

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO
FACULTAD DE INGENIERIA
DIVISION DE ESTUDIOS DE POSGRADO

**Un clasificador supervisado
para imágenes multispectrales y texturales
basado en un algoritmo automatizado de
crecimiento de regiones**

TESIS

que para obtener el grado de
Maestra en Ingeniería Eléctrica
(Comunicaciones)

PRESENTA

Gabriela Mariel Maletti

Dirigida por:

Dr. Jorge Lira Chávez

Ciudad Universitaria, México D.F., 1998



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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



DEPM

T. UNAM

1 9 9 8

MAL

Ej. 2

A
Esther y Carlos
Mis padres

A
Graciela y Roberto
Mis hermanos

Agradecimientos

A mis abuelos, por haberme heredado su espíritu de lucha.

Al Opa Carlos, por haberme dado su bendición antes de partir.

A mis padres Esther y Carlos, por enseñarme los buenos valores y haberlos defendido con el ejemplo, por su apoyo constante y su gran amor.

A mi hermano Roberto, por sus lindas cartas y su ejemplo de constancia y dedicación. A mi hermana Graciela, por haberme venido a visitar y apoyar.

A Jorge Lira y a su esposa Martha Lucía por su calidez y respaldo continuo.

A Dubhe y a Ricardo, por abrirme sus corazones al llegar y llenarme de energías.

A José y a Francisco, por su linda amistad.

A Ericka, la amiga que me acompañó cada uno de los días de mi tesis.

A Ernesto, Antonio, Fernando, Cris y Vero, por sus buenas ondas.

A Sergio, mi amor.

Al Dr. Jorge Lira, por haberme abierto las puertas de este fascinante mundo del procesamiento de imágenes.

Al Instituto de Geofísica, por permitir que empleara sus instalaciones para realizar esta tesis.

A DGAPA, por la beca otorgada.

Indice

CAPITULO I

LA INTRODUCCION AL TEMA DE TESIS DESDE EL PUNTO DE VISTA DE SISTEMAS

I.1 DEFINICIONES Y TERMINOLOGIA	
I.1.1 Introducción	1
I.1.2 La caracterización de una escena	3
I.1.3 La imagen digital multispectral	5
I.1.4 El realce	8
I.1.5 La segmentación	8
I.1.6 El reconocimiento de patrones	10
I.1.7 El modelo de la escena	12
I.2 ANALISIS	
I.2.1 La organización formal	14
I.2.2 La definición del problema	16
I.2.3 Partes del sistema de modelado de la escena	17
I.2.4 Necesidades y limitaciones	32

CAPITULO II

EL CLASIFICADOR CONTEXTUAL

II.1 EL ESQUEMA DE CRECIMIENTO DE REGIONES	
II.1.1 La descripción del algoritmo de crecimiento de regiones	34
II.1.2 La formalización del algoritmo de crecimiento de regiones	34
II.1.3 La implementación del algoritmo de crecimiento de regiones	35
II.2 LA ESTIMACION DE LOS PARAMETROS OPTIMOS PARA EL CRECIMIENTO DE REGIONES	
II.2.1 El criterio de homogeneidad	35
II.2.2 La función objetivo	36
II.2.3 La ventana óptima de una clase	41
II.2.4 La función de separabilidad	50
II.2.5 El umbral óptimo	52
II.2.6 Notas acerca de posibles caminos de estudio	53
II.3 FUNDAMENTOS DEL CLASIFICADOR	
II.3.1 La estimación de los parámetros para la clasificación	54
II.3.2 La clasificación	54

II.3.3 Clasificadores espectrales	56
II.3.4 Clasificadores texturales	60

CAPITULO III

RESULTADOS Y DISCUSION

III.1 LAS BANDAS SINTETICAS	
III.1.1 Los criterios de evaluación	64
III.1.2 La generación de las imágenes sintéticas	65
III.1.3 El simulador	67
III.1.4 Las simulaciones realizadas	69
III.2 LA IMAGEN SPOT	
III.2.1 Una banda con segmentación por medias espectrales	85
III.2.2 Una banda con segmentación por histogramas conjuntos	86
III.2.3 Dos bandas con segmentación por medias espectrales	94
III.2.4 Dos bandas con segmentación por histogramas conjuntos	97
III.3 LA IMAGEN SEASAT	
III.3.1 Segmentación por matrices de co-ocurrencia	98
III.3.2 Segmentación por centros de masa de matrices de co-ocurrencia	100
III.4 CONCLUSIONES Y PERSPECTIVAS	
III.4.1 Conclusiones	104
III.4.2 Caminos a seguir	105

Introducción

En un clasificador contextual se clasifica un pixel no solamente en función de sus propiedades espectrales puntuales sino también en función del contexto espacial donde se encuentra. Un clasificador puntual clasifica los pixels únicamente en función de sus valores a través de las bandas. Los clasificadores contextuales son clasificadores del entorno, puesto que consideran una vecindad alrededor de cada pixel. La bibliografía reporta sobre la necesidad de incluir tal información de la vecindad con información espectral [Wilson, 1992] o información textural [Jensen, 1990].

Un clasificador supervisado emplea información a priori de cada clase determinada: esto se hace usualmente por medio de conjuntos de entrenamiento o pixels prototipo. Se tienen antecedentes de un algoritmo automatizado de crecimiento de regiones que a partir de un pixel sembrado por cada clase utiliza un criterio de entropía para calcular sus parámetros: el tamaño de ventana óptimo y los umbrales óptimos. El proceso de crecimiento de regiones consiste en tomar una medida de similaridad e incorporar a la región crecida aquellos pixels que satisfacen un criterio de homogeneidad. La región crecida constituye una muestra estadísticamente válida [Chavira, 1996].

En este trabajo se propone un nuevo clasificador contextual donde se determinan muestras estadísticas de las clases definidas como resultado de un algoritmo automatizado de crecimiento de regiones [Lira & Frulla, 1998]. Un pixel es clasificado comparando la forma, la media y/o la correlación de la función de densidad asociada a la vecindad del pixel con aquella de las clases. La comparación se lleva a cabo por medio de una medida de similaridad entre las funciones de densidad tanto para la respuesta espectral como para el contenido textural. Este esquema de clasificación es válido para ambos clasificadores espectral y textural tratados en el presente trabajo.

El trabajo de tesis se desarrolla bajo la dirección del Dr. Jorge Lira Chávez en el Laboratorio de Percepción Remota del Instituto de Geofísica de la Universidad Nacional Autónoma de México. Este Laboratorio forma parte de la estructura de investigación del Instituto de Geofísica, desarrollándose en él una línea de investigación llamada "Modelos Matemáticos para el Análisis Digital de Imágenes". La presente tesis forma parte de esta línea, específicamente dentro del marco del proyecto IN102797 intitulado: "Diseño e Implantación de un nuevo clasificador supervisado para imágenes multiespectrales y texturales basado en un algoritmo automatizado de crecimiento de regiones".

La tesis está compuesta de tres capítulos. El primero consiste en "La introducción al tema de tesis desde el punto de vista de sistemas", en el cual se realiza una revisión de antecedentes. El segundo capítulo se denomina "El clasificador contextual" y contiene la descripción del esquema de crecimiento de regiones, de la estimación de los parámetros óptimos para el crecimiento de regiones y los fundamentos de los clasificadores implantados. Por último, en el tercer capítulo, se presentan los "Resultados y Discusiones" en imágenes sintéticas y dos imágenes reales. Al final del tercer capítulo se encuentran las conclusiones de la presente tesis.

CAPITULO I

LA INTRODUCCION AL TEMA DE TESIS DESDE EL PUNTO DE VISTA DE SISTEMAS

I.1 DEFINICIONES Y TERMINOLOGIA

I.1.1 Introducción

Con el objeto de emplear un lenguaje común, se introduce a continuación terminología propia de la teoría de sistemas, la cual será empleada a lo largo de este capítulo. Se lista una serie de proposiciones y una explicación.

Un sistema es un conjunto de entidades con sus atributos relacionados entre sí que persiguen un objetivo.

Analizando palabra por palabra esta definición, se entiende por **conjunto** a una reunión de elementos. Una **entidad** es todo lo que es. Por ejemplo, en una escena dada, se tienen las siguientes entidades: bosque, ciudad, cultivo, carretera, lago. Una entidad puede tener existencia real o ideal y es un conjunto, donde cada uno de sus elementos se denominan **instancias de entidad**. El Bosque de Chapultepec es una instancia de la entidad bosque, es decir, es un elemento del conjunto bosque, es uno de los tantos bosques que existen. La carretera Federal a Cuernavaca es un elemento de la entidad carretera. La ciudad Toluca es un elemento de la entidad ciudad.

Los **atributos** son las propiedades o características de las entidades. Por ejemplo, el área cubierta por el bosque es un atributo del mismo; la longitud de la carretera es un atributo de la entidad carretera; el estado al que pertenece una ciudad es un atributo de la entidad ciudad. El **dominio de un atributo** es el conjunto de valores posibles del atributo. Por ejemplo, todos los estados de México forman el dominio del atributo estado de la entidad ciudad. El dominio del atributo área de la entidad bosque está formado por todos los valores que van desde 0 hasta una cantidad determinada de hectáreas al cuadrado. Un **dato** es el valor de un atributo de una instancia de entidad. Morelos es el valor del atributo estado de la instancia de entidad Cuernavaca perteneciente a la entidad ciudad.

La **relación entre las entidades** es el atributo fundamental de un sistema, puesto que permite que un sistema sea más que la suma de sus partes. Siguiendo con el ejemplo, tener las entidades ciudad, carretera, lago, cultivo no relacionadas, no es suficiente para tener un sistema. Si tal carretera pasa por tal ciudad se tiene una relación entre la entidad carretera y la entidad ciudad. La relación entre entidades es a su vez una nueva entidad y por eso un sistema es más que la suma de sus partes.

El **objetivo** es lo se desea obtener.

Un sistema está incluido en un universo e interactúa con él.

El **universo** es el espacio que contiene al sistema. Los sistemas no son cerrados, necesitan interactuar con el espacio al que pertenecen. Esto se realiza a través de una relación funcional de entradas y salidas del sistema (Figura 1.1). Las entradas están constituidas por los datos que alimentan el sistema. Las salidas son el resultado de aplicar un conjunto de funciones a las entradas. Si las entradas están caracterizadas por una función **f** y las salidas, por una función **g**, el sistema que genera la salida está caracterizado por un operador **O**, tal que:

$$g = O[f] \quad [1.1]$$

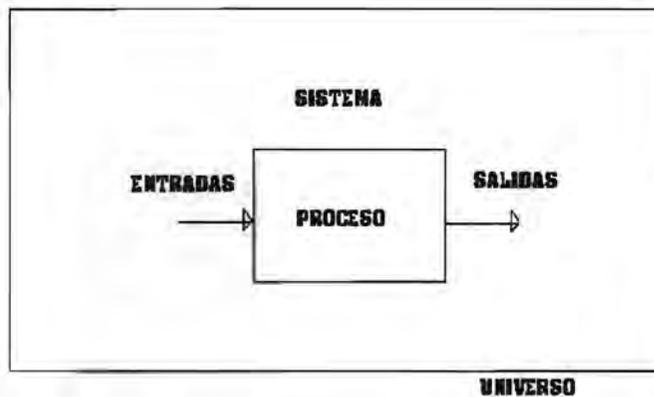


Figura 1.1

Todo sistema contiene subsistemas.

Este concepto nos da una idea de la relatividad de la definición de sistema. Una entidad incluida en un sistema dado, puede ser vista como sistema propiamente dicho en otro problema. En el ejemplo anterior, se consideró a ciudad como entidad; sin embargo, en otro problema, puede ser vista como un sistema formado por las entidades barrio, calle, plaza, habitante, etc.. El sistema ciudad es entonces un subsistema de la escena dada.

Un modelo es la representación de una realidad.

La realidad es el sistema. Es necesario representar la realidad porque ésta no se puede retener. Quien registra los pagos al personal de la UNAM no puede poner a todos los investigadores, académicos y administrativos con sus sueldos en un fichero. Toma una representación de esa realidad, registra un nombre, un número de expediente, el importe de su sueldo, etc.. El huracán Paulina fue una realidad que no se pudo retener. Y por más que se hubiera podido retener el huracán Paulina, no hubiese sido posible retenerlo en todos los instantes del tiempo. Por eso se tomó una representación de esa realidad: una sucesión de imágenes de satélite. Aún más, es necesario representar la realidad, porque la realidad es compleja y se desea predecir y controlar su comportamiento, se necesita tomar decisiones y para ello, se necesita contar con algo más simple que la realidad misma. En el caso de acercamiento de un huracán, se necesitan tomar decisiones como organizar un plan de

evacuación y para ello se debe conocer el tiempo en que tardará en llegar, etc.. Nada se gana con estar en el ojo del huracán, con tener la realidad encima de uno. Para abstraer o representar una realidad se deben definir entidades y atributos que contengan su esencia, es decir, aquellos que son relevantes, no tomando en cuenta aquellos que no lo son.

I.1.2 La caracterización de una escena

Podemos considerar la caracterización de una **escena** como un ejemplo de modelado. Una escena constituye en este caso el sistema, la realidad a representar. La escena forma parte de un **sistema físico**, que juega el rol de universo. Vemos aquí un ejemplo del concepto de un sistema que contiene subsistemas, puesto que la escena es un subsistema del sistema físico. Se tiene también una **fuerza de radiación**, que no necesariamente tiene que estar separada de la escena. En la escena se definen una colección de subáreas o elementos físicos que se llaman **campos instantáneos de vista**, cada uno de los cuales constituye una instancia. A la discretización espacial de la escena se le denomina **digitización**. La escena en sí es también una entidad que tiene **atributos espectrales y texturales** que se desean cuantificar, para poder manipularlos de una manera más sencilla.

La radiación de la fuente interacciona con el campo instantáneo de vista. El resultado de tal interacción se mide con arreglos matriciales de detectores. Cada detector integra la radiación del campo instantáneo de vista y genera un voltaje a la salida, el cual se discretiza a través de un convertidor analógico-digital en una cantidad numérica. A la discretización radiométrica proveniente de los campos instantáneos de vista se le llama **cuantización**. En resumen, el proceso de caracterización de la escena se realiza a través de dos sub-procesos: la digitización y la cuantización.

El resultado de la medida obtenida por el convertidor analógico-digital se almacena en una colección de matrices denominadas **bandas**. El tamaño de cada una de estas matrices es el número de campos instantáneos de vista en que fue subdividida la escena. A la colección de bandas se le llama **imagen digital multiespectral** y es la caracterización numérica de la distribución espacial de los atributos espectrales de la escena. Sin embargo, siendo una caracterización numérica de la escena, no tiene de por sí dimensión espacial ni representación visual. Un **pixel** es un elemento lógico de la imagen que porta información de un elemento físico de la escena: de un campo instantáneo de vista. Un pixel en la imagen es la caracterización numérica de las propiedades radiométricas promedio de un campo instantáneo de vista: es una muestra estadística de la respuesta promedio de la radiación proveniente del campo instantáneo de vista. (Figura 1.2)

Utilizando la primera formalidad matemática introducida con anterioridad (Ecuación 1.1), podemos decir que si la escena está caracterizada por una función $f(x,y)$, la imagen por una función $g(x,y)$ y el sistema que genera la imagen por un operador O :

$$g(x,y) = O[f(x,y)] \quad [1.2]$$

El número óptimo de pixels para una escena dada puede establecerse utilizando el teorema de muestreo en dos dimensiones.

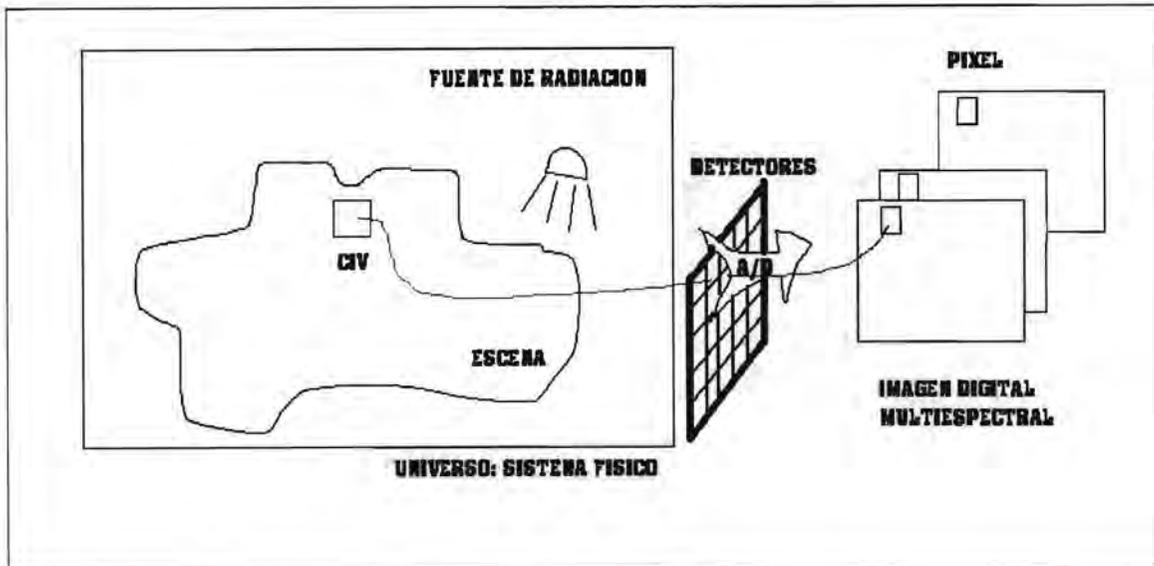


Figura 1.2

En Resumen:

Una imagen digital multispectral es entonces el resultado del proceso de caracterización de la escena, es decir, un modelo de la misma. Esta primera representación de la escena, constituye ahora una nueva realidad, todavía compleja, que necesita pasar por un proceso de **análisis digital** que produzca datos estructurados (**información**) útil para la toma de decisiones.

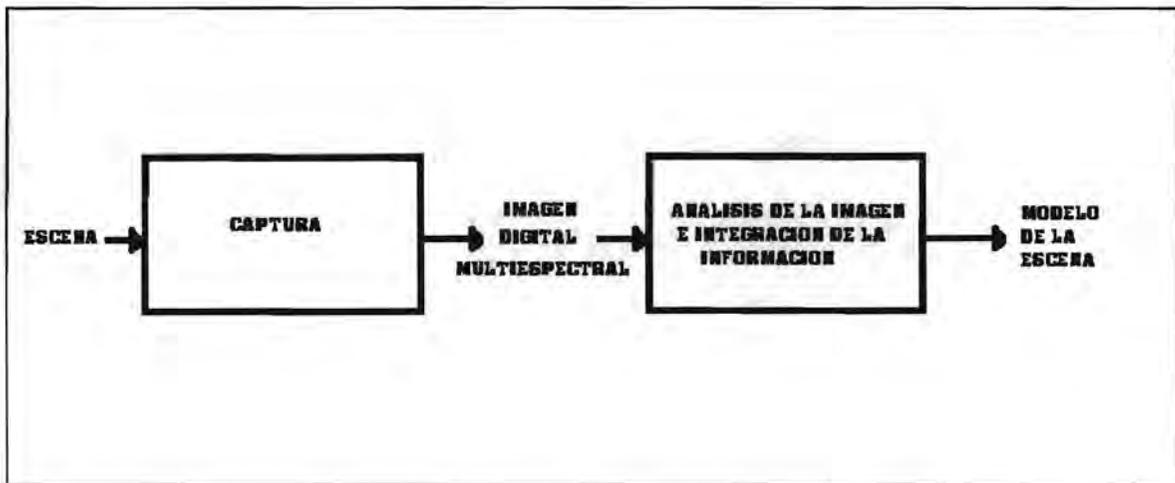


Figura 1.3

A grandes rasgos, la entrada al sistema de análisis digital de imágenes es entonces una imagen digital multispectral y la salida de este sistema es un modelo más sencillo y manipulable de la escena. El sistema de análisis digital de imágenes "ve" a la **imagen como el sistema** a modelar, o más específicamente, como el modelo de la escena (o sistema) que

hay que pulir y refinar. El proceso de análisis digital de imágenes se realiza a través de tres etapas: el realce, la segmentación y el reconocimiento de patrones.

