

Fundamentos del procesamiento digital de imágenes

Mehl, H.¹; Peinado, O.^{2,3}

¹Deutsches Zentrum für Luft- und Raumfahrt (DLR) - Alemania

²Arbeitsgruppe Fernerkundung (AGF) Ludwig Maximilians Universität München - Alemania

³Comisión Nacional de Actividades Espaciales (CONAE) Mendoza – Argentina

Introducción

Todos los elementos de la corteza terrestre, rocas, suelos, vegetación y agua, así como los objetos que la recubren, absorben, reflejan o emiten una cantidad de energía que depende de la longitud de onda, de la intensidad y tipo de radiación incidente, así como también de las características de absorción de los objetos y de la orientación de estos respecto al sol o fuente de iluminación

El ojo humano tiene ciertas limitaciones en lo que a la percepción se refiere. Por un lado sólo podemos captar un cierto tipo de energía dentro de lo que llamamos el rango visible, el calor y las microondas que son otras formas de energía, no podemos percibirlos directamente, y por otro lado no podemos ver más allá de lo que nuestra estatura o desde plataformas naturales podemos ver, de todos modos siempre tenemos una visión oblicua y no podemos captar grandes extensiones como por ejemplo grandes incendios o inundaciones, por eso necesitamos “ojos artificiales” o sensores que nos permiten solucionar estos inconvenientes. Se han creado sensores que pueden captar otros tipos de energía como la ultravioleta, infrarroja o microondas desde plataformas a una cierta altitud, desde las cuales se puede tener una visión vertical y panorámica, lo cual es una gran ventaja. El objetivo básico del sensoramiento remoto o teledetección es obtener información a distancia de los objetos que vemos en la superficie de la tierra.

El término sensoramiento remoto proviene de la traducción del inglés remote sensing, pero en la lengua castellana usamos normalmente las palabras teledetección o percepción remota o simplemente sensores remotos. La palabra “sensoramiento remoto” envuelve todo el proceso de adquirir la información mientras que el procesamiento digital de imágenes implica un conjunto de métodos para obtener esa información. Existen también otros métodos, por ejemplo las imágenes pueden ser tratadas con métodos fotogramétricos para su interpretación visual o los datos de reflexión espectral pueden ser comparados con espectros ya conocidos de rocas o suelos. Una de las razones por la cual el procesamiento digital de imágenes se ha transformado en una herramienta tan importante es que los datos son almacenados en forma digital, lo cual permite una manipulación mucho más amplia de los mismos que no son posibles por medios fotográficos.

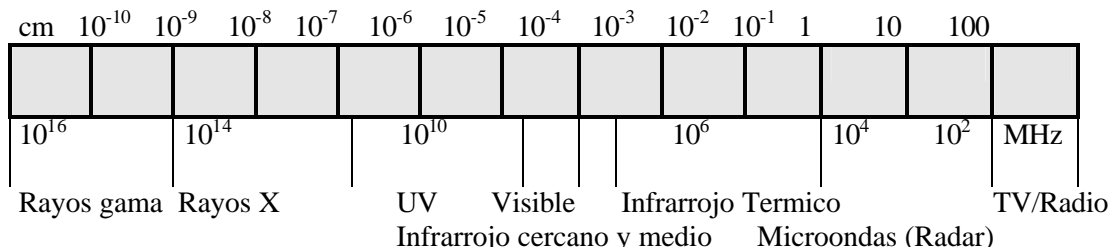
Principios físicos de la teledetección

Para que haya percepción remota debe existir algún tipo de interacción entre los objetos observados y el receptor. Por ejemplo nuestros ojos pueden

ver un árbol porque son capaces de descifrar la información que éste les envía, pero la energía no proviene de él mismo sino que es la reflejada de la luz solar.

Hay tres formas de adquirir información por medio de los sensores remotos, por reflexión, emisión y por emisión-reflexión. La primera es la más importante en teledetección y es la energía que reflejan los objetos proveniente de la luz solar. La segunda es la energía emitida por los propios objetos, como una ciudad iluminada por la noche y la tercera es aquella que el mismo sensor emite y luego capta la reflexión como es el caso de los sensores activos (radar). En todos los casos lo que capta el sensor es una forma de energía electromagnética. Esta energía se desplaza siguiendo un movimiento armónico a la velocidad de la luz y se la mide o reconoce por dos parámetros inversamente proporcionales entre sí, frecuencia y longitud de onda. Cuando la longitud de onda es mayor (o la frecuencia es menor) el contenido de energía de la onda electromagnética es menor, por eso es más difícil de detectar, en cambio las longitudes de onda corta son más fáciles de detectar con sistemas comunes por su mayor contenido energético. Si bien las longitudes de onda varían en forma continua se suelen establecer una serie de bandas donde el comportamiento de las ondas es similar, esto es lo que se conoce como el espectro electromagnético. Allí se grafican desde las longitudes de onda más cortas (rayos gamma, rayos X) hasta las kilométricas (Telecomunicaciones). Para medir las más cortas se usan los micrones $1 \mu\text{m} = 10^{-6} \text{ m}$ y las más largas cm o m, otra forma de medirlas o designarlas es por la frecuencia en GHz, donde $1 \text{ GHz} = 10^8 \text{ Hz}$.

Gráfico espectro electromagnético



Desde el punto de vista de la teledetección conviene destacar la serie de bandas más frecuentemente empleadas que son:

Espectro visible (0,4 a 0,7 μm), se denomina así porque se trata de la radiación electromagnética que pueden percibir nuestros ojos, coincidiendo con las longitudes de onda donde es la máxima radiación solar. Suelen distinguirse tres bandas elementales, que se denominan azul (0,4 a 0,5 μm), verde (0,5 a 0,6 μm) y rojo (0,6 a 0,7 μm) en razón de los colores elementales asociados a estas longitudes de ondas.

Infrarrojo cercano (0,7 a 1,3 μm) a veces también se lo llama infrarrojo reflejado o fotográfico, porque se puede captar con películas especiales.

Resulta de especial importancia para discriminar masas vegetales o concentraciones de humedad.

Infrarrojo medio (1,3 a 8 μm) en donde se entremezclan los procesos de reflexión de la luz solar y la emisión de la superficie terrestre.

Infrarrojo lejano (8 a 14 μm) es la porción emisiva del espectro terrestre

Microondas (a partir de 1 mm) es de gran interés porque es una energía transparente a la superficie nubosa.

Los sensores se pueden clasificar según el tipo de energía que captan en pasivos y activos. Los sensores pasivos se limitan a recoger la energía electromagnética procedente de las cubiertas terrestres, ya sea esta reflejada de los rayos solares o emitida en virtud de su propia temperatura.

Los sensores pasivos pueden ser a su vez sensores fotográficos o de barrido. Los rastreadores de barrido (scanners) barren la superficie terrestre en franjas a lo largo del recorrido, por ejemplo la serie Landsat dispone de un espejo que oscila en forma perpendicular a la trayectoria y una serie de sensores que captan la radiación recibida en el espejo y la convierten en información digital para ser enviada a tierra.

La señal recibida se descompone en varias longitudes de onda, cada una de las cuales se envía a un grupo de detectores sensibles a ellas, por eso se habla de scanners multiespectrales porque detectan la misma superficie del terreno en distintas bandas espectrales. Existen otros sensores pasivos denominados de empuje (pushbroom) que son una cadena de detectores que captan toda una línea de terreno por vez, son los CCD (Charge Couple Devices) detectores sólidos que directamente digitalizan la radiación que reciben. Los sensores activos por otro lado son aquellos que emiten su propia energía y reciben la energía reflejada por la superficie de la tierra, como es el caso del RADAR, cuya principal ventaja es la independencia de la luz solar y la posibilidad de penetrar las cubiertas de nubes, la desventaja es que trabajan en un solo rango del espectro, el de las microondas, y que tienen mucho ruido que es propio del sistema (speckle). A continuación vamos a describir brevemente las características del sensor Thematic Mapper de Landsat que es el más usado para cartografía temática.

