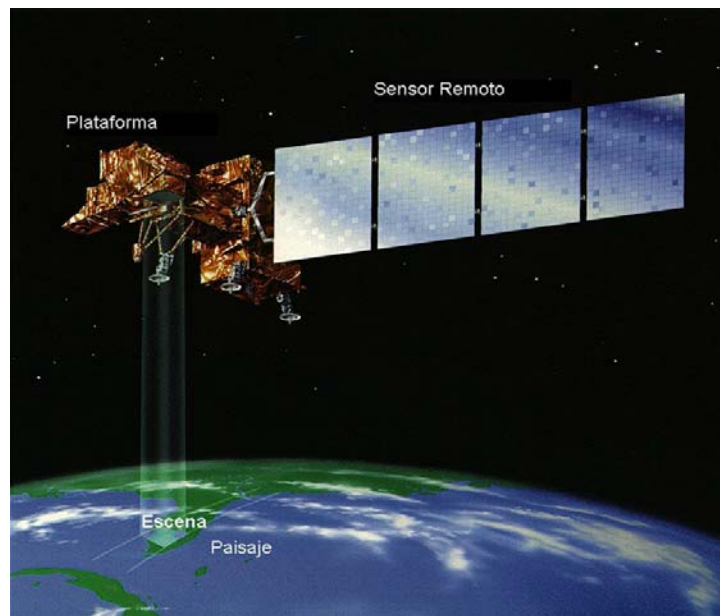


Ilustración 2 Elementos Básicos de la Percepción Remota

Fuente de Iluminación: Nos permite iluminar el paisaje y puede ser de tipo externa o interna.

Fuentes de iluminación externa: Hay sistemas en percepción remota que utilizan una fuente de iluminación externa sobre la cual no se puede tener control. Estos sistemas son llamados pasivos, y pueden utilizar al Sol o bien a una estrella lejana como fuente de iluminación. Como es un tipo de fuente en la que no se puede tener control, es necesario esperar a que las

condiciones de iluminación sean las adecuadas para efectuar la observación correspondiente de la escena.

Fuente de iluminación interna: Hay sistemas en percepción remota que utilizan su propia fuente de iluminación, y por lo tanto sí se puede tener control sobre ella. Estos sistemas son llamados activos. Estos sistemas son los que han permitido que se puedan estudiar fenómenos en donde no se tiene una fuente de iluminación externa. Estos sistemas pueden estar provistos de luz ultravioleta o infrarroja, así como de rayos gamma o rayos X, por lo que las aplicaciones pueden ser diversas.

Paisaje: Es una porción de la superficie terrestre que se encuentra bajo estudio. Este elemento no se puede controlar porque generalmente es una parte muy compleja donde intervienen muchos factores. En el análisis de un sistema de percepción remota se pueden llegar a subestimar la complejidad del paisaje ocasionando errores de operación y diseño.

Escena: Es una sección o superficie del paisaje donde enfocamos nuestro interés. Puede ser una superficie tanto física como virtual. La escena es una superficie que es parte del paisaje. Para estudiar un paisaje podemos estudiar varias escenas.

Sensor Remoto: Es un dispositivo que captura la radiación proveniente de la escena y permite generar una representación numérica de la misma, es decir, una imagen digital. El sensor realiza medidas a distancia de la radiación proveniente de la escena por unidad de ángulo sólido por unidad de tiempo.

Plataforma: Está compuesta por el lugar donde se encuentra montado el dispositivo o sensor remoto para poder obtener en conjunto la imagen de la escena, las hay de dos tipos:

a) Orbitales, las cuales se encuentran girando alrededor de la Tierra a cientos de kilómetros y de acuerdo a su trayectoria se clasifican en dos tipos: Polares y Geoestacionarias (Ilustración 3). Los satélites que tienen órbitas polares se les conocen como satélites de percepción remota. Los satélites que tienen órbitas geoestacionarias se les conocen como satélites del medio ambiente

b) Aerotransportadas, las cuales se encuentran montadas en aeronaves y por lo tanto se encuentran a poca altura de la superficie terrestre.

La imagen que se observe dependerá de la distancia de observación a la cual esté colocada la plataforma, así mismo, el detalle o grado de discernimiento que se pueda tener en la imagen. A medida que la plataforma de observación se coloca a una mayor distancia de la

escena, la visión de conjunto crece, pero no así el detalle, que consecuentemente se hace menor.

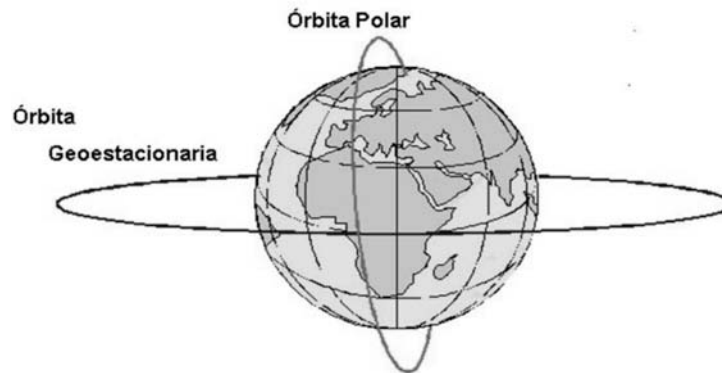


Ilustración 3 Trayectorias de las plataformas Orbitales

Para más información sobre plataformas satelitales, ver Anexo A *Plataformas Orbitales empleadas en percepción remota*

Sistema de procesamiento de Imágenes: Está formado por el grupo de dispositivos electrónicos y computacionales que permiten analizar cuantitativamente las imágenes que genera el sensor remoto (Ilustración 4).

Estos dispositivos procesan de manera cuantitativa o cualitativa, los datos que han sido proporcionados por el sensor remoto, ya sea en forma analógica o digital, para tener así una representación continua o discreta de la escena. Existen sistemas de procesamiento con un rango muy amplio de algoritmos para el análisis de imágenes, las cuales pueden ser analizadas de manera digital y obtener información que puede ser extraída, representada o manipulada de alguna manera para realizar una toma de decisiones con respecto a uno o varios aspectos de la imagen.

Existen dispositivos de procesamiento de entrada y dispositivos de procesamiento de salida.

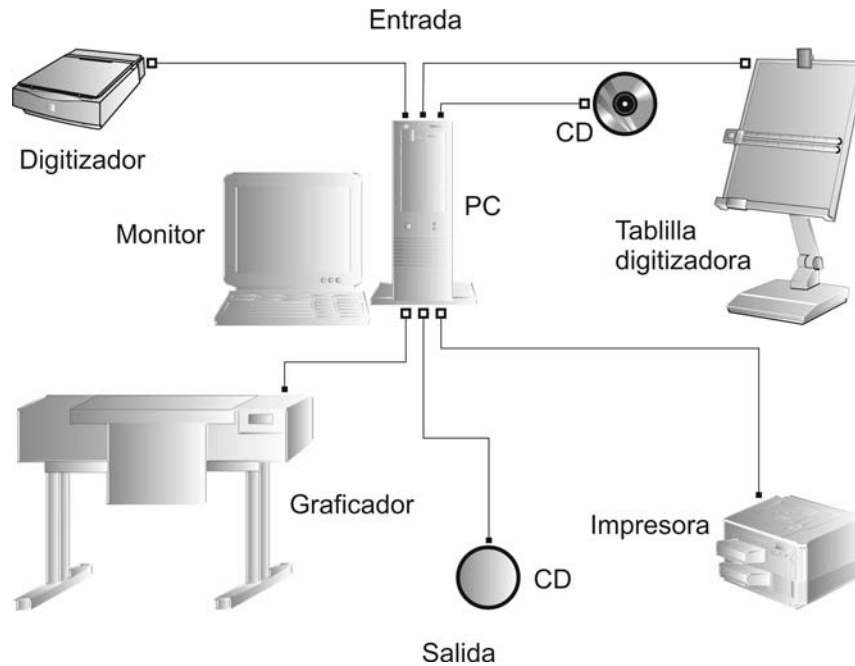


Ilustración 4 Dispositivos de entrada y salida en el campo de la percepción remota

Apoyo de campo: Son las medidas adicionales que se generan en el área de interés y complementan la información del sensor remoto, son las medidas realizadas directamente en el campo, es decir los datos complementarios de la escena, los cuales son llamados datos ancilares.

1.3 Radiación electromagnética

Energía Electromagnética se refiere a toda la energía que se mueve a la velocidad de la luz con un patrón ondulatorio armónico (Sabins, 2000).

Un patrón armónico consiste en ondas que ocurren en intervalos de tiempo iguales. El concepto de onda explica como la energía electromagnética se propaga pero esta energía solo puede ser detectada cuando interacciona con la materia.

Históricamente, las propiedades de la radiación electromagnética se han explicado por dos teorías aparentemente contrapuestas: la teoría ondulatoria que la concibe como un haz ondulatorio, es decir que se transmite de un lugar a otro siguiendo un modelo armónico y continuo, a la velocidad de la luz (c) y conteniendo dos campos ortogonales entre sí: eléctrico y magnético, y la teoría corpuscular que la considera como una sucesión de unidades discretas

de energía, fotones o cuantos, con masa igual a cero, pero que tienen propiedades asociadas a las partículas.

En cierta forma ambas teorías se complementan, ya que el concepto de onda explica cómo la energía electromagnética se propaga, sin embargo, esta energía sólo puede ser detectada cuando interactúa con la materia. En esta interacción, la energía electromagnética se comporta en forma corpuscular. A medida que la luz se propaga a través de medios de diferentes densidades ópticas, se desvía (refracta), adquiriendo un comportamiento asociado con las ondas.

Cuando un sensor mide la intensidad de la luz, la interacción de los fotones con el material fotosensible del fotodetector, produce una señal eléctrica cuya magnitud varía en forma directamente proporcional con el número de fotones que sobre ella inciden.

1.3.1 Propiedades de las ondas electromagnéticas

Las ondas electromagnéticas pueden ser descritas en términos de su velocidad, longitud de onda y frecuencia. Todas las ondas electromagnéticas viajan a la misma velocidad (c). Esta velocidad es conocida comúnmente como la velocidad de la luz y depende del medio a través del cual la onda electromagnética se propaga. Si la onda electromagnética se propaga a través del vacío, $c = 299\,793\,000$ m/s.

La longitud de onda (λ) de las ondas electromagnéticas es la distancia de cualquier punto de un ciclo u onda a la misma posición sobre el siguiente ciclo u onda. El micrómetro (μm) es una unidad conveniente para designar longitudes de onda tanto de la radiación visible como de la radiación infrarroja. Sin embargo, en el ámbito científico es muy común emplear nanómetros (nm) para mediciones de longitudes de onda muy cortas, como la luz visible.

A diferencia de la velocidad y de la longitud de onda, las cuales cambian a medida que la energía electromagnética se propaga a través de medios de diferentes densidades, la frecuencia (ν) permanece constante y por tanto es una propiedad fundamental.

La frecuencia es el número de crestas de onda pasando a través de un punto dado en un período de tiempo. La frecuencia es expresada formalmente en “ciclos por segundo” hertz (Hz).

La longitud de onda (λ) y la frecuencia (ν) se encuentran relacionadas por medio de la velocidad de la luz (c) en la siguiente ecuación:

$$c = \lambda \nu \quad (\text{I.1})$$

1.4 El espectro electromagnético

El espectro electromagnético es una secuencia continua de energía en rangos desde metros hasta nanómetros en longitudes de onda, que viaja a la velocidad de la luz y se propaga a través de diversos medios incluyendo el vacío. Toda la materia irradia un rango de energía electromagnética de tal manera que la intensidad máxima presenta un corrimiento en el espectro hacia longitudes de onda pequeñas con un incremento de temperatura de acuerdo con la Ley de Planck.

El espectro electromagnético es la base principal de los métodos y sistemas de la Percepción Remota, ya que mide los niveles de variación de energía de la unidad fundamental de la fuerza electromagnética: el fotón.

La Ilustración 5 muestra el espectro electromagnético, el cual es dividido en regiones con base en diferentes longitudes de onda, como se describen en la tabla 1.

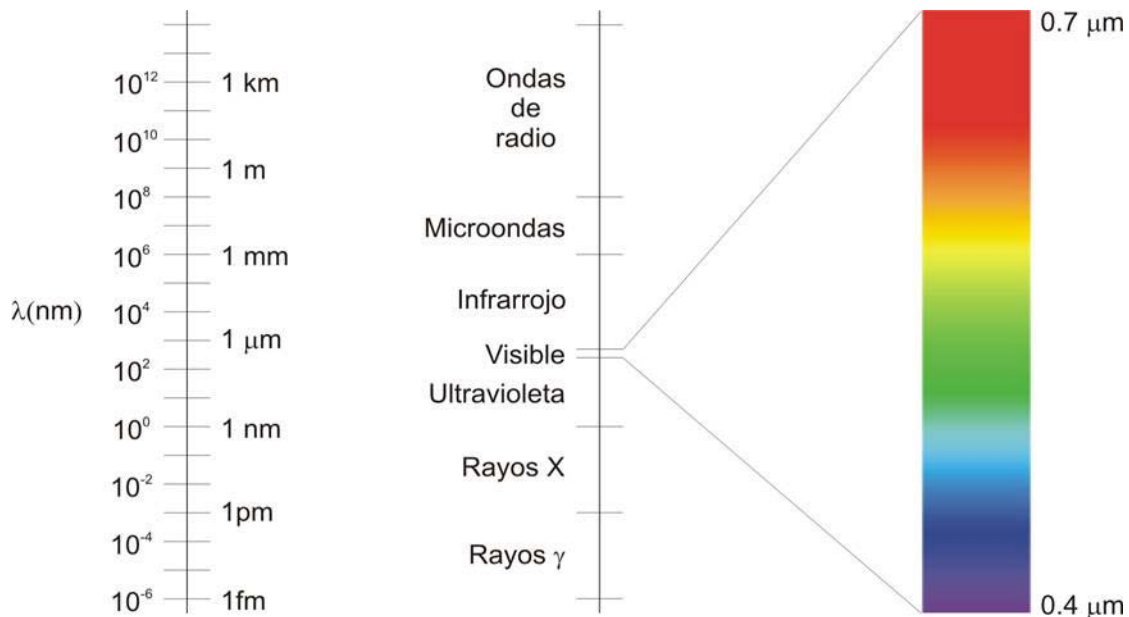


Ilustración 5 Espectro electromagnético
(Ilustración adaptada de Lira 2009)

Tabla 1. Regiones Espectrales Electromagnéticas

Región	Longitud de Onda	Comentarios
Rayos γ	< 0.03 nm	La radiación que llega es completamente absorbida en la atmósfera superior y no está disponible para la percepción remota.
Rayos X	0.03 a 30 nm	Completamente absorbida por la atmósfera. No empleada en el sensado remoto.
Ultravioleta	0.03 a 0.4 μm	Las longitudes de onda entrantes menores de 0.3 μm son completamente absorbidas en la atmósfera superior.
Visible	0.4 a 0.7 μm	Disponible para fotodetectores, pero la dispersión atmosférica es considerable
Infrarrojo	0.7 a 1000 μm	La interacción con la materia varía con las longitudes de onda.
Microondas	0.1 a 100 cm	Longitudes de onda más grandes que pueden penetrar en nubes, bancos de niebla y lluvia. Las imágenes pueden ser adquiridas en modos activos o pasivos.
Ondas de radio	> 100 cm	Porción del espectro electromagnético con mayor longitud de onda.

Desde el punto de vista de la percepción remota, conviene destacar una serie de regiones espectrales, que son las más comúnmente utilizadas por la tecnología actual. Su denominación y amplitud varían, según distintos autores, pero pueden resumirse de la siguiente forma (Barilla y Franco, 2001):

a) Espectro visible: se denomina así por tratarse de la única radiación electromagnética que pueden percibir nuestros ojos, coincidiendo con longitudes de onda en donde es máxima la radiación solar. Suelen distinguirse tres bandas elementales, que se denominan azul (492-455 nm), verde (577-492 nm) y rojo (770-622 nm), en razón de los colores primarios que nuestra psicovisión perciben en esas longitudes de onda.

b) Infrarrojo cercano (IR): a veces se denomina también Infrarrojo reflejado o fotográfico, puesto que parte de él puede detectarse a través de películas fotográficas dotadas de emulsiones especiales. Resulta de suma importancia por su capacidad para discriminar masas vegetales y concentraciones de humedad.

c) Infrarrojo medio: en donde se entremezclan los procesos de reflexión de la luz solar y de emisión de la superficie terrestre. Resulta idóneo para estimar el contenido de humedad en la vegetación y para la detección de focos de alta temperatura.

d) Infrarrojo lejano o térmico: incluye la porción emisiva del espectro terrestre, en donde se detecta el calor proveniente de la mayor de las cubiertas terrestres.

e) Microondas: son de gran interés por ser un tipo de energía de considerable transparencia a las nubes.

1.5 Obtención de la información

El objetivo de estudio y principio de análisis en percepción remota es 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 encuentran los objetos de interés. Esta superficie puede encontrarse en cualquier parte del paisaje. Sobre la escena se define un elemento con dimensiones físicas, denominado campo instantáneo de vista (CIV), cuyo tamaño depende de las dimensiones del paisaje, la resolución del sensor remoto y de la distancia de ésta al CIV (Ilustración 6).

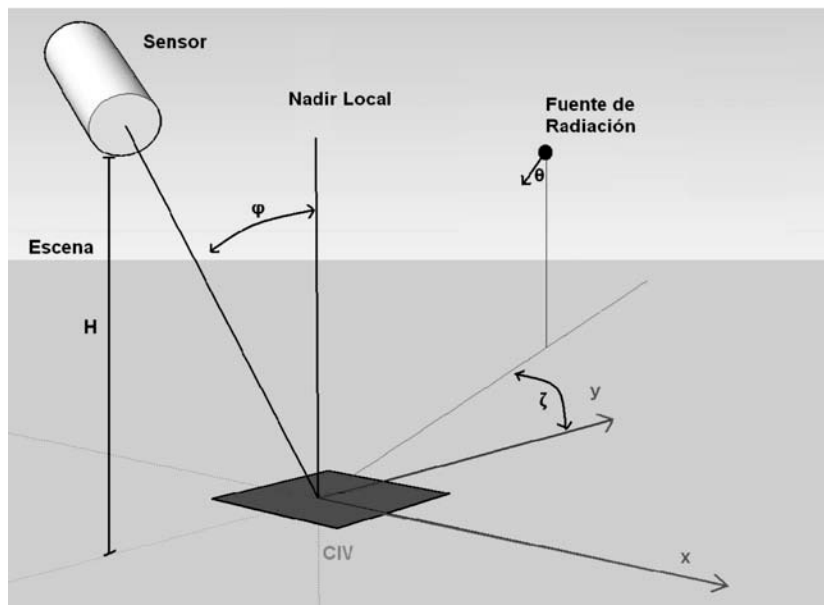


Ilustración 6 Geometría de observación en percepción remota
(Adaptado de Lira, 2002)

En el esquema anterior (Ilustración 6) observamos que la escena se encuentra iluminada por una fuente de radiación, cuyo ángulo de elevación es θ ; además la fuente de radiación se

encuentra a un cierto ángulo de azimut ζ con respecto a una dirección azimutal dada por el eje y . El sensor remoto se encuentra observando la escena a un ángulo de observación φ con respecto al nadir local.

La escena es dividida virtualmente en una rejilla de arreglo rectangular y homogéneo de campos instantáneos de vista. El sensor remoto mide la radiación proveniente de cada campo de vista por unida de tiempo.

Existen algunos conceptos que permiten entender el modo en el que se utiliza la radiación electromagnética en la Percepción remota para obtener información.

Para que pueda producirse el sensado remoto de un sistema es necesario que el sensor detecte un flujo energético proveniente de este. Por ello es importante definir las magnitudes utilizadas.

Energía Radiante (Q) es la cantidad de energía transportada por una onda electromagnética a través de un medio por unidad de tiempo, este tipo de energía puede incidir, atravesar o emerger de la superficie.

A la cantidad de energía radiante transportada por una onda electromagnética, por unidad de volumen se le denomina **Densidad de energía radiante** cuya magnitud proporciona información sobre la distribución de la radiación

$$(W = \partial Q / \partial V), \quad (I.2)$$

La tasa a la cual la energía radiante pasa a través de un cierto lugar se le llama **Flujo radiante**

$$(\Phi = \partial Q / \partial t). \quad (I.3)$$

Un sensor remoto integra el flujo radiante por unidad de tiempo, por unidad de ángulo sólido, proveniente de un elemento de la superficie (Ilustración 7). Es decir el flujo radiante representa la energía radiante emitida, transportada o recibida por unidad de tiempo.

Al flujo radiante por unidad de área se le llama **Densidad de flujo radiante**

$$(E = M = \partial \Phi / \partial A). \quad (I.4)$$

Cuando la densidad de flujo radiante incide sobre una superficie se le llama **Irradiancia** (E), en cambio si la densidad de flujo radiante emerge de la superficie, se le llama **Emitancia** o **Exitancia** (M).

El flujo radiante por unidad de ángulo sólido es una cantidad llamada **Intensidad radiante**

$$(I = \partial\Phi / \partial\Omega) \quad (\text{I.5})$$

La intensidad radiante es función de la dirección, desde o hacia el punto para la cual es definida. Un detector de percepción remota mide la intensidad radiante.

Al flujo radiante por unidad de ángulo sólido por unidad de área proyectada en una dirección dada, incidiendo, atravesando o emergiendo en una dirección específica de un punto de una superficie, es llamada **Radiancia**

$$(L = \frac{\partial^2\Phi}{\partial\Omega\partial A\cos\varphi}). \quad (\text{I.6})$$

Si en un punto considerado existe una fuente cuya radiancia no cambia como una función de la dirección de la emisión, la fuente es llamada Lambertiana.

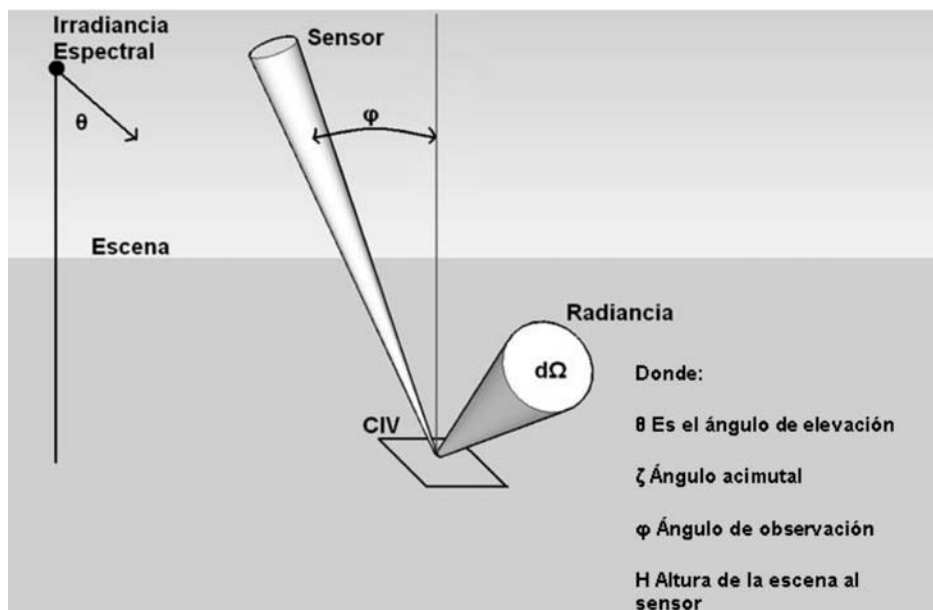


Ilustración 7 Componentes Radiométricos
 (Adaptado de Lira, 2002)

1.6 Interacción radiación - materia

La radiación electromagnética es generada por la transformación de energía cinética, química, nuclear o química. Una variedad de mecanismos de transformación conduce a la generación de ondas electromagnéticas sobre diferentes regiones del espectro electromagnético.

Las ondas en la región de radiofrecuencia son generalmente producidas por corrientes periódicas de cargas eléctricas en alambres, haces de electrones o superficies de antenas. En la región de microondas, la radiación electromagnética es generada por medio de tubos electrónicos que emplean el movimiento de electrones de alta velocidad en estructuras especialmente diseñadas para producir un campo electromagnético variable, el que es entonces conducido por guías de onda hacia una estructura radiante (antenas).

Las ondas en el infrarrojo y el visible son producidas por excitación molecular, vibracional o de cambio de nivel de energía. En términos generales la interacción radiación – materia depende de la longitud de onda de la radiación electromagnética y comprende cuatro mecanismos básicos: reflexión, absorción, transmisión y dispersión. Tales mecanismos dependen de la longitud de onda. Para cada región del espectro electromagnético existe un mecanismo de interacción que puede comprender procesos sub-atómicos, atómicos y moleculares.

1.7 Imágenes obtenidas en percepción remota

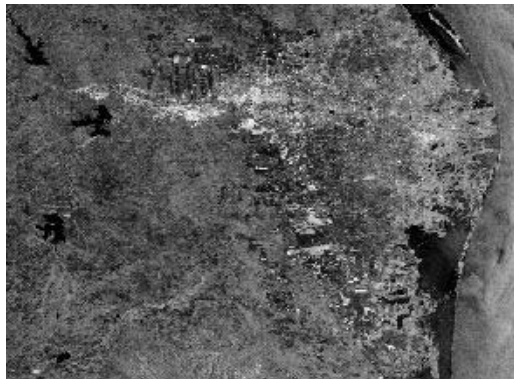
Existen dos clases de imágenes en la Percepción remota

Imágenes ópticas: Las bandas que las componen se encuentran en la región óptica del espectro electromagnético, generalmente son imágenes multispectrales, obtenidas por sistemas pasivos (Ilustración 8).



**Ilustración 8 Imagen óptica de la Ciudad de México
(Composición a color RGB de las bandas 4,5 y3)**

Imágenes de radar: Son imágenes producidas en la región de microondas del espectro electromagnético, generalmente son imágenes monoespectrales, obtenidas por sistemas activos (Ilustración 9).



**Ilustración 9 Imagen de radar ASAR del norte del golfo de México,
estado de Tamaulipas. Obtenida de: <http://earth.esa.int>**

Con todos los elementos anteriores se puede tener una mejor comprensión de lo que es la percepción remota y entender su importancia en el estudio del paisaje (Ilustración 10) ya que tiene múltiples aplicaciones en geología, exploración de otros planetas, urbanismo, ciencias del mar, silvicultura, hidrología, estudios de contaminación, agricultura, entre otros.

Capítulo 2

2. Imagen Digital



Ilustración 10 Representación visual de una imagen digital a través de una escala de niveles de gris

2.1 Percepción visual

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

Es por tanto importante entender la física de la formación de una imagen por sensores y sistemas ópticos incluyendo la percepción de la visión humana, ya que en las aplicaciones del procesamiento de imágenes, uno de los objetivos es mejorar la calidad visual de las mismas adaptándola lo mejor posible a la capacidad perceptiva del sistema de visión humana (Ilustración 11).



Ilustración 11 El ojo humano

Los humanos son criaturas principalmente visuales. No todos los animales dependen de sus ojos como nosotros lo hacemos para más del 99% de la información recibida acerca de nuestro entorno. Los murciélagos emplean sonidos de alta frecuencia, los gatos tienen una pobre visión pero su sentido del olfato está muy desarrollado, las serpientes localizan a sus presas por las emisiones de calor, los peces tienen órganos que detectan (y en algunos casos generan) campos eléctricos. Incluso los pájaros, quienes tienen un sentido de la vista altamente desarrollado, tienen una configuración visual diferente a la nuestra.

Con pocas excepciones por ejemplo, los búhos cuyos ojos están en lados opuestos de su cabeza proporcionando casi 360 grados de cobertura, pero pocos de manera estereoscópica, los búhos tienen de cuatro a cinco diferentes receptores de color; en cambio, nosotros los humanos tenemos tres: el Rojo, Verde y Azul.

2.1.1 Elementos de la percepción visual

Percepción de la luz.