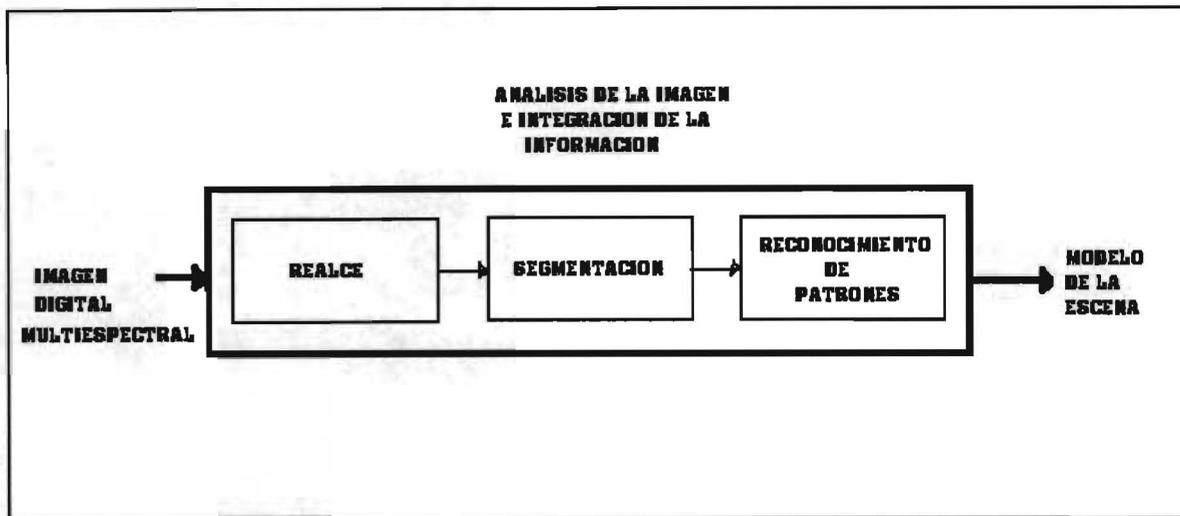


Figura 1.4

Se procede a continuación a describir más detalladamente cada uno de los elementos mencionados en el párrafo anterior. (Tomado de Notas de los cursos de maestría Procesamiento Digital de Imágenes y Reconocimiento de Patrones dictados por el Dr. J. Lira, DEPMI, UNAM).

1.1.3 La imagen digital multiespectral

Una imagen digital multiespectral puede definirse desde distintos puntos de vista. Se introducirán aquellas definiciones que sean empleadas a lo largo de este trabajo.

La imagen digital multiespectral es un campo vectorial.

En un espacio de tantas dimensiones como bandas tenga la imagen digital multiespectral pueden graficarse todos los pixels, generándose una nube de datos. Dicho espacio, es un espacio vectorial. Cada uno de los ejes de este espacio representa a una de las bandas de la imagen digital multiespectral y los valores que toman los pixels en cada banda, son las coordenadas de los puntos que pertenecen a la nube de datos. La imagen digital multiespectral constituye un campo vectorial y gráficamente se representa entonces, en el espacio vectorial (Figura 1.5).

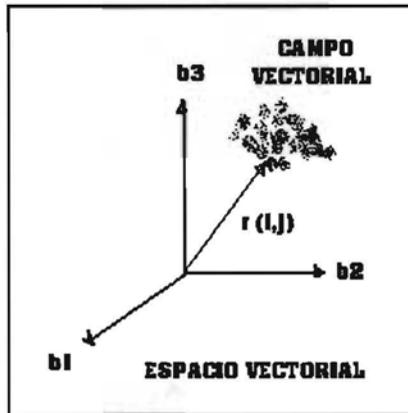


Figura 1.5

Cada pixel del espacio está definido por su vector de posición r_{ij} , donde i es la coordenada de la fila y j es la coordenada de la columna dentro de las bandas. Un pixel está incrustado en cierto contexto espacial y guarda, por ende, una relación de vecindad con pixels próximos vecinos bajo una cierta conectividad. Por la naturaleza digital de la imagen distinguimos ocho direcciones, que según el Código de Freeman se denominan 0, 1, 2, 3, 4, 5, 6 y 7 (Figura 1.6).

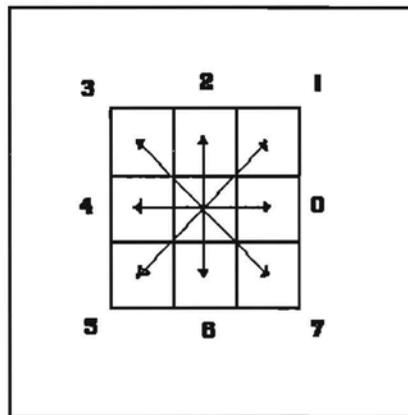


Figura 1.6

Los primeros próximos vecinos ocurren en las direcciones 0, 2, 4 y 6 del código de Freeman. Esta relación de vecindad se denomina conectividad 4. Los segundos próximos vecinos ocurren en las direcciones 1, 3, 5 y 7 del código de Freeman. Los primeros y los segundos próximos vecinos guardan una relación de vecindad con el pixel considerado llamada conectividad 8.

La imagen digital multispectral es una función paramétrica.

Una imagen digital multispectral es una función $g(\vec{r}_b)$, donde \vec{r} es un vector de posición que se encuentra en un espacio vectorial de b dimensiones. El vector \vec{r} es, a su

vez, función de sus coordenadas fila y columna a lo largo de las b bandas, llamadas x e y respectivamente. Esto es:

$$\bar{r}_b = \bar{r}_b(x, y) \quad [1.3]$$

$$g(\bar{r}_b) = g(\bar{r}_b(x, y)) \quad [1.4]$$

El vector \bar{r}_b depende del número de bandas, es decir, del parámetro b , por lo que la función g es una función paramétrica. El espacio vectorial en el que se representa una imagen digital multiespectral es paramétrico también, puesto que su dimensión depende del número de bandas b .

La imagen digital multiespectral es una función que se puede expresar como una suma finita de puntos.

Sea $f(x,y)$ una función bidimensional continua que representa una escena.

$$f(\bar{r}) = f(x, y) \quad [1.5]$$

Utilizando la propiedad de sifting, puede expresarse la función que representa la escena como una suma infinita de puntos:

$$f(x, y) = \iint f(\bar{r}') \delta(\bar{r} - \bar{r}') d\bar{r}' = f(\bar{r}) \quad [1.6]$$

Sean $f(i,j)$ las muestras correspondientes tomadas en las posiciones (i,j) .

$$f(\bar{r}_{ij}) = f(i, j) \quad [1.7]$$

Como ya se ha mencionado anteriormente, la función $g(x,y)$ que representa la imagen es el resultado de aplicar un operador O a la función $f(x,y)$ que representa la escena. Este operador está caracterizando el detector.

$$g(\bar{r}) = O[\iint f(\bar{r}') \delta(\bar{r} - \bar{r}') d\bar{r}'] \quad [1.8]$$

Intercambiando el operador O con los operadores de integración se tiene:

$$g(\bar{r}) = \iint O[f(\bar{r}') \delta(\bar{r} - \bar{r}') d\bar{r}'] \quad [1.9]$$

Si el sistema generador de la imagen funcionará igual no importa adonde se encuentre, el operador O no depende de r' , por lo que la expresión anterior es igual a:

$$g(\bar{r}) = \iint f(\bar{r}') O[\delta(\bar{r} - \bar{r}')] d\bar{r}' \quad [1.10]$$

Se define la función de respuesta al impulso que caracteriza al sistema sensor como:

$$h(\bar{r}, \bar{r}') \equiv O[\delta(\bar{r} - \bar{r}')] \quad [1.11]$$

Entonces, para un sistema lineal e invariante frente al tiempo:

$$g(\bar{r}) = \iint f(\bar{r}') h(\bar{r} - \bar{r}') d\bar{r}' = f * h \quad [1.12]$$

Se está considerando que el espacio en el que se encuentra \bar{r} es discreto, por tanto:

$$g(\bar{r}) = \sum_{m=1}^{\bar{M}} f(\bar{r}') h(\bar{r}, \bar{m}\Delta\bar{r}') \Delta\bar{r}' \quad [1.13]$$

donde \bar{M} es el vector que contiene el tamaño de la imagen digital. De la misma manera que en el caso de la escena, puede expresarse entonces la función discreta que representa la imagen como una suma finita de puntos.

Nota: La función de respuesta al impulso que caracteriza al sistema sensor, si depende de los valores de la escena, se define de la siguiente manera:

$$h(\bar{r}, \bar{r}', f(\bar{r}')) \equiv O[\delta(\bar{r} - \bar{r}'), f(\bar{r}')]]$$

1.1.4 El realce

El realce está formado por el proceso de filtrado y restauración de la imagen. Consiste, entre otras, en las siguientes transformaciones:

- a) Eliminación de ruido: El ruido es un conjunto de valores espúreos que se agregan a la imagen de manera aditiva o multiplicativa. Es producto de un proceso estocástico, pero regido por leyes probabilísticas que están dadas de acuerdo al mecanismo que genera la imagen. El ruido es previsible.
- b) Eliminación de artefactos: Un artefacto es uno (o varios) patrones que ocurren en la imagen de manera azarosa, sin estar asociado ni con la imagen ni con la escena. Puede ocurrir en la generación de la imagen o en alguna transformación aplicada a ésta. No tiene comportamiento estadístico aparente, y por más que lo tuviera, no es previsible.
- c) Modificación del contraste: El contraste es una relación de valor de un pixel o varios de ellos con sus próximos vecinos.
- d) Eliminación de borrosidad.

1.1.5 La segmentación

En términos prácticos, segmentar consiste en reconocer objetos diferentes y separarlos, con fines de una cuantificación ulterior. A continuación se proporciona una definición conceptual.

La segmentación es la partición de la imagen en regiones que pueden tener una relación con la escena: un significado [Bealieu & Goldberg, 1989].

Las regiones en que se particiona la imagen no necesariamente deben ser excluyentes. Pueden tenerse particiones de la imagen en las cuales las clases se encuentran traslapadas. Por este motivo se introducirán a continuación definiciones formales de segmentación con y sin traslape de regiones.

Cabe acotar que en la definición se expresa que las regiones “pueden” tener relación con la escena; esto se debe a que separar objetos de ruido también es segmentar, y el ruido no forma parte de la escena.

I.1.5.1 Segmentación sin traslape de regiones:

La segmentación de una imagen digital $\{g(n,m)\}$ es la división de g en subáreas g_1, g_2, \dots, g_k disjuntas y no vacías que satisfacen un cierto criterio de uniformidad E de acuerdo a las siguientes condiciones[Fu & Mui, 1979]:

a) La unión de las regiones es la imagen original.

$$\bigcup_{i=1}^k g_i = g \quad [1.14]$$

g_i es una región conectada simple para todo i bajo una cierta conectividad. Una región es conectada simple si se pueden tener dos puntos que pertenecen a la región conectados por una línea que también pertenece a la región.

b) El criterio de uniformidad se satisface para toda región.

$$E(g_i) = \text{verdadero} \quad \forall i \quad [1.15]$$

c) El criterio de uniformidad no se satisface para la unión de dos regiones g_i y g_j adyacentes entre sí.

$$E(g_i \cup g_j) = \text{falso} \quad [1.16]$$

d) La suma de las regiones es igual a la imagen original.

$$\sum_i g_i = g \quad [1.17]$$

I.1.5.2 Segmentación con traslape de regiones:

La segmentación de una imagen digital $\{g(n,m)\}$ es la división de g en subáreas g_1, g_2, \dots, g_k no vacías y no necesariamente disjuntas que satisfacen un cierto criterio de uniformidad E de acuerdo a las siguientes condiciones.

a) La unión de las regiones es la imagen original.

$$\bigcup_{i=1}^k g_i = g \quad [1.18]$$

g_i es una región conectada simple para todo i bajo una cierta conectividad.

b) El criterio de uniformidad se satisface para toda región.

$$E(g_i) = \text{verdadero} \quad \forall i \quad [1.19]$$

c) Dada una función de membresía, el criterio de uniformidad puede satisfacerse para la unión de dos regiones g_i y g_j adyacentes entre sí.

$$E(g_i \cup g_j) = \text{verdadero} \quad [1.20]$$

d) La suma de las regiones es mayor que la imagen original.

$$\sum_i g_i > g \quad [1.21]$$

1.1.6 El reconocimiento de patrones

Antes de introducir una definición de Reconocimiento de Patrones veamos qué quieren decir por separado cada uno de los siguientes vocablos: “reconocimiento” y “patrón”.

Reconocer significa conocer de nuevo.

“Conocer” es averiguar atributos de y relaciones entre las entidades y su objetivo. Este último vocablo, tiene una similitud muy grande con la definición de sistema. Conocer es comprender el sistema.

Un patrón es un arreglo de entes de la misma naturaleza.

Es una entidad formada por elementos cuyos valores de atributo pertenecen a los respectivos dominios de los atributos de dicha entidad. El vocablo arreglo da idea de que existe una relación entre los elementos: un orden.

Sea, por ejemplo, la entidad pixel que tiene los siguientes atributos:

- a) Nivel de gris en la banda 1.
- b) Nivel de gris en la banda 2.
- c) Nivel de gris en la banda 3.
- d) Coordenada de la fila
- e) Coordenada de la columna

Y sean los dominios de los atributos mencionados respectivamente los siguientes:

- a) [0, 255]
- b) [0, 255]

- c) [0, 255]
- d) [1, 100]
- e) [1, 200]

Sea $p = (32, 24, 46, 50, 150)$ un elemento "candidato" a pertenecer al conjunto entidad pixel. Los valores que se encuentran entre paréntesis están dados en el mismo orden que los atributos mencionados, y son los datos de p para dichos atributos. Los valores de los atributos pertenecen a los dominios de la entidad pixel:

- 32 \in [0,255] que es el dominio de nivel de gris de la banda 1.
- 24 \in [0,255] que es el dominio de nivel de gris de la banda 2.
- 46 \in [0,255] que es el dominio de nivel de gris de la banda 3.
- 50 \in [1,100] que es el dominio de coordenada de la fila.
- 150 \in [1,200] que es el dominio de coordenada de la columna.

Por tanto, p pertenece al conjunto entidad pixel.

Sean ahora las entidades carretera y bosque que tienen cada una el siguiente atributo:

- a) Nivel de gris

El dominio del atributo nivel de gris de la entidad carretera es el siguiente:

- a) [0, 4]

El dominio del atributo nivel de gris de la entidad bosque es:

- a) [20, 40]

Sean $p = (32)$ y $q = (2)$ dos pixels o instancias "candidatas" a pertenecer al conjunto entidad carretera. El dato 32 no pertenece al dominio del atributo nivel de gris de la entidad carretera, por lo cual el pixel p no es una instancia de entidad carretera. El dato 2 pertenece al dominio mencionado, por lo que el pixel q sí es una instancia de la entidad carretera. El elemento p es una instancia de la entidad bosque.

Nota: Obsérvese que en este ejemplo un pixel es una instancia de entidad, mientras que en el ejemplo anterior, un pixel es una entidad.

Sea, como último ejemplo, la entidad gota que tiene los siguientes atributos:

- a) área
- b) perímetro

Los dominios de los atributos son respectivamente los siguientes

- a) [10, 15]
- b) [25, 40]

La instancia $p = (12, 24)$ no pertenece a la entidad gota, porque el valor del atributo perímetro no pertenece al dominio del atributo perímetro de la entidad gota.

Aquellos patrones que tienen atributos similares, se reúnen a su vez en clases, formando de esta manera, nuevas entidades.

Una clase de patrones es un conjunto de patrones con características similares. La similaridad se define de acuerdo a un cierto criterio.

Se ha mencionado al principio que el concepto de sistema es relativo, puesto que un sistema contiene subsistemas. De la misma manera podemos pensar que lo que es entidad en un sistema, es instancia de entidad en otro sistema. Tal es el caso de los patrones, de los que se ha dicho que son una entidad. Cuando se habla de clases de patrones, los patrones juegan el rol de instancias de entidad, es decir, son elementos de un conjunto de jerarquía superior. De la misma manera, lo que es atributo en un sistema, puede ser dato en otro sistema que contiene al anterior. Lo que es atributo de un patrón es valor de un

atributo(dato) de una instancia de entidad (patrón) de una clase de patrones. En general, para pertenecer una instancia a un conjunto entidad, sus datos deben estar dentro de los dominios de los atributos de la entidad dada. Entonces, dada una clase de patrón, deben definirse los dominios de sus atributos.

Podemos considerar a una clase de patrones como una nueva entidad cuyas instancias son aquellas entidades patrones cuyos atributos son valores de atributos pertenecientes a los dominios de los atributos de la nueva entidad.

Se consideran a continuación dos clases de patrones: los patrones espectrales y los patrones espaciales.

Un patrón espectral es una sucesión de valores de un mismo pixel sobre las diferentes bandas que componen la imagen digital multiespectral.

Ya se ha mencionado que los valores que toma un pixel sobre las diferentes **bandas** son los **niveles de gris** asignados durante el proceso de cuantización. Un pixel tiene una posición dentro de la imagen digital multiespectral, la cual está dada por sus **coordenadas fila y columna**. De manera que, un patrón espectral puede verse como una nueva entidad cuyas instancias son aquellas entidades patrones cuyos atributos son: posición dada por fila y columna, número de banda y nivel de gris. Tal es el caso de la entidad pixel, que es entonces, una instancia de la entidad patrón espectral.

Un patrón espacial está determinado por cualquier arreglo físico de pixels; puede estar compuesto desde uno sólo hasta todos los pixels que componen la imagen, no importando su valor o vecindad con otros.

Un patrón espacial puede ser definido al medirle la siguiente lista de atributos: área, perímetro, rugosidad del perímetro, forma, homogeneidad y textura, entre otros. Una gota, un objeto en la imagen, es una instancia de entidad patrón espacial.

Teniendo ya estas definiciones previas, se introduce a continuación la definición de Reconocimiento de Patrones.

El reconocimiento de patrones es la clasificación de los objetos en función de sus atributos espectrales y morfológicos esenciales.

Reconocer patrones es asignar instancias a entidades en función de los valores de sus atributos. Es buscar estructuras en los datos [Bezdek, 1987]. Es el objetivo último al analizar una imagen, puesto que se obtiene información de la escena.

1.1.7 El modelo de la escena

El modelo de la escena es la salida del sistema de análisis de la imagen.

De la etapa de captura de la imagen se tenía la siguiente expresión para la función que representa a la imagen

$$g(\bar{r}) = \sum_{m=1}^{\bar{M}} f(\bar{r}')h(\bar{r}, \bar{m}\bar{\Delta r}')\bar{\Delta r}' = g(x, y) \quad [1.22]$$

\bar{M} es el vector que contiene el tamaño de la imagen digital.