Usa la técnica del espejo de barrido, tiene 100 sensores, y codifica en 256 niveles de gris. Tiene una resolución radiométrica de 8 bits, una espacial de 30 m y espectral de 7 bandas y una temporal de 15 días volando a una altura de 917 km. Con esta configuración permite la detección de aguas turbias o contaminadas (canal azul), exploración minera y contenido de humedad en plantas y suelo (infrarrojo medio), detección de focos de calor (infrarrojo medio y térmico) y en las visibles (3,4 y 5) áreas urbanas y vías de comunicación, calidad del agua. Las longitudes de onda y resoluciones de las bandas son las siguientes

Banda	Longitud de onda
1	(0,45 0,52 μm)
2	(0,52 0,60 μm)
3	(0,63 0,69 μm)
4	(0,76-0,90 μm)
5	(1,55 1,75 μm)
6	(10,40 - 12,50 μm)
7	(2,08 - 2,35 μm)
	Resolución espacial
1-5 y 7	30 m

6 120m
 8 (pan) 15 m

Los sensores de los satélites proveen datos y para extraer información de ellos deben establecerse relaciones con otros grupos de datos de sensores remotos o de campo o de pruebas de laboratorio. En la figura 1 se muestra una relación entre los datos de sensores remotos y la extracción de información a partir de éstos. El „mundo real“ es donde vivimos y cómo lo perciben nuestros sentidos. Datos, como datos de imagen propiamente dichos, en el mundo real son adquiridos por algún sistema de sensoramiento remoto, como son nuestros ojos. Como sistema de teledetección el ojo humano recibe un haz de energía proveniente de los objetos exteriores, por reflejo de la luz solar o artificial, forma una imagen que se transmite al cerebro que la percibe, e interpreta así, el mundo real que lo rodea.

Los datos de sensores remotos pueden ser reproducidos en algo que llamamos “espacio de imágenes” (image space) como una función de las características físicas de los objetos y el sensor usado para captarlos. También se pueden extraer características propias de los datos como reflectancia o textura y a estos podemos ubicarlos en un espacio de características (feature space) en este proceso es posible obtener „clases espectrales“ que pueden o no coincidir con las clases temáticas definidas por el ojo humano en el mundo real como por ejemplo campos de arroz o papas. Estos conceptos de image space y feature space son muy utilizados por los nuevos software comerciales de procesamiento digital de imágenes, por lo que es bueno conocer las diferencias y a qué se refiere cada uno de ellos para su mejor entendimiento.

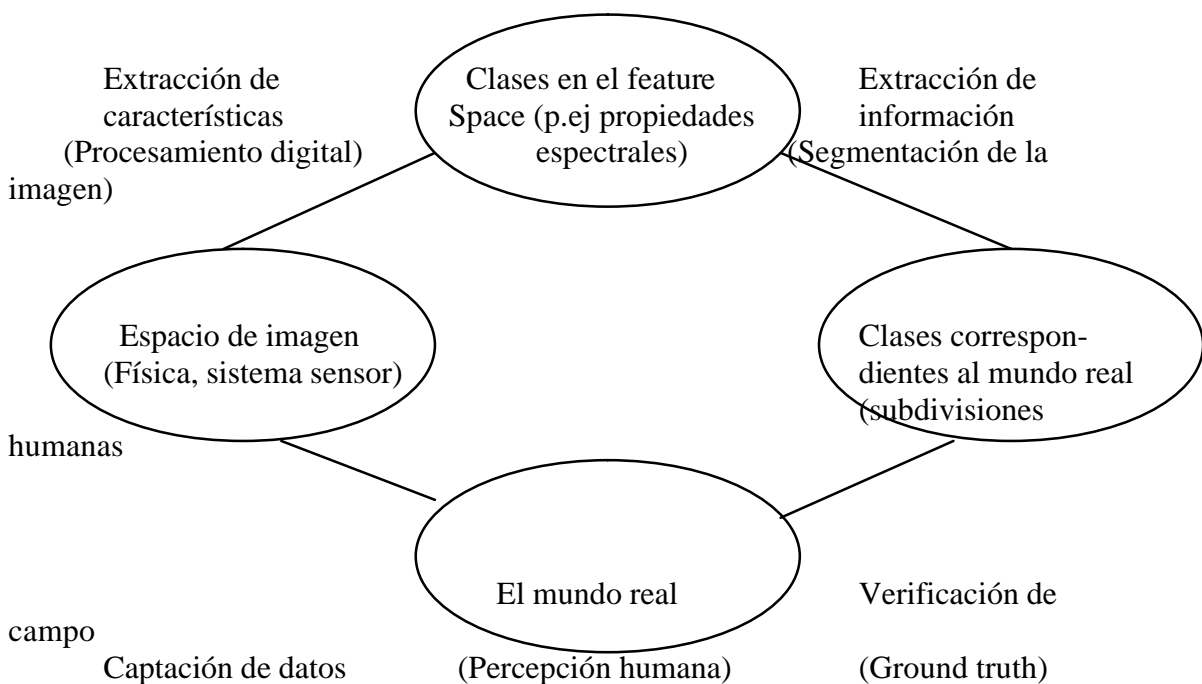


Figura 1: Relación entre los datos de los sensores remotos y la extracción de información

El próximo paso, y el más crítico es la extracción de la información y consiste en la clasificación de los datos satelitales de acuerdo a la definición de nuestro ojo humano. Esto resulta ser bastante dificultoso e incluso a veces prácticamente imposible de diferenciar cuando nuestras clases temáticas no están bien separadas en nuestro feature space, además los resultados de la clasificación deben ser comparados con la verdad de campo, en el mundo real y normalmente hay que corregir y rehacer el proceso varias veces para tener un buen grado de veracidad entre los resultados obtenidos del procesamiento digital y la realidad.

Interpretación visual o digital de imágenes

De los datos de sensores remotos se puede obtener información mediante 2 formas, ya sea por medio de la interpretación visual o de la clasificación digital. La primera es la más difundida ya sea tanto en imágenes satelitales como aéreas. La fotogeología como ciencia que se basa en la interpretación visual de pares estereoscópicos de fotos aéreas se remonta a fines de 1930 cuando buscando petróleo, los geólogos usaron este método para buscar potenciales estructuras bajo la densa selva en Guinea.

Actualmente la interpretación visual de imágenes sigue siendo el aspecto más usado de los datos de sensores remotos.

La clasificación digital funciona bien si la información requerida está bien definida en los datos de teledetección, como por ejemplo en la clasificación de diferentes tipos de vegetación o suelos. Pero el procesamiento digital de imágenes no se limita sólo a la clasificación digital, también se usa para ensanche de contraste o diversos filtrados, para destacar elementos en la imagen para su posterior interpretación visual o para hacer mosaicos de imágenes. La gran ventaja del procesamiento digital reside en que la mayoría de los datos de sensores remotos vienen en forma digital y por lo tanto el procesamiento es mucho más sencillo y rápido, además de que los costos a una escala media (1:200.000) son sensiblemente inferiores por unidad de área a la fotografía tradicional. Esto siempre y cuando se disponga del equipo de computación adecuado, y dependiendo de la duración del proyecto y del grado de adiestramiento del personal, sino resulta más conveniente el método tradicional.

Imágenes digitales y su procesamiento

Una imagen en forma genérica es una representación en 2 dimensiones de un objeto de 2 o 3 dimensiones. Esta representación puede ser definida por colores o diferentes niveles de gris. Una imagen puede ser obtenida por muchos medios, ya sea un sensor electrónico o radar mientras que fotografías se refieren a una imagen que fue obtenida sólo mediante medios fotográficos sobre un film, hay que tener en cuenta la diferencia entre estos 2 términos. Por lo tanto los datos de satélites como Landsat, SPOT, ERS o MOMS son

imágenes mientras que los datos tomados por cámaras del Apolo, o la cámara métrica del Spacelab o la KFA-1000 rusa son fotos espaciales.

