

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

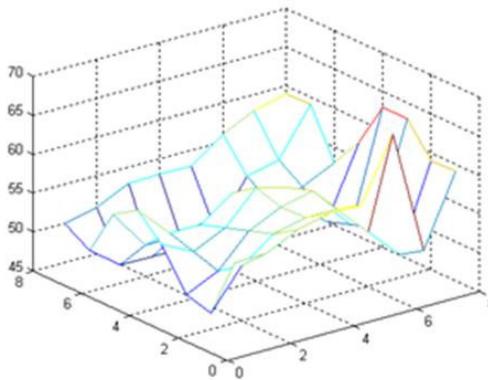


Ilustración 1 Textura de una escena

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

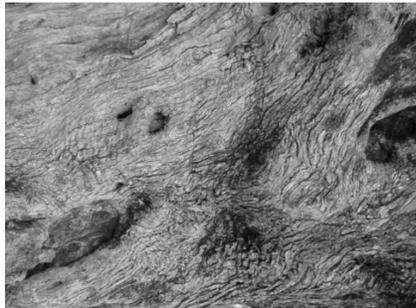
4.1 Textura

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.

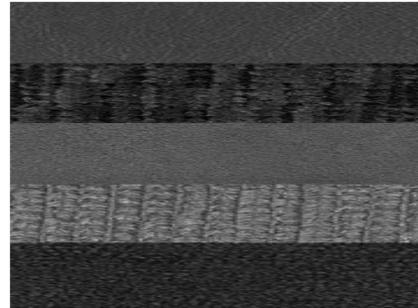
Estas propiedades son el suavizado, la rugosidad, la regularidad, la granularidad, etc.

Las texturas son frecuentemente definidas como clases de “patrones” visuales homogéneos y las podemos observar en cierto tipo de materiales tales como telas, 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 (Ilustración 42).

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 2 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 el 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 píxeles 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 “anisotropía” 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.

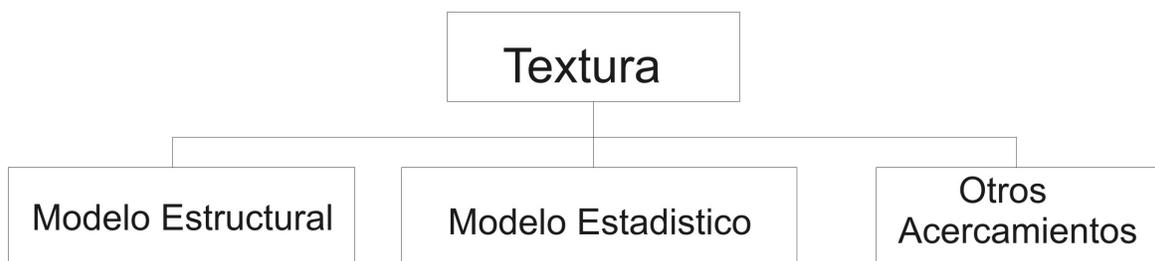
(f) Rugosidad

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

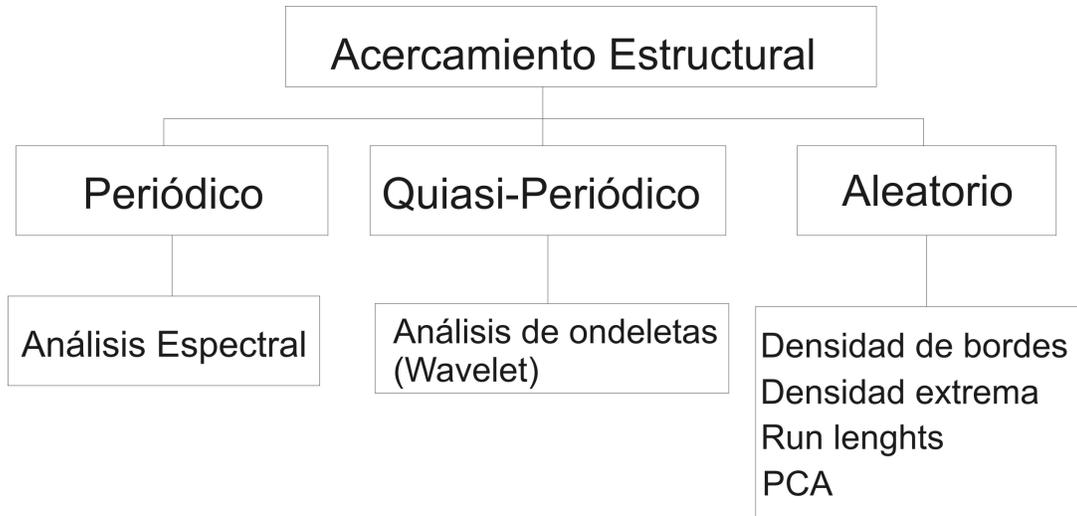
Cuando podemos identificar o existe uno o más primitivas de textura llamados textels para elementos de textura (o también llamados texons, textons, para átomos texturales) los cuales consisten de cualquier patrón el cual se repite de acuerdo a algunas reglas:

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

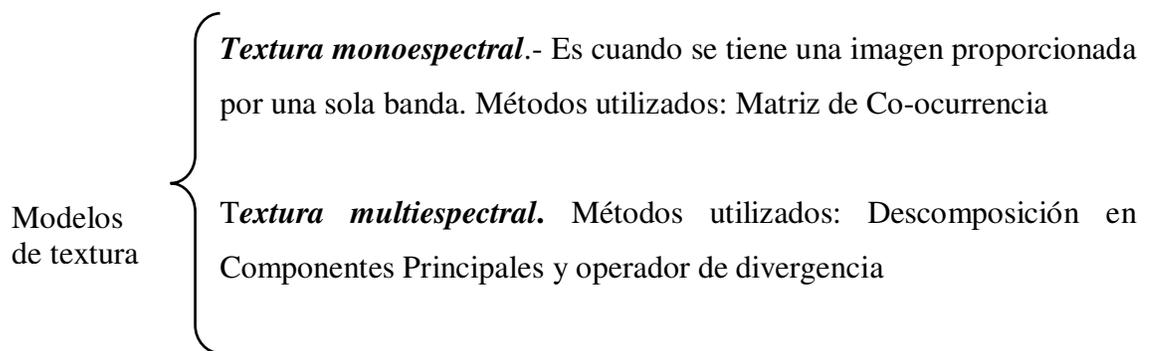
El problema del tratamiento de texturas se enfoca como un problema de reconocimiento de patrones, esto es debido a que muchas texturas no tienen una distribución geométrica regular (imágenes de escenas naturales), por el contrario presentan variaciones que no pueden ser descritas por un modelo satisfactorio 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 pixeles vecinos.

Para ello tenemos que revisar la función de densidad conjunta, la cual expresa la probabilidad de que ocurran pares de pixeles 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 pixeles 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 co-ocurrencia proporciona la probabilidad de encontrar un par de valores de pixeles 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 píxeles de la ventana.

La matriz c de co-ocurrencia es la probabilidad estimada de encontrar un par de valores de píxeles (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 píxeles.

Debido a que la función de correlación decae para píxeles vecinos, es suficiente tomar la magnitud de β como para considerar píxeles 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 píxeles en la posición relativa dada por β , por esta razón la matriz de co-ocurrencia, establece la estructura espacial de los píxeles 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.

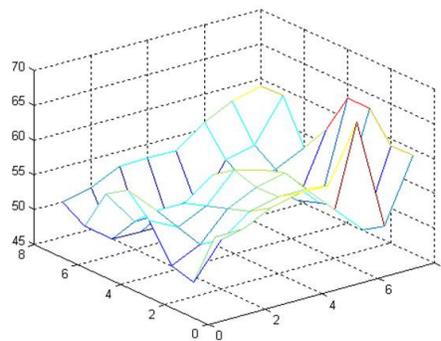


Ilustración 3 Superficie generada por los valores de los píxeles de la Banda 1, la cual muestra la textura que representa dicha banda

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

Si los valores de los píxeles 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 (píxeles) 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 ésta 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.