La función $e(x, y)$ que representa al modelo de la escena es el resultado de aplicar un operador Q a la función $g(x, y)$ que representa a la imagen. Este operador está caracterizando el sistema de análisis digital de imágenes.

$$e(\bar{r}) = Q[g(\bar{r})] \quad [1.23]$$

$$e(\bar{r}) = Q\left[\sum_{m=1}^{\bar{M}} g(\bar{r}')h(\bar{r}, \bar{m}\bar{\Delta r}')\bar{\Delta r}'\right] \quad [1.24]$$

Intercambiando el operador Q con el operadores de sumatoria se tiene:

$$e(\bar{r}) = \sum_{m=1}^{\bar{M}} Q[g(\bar{r}')h(\bar{r}, \bar{m}\bar{\Delta r}')\bar{\Delta r}'] \quad [1.25]$$

Si el sistema de análisis digital de la imagen funciona igual sin importar de qué imagen se trate, el operador Q no depende de r' , por lo que la expresión anterior es igual a:

$$e(\bar{r}) = \sum_{m=1}^{\bar{M}} g(\bar{r}')Q[h(\bar{r}, \bar{m}\bar{\Delta r}')]\bar{\Delta r}' \quad [1.26]$$

Se define la función de respuesta al impulso que caracteriza al sistema de análisis digital de imágenes como:

$$k(\bar{r}, \bar{r}') = Q[h(\bar{r}, \bar{m}\bar{\Delta r}')] \quad [1.27]$$

Entonces, si el sistema es lineal e invariante frente al tiempo:

$$e(\bar{r}) = \sum_{m=1}^{\bar{M}} g(\bar{r}')k(\bar{r} - \bar{r}') = g * k \quad [1.28]$$

Si la función de respuesta al impulso depende de los valores de la imagen, debe usarse la siguiente expresión:

$$k(\bar{r}, \bar{r}', g(\bar{r}')) \quad [1.29]$$

Objetivo del trabajo

Analizar, diseñar y programar un clasificador supervisado para imágenes multispectrales y texturales basado en un algoritmo automatizado de crecimiento de regiones.

I.2.1 La organización formal

I.2.1.1 El Laboratorio de Percepción Remota

El trabajo de tesis se desarrolla en el Laboratorio de Percepción Remota del Instituto de Geofísica de la Universidad Nacional Autónoma de México. Este Laboratorio forma parte de la estructura de investigación del Instituto de Geofísica, desarrollándose en él una línea de investigación llamada "Modelos Matemáticos para el Análisis Digital de Imágenes". La presente tesis forma parte de esta línea, específicamente dentro del marco del proyecto IN102797 intitulado: "Diseño e Implantación de un nuevo clasificador supervisado para imágenes multispectrales y texturales basado en un algoritmo automatizado de crecimiento de regiones".

I.2.1.2 El software existente

En el Laboratorio de Percepción Remota se ha desarrollado un paquete de análisis digital de imágenes llamado SANDI (Sistema de Analisis Digital de Imágenes [Lira & Landgrave, 1992]), el cual corre bajo el sistema operativo MS-DOS versión 6.00 o superior. Es un software de uso didáctico y de apoyo a la investigación. Las rutinas fueron escritas en Turbo Pascal, son totalmente modulares y de propósito general, lo que hace al paquete fácilmente expandible y actualizable.

El paquete SANDI cuenta con imágenes de radar de apertura sintética e imágenes de satélite LANDSAT, SPOT e imágenes médicas, todas sin pre-procesamiento.

I.2.1.3 El hardware existente

El Laboratorio cuenta con tres PC 586, dos de ellas a 200 MHz y una a 120 MHz, cada una con disco duro de 5 GB y 64 MB de RAM, con disketteras de 3 y ½ y CD-ROM, monitor a color con 1280 pixels por 1024 pixels una impresora LaserJet 4P HP, una impresora EPSON Stylus 800 a color, una impresora Photo Smart de Hewlet Packard de calidad fotográfica y una estación de trabajo SUN Sparc Startion 5, con 5 GB de discos duros, 64 MB de RAM, CD-ROM, diskettera de 3 y ½ y lectora de cinta magnética.

I.2.1.4 El marco teórico

Se introducirán a continuación distintos grupos de clasificadores existentes, mencionándose sus características principales.

I.2.1.4.1 La clasificación booleana y la clasificación difusa

Clasificadores booleanos

En un clasificador booleano una instancia pertenece o no pertenece a una entidad, es decir, un pixel pertenece o no pertenece a una clase. El grado de membresía es cero o uno.

Una vez que se definan las regiones, todos los pixels quedarán clasificados, es decir, la unión de las clases es igual a la imagen. Poniendo algún parámetro extra, pueden dejarse pixels sin clasificar.

Todos los pixels pertenecen a una clase y sólo una.

Clasificadores difusos

En un clasificador difuso un pixel pertenece a una clase con una cierta probabilidad. Esta probabilidad se modela a través de una función de membresía. La membresía de un pixel a una clase dada no es cero o uno, sino un valor real en el rango $[0,1]$.

Un pixel puede tener diferentes grados de membresía a varias clases. Si se clasifican todos los pixels, algunos quedarán en más de una clase, por lo que la suma de las clases será mayor que la imagen. Si aquellos pixels que pertenecen a más de una clase se dejan sin clasificar, la suma de las regiones será menor que la imagen.

I.2.1.4.2 La clasificación puntual y la clasificación contextual

La clasificación puntual

Los clasificadores puntuales etiquetan los pixels de la imagen en una clase usando solamente atributos espectrales e ignoran el contexto [Jensen, 1986; Gong, 1992].

Este enfoque, también denominado clasificación por pixel, demostró ser de naturaleza limitada y aplicable solamente a casos espectralmente bien diferenciados.

La clasificación contextual

En los últimos años comenzaron a estudiarse clasificadores contextuales, a los cuales se les llama también clasificadores del entorno.

En este enfoque, un pixel es etiquetado en una clase tomando en cuenta sus propiedades espectrales y el contexto de su ubicación.

Comparación entre ambos clasificadores

Un clasificador contextual consistentemente produce mayor exactitud en la clasificación que un clasificador puntual. La ventaja que tiene el clasificador contextual es que para cada pixel cuenta con mayor información para tomar la decisión de en qué clase clasificarlo, puesto que no sólo cuenta con su información puntual sino con la información del entorno de dicho pixel. Y a mayor información, menor la incertidumbre.

I.2.1.4.3 La clasificación no supervisada y la clasificación supervisada.

La clasificación no supervisada

En esta clasificación primero se segmenta la imagen y luego se etiquetan las clases: primero se asignan las instancias a conjuntos, y luego se define qué entidades son esos conjuntos.

Los pasos básicos de la clasificación no supervisada son los siguientes [Chubieco, 1990]:

- . Seleccionar las bandas que intervienen en la clasificación, ya sea originales o fruto de alguna transformación.
- . Seleccionar un criterio para medir la similitud o distancia entre pixels que componen la imagen.
- . Seleccionar un criterio para agrupar pixels similares.
- . Generar un mapa temático.

La clasificación supervisada

En esta clasificación primero se etiquetan las clases y luego se segmenta la imagen: primero se definen las entidades con sus atributos y los dominios de los mismos, y luego se asignan las instancias a las entidades.

Los pixels de la imagen multiespectral son asignados a clases de las cuales se tiene conocimiento a priori a través de lo que se conoce como pixels prototipo. Para ello se necesita hacer una interacción con la imagen y marcar los polígonos que contienen pixels representativos de la clase.

Los pasos básicos de la clasificación supervisada son los siguientes [Richard, 1993]:

- . Decidir el conjunto de clases en las cuales se va a segmentar la imagen.
- . Escoger en forma interactiva los pixels representativos o prototipo de cada clase.
- . Utilizar los pixels prototipo para determinar los parámetros del clasificador.
- . Emplear el clasificador calibrado (con los parámetros determinados), para etiquetar cada pixel de la imagen.
- . Generar un mapa temático.

I.2.2 La definición del problema

I.2.2.1 Signos y síntomas del problema

Ya se ha mencionado que un clasificador supervisado emplea información a priori de cada clase determinada: esto se hace usualmente por medio de conjuntos de entrenamiento o pixels prototipo. Los conjuntos de entrenamiento se definen a través de polígonos cerrados marcados por algún procedimiento interactivo en la imagen. En esta definición no hay un criterio claro para asumir que los conjuntos de entrenamiento sean muestras estadísticas válidas de las clases.

Entre las premisas básicas para un clasificador se enuncia, primero, que debe establecerse un procedimiento que asegure que los conjuntos de entrenamiento sean

muestras representativas de las clases. Segundo, la clasificación de un pixel debe llevarse a cabo por comparación directa entre el comportamiento estadístico de las clases y aquel de la vecindad del pixel.

En los clasificadores no supervisados el control que se tiene sobre los resultados no sólo está dado por la exactitud en los parámetros especificados que en muchos casos son difíciles de estimar. Por otro lado, existen problemas en los cuales es necesario segmentar sólo una parte de la imagen que contiene un grupo de clases de interés. La región que no es de interés conforma una región muy heterogénea cuyos parámetros son difíciles de estimar.

En resumen, se requiere de un clasificador supervisado cuyos parámetros sean sencillos de especificar, que se ingresen interactivamente, y que sean suficientes como para iniciar un proceso de generación de conjuntos de entrenamiento que contengan muestras estadísticamente válidas de las clases, y proceder luego a la clasificación de los pixels.

1.2.2.2 La propuesta

Se desea implantar un clasificador supervisado en el cual los conjuntos de entrenamiento sean muestras estadísticas de las clases delineadas por un algoritmo automatizado de crecimiento de regiones, y clasificar los pixels en función de las estadísticas obtenidas a partir de dichas muestras. Se estudiarán distintos criterios de segmentación.

1.2.3 Partes del sistema de modelado de la escena

Se enuncian a continuación las partes que se pretende tenga el clasificador supervisado. Se asume que la imagen digital multiespectral a utilizar ya pasó por la etapa de realce, de modo que el clasificador comprende exclusivamente los procesos de segmentación y reconocimiento de patrones.

En primer lugar, se describen las entradas al clasificador, en segundo lugar, los procesos a realizar y por último las salidas del sistema. Aspectos de los procesos a realizar se profundizarán en el capítulo II.

1.2.3.1 Entradas

Se detallan a continuación las siguientes entradas: la imagen, las entidades relevantes, los atributos relevantes y los parámetros de acotación. La imagen juega el rol de universo. Las entidades relevantes se definen a través de una instancia relevante denominada pixel prototipo. Los atributos relevantes están dados por un criterio de homogeneidad que se debe especificar. Los parámetros de acotación se emplean en el cálculo de parámetros que produce el algoritmo de segmentación.

1.2.3.1.1 La imagen: el universo

Una imagen digital multiespectral contiene información espectral, contextual y textural relacionada a la escena de interés. Pero ya no se tiene la escena propiamente dicha,

por lo que en este punto, la imagen digital multiespectral juega el rol de universo que contiene al sistema que se desea modelar. Este sistema es un conjunto de clases con atributos esenciales que un poco más adelante se enumerarán. La instancia de menor jerarquía que se considera es un pixel, el cual es un elemento del universo que se deberá asignar a una de las entidades o clases. Estas clases tienen entre sí una relación de distancia que se desea cuantificar.

Para imágenes ópticas el detalle de información depende de:

- número de bandas espectrales
- tamaño del pixel
- número de niveles de cuantización
- relación señal ruido.

Para imágenes de radar, el detalle de información depende de:

- tamaño del pixel
- polarización
- ancho de banda
- aspecto geométrico
- relación señal-ruido.

I.2.3.1.2 Las entidades relevantes

Se desea especificar cuáles son las clases a reconocer de la manera más sencilla posible, esto es, señalando para cada clase a uno de sus elementos o pixels prototipos. Se supone que estos pixels satisfacen hipótesis que se enuncian a continuación.

La instancia prototipo es un elemento del universo que cumple con los requisitos para pertenecer a un conjunto entidad, puesto que sus datos se encuentran dentro de los dominios de los atributos de la entidad considerada, y además, en un radio suficientemente grande dicha instancia está rodeada de instancias de la misma entidad a la que pertenece.

En el párrafo anterior identificamos dos hipótesis. La primera nos dice que el pixel prototipo de una clase es un elemento de dicha clase y la segunda, que en un radio suficientemente grande, sus vecinos son de su misma clase. Con esto se garantiza que muy cerca del pixel prototipo, la energía de la clase es muy grande, puesto que la interferencia de instancias de otras entidades es mínima. Existe un radio para el cual la energía asociada al pixel prototipo se reduce lo suficiente como para poder establecer el comienzo de la interferencia de otras entidades. Por la naturaleza digital de la imagen, en vez de un círculo alrededor del pixel prototipo, se emplea una **ventana**, término que se define a continuación:

Sea $V(y_{ini}, x_{ini}, y_{fin}, x_{fin})$ una ventana de una imagen g , donde:

y_{ini} es la coordenada de la fila de su esquina superior izquierda,

x_{ini} es la coordenada de la columna de su esquina superior izquierda,

y_{fin} es la coordenada de la fila de su esquina inferior derecha y

x_{fin} es la coordenada de la columna de su esquina inferior izquierda.

La ventana V es un subconjunto de la imagen g formado por todos aquellos pixels $p(x,y)$ pertenecientes a la imagen g tales que sus coordenadas pertenecen a los dominios determinados por los extremos de V .

$$p(x,y) \in V(x_{ini}, y_{ini}, x_{fin}, y_{fin}) \subset g \Leftrightarrow p(x,y) \in g \wedge x \in [x_{ini}, x_{fin}] \wedge y \in [y_{ini}, y_{fin}]$$

Con el objeto de justificar la segunda hipótesis, se muestra en Figura 1.7 tres gráficas. En la primera, se calculó una función de densidad $f(x)$ con datos extraídos de una ventana de 32×32 pixels sobre una clase A. En la segunda figura, se calculó la misma función $f(x)$ con datos de una ventana del mismo tamaño, ubicada sobre la frontera entre dicha clase A y una clase B. Por último, se tiene la gráfica de la función $f(x)$ calculada para una ventana de 32×32 pixels sobre la clase B. Claramente puede apreciarse que la energía de las clases A y B disminuye en la frontera, es decir, cuando ambas clases comienzan a interferirse mutuamente.

En términos de operación del sistema, estas hipótesis implican que se está aceptando que el pixel prototipo estará correctamente colocado. La validación de su especificación se hará únicamente mediante apreciación visual.

Las entidades tienen el mismo nivel de jerarquía y no existen entidades mellizas.

Con esta hipótesis se desea garantizar la separabilidad entre las entidades. Existe la necesidad de reducir la incertidumbre en la asignación de las instancias a las entidades consideradas como relevantes. Para ello, es necesario que las entidades tengan la máxima separabilidad, es decir, el mínimo traslapamiento: se requiere que éstas tengan el mismo nivel de jerarquía. Por ejemplo: Sería incorrecto definir para un mismo problema, la entidad vegetación y la entidad pinos, puesto que la segunda está incluida en la primera. [Wookcock & Harwar, 1991, hablaban de jerarquías.]

Que no existen entidades mellizas significa que los pixels prototipo son de clases diferentes.

La imagen es una realización de un campo aleatorio definido jerárquicamente en términos de un campo aleatorio subyacente. El campo aleatorio subyacente es un proceso que genera las regiones según atributos comunes (o atributos relevantes) cuya combinación de dominios es diferente para cada clase.

Con esta hipótesis se asegura, en primer lugar, que la imagen es segmentable, es decir, que las regiones existen. En segundo lugar, se dice que las clases comparten por lo menos un atributo que permite distinguirlas. En este trabajo se considera que las entidades presentes en la imagen son clases espectrales o clases texturales, pero no ambas a la vez. Con base en esta hipótesis, los criterios de homogeneidad que se utilicen para reconocer patrones están fundamentados en conjuntos de atributos espectrales o conjuntos de atributos texturales, pero no una combinación de ambos. La clasificación sí es contextual, pero no es espectral y textural a la vez. Los atributos mencionados se describirán en I.3.1.3 y esta hipótesis se enunciará formalmente en esa sección bajo el nombre de "hipótesis del atributo relevante".

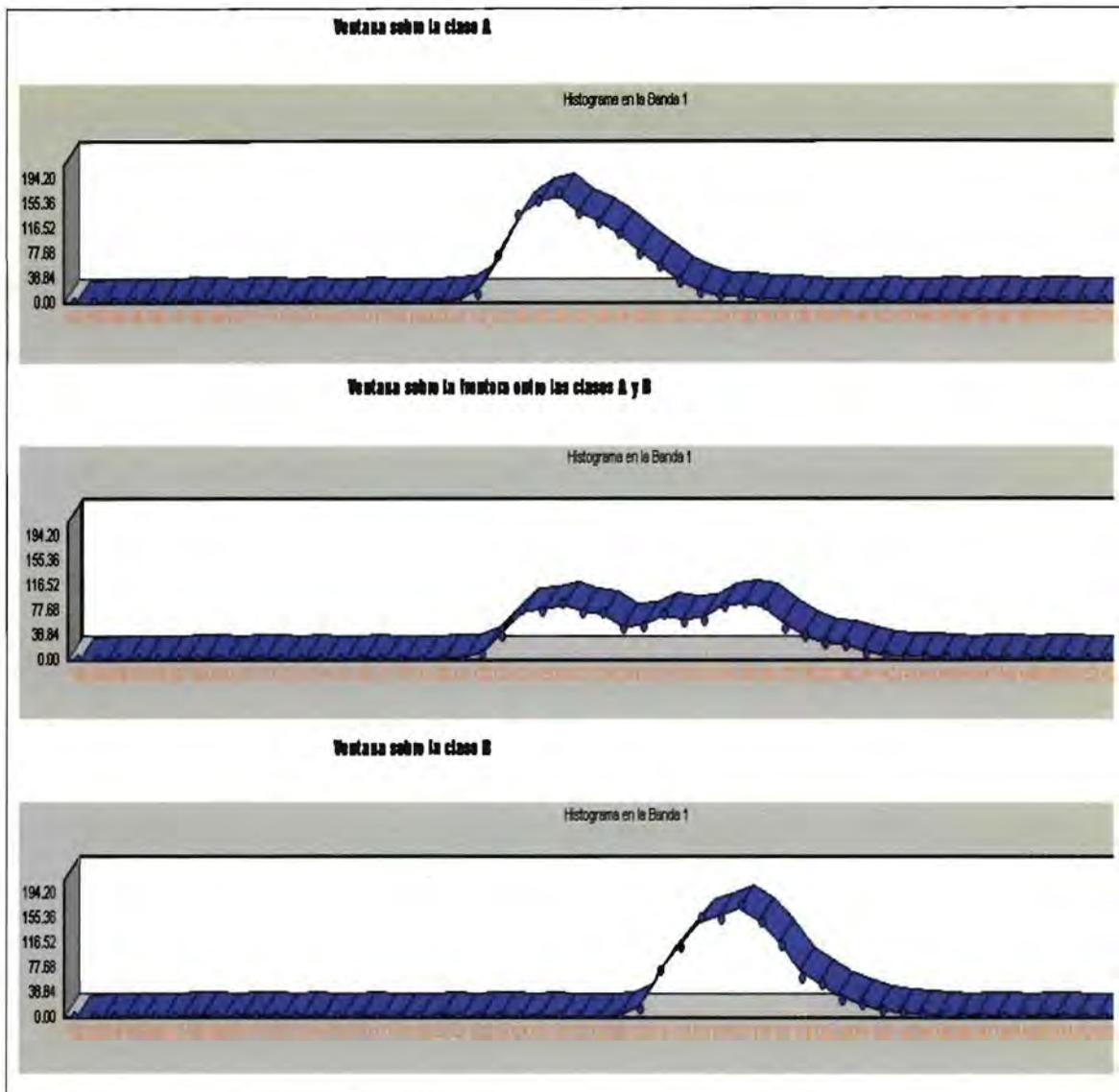


Figura 1.7

I.2.3.1.3 Los atributos relevantes:

Los atributos relevantes intervienen implícitamente en una función que se denomina “criterio de homogeneidad”, la cual será descrita en el próximo capítulo. Aquí, nada más, se manifiesta que al seleccionar un criterio de homogeneidad, se están determinando aquellos atributos que se consideran relevantes durante el proceso de modelización de la escena, y, por lo tanto, en algún momento los dominios de dichos atributos deben quedar definidos, para poder efectuar la asignación de las instancias a cada una de las entidades

relevantes especificadas por medio de la instancia prototipo. En esta parte se procede a enunciar formalmente la hipótesis del atributo relevante.