Las imágenes pueden ser divididas en tres tipos:

Análogas
Digitales
Vectoriales

Las primeras son las que más conocemos, porque que todas las imágenes fotográficas, ya sean en color o blanco y negro son analógicas. El sistema fotográfico consistente de lentes, filtros, películas y el sistema de revelado y copiado, lo que hace es transformar la reflexión espectral de las superficies de los objetos en valores de gris o color de acuerdo a la intensidad de la radiación recibida. El mayor problema de las imágenes analógicas resulta en el proceso de copiado, ya que se produce una degradación de la información original y no hay forma de conservarla y mientras más copias se hacen más se degrada. Esa es una gran ventaja de las imágenes digitales que al ser datos digitales pueden ser copiadas en computadoras todas las veces que sea necesario y siempre se mantienen en cada copia la información original.

El sensor explora la superficie terrestre adquiriendo a intervalos regulares la radiación proveniente de la misma, al barrer la superficie el sensor traduce la radiancia media de una parcela correspondiente a un valor de pixel que es la unidad visual más pequeña que aparece en una imagen en un valor numérico entero. Una imagen digital es la digitalización de la energía radiante recibida. La imagen resultante es una matriz discreta de valores numéricos que representan niveles de gris o color, estos puntos representan áreas pequeñas llamados elementos de imagen o píxeles, debido a que estos píxeles representan valores numéricos o valores digitales se los suele llamar DN (del inglés digital numbers).

La matriz de puntos que forma la imagen se parece a un gran tablero de ajedrez, y dado que ésta tiene un cierto orden, que es el barrido de línea tras línea del scanner se las llama imágenes raster (del inglés barrido). Cuando un pixel tiene sólo información sobre blanco o negro hablamos de una imagen de 2 valores, una silueta o una línea por ejemplo. Normalmente un pixel contiene información en un rango de 128 o 256 valores de gris por lo que se puede tener así una imagen monocromática.

Cada banda espectral de un sensor multispectral es una imagen monocromática que podemos ver en tonos de gris o combinar con otras bandas para producir una imagen color o falso color, esto se hace asignando a cada uno de los canales (RGB) de los tubos de rayos catódicos (CRT) de la pantalla una banda en particular. El máximo número de canales que podemos ver simultáneamente en una pantalla son 3 y la combinación de colores dependerá de que bandas espectrales asignemos a los cañones rojo, verde y azul del visor. El tercer tipo de imágenes, las vectoriales, consisten en almacenar no todo el conjunto de píxeles sino puntos o vértices que formen una línea que sea representativa de algo, por ejemplo bordes del resultado de una clasificación,

caminos, autopistas, vías de ferrocarril, separación entre cultivos, etc. Este tipo de imágenes es especialmente útil cuando trabajamos con sistemas de información geográfica (SIG). De aquí en adelante trabajaremos sólo con imágenes raster que son las que se usan para el procesamiento digital.

Estadística de una imagen

Un conocimiento básico de la estructura de los datos es fundamental para manejar las operaciones de procesamiento digital como ensanche de contraste, clasificaciones o correcciones y la estadística provee esa información necesaria para el análisis de los datos. Normalmente una imagen está compuesta por píxeles con niveles de gris que van de 0 a 255, uno de los primeros pasos del procesamiento digital es producir un histograma de esos valores en cada banda espectral. Un histograma es normalmente un gráfico de barras que representa la cantidad de puntos existentes con el mismo valor numérico en una imagen dentro del rango de valores (0 - 255). Este histograma puede considerarse como una distribución discreta donde el peso relativo de cada barra representa la posibilidad de encontrar un píxel con ese valor dentro de la imagen.

Existen diferentes parámetros que sirven para caracterizar la frecuencia de distribución de una imagen , ellos son:

Máximo y mínimo: Son el mayor y el menor valor numérico dentro de la distribución, como ya dijimos en una imagen de 8 bits, son 0 y 255. En la práctica el contraste de los objetos en la superficie es bastante menor a todo el rango disponible.

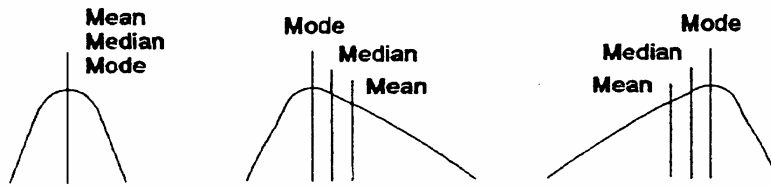
Media: Esta designa, en la mayoría de las aplicaciones la media aritmética de una muestra. La media es muy sensible a los valores extremos.

Mediana: Es el valor digital para el cual el 50% de todos los píxeles en el histograma son mayores y 50% son menores, al contrario de la media ésta no es influenciada por los valores máximos o mínimos.

Moda: Es el valor digital con mayor ocurrencia dentro de la muestra.

Varianza: Es una medida de la diferencia existente entre todos los valores de la muestra y la media, se define matemáticamente como el promedio del cuadrado de esas diferencias.

Desviación estándar: Es la raíz cuadrada de la varianza, es una forma más fácil de interpretar esa medida entre la media y todos los valores de la muestra, sirve para conocer qué forma tiene la curva, por ejemplo en una distribución gauseana, cuánto se separan de la media algunos valores. En el siguiente ejemplo se ven distribuciones simétricas y desplazadas con sus respectivas media, mediana y modas.



Una distribución normal (gausiana) es un caso muy especial porque la mayoría de las veces los datos están desviados hacia algún extremo y dependiendo de las características de los objetos en la imagen pueden una distribución bimodal o multimodal como se ve en la siguiente figura:



Ejemplos de distribución unimodal, bimodal y multimodal

Resolución

En el caso de las imágenes fotográficas tiene un significado especial. La resolución fotográfica se refiere al número de líneas o más precisamente de pares de líneas que pueden ser grabados en un film o un papel. En teledetección el término es mucho más amplio y tiene distintos factores involucrados, la resolución de un sensor depende de todos sus componentes, por ejemplo una mejora en la lente no implica necesariamente una mejora en la resolución sino se la acompaña con una mejor resolución del conversor análogo-digital o las condiciones de exposición, podemos así hablar de una resolución espacial, espectral, radiométrica y temporal. A continuación se explican cada una de ellas

Resolución espacial: Este concepto designa al objeto más pequeño que se puede distinguir en la imagen. Está determinada por el tamaño del pixel medido en metros sobre el terreno, esto depende de la altura del sensor con respecto a la tierra, el ángulo de visión, velocidad de barrido y las características ópticas del mismo.

Por ejemplo las imágenes Landsat TM tienen una resolución espacial de 30x30 m de las bandas 1 a 5 y 7 y de 120x120m en la 6 (térmica) y MOMS-02 tiene 4,5x4,5 m en el modo pancromático, SPOT - HRV de 10x10m mientras que satélites como NOAA-AVHRR el pixel tiene desde 500 a 1100m de lado.

Resolución espectral: Es el número de canales espectrales (y su ancho de banda) que es capaz de captar el sensor. Por ejemplo SPOT tiene una resolución espectral de 3, Landsat de 7. Los nuevos sensores, llamados también espectrómetros o hiperespectrales llegan a tener hasta 256 canales con un ancho de banda muy angosto (unos pocos nm) para poder separar en

forma bastante precisa distintos objetos por su radiación. Se pueden distinguir características lineales inferiores a la resolución nominal del pixel , esto se debe a que el pixel es un promedio de las radiaciones recibidas en esa área, por ej., 30x30 pero si el contraste es suficientemente alto se pueden distinguir como una ruta que cruza un campo claro o un desierto.

Siempre es conveniente que las bandas estén ubicadas en un sector coherente del espectro y de acuerdo al objetivo del sensor, mientras más angostas sean es mejor ya que sino lo que se tiene es un promedio de todas las radiaciones en ese rango.