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

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

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

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

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

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

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

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

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

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

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

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

2.1.2 Formación de la imagen en el ojo

Estructura del ojo humano

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

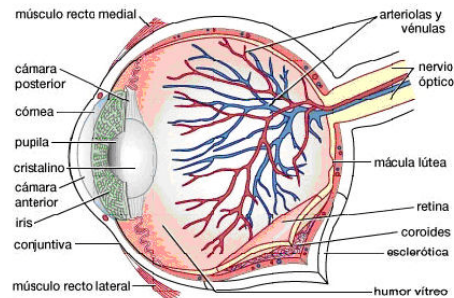


Ilustración 12 Corte horizontal del ojo humano

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

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

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

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

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

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

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

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

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

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

La visión es posible en gran medida debido a la retina. La retina ocupa un área de aproximadamente 1000 mm^2 , y está compuesta de aproximadamente 100 millones de sensores. Cada sensor convierte sólo una porción de la luz radiante de la imagen que incide en la retina en una señal eléctrica que es procesada por el cerebro.

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

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

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

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

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

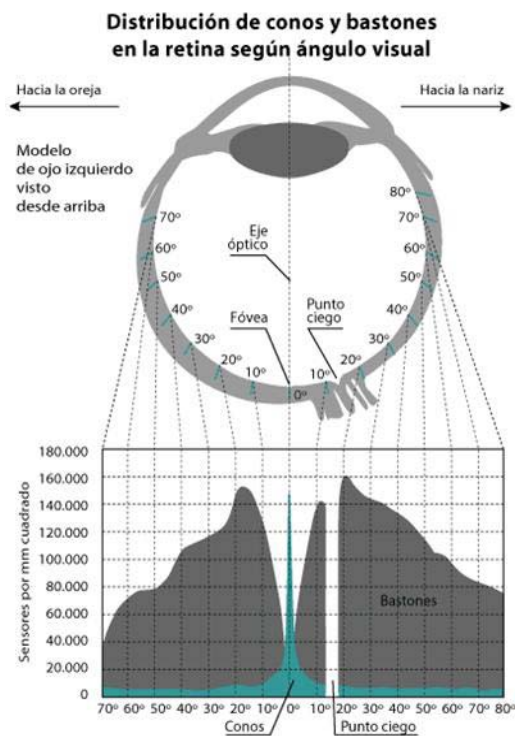


Ilustración 13 Densidad de conos y bastones en el ojo humano

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

2.2 Imagen

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

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

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

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

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

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

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

2.3 Tipos de imágenes

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

Imágenes binarias

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

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

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

Ejemplos:



Ilustración 14 Ejemplos de imágenes binarias

Imágenes en escala de grises

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

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



Ilustración 15 Ejemplos de imágenes en escala de grises

Imágenes de color

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

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

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

Ejemplos:

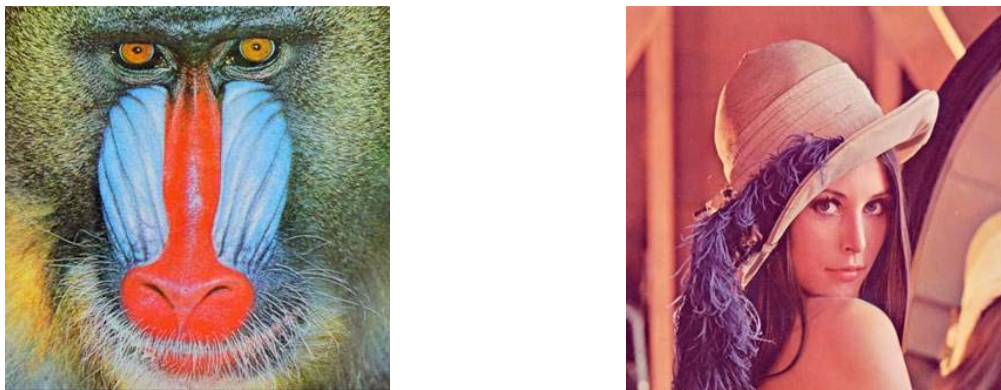


Ilustración 16 Ejemplos de imágenes en color

2.4 Imagen digital

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

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

$$0 < f(x, y) < \infty \quad (\text{II.3})$$

La función puede ser caracterizada por dos componentes:

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

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

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

2.5 Muestreo y cuantización

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

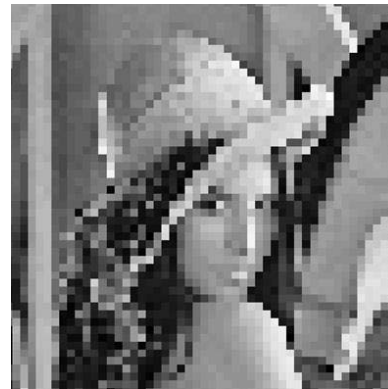
Una imagen puede ser continua respecto al eje de coordenadas x , y , pero también puede ser continua en amplitud. Para convertir una imagen continua a su forma digital, como se vio anteriormente, a la digitalización de los valores de las coordenadas se le llama muestreo mientras que el proceso de digitalizar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

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

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



(a)



(b)

Ilustración 17 Muestreo y cuantización de imágenes

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

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

2.6 Modelos de color

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

Existen diversos modelos de color, los más importantes son:

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

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

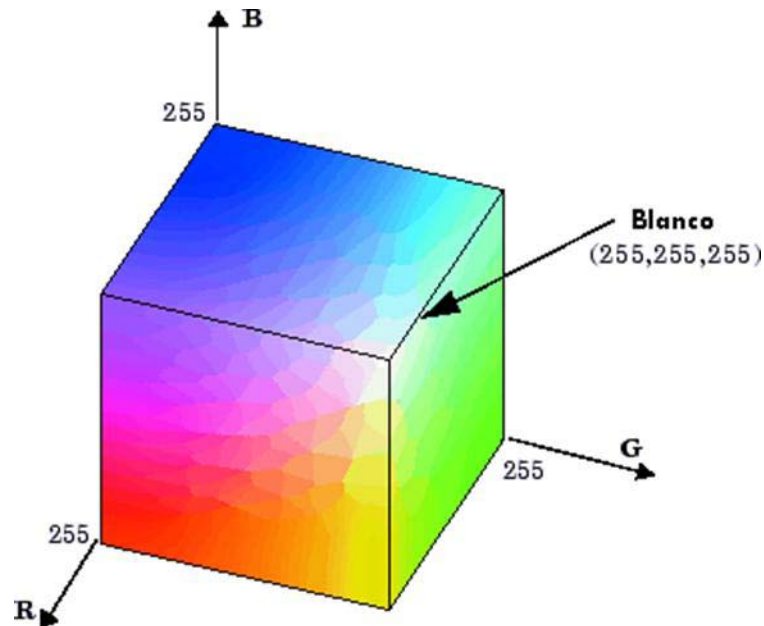


Ilustración 18 Modelo de color RGB

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

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

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

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

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

Normalizadas de [0,1]

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

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

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

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

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

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

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

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

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

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

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

$$\begin{aligned} &\text{si } B \leq G \\ &\text{si } 360 - \theta \geq G \end{aligned}$$

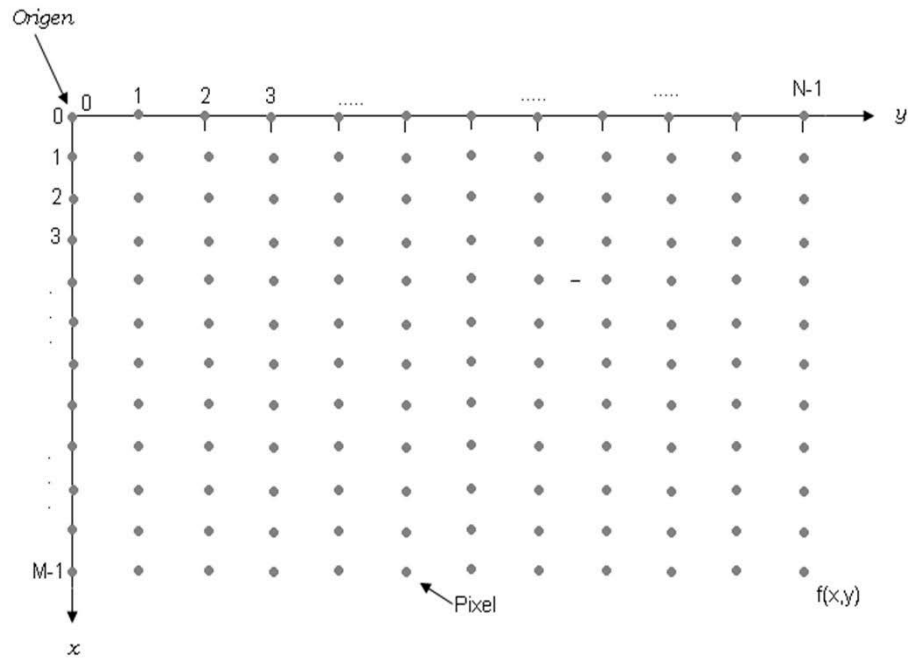
2.7 Representación de imágenes digitales

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

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

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

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



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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

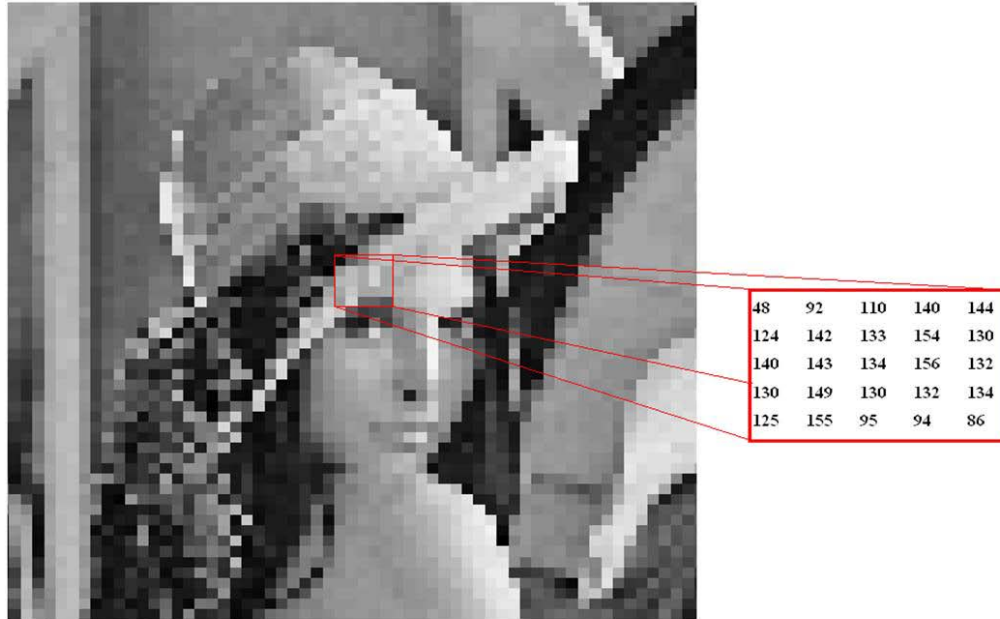


Ilustración 19 Representación visual de una imagen digital con los valores numéricos de los pixels correspondientes

2.8 Procesamiento digital de imágenes

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

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

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

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

El análisis de imágenes se descompone básicamente en dos grandes grupos de transformaciones: El procesamiento Digital y el Reconocimiento de Patrones.

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

El Procesamiento digital se divide a su vez en:

Realces

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

◆ Los métodos de realce en el dominio espacial

Estos métodos consisten en la manipulación directa de los pixels.

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

1. El *realce radiométrico* donde las operaciones son efectuadas directamente sobre un pixel sin importar o tomar en cuenta a los pixeles vecinos. Sirve para mejorar condiciones de bajo contraste, baja luminosidad o demasiada oscuridad. Ejemplo: ecualización del histograma.
2. El *realce con operaciones entre vecinos* que, a diferencia del realce radiométrico, las operaciones son efectuadas sobre un pixel pero tomando en cuenta a los pixeles que lo rodean. Sirve para eliminar ruido o para el mejoramiento de la nitidez. Ejemplo: kernel correspondiente a un filtro paso-bajas usando la convolución para realizar un filtrado espacial.

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

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



(a)



(b)

(a) Imagen sin realce (b) Imagen con realce por medio de la técnica normalización del histograma.

Ilustración 20 Realce de imágenes

a) Restauraciones

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



(a)



(b)

(a) Imagen con ruido (b) Imagen restaurada

Ilustración 21 Restauraciones de imágenes

b) Segmentación

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



Ilustración 22 Segmentación de imágenes

(a) Imagen Original (b) Imagen segmentada, la segmentación se realizó empleando el algoritmo isodata, el cual es un algoritmo iterativo. A partir de la primera estimación del umbral, se van recalculando los parámetros y nuevos umbrales sucesivamente.

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

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

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

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

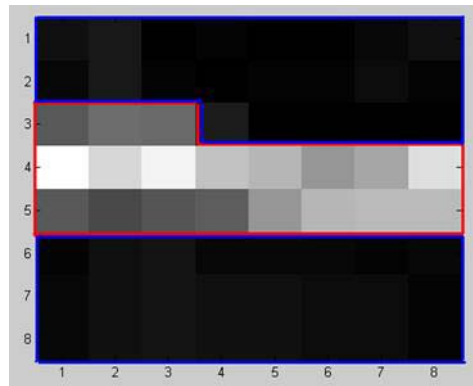


Ilustración 23 Región a segmentar delimitada por las líneas azul y roja

Podríamos delimitar dos Zonas:

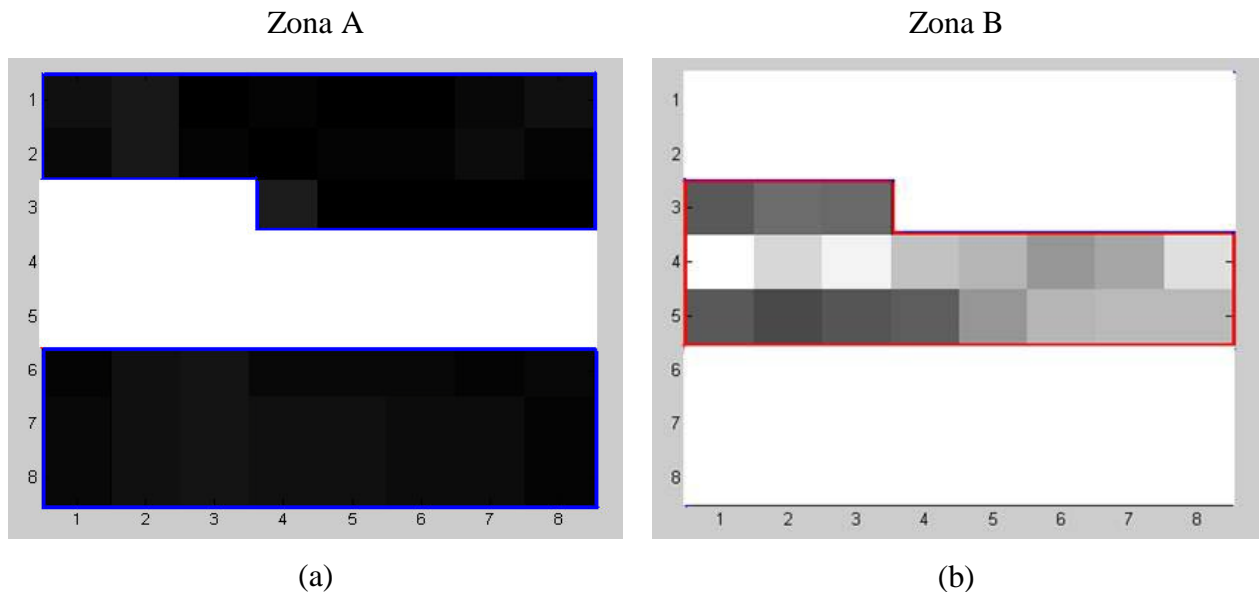


Ilustración 24 Segmentación de zonas

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

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

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

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

Ilustración 25 Segmentación espectral

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

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

Ilustración 26 Ejemplo de una segmentación de tipo contextual

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

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

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

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

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

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

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

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

Ilustración 27 Representación de una segmentación de tipo contextual

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

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

2.9 Imágenes digitales en percepción remota

2.9.1 Elementos de una imagen digital

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

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

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

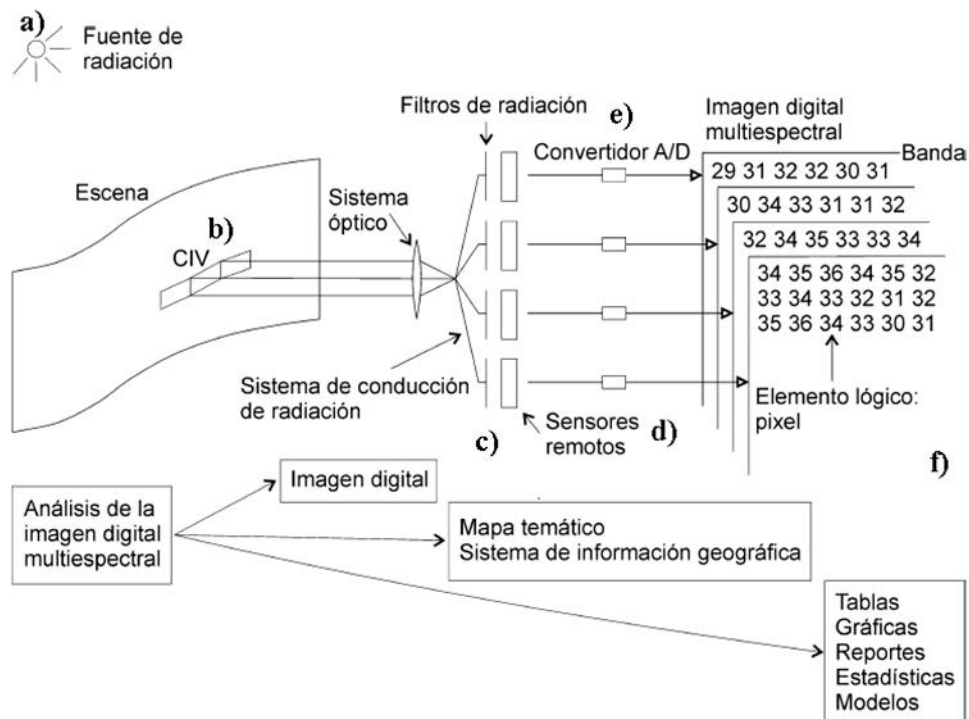


Ilustración 28 Esquema general que muestra la generación de una imagen digital multispectral a partir de una escena dada

(Adaptado de Lira, 2002)

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

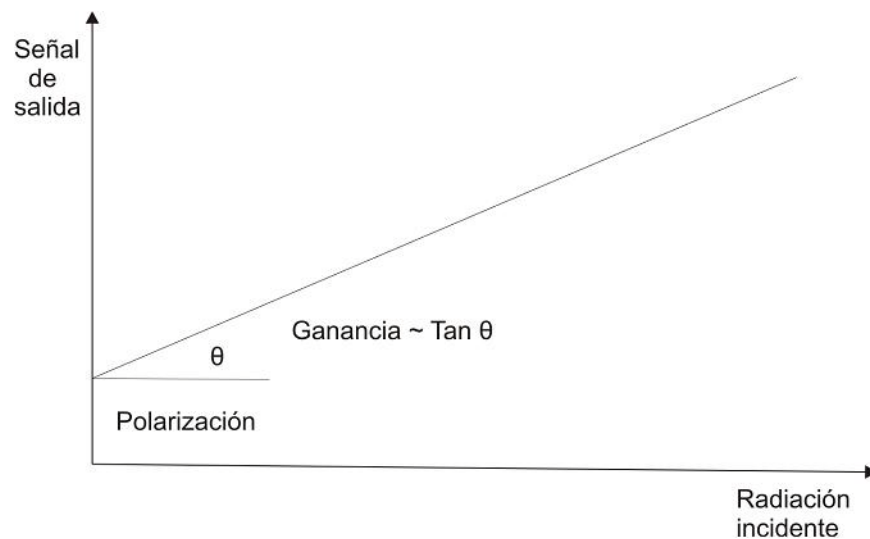


Ilustración 29 Respuesta lineal de un sensor remoto frente a una cierta radiación incidente
(Adaptada de Lira, 2002)

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

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

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

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

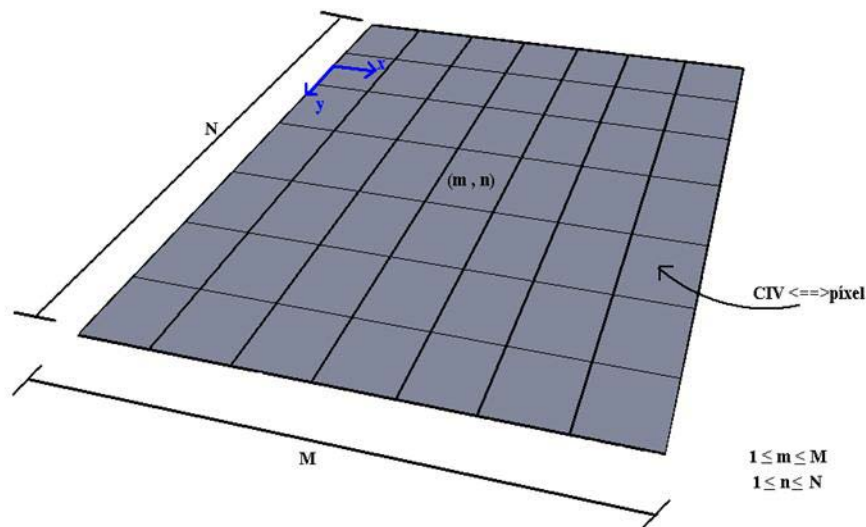


Ilustración 30 Elementos de una Imagen digital

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

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

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

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

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

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

2.9.2 Características de una imagen digital

Para generar una Imagen Digital se llevan a cabo procesos de discretización, estos pueden ser de dos tipos, espacial y radiométrica (Ilustración 31).

◆ Discretización espacial.- Se realiza una división de la escena en CIV empleando una rejilla virtual. Digitización o Digitalización: produce o define la medición del arreglo matricial de cada banda.

◆ Discretización radiométrica. Conversión a un número de la Intensidad de flujo radiante integrada por el sensor remoto. Cuantización: define el número de niveles o tonos de la imagen.

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

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

Capítulo 3

3. Imagen Digital Multiespectral

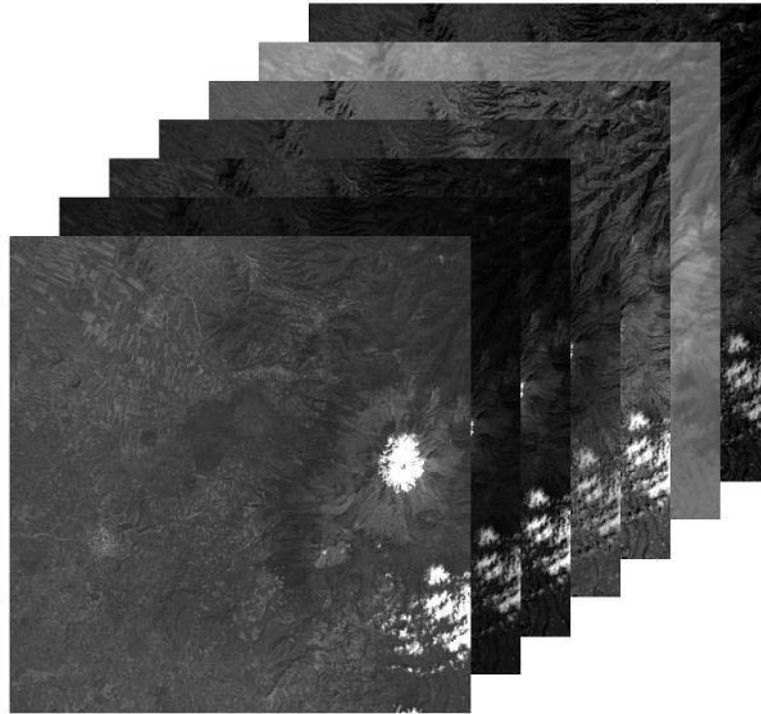


Ilustración 31 Imagen Digital Multiespectral

Una imagen digital multiespectral es un conjunto de bandas para una escena dada las cuales caracterizan la distribución espacial y espectral de la radiación proveniente de ella. Esta es una imagen multiespectral Landsat TM del estado de Puebla.

3.1 Imagen digital multiespectral

El objetivo de estudio y análisis en percepción remota es el paisaje (sistema físico tridimensional) iluminado por una fuente de iluminación.

Por medio del paisaje se encuentra la escena, sobre la cual se ubican o localizan los objetos de interés, sobre la escena se define un elemento con dimensiones físicas, también llamado campo instantáneo de vista (CIV) cuyo tamaño depende de las dimensiones del paisaje, la resolución del sensor remoto y de la distancia de este sensor al CIV (Ilustración 28).

La energía luminosa reflejada o emitida por el CIV incide sobre el sensor opto-electrónico y por un conjunto de dispositivos de conducción que hacen llegar esta energía hasta un conjunto de filtros (Ilustración 28).

Todo lo anterior se realiza con el fin de obtener la información proveniente de la escena simultáneamente a través de los filtros, así se registra una imagen digital por cada filtro.

Este proceso se puede entender si se toma una sucesión de fotografías de la misma escena, pero cada vez interponiendo un filtro diferente; el resultado es una serie de imágenes mostrando aspectos diferentes de la misma escena.

Así, cada objeto tiene una manera diferente de manifestarse cuando es vista bajo diferentes filtros. Se obtienen para una misma escena tantas representaciones numéricas como filtros se hayan empleado. Por lo anterior podemos denominar al conjunto de imágenes digitales como **“imagen multiespectral”** y al conjunto de números para cada CIV se les conoce como **“firma espectral”**

Una imagen digital multiespectral está formada por un conjunto de imágenes digitales (también llamadas bandas) para una escena dada y caracteriza la distribución espacial y espectral proveniente de aquella.

Una imagen multiespectral, está entonces formada por un conjunto de imágenes digitales, y cada una contiene diversos aspectos de la escena.

Una imagen multiespectral se obtiene a través de un mecanismo de captura de datos de la escena. Dependiendo del método de captura, los mecanismos se dividen en dos grupos: directos (de contacto) e indirectos (percepción remota)

Existen dos enfoques para la caracterización matemática de imágenes multiespectrales: una descripción determinista y una descripción estocástica.

En una descripción determinista se supone que se conocen las relaciones de valor que genera cada pixel de la imagen en relación con sus vecinos y a partir de éstas se define una función matemática donde se incluyen las relaciones puntuales y regionales de la imagen:

$$z = g(x,y,t) \quad (\text{III.1})$$

Donde t = diferentes bandas de la imagen, z = definida para toda $(x,y) \in D_j$. $g(x,y,t)$ define la imagen relativa a una escena dada

En una descripción estocástica no se conocen o no se emplean las relaciones de valor que den una función única y definida de los pixels de la imagen, la imagen se describe a través de propiedades promedio utilizando funciones probabilísticas como la de densidad y distribución. En una distribución estocástica el valor resultante en el momento de realizar las mediciones sobre los CIV de la escena es indeterminado. En una descripción estocástica z es una variable aleatoria (expresión III.1).

Si en el proceso de captura el valor de z está bien determinado, el proceso es determinista, pero si se encuentra regido por leyes de probabilidad entonces es un proceso estocástico. El proceso de captura es una operación realizada sobre la escena por lo que el sistema de captura es un operador que transforma la información del marco de referencia de la escena al de la imagen.

Por otro lado, una imagen multiespectral es la representación espacial de la distribución de energía proveniente de la escena para un conjunto de intervalos de energía, donde cada intervalo corresponde a una banda de la imagen multiespectral, por lo que la energía radiante depende de las coordenadas espaciales (x,y) , del tiempo t y de un parámetro ι .

Una función que representa una imagen multiespectral es la siguiente:

$$0 \leq g_i^{\min} \leq g_v(k\Delta x, l\Delta y, \iota, t) \leq g_i^{\max} \quad (\text{III.2})$$

Los valores asociados a g dependen de la escena utilizada para representarlos y son finitos. Los valores de las coordenadas (x,y) también están acotados y su intervalo de validez depende del tamaño espacial asociado a la imagen y del sistema de coordenadas empleado.

La función g puede ser vista como un campo escalar o como un campo vectorial. Si g está definida para un conjunto de valores ι , el campo es vectorial; si en cambio g está definida para un solo valor de ι , entonces el campo es escalar.

Barredores multiespectrales

Un *barredor multiespectral* es un dispositivo opto-electrónico formado por un conjunto de filtros y detectores, que se ocupa de la adquisición de imágenes simultáneas.

Todo ello conlleva un diseño avanzado de los sistemas de percepción remota ya que se necesita obtener la respuesta espectral y la resolución temporal y espacial de la escena.