Hipótesis del atributo relevante

Sea $z = \{z(i, j)\}$ definida sobre el dominio $L = \{i, j / 1 \leq i \leq N, 1 \leq j \leq M\}$ una imagen específica con t regiones. Se supone que \bar{z} es una realización de un campo aleatorio bidimensional definido jerárquicamente en términos de un campo aleatorio subyacente $\bar{x} = \{x(i, j)\}$.

Se supone que la imagen está formada por distintas regiones, donde $z = \{z(i, j)\}$ representa la partición del dominio L en regiones de distintos tipos. El campo subyacente x es el proceso que está generando las distintas regiones. Es subyacente porque está jerárquicamente debajo de la imagen digital multispectral. Cada pixel $x(i, j)$ deberá ser asignado a una de las t regiones.

Cada variable aleatoria $x(i, j)$ toma valores en el conjunto de etiquetas $\chi = \{1, 2, \dots, t\}$ donde $x(i, j) = t$ indica que la coordenada (i, j) de un pixel pertenece a la región t . En cada región de tipo t se asume la presencia de una muestra particular de la región h_t que corresponde a la realización de un campo aleatorio discreto y homogéneo H_t . El campo aleatorio que representa las imágenes con las regiones se expresa entonces como:

$$\begin{aligned} \forall t \in \chi, z(i, j) &= H_t(i, j) \\ \forall (i, j) / x(i, j) &= t \end{aligned}$$

Es decir, a los pixels se la región se les cambia de valor, etiquetándolos como t . $z(i, j) = H_t(i, j)$ ya es la segmentación; z se parte en las regiones que se denominan H_t .

Se supone que H existe y es el atributo relevante.

Se enumeran a continuación tres conjuntos de atributos que pueden utilizarse para describir aspectos esenciales de las clases como entidades. En primer lugar, se mencionan atributos espectrales, en segundo lugar, atributos contextuales y por último, una combinación de ambos.

1.2.3.1.3.1 Atributos espectrales:

Nivel de gris y banda

El nivel de gris es una relación con el nivel de cuantización asignado a un pixel en una banda dada.

Se establece una relación entre el nivel de gris, que es pictórico y el de cuantización que es numérico numérico. El número de niveles de cuantización se corresponde con el número de niveles de gris. El dominio del atributo nivel de gris es el intervalo $[0, \text{número de niveles de gris}-1]$.

Cada nivel de gris se representa con una cantidad fija de bits. Generalmente se utiliza una cantidad de bits igual a 8 ($2^8 = 256$), teniéndose el rango de valores: [0, 255].

I.2.3.1.3.2 Atributos contextuales:

El histograma

El histograma es la distribución de frecuencias en una parte o toda la imagen de los valores de pixels en una banda dada[Jain, 1989].

El histograma es una función que depende de los niveles de gris; el dominio de dicha función es el intervalo [0, número de niveles de gris-1]. El número total de ocurrencias toma valores entre cero y el tamaño de la imagen.

Si se calcula un histograma de ocurrencia conjunta de valores de pixels a través de las bandas, el histograma opera además como atributo espectral, y su dominio está determinado también por el número de bandas.

El histograma tiene una connotación probabilística porque nos dice qué tan frecuentemente ocurre un valor de pixel [Richard, 1986]. Por lo tanto, al ser normalizado dividiendo cada frecuencia de ocurrencia de un nivel de gris por el tamaño de la imagen, se obtiene una función de densidad de probabilidad. Toda función de densidad tiene un área bajo la curva igual a uno.

El histograma es un descriptor global de la brillantez y el contraste de una imagen. Cuando los valores de los niveles de gris están muy concentrados cerca de 255 se interpreta que la imagen es muy brillante y tiene poco contraste. Cuando los valores de los niveles de gris están muy concentrados cerca de 0, se interpreta que la imagen es muy oscura y tiene poco contraste. Cuando los valores de los pixels están distribuidos a lo largo de todos los niveles posibles, la imagen tiene buen contraste y es de brillo regular.

La textura

La textura es una manifestación espacial organizada de un conjunto de pixels de la imagen.

Intuitivamente esto significa que si se toma una banda de una imagen digital multispectral y se grafican los valores de los pixels en sus respectivas posiciones se obtiene una superficie. Si esta superficie es accidentada, su textura es gruesa. Si esta superficie es suave, su textura es uniforme.

Algunos autores hablaban de tonos en vez de pixels [Gurney & Townshend, 1983].

Las características básicas de la textura son las siguientes:

- Un cierto orden local se repite sobre una región que es relativamente grande comparada con el tamaño del orden. El orden local implica que se tiene un modelo matemático que aplicado sobre una subregión presenta las mismas características cuantitativas que otra subregión.

- El orden consiste en un cierto arreglo de partes elementales. Es decir que los pixels están arreglados de una cierta manera, con una cierta estructura geométrica.

- Las partes están compuestas por entidades aproximadamente uniformes teniendo más o menos las mismas dimensiones dentro de la región de textura. Esto es cuantitativo, se sabe que son entidades aproximadamente uniformes hasta tanto se haya aplicado el modelo matemático. Será uniforme cuando los valores sean cercanos entre sí dentro de cierto umbral.

- Las partes pueden ocurrir en cualquier dirección. Con esto se asegura la invariancia frente a rotaciones.

- La textura depende de la escala a la cual se observa.

Un clasificador textural [Hsv, 1977; Haralick, 1979; Woodcock, 1980] tiene que se contextual por naturaleza. En el presente trabajo, el clasificador textural considera a las texturas como clases o entidades.

La matriz de co-ocurrencia

Para cuantificar la textura se emplea la función de densidad conjunta o matriz de co-ocurrencia [Haralick, 1973], la cual se define formalmente a continuación.

Se considera un operador vectorial de posición $\bar{\beta}$ que establece la posición relativa que ocupan un par de pixels λ_p y λ_q sobre una ventana o subregión δ para una banda dada. Debido a que en una imagen digital hay ocho posibles direcciones, $\bar{\beta}$ puede asumir estas posibles direcciones. El tamaño de $\bar{\beta}$ puede ser cualquiera, pero la función de autocorrelación es muy pequeña más allá de los ocho primeros vecinos y por ello, no es necesario que su módulo sea más grande que bajo conectividad ocho.

Para una $\bar{\beta}$ y una δ dadas ($\bar{\beta}$ dado tanto en dirección como en magnitud) se cuentan el número de pares de pixels (λ_p, λ_q) y se construye una matriz M cuyos elementos tienen un valor igual al número de pares. La posición de estos elementos es la definida por los valores de los pixels. La dimensión de M es igual al rango posible de valores que pueden asumir los pixels. Se cuenta el número total de pares de pixels que ocurren en δ y se divide cada elemento de M por este número de pares; el resultado es una matriz llamada de co-ocurrencia C.

En la matriz de co-ocurrencia C los valores pertenecen al intervalo [0,1]. Si los elementos de C son $c(i,j)$ se tiene que:

$$0 \leq c(i,j) \leq 1 \quad \forall(i,j)$$

De acuerdo a la definición anterior, los elementos de la matriz de co-ocurrencia están dados entonces por:

$$c(\lambda_p, \lambda_q, \bar{\beta}, \delta) = \frac{\vartheta\{\bar{r}/\bar{r} + \bar{\beta} \in \sigma(\delta), g(\bar{r}) = \lambda_p, g(\bar{r} + \bar{\beta}) = \lambda_q\}}{\vartheta\{\bar{r}/\bar{r} + \bar{\beta} \in \delta\}}$$

donde $\sigma(\delta)$ es una isometría de traslación que implica que el operador vectorial de posición $\bar{\beta}$ se traslada por la ventana δ con la condición de que $\bar{r} + \bar{\beta}$ esté dentro de la ventana, siendo \bar{r} el vector de posición del pixel y ϑ es el orden o número de elementos del conjunto.

La matriz de co-ocurrencia es una cuantificación de la definición de textura, aunque de por sí no genera un escalar [Haralick, 1979] asociado al contenido de textura de una ventana. Para generar tal escalar se emplea un descriptor de la matriz de co-ocurrencia. A continuación se proporciona una lista de los principales descriptores.

La probabilidad máxima

La probabilidad máxima es el mayor valor de la matriz de co-ocurrencia.

La probabilidad máxima es un descriptor que se obtiene a partir de la información textural presente en la matriz de co-ocurrencia y se obtiene buscando el máximo valor de la misma.

$$P_m = \max(p(i,j))$$

La entropía

La entropía cuantifica el grado de orden o desorden que tiene un sistema físico cerrado.

La entropía es uno de los mejores descriptores que cuantifican el contenido de información textural presente en la matriz de co-ocurrencia, puesto que tiene base física. A mayor desorden, mayor entropía.

En un sistema ordenado se tiene una entropía baja y esto se asocia con una textura uniforme. En un sistema desordenado, en cambio, se tiene una entropía alta, que se asocia con una textura heterogénea.

$$\varepsilon = - \sum_i \sum_j c(i,j) \log[c(i,j)]$$

La homogeneidad

La homogeneidad cuantifica el nivel de heterogeneidad de la matriz de co-ocurrencia.

La homogeneidad toma un valor menor cuanto menos celdas de la matriz de co-ocurrencia tienen valores mayores que cero, lo cual significa que la textura es más homogénea. A medida que aumenta la cantidad de celdas con valores mayores que cero en

la matriz de co-ocurrencia, el valor de esta variable aumenta, reflejando una mayor heterogeneidad de la textura.

$$H = \sum_i \sum_j c(i, j)c(i, j)$$

El contraste

Es la relación de valor de un pixel con sus próximos vecinos.

Puede calcularse a partir de la matriz de co-ocurrencia mediante la siguiente expresión:

$$C_N = \sum_i \sum_j (i - j)(i - j)c(i, j)$$

I.2.3.1.3.3 Los atributos espectrales – contextuales:

En los atributos espectrales-contextuales se asume que un pixel se encuentra en un contexto espectral.

Los atributos espectrales-contextuales tienen asociados vectores de dominios de tantas dimensiones como bandas tiene la imagen digital multispectral y sus modelos matemáticos se aplican sobre una ventana o región δ de la imagen para cada una de sus b bandas.

Sea α un atributo espectral-contextual calculado sobre una ventana o región δ de una imagen digital multispectral de b dimensiones. Para cada una de las b bandas de la imagen, el atributo α toma un conjunto de valores, es decir, tiene una cantidad de b dominios $D_i(\alpha)$ asociados. Estos dominios se expresan vectorialmente de la siguiente manera:

$$D(\bar{\alpha}) = [D_1(\alpha), D_2(\alpha), \dots, D_b(\alpha)]$$

Algunos atributos contextuales citados pueden emplearse de esta manera espectral, y otros más como ser la media, el desvío estándar, la media de correlación cruzada, el desvío estándar de correlación cruzada, entre otros.

I.2.3.1.3.4 La definición implícita de los atributos relevantes a través de los criterios de homogeneidad

Ya se ha señalado que al especificar un criterio de homogeneidad, implícitamente se determinan los atributos que se desean utilizar para asignar las instancias a las entidades o clases correspondientes. Se listan a continuación los criterios de homogeneidad estudiados, con los respectivos atributos implícitos en los mismos. El capítulo II tratará con detalle estos aspectos.

- * La diferencia de medias
 - Media espectral.
 - Desvío estándar espectral.
- * El producto escalar de medias
 - Media Espectral
- * Medias de correlación cruzada.
 - Media espectral de correlación cruzada.
 - Devío standard espectral de correlación cruzada.
- * El producto escalar de medias de correlación cruzada.
 - Media espectral de correlación cruzada.
- * La diferencia de histogramas conjuntos normalizados.
 - Histograma conjunto normalizado.
 - Media espectral.
 - Desvío estándar espectral.
- * La máxima correlación de histogramas normalizados.
 - Histograma conjunto normalizado.
 - Media espectral.
 - Desvío estándar espectral.
- * La diferencia de matrices de coocurrencia .
 - Matriz de coocurrencia.
- * La diferencia de matrices de coocurrencia en sus centros de masa.
 - Matriz de coocurrencia.
 - Centro de masa o media de la matriz de coocurrencia.
 - Desvío estándar de la matriz de coocurrencia.

I.2.3.1.4 Los parámetros de acotación.

El clasificador supervisado necesita de parámetros de acotación para poder calibrarse. La calibración consiste en la estimación de parámetros con datos propios de la imagen para poder iniciar el proceso de entrenamiento, esto es, la generación del conjunto de muestras de cada una de las clases. Los parámetros de acotación constituyen el factor de acotación a ser empleado en el umbral del sistema de modelado y el umbral de reducción suficiente de la energía local.

El factor de acotación η a aplicar al umbral del sistema de modelado

Este factor, según el criterio de homogeneidad empleado, toma distintos rangos de valores; en todos los casos es mayor que cero. Existe una relación directa entre el factor de acotación y el tamaño de los dominios de los atributos implícitamente utilizados en el criterio seleccionado. A mayores valores, mayores los dominios de los atributos de las clases. Por ejemplo, que el factor de acotación tenga el valor 1.5 en un algoritmo de diferencia de medias, implica que el atributo media de cada clase tendrá como dominio una media inicial más menos 1.5 del desvío estándar inicial.

Se asume que las clases están separadas, y que si hay traslapamiento, este es mínimo. Pero no se asume nada respecto de qué tanto están dispersos los valores de sus atributos.

Con este factor se puede regular la dispersión de los valores de los atributos de las clases. El factor de acotación tiene una relación directa con la dispersión de los valores de los atributos de las clases. A mayores valores del factor de acotación, mayores los dominios de los atributos de las clases y a valores mayores aún, mayor traslapamiento entre los dominios de los atributos de las clases.

Según el criterio de homogeneidad seleccionado, este factor puede afectar a valores tomados a partir de una función de energía asociada a la clase, o a valores obtenidos a partir de una función de separabilidad asociada a todas las clases. Sin embargo, en sendos casos, su efecto es el mismo: acotar los dominios de los atributos de las clases.

El umbral de reducción suficiente de la energía local en la derivada de la función objetivo o parámetro de estabilidad Ω .

Debido a que no se ha introducido hasta este momento la función objetivo, se completará en el segundo capítulo el concepto de este parámetro.

I.2.3.2 El proceso:

El proceso está constituido por el conjunto de funciones que transforman las entradas en salidas. En primer lugar se enuncia su objetivo, y en segundo lugar, las etapas que lo componen.

I.2.3.2.1 Objetivo:

Establecer cuáles son las poblaciones de las clases espectrales o texturales y obtener los valores de sus características esenciales o descriptores de acuerdo a un modelo.

I.2.3.2.2 Etapas:

Calibración o ajuste del sistema de modelado

En esta primera etapa se estiman dos parámetros para cada una de las clases a partir de información relacionada con sus pixels prototipo: la ventana óptima y el umbral óptimo. Con ellos se tiene una primera estimación de los parámetros de las clases.

En la etapa de calibración el problema consiste en: dadas las clases a través de su pixel prototipo y dado un atributo para clasificar los pixels (denominado función criterio de homogeneidad), se desea definir el dominio de este atributo para cada clase.

Para ello, se necesita contar con información de la clase y opcionalmente de la relación entre las clases. La información de una clase se extrae del contexto del pixel prototipo, cuyo tamaño se tiene que estimar: ésta es la ventana óptima. Teniendo información de cada clase se está en condiciones de definir los dominios de cada una de

ellas. Hasta este punto, un pixel pertenece a una clase si su valor para el atributo utilizado como función criterio de homogeneidad pertenece al dominio correspondiente a dicha clase.

Debido a que los dominios de los atributos de las clases pueden estar traslapados se efectúa una interacción entre las clases a través de una función de separabilidad que se obtiene a partir de la función criterio de homogeneidad y se modifican los dominios a rangos lo menos superpuestos posible. Con esto se evita en la etapa de muestreo siguiente el agregado de pixels dudosos que pudieran alterar la estimación de los parámetros óptimos de las clases.

La ventana óptima se calcula para poder extraer información contextual relacionada a cada pixel prototipo. Se utiliza una función de energía para determinar para cada clase el menor radio de mínima interferencia por otras clases en base al cual se obtiene información contextual suficiente de la clase. Por la naturaleza digital de la imagen, en vez de círculos generados a partir de dichos radios, se trabaja con cuadrados o ventanas, por lo que de aquí en adelante se empleará este último vocablo.

El umbral óptimo es un parámetro que determina el tamaño del dominio del atributo que se utiliza como función criterio de homogeneidad. Este parámetro depende de la ventana óptima, de una función de separabilidad que se obtiene a partir de la función criterio de homogeneidad o de la función criterio de homogeneidad propiamente dicha, y del factor a emplear en el umbral. Se calcula a partir de datos propios de cada clase sin o con interacción entre clases. La interacción entre las clases se efectúa por medio de una función de separabilidad.

Entrenamiento

Hasta este punto, cada clase es un conjunto unitario, puesto que su único elemento es el pixel prototipo. Una muestra permite generar una aproximación cruda de las estadísticas de una clase, pero no es suficiente, por lo que es necesario incorporar más elementos a cada una de las mismas. Se requiere de una herramienta de muestreo que sea capaz de reconocer aquellos pixels que pertenecen a determinada clase y de incorporarlos al conjunto de muestras de entrenamiento para dicha clase. El crecimiento de regiones es una herramienta [Zucker, 1976] que satisface este requerimiento porque evalúa para cada pixel vecino a una región dada un criterio de homogeneidad. Subsecuentemente este procedimiento incorpora dicho pixel a la región y se continúa hasta que el criterio de homogeneidad no se satisface para ningún pixel vecino a la misma.

Una vez finalizado el crecimiento de regiones se tiene una muestra estadísticamente válida para cada clase y de esta manera, una aproximación de las estadísticas de las clases. Se redefinen los dominios de los atributos relevantes especificados a través del criterio de segmentación seleccionado y con esto se procede a iniciar la tercera etapa.

En la etapa de entrenamiento el problema consiste en: dadas las clases a través de un conjunto de muestras cuyos valores de atributo pertenecen a los respectivos dominios correspondientes a dichas clases, y dado el atributo que se va a utilizar para clasificar los pixels (denominado función criterio de homogeneidad), se propone redefinir el dominio de este atributo para cada clase.

Obtención del modelo de la escena

Como ya se han ajustado los dominios de los atributos relevantes especificados, en esta etapa se está en condiciones de generar las poblaciones de las clases y sus respectivos descriptores. Los descriptores son aspectos cuantificables del comportamiento de la escena. Con esto se tiene una representación de la escena.

En la etapa de obtención del modelo de la escena el problema consiste en: dados para cada clase los dominios del atributo utilizado como función criterio de homogeneidad, clasificar los pixels y determinar, entre otros descriptores, el valor de este atributo para cada clase.

I.2.3.3 Las salidas

Se desea obtener un modelo de la escena, donde queden definidas las clases espectrales o texturales, los atributos relevantes y las relaciones entre las clases. Para ello se necesita conocer, en primer lugar, cómo están organizados y dónde están ubicados los objetos en la imagen. Un mapa temático proporciona esta información. En segundo lugar y a partir de dicho mapa, se desea obtener un conjunto de estadísticas referidas a dichos objetos, las cuales se calculan por medio de descriptores.

I.2.3.3.1 El mapa temático

Es el producto final de la segmentación y es una herramienta útil para comprender la escena. La segmentación es la partición de la imagen en regiones, cada una de las mismas relacionadas a un patrón espacial de la escena. Las regiones pueden ser etiquetadas como pertenecientes a cierta clase de objetos, generando una clasificación y, en consecuencia, un mapa que muestre la organización y ubicación de las clases.

I.2.3.3.2 Los descriptores

Por medio del mapa temático se pueden obtener aspectos cuantificables del comportamiento de la escena, los cuales se denominan descriptores [Lira, 1997]. A continuación se enumeran los siguientes descriptores: área, perímetro, compacidad por compactación, rugosidad del perímetro, dimensión fractal, excentricidad y diversidad o estructura del paisaje. Cabe acotar que los atributos mencionados en párrafos anteriores también pueden operar en este punto como descriptores, ya que teniendo las clases con sus poblaciones completas, dichos atributos de naturaleza variable se convierten en variables de abstracción fija o parámetros o descriptores de las clases.