Resolución radiométrica: Se la llama a veces también resolución dinámica, y se refiere a la cantidad de niveles de gris en que se divide la radiación recibida para ser almacenada y procesada posteriormente. Esto depende del conversor analógico digital usado. Como sabemos la imagen captada por el sensor es una imagen analógica y por ende continua pero las computadoras trabajan con elementos discretos, bytes y bits por lo que esa señal es dividida en una cierta cantidad de niveles de gris, siempre múltiplos de 2 , ya que un bit sólo puede almacenar 2 valores 0 y 1. Así por ejemplo Landsat MSS tiene una resolución espectral de $2^6 = 64$ niveles de gris en el canal 6, y Landsat MSS en las bandas 4 a 7 de $2^7 = 128$ niveles de gris mientras que en Landsat TM es de $2^8 = 256$, esto significa que tenemos una mejor resolución dinámica en el TM y podemos distinguir mejor las pequeñas diferencias de radiación. Hay que tener en cuenta que el ojo humano difícilmente pueda distinguir más de 64 niveles de gris y 200.00 tonalidades, por lo que parece superfluo contar con 256 niveles de gris y 16 millones de tonalidades para una imagen color, pero es con la clasificación digital donde se aprovecha esta altísima resolución.

Resolución temporal: Es la frecuencia de pasadas por una misma área. Esto significa cada cuanto tiempo pasa el satélite por el mismo sector de la tierra, esto depende de las características de la órbita fundamentalmente. Muchas veces esta frecuencia está obstruida por otros factores como por ejemplo cubierta de nubes o nieve en la cercanía de los polos o zonas montañosas. Los datos de Landsat MSS vienen siendo colectados y almacenados regularmente desde 1972. El ciclo de repetición de Landsat-1 hasta Landsat -3 es de 17 días, desde el lanzamiento de Landsat 4 en 1984 el ciclo de repetición se redujo a 15 días para MSS y TM. SPOT permite un ciclo de 3 días usando el máximo ángulo de deflexión del espejo o 26 días en un ciclo normal.

Representación en colores

Para la visualización de los datos de teledetección es mejor tener una representación en colores ya que el ojo humano capta mejor el color que los niveles de gris. El color es una función de la longitud de onda de la luz visible dentro del espectro desde el rojo como la mayor longitud de onda pasando por el naranja, amarillo, verde, azul y violeta como la más corta. El blanco y el gris se pueden obtener mezclando todos los otros colores, mientras que los demás colores se pueden obtener mezclando rojo, azul y verde (RGB, del inglés Red, Green, Blue). Este proceso llamado „mezcla de colores aditiva“ es el principio

de funcionamiento del tubo de rayos catódicos (CRT en inglés) para televisión y computadoras. Existe otro método que es la „mezcla de colores substractiva“ donde los colores complementarios de rojo, verde y azul son cian, magenta y amarillo, que también se pueden combinar para tener la paleta completa de colores incluyendo el negro. Este es el principio de funcionamiento de casi todos las impresoras color y plotters de inyección de tinta, se lo llama sistema YMC. Muchas veces el negro se obtiene con un chorro aparte. Esto podemos verlo resumido en los siguiente gráficos.

Gráfico RGB

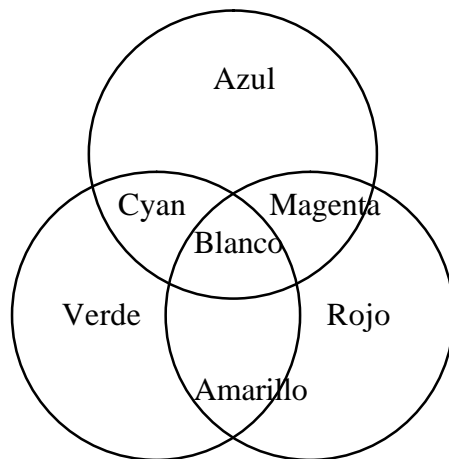
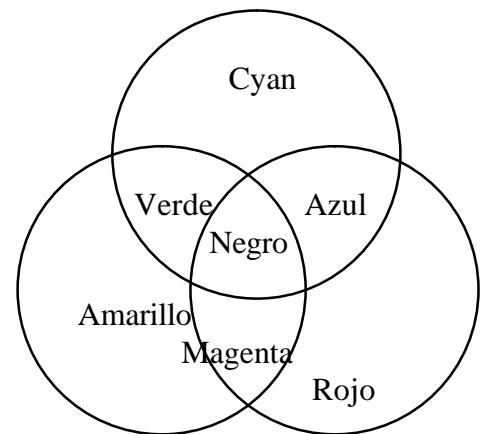


Gráfico YMC



Otra forma de definir el color es usando los parámetros de intensidad, tono y saturación (IHS en inglés, de Intensity, Hue y Saturation). En el procesamiento digital de imágenes es factible hacer la conversión RGB a IHS y viceversa lo cual es muy útil para la diferenciación de pequeñas variaciones de tonos que a veces es casi imposible distinguir en RGB y es fácilmente separable en IHS y que no se puede hacer con otros métodos, como por ejemplo distinguir cultivos con una reflexión espectral muy similar. Los datos de sensores remotos son almacenados normalmente en distintas bandas, y para verlos en colores se hace una combinación de tres bandas, la combinación de cualquiera de estas bandas se lo llama imagen de color compuesta. Una composición de bandas en las cuales se asigna a los cañones rojo, verde y azul del CRT las longitudes de onda correspondiente al código de colores (o sea rojo en rojo, verde en verde y azul en azul) nos da una imagen en color natural o colores verdaderos, usando diferentes bandas o códigos de colores obtenemos imágenes en falso color.

Esto es a veces un poco confuso, ya que normalmente las imágenes en falso color tienen mayor información y es más fácil distinguir objetos que en las imágenes de color natural. Se pueden tener también imágenes color partiendo de una sola banda, y asignando ésta a un canal del CRT y en otro por ejemplo el resultado de una división y al tercero otra operación matemática. Los resultados de una clasificación digital usualmente son imágenes con códigos de color, cada clase se representa por un número entero y se le asigna a un determinado color, la selección de colores depende del tipo de clases, por

ejemplo azul para el agua, verde para cultivos, verde oscuro para bosques, etc. pero puede ser también arbitraria.

Hace unos 20 años atrás el trabajo de procesamiento digital de imágenes era sumamente costoso por los equipos y la necesidad de escribir prácticamente todo el software para procesarlas en complicados sistemas operativos. Con la evolución de las PC y las Workstation, con las constantes mejoras de velocidad de procesamiento y capacidad de almacenamiento así como la baja en los costos hace que cualquier compañía o universidad pueda acceder a estos equipos, además que existen actualmente en el mercado gran cantidad de software de procesamiento de imágenes a costos razonables.

Para la entrada de datos se usan varios medios, los más difundidos son las cintas Exabyte y los CD-ROM que es el medio más barato, y rápido que se usa. Las imágenes vienen grabadas en cuatro formatos básicos:

BIL (Band interleaved by line)

En este formato cada línea corresponde a una determinada banda seguida por otra de la siguiente banda, es decir si tomamos una imagen Landsat TM, la primera línea corresponde a la primera banda, la segunda línea es la primera línea de barrido de la segunda banda, la tercera línea corresponde a la primera línea de la tercera banda y así hasta la séptima, la octava línea corresponde a la segunda línea de barrido de la primera banda, la novena a la segunda de la segunda banda y así sucesivamente. O sea que las bandas están intercaladas línea a línea.

BSQ (Band Sequential)

En este formato cada banda es grabada como un archivo separado uno a continuación del otro independientemente. En los datos BSQ se usa a veces un factor de bloqueo y la información está incluida en headers files al principio de la cinta.

BIP (Band Interleaved by pixel)

En este formato las bandas se intercalan por pixel, es similar al BIL sólo que en vez de líneas se lo hace con pixeles, o sea que cada columna corresponde a una banda determinada. Prácticamente ya no se lo usa.