El sensor remoto tiene que ser capaz de responder rápidamente a la energía luminosa proveniente del CIV, además de operar a niveles de ruido muy bajos y altos de eficiencia.

Existen algunos diseños fundamentales de barredores, como los siguientes:

- Los de diseños opto-mecánico. Ejemplos de este tipo de barredor son el Barredor Multiespectral MSS y el Mapeador temático TM
- Los de arreglo lineal. Ejemplo es el barredor que utiliza el satélite SPOT.
- Los de arreglo opto-mecánico lineal
- Los de arreglo matricial.

3.2 Despliegue de imágenes

Después de obtener los datos digitales del sensado remoto, es decir, que la escena ha sido discretizada, estos datos son almacenados en dispositivos ya sean magnéticos u ópticos para después ser convertidos en información útil.

Los datos digitales de un sensado remoto son usualmente guardados como una matriz de números (arreglos matriciales).

Ya que el conjunto de datos puede consistir en n bandas individuales, es posible identificar los valores de brillantez de un elemento particular de la imagen (píxel) en un conjunto de datos multiespectral especificando su fila (i) y su columna (j), y banda (k).

Para poder visualizar la imagen digital es necesario contar con tarjeta de video y un monitor de despliegue a color, estos dispositivos deben ser capaces de leer un punto en un tono de gris o de color de intensidad proporcional a su valor.

Primero se lee el valor de uno o varios píxels en una o varias bandas y se envían al monitor de color en las mismas coordenadas relativas que ocupan el archivo donde están almacenados. A cada píxel se le asigna un área en la pantalla y a cada valor un tono ya sea de color o de gris.

Cuando se trabaja con imágenes multiespectrales se debe seleccionar alguna de las bandas para ser desplegada o bien, si el *hardware* lo permite manejar los tres cañones del monitor (rojo, verde y azul) independientemente. Es posible desplegar tres bandas cualesquiera al mismo tiempo.

Esta representación visual de la imagen digital permite apreciar en tonos diferentes los patrones espaciales y espectrales si se utilizan los tres cañones del monitor

La representación visual de una imagen multiespectral, tiene las siguientes modalidades:

Tonos de gris

Se despliega una sola banda y consiste en usar una escala de gris para generar una imagen digital, generalmente se asigna un tono negro a un valor 0, un tono blanco a un valor 2^k-1 , (k = numero de bits utilizados para representar un píxel, generalmente $k = 8$) y tonos grises a los valores intermedios (Ilustración 32).

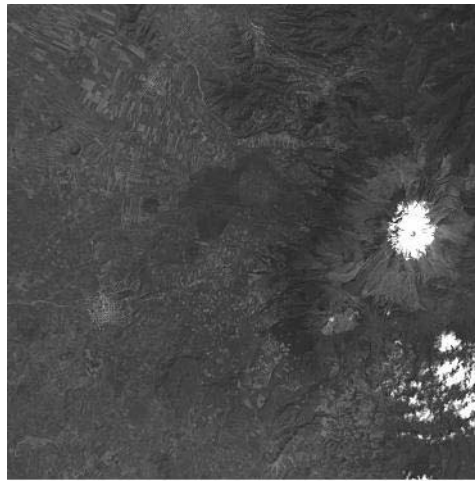


Ilustración 32 Imagen en tonos de gris

Pseudocoloración

Se despliega una sola banda y los valores de los píxels se subdividen en intervalos asignando a cada intervalo un conjunto de tonos de un cierto color. Por ejemplo si tenemos el intervalo $[0$ a $2^k-1]$, lo dividimos en segmentos, y a los colores de cada segmento le asignamos un tono de color de acuerdo con una cierta función (Ilustración 33)

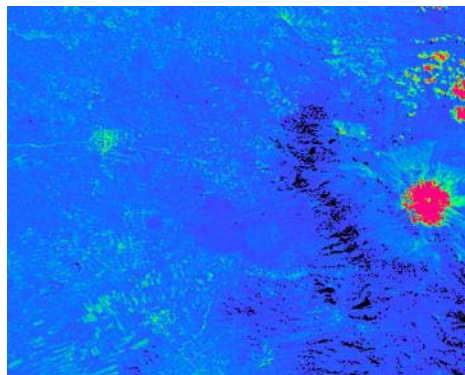


Ilustración 33 Pseudocoloración de una imagen

Falso color

Son seleccionadas 3 bandas cualesquiera de la imagen, y se asigna a cada una de estas bandas tonos de uno de los colores primarios. A esta reconstrucción también se le llama compuesto RGB (Ilustración 34).

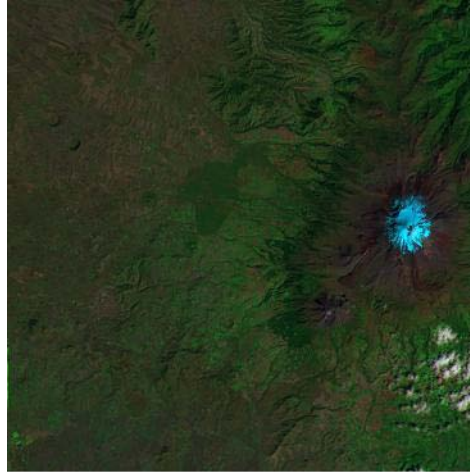


Ilustración 34 Composición RGB de la imagen multiespectral (Bandas 7,4 y2)

En cualquiera de las modalidades anteriores para la representación visual de una imagen multiespectral, la asignación de tonos ya sea en gris o color se hace por medio de una función, también conocida como función de color la cual no necesariamente es lineal y se representa por medio de una tabla.

3.3 Patrones y clase de patrones

Un patrón es un arreglo de entes de la misma naturaleza (Lira, 2002). A pesar de que no se especifica qué tipo de arreglo o qué tipo de entes lo constituyen, la única restricción es en cuanto a su naturaleza.

Existen dos posibles clases de arreglos, los de orden lógico y los de orden físico.

Un arreglo de orden físico se manifiesta siempre en una dimensión espacial por lo que es necesario emplear entes con un tamaño o área definidos.

Un ejemplo de arreglo lógico es el siguiente: ABC.

Cuando se tiene una imagen multiespectral cada uno de los pixels que forman la imagen es un patrón de orden lógico, y representan la respuesta espectral de los CIV de una escena dada. Un píxel es, entonces, un patrón espectral de orden lógico.

Si deseamos considerar un patrón de orden físico, tenemos que recordar la relación CIV \Leftrightarrow pixel para poder asignarle un tamaño al píxel.

Por si misma la imagen digital no tiene dimensión asociada para poder visualizarla por lo que hay que asociarle un tamaño físico a los pixels para que así se tenga una representación lógica y visual o física.

Cuando realizamos lo anterior es posible considerar arreglos de pixels con una estructura física que forman un patrón espacial.

Así entonces a cada píxel le asignamos un tamaño igual para todas las bandas y valores diferentes para cada una de ellas.

Por ejemplo, la siguiente es una Imagen multiespectral esquematizada de cuatro bandas, donde se ha asignado una cierta área cuadrada a los píxels:

55	55	57	58	58	55	50	59
47	52	55	56	56	52	48	59
48	50	54	56	57	53	50	63
53	51	53	57	56	53	49	63
56	52	52	55	52	51	48	57
54	49	49	50	48	50	51	53
48	45	45	45	46	52	53	59
50	51	53	53	53	56	58	59

Banda 1

50	51	51	51	51	50	50	51
50	50	50	50	50	49	49	53
50	50	50	51	51	49	50	54
50	51	51	52	50	50	49	51
51	50	51	51	49	49	48	49
50	49	49	49	48	49	48	49
48	47	47	47	48	49	51	52
48	47	47	47	49	50	52	52

Banda 2

56	56	55	56	56	57	57	51
53	56	54	56	56	54	55	54
55	56	55	56	59	55	57	58
58	56	56	56	57	55	56	54
55	57	55	56	54	55	54	49
53	55	53	56	53	56	55	48
52	52	53	52	53	57	57	57
45	45	46	46	49	54	54	60

Banda 3

30	30	31	33	32	38	44	58
68	46	31	28	31	37	46	55
72	58	35	29	31	35	49	53
52	54	37	31	31	37	50	53
41	49	42	35	39	49	54	59
39	50	53	44	53	57	48	60
47	65	66	65	64	41	43	54
56	68	65	65	60	40	48	47

Banda 4

Ilustración 35 Imagen multiespectral de cuatro bandas

La imagen multiespectral se encuentra representada en una dimensión espacial en donde se asocia a cada píxel una posición y un área dentro de cada banda (retícula). De esta manera cuando seleccionamos cualquier arreglo (físico) de pixels se forma un patrón espacial que puede estar compuesto por uno o hasta por todos los pixels que constituyen la imagen sin importar su valor o vecindad con otros. El patrón anterior tendrá una dimensión espectral (profundidad espectral) ya que se manifiesta a través de las cuatro bandas de la imagen

En el análisis de imágenes digitales existe un gran interés por estudiar el reconocimiento de patrones tanto espaciales como espectrales. Por lo que es importante revisar la definición de Clase de patrones.

Una clase de patrones es un conjunto de patrones con características similares, formada tanto de patrones espaciales como espectrales, con propiedades morfológicas o espectrales que son comunes.

Cuando se define una clase de patrones para identificar un conjunto de patrones espaciales o espectrales se reduce el universo de patrones a uno manejable lo cual hace que se mantenga una relación directa con la escena en estudio, al momento de definir las es importante elegir el modelo matemático adecuado para cuantificar las propiedades de similitud que se establecen en la definición de patrón.

Cuando se define un **patrón espacial** es muy importante definir las relaciones contextuales que guardan cada píxel con sus vecinos y el valor que tienen.

Cuando se define un **patrón espectral** es esencial la manera como corresponden espectralmente los CIV de la escena. Cada objeto de la escena, dentro de ciertos límites, tiene una firma o respuesta espectral que lo distingue de los demás, de esta manera al agrupar pixels con base en su respuesta espectral, nos permite detectar objetos de la escena y asignarlos a una clase, este proceso se denomina clasificación.

Podemos entonces considerar una clase espectral y una clase espacial, dependiendo de las propiedades similares que se hayan definido

En un patrón espacial importan las relaciones contextuales que guardan cada píxel con sus vecinos y el valor que tienen. Por ejemplo, un objeto puede contener pixels con diferente color pero estar agrupados en un patrón o clase espacial de acuerdo con las características morfológicas de dicho objeto.

En la clase espectral, las características similares están relacionadas únicamente con las propiedades espectrales de los pixels.

Los diferentes objetos que se encuentran presentes en una imagen están caracterizados por patrones espaciales o espectrales de tal manera que los objetos del mismo tipo forman una clase.

3.4 Ruido y artefactos

Cuando se captura o se aplica una transformación a una imagen es posible que se generen patrones ajenos a la escena respectiva, dichos patrones son el Ruido y el Artefacto.

a) Ruido

Es un patrón espacial, que no tiene relación con la escena cuyo comportamiento y propiedades estadísticas son obtenibles a partir de la imagen o del proceso de generación de esta última. El ruido es el resultado de un proceso estocástico con una función de probabilidad conocida y asociada al sistema que genera la imagen digital.

b) Artefacto

Es un patrón espacial o espectral que se introduce en la imagen digital y que no tiene relación alguna con los objetos del sistema físico que se desea estudiar.

Un artefacto se introduce en la imagen por algún mal funcionamiento del sistema sensor o por error o limitación de una transformación aplicada a la imagen. Un artefacto ocurre de manera aleatoria.

3.5 Relaciones de vecindad

La estructura de una imagen digital hace que las direcciones en ésta se encuentren discretizadas, únicamente ciertas de ellas son dadas de acuerdo con el arreglo matricial de los pixels.

Código de Freeman

El código de Freeman es una roseta de ocho direcciones principales, entendiendo como dirección principal a aquella que únicamente pasa por el centro de un conjunto de pixels

alineados y contenidos en la imagen (Ilustración 36), cualquier otra situación diferente determina una dirección secundaria no especificada en el código de Freeman.

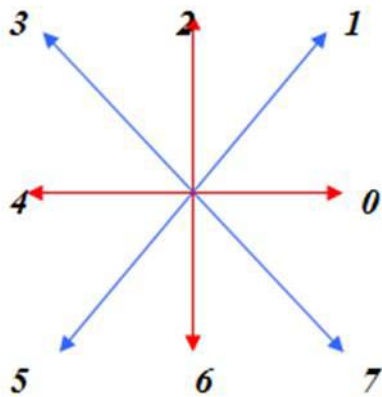


Ilustración 36 Código de Freeman

Conectividad

La conectividad es una relación de vecindad entre dos píxels. Es posible cuantificar la relación de vecindad de un grupo de píxels espacialmente arreglados marcando el contorno o borde de un objeto. Así el contorno definido por ciertos valores de píxels puede ser codificado empleando el código de Freeman.

Si se comenzara por cualquier extremo de una imagen, primero se anota la dirección en la cual se encuentra el siguiente píxel, escribiendo el número correspondiente de acuerdo con el código de Freeman; así al finalizar se obtendrá una sucesión de dígitos, la cual es conocida como *cadena de Freeman*. Hay que tener en cuenta que al momento de codificar un contorno es necesario adoptar convenciones como: ¿dónde se comienza? y ¿en qué sentido se hace el seguimiento?

Conectividad 4 y conectividad 8

Sea un píxel p con coordenadas (i,j) , sus **primeros próximos vecinos son aquellos con coordenadas $(i+1, j)$, $(i, j-1)$, $(i-1, j)$, $(i, j+1)$** , es decir aquellos que se encuentran en las direcciones horizontales y verticales especificadas por las direcciones 0, 2, 4 y 6 del código de Freeman y se denotan por $N_4(p)$ (Ilustración 37), es decir, $N_4(p) = \{(i+1,j), (i,j-1), (i-1,j), (i,j+1)\}$.

Los segundos próximos vecinos de p son aquellos con coordenadas $(i+1, j-1)$, $(i-1, j-1)$, $(i-1, j+1)$, $(i+1, j+1)$, es decir, aquellos que se encuentran localizados en las diagonales especificadas por las direcciones 1, 3, 5 y 7 del código de Freeman. A este grupo de pixels se les denota como $N_D(p)$, es decir, $N_D(p) = \{ (i+1, j-1), (i-1, j-1), (i-1, j+1), (i+1, j+1) \}$

$(i-1, j-1)$	$(i, j-1)$	$(i+1, j-1)$
$(i-1, j)$	p (i, j)	$(i+1, j)$
$(i-1, j+1)$	$(i, j+1)$	$(i+1, j+1)$

Ilustración 37 Tipos de conectividad

La unión de los grupos $N_D(p)$ y $N_4(p)$ forman los ocho primeros vecinos de p y son designados como $N_8(p)$, es decir

$$N_8(p) = N_D(p) \cup N_4(p) \quad (\text{III.3})$$

La conectividad constituye una relación de vecindad espacial entre un píxel y sus próximos vecinos. Sea N_g el conjunto de valores posibles (256 en muchos casos) que puede asumir un píxel en una banda dada, entonces:

Conectividad 4: Dos pixels p y q con valores N_g tienen conectividad 4 si q pertenece al grupo $N_4(p)$

Conectividad 8: Dos pixels p y q con valores N_g tienen conectividad 8 si q pertenece al grupo $N_8(p)$

Sistemas de vecindad

Es posible cuantificar la relación de vecindad de un grupo de pixels espacialmente arreglados que marcan el contorno o borde de un objeto.

Con base en los conceptos de vecindad y conectividad es posible introducir el concepto de Sistemas de Vecindad.

Sea $N = \{r_{11}, r_{12}, \dots, r_{MN}\} \in D$ el conjunto finito de nodos de una *lattice*, llamado también el conjunto de sitios de un dominio finito discreto D , donde éste dominio es un subconjunto finito del espacio bidimensional discreto \mathbb{Z}^2 : $D \subset \mathbb{Z}^2$. La definición de *lattice* es la siguiente: Una *lattice* es una partición del espacio continuo (\mathbb{R}^n) en células elementales, donde toda la partición es recubierta por células que no se interceptan. También se les conoce como “pixels (\mathbb{Z}^2)” y/o “voxels” (\mathbb{Z}^3).

En una imagen digital, los nodos r_{ij} son vectores de posición que apuntan a los centros de los pixels.

El conjunto de sitios define una malla finita y discreta de nodos (*lattice*) de acuerdo con un cierto arreglo que puede ser rectangular o hexagonal. Si se asocia a N una familia de sitios de vecindad denotada como

$V = \{v_{11}, v_{12}, \dots, v_{MN}\}$ tal que $v_{ij} \in N$, donde $r_{ij} \in v_{ij}$ y $r_{lm} \notin v_{st}$ sí y sólo si $r_{lm} \notin v_{st}$.

El conjunto V es llamado un sistema de vecindad para el conjunto de sitios N , y los puntos v_{ij} son vecinos de r_{ij} . La dupla de conjuntos $[N, V]$ denota un conjunto de N sitios denotados por un sistema de vecindad N . Los sitios correspondientes a una *lattice* N se relacionan entre sí a través de lo que conocemos como un sistema de vecindad V .

Un **sistema de vecindad de V sobre N** se define como

$$V = \{V_i \mid \forall i \in N\} \quad (\text{III.4})$$

Donde:

V_i es el conjunto de sitios vecinos al sitio i (píxel)

Sistema de vecindad homogéneo y orden de vecindad

Sea una *lattice* rectangular N , el conjunto vecino del píxel i (V_i), se define como el conjunto de los sitios más cercanos dentro de un círculo de radio r :

$$V_i = \{i \in N \mid [\text{dist}(\text{pixel}_i, \text{pixel}_j)]^2 \leq r, i \neq j\} \quad (\text{III.5})$$

Donde $\text{dist}(\text{pixel}_i, \text{pixel}_j)$ denota la distancia euclidiana entre los pixels i y j de tal forma que r toma un valor entero (Ilustración 38). Es importante notar que los sitios cercanos a los límites de la circunferencia tienen un número menor de vecinos

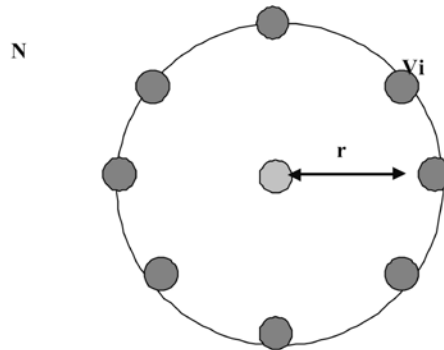
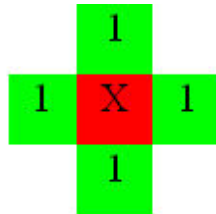


Ilustración 38 Sistema de vecindad circular

Vecindad de primer orden

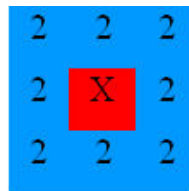
Llamado también sistema de 4-vecinos, cada sitio (interior) tiene cuatro vecinos.

X denota el sitio considerado y los marcados con 1 son sus vecinos.



Vecindad de segundo orden

Llamado también sistema de 8-vecinos, hay ocho vecinos por cada sitio (interior)



La forma de un conjunto de vecinos V_i puede ser descrita como la envolvente de todos los sitios del conjunto.

Los sitios más alejados del sitio central se muestran en la siguiente Ilustración en donde el número que aparece en cada casilla indica el orden del sistema de vecindad al que corresponde

5	4	3	4	5
4	2	1	2	4
3	1	X	1	3
4	2	1	2	4
5	4	3	4	5

3.6 Campos vectoriales en una imagen multiespectral

Los modelos matemáticos dan la oportunidad de probarlos todos minuciosamente, realizando desde modificaciones pequeñas hasta cambios radicales.

La simulación de un túnel de viento puede rastrear la velocidad del aire en cada punto del automóvil y alrededor de este. Una función que asigna un vector, por ejemplo, un vector de velocidad a cada punto del espacio se denomina **Campo Vectorial**.

Para determinar dónde ocurren remolinos y turbulencias en el flujo del fluido se deben calcular integrales de línea, por otro lado el rotacional y la divergencia permiten analizar las propiedades rotacionales y lineales del flujo de un fluido.

Campo vectorial.

Un campo vectorial en el plano es una función $V(x,y)$ que aplica puntos en \mathbb{R}^2 en el conjunto de vectores bidimensionales.

Se escribe:

$$V(x,y) = (v_1(x,y), v_2(x,y)) = v_1(x,y)\mathbf{i} + v_2(x,y)\mathbf{j} \quad (\text{III.6})$$

Para funciones escalares $v_1(x,y)$ y $v_2(x,y)$.

En \mathbb{R}^3 , un campo vectorial es una función $V(x,y,z)$ que aplica puntos en \mathbb{R}^3 en el conjunto de vectores tridimensionales. En este caso se escribe:

$$V(x,y,z) = (v_1(x,y,z), v_2(x,y,z), v_3(x,y,z)) = v_1(x,y,z)\mathbf{i} + v_2(x,y,z)\mathbf{j} + v_3(x,y,z)\mathbf{k} \quad (\text{III.7})$$

Para funciones escalares $v_1(x,y,z)$, $v_2(x,y,z)$ y $v_3(x,y,z)$

En general, si a una región del espacio en un sistema cartesiano de n dimensiones se le asocia a cada punto $(x_1, x_2, x_3, \dots, x_n)$ un vector dado por $u(x_1, x_2, x_3, \dots, x_n)$ se dice que se tiene un campo vectorial representado por $U = \{u(x_1, x_2, x_3, \dots, x_n)\}$.

Gráficamente, un campo vectorial se representa en una región del espacio por un grupo de vectores de magnitud y dirección de acuerdo con la expresión del campo (Ilustración 39).

Cualquier curva de esta región tal que el campo vectorial sea tangente a los puntos que la componen es llamada **línea de flujo de campo**.

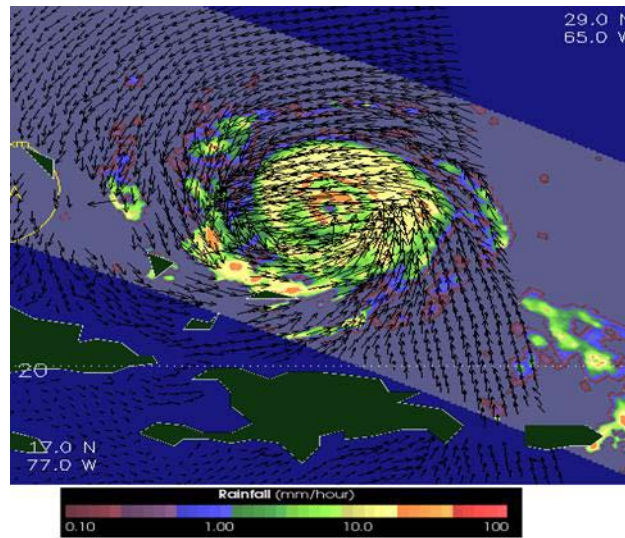


Ilustración 39 Ejemplo de campo vectorial

Imagen que muestra al Huracán Floyd en la cual se observa el campo vectorial del comportamiento de éste.

(Imagen obtenida de <http://visibleearth.nasa.gov>)

Campo vectorial de una imagen digital multiespectral

Si consideramos una *lattice* rectangular discreta $\mathcal{L} = \{1, \dots, N\} \times \{1, \dots, M\}$ que centrada en cada nodo de \mathcal{L} se localiza un CIV (subregión rectangular de una escena).

Cada CIV tiene asociado un vector n -dimensional $\{b_1, b_2, \dots, b_n\}$ que es generado por medio de un conjunto de sensores multiespectrales. Cada vector $\{b_1, b_2, \dots, b_n\}$ representa las propiedades espectrales promedio de un CIV de la escena y se le conoce como píxel de una imagen multiespectral: píxel vector = $\{b_1, b_2, \dots, b_n\}$.

Por tanto, el CIV es un elemento físico de la escena, mientras que el píxel es un valor entero o número digital (ND) en la imagen.

Sea $g = \{g_i\}$ una imagen multiespectral formada por un grupo de píxels de acuerdo con el conjunto $g_i = \{b_j(k,l)\}_i$, $\forall i \in \mathbb{N}$, $i = 1, 2, \dots, n$, donde i representa la colección de bandas de una imagen multiespectral y (k,l) es la pareja de valores enteros que representan las coordenadas espaciales de un píxel perteneciente a la imagen, es decir

$$(k,l) \in A \equiv \{(s,t) \mid (s,t) \in \mathbb{N}, 1 \leq s \leq M; 1 \leq t \leq N\} \quad (\text{III.8})$$

Los valores de los píxels en la imagen multiespectral pueden asumir valores enteros denotados por b_i :

$$b_i \in B \equiv \{p \mid p \in \mathbb{N}, 0 \leq p \leq 2^m - 1\}, m \in \mathbb{N}, 7 \leq m \leq 12 \quad (\text{III.9})$$

En resumen, una imagen digital multiespectral g es una función tal que $g: \mathcal{L} \rightarrow B$.

Entonces, si X_i es el conjunto $X_i \equiv \{x_i \mid x_i \in \mathbb{N}, 0 \leq x_i \leq 2^m - 1\}$, $\forall i \in \mathbb{N}$. El producto cartesiano $X^n = X_1 \times X_2 \times \dots \times X_n$ define el conjunto de n -tuplas ordenadas $(x_1, x_2, x_3, \dots, x_n)$, donde $n \in \mathbb{N}$ es el número máximo de bandas en la imagen multiespectral. Por otra parte, si se tiene que (b_1, b_2, \dots, b_n) es un píxel en el sistema cartesiano de coordenadas, para cada n -tupla (b_1, b_2, \dots, b_n) se tiene asociado un vector u :

$$\mathbf{u} = (x_1, x_2, \dots, x_n) \Leftarrow (b_1, b_2, \dots, b_n) \quad (\text{III.10})$$

Sea E un espacio Euclidiano n -dimensional discreto, $E \in \mathbb{N}^n$, la dimensión del vector u está dada por la métrica euclidiana:

$$\|\mathbf{u}\| = \sum_{i=1}^n x_i^2 \quad \frac{1}{2} \quad (\text{III.11})$$

El conjunto de vectores $\{(x_1, x_2, \dots, x_n)\}$ es el resultado de la proyección de la imagen multiespectral sobre un campo vectorial (Ilustración 40).

Si se observa el conjunto de n-tuplas (x_1, x_2, \dots, x_n) del dominio $D(X^n) \square E$, sólo una parte de ellas tienen un vector asociado, dependiendo de la información espectral contenida en la imagen. Sin embargo, es posible que para alguna n-tupla (x_1, x_2, \dots, x_n) se tenga más de un vector asociado, debido a la redundancia de información espectral.

En conclusión, el conjunto de vectores asociados con la imagen multiespectral forman el espacio vectorial $U = \{(x_1, x_2, \dots, x_n)\}$ en el dominio $D(X^n) \square E$,

Modelar la imagen como un campo vectorial trae como consecuencia la evaluación de texturas en la imagen (Ilustración 40).