El área

El área A de un objeto es el número de pixels que lo forman.

Debe conocerse el área del pixel, es decir, tener información de la escala a la que se generó la imagen correspondiente.

El perímetro

Debe definirse primero una conectividad.

Bajo conectividad 4, *el perímetro P de un objeto es el número de lados que forman la frontera*; el resultado es una cantidad entera.

Bajo conectividad 8, *el perímetro P de un objeto es la distancia de los centros entre los pixels que forman la frontera*; el resultado es una cantidad real.

La compacidad por compactación

La compacidad por compactación de un objeto es el promedio de las distancias más cortas de cada uno de sus puntos a la frontera.

Se parte de la premisa que lo más compacto que puede estar un objeto es un círculo y que la compacidad de un círculo es igual a uno. Para el espacio continuo se tiene la siguiente fórmula:

$$c = \frac{P^2}{2\pi A}$$

donde:

P = perímetro del objeto

A = área del objeto

Como la fórmula continua no es buena para el espacio discreto, porque nunca se tiene un círculo realmente, se propone la siguiente variante:

$$C_d = \frac{\sum_i d_{\pi^i}}{A}$$

donde:

d_{π^i} = distancia más corta del pixel i-ésimo del objeto a la frontera de éste.

A = área del objeto.

Este atributo no es un buen descriptor de forma, puesto que genera ambigüedades.

La rugosidad del perímetro

La rugosidad del perímetro de un objeto está dada por la estructura de un histograma de cambios angulares entre pixels vecinos pertenecientes a su perímetro.

Debe definirse primeramente una conectividad. La curvatura del perímetro es el cambio angular que hay entre dos pixels vecinos. Con estos cambios angulares se construye un histograma. La rugosidad del perímetro está dada por la estructura de este histograma. Cuantas más barras tenga el histograma más rugoso es el perímetro del objeto. Cuantas menos barras tenga el histograma menos rugoso es el perímetro del objeto.

La rugosidad del perímetro no necesariamente se aplica a un objeto que tenga el perímetro cerrado.

La dimensión fractal

La dimensión fractal proporciona una cuantificación de la forma de un objeto, es decir, su organización estructural [Olesen, et al, 1993].

El ser humano no es apto para distinguir formas irregulares, o pequeños cambios de forma. La dimensión fractal es un descriptor muy sensible a cambios de forma.

$$D = \frac{2 \ln\left(\frac{P}{4}\right)}{\ln(A)}$$

La excentricidad

La excentricidad es una medida de la elongación de un objeto.

Para calcular este descriptor, se aplica una transformación de distancia. Para una conectividad, se obtienen las longitudes $d_{\pi}(i,j)$ de las líneas rectas que pasan por cada pixel y que van de borde a borde. En cada pixel se anota entonces la longitud máxima $d_{\pi}^{\text{máx}}$ y la mínima $d_{\pi}^{\text{mín}}$ y a partir de aquí se define entonces la excentricidad.

$$E = \frac{\text{máx}(d_{\pi} \text{ máx})}{\text{máx}(d_{\pi} \text{ mín})}$$

La diversidad o estructura del paisaje

La diversidad o estructura del paisaje es una medida de la heterogeneidad en la distribución de las clases.

Supóngase que la imagen ha sido segmentada en una colección de regiones $\{R_1, R_2, \dots, R_c\}$, cada una definiendo una clase. Si P_i y A_i son el perímetro y el área de la región R_i y N_i es el número de regiones adyacentes a R_i y N_c es el número total de regiones, entonces se define un parámetro de estructura que está dado por:

$$E^i = 2 \frac{\ln(p_i^m)}{\ln(A_i)}$$

donde:

P_m^i = perímetro modificado dado por

$$P_m^i = P_i + 2 \frac{N_i(A_i - 1)}{N_c - 1}$$

La diversidad o estructura está dada por el promedio de los parámetros de estructura.

$$E_s = \frac{1}{N_c} \sum_i E_{s_i}$$

I.2.4 Necesidades y limitaciones

I.2.4.1 De hardware:

Teniendo un solo procesador con suficiente RAM como para evitar accesos temporales a disco disminuiría el costo computacional. Sin embargo, el uso de procesadores paralelos reduciría el costo computacional de algoritmos que procesan secuencialmente datos que son independientes entre sí. Tal es el caso del cálculo de los parámetros óptimos de cada clase, durante las etapas de calibración y entrenamiento.

I.2.4.2 De software:

En cuanto al entorno de desarrollo:

La tendencia a cambiar de sistema operativo lleva a la necesidad de migrar el paquete SANDI para DOS a una versión para Windows, con el mínimo de modificaciones en las rutinas ya programadas en Turbo Pascal, por lo que se considera conveniente la utilización de un compilador con una interfaz de desarrollo amigable en el nuevo entorno, pero con una recompilación sencilla de los módulos de SANDI para DOS. El compilador seleccionado es DELPHI, que tiene un lenguaje de programación orientado a objetos, con la sintaxis de Pascal y cumple con los requisitos antes mencionados.

Respecto a la forma de almacenar las imágenes, se conserva el formato Raster para una tener etapa de transición más cómoda hasta que el paquete pueda ser recompilado totalmente.

En cuanto a los algoritmos:

Los algoritmos deben ser diseñados de manera que la sencillez en su lectura y la eficiencia en su rendimiento queden compensadas.

En cuanto a los datos de entrada:

- El ruido y los artefactos de la imagen a segmentar deben ser filtrados previamente.
- En imágenes multiespectrales las bandas decorrelacionadas.
- La correcta especificación de los pixels prototipo es crucial para una buena segmentación.
- La selección del criterio de homogeneidad es un aspecto que merece investigación a fondo, puesto que no hay un esquema matemático para definir un espacio de criterio de homogeneidad.

CAPITULO II

EL CLASIFICADOR CONTEXTUAL

En un clasificador contextual se clasifica un pixel no solamente en función de sus propiedades espectrales puntuales sino también en función del contexto espacial donde se encuentra. Como ya se ha mencionado, un clasificador puntual clasifica los pixels únicamente en función de sus valores a través de las bandas. Los clasificadores contextuales son clasificadores del entorno, puesto que consideran una vecindad alrededor de cada pixel.

En este trabajo se propone un nuevo clasificador contextual donde se determinan muestras estadísticas de las clases definidas como resultado de un algoritmo automatizado de crecimiento de regiones. Un pixel es clasificado comparando la forma, la media y/o la correlación de la función de densidad asociada a la vecindad del pixel con aquella de las clases. La comparación se lleva a cabo por medio de una medida de similaridad entre las funciones de densidad tanto para la respuesta espectral como para el contenido textural. Este esquema de clasificación es válido para ambos clasificadores espectral y textural tratados en el presente trabajo.

En este capítulo se describirán los siguientes aspectos: el esquema de crecimiento de regiones, la estimación de los parámetros óptimos para el crecimiento de regiones y los fundamentos físicos y matemáticos de los clasificadores espectral y textural.

II.1 EL ESQUEMA DE CRECIMIENTO DE REGIONES

II.1.1 La descripción del algoritmo de crecimiento de regiones

El clasificador contextual usa los conjuntos de entrenamiento determinados por un algoritmo optimizado de crecimiento de regiones (Lira y Frulla, 1998). Los conjuntos son representaciones estadísticas de las clases definidas, siendo éstas espectrales o textuales.

El algoritmo comienza al sembrar pixels en lugares convenientes de la imagen donde se conoce la existencia de la clase. Esta tarea se lleva a cabo manualmente por inspección visual de la imagen con el soporte del conocimiento de la escena real. Una vez que se determinaron las semillas, una por clase, comienza el algoritmo de crecimiento de regiones. El crecimiento se lleva a cabo por agregación de pixels que satisfacen un criterio de homogeneidad. El criterio se evalúa en una ventana de tamaño óptimo. Un pixel es agregado dentro de la región si la diferencia entre el valor de homogeneidad en una ventana centrada sobre la semilla sembrada y el valor de homogeneidad en una ventana centrada sobre el pixel en cuestión, está dentro de cierto umbral. El algoritmo de crecimiento de regiones termina cuando este criterio de homogeneidad ya no se satisface [Kai & Muller, 1991].

II.1.2 La formalización del algoritmo de crecimiento de regiones

Sea g una imagen y un pixel $p_{ij}^0 \in g$ y sea $R_0 = \{ p_{ij}^0 \}$ la subregión inicial. A p_{ij}^0 se le conoce como el pixel sembrado relativo a la subregión R_0 .

Se define a R_1^0 como el conjunto de todos aquellos pixels que no están en R_0 pero que tienen al menos un vecino con R_0 bajo cierta conectividad.

Sea $E(R_0)_n$ el valor del criterio de uniformidad aplicado a una vecindad n de R_0 .

Se define R_1 como aquella región formada conjuntamente por R_0 y por todos los pixels $p_{kl}^1 \in R_1^0$ para los cuales el valor del predicado de uniformidad $E(p_{kl}^1)_n$ evaluado sobre una vecindad n difiere de $E(p_{ij}^0)_n$ en a lo sumo una cierta cantidad ϵ llamada umbral, es decir, R_1 es el siguiente conjunto:

$$R_1 \equiv \{ p_{kl}^1 : |E(p_{kl}^1)_n - E(p_{ij}^0)_n| \leq \epsilon \} \quad [2.1]$$

El número real ϵ es referido como el parámetro de uniformidad.

El proceso continúa y en general:

$$R_m \equiv \{ p_{kl}^m : |E(p_{kl}^m)_n - E(p_{ij}^0)_n| \leq \epsilon \} \quad [2.2]$$

Donde:

$$R_m \supset R_{m-1} \supset \dots \supset R_1 \supset R_0 \quad [2.3]$$

El proceso continúa iterativamente hasta que el nuevo conjunto sea igual al anterior.

$$R_{k+1} = R_k \quad [2.4]$$

Generalización a n regiones o n clases:

Un pixel p_{kl} es agregado a la clase t si el predicado de uniformidad se satisface para una vecindad n_t de éste.

$$p_{kl} \rightarrow \text{clase } t \Leftrightarrow |E(p_{kl}^t)_n - E(p_{ij}^0)_n| \leq \epsilon \quad [2.5]$$

Donde $n = n_t$

II.1.3 La implementación del algoritmo de crecimiento de regiones

El algoritmo de crecimiento de regiones empleado en el presente trabajo está fundamentado en "A Scan Line Seed Fill Algorithm". Para detalles se puede consultar el libro "Procedural Elements For Computer Graphics" de David F. Rogers, página 88.

II.2 LA ESTIMACION DE LOS PARAMETROS OPTIMOS PARA EL CRECIMIENTO DE REGIONES

II.2.1 El criterio de homogeneidad

II.2.1.1 La familia de funciones criterio de homogeneidad

Dado un universo U , dado un conjunto de instancias que pertenecen al universo U y dado un conjunto de c entidades con sus atributos a las cuales se tiene que asignar unívocamente las instancias, la familia \mathfrak{S}_U de funciones criterio de homogeneidad h es el conjunto de atributos a de dichas entidades que satisfacen la siguiente condición:

Si las entidades tienen el atributo a en común [Acton, 1996] y los dominios $D^i(a)$ de este atributo para cada una de las entidades i están separados de tal manera que es posible efectuar la asignación de las instancias a las entidades con el mínimo error, dicho atributo es una función criterio de homogeneidad h .

Formalmente una familia de funciones criterio de homogeneidad se define de la siguiente manera:

Sean c clases C_i incluídas en una imagen g tales que $C_i \cap C_j = \{\}$ y sea a un atributo de dichas clases y sea $D^i(a)$ el dominio del atributo a para la clase i -ésima:

$$\mathfrak{S}_U = \{h=a / D^i(a) \cap D^j(a) \equiv \{\} \quad \forall i \neq j \quad i, j = 1, 2 \dots c\} \quad [2.6]$$

Nótese que la pertenencia o no de un atributo a la familia de funciones criterio de homogeneidad depende de cada problema o universo U .

Sea, por ejemplo, el atributo media. Si en una imagen g las medias de las clases son diferentes, el atributo media pertenece a la familia de funciones criterio de homogeneidad de g . En una imagen en las cuales todas las clases tienen la misma media, pero distinto desvío standard, la media no es una función criterio de homogeneidad para dicha imagen, en cambio, el desvío estándar sí lo es [Arai, 1993].

En resumen, todos aquellos atributos que para una imagen dada permiten reconocer cada elemento en sus clases, son elementos de la familia de funciones criterio de homogeneidad.

II.2.2 La función objetivo

II.2.2.1 El dominio de la función objetivo

En general

Sea p_0 un pixel perteneciente a una imagen g y R un conjunto de regiones R_i formadas por pixels que pertenecen a dicha imagen tales que:

$$R_0 = \{p_0\} \quad [2.7]$$

$$R_i \subset g \quad \forall i \quad [2.8]$$

$$R_0 \subset R_1 \subset \dots \subset R_m \quad [2.9]$$

y

$$R = \bigcup_{i=0}^m R_i = R_m \quad [2.10]$$

Sea $W = \{w_0, w_1, \dots, w_m\}$ un conjunto de etiquetas incluido en el conjunto N de los naturales y Q una función que asocia a cada subconjunto R_i de R un elemento del conjunto W tal que al subconjunto R_i le corresponde el elemento w_i , se dice que W es el dominio de la función objetivo.

$$W = Q(R) \quad [2.11]$$

$$w_i \in N$$

Estas regiones R_i no tienen ningún requisito de forma. En particular, cuando sean cuadradas y todas estén centradas en un pixel p_0 , se utilizará la notación V_i .

II.2.2.2 La función objetivo

La función objetivo B es un operador que aplica la función criterio de homogeneidad h normalizada a cada elemento del conjunto W .

$$B: w \rightarrow h(w) \quad [2.12]$$

$$h = h(w) = h[Q(R)] = m(R) \quad [2.13]$$

$$h \in [0, 1] \quad [2.14]$$

La función objetivo es discreta.

Si W^c es el dominio de la función objetivo B asociado a una clase c dada por su pixel prototipo p_0^c , entonces en lugar de B se utilizará la notación B^c .

Si la función objetivo aplica la función criterio de homogeneidad h normalizada a la imagen, se obtiene un parámetro de la imagen.

Sea g una imagen, y R_m una región igual a la imagen g , la función criterio de homogeneidad aplicada sobre R_m genera un parámetro θ (o descriptor) de la imagen.

$$H[Q(R_m)] = \theta \quad [2.15]$$

II.2.2.3 La función objetivo continua

La función que interpola los puntos de la función objetivo discreta se denomina función objetivo continua en el intervalo continuo $(w_0, w_m) \subset \mathfrak{R}$. El límite del cociente incremental entre el incremento de la función objetivo Δb y el incremento Δw , cuando este último tiende a cero existe dentro de dicho intervalo, por lo que la función objetivo continua es derivable en el intervalo $(w_0, w_m) \subset \mathfrak{R}$.

A efectos de no introducir más notaciones se la distinguirá de la función discreta por medio un índice. Se nombrará B^* a la función objetivo continua y solamente B a la función objetivo discreta.

Sea g una imagen y w_m el elemento de W correspondiente a la región R_m de tamaño igual al tamaño de dicha imagen g , el límite de la función objetivo continua B^* cuando w tiende a w_m por la izquierda es un parámetro de la imagen.

$$\lim_{w \text{ (izq)} \rightarrow w_m} B^*[h(w)] = h[Q(R_m)] = \theta \quad [2.16]$$

Sean g una imagen, c una clase incluida en g que tiene el atributo ϕ , R_i^c una región conectada simple formada por pixels que pertenecen a dicha clase c y que constituyen muestras suficientes de c , y sea p_0^c el pixel prototipo de la clase c tal que $R_0^c = \{p_0^c\}$. Sea W^c el conjunto de etiquetas asociado al conjunto R^c . El límite de la función objetivo continua B^* cuando w^c tiende a w_i^c es un parámetro $\hat{\phi}^c$ estimado de la clase. (Nótese que $\hat{\phi}$ es un valor de la variable ϕ).

$$\lim_{w^c \rightarrow w_i^c} B^*[h(w^c)] = \hat{\phi}^c \quad [2.17]$$

$$h(w_i^c) = \hat{\phi}^c \quad [2.18]$$

La formalización de la hipótesis respecto del pixel sembrado

Para el pixel prototipo p_0^c de toda clase c existe una región conectada simple R_i^c formada por pixels que pertenecen a dicha clase c y que constituyen muestras estadísticas suficientes de c .

$$\exists R_i^c \forall p_0^c \quad [2.19]$$

II.2.2.4 Características de la función objetivo para todo $w_i^c \in [w_0^c, w_i^c]$, es decir, para toda región R_i^c incluida en R^c :

- 1) Arranca con una zona denominada inestable, puesto que si bien las entradas están acotadas, las salidas no lo están: los elementos de las regiones pertenecen a la clase, pero el criterio de homogeneidad aplicado sobre la región genera un valor

no representativo de la clase. Esto es consecuencia del submuestreo: las muestras son válidas, pero no son suficientes para que sus estadísticas sean válidas también.

La inestabilidad se define formalmente para el caso discreto de la siguiente manera: existe un intervalo $[w_0, w_t)$ del dominio de la función objetivo discreta sobre el cual la sumatoria de los valores de la función objetivo discreta es igual a infinito.

$$\exists [w_0, w_t) / \sum_{i=0}^t B[h(w_i)] = \infty \quad [2.20]$$

La definición formal de esta zona para el caso continuo es la siguiente: existe un intervalo $[w_0, w_t)$ del dominio de la función objetivo continua sobre el cual la integral de la función objetivo continua es igual a infinito.

$$\exists [w_0, w_t) / \int_{w_0}^{w_t} B[h(w)]dw = \infty \quad [2.21]$$

Nota: en términos de aplicación este valor ∞ se traduce en un valor numérico grande.

- 2) A valores de su dominio mayores que w_t y cuando el tamaño de la región tiende al tamaño de la región R_i^c (es decir, cuando w tiende a w_{tf}), la función objetivo tiende a hacerse asíntota a una función $\hat{\phi}$ constante en w , la cual es un parámetro estimado de la clase. En consecuencia, se tiene en el intervalo $[w_t, w_{tf}]$ una zona estable de la función objetivo, puesto que tanto las entradas como las salidas son acotadas: los elementos de las regiones pertenecen a la clase y el criterio de homogeneidad aplicado sobre la región genera un valor representativo de la clase. En otras palabras, en esta zona las muestras de la clase son válidas y suficientes como para generar estadísticas válidas de la clase. Para el caso discreto se dice que existe un intervalo $[w_0, w_t)$ del dominio de la función objetivo discreta sobre el cual la sumatoria de los valores de la función objetivo discreta es menor que infinito.

$$\exists [w_t, w_{tf}] / \sum_{i=t}^{tf} B[h(w_i)] < \infty \quad [2.22]$$

La definición formal de esta zona para el caso continuo es la siguiente: existe un intervalo $[w_t, w_{tf}]$ del dominio de la función objetivo continua sobre el cual la integral de la función objetivo continua es menor que infinito.

$$\exists [w_t, w_{tf}] / \int_{w_t}^{w_{tf}} B[h(w)]dw < \infty \quad [2.23]$$

- 3) La función objetivo sufre un cambio brusco en su pendiente al salir de R_{if}^c . En esta zona se tienen muestras inválidas de la clase c .

Notas: En la parte II.2.3 se presenta un ejemplo que permite apreciar el comportamiento descrito de la función objetivo. [La definición de estabilidad se extrajo de Oppenheim & Schafer, 1975].