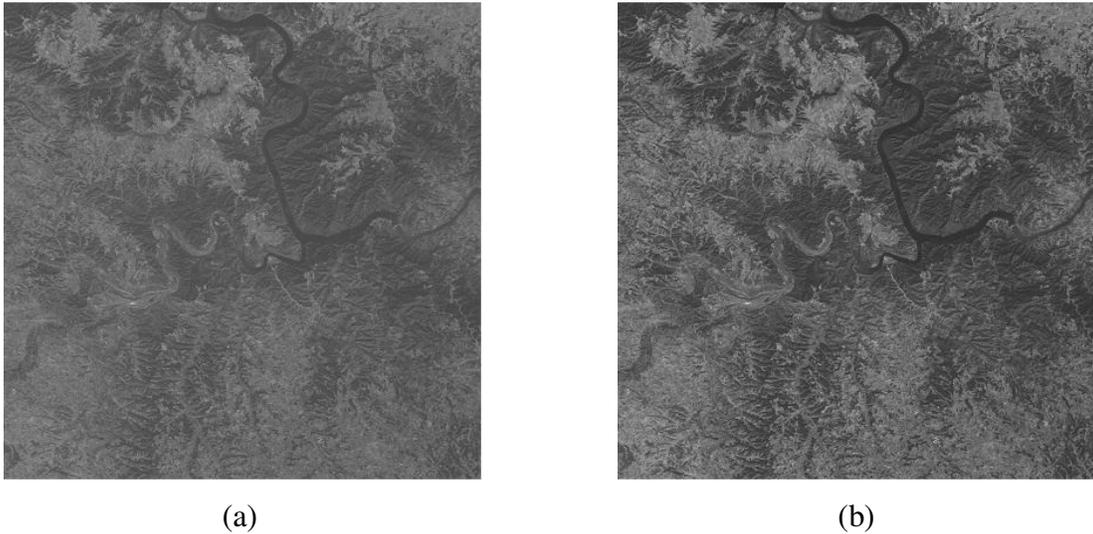


Ilustración 4 Primeras dos bandas de una imagen multiespectral Landsat TM

Tomando sólo un área de 5×5 píxeles 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.

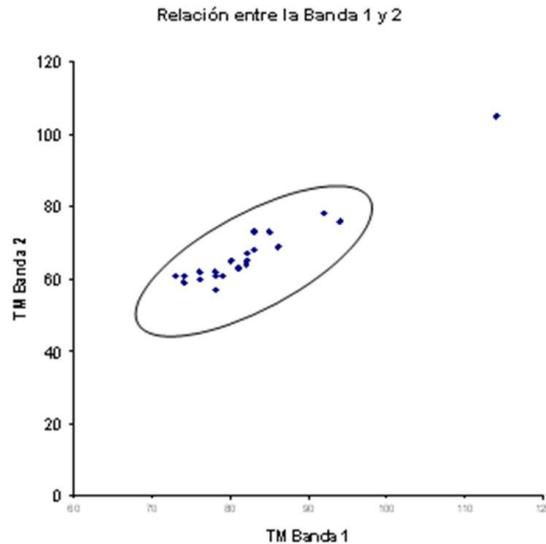


Ilustración 5 Representación de la descomposición en componentes principales

Gráfica de los valores digitales de la banda1 (eje x_1) contra los valores digitales de la banda2 (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, incrementarán 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 dimensión larga 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 se definen secuencialmente. Cada nueva coordenada se orienta perpendicularmente hacia todas las direcciones previamente definidas y en la dirección de la densidad máxima restante de los puntos definidos por los píxeles.

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

Estos valores digitales se usan después 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 píxeles 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:

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

Donde γ es el número de bandas de la imagen

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

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

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 píxeles 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:

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

Donde b_1' son las direcciones del nuevo espacio y representan los valores de los píxeles 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 píxeles 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 píxeles, en niveles de gris.

La imagen multiespectral original se muestra en la Ilustración 46.