Dadas dos bandas de una imagen multiespectral se puede obtener un campo vectorial de la misma como se muestra a continuación:

53	55	49	50	49	49	51	53
51	55	50	48	50	50	52	50
71	76	75	56	48	49	49	49
114	103	110	98	95	87	91	105
71	67	70	72	87	95	96	96
50	53	54	51	51	51	50	51
51	53	54	53	53	52	52	50
51	53	54	53	53	52	52	50

21	22	20	21	19	20	21	23
21	22	20	19	20	22	22	22
32	37	37	23	20	21	20	19
59	51	56	48	48	43	47	51
31	29	31	34	42	47	46	48
21	21	22	21	21	21	21	21
20	21	22	22	22	22	22	20
20	21	22	21	22	21	22	20

Valores lógicos de un subconjunto de pixeles de una imagen multiespectral, empleando las bandas 1 y 7

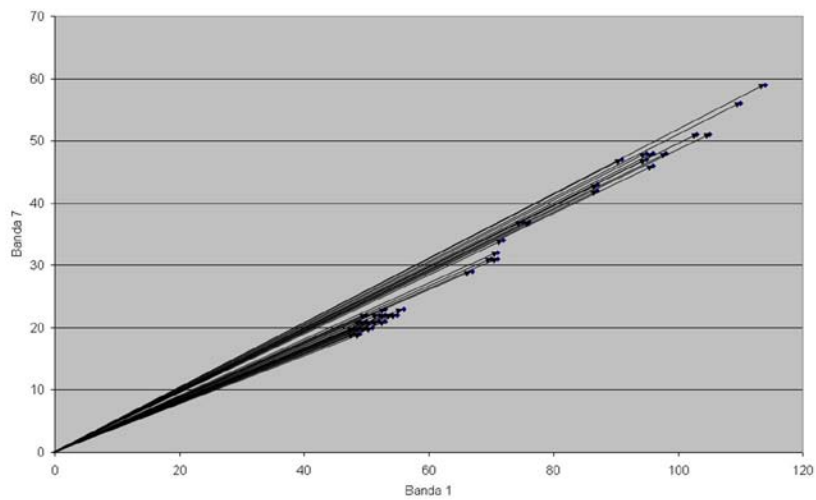


Ilustración 40 Campo vectorial formado por un subconjunto de valores de pixeles de una imagen multiespectral empleando las bandas 1 y 7

Capítulo 4

4. Textura Multiespectral

Banda 1

55	55	57	58	58	66	50	59
47	52	55	56	56	52	48	59
48	50	54	56	57	53	50	63
53	51	53	57	56	53	49	63
56	52	52	55	52	51	48	57
54	49	49	50	48	50	51	53
48	45	45	45	46	52	53	59
50	51	53	53	53	56	58	59

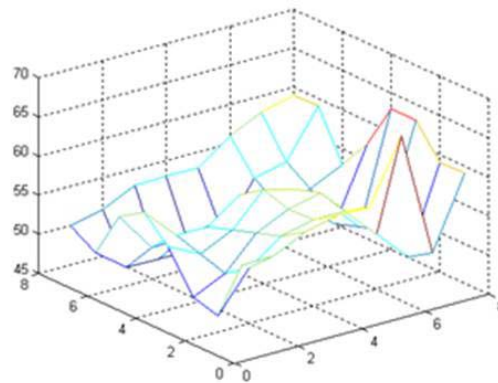


Ilustración 41 Textura de una escena

Superficie generada por los valores de los pixels de la Banda 1, la cual muestra la textura de la escena que representa dicha banda

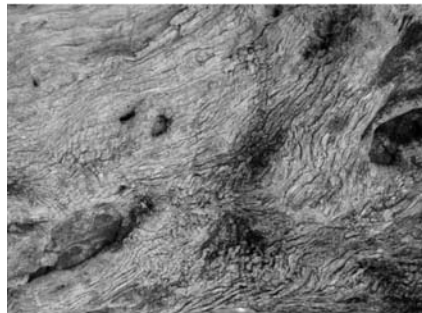
4.1 Textura

La textura es una propiedad asociada con las superficies, ya sea que tengamos contacto físico con ellas a través de nuestro sentido del tacto o las observemos en una imagen.

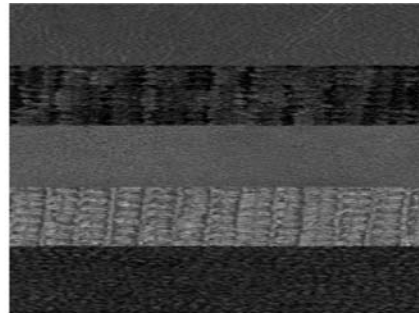
Estas propiedades son el suavizado, la rugosidad, la regularidad, la granularidad, etc.

Las texturas son frecuentemente definidas como clases de “patrones” visuales homogéneos y las podemos observar en cierto tipo de materiales tales como telas (Ilustración 42), por ejemplo las puntadas de algún suéter, las piedras que observamos en un camino, los granos de arena en la playa y la corteza de los árboles.

Las imágenes de la Ilustración 42 representan texturas que podemos observar en nuestra vida diaria.



Textura de la corteza de un árbol



Textura en diferentes telas



Textura de piedras



Textura de piedras de río

Ilustración 42 Diferentes ejemplos de textura

Observamos en las imágenes anteriores que existen patrones en las texturas, estos pueden ser periódicos como en el caso de las telas (enfoque sintáctico, por primitivas) o estadísticos como se muestra en el caso de las piedras (enfoque estadístico, global).

A pesar de que encontramos diferentes tipos de texturas en nuestro entorno, la visión humana es relativamente limitada para distinguir y clasificar texturas. Nuestra visión es capaz sólo de distinguir texturas hasta el tercer momento estadístico, más allá son confundidas por ser humano.

En el campo del procesamiento digital de imágenes la textura es un elemento importante para obtener las características de las mismas. A pesar de que actualmente no hay una

definición matemática universal aceptada, desde un punto de vista analítico podemos considerar la textura como una región elemental cuya estructura se atribuye a la repetición de patrones, con elementos (primitivas) ordenados.

4.2 Primitivas de textura

La noción de primitiva es clave en el tratamiento de texturas, un textel es una primitiva visual. Un textel es una primitiva visual que ocurre repetidamente en diferentes posiciones, deformaciones y orientaciones en una región dada, un textel cuenta con ciertas propiedades invariantes, por ejemplo sus pixels tienen un nivel de gris constante.

El aspecto más relevante de un textel es que las primitivas ocurren repetidamente dentro de un área dada. Consideremos una ventana de la imagen que corresponda a nuestro campo de vista superpuesto en un área grande, a medida que esta ventana se va haciendo más pequeña, es decir al acercarnos a la textura, aparecen menos textels contenidos en ella. Para alguna distancia, la imagen en la ventana ya no aparece texturada, o si persiste, el hecho de cambiar el tamaño de la ventana cambia la percepción de la textura drásticamente. En este momento desaparece la textura de la imagen. Un efecto similar ocurre si hacemos la ventana demasiado grande, correspondiente a un alejamiento del campo de vista de la imagen, en ese caso los detalles de textura desaparecerán y la imagen mostrará otra textura ya que ésta depende de la escala de observación.

Por lo anterior podemos concluir que si la resolución es apropiada, la textura se hace evidente a medida que el campo visual se mueve a lo largo del área texturada.

4.3 Características texturales

Se pueden definir seis medidas matemáticas las cuales se consideran necesarias para clasificar una textura (Tamura, 1978).

Estas son:

(a) Aspereza

Áspero versus suavidad (Coarseness-Coarse versus Fine).

La aspereza es la principal característica de una textura. Cuando dos patrones difieren solo en la escala, la que resalta es la aspereza.

(b) Contraste

Alto Contraste versus Bajo Contraste (Contrast-High Contrast versus Low Contrast).

El método más simple para cambiar el contraste de una imagen es elongado o comprimiendo la escala de grises. Cuando cambiamos el nivel de contraste de una imagen, alteramos su calidad visual, no su estructura. Cuando dos patrones difieren solo en la distribución de sus niveles de gris, la diferencia en su contraste puede ser medido.

Los factores que influyen la diferencia de contraste entre dos patrones de textura son:

- ◆ Rango dinámico de los niveles de gris.
- ◆ Polarización de la distribución de negro y blanco en histogramas de niveles de gris o radios de áreas negras y blancas.
- ◆ Filo de los bordes.
- ◆ Periodos de repetición de los patrones.

En un amplio sentido, en el contraste se encuentra la calidad de la imagen. Los Primeros dos factores citados arriba son usados por imágenes de televisión e impresoras fotográficas. El tercer factor puede ser visto como la comparación entre dos imágenes con la misma distribución en los niveles de gris. Una imagen con bordes afilados tiene mayor contraste. Cuando vemos dos patrones como tableros de ajedrez, que difieren solo en la escala, podemos percibir la diferencia de contraste causada por el cuarto factor.

(c) Direccionalidad

Direccionalidad contra no direccionalidad (Directionality-Directional versus Nondirectional).

Esta es una propiedad global sobre una región dada. La direccionalidad envuelve la forma y reglas de colocación.

Para Bajcsy (1973) la direccionalidad puede ser dividida en dos grupos : Monodireccional y Bidireccional; en cambio, para Tamura (1978) la direccionalidad se mide únicamente en grados.

Existe una ligera diferencia entre direccionalidad y la coherencia, esta última mide el tipo de correlación espacial donde existe una direccionalidad local.

Una herramienta de estudio para la orientación y la direccionalidad es el Tensor local de inercia. La direccionalidad es llamada también “asintropia” y una direccionalidad no dominante es característica de una textura isotrópica.

(d) Parecido a líneas versus parecido a manchas

(Line-Likeness-Line-Like versus Blob-Like)

Este concepto es concerniente solo con la forma de un elemento de textura. Esta característica de textura puede complementar las otras antes mencionadas, especialmente cuando en dos patrones no se puede distinguir la direccionalidad.

(e) Regularidad

Regular versus Irregular (Regularity Regular versus Irregular).

Esta es una propiedad para variaciones de una regla. Para elementos de textura, se puede suponer que la variación entre elementos, especialmente en el caso de texturas naturales, reduce el promedio de regularidad. Además de que estructuras finas tienden a ser percibidas como regulares.

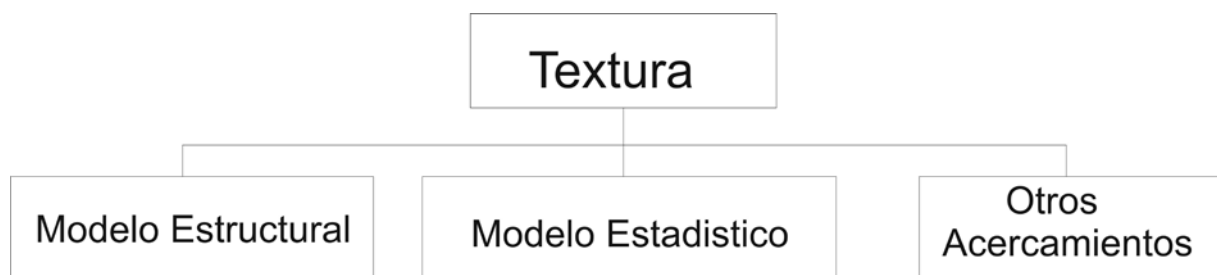
(f) Rugosidad

Rugoso versus suave (Roughness Rough versus Smooth).

Esta característica fue originalmente pensada para texturas táctiles, no para texturas visuales, sin embargo, cuando observamos texturas naturales como una tela de yute, somos capaces de compararla en términos de rugosidad y suavidad.

4.4 Clasificación de textura

Los modelos de textura pueden ser clasificados en tres amplias categorías:



a) Acercamiento Estructural

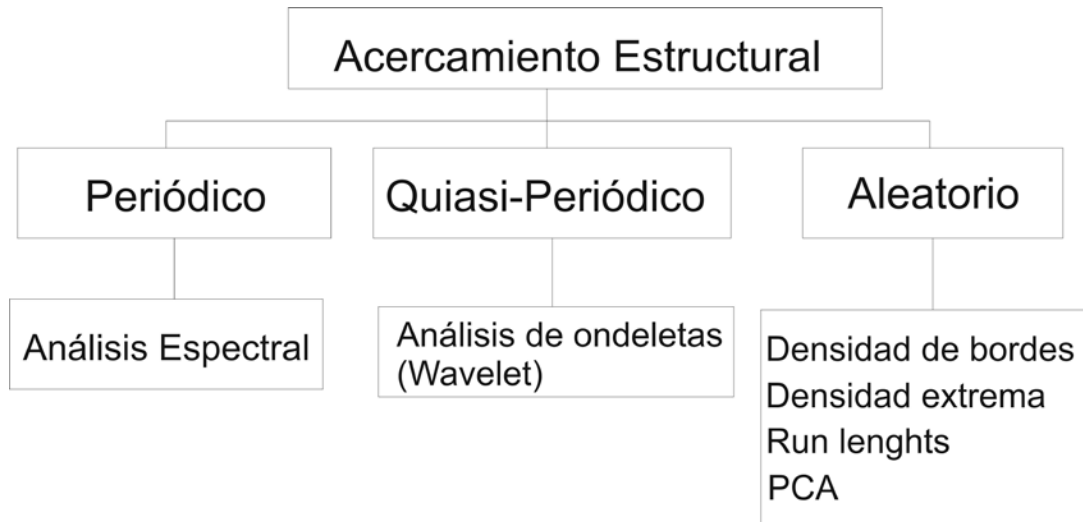
Cuando podemos identificar o existe uno o más primitivas de textura llamados textels para elementos de textura (o también llamados texons, textons, para átomos texturales) los cuales consisten de cualquier patrón el cual se repite de acuerdo a algunas reglas:

Textura = textel(s) + reglas de de posición (desde determinístico hasta aleatorio).

Éste último incluye: translación, rotación, escalamiento, superposición, combinación u otros métodos de combinación de textels.

El posicionamiento puede ser periódico, quiasi-periódico o cambiante; tan discretas sean las sub-imágenes, los textels pueden ser pensados como cualquier grupo pequeño (arreglo) de valores de píxel (binario, niveles de gris o valores de color).

Los métodos para analizar texturas desde el punto de vista estructural son:



b) Acercamiento Estadístico

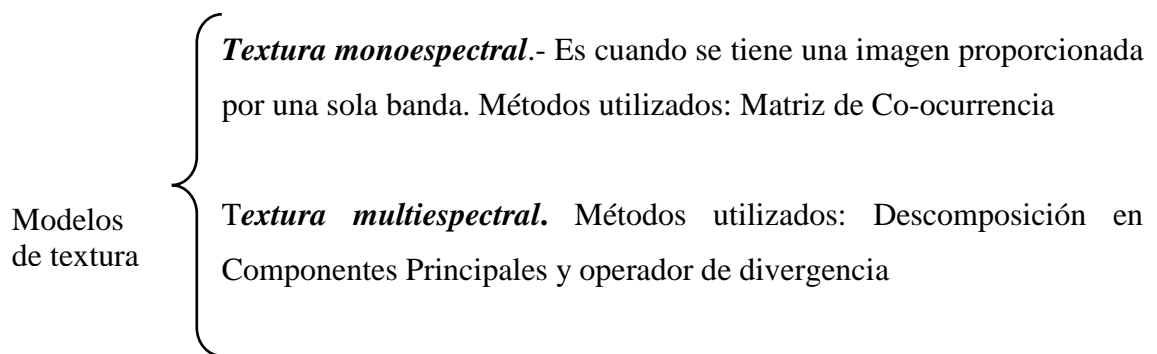
El problema del tratamiento de texturas se enfoca como un problema de reconocimiento de patrones, esto es debido a que muchas texturas no tienen una distribución geométrica regular (imágenes de escenas naturales), por el contrario presentan variaciones que no pueden ser descritas por un modelo satisfactorio mediante las formas, sin embargo, pueden describirse mejor a través de modelos estadísticos. Entre los análisis estadísticos de textura se encuentran:

- Histogramas
- Análisis de superficies
- Momentos
- Matriz de Co-Ocurrencia
- Descriptores de textura
- Descomposición en Componentes Principales
- Variogramas
- Función de Autocorrelación
- Operadores de Bordes
-

c) Otros Acercamientos:

- Aproximación híbrida
- Modelos de Mosaicos-Voroni
- Modelos autorregresivos
- Dimensión Fractal
- Estereológicos (sampling)

Se puede también dividir a la textura en dos tipos, monoespectral y multiespectral



4.5 Matriz de co-ocurrencia

Para cuantificar la textura se requiere de un modelo matemático que exprese la relación que existe entre los pixels vecinos.

Para ello tenemos que revisar la función de densidad conjunta, la cual expresa la probabilidad de que ocurran pares de pixels en direcciones pre-especificadas.

La función de densidad mide el contraste global de una imagen y la función de densidad conjunta establece la heterogeneidad espacial de los pixels en una vecindad dada.

Una aproximación a la función de densidad conjunta es la que se obtiene por medio de la matriz de co-ocurrencia.

La matriz de concurrencia proporciona la probabilidad de encontrar un par de valores de pixels en una dirección relativa determinada por un operador vectorial de posición. A la matriz de co-ocurrencia también se le llama Función de Densidad conjunta.

Formalismo teórico

Sea δ un conjunto convexo y compacto establecido en el espacio de la imagen como una ventana rectangular de tamaño impar.

Sea β un operador vectorial de posición que relaciona la localización espacial relativa de un par de pixels de la ventana.

La matriz c de co-ocurrencia es la probabilidad estimada de encontrar un par de valores de pixels (λ_p, λ_q) en la posición relativa dada por β ; los elementos de esta matriz están dados por:

$$c(\lambda_p, \lambda_q, \beta, \delta) = \frac{O\{r | r, r + \beta \in \delta, g(r) = \lambda_p, g(r + \beta) = \lambda_q\}}{O\{r | r + \beta \in \delta\}} \quad (\text{IV. 1})$$

Donde

$g(r)$ es la imagen

$\sigma(\delta)$ es una isometría de translación sobre la ventana

O es el orden del conjunto

$R = (i, j)$ vector de posición de un píxel de la imagen.

La dimensión de la matriz de co-ocurrencia es igual al intervalo permitido de valores de los pixels.

Debido a que la función de correlación decae para pixels vecinos, es suficiente tomar la magnitud de β como para considerar pixels próximos bajo conectividad 8.

Las posibles direcciones de β son las definidas por el código de Freeman.

La matriz de co-ocurrencia es una medida de la probabilidad de ocurrencia de un par de pixels en la posición relativa dada por β , por esta razón la matriz de co-ocurrencia, establece la estructura espacial de los pixels de la ventana.

Cuando en la matriz de co-ocurrencia observamos un solo elemento y este es distinto de cero, la textura de la ventana es uniforme. En cambio a medida que los elementos en la matriz de co-ocurrencia sean diferentes de cero, la textura se hace heterogénea.

La distribución y el número de elementos en la matriz de co-ocurrencia es un reflejo del tipo de textura que se encuentra en cada ventana.

Para caracterizar el contenido de información textural de la matriz de co-ocurrencia se emplean los llamados descriptores, los más usuales son:

Homogeneidad

$$H = \sum_i \sum_j c^2(i, j) \quad (\text{IV. 2})$$

Contraste

$$C_n = \sum_i \sum_j (i - j)^2 c(i, j) \quad (\text{IV. 4})$$

Entropía

$$E = \sum_i \sum_j c(i, j) \log[c(i, j)] \quad (\text{IV. 3})$$

Probabilidad máxima

$$P_m = \max_{i,j} [c(i, j)] \quad (\text{IV. 5})$$

La matriz de co-ocurrencia refleja la situación de textura de una imagen, a partir de la matriz de co-ocurrencia podemos cuantificar la textura en una región dada.

Hay ocasiones en que no basta con analizar solo un descriptor para cuantificar de manera adecuada el contenido textural de una imagen, en estas situaciones es conveniente construir un vector de descriptores que conjuntamente miden el contenido textural. A partir de un vector de descriptores se puede generar una imagen multiespectral sintética, donde cada banda representa el valor espacial de cada descriptor. En una imagen sintética puede existir redundancia entre bandas, pero ésta puede ser eliminada empleando la descomposición en componentes principales. Teniendo una imagen sintética sin redundancia es posible aplicarle análisis de cúmulos obteniendo así un mapa de texturas.

4.6 Textura multiespectral

Si se tiene un espacio de tres dimensiones y se grafica una banda cualquiera de una imagen multiespectral, colocando en el plano x - y la posición de cada píxel, y en el eje z su valor, obtenemos entonces una superficie, como la que se observa en la Ilustración 43.

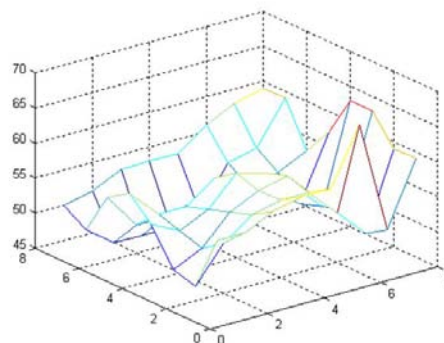


Ilustración 43 Superficie generada por los valores de los píxels de la Banda 1, la cual muestra la textura que representa dicha banda

La superficie resultante tiene las siguientes características espaciales: es suave donde los valores de los pixels cambian poco en su valor y rugosa donde cambian rápidamente.

Si los valores de los pixels cambian de manera rápida se tiene una superficie con sucesión de crestas y valles cuya amplitud y alternancia determinan la rugosidad del área. Por lo anterior, es posible asociar rugosidades con textura y así obtener una definición:

“Textura es la manifestación organizada de un arreglo local de tonos (pixels) de la imagen”. La textura es local cuando cambia de un lugar a otro de la imagen y depende de la escala de observación. Varias subtexturas pueden formar otra textura. También es diferente de banda a banda de una imagen multiespectral.

A pesar de que la textura es un concepto difícil de definir se pueden encontrar tres características básicas:

- 1) Un cierto orden local se repite sobre una región que es relativamente grande comparada con el tamaño del orden
- 2) El orden consiste de un cierto arreglo de partes elementales
- 3) Las partes están compuestas por entidades aproximadamente uniformes, teniendo más o menos las mismas dimensiones dentro de la región de textura, las partes pueden ocurrir en cualquier dirección

Algunos modelos de textura multiespectral son derivados de operadores que trabajan directamente en el dominio espacial de una imagen

4.6.1 Descomposición en componentes principales

La descomposición en componentes principales es una transformación al conjunto de datos que se aplica simultáneamente a todas las bandas que componen una imagen y como resultado se obtiene entonces una imagen multiespectral también.

La imagen de salida se encuentra referida a un espacio generado por vectores base cuyas direcciones son las de las componentes principales de la imagen original.

Esta transformación se conoce por algunos autores como Karhunen – Loeve (KL), Hotelling o transformación de eigenvector.

Esta transformación es unitaria y se considera como un realce global de la imagen. Tiene como ventaja que, no requiere de parámetros de entrada, sino que hace uso únicamente de las

propiedades estadísticas del conjunto de la imagen multiespectral. Además, cada banda de salida puede tener una interpretación física, ya que los eigenvectores y eigenvalores asociados la tienen en principio.

En la captura de una imagen satelital la variación del flujo del espectro radiante es medida por un sensor que depende de aspectos relativos a la iluminación solar y los efectos del albedo en la superficie.

El análisis de componentes principales es un medio poderoso de suprimir los efectos de la irradiancia que dominan todas las bandas así que los rasgos espectrales interesantes de la superficie de los materiales pueden ser examinados.

La descomposición en componentes principales es útil para reducir la dimensionalidad de los datos multiespectrales y la alta correlación banda a banda que es inherente en conjuntos semejantes de datos.

Una transformación de componentes principales supone la rotación y translación de datos hacia un nuevo conjunto de ejes ortogonales estadísticamente independientes los cuales son los principales componentes (eigenvectores de la matriz de covarianza) del conjunto de datos.

El origen del nuevo sistema de coordenadas es el significado del conjunto de datos originales.

El número de componentes principales es el mismo que el número de bandas espectrales usadas en la transformación.

Las características de la reflectancia espectral sobre la superficie de los materiales y el ruido aparecen en un alto número de componentes principales. Como la descomposición en componentes principales es un procedimiento estadístico, la información de cada componente principal es dependiente de la escena. Por ejemplo la primera componente principal es usualmente denominada por los efectos de iluminación surgidos de las diferencias de pendientes en el aspecto de relieve hacia el sol y provee una buena representación topográfica libre de ruido.

La descomposición en componentes principales ha sido usada extensamente para extraer información relacionada con rocas, mineralización y alteraciones de imágenes Landsat TM.

Para el desarrollo de esta tesis esta transformación a la imagen es de vital importancia ya que nos permite extraer información importante proveniente de la escena que estamos estudiando.

En la primera etapa al analizar la imagen multiespectral del área de estudio lo que nos interesa es obtener un realce global de la imagen para tener una apreciación general de esta y esto es lo que precisamente obtenemos con la descomposición en componentes principales.

Si comparamos dos bandas de manera individual de una subescena por ejemplo la Banda 1 y la Banda 2 mostradas en la Ilustración 44, notamos una fuerte similitud.

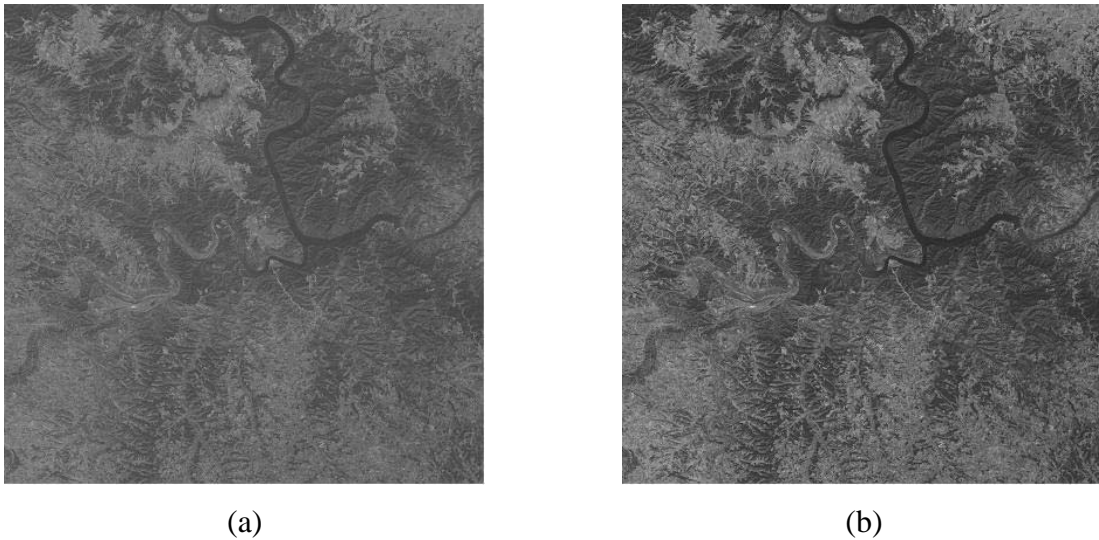


Ilustración 44 Primeras dos bandas de una imagen multiespectral Landsat TM

Tomando sólo un área de 5×5 pixels de la imagen multiespectral anterior, podemos observar los valores digitales para cada banda:

a) Valores digitales (Banda 1)	b) Valores digitales (Banda 2)
78 76 74 78 79	57 62 59 62 61
82 74 73 76 78	64 61 61 60 61
86 82 82 81 78	69 67 65 63 61
92 94 85 80 83	78 76 73 65 68
83 86 83 83 114	68 69 68 73 105

La relación de los valores digitales es mostrada en la Ilustración 45.

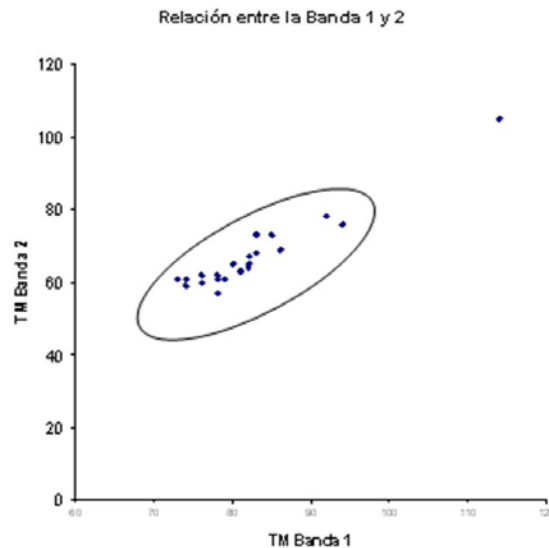


Ilustración 45 Representación de la descomposición en componentes principales

Gráfica de los valores digitales de la banda 1 (eje x_1) contra los valores digitales de la banda 2 (eje x_2) la cual muestra la correlación entre estas bandas. La descomposición en componentes principales es usada para generar un nuevo sistema de coordenadas (y_1, y_2).

En la gráfica se muestran los números digitales de la banda 1 los cuales son graficados contra la banda 2. Los puntos de los datos están distribuidos en una banda elongada, lo cual muestra que si los valores digitales incrementan para una banda, incrementaran también los valores para la otra banda.

Si por cualquier píxel que conocemos del valor digital para la banda 1, podemos predecir el valor aproximado de la otra banda, decimos entonces que los datos están fuertemente correlacionados. Esta correlación significa que hay mucha redundancia de información en un conjunto de datos multiespectral.

Si esta redundancia la reducimos, la cantidad de datos requeridos para describir una imagen multiespectral puede ser reducida.

La transformación de componentes principales es usada para comprimir conjuntos de datos multiespectrales calculando un nuevo sistemas de coordenadas.

Para las dos bandas de los datos de la Ilustración 44, la transformación define un nuevo eje (y_1) orientado en la larga dimensión de la distribución y un segundo eje (y_2) perpendicular a y_1 .