II.2.2.5 La primera derivada de la función objetivo

La primera derivada de la función objetivo B^* continua se define como:

$$b^* = \frac{dB}{dw} \quad [2.24]$$

Para el caso discreto, y con fines de aplicación, se emplea la siguiente expresión:

$$b_i = B[h(w_{i+1})] - B[h(w_i)]$$

Para una clase c dada, la primera derivada de la función objetivo B^c se denomina $b^c(w^c)$. Esta función es una herramienta útil para obtener el tamaño de ventana óptimo de la clase, puesto que permite visualizar el cambio que se produce en la vecindad del pixel sembrado. Además, el límite de la primera derivada de la función objetivo b^c cuando la región tiende a ser la región R_1^c es "aproximadamente" cero. Esto se debe a que la función criterio de homogeneidad tiende, también en este caso, a una constante.

$$\begin{aligned} \lim_{w^c \rightarrow w_1^c} b^{*c}(w^c) &= \lim_{w^c \rightarrow w_1^c} \frac{dB^{*c}}{dw} = \frac{dB^{*c}(w_1^c)}{dw} = \\ &= \lim_{\Delta w^c \rightarrow 0} (\hat{\phi}^c - \hat{\phi}^c) \equiv 0 \end{aligned} \quad [2.25]$$

Característica relevante de la función primera derivada de la función objetivo para una clase c dada por su pixel prototipo p_0^c :

La energía local tiende a cero cuando la región tiende a ser R_1^c .

Para el caso discreto se expresa de la siguiente manera:

$$|b(w_t^c)|^2 \cong 0 \quad [2.26]$$

Para el caso continuo:

$$|b^*(w^c)|^2 \cong 0 \quad [2.27]$$

Para una clase dada, se tiene la necesidad de establecer una región que contiene muestras estadísticas válidas de la misma. Con base en esta propiedad, debe ser posible establecer un umbral que determine en qué w se reduce suficientemente la energía local en la derivada de la función objetivo y con ello, establecer tal región.

Cuando más muestras se agreguen a la región, menos tiende a cambiar el estimador de la clase, por lo que para el caso continuo:

$$\lim_{w^c \rightarrow w_{tr}^c} b^*(w^c) = 0 \quad [2.28]$$

y para el caso discreto:

$$\lim_{w^c \rightarrow w_{tr}^c} b(w^c) = 0 \quad [2.29]$$

II.2.2.6 La segunda derivada de la función objetivo

La segunda derivada de una función permite apreciar los cambios de población de clases. La función objetivo tiene la peculiaridad que cuando las regiones centradas en el pixel prototipo de una clase comienzan a tomar pixels de otras clases, ésta sufre un cambio brusco en su concavidad. Esto no implica que necesariamente se presente un punto de inflexión. Dicho cambio de concavidad permite distinguir el fin de la clase, es decir, encontrar R_{tr}^c .

II.2.3 La ventana óptima de una clase

Dada una clase c por su pixel prototipo p_0^c , su ventana óptima se obtiene a partir de su función objetivo B^c asociada. Para ello, debe definirse previamente el dominio W^c correspondiente. Este dominio tiene la peculiaridad que está asociado a regiones rectangulares, más precisamente cuadradas, las cuales están centradas en el pixel prototipo de la clase dada.

Es necesario definir, en primer lugar, el concepto de centro de ventana. El concepto de ventana ya ha sido definido en la parte I.2.3.1.2.

II.2.3.1 Centro de una ventana

Sea $p_0(x,y)$ un pixel que pertenece a una ventana $V(y_{ini}, x_{ini}, y_{fin}, x_{fin})$.

$$p_0(x,y) \text{ es el centro de la ventana } V \Leftrightarrow x = \frac{x_{fin} - x_{ini}}{2} \wedge y = \frac{y_{fin} - y_{ini}}{2} \quad [2.30]$$

II.2.3.2 El dominio de la función objetivo asociada a una clase c para el cálculo de la ventana óptima

Sea g una imagen, c una clase incluida en la imagen g y p^c_0 el pixel prototipo de tal clase.

El dominio de la función objetivo asociada a la clase c incluida en una imagen g y dada por su pixel prototipo p^c_0 es el conjunto de ventanas V^c_i , de tamaño impar con centro en el pixel sembrado p^c_0 , que están incluidas en la imagen g , donde i varía desde 1 hasta un máximo que depende del tamaño de la imagen g y de la posición de p^c_0 .

Si se consideran ventanas cuadradas, la ventana V^c_1 de tamaño 1×1 , con centro en el pixel sembrado p^c_0 , es un conjunto de un solo elemento: el pixel sembrado p^c_0 .

$$V^c_1 = \{p^c_0\} \quad [2.31]$$

El máximo tamaño de ventana $máx \times máx$ contenida en la imagen con centro en el pixel sembrado está dado por dos veces la menor distancia del pixel sembrado a cualquiera de los bordes de la imagen más uno. Se están considerando únicamente ventanas de tamaño impar.

La ventana dominio asociada a la clase c es una unión de conjuntos

$$V^c = \bigcup_{i=1}^{(máx-1)/2} V_{2i-1} \text{ tal que } V_{2i-1} \text{ es una ventana de tamaño } (2i-1)^2 \text{ con centro en el pixel } p^c_0, \text{ es decir:}$$

$$V^c = V_1 \cup V_3 \cup \dots \cup V_{máx} \quad [2.32]$$

Como:

$$V_1 \subset V_3 \subset \dots \subset V_{máx} \quad [2.33]$$

Entonces:

$$V^c = V_{máx} \quad [2.34]$$

Sea $W^c = \{w^c_1, w^c_3, \dots, w^c_{máx}\}$ un conjunto incluido en el conjunto de los números naturales y sea Q una función que asocia a cada subconjunto V^c_i de V^c un elemento del conjunto W^c tal que al subconjunto V^c_i le corresponde el elemento w^c_i , se dice entonces que W^c es el dominio de la función objetivo asociada a la clase c para el cálculo de la ventana óptima.

$$w=Q(v) \quad [2.35]$$

$$w_i^c \in N$$

Nota: Es válido usar ventanas rectangulares, siempre y cuando se cumpla la siguiente condición:

$$V_i \subset V_{i+2} \quad \forall i \quad [2.36]$$

La función objetivo asociada a una clase c para el cálculo de su ventana óptima

La función objetivo B^c asociada a una clase c para el cálculo de la ventana óptima es el operador B que aplica la función criterio de homogeneidad normalizada a cada elemento del conjunto W^c , donde cada w_i^c se corresponde con una ventana de tamaño $i \times i$ centrada en el pixel prototipo p_0^c .

$$B^c=B:Q(v^c) \rightarrow h(w^c) \quad [2.37]$$

$$h^c = h(w^c) = h[Q(v^c)] \quad [2.38]$$

$$h^c \in [0, 1] \quad [2.39]$$

Ejemplos inductivos para el cálculo de la ventana óptima:

Se tienen cuatro imágenes sintéticas de 512 pixels por 512 pixels, en las cuales se presenta centrada una clase determinada. El fin de la clase de la imagen correspondiente a la figura 2.1 se tiene con una ventana igual 256 por 256 pixels. El fin de la clase de la imagen correspondiente a la figura 2.2 se tiene con una ventana igual 128 por 128 pixels. En las figuras 2.3 y 2.4 se tienen clases cuyo fin se encuentra en ventanas de 256 por 256 pixels.

Sea $h(w)$ la función media normalizada de la ventana de tamaño $v \times v$ correspondiente a w centrada en el centro de la clase. Se presentan a continuación las gráficas del comportamiento de la función objetivo $B[h(w)]$ sobre las imágenes mencionadas.

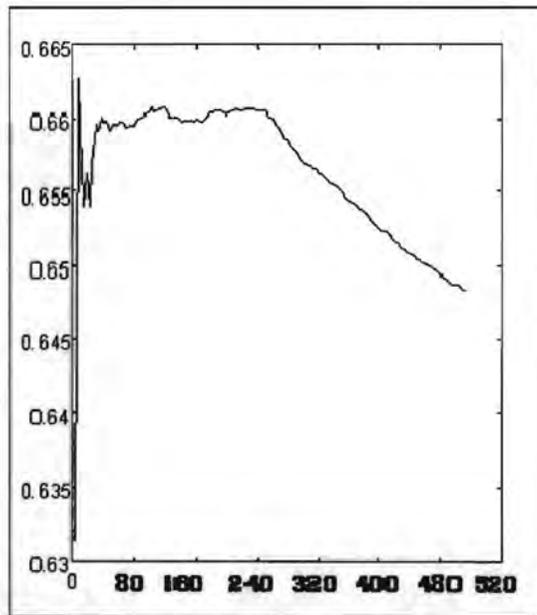


Figura 2.1

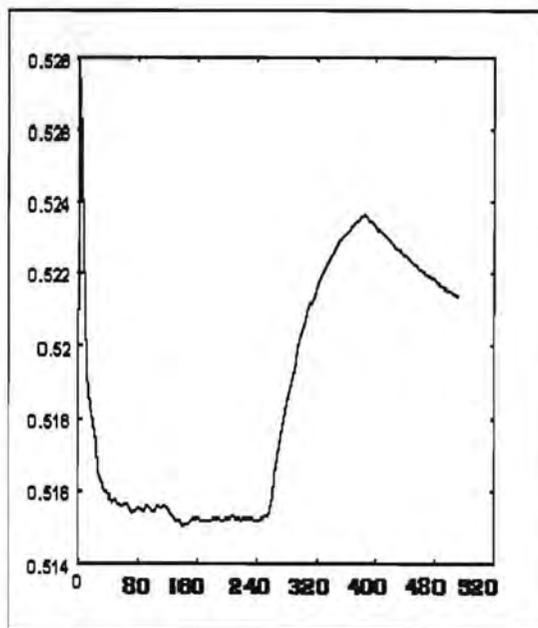


Figura 2.2

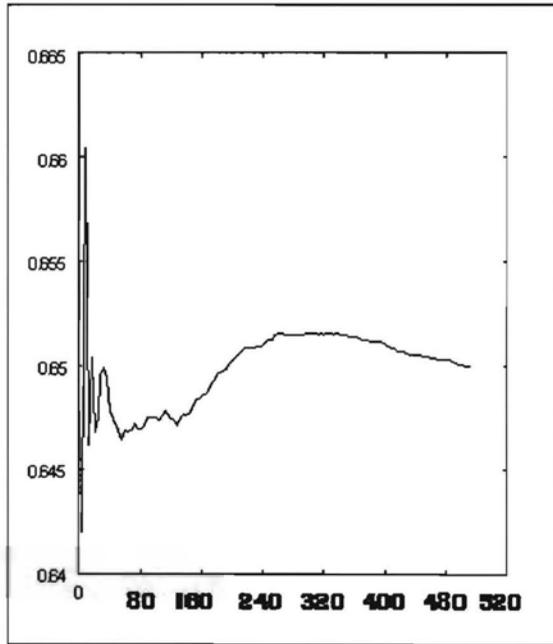


Figura 2.3

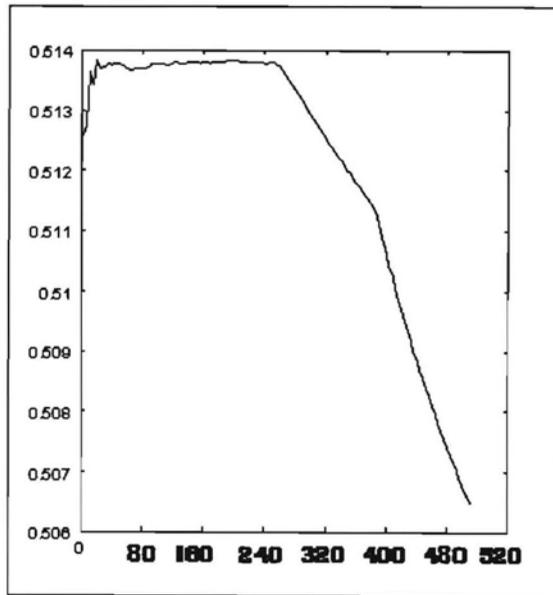


Figura 2.4

Las gráficas de las respectivas primeras derivadas de la función objetivo son las siguientes:

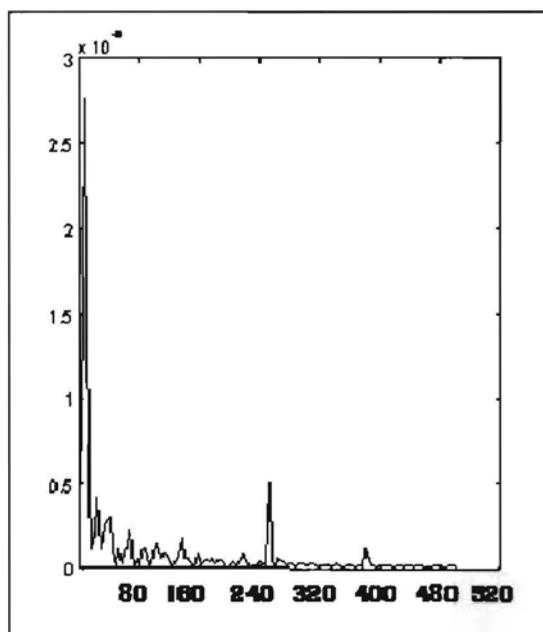


Figura 2.5

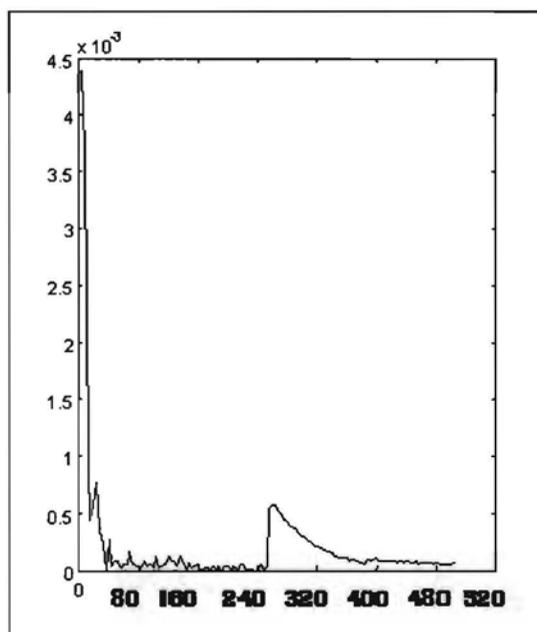


Figura 2.6

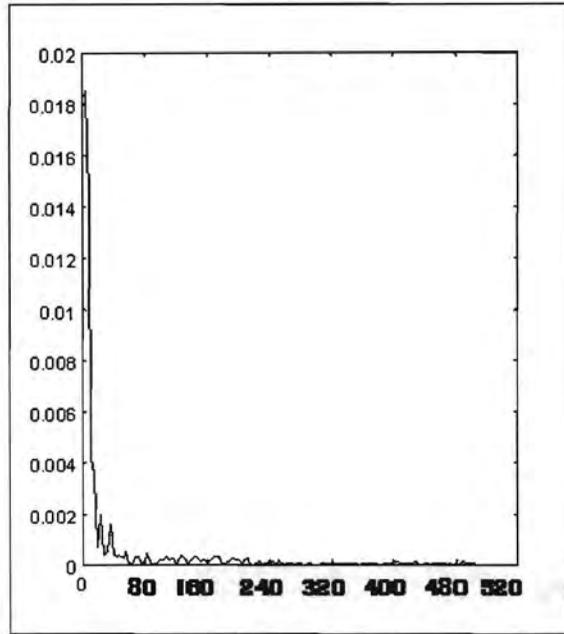


Figura 2.7

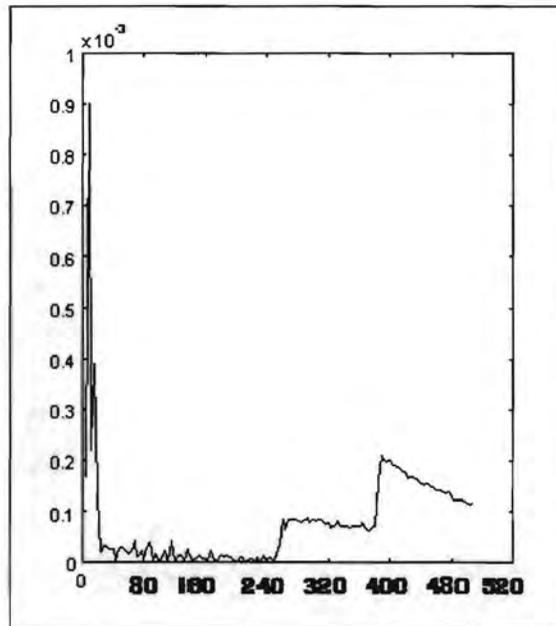


Figura 2.8

Comentarios:

Las inestabilidades en la función objetivo introducidas por el submuestreo

Cuando se toman tamaños de ventana pequeños, en la mayoría de los casos, las estadísticas que se obtienen no son representativas de la clase. Cuando comienzan a tenerse muestras suficientes de la clase, las estadísticas se estabilizan hasta que, al salir de la clase, se observan nuevas inestabilidades. En consecuencia, al principio, la curva de la función objetivo sufre una sucesión de variaciones bruscas que están reflejando los grandes cambios que van sufriendo las estadísticas de la clase al incorporar unas pocas muestras más. Estas variaciones deben ser ignoradas.

La ventana óptima versus la ventana fin de la clase

Como ya se ha mencionado, en el fin de la clase la función objetivo sufre un cambio brusco en su concavidad. En este punto, se tiene la ventana fin de la clase. Se desea obtener una ventana suficientemente pequeña como para tener aproximadamente las mismas estadísticas que la ventana fin de la clase: ésta es la ventana óptima.

El problema consiste en detectar el fin de las variaciones bruscas introducidas por el submuestreo en la función objetivo, o en otras palabras, el principio de la zona de estabilidad de dicha función.

El umbral de estabilidad Ω

Para reconocer el comienzo de la región de estabilidad se utiliza el parámetro Ω , denominado umbral de estabilidad. En este trabajo se han utilizado valores empíricos para Ω , pero se entiende que un camino a seguir es optimizar este parámetro en función de los datos mismos de la clase. (Véase Figura 2.9)

Formalmente:

Dada una clase c señalada en una imagen por su pixel prototipo p^c_0 , sea W^c el dominio de su función objetivo B^c asociado a un conjunto V^c con centro en el pixel mencionado, sea B^c continua y derivable dentro de W^c asociada a una función criterio de homogeneidad h , entonces v^c_t es la ventana óptima para la clase c si t está asociado al menor elemento de W^c para el cual la función objetivo es un parámetro estimado $\hat{\phi}$ representativo de la clase c .

$$v^c_t = Q^{-1}(\min(w^c_t)) / w^c_t = Q(v^c_t) \in W^c \wedge B^c = h(w^c_t) = \hat{\phi} \quad [2.40]$$

Q^{-1} es la función inversa de la función Q que relaciona biunívocamente a cada elemento de V^c con un elemento de W^c .