X-Format (Band Interleaved by pixel pairs)

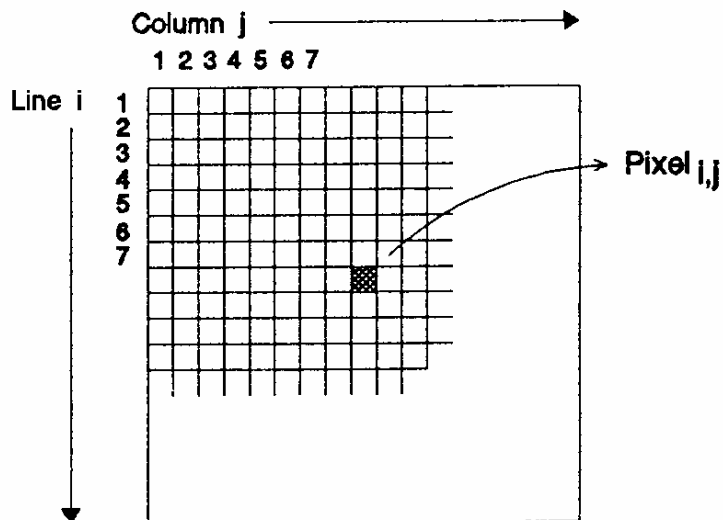
Muchas de las últimas imágenes de Landsat MSS han sido almacenadas en este formato, este consiste de 4 archivos, donde cada archivo es referenciado como un strip (tira o faja) El primer archivo es la parte izquierda de la escena, las fajas 2,3 y 4 siguen de izquierda a derecha, para las imágenes de una sola banda sin factor de bloqueo los formatos BIL, BSQ o BIP son idénticos.

Ensanche de contrastes

Imágenes Raster

El propósito del procesamiento digital de imágenes es poder distinguir objetos que no se pueden o es difícil ver directamente en la imagen raster. Para poder ser procesados los datos deben estar en formato digital, lo cual es normal para las imágenes de satélite pero no siempre lo es para las fotos aéreas por lo tanto estas deben ser escaneadas para tener la información digital. Existen cámaras para fotografía aérea que ya entregan los datos multispectrales en formato digital como la AMS (Alemania-Argentina).

La forma que tiene una imagen raster es una secuencia de líneas ordenadas en columnas como se muestra en la figura.



Como se ve es una matriz bidimensional donde cada píxel representa un valor de intensidad y es un número digital entero, con estos números se pueden realizar operaciones matemáticas que las podemos dividir en dos tipos, según el número de píxeles involucrados en la operación, como sigue:

Operaciones globales

La operación se aplica de la misma forma a todos los píxeles de la imagen, uno tras otro, por eso a veces se las llama operaciones puntuales, porque el nuevo valor de un píxel depende exclusivamente del valor que tenía anteriormente, esto se usa por ejemplo para un ensanche de radiometría.

Operaciones locales

En este caso todos los píxeles vecinos al píxel a ser trabajado son tenidos en cuenta para obtener el nuevo valor, los filtros espaciales con una ventana que se desplaza por la imagen es un caso típico.

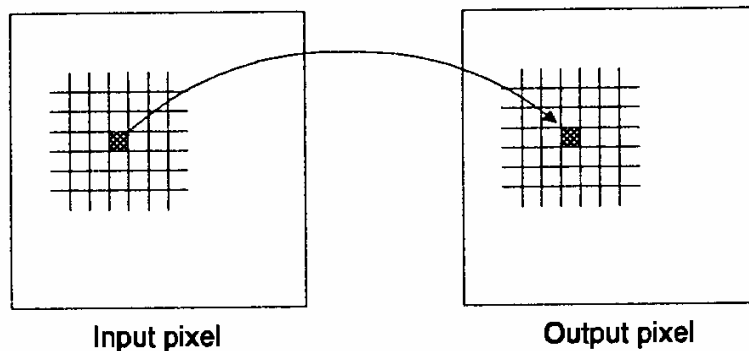
Imágenes de una sola Banda

Todos los conjuntos de datos de teledetección consisten de bandas espectrales simples, los datos multispectrales o multitemporales consisten en varias bandas compuestas por bandas simples. El procesamiento de imágenes

de una banda es fundamental para el procesamiento de cualquier tipo de imágenes raster, además el procesamiento de imágenes multitemporales se hace procesando varias bandas secuencialmente.

Operaciones globales

Normalmente los datos de imágenes satelitales o fotos aéreas tienen un contraste poco nítido y el brillo de la imagen es bajo, por lo que es difícil de interpretar directamente, por este motivo lo primero que se hace es modificar el brillo y contraste de la imagen, para esto es necesario conocer las estadísticas de la imagen.

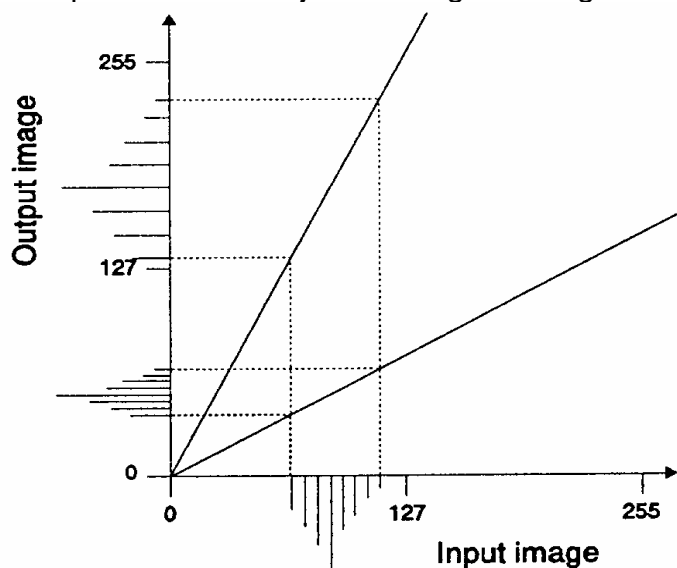


Este proceso de mejora de contraste suele llamarse ensanche de contraste. El método básico consiste en estirar (stretch) el rango de valores de gris al máximo posible (256), esto se puede hacer con una simple función lineal (linear stretch) que usa la ecuación de una recta

$$y = ax + b$$

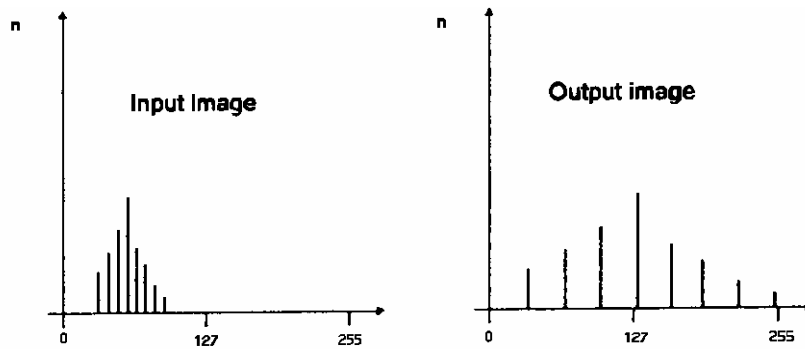
donde x es el valor del pixel de entrada (los datos originales); y el de salida (imagen corregida); a es la pendiente de la línea y b es el offset. La pendiente determina el contraste y el offset el brillo de salida.