La operación matemática realiza una combinación lineal de los valores de píxel en el sistema original de coordenadas que da como resultado valores de píxel en un nuevo sistema de coordenadas:

$$y_1 = a_{11} x_1 + a_{12} x_2 \quad (\text{IV. 6})$$

$$y_2 = a_{21} x_1 + a_{22} x_2 \quad (\text{IV. 7})$$

Donde:

(x_1, x_2) = coordenadas de píxel en el sistema original

(y_1, y_2) = coordenadas en el nuevo sistema

$a_{11}, a_{12}, a_{21}, a_{22}$ = constantes

En la Ilustración 45 hay que notar que el intervalo de los valores de píxel para y_1 es más grande que los intervalos para cualquiera de las dos coordenadas originales, x_1 o x_2 , y el intervalo de los valores para y_2 es relativamente pequeño.

La misma transformación puede cumplirse para conjuntos de datos multiespectrales consistentes de cualquier número de bandas. Adicionalmente las direcciones en las coordenadas son definidas secuencialmente. Cada nueva coordenada es orientada perpendicularmente hacia todas las direcciones previamente definidas y en la dirección de la densidad máxima restante de los puntos definidos por los píxels.

Para cada píxel los nuevos valores digitales son determinados relativamente para cada uno de los nuevos ejes coordenados. Un grupo de valores digitales es determinado para cada píxel relativo hacia la primera componente principal.

Estos valores digitales son después usados para generar una imagen de la primera componente principal. El mismo procedimiento es usado para producir imágenes para el resto de las componentes principales.

Formalismo Teórico

Supongamos que la imagen de entrada es un campo aleatorio resultado de un proceso estocástico. Consideremos la clase de imágenes formada por g bandas donde cada banda $f_i(r)$ representa una imagen digital resultado de la realización de un proceso estocástico. Sea $f(r)$ la imagen multiespectral compuesta por las g bandas, entonces deseamos encontrar una transformación de tipo vectorial que se forme como sigue:

$$g(s) = A[f(r)] \quad (\text{IV. 8})$$

El sentido vectorial de la ecuación (IV. 8) consiste en que el operador A actúa sobre los valores del campo que forman los pixels de la imagen multiespectral de entrada para producir el campo vectorial g de la misma dimensión que el original.

La transformación (IV. 8) está sujeta a la restricción de que la matriz de covarianza K_g de la imagen g sea diagonal. En el espacio de salida la matriz de covarianza es, por definición

$$K_g = E\{[g(s) - \mu_g(s)][g(s) - \mu_g(s)]^t\} \quad (\text{IV. 9})$$

La razón de ser de esta restricción sobre K_g consiste en lo siguiente: En general existe una correlación entre las bandas de una imagen multiespectral, sobre todo cuando éstas son contiguas, ya que una banda es adquirida para un cierto intervalo de energías; es de esperarse, entonces que para intervalos cercanos o próximos la información correspondiente sea similar.

Es decir, entre bandas contiguas existe una redundancia de información que se traduce en que la matriz de covarianza de la imagen original tiene elementos diferentes de cero fuera de la diagonal.

Imponer una matriz de covarianza diagonal para la imagen de salida implica que se elimina la redundancia de información o correlación entre bandas, lo que conduce a una reducción en la dimensionalidad de la imagen multiespectral.

La matriz A de la relación (IV. 8) es una matriz unitaria, cuyos renglones están compuestos por los eigenvectores de la matriz de covarianza K_f de las bandas originales.

La matriz A realiza una diagonalización de la matriz K_f de tal manera que la matriz de covarianza de la imagen transformada

$$K_g = A K_f A^t = \Lambda \quad (\text{IV. 10})$$

Es una matriz diagonal cuyos elementos son los eigenvalores de K_f arreglados en forma decreciente. A la relación (IV. 10) se le conoce como transformación de similaridad y A es un operador hermitiano y real. La matriz Λ , es la matriz de eigenvalores de K_f .

La transformación KL produce un conjunto de bandas decorrelacionadas que forman una imagen con dimensionalidad menor a la original. Este proceso requiere del conocimiento de la matriz de covarianza definida por las bandas de la imagen multiespectral original. Dicha matriz debe ser modelada, estimada o medida, según el caso.

En forma explícita la transformada KL se escribe como:

$$g_i(k,l) = \sum_{j=1}^{\gamma} f_j(k,l) A(i,j) \quad \square i = 1, 2, \dots, \gamma \quad (\text{IV. 11})$$

Donde γ es el número de bandas de la imagen

El kernel A satisface la siguiente ecuación de eigenvalores y eigenvectores:

$$\lambda_i A(i,j) = \sum_{k=1}^{\gamma} K_f(i,k) A(i,k) \quad \square ij = 1, 2, \dots, \gamma \quad (\text{IV. 12})$$

donde λ_i representa los eigenvalores de la función de covarianza y A los eigenvectores de la misma. En otras palabras, la matriz de transformación A satisface la relación

$$A K_f = \Lambda A \quad (\text{IV. 13})$$

Es importante hacer notar que la matriz K_f es real simétrica y por tanto diagonalizable, con valores reales. Es decir que los eigenvalores λ_q deben cumplir con la ecuación secular

$$|K_f - \lambda_i I| = 0 \quad (\text{IV. 14})$$

Donde I es la matriz identidad. Los eigenvectores correspondientes a la ecuación (IV. 12) (5) satisfacen por tanto la relación:

$$K_f \hat{e}_i = \lambda_i \hat{e}_i \quad \forall i = 1, 2, \dots, \gamma \quad (\text{IV. 15})$$

Estos eigenvectores forman una base ortonormal: $\hat{e}_i \cdot \hat{e}_j = \delta_{ij}$

Cabe agregar que se cumplen las siguientes relaciones:

$$\Lambda_{ij} = \lambda_i \delta_{ij} \quad (\text{IV. 16})$$

Y además

$$A_{ij} = (e_i)_j \quad (\text{IV. 17})$$

Donde δ_{ij} es la delta de Kronecker. A los eigenvectores se les llama **componentes principales** y como los eigenvalores están arreglados de manera decreciente, entonces el contenido de información de la imagen es redistribuido entre estas componentes (o direcciones principales), de manera que la mayor información está concentrada en las primeras componentes o bandas de la imagen de salida, mientras que en las restantes se tiene muy poca información y prácticamente todo el ruido en las últimas.

Visualización de la descomposición en Componentes Principales

En forma gráfica podemos visualizar la descomposición en componentes principales. Supongamos que tenemos una imagen de tres bandas, por lo que en un espacio tridimensional se pueden representar el total de datos en forma de nube. Esto se logra al graficar el valor que tiene cada píxel en las tres bandas de este espacio tridimensional, el cual está construido por tres ejes normales entre sí y cuyas respectivas escalas son iguales

a los rangos dinámicos que pueden tener los pixels en cada banda. Así, un píxel cuyo valor en las tres bandas sea, por ejemplo (63, 65, 72) tiene justamente las coordenadas (63, 64, 72) del mencionado espacio; estas consideraciones son fácilmente generalizables a imágenes de γ bandas. El conjunto de datos graficados de esta manera forma lo que se conoce como una nube de datos.

Los datos así representados son referidos a un espacio de tantas dimensiones como bandas tenga la imagen. Si la correlación entre bandas es alta, podemos esperar que la nube de datos presente una elongación que denote dicha correlación.

Si los ejes originales (b_1, b_2, b_3) son rotados con la inclusión al origen de tal manera que el eje b_1 se coloque a lo largo de la máxima elongación de la nube de datos, el b_2 en la siguiente en importancia y b_3 en la siguiente, se tendrá un nuevo sistema de coordenadas (b_1', b_2', b_3').

En el eje b_1' queda definida la dirección de mayor elongación de los datos, en un sistema de coordenadas de 3 dimensiones solo se puede reorientar el eje b_2 porque b_3 queda reorientado automáticamente.

Las ecuaciones que describen esta multirrotación de ejes son las siguientes:

$$b_i' = \sum_{j=1}^{\gamma} a_{ij} b_j, \quad \square i \quad (\text{IV. 18})$$

Donde b_1' son las direcciones del nuevo espacio y representan los valores de los pixels de la imagen de salida para cada banda de ésta, b_j son los valores de entrada y a_{ij} son los coeficientes de la transformación. Esta última relación nos dice básicamente que la descomposición en componentes principales es una transformación reversible que se forma por medio de una combinación lineal de las bandas originales tomando el valor de cada píxel como entrada al proceso. Los coeficientes a_{ij} se determinan por el grado de rotación aplicado a los ejes del espacio original.

Por medio de esta transformación el intervalo dinámico de valores para el nuevo eje b_1' es mucho mayor que el original, el intervalo para b_2' es mayor que el de b_2 y el que se refiere a b_3' es menor que el b_3 . Así observamos que la mayor parte de información ha

sido puesta sobre el nuevo eje b_1' , una pequeña parte sobre el b_2' y muy poca información y ruido sobre el nuevo eje b_3' .

Estos nuevos ejes o direcciones (b_1' , b_2' y b_3') son conocidos como componentes principales, de ahí el nombre de esta transformación.

Si la correlación entre bandas es alta, la dimensionalidad de la imagen de salida puede reducirse a casi uno.

Para imágenes Landsat, generalmente de la primera a la tercera componentes principales acumulan casi el 100% de la información de la imagen multiespectral, por lo que se dice que Landsat tiene una dimensionalidad intrínseca de 3.

Para imágenes SPOT la dimensionalidad intrínseca es de 1.5, mientras que para imágenes ASTER la dimensionalidad intrínseca es de 3.5.

La descomposición en componentes principales emplea todos los pixels de la imagen original para generar un kernel de la transformación A a partir de la matriz de covarianza K_f , produce entonces un realce general, colocando la información relevante en las primeras bandas de la imagen resultante y separando el ruido contenido en la imagen original.

Cuando se requiere analizar la textura de una imagen multiespectral la descomposición en componentes principales es de gran ayuda ya que la imagen de salida nos muestra la información más relevante y sobre todo libre de ruido, por lo que en esta la textura resalta y la podemos ver a simple vista.

Para Ilustrar la descomposición en componentes principales he ocupado Imágenes Landsat del Valle de México. Estas imágenes son de tamaño 2048 x 1800 pixels, en niveles de gris.

La imagen multiespectral original se muestra en la Ilustración 46.

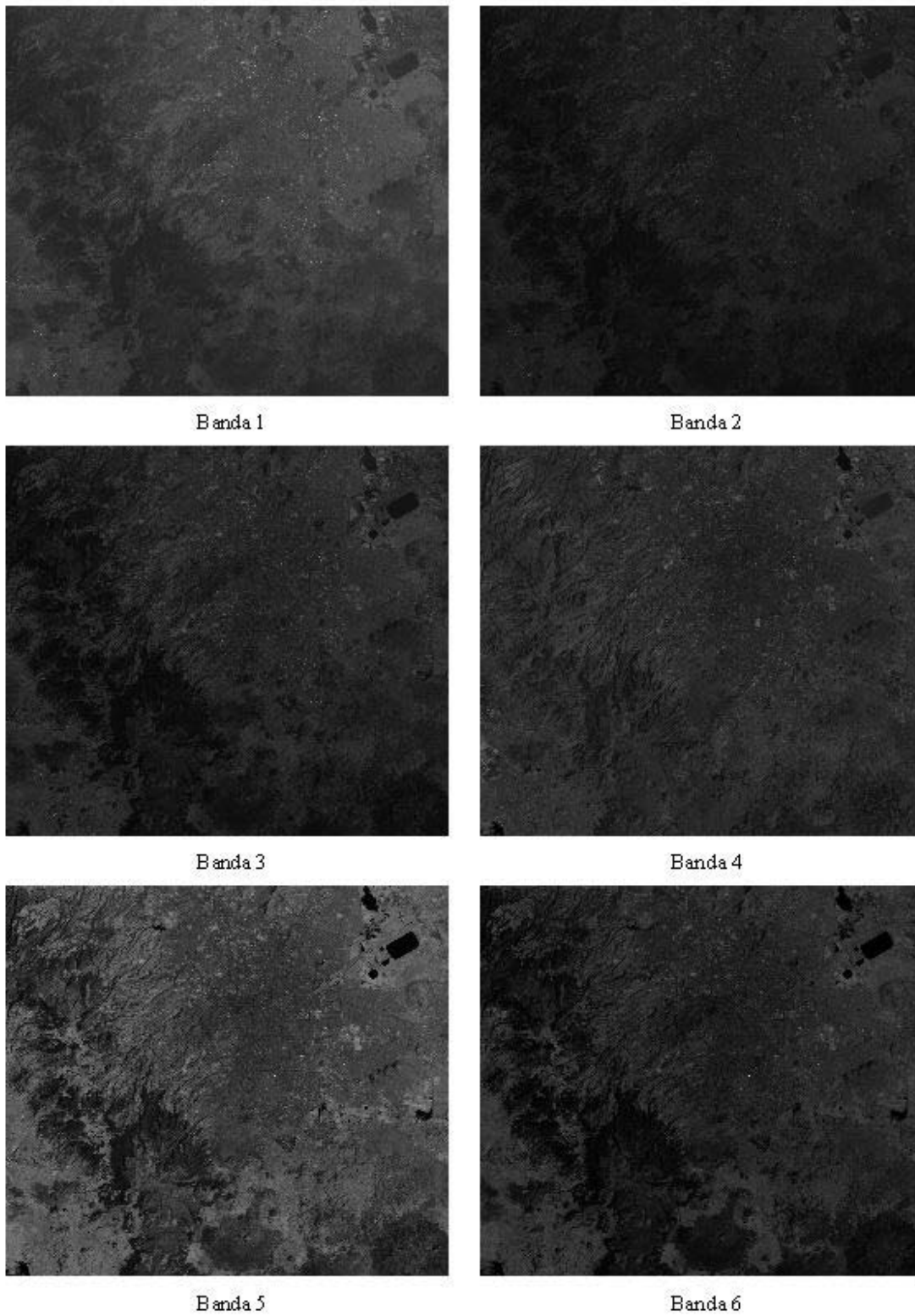


Ilustración 46 Imagen multiespectral del valle de México

Para aplicar el Algoritmo de Descomposición en Componentes Principales se ocupó el software Projectdcp de SANDI realizado en el Laboratorio de Percepción Remota del Instituto de Geofísica, dando el siguiente resultado:

TRANSFORMACION K-L.

Media de la imagen multiespectral:

```
media[ 1] = 59.29340
media[ 2] = 27.12489
media[ 3] = 32.97218
media[ 4] = 42.59798
media[ 5] = 64.52397
media[ 6] = 33.98060
```

Matriz de Covarianza:

222.40569	126.78917	179.92456	72.75105	205.70603	170.68203
126.78917	77.38230	110.50840	53.26047	141.85726	109.53056
179.92456	110.50840	167.00809	83.63993	235.04465	173.25191
72.75105	53.26047	83.63993	112.91132	185.56745	98.24393
205.70603	141.85726	235.04465	185.56745	559.61972	335.31037
170.68203	109.53056	173.25191	98.24393	335.31037	234.83278

Matriz de Correlación:

1.00000	0.96647	0.93357	0.45909	0.58308	0.74685
0.96647	1.00000	0.97209	0.56979	0.68169	0.81252
0.93357	0.97209	1.00000	0.60908	0.76884	0.87484
0.45909	0.56979	0.60908	1.00000	0.73822	0.60333
0.58308	0.68169	0.76884	0.73822	1.00000	0.92496
0.74685	0.81252	0.87484	0.60333	0.92496	1.00000

Eigenvalores:

1128.48637	173.10454	55.81678	9.25783	5.86148	1.63290
------------	-----------	----------	---------	---------	---------

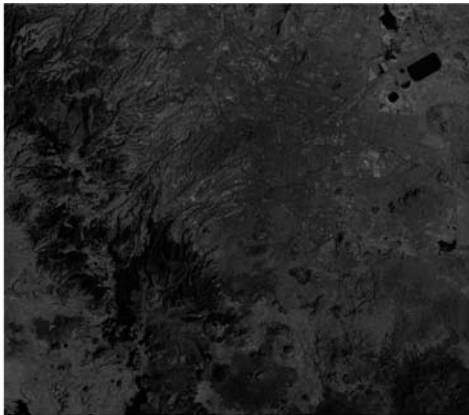
Kernel de la Transformación:

0.35596	0.22822	0.35581	0.23153	0.66816	0.44109
0.66527	0.30982	0.34090	0.20010	0.55152	0.03168
0.06346	0.09205	0.05864	0.89351	0.13619	0.40884
0.28532	0.00512	0.03429	0.29948	0.47893	0.77352
0.54628	0.18277	0.78348	0.12390	0.01575	0.19680
0.21645	0.90002	0.37247	0.05435	0.03483	0.01477

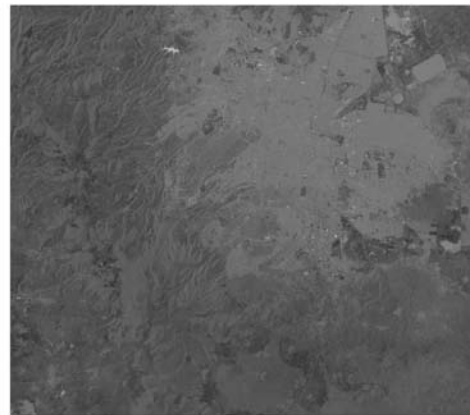
 Porcentaje de información significativa en cada banda de transformación:

Banda 1:	82.12191 %
Banda 2:	12.59712 %
Banda 3:	4.06188 %
Banda 4:	0.67371 %
Banda 5:	0.42655 %
Banda 6:	0.11883 %

:
 Observamos que las tres primeras bandas contienen el mayor porcentaje de información significativa (Ilustración 47).



Primera Componente Principal
 (82.12191% de información significativa)



Segunda Componente Principal
 (12.59712% de información significativa)



Tercera Componente Principal
 (4.06188% de información significativa)

Ilustración 47 Primeras tres componentes principales de la imagen multiespectral del valle de México

La descomposición en componentes principales es una técnica que hace uso de γ dimensiones (bandas), en la cual tenemos imágenes multiespectrales con datos fuertemente correlacionados entre banda y banda, los ejes en estas bandas no son ortogonales, es decir, las variables de los ejes no son independientes. La descomposición en componentes principales es una técnica que genera nuevas imágenes las cuales son ortogonales, es decir, independientes entre sí.

Esto es gracias a una transformación lineal de variables que corresponden a la rotación y translación del sistema de coordenadas original.

Una propiedad importante de la descomposición en componentes principales es que la separación entre cúmulos en una imagen puede ser incrementada. Cuando queremos segmentar una imagen podemos formar cúmulos con los pixels pertenecientes a ésta, los pixels que cumplen con ciertas propiedades pueden ser localizados en alguna nube de pixels o cúmulos. La separación entre estos es importante para saber en dónde poner a cada pixel y por lo tanto ayuda a una mejor segmentación de la imagen.

Se usa también el método de descomposición en componentes principales para remover el ruido de imágenes multiespectrales, ya que esta técnica se basa en la premisa que el ruido corresponde a una parte relativamente pequeña del total de la varianza.

De la aplicación del programa mencionado se puede observar que la primera componente principal contiene el 97.08467 % de la varianza.

En comparación, la varianza en la última componente parece ser insignificante, por lo que es de esperarse que en esta componente aparezca casi la totalidad del ruido

La descomposición en componentes principales tiene muchas ventajas, tales como:

- Una porción significativa de la variación en un conjunto de datos multiespectrales es comprimida en las primeras bandas.
- El ruido es generalmente relegado hacia las bandas menos correlacionadas.
- Las diferencias espectrales entre materiales pueden ser más aparentes en las componentes que en las bandas individuales.
- Se requiere de un espacio menor para almacenar o transmitir la imagen.
- El tiempo de cómputo para transformaciones a la imagen es menor.
- Existe una mayor calidad de resultados debido a la eliminación de ruido.
- La textura de la imagen original resalta de manera significativa.

4.6.2 Divergencia de una Imagen digital multiespectral

Divergencia

El rotacional y la divergencia son generalizaciones de la noción de derivada aplicados a los campos vectoriales (Ilustración 48). Ambas miden directamente cantidades físicas importantes relacionadas con el campo vectorial $F(x, y, z)$

La divergencia de un campo vectorial de tres dimensiones

$$F(x, y, z) = \langle F_1(x, y, z), F_2(x, y, z), F_3(x, y, z) \rangle \quad (\text{IV. 19})$$

Se define como sigue:

$$\text{div}F(x, y, z) = \frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} + \frac{\partial F_3}{\partial z} \quad (\text{IV. 20})$$

Definida en todos los puntos en que existen todas las derivadas parciales indicadas. La divergencia de un campo vectorial puede escribirse empleando la noción de producto punto:

$$\square \quad F = \left\langle \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right\rangle \cdot \langle F_1, F_2, F_3 \rangle = \frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} + \frac{\partial F_3}{\partial z} \quad (\text{IV. 21})$$

Por tanto la divergencia de un campo vectorial es una función escalar.

Es importante hacer notar que la divergencia de $F(x, y, z)$ no es afectada por los términos de la componente i de F que no contengan x ni los términos de la componente j de F que no contengan y ni los términos de la componente k de F que no contengan z , siendo i, j, k los ejes coordenados en el sistema cartesiano en el espacio de tres dimensiones.

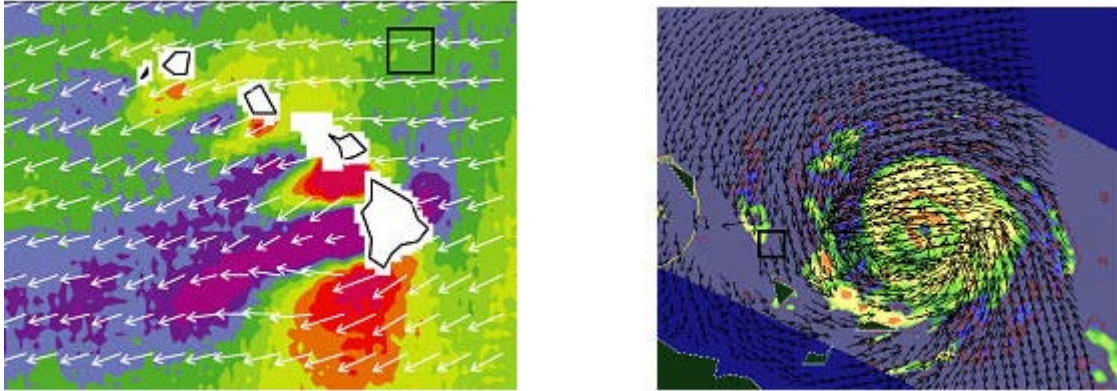


Ilustración 48 Campos vectoriales

Si observamos las cajas superpuestas en la gráfica de cada campo vectorial y empleando el concepto de divergencia, la divergencia de un campo vectorial en un punto (x, y, z) , corresponde al flujo neto del campo afuera de esta caja centrada en (x, y, z) (Tabla 2).

Si $\nabla \cdot F(x, y, z) > 0$	La cantidad de fluido que sale de la caja es mayor que la cantidad que entra
Si $\nabla \cdot F(x, y, z) < 0$	La cantidad de flujo que entra en la caja es mayor que el que sale y a ese punto (x, y, z) se le puede denominar sumidero.
Si $\nabla \cdot F(x, y, z) = 0$	El campo vectorial F es una Fuente libre o incomprensible

Tabla 2. Comportamiento de la Divergencia en un campo vectorial

El operador Divergencia para imágenes Multiespectrales

El operador Divergencia de un campo vectorial U se define como

$$\square \quad U = \sum_{i=1}^n \frac{\partial u_{.xi}}{\partial x_i} \quad (\text{IV. 22})$$

Donde:

x_i es la i -ésima coordenada de un sistema de coordenadas cartesianas en un espacio euclidiano. $U_{.xi}$ es la componente del campo vectorial $U = u(x_1, x_2, \dots, x_n)$ a lo largo de la dirección definida por x_i .

Las derivadas involucradas en la ecuación no pueden ser definidas en forma analítica, por lo tanto se tienen que aproximar por medio de diferencias finitas.

El operador divergencia mide el flujo de un campo vectorial. El flujo es la intensidad de líneas de campo que pasan a través de un área unitaria δa . Si \hat{n} es un vector unitario, normal al área de δa para una superficie cerrada S , el flujo está dado por:

$$\phi = \int_S \hat{n} \cdot \mathbf{U} \, \delta a \quad (\text{IV. 23})$$

Si se considera un hipercubo de tamaño $\delta x_1 \delta x_2 \delta x_3 \dots \delta x_n$ donde δx_i es el lado que se encuentra en la dirección de la coordenada x_i . Como primera aproximación, el flujo a través de los lados de un hipercubo se puede expresar como:

$$\phi_{x_i} = [\mathbf{u}_{x_i}(x_1, \dots, x_i + \delta x_i, x_{i+1}, \dots, x_n) - \mathbf{u}_{x_i}(x_1, \dots, x_i, x_{i+1}, \dots, x_n)] \delta x_1 \dots \delta x_{i-1} \delta x_{i+1} \dots \delta x_n, \quad \forall x_i \quad (\text{IV. 24})$$

Acumulando las contribuciones de flujo ϕ_{x_i} , $\forall x_i$ y tomando el límite cuando $\delta x_i \rightarrow 0$, se tiene que:

$$\phi = \sum_{i=1}^n \frac{\partial \mathbf{u}_{x_i}}{\partial x_i} \quad (\text{IV. 25})$$

La divergencia es una medida del flujo del campo vectorial $\mathbf{U} = \mathbf{u}(x_1, x_2, \dots, x_n)$ y puede ser calculada por medio de la aproximación por diferencias infinitas.

De acuerdo con lo anterior el flujo puede ser positivo, negativo o cero (Tabla 3).

Si $\phi > 0$	Se tiene una superficie en donde la textura cambia de suave a rugosa
Si $\phi < 0$	Se tiene una superficie en donde la textura cambia de rugosa a suave
Si $\phi = 0$	Indica que el flujo neto de las líneas que atraviesan el hipercubo es cero, por lo tanto se tiene una región de textura suave.

Tabla 3. Comportamiento de la Textura dependiendo del flujo de un campo vectorial

Si se considera que

$$\mathbf{U} = \rho \mathbf{V} \quad (\text{IV. 26})$$

Donde:

ρ es la densidad de líneas de flujo del campo vectorial U

V es la unidad de volumen del hipercubo

Entonces puede considerarse la ecuación de continuidad para un hipercubo:

$$\frac{\partial^n \rho}{\partial x_1 \partial x_2 \dots \partial x_n} \square (\rho V) = 0 \quad (\text{IV. 27})$$

La ecuación anterior nos señala que el flujo del volumen indica un decremento de la densidad en el interior del volumen. El parámetro ρ representa la densidad de líneas de flujo y es un parámetro de la rugosidad de la imagen de textura.

Cambios grandes de ρ se asocian con texturas rugosas y viceversa, en cambio cambios pequeños de ρ se asocian con texturas suaves y viceversa.

El nivel de intensidad de gris en la imagen resultante es directamente proporcional al valor de $\partial\rho$, es decir:

Niveles de gris $\propto \partial\rho$.

Lo anterior significa que texturas suaves presentan niveles de gris de baja intensidad, es decir, tonos de gris oscuros con valores en sus pixels cercanos a 0, en cambio las texturas rugosas presentan niveles de gris de mayor intensidad, tonos de gris claros con valores en sus pixels cercanos a 255.

En resumen, la divergencia es la suma de los gradientes en las n direcciones ortogonales de un campo vectorial n -dimensional, es decir, el operador divergencia produce una medida de la textura asociada con el campo vectorial de una imagen multiespectral.

En regiones con textura suave, la densidad de líneas es baja y la divergencia proporciona valores cercanos a cero. De manera similar al ir aumentando el grado de rugosidad de textura, tanto la densidad como el valor de la divergencia se incrementan.

Capítulo 5

5. El Algoritmo Lbp



Ilustración 49 Imagen obtenida al aplicar el algoritmo LBP a una imagen monoespectral

5.1 Análisis de texturas con patrones locales binarios

La textura en una imagen puede proveer información invaluable al momento de identificar objetos presentes en ésta. El sistema de visión humano no sólo puede distinguir objetos basándose en el color; también puede distinguir la textura.

La principal característica de la textura es la repetición de un patrón o de varios patrones contenidos en una región. El patrón puede repetirse exactamente, o puede comportarse como un conjunto de pequeñas variaciones, esto lo realiza posiblemente como una función de posición. (Lucieer, 2005).