Debido a que se muestrea con el objeto de obtener precisamente este parámetro $\hat{\phi}$, se utiliza el umbral de estabilidad Ω para encontrar t , es decir, el principio de la zona de

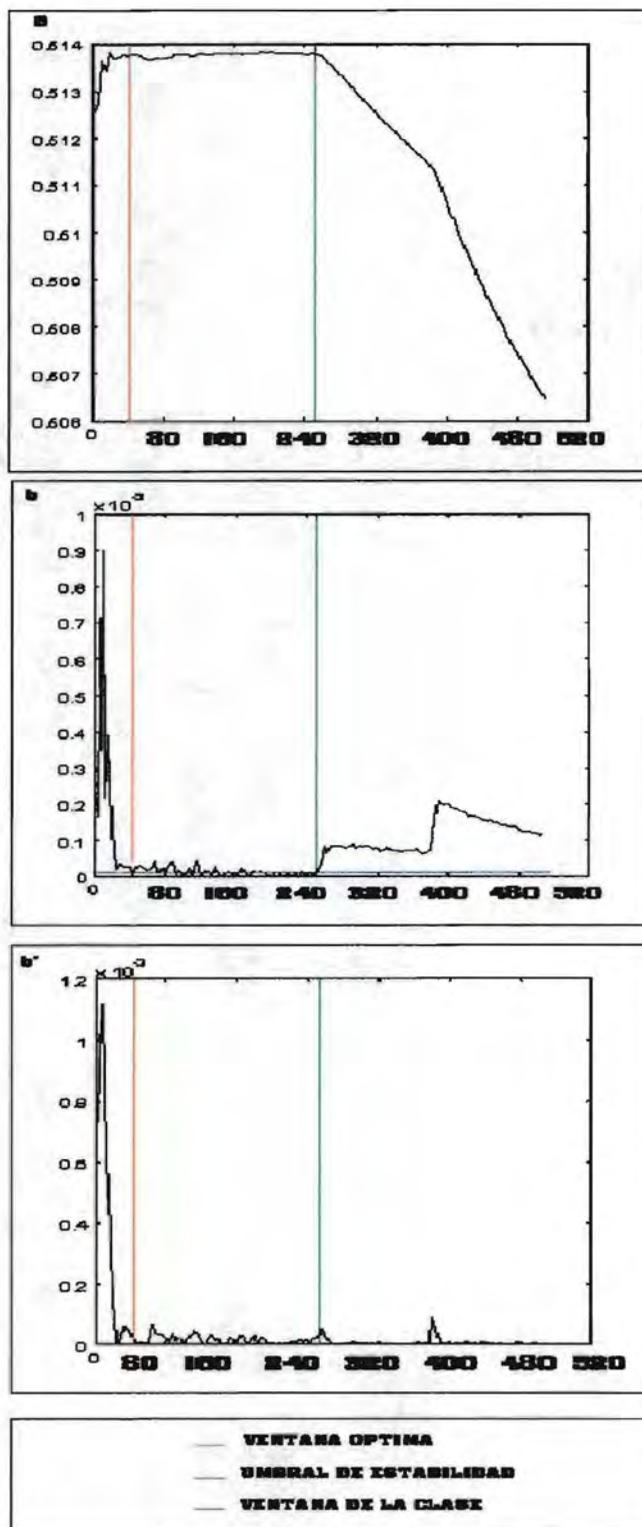


Figura 2.9
49

estabilidad de la función objetivo B^c , y con ello, la ventana óptima para generar las estadísticas de las muestras. Entonces, la definición de aplicación que se utiliza para el cálculo de la ventana óptima es la siguiente:

Dada una clase c incluida en una imagen por su pixel prototipo p^c_0 , sea W^c el dominio de su función objetivo B^c asociado a un conjunto V^c con centros en el pixel mencionado, sea B^c continua y derivable dentro de W^c , entonces la ventana óptima v^c_i para la clase c es la ventana asociada al menor elemento de W^c para el cual el valor absoluto de la derivada de la función objetivo es menor que un umbral de estabilidad Ω , es decir:

$$v^c_i = Q^{-1}(\min(w^c)) / w^c = Q(v^c) \in W^c \wedge \left| \frac{dB(w)}{dw} \right| = |b^c(w)| < \Omega \quad [2.41]$$

Experimentalmente se prueba que utilizando $\Omega \leq 0.3$ se obtienen buenos resultados.

II.2.4 La función de separabilidad

Una primera relación entre las entidades

Con el objeto de minimizar el traslape entre las clases en el crecimiento de regiones, se efectúa una primera interacción entre las mismas, comparando sus estadísticas estimadas. Aquí no se están validando los niveles de jerarquía de las entidades. Lo que se desea es llevar a cabo crecimientos no invasivos de regiones, es decir, no incorporar a una clase pixels que, a pesar de tener estadísticas cercanas a la misma, estén también muy cercanos a otra clase y que pudiesen afectar las nuevas estadísticas de la clase que se generen con el crecimiento de regiones.

El objetivo de la interacción entre las clases es establecer los umbrales óptimos para el crecimiento de regiones. El procedimiento consiste en: comparar a través de una función de separabilidad las estadísticas de cada clase con las demás, buscar la clase que se encuentra más cercana a cada clase en cuestión y establecer los respectivos umbrales como función de la distancia de cada clase con su clase más cercana.

La función de separabilidad

La función de separabilidad depende de las distancias entre clases. Por ello se define, en primer lugar, este concepto.

La distancia entre dos clases

Sean i y j dos clases, la distancia entre la clase i y la clase j es la diferencia absoluta entre el valor que toma la función criterio de homogeneidad para la clase i y el valor que toma la función criterio de homogeneidad para la clase j , en los valores asociados a regiones representativas para cada una de ellas

$$d_{i,j} = \left| h^i(w_{ti}^i) - h^j(w_{tj}^j) \right| \quad [2.42]$$

$w_{ti}^i \in W^i$
 $w_{tj}^j \in W^j$

La distancia de una clase i a las demás clases

Para calcular la distancia de una clase i a las demás clases deben definirse factores de peso $P(i,j)$ o probabilidades. La distancia de una clase i a las demás clases se calcula mediante la siguiente expresión:

$$d_i = \sum_{j=1}^{N_c} P(j) d_{i,j} \quad [2.43]$$

donde:

$$\sum_{j=1}^{N_c} P(j) = 1 \quad [2.44]$$

donde N_c es la cantidad de clases.

Si todas las distancias entre clases son afectadas por el mismo factor de peso $P(j) = \frac{1}{(N_c - 1)}$ excepto cuando se mide la distancia de una clase consigo misma (es decir, cuando j es igual a i), caso en que se toma $P(j)=0$, el valor d_i resultante es el promedio de las distancias entre clases.

Si se considera que $P(j)$ está definido de la siguiente manera: $P(j)$ es igual a uno si $d_{i,j}$ es la mínima distancia de la clase i a cualquiera de las otras clases, y $P(j)$ es igual a cero cuando se calcula la distancia de una clase a sí misma, entonces d_i es la mínima distancia a cualesquiera de las $(N_c - 1)$ clases restantes. O dicho de otra manera:

$$C_{\text{más cercana a la clase } i} = j / d_{i,j} = \min(d_{i,j}) \quad \forall j \neq i \in [1, N_c] \subset N \quad [2.45]$$

donde N es el conjunto de los número naturales.

La distancia promedio entre las clases

$$\bar{d} = \frac{1}{N_c} \sum_{i=1}^{N_c} d_i \quad [2.46]$$

Como función de separabilidad se emplea entonces, cualquier expresión que utilice distancias entre las clases.

El problema en el uso de promedios

Una función de separabilidad puede emplearse para el cálculo de la ventana óptima (promedio), pero al utilizar distancias promedio, los valores se compensan y algunas estadísticas obtenidas pueden terminar no siendo representativas. Pueden manejarse valores más puros, como las distancias de una clase a las demás, en vez de trabajar con promedios generales de distancias entre clases. El costo computacional aumenta, en consecuencia.

II.2.5 El umbral óptimo

Una vez obtenido el valor $\hat{\phi}$ representativo de una clase c debe definirse el rango de variación del estimador ϕ , es decir, el dominio del atributo empleado como criterio de homogeneidad. Contando con este rango, se puede iniciar el proceso de muestreo, puesto que se está en condiciones de tomar un pixel de la imagen, calcular las estadísticas de su contexto y ver si las mismas están dentro del rango mencionado; en otras palabras, si están dentro de un umbral ϵ denominado umbral óptimo.

El problema del cálculo del umbral se puede resolver empleando los datos propios de la clase. A efectos de no incorporar en el muestreo aquellos pixels cuyas estadísticas encajan en los rangos de variación de más de una clase, pueden calcularse los extremos del intervalo de variación en función de los valores de las demás clases, más específicamente, utilizar las distancias antes mencionadas y con esto evitar traslapes de muestras de clases en el crecimiento de regiones.

$$\hat{\phi} - \epsilon \leq \phi(w) \leq \hat{\phi} + \epsilon \quad [2.47]$$

O lo que es lo mismo:

$$|\phi(w) - \hat{\phi}| \leq \epsilon \quad [2.48]$$

En el caso del umbral óptimo de la clase c para el crecimiento de regiones se entiende que:

$$\hat{\phi}^c = h(w^c) \quad [2.49]$$

donde:

$$\phi(w) = h(w) \quad [2.50]$$

Los valores de umbral obtenidos, ya sea a partir de las estadísticas propias de la clase o por la interacción entre las clases, son el resultado de aplicar una función f a los mismos, la cual puede ser acotada por un umbral de acotación η .

$$\epsilon = \eta f \quad [2.51]$$

Por ejemplo, si la función f es el desvío estándar de la clase, se sabe que utilizando un parámetro de acotación η igual a uno se tienen el 68% de los datos, con η igual a dos, el 95% de los datos, y con η igual a tres, prácticamente el 100% de los datos que se distribuyen bajo una Gaussiana.

En el caso en que la función f es una función de separabilidad de mínima distancia de una clase a las demás, entonces el parámetro η es el porcentaje de dicha distancia.

II.2.6 Notas acerca de posibles caminos de estudio

En cuanto a la validación de la hipótesis respecto del pixel prototipo

Se está suponiendo que para el pixel prototipo de cada clase c existe una región o vecindad R_t^c que contiene muestras representativas suficientes de la clase cuando el valor absoluto de la derivada de la función objetivo comienza a ser menor que un umbral de estabilidad preestablecido. Un camino a seguir sería encontrar la zona de estabilidad de la función objetivo asociada a cada clase calculando el umbral de estabilidad a partir de los datos propios de la imagen. Es decir, establecer automáticamente que un pixel sembrado es un pixel prototipo válido de una clase c si existe una zona de estabilidad no vacía de la función objetivo asociada, dentro de la cual muestras válidas de la clase c producen estadísticas válidas de tal clase:

$$P_0^c \text{ es válido para la clase } c \Leftrightarrow \exists [w_t^c, w_{tf}^c] \neq \{\} / \sum_{i=t}^{tf} B[h(w_i)] \leq \infty \quad [2.52]$$

donde:

$$h(w) \equiv \hat{\phi} \quad \forall w \in [w_t, w_{tf}] \quad [2.53]$$

Es decir, si existe el intervalo $[w_t^c, w_{tf}^c]$, en él se puede encontrar muestras suficientes de la clase, siempre y cuando, éstas sean muestras correctas de la clase. Por ejemplo, sea una imagen que contiene dos clases. Si el pixel prototipo de una de ellas está exactamente sobre la frontera de estas clases y ésta es una línea recta, entonces las ventanas asociadas a la función objetivo contienen exactamente la misma cantidad de pixels de ambas. Existe, por tanto, un intervalo en el cual la función objetivo es estable, pero las estadísticas de la clase que se obtienen no son correctas.

Al ingresar un parámetro de estabilidad se corre el riesgo de que éste sea demasiado grande o demasiado pequeño. En el primer caso, la ventana óptima de la clase será demasiado pequeña y las muestras no serán representativas de la clase. Sin embargo, también puede suceder que dicho parámetro sea tan pequeño que pase por alto a la región de estabilidad que contiene estadísticas válidas para la clase y corte a la función objetivo en una segunda zona de estabilidad donde las muestras ya no son representativas tampoco. Como ya se ha mencionado, la función objetivo, cuando más pueda asociarse su dominio con ventanas que sean del tamaño de la imagen, tiende a hacerse asíntota a una función

constante que es un parámetro de la imagen. Es decir, tiende a una región estable que no es en la mayoría de los casos, la “primera” región estable que se desea encontrar en el cálculo de la ventana óptima de una clase.

II.3 FUNDAMENTOS DEL CLASIFICADOR

II.3.1 La estimación de los parámetros para la clasificación

Una vez reunidas las muestras estadísticamente válidas para una clase c dada por medio del algoritmo de crecimiento de regiones, debe obtenerse el nuevo valor del estimador ϕ para este conjunto de muestras que se denota como $\hat{\phi}^{(g)}$. La región crecida está asociada al nuevo dominio $W^{(g)c}$ de la función objetivo B , donde $w^{(g)c}_0$ es el elemento de $W^{(g)c}$ asociado a la región crecida inicial $R^c_0 = \{p^c_0\}$, $w^{(g)c}_1$ es el elemento de $W^{(g)c}$ asociado a la región crecida R^c_1 y así sucesivamente hasta $w^{(g)c}_m$ asociado a R^c_m , donde m es la cantidad de pixels que pertenecen a la región crecida de la clase c .

El límite de la función objetivo, cuando el número de muestras de una clase c dada tiende a la cantidad de muestras m de la región crecida de dicha clase, es un parámetro estimado de la función criterio de homogeneidad para ese conjunto de muestras

$$\lim_{w^{(g)} \rightarrow w^{(g)}_m} B(w^{(g)}) = \hat{\phi}^{(g)} \quad [2.54]$$

$$\hat{\phi}^{(g)} = h(w^{(g)}_m) \quad [2.55]$$

Para una clase c dada por su muestra de m_c pixels obtenidos en el crecimiento de regiones :

$$\lim_{w^{(g)c} \rightarrow w^{(g)c}_{m_c}} B(w^{(g)c}) = \hat{\phi}^{(g)c} \quad [2.56]$$

$$\hat{\phi}^{(g)c} = h(w^{(g)c}_{m_c}) \quad [2.57]$$

II.3.2 La clasificación

Para cada clase se conserva el tamaño de su ventana óptima v^c_i asociada a su respectivo w^c_i perteneciente al dominio W^c de la función objetivo B . Se conservan además las estadísticas obtenidas a partir del crecimiento de regiones, es decir, las $h(w^{(g)c}_{m_c}) = \hat{\phi}^{(g)c}$. En la clasificación ya no se emplea la función objetivo B . El propósito de evaluar la función B es encontrar los $\hat{\phi}^{(g)c}$.

En esta etapa se busca minimizar una función de energía E que depende de una función potencial \mathcal{P} y del parámetro $\hat{\phi}^{(g)}$. En primer lugar, se define el dominio de la

función potencial \mathcal{P} , luego la función potencial propiamente dicha y por último se enuncia la función de energía E .

II.3.2.1 La función potencial

Sea $p_b(x,y)$ un pixel perteneciente a una imagen $g_b(x,y)$ de tamaño $N \times M$ cuyas coordenadas son (x,y) tales que $x \in [1,N] \subset \mathbb{N}$, $y \in [1,M] \subset \mathbb{N}$. Sean c clases incluidas en la imagen g representadas por el vector $\hat{\phi}^{(g)}$.

Sea el conjunto V_i formado por los tamaños de ventana óptimos asociados a cada una de las c clases.

$$V_i = \{v_i^1, \dots, v_i^c\}$$

Sea el vector $v(x,y) = [v^1(x,y), \dots, v^c(x,y)]$ tal que el elemento $v^i(x,y)$ es la ventana con centro en las coordenadas x e y del pixel $p_b(x,y)$ del tamaño de la ventana óptima v_i^1 de la clase i .

Sea el arreglo $[v(x,y)]$ formado por todos los vectores $v(x,y)$ que son subconjuntos de la imagen $g_b(x,y)$, el dominio de la función potencial. A través de una función Q se asocia $[v(x,y)]$ al arreglo de etiquetas $[w(x,y)]$.

Sea $m(v) = h[Q(v)]$ el criterio de homogeneidad empleado para llevar a cabo la segmentación.

La función potencial \mathcal{P} se define de la siguiente manera:

$$\mathcal{P} = \mathcal{P}([v(x,y)]) = m([v(x,y)]) = h([w(x,y)]) = [p_{oi}(x,y)]_{N \times M} \quad [2.58]$$

donde $v(x,y)$, $w(x,y)$ y $p_{oi}(x,y)$ son vectores de c elementos.

II.3.2.2 La función de energía

A cada pixel $p_b(x,y)$ se le asocia un vector de energía local $\hat{E}_{p(x,y)}$ de tamaño c tal que:

$$\hat{E}_{p(x,y)} = [E_{p(x,y)}^{(c)}] = [|p_{oi}(x,y)^{(c)} - \hat{\phi}^{(g)c}|] = |p_{oi}(x,y) - \hat{\phi}^{(g)}|^2 \quad [2.59]$$

La energía global se define como:

$$E = E(\mathcal{P}, \hat{\phi}^{(g)}) = [E_{p(x,y)}] = \sum_{x=1}^N \sum_{y=1}^M E_{p(x,y)} \delta(x,y) \quad [2.60]$$

donde δ es la delta de Dirac.

II.3.2.3 La minimización global de la energía

El problema de la asignación de los pixels a la clase más cercana es un problema de minimización de energía. La energía global mínima se obtiene a través de la minimización de las energías locales.

El pixel $p(x, y)$ es asignado a la clase c si $E_{p(x, y)}^{(c)}$ es el menor elemento del vector de energía $E_{p(x, y)}$ asociado a $p(x, y)$

$$p \rightarrow \text{clase } c : E_{p(x, y)}^{(c)} = \min(E_{p(x, y)}) \quad [2.61]$$

La energía global mínima es entonces :

$$\text{Mín}(E) = [\min(E_{p(x, y)})] \quad [2.62]$$

II.3.2.4 Fundamentos del clasificador.

Los pasos básicos del clasificador son los siguientes:

- 1) Se cargan en memoria RAM las bandas elegidas para la clasificación. Se emplean usualmente bandas decorrelacionadas en el caso de imágenes multispectrales y filtradas de ruido en caso de imágenes de radar.
- 2) Se siembran los pixels en los lugares elegidos de cada clase definida para la segmentación.
- 3) La ventana óptima y el parámetro de uniformidad espectral son valores derivados de cada clase definida en función de los parámetros Ω y η .
- 4) Se aplica el algoritmo de crecimiento de regiones a los pixels seleccionados, empleando parámetros optimizados de acuerdo al paso anterior.
- 5) Para cada región crecida, se calcula la función criterio de homogeneidad. Cada región representa una clase y es una muestra estadística válida del comportamiento de la dicha clase.
- 6) Se consideran un conjunto de vecindades centradas en un pixel p y cuyos tamaños se corresponden con los tamaños de ventana óptimos de las clases para clasificar los pixels de la imagen. Para cada vecindad de dicho pixel se calcula la función criterio de homogeneidad.
- 7) Un pixel $p_{k,l}$ de la imagen multispectral se clasifica de acuerdo a lo siguiente:

$$P_{k,l} \rightarrow \text{clase } k : |p_{\alpha}(x, y)^{(k)} - \hat{\phi}^{(g)k}| \text{ es mínimo } \forall k \quad [2.63]$$

II.3.3 Clasificadores espectrales

II.3.3.1 El clasificador espectral por medias.

Sea M la función media normalizada que se emplea como función criterio de homogeneidad h . Para una clase i dada por su pixel prototipo p_0^i , se define el conjunto W^i cuyos valores están asociados a ventanas cuadradas centradas en p_0^i .

$$h(w^i) = M(w^i) \quad [2.64]$$

Se define la función objetivo para el cálculo de la ventana óptima de la clase i como:

$$B: w^i \rightarrow h(w^i)$$

Se asigna al parámetro de estabilidad Ω el valor 0.03.

La ventana óptima para la clase i es:

$$v_i^j = Q^{-1}(\min(w^i)) /$$

$$w_i^j = Q(v_i^j) \in W^i \wedge \left| \frac{dB(w)}{dw} \right| = |M(w^{i+1}) - M(w^i)| = |b^i(w^i)| < \Omega = 0.03 \quad [2.65]$$

Conociendo esta ventana se calcula el valor estimado de $\hat{\phi}^i = M(w^i)$ para la clase i y se definen los extremos del dominio del estimador ϕ para la clase i , o lo que es equivalente, el umbral ξ . Para ello, se emplea un umbral de acotación η igual a uno [Ayans & Biging, 1996] y una función f igual al desvío estándar D calculado sobre la ventana óptima de la clase i .

$$\begin{aligned} \varepsilon &= \eta f \\ f &= D(w^i) \\ \eta &= 1 \end{aligned} \quad [2.66]$$

Con esto, el criterio de homogeneidad para el crecimiento de regiones de la clase i es el siguiente:

$$|M(w^i) - M(w^i)| \leq D(w^i) \quad [2.67]$$

Al finalizar los crecimientos se tienen los valores $\hat{\phi}^{(g)i} = M(w^{(g)i})$ del estimador ϕ para cada clase i . Se define la función potencial:

$$P = P([v(x, y)]) = M([w(x, y)]) = [p_{oi}(x, y)]_{N \times M}$$

Y con esto la función de energía a minimizar durante el proceso de clasificación:

$$\bar{E}_{p(x, y)} = [\bar{E}_{p(x, y)}^{(c)}] = [| p_{oi}(x, y)^{(c)} - \hat{\phi}^{(g)c} |] = | p_{oi}(x, y) - \hat{\phi}^{(g)} |^2 = | M(w(x, y)) - M(w^{(g)i}) | \quad [2.68]$$

II.3.3.2 El clasificador espectral por histogramas.