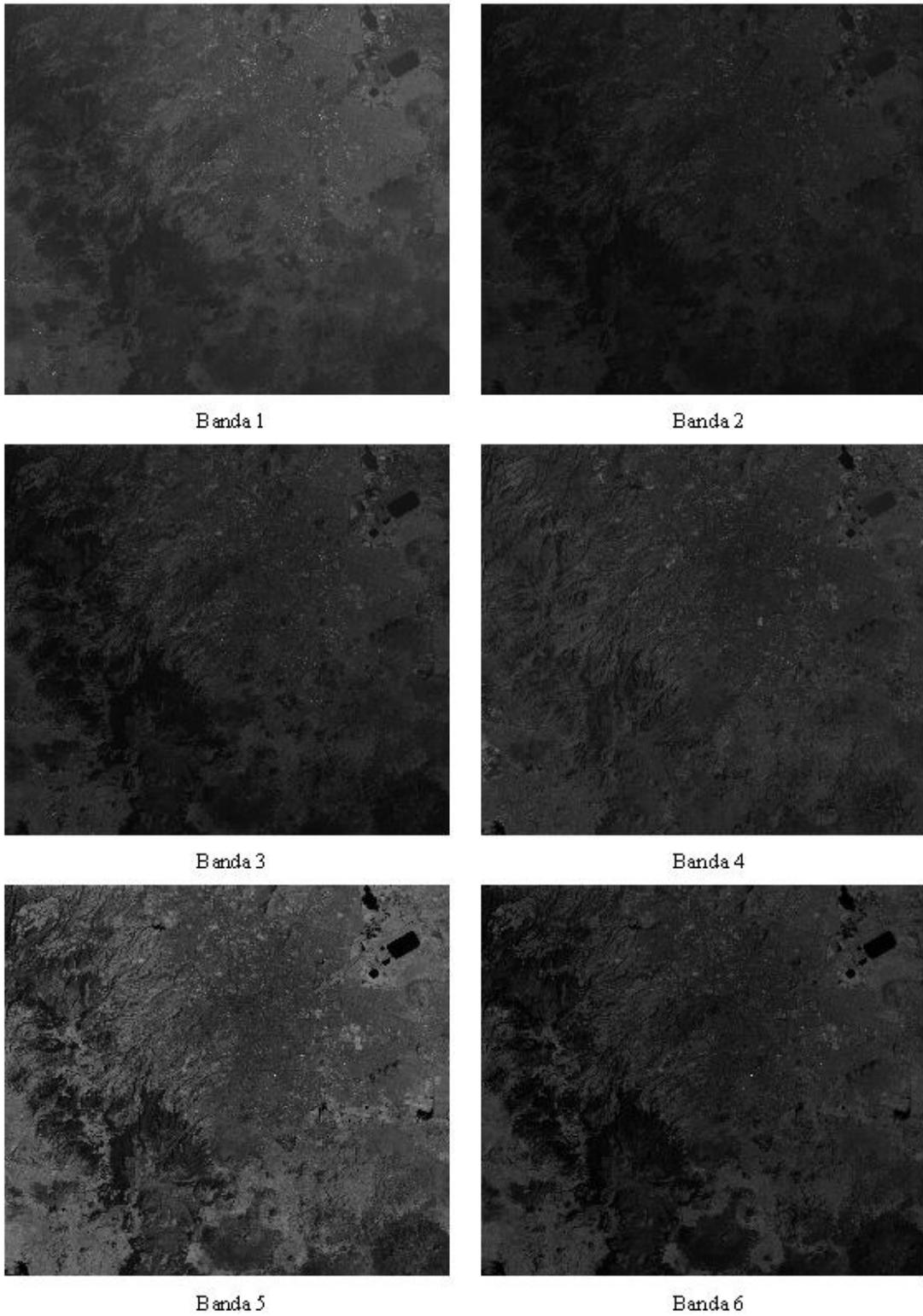


Ilustración 6 Imagen multiespectral del valle de México

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:

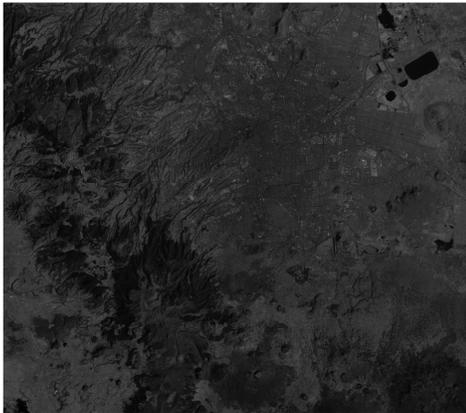
0.35596	0.22822	0.35581	0.23153	0.66816	0.44109
0.66527	0.30982	0.34090	0.20010	0.55152	0.03168
0.06346	0.09205	0.05864	0.89351	0.13619	0.40884
0.28532	0.00512	0.03429	0.29948	0.47893	0.77352
0.54628	0.18277	0.78348	0.12390	0.01575	0.19680
0.21645	0.90002	0.37247	0.05435	0.03483	0.01477

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 %

:

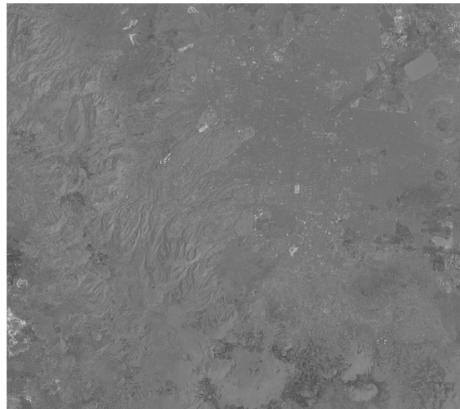
Observamos que las tres primeras bandas contienen el mayor porcentaje de información significativa (Ilustración 47).



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 7 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 píxeles pertenecientes a ésta, los píxeles que cumplen con ciertas propiedades pueden ser localizados en alguna nube de píxeles o cúmulos. La separación entre estos es importante para saber en dónde poner a cada píxel 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:

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

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.

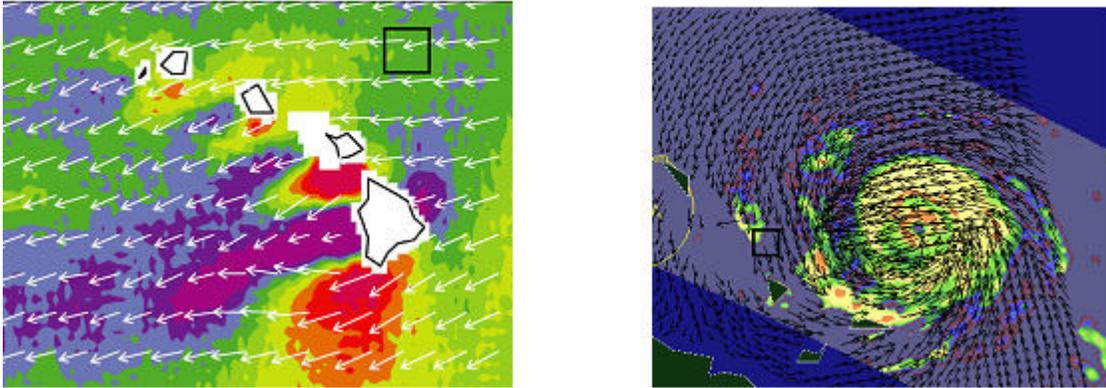


Ilustración 8 Campos vectoriales

Si observamos las cajas superpuestas en la ilustración 48 de cada campo vectorial y tomamos en cuenta que 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) es el comportamiento de la divergencia en un campo vectorial es el mostrado en la 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 (Fuente)
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 o pozo.
Si $\nabla \cdot F(x, y, z) = 0$	El campo vectorial F es una Fuente 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

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

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:

$$\phi = \iint_S \hat{n} \delta a \quad (\text{IV. 23})$$

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:

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

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

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

La divergencia es una medida del flujo del campo vectorial $U = u(x_1, x_2, \dots, x_n)$ 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

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

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:

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

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 píxeles cercanos a 0, en cambio las texturas rugosas presentan niveles de gris de mayor intensidad, tonos de gris claros con valores en sus píxeles 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.