Un simple y nuevo modelo de textura es el LBP (Ojala, 1996), el cual ha demostrado ser un adecuado método para analizar texturas en imágenes monoespectrales. El LBP se ha convertido una medida de la textura de una imagen, mostrando excelentes resultados en términos de precisión y complejidad computacional.

El operador LBP puede ser visto como un acercamiento unificador hacia los modelos estadísticos y los modelos estructurales del análisis de texturas.

En el modelo LBP la textura es descrita en términos de micro-primitivas (textones) y sus reglas estadísticas; opcionalmente, las primitivas pueden ser relacionadas como una medida complementaria del contraste local de la imagen, lo cual mide la fuerza de las primitivas.

Al principio el operador LBP de textura fue introducido como una medida complementaria para medir el contraste local de una imagen. En esta primera versión el LBP trabaja con los 8 vecinos más cercanos de un píxel y maneja los conceptos de umbral y pesos. El umbral es un valor que se toma como referencia y es propio de cada vecindad; los pesos son valores resultantes de trabajar con 2^k , donde $0 \leq k \leq 7$.

El valor del píxel central es el que se toma como umbral, posteriormente se produce un código LBP para cada vecindad al multiplicar los valores del umbral con los pesos dados a los píxels correspondientes y finalmente se suma el resultado.

Para ejemplificar lo anterior se tiene el ejemplo con la siguiente vecindad:

6	5	2
7	6	1
9	8	7

Primero se localiza al píxel central y su valor; para la vecindad del ejemplo este valor es de 6.

Los valores de los píxels de la vecindad (P_v) se comparan con el valor del píxel central (P_c) y se establece la siguiente condición:

$$\text{Si } \begin{cases} P_v \geq P_c \square P_v = 1 \\ P_v < P_c \square P_v = 0 \end{cases} \quad (\text{V. 1})$$

Se construye una nueva tabla, con los valores de los píxels de la vecindad umbralizada

Umbral		
1	0	0
1		0
1	1	1

Se crea una tabla de pesos ($\text{pesos}=2^k$, $0 \leq k \leq 7$) en donde la posición de cada uno se asigna arbitrariamente, para el ejemplo se asignó en dirección de las manecillas del reloj, comenzando en la esquina superior derecha.

Pesos		
1	2	4
128		8
64	32	16

Los valores de los pixels en la vecindad umbralizada son multiplicados por los pesos dados a los pixels correspondientes

Umbral		Pesos	=	LBP																											
<table style="background-color: #cccccc; padding: 5px;"> <tr><td>1</td><td>0</td><td>0</td></tr> <tr><td>1</td><td style="background-color: #cccccc;"></td><td>0</td></tr> <tr><td>1</td><td>1</td><td>1</td></tr> </table>	1	0	0	1		0	1	1	1	X	<table style="background-color: #cccccc; padding: 5px;"> <tr><td>1</td><td>2</td><td>4</td></tr> <tr><td>128</td><td style="background-color: #cccccc;"></td><td>8</td></tr> <tr><td>64</td><td>32</td><td>16</td></tr> </table>	1	2	4	128		8	64	32	16	=	<table style="background-color: #cccccc; padding: 5px;"> <tr><td>1</td><td>0</td><td>0</td></tr> <tr><td>128</td><td style="background-color: #cccccc;"></td><td>0</td></tr> <tr><td>64</td><td>32</td><td>16</td></tr> </table>	1	0	0	128		0	64	32	16
1	0	0																													
1		0																													
1	1	1																													
1	2	4																													
128		8																													
64	32	16																													
1	0	0																													
128		0																													
64	32	16																													

De aquí se obtiene el siguiente patrón binario: 11110001

El patrón se comienza a formar de izquierda a derecha el cual corresponde a los valores del umbral.

Los valores de los 8 pixels resultantes de multiplicar el umbral por los pesos son sumados para obtener el número LBP del vecindario:

128	64	32	16	8	4	2	1
1	1	1	1	0	0	0	1

$$LBP = 1 + 0 + 0 + 0 + 16 + 32 + 64 + 128 = 241$$

Se obtiene el valor de C (contraste), donde:

$$C = \frac{\text{Promedio de valores sin umbralizar con valor de } 1}{1} - \frac{\text{Promedio de valores sin umbralizar con valor de } 0}{0} \quad (V.2)$$

$$C = \frac{6+7+8+9+7}{5} - \frac{5+2+1}{3} = 4.7$$

El promedio de los valores de los niveles de gris debajo del píxel central es sustraído de todos los valores de los niveles de gris superiores o iguales al valor del píxel central.

Las distribuciones bidimensionales del LBP y las medidas del contraste local fueron usadas como características.

El operador fue nombrado LBP/C y se han reportado muy buenos índices de discriminación con texturas seleccionadas del álbum fotográfico de Brodatz.

El operador LBP está relacionado con métodos muy bien conocidos del análisis de texturas; estas relaciones se muestran en el esquema de la Ilustración 50.

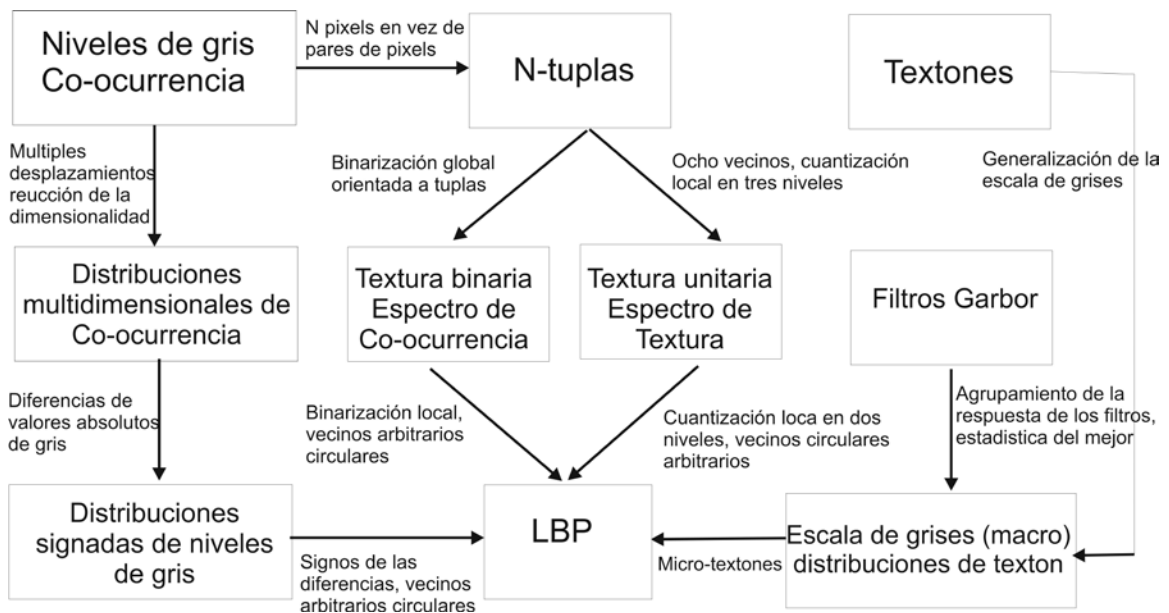


Ilustración 50 Relaciones del operador LBP con otros métodos de textura

5.2 Formalismo teórico

Se define la textura T en un vecindario local de una imágenes en niveles de gris, como la distribución conjunta de los niveles de gris de $P + 1$ ($P > 0$) pixels en una imagen.

$$T = t(g_c, g_0, \dots, g_{p-1}) \quad (V.3)$$

Donde g_c corresponde al valor de gris del píxel central de un vecindario local y g_p ($p = 0, 1, \dots, p-1$) corresponde a los valores de gris de P igualmente espaciados sobre un círculo de radio R ($R > 0$) que forma un conjunto de vecinos simétrico circular. Este conjunto de $P + 1$ pixels es después denotado por G_p .

En el dominio digital de una imagen, las coordenadas de los vecinos g_p están dadas por

$$x_c + R \cos \frac{2\pi p}{P}, \quad y_c + R \sin \frac{2\pi p}{P} \quad (V.4)$$

donde (x_c, y_c) son las coordenadas del píxel central

La Ilustración 51 ilustra tres conjuntos de vecinos circulares simétricos con diferentes valores de P y R . Los valores de los vecinos que no caen exactamente en los pixels son estimados por interpolación lineal. Puesto que la correlación entre pixels decrece con la distancia, la información textural de una imagen puede ser obtenida de vecindarios locales.

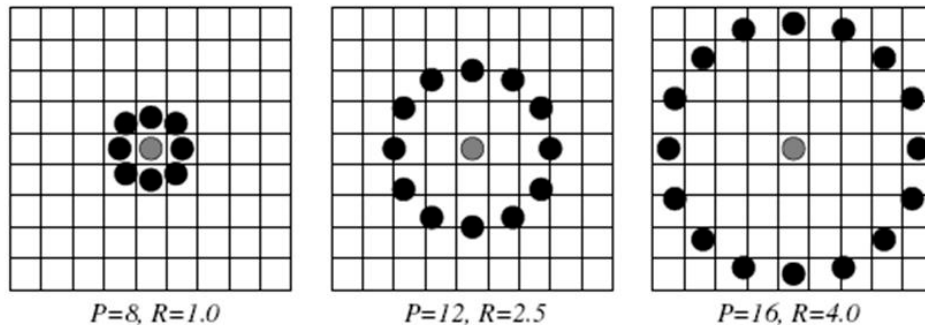


Ilustración 51 Diferentes conjuntos de vecinos circulares simétricos con valores de P y R diferentes

Si el valor del píxel central es restado a los valores de los vecinos, la textura local puede ser representada – sin perder información – como una distribución conjunta del valor del píxel central y las diferencias:

$$T = t(g_c, g_0 - g_c, \dots, g_{p-1} - g_c) \quad (\text{V.5})$$

Suponiendo que las diferencias son independientes de g_c , la distribución puede ser factorizada:

$$T \approx t(g_c)(g_0 - g_c, \dots, g_{p-1} - g_c) \quad (\text{V.6})$$

En la práctica, la suposición de independencia podría no siempre ser verdad. Debido a la naturaleza limitada de los valores en las imágenes digitales muy altos y muy pocos valores de g_c serán obviamente limitados abajo del rango de las posibles diferencias. Sin embargo, aceptando las pequeñas posibles pérdidas de información es posible obtener invariancia con respecto a los cambios en la escala de grises.

Debido a que $t(g_c)$ describe el total de la luminosidad de una imagen, la cual no tiene relación con la textura local de la imagen, $t(g_c)$ no provee información útil para el análisis de textura.

Por tanto, mucha de la información acerca de las características de textura en la distribución conjunta original es preservada en la distribución conjunta de las diferencias

$$T \approx t(g_0 - g_c, \dots, g_{p-1} - g_c) \quad (\text{V.7})$$

La distribución de la diferencia P-dimensional relaciona la ocurrencia de diferentes patrones de textura de cada píxel en el vecindario. Para una constante o lenta variación en las regiones, las diferencias del grupo son cercanas a cero.

En una mancha (*spot*), todas las diferencias son relativamente grandes. Sobre un borde, las diferencias en algunas direcciones son más grandes que en otras.

Aunque invariante contra del cambio en la escala de grises, las diferencias son afectadas al escalarlas. Para conseguir la invariancia respecto a cualquier transformación monotónica en niveles de grises, únicamente se consideran los signos de las diferencias.

$$T \approx t[s(g_0 - g_c), \dots, s(g_{p-1} - g_c)] \quad (V.8)$$

Donde

$$s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (V.9)$$

Ahora bien, pesos binomiales 2^p son asignados a cada signo $s(g_p - g_c)$ transformando las diferencias de un vecindario en un código LBP único. Este código caracteriza la textura local de la imagen alrededor de (x_c, y_c) :

$$LBP_{P,R}(x_c, y_c) = \sum_{p=0}^{P-1} s(g_p - g_c) 2^p \quad (V.10)$$

En la práctica la ecuación anterior significa que los signos de las diferencias en un vecindario son interpretados como números binarios P-bit, resultando en 2^P distintos valores de código LBP.

La distribución local en niveles de grises, por ejemplo la textura puede ser descrita con una distribución discreta 2^P -bin de código LBP.

$$T \approx t(LBP_{P,R}(x_c, y_c)) \quad (V.11)$$

Suponiendo que se tiene una imagen dada de dimensión $N \times M$, es decir $(x_c \in \{0, \dots, N-1\}, y_c \in \{0, \dots, M-1\})$. Calculando la distribución $LBP_{P,R}$ (vector característico) para esta imagen, sólo se considera la parte central porque un vecindario suficientemente grande no puede ser usado en los bordes. El código LBP es calculado para cada píxel en

una porción muy corta de la imagen, y la distribución del código es usada como un vector característico, denotado por S :

$$S = t(\text{LBP}_{P,R}(x, y)), x \in \{[R], \dots, N-1-[R]\}, y \in \{[R], \dots, M-1-[R]\} \quad (\text{V.12})$$

El LBP original es muy similar al $\text{LBP}_{8,1}$, con dos diferencias. Primero, el vecindario en la definición general es indexado circularmente, haciendo más fácil obtener la rotación invariante de los descriptores de textura. Segundo, los pixels en la diagonal en un vecindario de 3×3 son interpolados en $\text{LBP}_{8,1}$.

Ojalá et. al. (1996) encontró que no todos los patrones locales binarios describen de manera apropiada algunas propiedades de la textura. El algoritmo LBP captura la uniformidad del pixel central en relación con sus vecinos, pero no captura la uniformidad del vecindario por sí mismo.

El concepto de patrones “uniformes” fue introducido por Mäenpää, él observó que ciertos patrones parecen ser fundamentales en propiedades de la textura. Proporcionando una inmensa mayoría de patrones, algunas veces sobre el 90%. Esta proposición fue más allá, ya que fue confirmada con una gran cantidad de datos por Ojala y otros. Estos patrones son llamados “uniformes” porque ellos tienen una cosa en común, al menos existen dos transiciones en el código binario circular: Una transición de uno a cero o una transición de cero a uno.

Para definir formalmente la “uniformidad” de un vecindario G , se necesita una medida de uniformidad U :

$$U(G_P) = |s(g_{P-1} \ g_c) - s(g_0 \ g_c)| + \sum_{p=1}^{P-1} |s(g_p \ g_c) - s(g_{p-1} \ g_c)| \quad (\text{V.13})$$

Patrones con un valor de U menores o iguales a dos son designados como “uniformes”. Para un número binario P -bit, el valor de U puede ser calculado eficientemente con aritmética binaria como sigue:

$$U(x) = \sum_{p=0}^{P-1} F(x \text{ xor ROR}(x,1), p) \quad (\text{V.14})$$

Donde x es un número binario.

U corresponde al número de transiciones espaciales o cambios de bits de 0/1 en el patrón.

Cuando $g_p = g_o$, U_c (uniformidad del pixel central) es definida como:

$$U_c = \sum_{i=1}^P |\text{sign}(g_i - g_c) - \text{sign}(g_{i-1} - g_c)| \quad (\text{V. 13})$$

5.3 Contraste y patrones de textura

El contraste es una propiedad de la textura usualmente considerada como una sugerencia importante para nuestra visión de sistema, pero el operador LBP por si mismo ignora totalmente la magnitud de las diferencias en los niveles de gris. En muchas aplicaciones, especialmente en la inspección visual industrial, la luminosidad puede ser controlada de manera exacta. En tal situación, una escala de grises puramente invariante al operador de textura podría perder información útil, y adicionalmente la información dependiente de la escala de grises, podría elevar la precisión del método. Además, en aplicaciones tales como la segmentación de una imagen, los cambios graduales en la iluminación podrían no requerir de usar un método invariante en niveles de gris.

En una visión más general, la textura es distinguida no solo por patrones de textura, sino también por la fuerza de los patrones. La textura puede hasta ser incluso considerada como un fenómeno bidimensional caracterizado por dos propiedades ortogonales: estructura espacial (patrones) y contraste (la fuerza de los patrones). La información de los patrones es independiente de la escala de grises, mientras que el contraste no lo es. Por otro lado, el contraste no es afectado por la rotación, pero los patrones si lo son. Estas dos medidas se complementan a cada otra de una manera muy útil.

El operador LBP fue originalmente diseñado para este propósito: complementar una medida de escala de grises dependiente de la “cantidad” de textura. Ojala y otros usen

una distribución conjunta del código LBP y una medida del contraste local (LBP/C) como descriptores de textura.

La limitación más importante del operador original LBP es su pequeña área espacial de soporte. Las características calculadas en un vecindario local de 3×3 no pueden capturar estructuras a gran escala que tal vez sean las características dominantes de algunas texturas. Sin embargo, los códigos adyacentes LBP no son totalmente independientes uno del otro.

Como un muy simple operador de textura, el LBP es idealmente apropiado para aplicaciones que requieren una rápida extracción de características. Debido a su simplicidad y ejecución, el operador LBP se ha aplicado en:

- Inspección Visual Industrial
- Recuperación de Imágenes
- Análisis de escenas
- Análisis de rostros, etcétera.

5.4 Modificaciones al algoritmo LBP para imágenes multiespectrales

Con base en el principio del algoritmo LBP se realizó primero un programa en MATLAB para el análisis de imágenes multiespectrales. Su diseño y construcción está asociado con el número de bandas de la imagen multiespectral. Para imágenes Landsat TM se considera una dimensión de 6 bandas ya que no se incluye la banda del infrarrojo térmico. También se puede reducir el número de dimensiones a un tamaño menor al aplicar la descomposición en componentes principales a la imagen multiespectral, en tal caso el programa solo analizaría 3 bandas en donde la correlación entre éstas es de cero. Para otras imágenes multiespectrales la dimensionalidad puede ser diferente, por todo lo anterior, el algoritmo que se desarrolla en esta tesis emplea un número de N bandas multiespectrales.

Hay que mencionar que el resultado del algoritmo arroja también una imagen multiespectral, es decir el número de bandas de entrada es igual al número de bandas de salida.

Si se carga una imagen monoespectral, el algoritmo procesa con éxito su única banda, dando como resultado una imagen monoespectral.

El procedimiento de este algoritmo, resultado de modificar el algoritmo LBP para imágenes multiespectrales, se describe a continuación:

- a) Se carga la imagen multiespectral de N bandas (de 1 a 10) en memoria principal. Estas bandas pueden ser las bandas originales o bien bandas que hayan sido modificadas por algún tipo de transformación, esto lo decide el usuario. En el capítulo VI se proponen diferentes combinaciones para observar cuál es la combinación que proporciona el mejor resultado.
- b) Se determina la dimensión de la imagen multiespectral, es decir alto (Y) ancho (X), y número de bandas (N)
- c) Se asignan las coordenadas de los 8 vecinos más cercanos
$$[-1 -1; -1 0; -1 1; 0 -1; 0 1; 1 -1; 1 0; 1 1]$$
- d) Se localiza el origen la ventana inicial que tiene un tamaño de $3 \times 3 \times N$
- e) Se recorre la imagen desde la esquina superior derecha en el pixel que corresponde a la coordenada (2,2,N) de la imagen. Al momento que la imagen es recorrida por la ventana se resta el valor del centro de la ventana menos el valor de los vecinos
- f) Se realiza una nueva imagen resultado, la cual tiene una dimensión de (X-2,Y-2,N) la cual contiene el valor máximo resultado de la resta de los centros menos los vecinos, esta imagen contiene un mapa de textura de la imagen.
- g) La imagen resultado es al igual que la imagen original una imagen multiespectral, que está compuesta por el vector de máximas diferencias.

Capítulo 6

6.- Métodos y Materiales.



Ilustración 52 Composición RGB (7, 4,2) del valle de México

El Valle de México o Valle de Anáhuac (Ilustración 52) es una cuenca endorreica cerrada pues no cuenta con drenaje natural a las vertientes oceánicas. Está situada en el tramo central de Eje Neo volcánico Transmexicano, que atraviesa el territorio nacional a lo largo del paralelo 19° norte, desde el Golfo de México hasta el Océano Pacífico.

La cuenca del Valle de México se caracteriza por tener un fondo raso, formado por un relleno de depósitos aluviales y lacustres. El perímetro irregular del valle delimita una superficie de 8000 km² aproximadamente. Las montañas circundantes se cubren de bosques hasta los 4000 msnm y sus elevaciones, con casquetes nevados, superan los 5000 metros sobre el nivel del mar.

El clima se caracteriza por inviernos secos y temperaturas bajas en las noches y veranos húmedos y temperaturas templadas. La precipitación pluvial varía notablemente y el 80% ocurre en los meses de junio a octubre. Debido a la accidentada geografía y las

variaciones en la altura, pendiente, suelos y precipitaciones, el Valle de México alberga una amplia gama de zonas ambientales.

En el Valle de México se localiza la Ciudad de México, cuya plaza principal, El Zócalo, se encuentra en las coordenadas 19° 25' 58.53" N, 99° 07' 59.57" O .

La geografía del valle se halla dominada por los volcanes Popocatepetl e Iztacíhuatl, los cuales desde tiempos inmemoriales han sido reverenciados y temidos por los habitantes del valle de México. Por todas estas características, se seleccionó como zona de estudio el Valle de México. La imagen empleada es multiespectral Landsat TM (Ilustración 52).

Ya que la imagen es muy grande y comprende varias zonas, para aplicar el algoritmo y observar su comportamiento, se seleccionaron solamente tres, entre las cuales encontramos áreas urbanas, áreas de cultivo, áreas que contienen cuerpos de agua, áreas en donde están presentes diferentes formas de relieve. Estas tres zonas son:

- 1.- Aeropuerto de la Ciudad de México y Texcoco (Ilustración 53)
- 2.- Sur del Valle de México que comprende las delegaciones de Xochimilco y Tláhuac
- 3.- Volcanes Popocatepetl e Iztacíhuatl

Debido a que entre bandas contiguas existe redundancia de información como primer paso se realizó la descomposición en componentes principales (DCP), ésta se aplicó simultáneamente a todas las bandas que componen la imagen multiespectral del valle de México y obtuvimos a la salida otra imagen multiespectral.

Gracias a la DCP obtuvimos un conjunto de bandas decorrelacionadas que nos formaron una nueva imagen con dimensionalidad menor a la original la cual eliminó la redundancia, y en la cual entre las primeras tres componentes de salida se encuentra concentrada la información significativa; en las imágenes restantes tenemos muy poca información quedando en estas últimas el ruido contenido en la imagen. Con la DCP realizamos un realce global a la imagen. Por lo tanto para el análisis del comportamiento del algoritmo ocupamos las componentes principales de la imagen original.

Posteriormente a las componentes principales de la imagen original se les aplicó la modificación del algoritmo Lbp, obteniendo como resultado otra imagen multiespectral, en la cual se observan con claridad los bordes y la textura de la imagen. En los resultados que se presentan se puede observar la primera componente principal ya que como se

mencionó anteriormente es ésta la que presenta la mayor cantidad de información significativa de la imagen (82.33420% de información significativa).

Para obtener mejores resultados en cuanto al análisis de textura aplicamos el operador divergencia a la primera componente principal con bordes y también aplicamos el operador divergencia a la primera componente principal pero sin bordes para observar las diferencias.

El operador divergencia mide el flujo de un campo vectorial, el flujo es la intensidad de líneas de campo que pasan a través de un área unitaria, y puede ser positivo, negativo o cero. Si el flujo tiene un valor de cero quiere decir que no hay flujo neto y por lo tanto el cambio en la densidad de las líneas de flujo también es cero.

Una región con textura se encuentra formada por pixels que al ser representados en un campo vectorial están agrupados en nubes espectrales.

Un valor de la divergencia corresponderá entonces a un rasgo de textura suave cuando se esté atravesando una nube espectral con densidad homogénea. De lo anterior que la densidad de la nube espectral es un parámetro de rugosidad en la imagen de textura y nos indica que tan fina o que tan burda es la textura en la imagen analizada.

Cambios grandes en la densidad de la nube espectral son asociados con textura rugosa mientras que cambios significativamente pequeños son asociados con texturas suaves. Por lo tanto, los niveles en la intensidad de gris en el mapa de textura son directamente proporcionales a la densidad de la nube espectral.

Al visualizar el mapa de textura que resulta de aplicar el algoritmo del operador divergencia observaremos en tonos gris oscuro con valores de los pixels tendiendo a cero áreas que representarían texturas suaves, y en tonos de gris claros con valores de los pixels tendiendo a 256 a las áreas que representan texturas rugosas.

Por lo anterior aplicar el operador divergencia a la imagen que obtuvimos de bordes nos ayudará a obtener mejores resultados en cuanto a la textura en las imágenes que analizaremos. Los resultados se muestran a continuación.

6.1 Aeropuerto de la Ciudad de México



Ilustración 53 Composición RGB (7, 4, 2) del Aeropuerto de la Ciudad de México

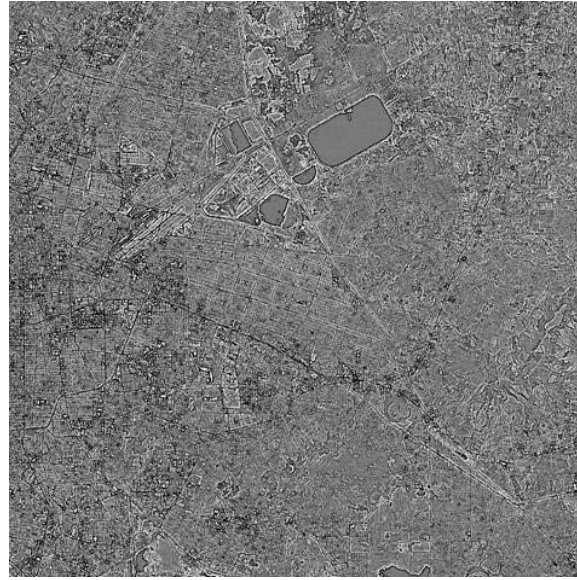
Localizado en el norte del Valle de México el Aeropuerto Internacional de la Ciudad de México "Benito Juárez" (Ilustración 53) fue construido cerca del Antiguo Aeródromo Militar de Balbuena. Se ubica en la delegación Venustiano Carranza y es el primer aeropuerto civil de México. Es el principal puerto aéreo nacional e internacional del país. Se decidió analizar esta área debido a que contiene un gran área urbana que es Ciudad Netzahualcóyotl y Chimalhuacán, también presenta cuerpos de agua pertenecientes al sistema de aguas de la Ciudad de México. Dentro de esta zona podemos observar el comportamiento del algoritmo:

- a) Bordes bien definidos para cuerpos de agua
- b) Delimitación del área urbana y visualización de carreteras

Primera Componente Principal sin bordes

**Ilustración 54** Primera Componente Principal sin bordes del aeropuerto de la Ciudad de México

Primera Componente Principal con bordes

**Ilustración 55** Primera Componente Principal con bordes del aeropuerto de la Ciudad de México

Divergencia a la Primera Componente Principal sin bordes

**Ilustración 56** Divergencia de la Primera Componente Principal sin bordes del aeropuerto de la Ciudad de México

Divergencia a la Primera Componente Principal con bordes

**Ilustración 57** Divergencia de la Primera Componente Principal con bordes del aeropuerto de la Ciudad de México



Ilustración 58 Mapa del área de Texcoco



Ilustración 59 Primera componente principal del área de Texcoco sin bordes



Ilustración 60 Primera componente principal del área de Texcoco con bordes

En las imágenes de la izquierda observamos los cuerpos de agua, como estos cuerpos de agua son artificiales se aprecia de mejor manera sus límites, bien definidos y regulares. En la imagen correspondiente a la primera componente principal sin bordes (Ilustración 59), se observa bien definido lo descrito anteriormente; para la imagen con bordes (Ilustración 60), de igual manera se observa con buena calidad los bordes y además se notan los sedimentos que tiene el vaso regulador “Cola de pato” que es parte del sistema de drenaje de la ciudad de México.

En la Ilustración 60 observamos en base a su textura y color cuerpos de agua e infraestructura (carreteras y avenidas) de color gris oscuro. Mientras que el color gris claro nos representa áreas construidas y de vegetación con actividad agrícola.



Ilustración 61 Divergencia de la Primera componente principal del área de Texcoco sin bordes

En la imagen correspondiente a la divergencia de la primera componente principal sin bores (Imagen 61) observamos de color gris oscuro tendiendo al negro las texturas suaves.