Esto puede verse mejor en la siguiente figura

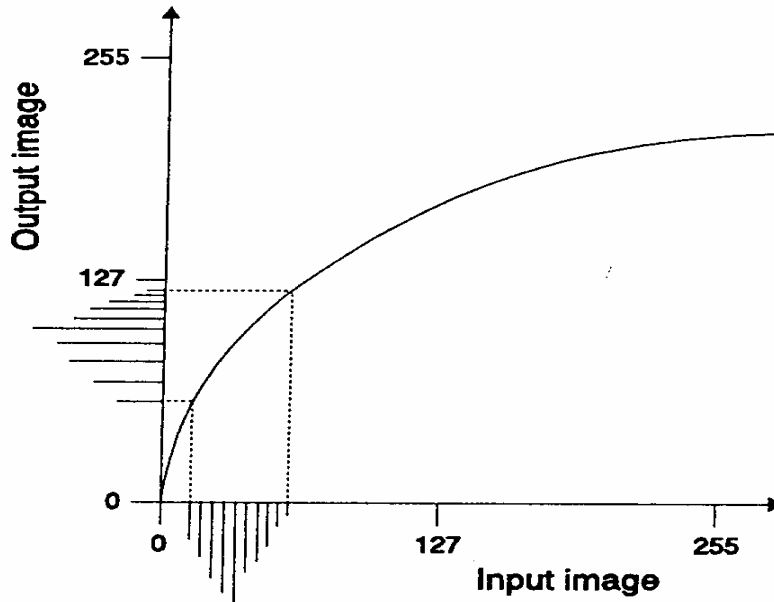


Con el ensanche no se generan nuevos valores de pixeles que anteriormente no tenían nada, sino que sólo se produce un corrimiento dentro del rango por lo

que el histograma de la imagen de salida tiene muchos espacios vacíos entre las barras como se ve en el gráfico.

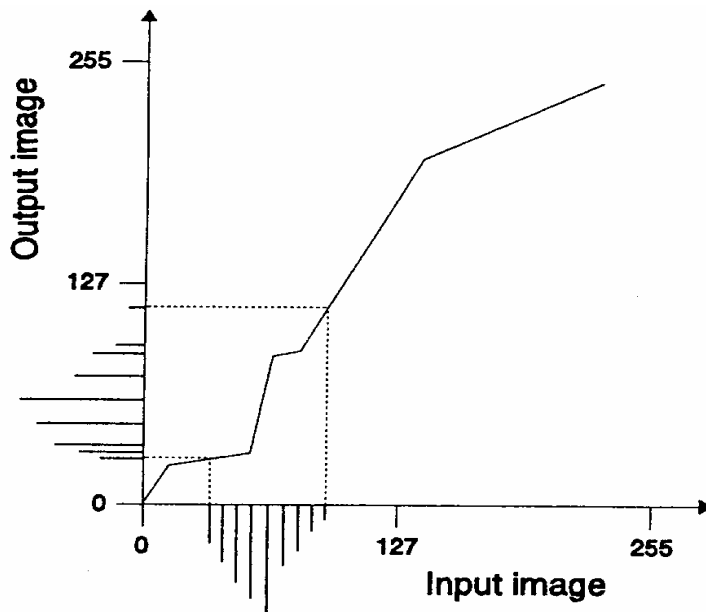


A veces el área de interés en la imagen está dentro de un rango radiométrico bastante pequeño, si ensanchamos sólo ese rango los valores máximos y mínimos del resto de la imagen pueden caer fuera del rango que puede mostrar en la pantalla y aparecen en blanco o en negro, eso es lo que se llama saturación por ensanche lineal de contraste. Cuando este método se lo aplica a toda la imagen se pierde parte de la información ya que muchos valores distintos de gris son mapeados con el mismo valor de salida. Pero es muy útil cuando se quiere destacar alguna característica en particular dentro de la imagen y su frecuencia de aparición. Normalmente los software comerciales de procesamiento de imágenes tienen una función para el ensanche automático de contraste, el punto de corte y la saturación dependen del histograma de los datos de entrada y pueden ser determinados por los valores máximo y mínimo o por la desviación estándar, usando esta última generalmente se obtienen mejores resultados.



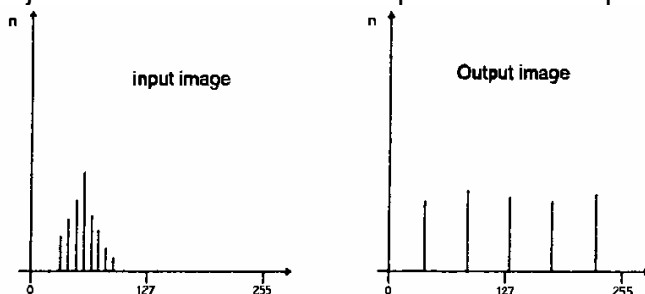
Ensanche logarítmico de un histograma

Para un ensanche logarítmico o exponencial se usa una función de orden mayor, no de primer orden como en el ensanche lineal, al usar estas funciones las zonas oscuras o claras se ven favorecidas con respecto al resto del ensanche. Existe otro método más flexible en el que hay más interacción con el usuario y se llama piecewise linear stretch, en este, el usuario define puntos de quiebre (break points) y automáticamente se varía la pendiente de la línea entre esos puntos, este método se usa mucho cuando la imagen presenta una distribución multimodal. Es como dibujar a mano la forma del polinomio de ensanche.



Ensanche lineal piecewise de un histograma

En esta transformación se usa los histogramas acumulados para tener igual probabilidad de transformación pero de todos modos no siempre la información es exactamente igual, ya que por ejemplo los pixeles con bajos niveles de gris son acumulados en la imagen de salida. Esta transformación depende mucho de la forma del histograma y generalmente se obtienen mejores resultados visuales que en una simple transformación lineal



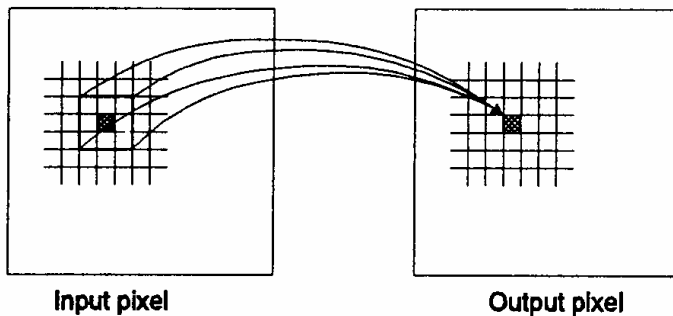
Todas estas transformaciones pueden representarse en gráficos o tablas que contienen la relación existente entre la imagen de entrada (old) y los valores de salida (new), estas tablas se las llama look up tables (LUT) y son un método

rápido para realizar las transformaciones ya que no se requiere de ningún cálculo de la computadora.

Filtrado

Operaciones locales

Las operaciones globales modifican cada pixel independientemente mientras que las operaciones locales lo modifican en función del valor de los pixeles vecinos.



Estas técnicas se usan en el espacio de imágenes (Image space) como filtrado espacial o de convolución. Es posible también trabajar en el dominio de la frecuencia por ejemplo usando las transformadas de Fourier. La idea de hacer un filtrado de la imagen es reforzar o suavizar los contrastes espaciales existentes entre los valores de los pixeles, la idea es transformar esos valores para que se asemejen o se diferencien más a los pixeles vecinos. Un punto a tener en cuenta al hacer un filtrado es que no conviene hacerlo si después se va a clasificar la imagen ya que los valores de los pixeles están cambiados, a menos que se tenga la seguridad que el filtrado ha realizar favorecerá la futura clasificación.

El proceso de la convolución espacial consiste en mover una pequeña ventana o núcleo del filtro (filter kernel) a través de toda la imagen y calculando el valor del nuevo pixel a través de todos los valores de los pixeles dentro de ese kernel. Este kernel es una matriz de números que se usa para calcular el valor del nuevo pixel en función de los valores de sus vecinos y esa matriz, conviene aclarar que la matriz se desplaza un pixel a cada paso en la imagen de entrada y que sólo se obtiene el valor para el pixel central de la imagen de salida y luego se aplica al segundo grupo de entrada y así sucesivamente va calculando los puntos de salida. Normalmente estas matrices son de 3x3 pixeles, 5x5 o 7x7 dependiendo del número de pixeles involucrados en la operación y mientras más aumente mayor es el efecto del filtrado. La convolución es una forma diferente de aplicar esa matriz. El efecto de estos filtros depende del tamaño de la matriz (número de filas y columnas) y de la distribución y del peso relativo que tengan los coeficientes dentro de esa matriz.

Filtro 1x2

1	1
---	---

Filtro 1x9

-1	-1	-1
-1	-9	-1
-1	-1	-1

Hay muchos tipos diferentes de filtros, que podemos dividir en filtros unidimensionales y filtros bidimensionales, y dentro de ellos en:

Filtros pasa bajo

Filtros pasa alto

Filtros de ensanche o detección de bordes

Filtros direccionales.