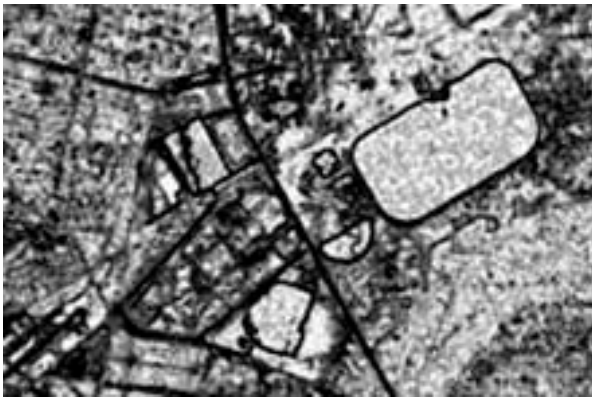


Ilustración 62 Divergencia de la Primera componente principal del área de Texcoco con bordes

En la imagen correspondiente a la divergencia de la primera componente principal con bordes (Ilustración 62) observamos de mejor calidad la textura, en color gris claro tendiendo al blanco para texturas rugosas como son los edificios y en un tono gris oscuro tendiendo al negro para texturas suaves como son las carreteras que se pueden ver, en esta imagen podemos observar la manera en que los bordes se ven con claridad.



Ilustración 63 Mapa de Ciudad Netzahualcóyotl

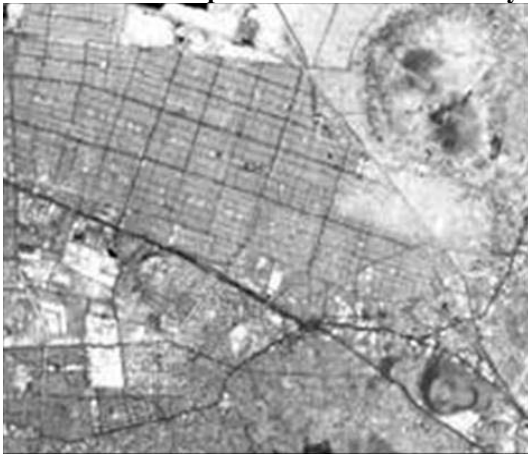


Ilustración 64 Primera Componente Principal del área de Netzahualcóyotl sin bordes

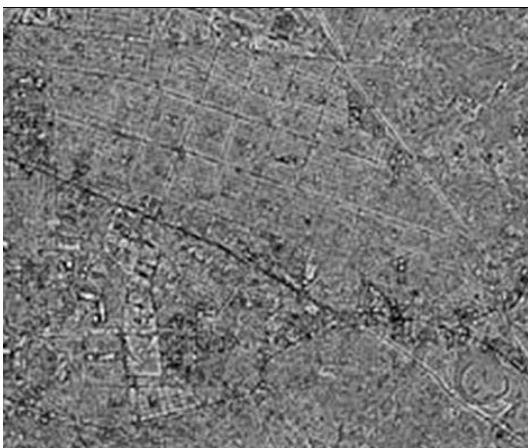


Ilustración 65 Primera Componente Principal del área de Netzahualcóyotl con bordes

Las tres imágenes de la izquierda corresponden a Ciudad Netzahualcóyotl y los Reyes Acaquilpan (Ilustración 63). Podemos observar también una parte de Chimalhuacán y el relleno sanitario de Santa Catarina, además de la carretera México-Puebla

En la imagen que corresponde a la primera componente principal sin bordes (Ilustración 64) se observan con claridad las áreas anteriormente descritas

En la tercera imagen, en la cual se aplicó el algoritmo Lbp (Ilustración 65), observamos de un color gris claro la textura que nos da la ubicación de una gran área urbana así mismo se observa el trazo de calles y avenidas con esa misma textura pero en color más oscuro, aunque se pierde la calidad en los edificios, ello se puede deber a la resolución de la imagen.



Ilustración 66 Divergencia de la Primera Componente Principal del área de Netzahualcóyotl sin bordes



Ilustración 67 Divergencia de la Primera Componente Principal del área de Netzahualcóyotl con bordes

En la imagen correspondiente a la divergencia de la primera componente principal sin bordes (Imagen 66) observamos de color gris oscuro tendiendo al negro con valores pequeños en los pixels las texturas suaves como son las carreteras, observamos que se pueden ver con claridad la delimitación de las manzanas .

En la imagen correspondiente a la divergencia de la primera componente principal con bordes (Ilustración 67) observamos con una mejor calidad la textura de la imagen. En color gris oscuro para objetos que presentan textura suave como las carreteras y de color gris claro para objetos con textura rugosa, observamos claramente el relleno sanitario de santa Catarina en color gris claro tendiendo al blanco al igual que la zona de Chimalhuacán por ser zonas con relieve no homogéneo.

6.2 Sur del Distrito Federal Xochimilco, Tláhuac

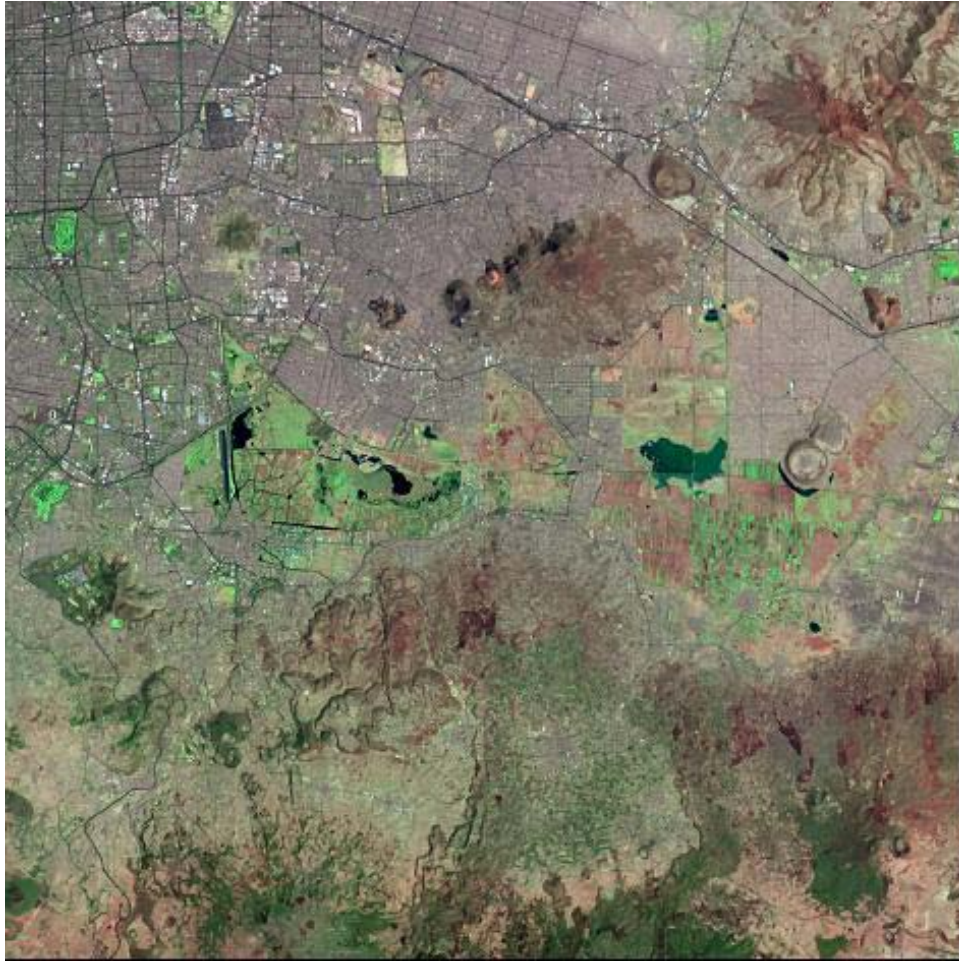


Ilustración 68 Composición RGB (7, 4,2) de la zona de Xochimilco y Tláhuac

Localizadas al sur del Valle de México se encuentran las delegaciones de Xochimilco y Tláhuac (Ilustración 68), se realizó el análisis de esta zona para ver de qué manera se comporta el algoritmo para las zonas de cultivo que aún estando en el distrito federal conservan su origen rural, además de ser un gran pulmón de aire para la Ciudad de México. En esta imagen podemos observar en la parte del centro los ejidos de Xochimilco y San Gregorio, en la parte derecha observamos el valle de Chalco, observamos también zonas de cultivo y cuerpos de agua. Dentro de esta zona podemos observar el comportamiento del algoritmo:

- a) Bordes bien definidos para Cuerpos de agua, relieve y áreas de cultivo.
- b) Caminos, carreteras y zona urbana bien definida

Primera Componente Principal sin bordes



Ilustración 69 Primera componente del sur del Valle de México sin bordes

Primera Componente Principal con bordes

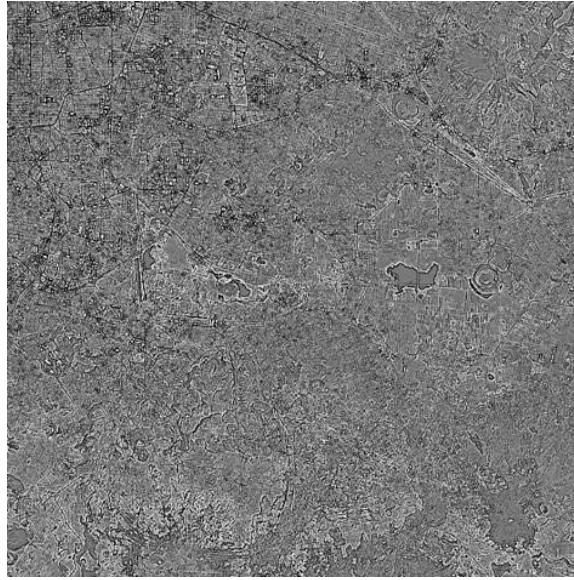


Ilustración 70 Primera componente del sur del Valle de México con bordes

Divergencia a la Primera Componente Principal sin bordes



Ilustración 71 Divergencia de la Primera componente del sur del Valle de México sin bordes

Divergencia a la Primera Componente Principal con bordes

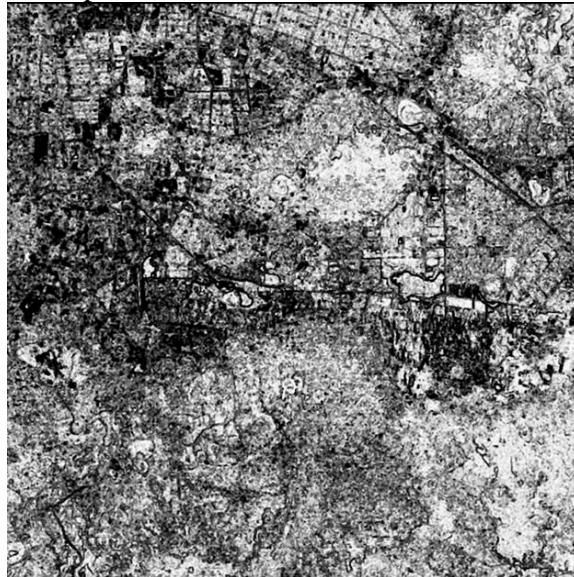


Ilustración 72 Divergencia de la Primera componente del sur del Valle de México con bordes



Ilustración 73 Mapa de la zona de Chalco

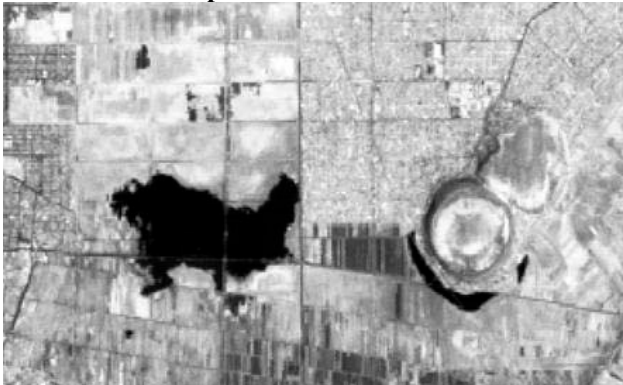


Ilustración 74 Primera Componente Principal de la zona de Chalco sin bordes



Ilustración 75 Primera Componente Principal de la zona de Chalco con bordes

En las imágenes de la izquierda observamos el área de Chalco, en la imagen correspondiente a la primera componente principal sin bordes (Ilustración 74) observamos claramente la laguna de Chalco, áreas de cultivo, avenidas principales y el cerro de Xico.

En la imagen que corresponde a la primera componente principal con bordes (Ilustración 75) observamos que la textura fina color gris nos representa áreas de vegetación definida así mismo se observa cuerpos de agua con el contorno más oscuro donde su textura es homogénea.

También se observa la textura fina de color más claro que representa áreas de construcción e infraestructura.

Existe otro elemento que se distingue de acuerdo a su textura la cual corresponde a una pequeña elevación que representa el cráter del cerro de Xico



Ilustración 76 Divergencia de la Primera Componente Principal de la zona de Chalco sin bordes

En la imagen correspondiente a la divergencia de la primera componente principal sin bordes (Ilustración 76) observamos en color gris oscuro casi negro zonas con textura suave como son los cuerpos de agua y en tono gris claro tendiendo al blanco áreas que presentan textura rugosa, como son las áreas de cultivo y el relieve.

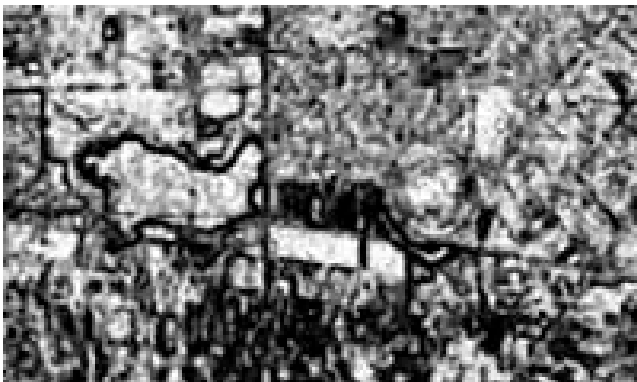


Ilustración 77 Divergencia de la Primera Componente Principal de la zona de Chalco con bordes

En la imagen correspondiente a la divergencia de la primera componente principal con bordes (Ilustración 77) observamos con mejor claridad los bordes del lago de Chalco y los canales, también observamos con claridad en tono gris claro las zonas que presentan una textura rugosa como en cerro de Xico.



Ilustración 78 Mapa del área de Xochimilco



Ilustración 79 Primera componente principal del área de Xochimilco sin bordes



Ilustración 80 Primera componente principal del área de Xochimilco con bordes

Las imágenes de la izquierda corresponden al área de Xochimilco, en la imagen correspondiente a la primera componente principal sin bordes (Ilustración 79) observamos con claridad el área urbana, las carreteras y avenidas principales, los cuerpos de agua, la zona chinampera y la red de canales.

En la imagen correspondiente a la primera componente principal con bordes (Ilustración 80) los elementos más representativos en base a su textura y color son; cuerpos de agua, áreas de cultivo, red de canales de la zona chinampera, infraestructura y construcciones.

Para la zona chinampera y el área de cultivo la textura es fina de color gris claro.

Para los cuerpos de agua la textura es fina, homogénea y de color gris oscuro así como para la red de canales.

Para la zona de infraestructura y construcciones la textura es fina y de color gris claro.



Ilustración 81 Divergencia de la Primera componente principal del área de Xochimilco sin bordes

En la imagen que muestra la divergencia de la primera componente principal sin bordes (Ilustración 81) se observan con claridad los cuerpos de agua como son las la laguna de Xochimilco y la pista de canotaje de Cuemanco, estos se observan en tono gris oscuro ya que estamos hablando de áreas que presentan textura suave, también se observa en un tono gris claro las áreas de cultivo.



Ilustración 82 Divergencia de la Primera componente principal del área de Xochimilco con bordes

En la imagen que muestra la divergencia de la primera componente principal con bordes (Ilustración 82) observamos con mayor claridad las áreas de textura rugosa que presentan valores de pixels que tienden a 256 y presentan un tono gris claro como son las áreas de cultivo o chinampas, observamos en tono gris oscuro los bordes de los cuerpos de agua y la carretera presente en la imagen

6.3 Volcán Popocatépetl e Iztacíhuatl

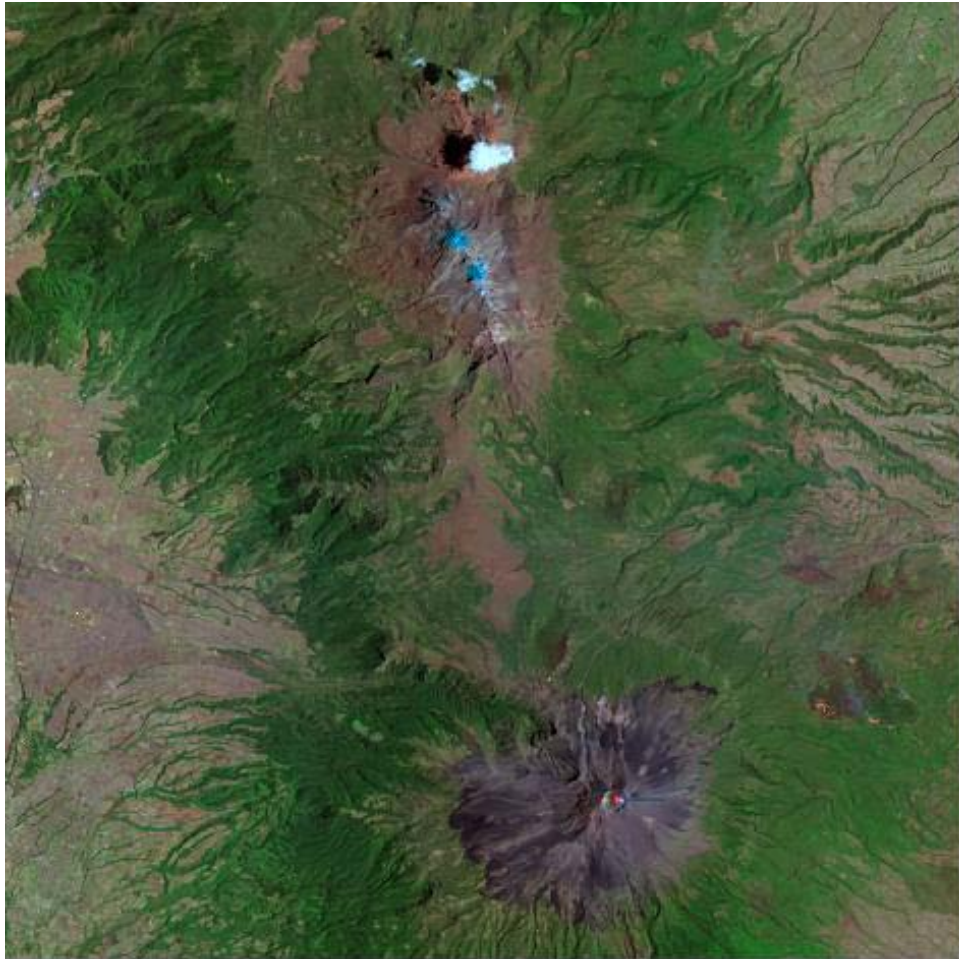


Ilustración 83 Composición RGB (7, 4,2) de los volcanes Popocatépetl e Iztacíhuatl

Los volcanes Popocatépetl e Iztacíhuatl (Ilustración 83) se encuentran localizados entre los límites territoriales de los estados de Morelos, Puebla y México. Se localizan a unos 55 Km del sureste de la Ciudad de México. El Popocatépetl es un volcán de forma cónica simétrica; está unido por la parte norte con el Iztacíhuatl mediante un paso montañoso conocido como Paso de Cortes. En la Ilustración 83 observamos podemos distinguir las zonas boscosas con algunas zonas de cultivo además de las faldas de los volcanes. Además, de manera curiosa observamos una nube sobre el volcán Iztacíhuatl. Dentro de esta zona podemos observar el comportamiento del algoritmo:

- a) Relieve bien definido
- b) Áreas verdes y áreas deforestadas bien definidas

Primera Componente Principal sin bordes



Primera Componente Principal con bordes

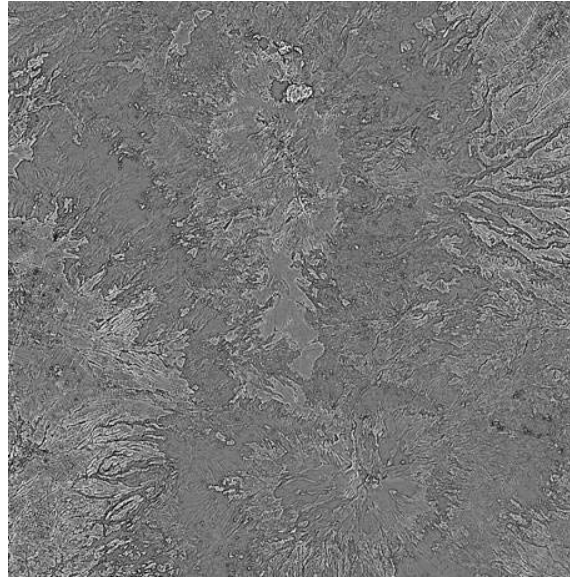
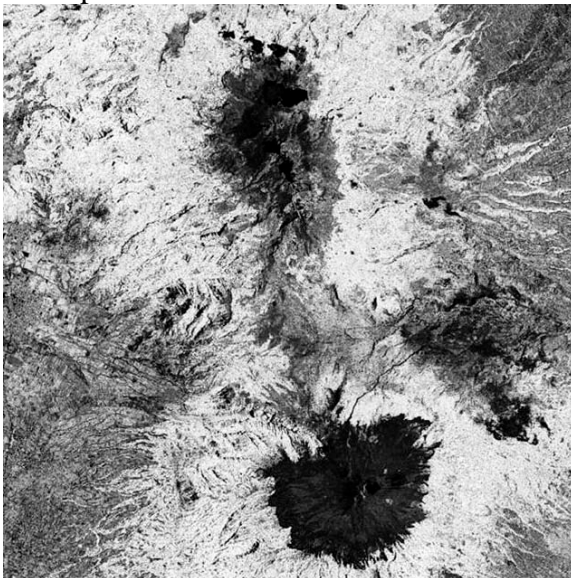


Ilustración 84 Primera componente principal de los volcanes Popocatépetl e Iztacihuatl sin bordes

Ilustración 85 Primera componente principal de los volcanes Popocatépetl e Iztacihuatl con bordes

Divergencia a la Primera Componente Principal sin bordes



Divergencia a la Primera Componente Principal con bordes

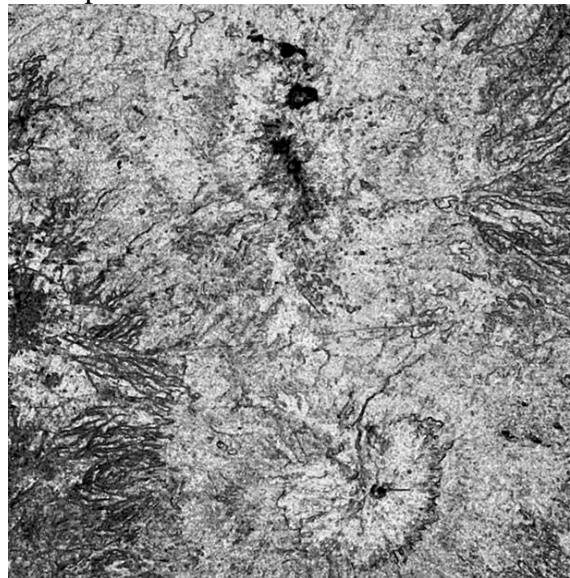


Ilustración 86 Divergencia de la Primera componente principal de los volcanes Popocatépetl e Iztacihuatl sin bordes

Ilustración 87 Divergencia de la Primera componente principal de los volcanes Popocatépetl e Iztacihuatl con bordes



Ilustración 88 Imagen del volcán Popocatépetl

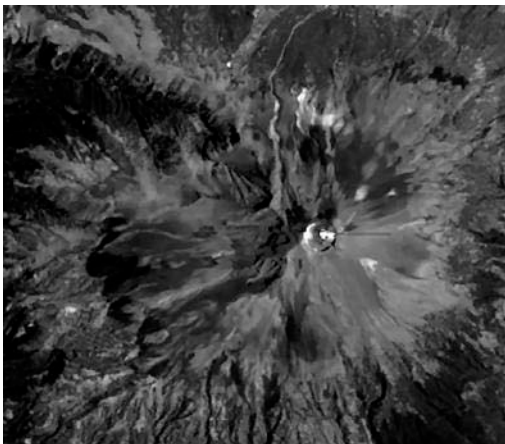


Ilustración 89 Primera componente principal del volcán Popocatépetl sin bordes

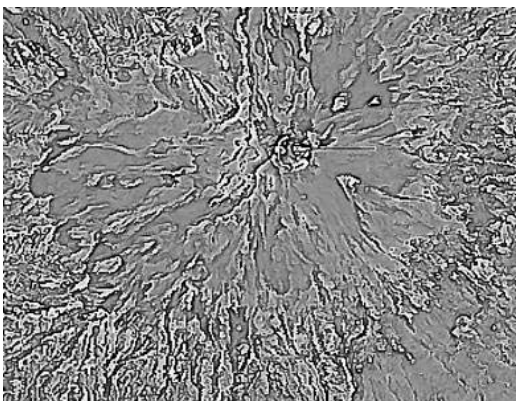


Ilustración 90 Primera componente principal del volcán Popocatépetl con bordes

En las imágenes de la izquierda observamos el relieve correspondiente al volcán Popocatépetl y áreas aledañas, apreciándose en el primer caso (Ilustración 88) el cráter y su edificio volcánico. En la imagen correspondiente a la primera componente principal sin bordes (Ilustración 89), se observa con mayor detalle las pendientes del cono. También se aprecian las áreas correspondientes a las formaciones rocosas formadas por las diferentes erupciones que fueron formando el cono volcánico. En la imagen correspondiente a la primera componente principal con bordes (Ilustración 90) se observa textura fina de color gris claro que representa las aéreas desprovistas de vegetación, estas no son uniformes debido a que son pendientes de ladera del volcán Popocatépetl. También se observa un gris con textura fina granular la cual representa pequeñas aéreas con vegetación de bosque y matorral, estas se ubican en ladera media del volcán Popocatépetl. Por último se observa esta misma textura pero de color gris oscuro que representan pequeñas depresiones tales como barrancas y cañadas.

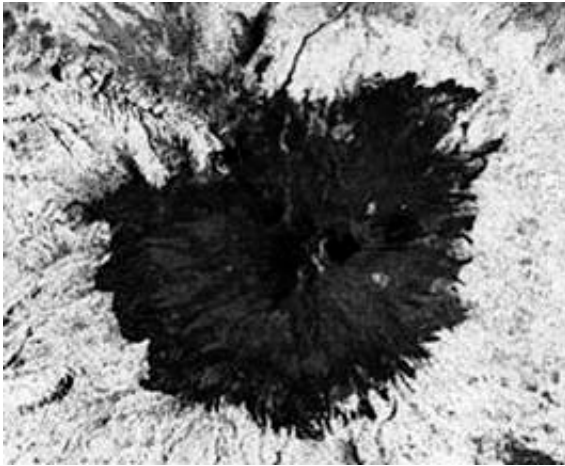


Ilustración 91 Divergencia de la Primera componente principal del volcán Popocatépetl sin bordes

En la imagen que corresponde a la divergencia de la primera componente principal sin bordes (Ilustración 91) observamos en tonos gris oscuros las aéreas desprovistas de vegetación, y en tono gris claro las áreas que presentan vegetación en la ladera media del volcán observamos en tono gris oscuro un camino al norte de la imagen.