Los filtros unidireccionales sólo pueden usarse en operaciones en filas o columnas, pero no las 2 en ambas direcciones, en ese caso se necesitan filtros bidimensionales. Para calcular el valor de salida de un pixel centrado, cada coeficiente del kernel se lo multiplica por el valor del pixel correspondiente, luego la suma de esos productos se la divide por la suma de los de los coeficientes del kernel.

Filtros pasa-bajo

Ejemplos de filtros unidimensionales pasa-bajo

1/2x

1	1
---	---

1/3x

1	1	1
---	---	---

En un filtro bidimensional el pixel de salida es el resultado del promedio de los pixeles de una ventana de 3x3 (kernel), la imagen de salida se ve como suavizada y un poco fuera de foco que la de entrada debido a que tiene una menor frecuencia espacial, el principio es tomar el promedio de los 9 pixeles del kernel y asignarle ese valor al pixel central, en otras palabras lo que se hace es aislar el componente de homogeneidad de la imagen seleccionando áreas donde la frecuencia de cambio es baja. Se usan para reducir el ruido inherente al sensor que está presente en la imagen y también en zonas urbanas donde hay muchos tipos distintos de cubiertas en un reducido espacio.

1x9

1	1	1
1	1	1
1	1	1

Filtro pasa-alto

Un filtro pasa-alto acentúa los bordes entre los pixeles homogéneos, como resultado la imagen tiene mayor contraste y se ve más áspera, para una interpretación visual detallada no se recomienda usar un filtro pasa-alto. Se los usa para evidenciar cualquier discontinuidad en los contornos de áreas homogéneas.

1x9

-1	-1	-1
-1	12	-1
-1	-1	-1

Filtros de ensanche de bordes

Hay muchos filtros distintos de este tipo, uno de los más usados es el de Laplace. Este se diferencia de un filtro pasa alto en que el peso del pixel central es 8 y no 9 o mayor. Este filtro también se lo llama filtro de suma cero, y lo que hace es acentuar los bordes entre pixeles con valores heterogéneos mientras que las áreas homogéneas de la imagen permanecen en negro

1x9

-1	-1	-1
-1	8	-1
-1	-1	-1

Filtros direccionales

Son filtros que en una determinada dirección hace un ensanche de bordes mientras que en las demás direcciones no afecta a la imagen original. Estos se los usa por ejemplo para detectar fracturas o líneas estructurales en alguna dirección en particular. Hay que tener cuidado al hacer la interpretación ya que también las rutas, caminos, o vías de ferrocarril son ensanchadas y puede prestarse a confusión.

A continuación se muestran 3 tipos de filtros

1x9

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

Filtro con ensanche de bordes horizontal

1x9

-1	-1	-1
-1	9	-1
-1	-1	-1

Filtro con ensanche de bordes vertical

1x9

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

Filtro para resaltar las características en dirección Nor-oeste

Filtros no Lineales

Es otro tipo de filtros, como por ejemplo el filtro de mediana, aquí el pixel central se lo sustituye por el valor de la mediana del kernel. Este tipo de filtro es

afectado en menor grado que un filtro de promedios por los valores extremos (0,255) que muchas veces se deben sólo a valores erróneos de los píxeles.

Imágenes multi-banda

Normalmente en el procesamiento digital de imágenes se trabaja con varias bandas espectrales, por ejemplo Landsat MSS tiene 4, Landsat TM 7 y SPOT 3 en el modo XS, pero también se puede trabajar en modo multisensor o multitemporal de modo que el número de bandas a trabajar puede ser mucho mayor. Cuando se trabaja en modo multibanda se pueden tener diferentes objetivos, por ejemplo:

- Extraer información que sólo puede verse combinando 2 o más bandas espectrales.
- Reducir la cantidad de datos eliminando la información redundante.

Sumas, diferencias y cocientes de imágenes

Transformaciones simples en la imagen se pueden hacer aplicando las operaciones matemáticas básicas, para ello necesitamos un mínimo de 2 bandas para poder procesarlas.

Las operaciones más usadas son :

Sumas: La suma de 2 o más bandas espectrales se usa para reducir el ruido propio de la imagen, se pueden generar imágenes pancromáticas usando las bandas rojo, verde y azul de Landsat TM (1,2 y 3).

Diferencias: Se usan básicamente para determinar cambios, por ejemplo imágenes tomadas en distintas fechas.

Cocientes: Los cocientes sirven para detectar o acentuar características especiales como el contenido de hierro en rocas o suelos o la vegetación. También se los usa para disminuir la influencia del relieve. Un punto a tener en cuenta al realizar un cociente es que no se produzca una división por cero, lo cual es bastante normal en el caso de lagos o superficies oscuras por lo que conviene adicionar 1 siempre al divisor.

Índices de vegetación

Una caso especial de la aplicación de cocientes son los índices de vegetación que dan una medida de la biomasa. Hay diferentes tipos de índices dependiendo de las relaciones de bandas que se usen , ya que la reflexión varía de una a otra. En el procesamiento de imágenes para geología los índices de vegetación se usan para distinguir entre suelo desnudo y vegetación, con este método es posible producir máscaras para luego trabajar sólo con áreas no cubiertas. La aplicación de diferentes índices de vegetación depende del tipo de vegetación a estudiar en las distintas partes del planeta y del estudio que se quiera realizar.

Los más usados son los siguientes:

$$Vi = \frac{NIR}{I}$$

Índice de vegetación diferencial normalizado

$$NDVI = \frac{NIR-R}{NIR+R}$$

Índice de vegetación transformado

$$TVI = \sqrt{Vi + 0,5}$$

donde NIR, es infrarrojo cercano, R= canal Rojo, I= Infrarrojo.

La vegetación muestra un claro contraste entre las bandas visibles, especialmente la roja y el infrarrojo cercano, mientras que los suelos ofrecen pocas variaciones, por eso se usa este cociente para resaltar estas diferencias entre suelos y vegetación.

Transformación RGB-IHS

Es un cambio de coordenadas de RGB a IHS, del sistema de colores al sistema intensidad, saturación y tono. La intensidad es el brillo, la fuerza del color, Hue o tono es el color propiamente dicho y la saturación es la cantidad de color, por ejemplo si tenemos un conjunto de fósforos con sus cabezas pintadas color amarillo, la intensidad sería que tan amarillo son las cabezas, el color es el amarillo y la saturación es la cantidad de fósforos que tenemos por unidad de superficie, a veces son conceptos difíciles de distinguir especialmente intensidad y saturación, ya que una mayor cantidad de fósforos, o sea mayor saturación puede tener el mismo color que si a cada uno de ellos se le aumenta la intensidad del color sin variar su número por unidad de superficie. La retransformación de nuevo en RGB después de hacer algún ensanche, por ejemplo en la saturación, produce imágenes con una muy buena diferenciación de colores para la interpretación visual. Esta transformación se usa mucho también para combinar imágenes de diferentes sensores, lo que se hace luego de la transformación RGB a IHS es reemplazar la intensidad por una banda de la otra imagen, por ejemplo SPOT pancromático y luego se retransforma a RGB.

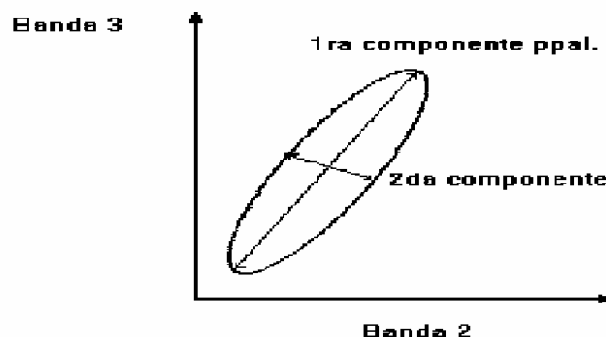
Componentes Principales

El método de análisis de las componentes principales se usa para reducir la cantidad de información existente en los datos originales, muchas veces existe información redundante en las distintas bandas del satélite y con este procedimiento se concentra la mayor parte de la información en las

primeras componentes principales. Siempre hay tantas componentes principales como bandas tenga la imagen original, las componentes principales se pueden considerar también como bandas, sólo que con otros valores de píxeles. Lo que se hace es una rotación y cambio de coordenadas, como en el caso de la transformación RGB a IHS, pero con otro criterio. Cuando se forma una imagen con las 3 primeras componentes principales se tienen características muy particulares, generalmente los datos originales tienen una distribución que se acerca a la normal y si se grafican los valores de las bandas en el dominio de las características (feature space) tiene la forma de una elipse en un espacio de dos dimensiones o de una elipsoide o hiperelipsoide en un espacio de n-dimensiones.

Como es sabido, el sentido y la fuerza de la correlación lineal entre 2 variables puede representarse gráficamente mediante un eje bivariado (scatterogram). En nuestro caso la nube de puntos indica la localización de los píxeles en las 2 bandas consideradas, la elipse que limita esos puntos expresa gráficamente la fuerza de correlación existente entre las 2 bandas y será mayor cuanto más se aproxime a una recta. Como en la transformación de componentes principales se produce una rotación de ejes, esto produce un cambio en los valores de los píxeles debido a los cambios en las coordenadas del espacio espectral, resultando los nuevos ejes paralelos a los ejes de la elipse.

La primera componente principal tiene la dirección y la longitud de la mayor transecta de la elipse, esta dirección es la del primer eigenvector y su longitud es el primer eigenvalor. La segunda componente principal es perpendicular a la primera y denota la mayor cantidad de variancia existente en los datos y que no hayan sido mostrados en la primera componente principal. La tercera componente principal es perpendicular a las dos anteriores y así sucesivamente. En el nuevo espacio espectral los ejes son paralelos a las componentes principales y los datos son almacenados en ese sistema. Hay básicamente n componentes principales en n dimensiones, pero normalmente en las primeras 4 componentes principales está contenido cerca del 98 % de la variancia de todos los datos de las n bandas de la imagen. Las componentes principales más altas muestran el ruido que contiene la imagen.



Correcciones Radiométricas

Las principales fuentes de errores en radiometría en las imágenes digitales son:

- La calibración de los detectores que convierten la radiancia recibida en valores digitales.
- Los efectos atmosféricos, principalmente la dispersión y absorción de la radiancia en su paso a través de la atmósfera.
- Las correcciones radiométricas consisten en reducir lo máximo posible esos errores que degradan la calidad de la imagen, por ejemplo para la preparación de un mosaico de imágenes esos efectos de la atmósfera y los efectos del sensor deben ser reducidos al máximo posible para obtener una imagen de buena calidad, porque en este caso podemos notar las diferencias entre las distintas imágenes.

También en imágenes simples se deben corregir algunos errores radiométricos. Dado que los datos barridos son siempre adquiridos usando más de un sensor a la vez, por ejemplo 6 sensores en Landsat MSS, 16 en Landsat TM y de 3000 a 6000 en SPOT las diferencias en las respuestas del sensor se pueden ver en diferencias de brillantez en filas y columnas. Errores de instrumentación como el bien conocido efecto de la sexta línea de Landsat 1,2 y 3 de MSS o líneas erróneas son típicas de los sistemas con scanner. Una comparación de la media y la desviación estándar de las líneas de cada detector permite la corrección y agregado de las líneas faltantes con la ayuda de un modelo de detector estándar.

A veces aparecen píxeles con valores defectuosos o líneas con valores realmente incorrectos que normalmente se deben a problemas de transmisión entre el satélite y la estación receptora en tierra. Esos píxeles o líneas se los reemplaza con el promedio de los valores de los píxeles o líneas circundantes. Los modelos de los efectos atmosféricos explican el por qué de la dispersión y absorción de la radiación recibida desde la tierra durante su paso a través de la atmósfera. Para una corrección precisa se requieren además muchos otros datos atmosféricos en el momento de la toma como temperatura, humedad relativa, presión atmosférica, visibilidad y otros datos que en la mayoría de los casos no están disponibles, por lo que normalmente se hacen correcciones aproximadas que dan resultados satisfactorios. La corrección de bruma se basa en el principio de que en cada banda hay píxeles que tienen valores cercanos a cero o cero, y la atmósfera suma un valor constante a todos los píxeles. Esta influencia se nota menos en las bandas infrarrojas y es bastante grande en las bandas visibles y aun más en la banda azul. Para corregirlo lo que se hace es correr los histogramas de cada banda restándole el offset entre cero y el menor valor de gris de cada píxel.

Conclusión

En este artículo se presenta una visión general de lo que es el procesamiento digital de imágenes. Se explican brevemente las más importantes técnicas que se usan para extraer información de los datos de teledetección. Esta es una ciencia que cada día crece más y más. Las aplicaciones se multiplican y la cantidad de especialistas en el mundo entero

también por eso quisimos presentar este resumen para la gente que se inicia el tema tenga una idea clara y global de los principales aspectos de la teledetección.

Bibliografía

- American Society of Photogrammetry, Manual of Remote Sensing, Volume I and II (The Sheridan Press)
- Chuvieco Emilio, Fundamentos de la Teledetección Espacial, (Ediciones RIALP, S. A. Madrid, 1979)
- Diccionario SELPER dictionary, dicionário, SELPER, Sociedad de Especialistas Latinoamericanos en Percepción Remota.
- Elachi C. 1987, Spaceborne Radar Remote Sensing: Applications and Techniques; IEEE Press, New York.
- Gierloff -Emden, H.G., Schröder-Lanz, H.
- Luftbildauswertung I, II, III BI Hochschultaschenbücher 358/358a, 367/367a/367b, 368
- GTZ (Deutsche Gesellschaft für Technische Zusammenarbeit GmbH). Procesamiento Digital de Imágenes, Aplicaciones en Fotogrametría y Teledetección Prof Hans-Peter Bähr, Universidad de Karlsruhe, 1991
- Jean - Yves Scanvic, Teledetección Aplicada (Paraninfo 1989)
- Jensen John R, Introductory : Digital Image Processing (Prentice-Hall, Englewood Cliffs, N.Y.)
- Lillesand/Kiefer, Remote Sensing and Image Interpretation, Wiley & Sons, New York, 1979 L
- Löffler E. Geographie und Fernerkundung, (Teubner Studienbücher)
- M. Duane Nellis, Ray Lougeay and Kamlesh Lulla Current Trends in Remote Sensing Education
- (Geocarto International Centre, Hong Kong 1989)
- Moik Johannes G. Digital Processing of Remotely Sensed Images (NASA Publication SP; 431)
- Siegal, B.S., Gillespie, A.R. Remote Sensing in Geology, Wiley & Sons, New York, 1980
- Ulaby F. T. et al., 1981, Microwave Remote Sensing; Active and Passive; Vol. I; Addison Wesley Publishing Company.
- Ulaby F. T. et al., 1982, Microwave Remote Sensing; Active and Passive; Vol. II; Addison Wesley Publishing Company.
- Ulaby F. T. et al., 1986 Microwave Remote Sensing; Active and Passive; Vol. III; Addison Wesley Publishing Company.
- Way, D. S. Terrain Analysis McGraw-Hill Book Company, New York, 1978

Información de los autores:

Dr. Harald Mehl

Deutsches Zentrum für Luft und Raumfahrt
(DLR) Oberpfaffenhofen / DFD
Weßlingerstraße
82234 Weßling
Tel.: 08153 281137
E-mail: harald.mehl@dlr.de

Dr. Osvaldo Peinado

Deutsches Fernerkundungsdatenzentrum (DFD)
Deutsches Zentrum fuer Luft- und Raumfahrt (DLR)
P.O.Box 11 16 . D-82234 Wessling Germany
Tel: +49-8153-28-3010
Fax: +49-8153-28-1445
E-mail: Osvaldo.Peinado@dlr.de