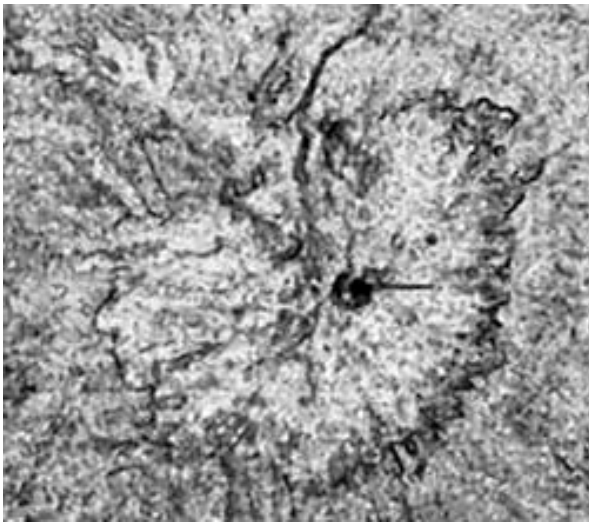


Ilustración 92 Divergencia de la Primera componente principal del volcán Popocatépetl con bordes

En la imagen correspondiente a la divergencia de la primera componente principal con bordes (Ilustración 82) observamos en tonos gris claros las áreas con textura rugosa y en tono gris oscuro los bordes de los límites entre el área que contiene vegetación y el área que carece de ella en el Volcán



Ilustración 93 Imagen del volcán Iztacihuatl

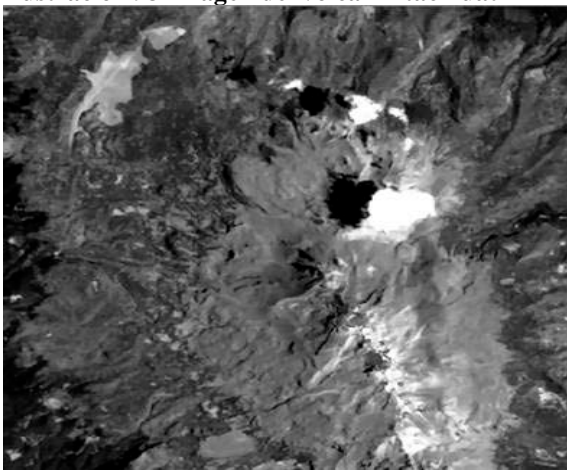


Ilustración 94 Primera componente principal del volcán Iztacihuatl sin bordes

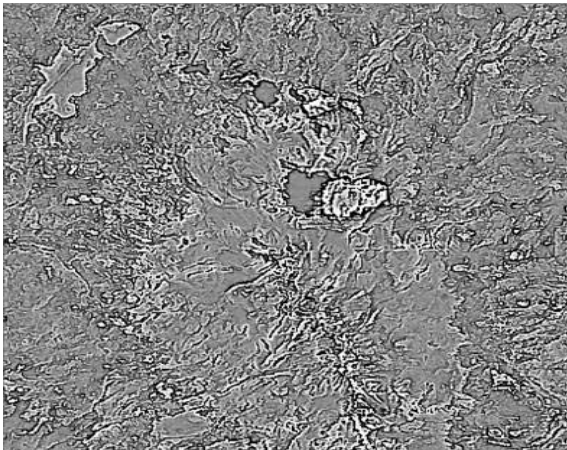


Ilustración 95 Primera componente principal del volcán Iztacihuatl con bordes

En las imágenes de la izquierda observamos el volcán Iztacihuatl (Ilustración 93), en la imagen correspondiente a la primera componente principal sin bordes (Ilustración 94) observamos las pendientes del volcán y el relieve. En la imagen correspondiente a la primera componente principal con bordes (Ilustración 95) observamos un gris con textura fina la cual representa pequeñas aéreas con vegetación y una textura fina de color gris claro y homogénea para zonas desprovistas de vegetación, así mismo observamos de color gris un poco más oscuro depresiones

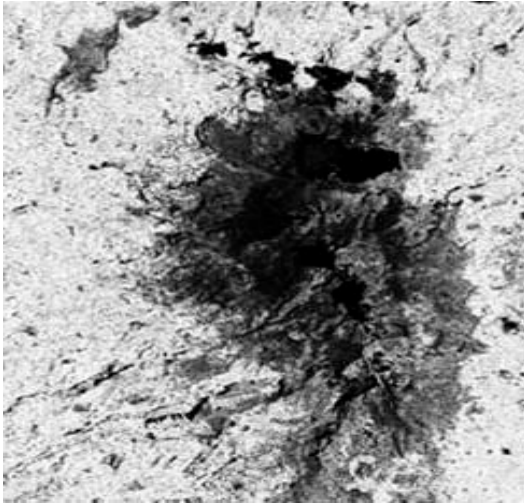


Ilustración 96 Divergencia de la Primera componente principal del volcán Iztacíhuatl sin bordes

En la imagen que corresponde a la divergencia de la primera componente principal sin bordes (Ilustración 96) observamos en color gris claro la zona que presenta vegetación y por lo tanto es una zona que presenta textura rugosa. En color gris oscuro observamos las áreas carentes de vegetación y que presentan caminos, cañadas, laderas o las depresiones del volcán.

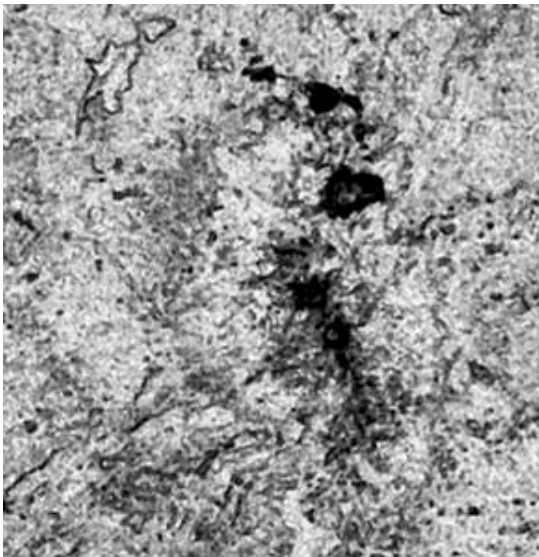


Ilustración 97 Divergencia de la Primera componente principal del volcán Iztacíhuatl con bordes

En la imagen que corresponde a la divergencia de la primera componente principal con bordes (Ilustración 97) observamos en color gris claro el área rugosa que presenta vegetación y en tono oscuro observamos el área que presenta más sinuosidades en el terreno del volcán Iztacíhuatl además observamos con claridad el borde que presenta un área que se encuentra deforestada en la parte superior izquierda de la imagen

Capítulo 7

7.- Conclusiones

La textura en una imagen provee información invaluable al momento de identificar objetos presentes en ésta. En el campo de la percepción remota, es muy importante localizar, identificar, clasificar y segmentar las diferentes zonas que componen una imagen. Debido a que la principal característica de la textura es la repetición de un patrón o de varios patrones contenidos en una región podemos realizar una identificación de las principales zonas de una imagen en base a su textura.

El algoritmo Local Binary Pattern (LBP) fue originalmente pensado para imágenes monoespectrales, se ocupó el principio de esta primera versión y se trabajó con los 8 vecinos más cercanos de un pixel; para esta nueva versión que propusimos se trabajó con el vector de máximas diferencias que resulta de una imagen multiespectral, el resultado es otra imagen multiespectral que muestra el realce en los bordes de la imagen además de la textura contenida en la misma.

La modificación al algoritmo LBP que realizamos muestra una mejora en los bordes de las imágenes, y esto nos ayuda a analizar la textura de la imagen, ya que la densidad de bordes nos representa la textura de la imagen.

Durante el desarrollo de este trabajo nos auxiliamos de dos herramientas muy poderosas la Descomposición en Componentes Principales y el operador Divergencia.

La descomposición en componentes principales nos ayuda a eliminar redundancia, además de eliminar el ruido y presentarnos en la primera componente la mayor información significativa de la imagen multiespectral, esto nos hizo tener una imagen multiespectral de muy buena calidad al momento de aplicar la modificación que realizamos del algoritmo Lbp

El operador divergencia es una herramienta poderosa para evaluar la textura de una imagen, nos ayuda a clasificar la textura de la imagen en suave para zonas en donde la densidad en las líneas de flujo en un campo vectorial es igual a cero y rugosa para zonas en donde la densidad de las líneas de flujo sufren cambios significativamente grandes.

Podemos dividir los resultados que obtuvimos en dos partes, una es el análisis de la modificación del algoritmo Lbp, en esta las imágenes que se analizaron de la zona del Valle de México mostraron una textura fina y/o suave pero con diferentes escalas en el gris, de manera general gris claro para infraestructura como construcciones, gris en tono medio para carreteras y gris en tono oscuro para cuerpos de agua. El algoritmo presentara mejores resultados para imágenes con una mayor resolución; la segunda es como nos ayuda el operador divergencia a la modificación que realizamos al algoritmo Lbp, en las imágenes a las que ya se les había aplicado la modificación del algoritmo Lbp observamos que los bordes se observan mejor que en aquellas en las que no se les aplicó el algoritmo Lbp.

Realizar un análisis de textura ayuda a los profesionales de la percepción remota como son geógrafos, geólogos, vulcanólogos, etc., para muchas aplicaciones como lo son monitoreo de deforestación, análisis de crecimiento urbano en ciudades, monitoreo de contaminación, entre muchas otras. Con la realización de la modificación al algoritmo LBP presentamos otra opción para el estudio de la textura en imágenes multiespectrales, además que comprobamos que con la ayuda del operador divergencia se pueden presentar mejores resultados en cuanto a bordes en la imágenes.

Anexos

A. Plataformas Orbitales

A.1 Plataformas geoestacionarias

A.1.1 Goes 9



Ilustración 98 Imagen Goes

Primera imagen visible obtenida del satélite GOES – 9 (Junio 12, 1995 17:45 GMT)

Ficha Técnica

Fecha de lanzamiento: 3 de Mayo de 1995

Altura promedio: 35 788 Km

Periodo de la órbita: 24 horas

Periodo de cobertura:

Bandas 1,3, 4: 30 minutos

Bandas 2,5: 60 minutos

Tamaño del Campo Instantáneo de

Vista: Banda 1: 1 x 1 km²

Bandas 2, 4,5: 4 x 4 km²

Banda 3: 8 x 8 km²

Tamaño de la Imagen:

Casi un hemisferio terrestre

Bandas espectrales:

1) 0.52 – 0.75 μm

2) 3.78 – 4.03

3) 6.50 – 7.00

4) 10.2 – 11.2

5) 11.5 – 12.5

Los satélites GOES (Geostationary Operational Environmental Satellite) forman parte de un sistema de satélites hemisféricos que fue comisionado por el gobierno de los EE.UU. con la finalidad de monitorear y prevenir eventos climáticos.

La NASA se ocupa del diseño, desarrollo y lanzamiento de estos satélites y la NOAA es la propietaria y operadora.

Se programaron cinco satélites a partir de los cuales se cubriría completamente la Tierra. La primera generación de la serie GOES empezó a operar a mediados de la década de los 70's.

Los satélites GOES son llamados satélites meteorológicos, ya que sus principales aplicaciones se encuentran en ese campo. Los satélites GOES cubren la Tierra según dos sectores: Este y Oeste.

A.2. Plataformas polares

A.2.1 Aster



Ilustración 99 Imagen Aster

Imagen de Dagze Co

Ficha Técnica

Fecha de lanzamiento: Diciembre, 1999

Altura promedio: 717 Km

Periodo de la órbita: 99 minutos

Periodo de cobertura: 16 días

Tamaño del Campo Instantáneo de

Vista: Bandas 1,2,3: 15 x 15 m²

Bandas 4 - 9: 30 x 30 m²

Bandas 10 - 14: 90 x 90 m²

Tamaño de la Imagen:

60 x 60 m²

Bandas espectrales:

- 1) 0.52 – 0.60 μm
- 2) 0.63 – 0.69
- 3) 0.76 – 0.86
- 4) 1.60 – 1.70
- 5) 2.145 – 2.185
- 6) 2.185 – 2.225
- 7) 2.235 – 2.285
- 8) 2.295 – 2.365
- 9) 2.360 – 2.430
- 10) 8.125 – 8.475
- 11) 8.475 – 8.825
- 12) 8.925 – 9.275
- 13) 10.25 – 10.95
- 14) 10.95 – 11.65

ASTER son las siglas en inglés del Radiómetro Espacial Avanzado de Reflexión de Emisión Termal, es un instrumento del visor que se encuentra a bordo del Terra, un satélite lanzado en diciembre de 1999 que forma parte del Sistema de Observación de la Tierra (EOS) de la NASA. ASTER es un esfuerzo conjunto entre la NASA y el Ministerio de Economía, Comercio e Industria (METI) de Japón y el Centro de Análisis de Datos de los Sensores Remotos de la Tierra (ERSDAC).

Contiene 14 bandas: Infrarroja termal (TIR) con 5 bandas a 90 metros de resolución, Infrarroja de Onda Corta (SWIR) con 6 bandas a 30 metros de resolución e Infrarroja Visible/Cercana (VNIR) con 4 bandas a 15 metros de resolución. Las imágenes ASTER son utilizadas para la interpretación geológica y ambiental, para obtener mapas de temperatura detallados de la superficie terrestre, emisiones, reflectancia y elevación.

A.2.2 Landsat 7



Ilustración 100 Imagen Landsat

Imagen del Monte Fuji y el Lago Fuji , Japón, Octubre de 2002

Ficha Técnica

Fecha de lanzamiento: Abril, 1999

Altura promedio: 705 Km

Periodo de la órbita: 99 minutos

Periodo de cobertura: 16 días

Tamaño del Campo Instantáneo de Vista:

Banda 1-5, 7: 30 x 30 m²

Bandas 6: 60 x 60 m²

Banda 8: 60 x 60 m²

Tamaño de la Imagen:

185 x 185 km²

Bandas espectrales:

- 1) 0.45 – 0.52 μm
- 2) 0.53 – 0.60
- 3) 0.63 – 0.69
- 4) 0.76 – 0.90
- 5) 1.55 – 1.75
- 6) 10.4 – 12.5
- 7) 2.09 – 2.35

Landsat es una serie de satélites construidos y puestos en órbita por EE.UU. para la observación en alta resolución de la superficie terrestre.

La meta fundamental de Landsat 7 es la de actualizar la base de datos de imágenes de todo el planeta. Aunque el programa Landsat es controlado por la NASA, las imágenes recibidas por el Landsat 7 son procesadas por la Agencia Geológica de los Estados Unidos (USGS por sus siglas en inglés)

Landsat 7 fue diseñado para que su vida útil fuese de 5 años y tiene la capacidad de recolectar y transmitir hasta 532 imágenes por día

A.2.3 Spot 5



Ilustración 101 Imagen Spot

Imagen de Copenhague, Dinamarca obtenida en octubre de 2003 resolución de 2.5 m

Ficha Técnica

Fecha de lanzamiento: Mayo, 2002

Altura promedio: 832 Km

Periodo de la órbita: 101 minutos

Periodo de cobertura: 26 días

Tamaño del Campo Instantáneo de Vista:

Pancromático	2.5 x 2.5 m ²
Multiespectral	10 x 10 m ²

Tamaño de la Imagen:

60 x 60 km ²	ó
120 x 120 km ²	

Bandas espectrales:

- 1) 0.50 – 0.59 μm
- 2) 0.61 – 0.68
- 3) 0.79 – 0.89
- 4) 1.58 – 1.75

SPOT 5 es el quinto satélite de la serie SPOT, puesto en órbita por el lanzador Ariane 4, diseñado por el CNES 8 (Centre National d'Etudes Spatiales), fue implementado por Francia en colaboración con Suecia y Bélgica.

La serie SPOT se encuentra en operación desde 1986, fecha del lanzamiento del primer satélite SPOT. SPOT 2 se le unió en órbita en enero de 1990, seguido por SPOT 3 en septiembre de 1993, SPOT 4 en marzo de 1998 y SPOT 5 en mayo de 2002.

Las imágenes de SPOT 5 tienen aplicaciones en la cartografía, la defensa, la agricultura, las telecomunicaciones, la planificación urbana, la gestión de riesgos entre otras áreas. Proveen una visión detallada del paisaje, ya que sus características son la alta resolución, la estereoscopia y la repetitividad, es decir, es posible adquirir diariamente una imagen de cualquier lugar del planeta.

A.2.4 Quickbird 2



Ilustración 102 Imagen Quickbird

Imagen de las pirámides de Egipto tomada en Febrero de 2002

Ficha Técnica

Fecha de lanzamiento: Octubre, 2001

Altura promedio: 450 Km

Periodo de la órbita: 93 minutos

Periodo de cobertura: 3.5 días

Tamaño del Campo Instantáneo de Vista:

Pancromático 0.61 x 0.61

m²

Multiespectral 2.44 x 2.44

m²

Tamaño de la Imagen:
16.5 x 16.5 km²

Bandas espectrales:

- 1) 0.45 – 0.52 μm
- 2) 0.52 – 0.60
- 3) 0.63 – 0.69
- 4) 0.76 – 0.90
- 5) 0.45 – 0.90

El QuickBird es un satélite comercial de teledetección que fue puesto en órbita el 18 de octubre de 2001 con un cohete Delta II, fabricado por Boeing desde la Base de la Fuerza Aérea de Vandenberg, en California. QuickBird es operado por Digital Globe quien también es su propietario.

En realidad QuickBird no fue el primero, sino que tuvo un hermano mayor que fue destruido al no alcanzar el cohete la órbita deseada y la misión fracasó, aumentando considerablemente los costos y dejando a Digital Globe en un punto muy incierto e inseguro, pero el segundo intento fue fructuoso y satisfactorio. En esta ocasión, el cohete Delta-7320-10 (288), colocó en la trayectoria esperada al QuickBird-2, construido sobre una plataforma BCP2000, de la empresa Bell.

Con este satélite se obtienen imágenes de muy alta resolución, éstas pueden ser aplicadas a industrias, incluyendo la de exploración y producción de petróleo y gas, ingeniería de la construcción y estudios ambientales.

A.2.5 Ikonos 2



Ilustración 103 Imagen Ikonos

Imagen de Jinhua China, obtenida el 30 de abril de 2003

Ficha Técnica

Fecha de lanzamiento:
Septiembre, 1999

Altura promedio: 681 Km

Periodo de la órbita: 98 minutos

Periodo de cobertura: 5 días

Tamaño del Campo Instantáneo de
Vista: Bandas 1-4: $4 \times 4 \text{ m}^2$
Bandas 5: $1 \times 1 \text{ m}^2$

Tamaño de la Imagen:
 $11 \times 13 \text{ km}^2$

Bandas espectrales:

- 6) $0.45 - 0.52 \mu\text{m}$
- 7) $0.52 - 0.60$
- 8) $0.63 - 0.69$
- 9) $0.76 - 0.90$
- 10) $0.45 - 0.90$

IKONOS es un satélite comercial de teledetección, siendo el primero en recoger imágenes de alta resolución. El IKONOS-1 estaba proyectado para ser colocado en órbita en 1998, pero su lanzamiento falló. IKONOS-2 estaba preparado para que su puesta en órbita se llevara a cabo en el año 2000 y debido al fallo de IKONOS-1, se le nombro simplemente IKONOS y se lanzó el 24 de Septiembre de 1999 desde California, Estados Unidos.

Las imágenes IKONOS se proponen a diferentes niveles de procesamiento geométrico:

- los productos Geo: se rectifican según un elipsoide específico y dentro de una proyección dada, a elección del usuario.

- los productos Ortho: se corrigen de las distorsiones debidas al relieve gracias a un MDE (Modelo Digital de Elevaciones) emitidas por IKONOS o provistas por el usuario. Sólo los productos Precisión y Precisión Plus utilizan puntos de control

B. El sistema Landsat

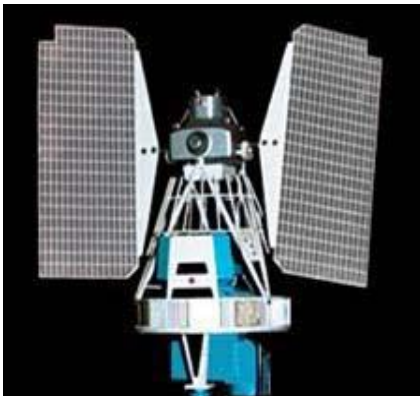


Ilustración 104 Satélite Landsat

Aunque muchos de los sistemas satelitales han sido usados para el monitoreo de la superficie terrestre, éstos no estaban optimizados para mapear detalladamente la superficie terrestre.

El primer satélite diseñado específicamente para monitorear la superficie de la tierra fue el Landsat – 1 el cual fue lanzado por la NASA en 1972. Inicialmente conocido como ERTS-1 (Earth Resources TEchnology Satellite), Landsat fue

diseñado como un experimento para probar que tan fácil era coleccionar datos multiespectrales de la tierra desde una plataforma satelital.

Desde aquel tiempo este fue un programa altamente exitoso que colecto un gran número de datos alrededor del mundo.

Originalmente manejado por NASA, la responsabilidad del programa Landsat fue transferido a NOAA en 1983. EN 1985, el programa empezó a ser comercializado, proveyendo de datos y aplicaciones a los usuarios.

El éxito de Landsat se debió a varios factores, incluyendo: una combinación de sensores con bandas espectrales, resolución espacial funcional, una buena área de cobertura.

Todos los satélites Landsat poseen una órbita polar síncrona con el sol.

Los primeros tres satélites (Landsat 1-3) estaban a altitudes de 900 km con periodos de 18 días mientras los últimos satélites estaban a una altitud de 700 km y contaban con periodos de 16 días.

Varios sensores han ido en la serie de los satélites Landsat, entre ellos el sistema RBV(Return Beam Vidicon) el sistema MSS(MultiSpectral Scanner), y el TM (Thematic Mapper). El sensor más popular fue el MSS y después el TM. Cada uno de estos sensores obtenía datos de escenas de 185 km x 185 km.

El sensor MSS capta radiación electromagnética de la superficie de la tierra en cuatro bandas espectrales. Cada banda tiene una resolución de 60 m x 80 m aproximadamente y una resolución radiométrica de 6 bits o 64 números digitales. El sensado se realiza con un dispositivo de escaneo en línea usando un espejo oscilador. Seis líneas de escaneo están colocadas con cada paso de oeste a este del espejo.

La siguiente tabla contiene los rangos de longitudes de onda para el MSS.

Bandas MSS

Canal		Longitud de onda (μm)
Landsat 1,2,3	Landsat 4,5	
MSS 4	MSS 1	0.5 – 0.6 (verde)
MSS 5	MSS 2	0.6 – 0.7 (rojo)
MSS 6	MSS 3	0.7 – 0.8 (infrarrojo cercano)
MSS 7	MSS 4	0.8 – 1.1 (infrarrojo cercano)

La obtención de datos del MSS cesó en 1992, cuando se empezó a usar el sensor TM. El sensor TM proveía de algunas mejoras sobre el sensor MSS incluyendo: alta resolución espacial y radiométrica, bandas espectrales más finas, siete contra cuatro bandas espectrales, y un incremento en el número de detectores por banda. 16 líneas de escaneo captaban simultáneamente para cada banda espectral no térmica y cuatro para la banda termal. Usaba un espejo oscilador con escáneres para ambos pasos de ida (este a oeste) y de reversa (oeste a este). Esta diferencia mejoraba la integridad de los datos. La resolución espacial del sensor TM es de 30m para todas las bandas excepto para la banda infrarroja la cual es de 120m. Todos los canales están sobre el rango de los 256 números digitales (8 bits).

La siguiente es una tabla que muestra la resolución espectral de las bandas individuales del sensor TM y algunas de las aplicaciones para cada una de ellas.

Bandas TM

Canal	Longitud de Onda (μm)	Aplicaciones
TM 1	0.45 – 0.52 (azul)	Discriminación de vegetación
TM 2	0.52 – 0.60 (verde)	Mapeado de vegetación
TM 3	0.63 – 0.69 (rojo)	Discriminación de especies vegetales
TM 4	0.76- 0.90 (infrarrojo cercano)	Identificación de plantas y vegetación, delineación de cuerpos de agua.
TM 5	1.55 – 1.75 (onda corta IR)	Sensible a la humedad, discriminación de áreas cubiertas de hielo y cenizas.
TM 6	10.4 – 12.5 (termal IR)	Mapeado termal
TM 7	2.08 – 2.35 (onda corta IR)	Discriminación de rocas, sensible al contenido de humedad en la vegetación

Los datos obtenidos de los sensores TM y MSS son usados para una amplia variedad de aplicaciones, incluyendo el manejo de recursos, el mapeado, el monitoreo ambiental.

La galería Canadiense incluye cerca de 350 000 escenas para el MSS y cerca de 200 000 para el sensor TM.

Bibliografía

- BAJCSY, RUZENA** (1973). *Computer description of textured surfaces*. Int. Joint Conf. Artificial Intelligence, 572-579
- CROSS G y JAIN A.** (1983) *Transactions on Pattern Analysis and Machine Intelligence* IEEE
- ESCALANTE B.**, (2007) Apuntes de la asignatura *Procesamiento Digital de Imágenes*, impartida en el Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas (IIMAS) de la UNAM durante el Semestre 2007-2, México.
- GONZALEZ R y WOODS R.**, (2002) *Digital Image Processing*, 2ª Edición, Prentice Hall, Estados Unidos de América.
- GOUSTIAS J** (2001)., Notas de la Clase *Image Processing and Análisis*, de la Universidad Jonh Hopkins, Estados Unidos de América
- HOFFMAN R. y MARKMAN A.**, (2001) *Interpreting Remote Sensing Imagery: Human Factors*, Lewis, Estados Unidos de América.
- JAIN A.**, (1989) *Fundamentals of Digital Image Processing*, 1ª Edición, Prentice Hall, Estados Unidos de América.
- LILLESAND T y KIEFER R.**,(2000) *Remote Sensing and Image Interpretation*, 4a. Edición, Wiley, Estados Unidos de América.
- LIRA J.**, (1987), *La percepción Remota: Nuestros ojos desde el espacio*, Colección la Ciencia desde México, vol. 33, Fondo de Cultura Económica, México.
- LIRA, J. y FRULLA, L.** (1998) *An automated region growing algorithm for segmentation of texture regions in SAR images*. International Journal of Remote Sensing
- LIRA J.**, (2002) *Introducción al tratamiento digital de imágenes*, Colección Ciencia de la Computación, FCE-IPN-UNAM, 1ª Edición, México.
- LIRA J., y RODRIGUEZ A.**, (2006) *A divergence operator to quantify texture from multi-spectral satellite images*, International Journal of Remote Sensing, Volume 27, Issue 13.
- LIRA J.**, (2007) Apuntes de la asignatura *Introducción a la Percepción remota*, impartida en el Instituto de Geofísica de la UNAM durante el Semestre 2007-2, México.

LUCIEER A, et. al., (2005) *Multivariate texture-based segmentation of remotely sensed imagery for extraction of objects and their uncertainty* International Journal of Remote Sensing Vol. 26 No. 14, 2917-2936.

MÄENPÄÄ T y PIETIKÄINEN (2004) *Texture Analysis with Local Binary Patterns*, Departamento de Ingeniería eléctrica e Información, Universidad de Oulu, Finlandia.

MARTINEZ E., (2005) *Apuntes de la Asignatura Visión por Computadora*, impartida en el Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas (IIMAS) de la UNAM, México

MARQUEZ J., (2007) *Apuntes del Curso Procesamiento Digital de Imágenes* Facultad de Ciencias UNAM, México

OJALA T, et. al. (1996) *Texture Discrimination with Multidimensional Distributions of Signed Gray Level Differences*, Unidad de Visión y Procesamiento, Universidad de Oulu, Finlandia.

PRATT W.,(2001) *Digital Image Processing*, 3a. Edición, Wiley - Interscience, Estados Unidos de América.

QUATROCHI D. y GOODCHILD M. (1997) *Scale in Remote Sensing and Gis*, 1a. Edición, Estados Unidos de América.

RENZ A., (1999) *Remote Sensing for the Earth Sciences: Manual of Remote Sensing*. Vol.3, Wiley, Estados Unidos de América.

RICHARDS J. y XIUPING J., (1999) *Remote Sensing Digital Análisis*, 1ª. Edición, Springer, Alemania

RODRIGUEZ A. (2003), *Filtro dinámico Newtoniano*, Tesis de Licenciatura, Universidad Nacional Autónoma de México, Facultad de Ingeniería, México.

RODRIGUEZ A. (2008), *Un operador Divergencia para cuantificar textura a partir de imágenes multiespectrales de satélite*, Tesis de Maestría, Universidad Nacional Autónoma de México, Facultad de Ingeniería, México.

RUSS, J., (2002) *The imagen Processing Handbook*, 4ª Edición, Crc. Press, Estados Unidos de América.

SABINS F., (2000) *Remote Sensing: Principles and Interpretation*, 3a. Edición, Freeman, Estados Unidos de América.

SIGAL B. y GILLESPIE A., (1980) *Remote Sensing in Geology*, 1ª Edición, Wiley. Estados Unidos de América.

TAMURA, H., MORI, S., YAMAEAKI, T., (1978). *Textural features corresponding to visual perception.* IEEE Trans. on Systems, Man and Cybernetics, SMC 8 (6), 460-472.