En el clasificador espectral por histogramas se emplea la función histograma conjunto normalizado [Cetin & Levandowski, 1991] como atributo relevante. El histograma porta información del desvío estándar, puesto que este último puede calcularse a partir del primero. La ventana óptima se calcula empleando únicamente el histograma, pero en el crecimiento de regiones, se define un criterio de homogeneidad que consta de ambos atributos. Los extremos del dominio del atributo histograma se definen en función de la distancia mínima de un histograma a otro. Si las clases no son equidistantes, se corre el riesgo de incorporar en las regiones crecidas pixels que no corresponden a las mismas, por ello necesita emplearse otro atributo que los acote: el desvío estándar.

Sea entonces H la función histograma normalizado que se emplea como función criterio de homogeneidad h_i . Para una clase i dada por su pixel prototipo p^i_0 , se define el conjunto W^i cuyos valores están asociados a ventanas cuadradas centradas en p^i_0 .

$$h_i(w^j) = H(w^j) \quad [2.69]$$

H = histograma normalizado

Se define la función objetivo para el cálculo de la ventana óptima de la clase i como:

$$B: w^j \rightarrow h_i(w^j)$$

Se asigna al parámetro de estabilidad Ω el valor 0.03.

La ventana óptima para la clase i es:

$$v^i_t = Q^{-1}(\min(w^i_t)) /$$

$$w^i_t = Q(v^i_t) \in W^i \wedge \left| \frac{dB(w)}{dw} \right| = |H(w^{i+1}) - H(w^i)| = |b^i(w^i)| < \Omega = 0.03 \quad [2.70]$$

Conociendo esta ventana se calculan los umbrales óptimos para cada clase. En este clasificador se trabaja con dos umbrales para cada clase. El primer umbral es una función de distancia de cada clase con las demás y el segundo umbral se obtiene de datos propios de la clase.

Se calculan entonces el valor estimado de $\hat{\phi}_1^i = h_i(w^i_t)$ para la clase i , y el valor $\hat{\phi}_2^i = h_2(w^i_t) = D(w^i_t)$ (donde D es el desvío estándar) y se definen los extremos de los dominios de los estimadores ϕ_1 y ϕ_2 para la clase i , o lo que es equivalente, los umbrales ξ_1 y ξ_2 . Para ello, se emplea un umbral de acotación η igual a 0.75 y funciones f_1 e f_2 que a continuación se definen:

$$\eta = 0.75$$

$$\varepsilon_1 = \eta f_1 \quad [2.71]$$

$$f_1 = \min(d_i) = \min(|H(w^i_t) - H(w^j_t)|) \quad \forall j \neq i \in [1, c]$$

$$\varepsilon_2 = (1 - \eta) f_2$$

$$f_2 = D(w^i_t)$$

Con esto, los criterios de homogeneidad para el crecimiento de regiones de la clase i son los siguientes:

$$\begin{aligned} |H(w^j) - H(w^i)| &\leq 0.75 \min(|H(w^j) - H(w^i)|) \quad \forall j \neq i \in [1, c] \\ |D(w^j) - D(w^i)| &\leq 0.25 D(w^i) \end{aligned}$$

La función objetivo que se emplea en el crecimiento de regiones es la siguiente:

$$B^{(g)}: w^{(g)i} \rightarrow H(w^{(g)i})$$

Los parámetros obtenidos en el crecimiento de regiones son los histogramas de las clases crecidas y se denotan de la siguiente manera:

$$\hat{\phi}_1^{(g)i} = h_1(w^{(g)i}) = H(w^{(g)i})$$

Se define la función potencial:

$$P = P([v(x, y)]) = H([w(x, y)]) = [p_{oi}(x, y)]_{N \times M}$$

Y con esto la función de energía a minimizar durante el proceso de clasificación:

$$\bar{E}_{p(x, y)} = [\bar{E}_{p(x, y)}^{(c)}] = [| p_{oi}(x, y)^{(c)} - \hat{\phi}^{(g)c} |] = | p_{oi}(x, y) - \hat{\phi}^{(g)c} |^2 = | H(w(x, y)) - H(w^{(g)c}) | \quad [2.72]$$

Ejemplo de función objetivo

Se tiene una imagen g de 512 pixels por 512 pixels de 256 niveles de gris, en cuyo centro existe una clase C que tiene un tamaño de 256 pixels por 256 pixels. Alrededor de dicha clase, existen otras clases. El pixel central de la imagen es el pixel prototipo p^c_0 de esta clase c . Sea $B(w)$ la función objetivo que calcula el histograma normalizado de cada ventana de tamaño v centrada en p^c_0 asociada a w .

La imagen que se presenta a en la Figura 2.10 muestra de menor a mayor tamaño de ventana, siguiendo desde el borde superior al borde inferior, los histogramas normalizados y reescalados a 256 para cada ventana v centrada en p^c_0 .

Puede apreciarse muy cerca del borde superior de la imagen, que los cambios en los histogramas son grandes. Luego, comienza una zona en la cual los valores comienzan a mantenerse estables. Al final de dicha zona, se nota un cambio brusco en los valores de los histogramas. Hacia el borde inferior de la imagen, los valores de los histogramas tienden a suavizarse. En la última línea se tiene el histograma de la imagen: un parámetro de la misma.

En la imagen de la Figura 2.11 se presenta la primera derivada $b^c(w)$ de la función objetivo de la clase c tomada en valor absoluto y reescalada a 256. Los grandes cambios que se producen en los histogramas normalizados al pasar de un tamaño de ventana

En la imagen de la Figura 2.11 se presenta la primera derivada $b^c(w)$ de la función objetivo de la clase c tomada en valor absoluto y reescalada a 256. Los grandes cambios que se producen en los histogramas normalizados al pasar de un tamaño de ventana pequeño a otro tamaño de ventana pequeño se pueden apreciar en los colores claros del borde superior. Más cerca del fin de la clase, se ve que los cambios tienden a ser mínimos: en estos tamaños de ventana, la energía tiende a minimizarse. Al comenzar a tomar las ventanas pixels de las clases vecinas, se producen cambios relativamente bruscos entre histogramas normalizados consecutivos. Esto se puede apreciar en la imagen siguiente si se considera una recta horizontal imaginaria que pasa a través de la misma por su parte central. Luego, más cerca del borde inferior, ya no hay cambios, puesto que los datos tienden a representar el histograma de la imagen: uno de sus parámetros.

Por último, la imagen que se presenta en la Figura 2.12 se tiene la segunda derivada de la función objetivo de la clase c (reescalada a 256 niveles), en la cual puede apreciarse el fin de la clase c , puesto que se ve claramente un cambio brusco de concavidad en esa posición.



Figura 2.10

Figura 2.11

Figura 2.12

II.3.4 Clasificadores texturales

II.3.4.1 El clasificador textural por matrices de co-ocurrencia

El clasificador textural por matrices de co-ocurrencia no es más que una generalización del clasificador por histogramas.

Sea C el atributo matriz de co-ocurrencia que se emplea como función criterio de homogeneidad h . Para una clase i dada por su pixel prototipo $p_{i_0}^i$, se define el conjunto W^i cuyos valores están asociados a ventanas cuadradas centradas en $p_{i_0}^i$.

$$h_1(w^i) = C(w^i) \quad [2.73]$$

donde C es la matriz de co-ocurrencia.

$$B: w^i \rightarrow C(w^i)$$

Se asigna al parámetro de estabilidad Ω el valor 0.03.
La ventana óptima para la clase i es:

$$v_i^i = Q^{-1}(\min(w^i)) /$$

$$w_i^i = Q(v_i^i) \in W^i \wedge \left| \frac{dB(w)}{dw} \right| = |C(w^{i+1}) - C(w^i)| = |b^i(w^i)| < \Omega = 0.03 \quad [2.74]$$

Conociendo esta ventana se calcula el valor estimado de $\hat{\phi}_1^i = h_1(w^i)$ para la clase i , y el valor $\hat{\phi}_2^i = h_2(w^i) = D(w^i)$ (donde D es el desvío estándar) y se definen los extremos de los dominios de los estimadores ϕ_1 y ϕ_2 para la clase i , o lo que es equivalente, los umbrales ξ_1 y ξ_2 . Para ello, se emplea un umbral de acotación η igual a 0.75 y funciones f_1 e f_2 que a continuación se definen:

$$\Lambda = 0.75 \quad [2.75]$$

$$\epsilon_1 = \eta f_1$$

$$f_1 = \min(d_i) = \min(|H(w^i) - H(w^j)|) \quad \forall j \neq i \in [1, c]$$

$$\epsilon_2 = (1 - \eta) f_2$$

$$f_2 = D(w^i)$$

Con esto, los criterios de homogeneidad para el crecimiento de regiones de la clase i son los siguientes:

$$|C(w^i) - C(w^j)| \leq 0.75 \min(|C(w^i) - C(w^j)|) \quad \forall j \neq i \in [1, c]$$

$$|D(w^i) - D(w^j)| \leq 0.25 D(w^i)$$

La función objetivo que se emplea en el crecimiento de regiones es la siguiente:

$$B^{(g)}: w^{(g)i} \rightarrow C(w^{(g)i})$$

Los parámetros obtenidos en el crecimiento de regiones son los histogramas de las clases crecidas y se denotan de la siguiente manera:

$$\hat{\phi}_1^{(g)i} = h_1(w^{(g)i}) = C(w^{(g)i})$$

Se define la función potencial:

$$P = P([v(x, y)]) = C([w(x, y)]) = [p_0(x, y)]_{N \times M}$$

Y con esto la función de energía a minimizar durante el proceso de clasificación:

$$\bar{E}_{p(x,y)} = [\bar{E}_{p(x,y)}^{(c)}] = [[p_{\alpha}(x,y)^{(c)} - \hat{\phi}^{(g)c}]] = | p_{\alpha}(x,y) - \hat{\phi}^{(g)} |^2 = | C(w(x,y)) - C(w^{(g)}) | \quad [2.76]$$

II.3.4.2 El clasificador textural por centros de masa de matrices de co-ocurrencia

El centro de masa de una función de densidad es aquel valor de abscisa para el cual la probabilidad acumulada tiene el valor 0.5. En el caso de una función de densidad bidimensional, se tiene un vector cuyos valores son las medias o centros de masa para cada uno de los ejes.

Este clasificador se basa en la propiedad de simetría de la matriz de coocurrencia: la función de densidad bidimensional que se obtiene es simétrica respecto a la diagonal principal de tal matriz. Se refleja la parte inferior izquierda de la matriz de co-ocurrencia sobre su diagonal principal, y se suma tal parte sobre la parte superior derecha. De este modo, se tiene una nueva función de densidad, de la que se obtienen para cada eje los centros de masa.

En términos prácticos, el procedimiento consiste en armar la matriz de co-ocurrencia de la siguiente manera: se cuentan los pixels en aquellas celdas cuyas posiciones están dadas por subíndices en los cuales el primero siempre es menor o igual que el segundo subíndice. Toda vez que se tenga un par de pixels (u,v) donde u es mayor que v, éste se cuenta en la celda de subíndice (v,u).

Este clasificador tiene fundamentos similares al clasificador espectral por medias, puesto que los umbrales se calculan también empleando el desvío estándar. En este caso específico se emplean las medias o centros de masa de la matriz de co-ocurrencia y el desvío estándar de la matriz de co-ocurrencia.

Sea entonces M la función centro de masa normalizado que se emplea como función criterio de homogeneidad h. Para una clase i dada por su pixel prototipo p^i_0 , se define el conjunto W^i cuyos valores están asociados a ventanas cuadradas centradas en p^i_0 .

$$h(w^i) = S(w^i) \quad [2.77]$$

Se define la función objetivo para el cálculo de la ventana óptima de la clase i como:

$$B: w^i \rightarrow h(w^i)$$

Se asigna al parámetro de estabilidad Ω el valor 0.03.

La ventana óptima para la clase i es:

$$v^i = Q^{-1}(\min(w^i)) / w^i = Q(v^i) \in W^i \wedge \left| \frac{dB(w)}{dw} \right| = |S(w^{i+1}) - S(w^i)| = b^i(w^i) < \Omega = 0.03 \quad [2.78]$$

Conociendo esta ventana se calcula el valor estimado de $\hat{\phi}^i = M(w^i)$ para la clase i y se definen los extremos del dominio del estimador ϕ para la clase i, o lo que es

equivalente, el umbral ξ . Para ello, se emplea un umbral de acotación η igual a uno y una función f igual al desvío estándar D calculado sobre la matriz de co-ocurrencia de la clase i .

$$\begin{aligned} \varepsilon &= \eta f \\ f &= D(w^i) \\ \eta &= 1 \end{aligned}$$

Con esto, el criterio de homogeneidad para el crecimiento de regiones de la clase i es el siguiente:

$$|S(w^j) - S(w^i)| \leq D(w^i) \quad [2.79]$$

Al finalizar los crecimientos se tienen los valores $\hat{\phi}^{(g)i} = M(w^{(g)i})$ del estimador ϕ para cada clase i . Se define la función potencial:

$$P = P([\varphi(x, y)]) = S([\varpi(x, y)]) = [p_{\alpha}(x, y)]_{N \times M}$$

Y con esto la función de energía a minimizar durante el proceso de clasificación:

$$\bar{E}_{p(x, y)} = [\bar{E}_{p(x, y)}^{(c)}] = [p_{\alpha}(x, y)^{(c)} - \hat{\phi}^{(g)c}] = |p_{\alpha}(x, y) - \hat{\phi}^{(g)}|^2 = |S(\varpi(x, y)) - S(w^{(g)})| \quad [2.80]$$

CAPITULO III

RESULTADOS Y DISCUSION

Para evaluar el potencial de los clasificadores propuestos en este trabajo, se construyó una serie de ejemplos consistente básicamente de imágenes sintéticas e imágenes de radar. El desglose de estos ejemplos es el siguiente:

a) Un conjunto de treinta y seis imágenes sintéticas cuyos parámetros estadísticos son conocidos. Las imágenes sintéticas son de una sola banda y contienen seis clases cada una. El tamaño de estas imágenes es de 192 pixels por 256 pixels.

b) Una imagen multispectral SPOT a la cual se le aplicó previamente la descomposición en componentes principales. El tamaño de esta imagen es de 512x512 pixels; el tamaño del pixel es de 20x20 m² y cubre una porción de México Central.

c) Una imagen de radar SEASAT filtrada de speckle con el filtro geométrico [Pastrana, 1996]. El tamaño de esta imagen es de 512x512 pixels; el tamaño de cada pixel es de 25x25m² y cubre una porción del norte de México.

III.1 Las bandas sintéticas

Las bandas sintéticas empleadas tienen parámetros estadísticos conocidos. El objetivo del empleo de tales bandas es evaluar, de una manera controlada, los algoritmos implantados. Existen varios criterios de evaluación, entre ellos se destacan la precisión, o porcentaje de aciertos sobre número de casos, y el coeficiente Kappa, que depende de una matriz de confusión. Ambos criterios se detallan a continuación.

III.1.1 Los criterios de evaluación

III.1.1.1 La precisión calculada sobre pixel clasificado

Se determina para cada pixel clasificado si es un acierto o un fallo. Un pixel es acertadamente clasificado si es etiquetada a la clase para la cual fue designado; en caso contrario, es un fallo. Los fallos son considerados confusiones con otras clases. Si un pixel fue asignado a una clase diferente a la que pertenece, se determina a cuál de ellas. Se cuenta el número de confusiones de una clase con otra, para todos los pares de clases. Con esto, se tiene una matriz denominada "de confusión" de dimensión igual a la cantidad de clases. Las celdas de su diagonal principal contienen valores que representan los aciertos para cada clase.

Se cuentan para cada clase el número de aciertos y se divide por la cantidad de pixels correspondientes a dicha clase que fueron clasificados, es decir, el número de aciertos más el número de fallos. Este valor es la precisión resultante para una clase dada. Se calcula un promedio de precisiones entre todas las clases y esta es la precisión global del algoritmo.

III.1.1.2 El coeficiente Kappa

El coeficiente Kappa [Cohen, 1960; Congalton, 1983; Hudson & Ramm, 1987] es un valor que se obtiene a partir de la matriz de confusión. Los valores de la matriz de confusión se normalizan dividiendo por el número total de muestras N contadas en la matriz de confusión, de manera que la suma de todos los elementos de la matriz de confusión se hace igual a uno.

Se suman los valores de las celdas de la matriz de confusión por fila y por columna en vectores que contienen estos totales. Se efectúa el producto escalar de dicho vectores y el resultado se denota con la letra q. La suma de los elementos de la diagonal de la matriz de confusión se denota con d. El coeficiente Kappa se define de la siguiente manera:

$$K = \frac{d - q}{1 - q} \quad [3.1]$$

Una clasificación perfecta arroja un coeficiente Kappa igual a uno. Una clasificación completamente errónea produce un coeficiente Kappa negativo.

III.1.2 La generación de las imágenes sintéticas

III.1.2.1 El generador de funciones de densidad Rayleigh y la generación de las clases

Las clases se obtienen a partir de un generador de funciones de densidad Rayleigh. Se emplea este tipo de distribución debido a que la bibliografía reporta que, en general, los histogramas de las clases tienen una forma similar a la curva Rayleigh.

Al generador de funciones de densidad Rayleigh deben ingresarse dos parámetros: el valor de la abcisa para el cual la función de densidad Rayleigh es igual a cero, y el desvío estándar, que se utiliza también para calcular el incremento de abcisa. Dado un valor de abcisa, comenzando desde cero, el valor de la ordenada se calcula mediante las siguientes expresiones:

$$k = -0.5\left(\frac{x}{\text{desvio}}\right)^2 \quad [3.2]$$

$$y = \frac{x}{100.\text{desvio}} e^k \quad [3.3]$$

Se generan cuatrocientos valores de abcisa "x" y se calculan sus respectivas ordenadas "y". Para cada valor x de la abcisa se guarda el valor acumulado "a" de las ordenadas y. Al vector x se le suma el valor de la abcisa para el cual la función de densidad Rayleigh es igual a cero, con el fin de efectuar una traslación de la función Rayleigh sobre el eje de las abcisas. Los vectores de abcisas x y de acumulados a constituyen la salida del generador de funciones Rayleigh.

Una clase se genera de la siguiente manera:

- 1) Se introducen en el generador los dos parámetros antes mencionados, que representan la clase que se desea generar.
- 2) El generador calcula el vector "a", que representa la función de distribución Rayleigh, y el vector de abcisas "x" correspondientes.
- 3) Se genera un número aleatorio n entre 0 y 1 empleando una distribución uniforme.
- 4) El valor de "x" correspondiente al menor valor del vector de acumulados "a" que es mayor que el número n es el valor del nivel de gris asociado al pixel a generar.
- 5) Se repiten los pasos anteriores hasta que se haya completado el número de pixels que se desean generar para la clase dada.

II.1.2.2 Los parámetros estadísticos empleados para las clases

Se generan treinta y seis bandas con seis clases cada una, es decir, un total de 216 clases con parámetros estadísticos diferentes. Cada banda tiene un tamaño de 192 pixels por 256 pixels y el tamaño de cada clase es de 64 pixels por 128 pixels. En cada banda, las clases tienen la misma separación entre medias y el mismo desvío estándar. Considerando un orden ascendente de media, se numeran de uno a seis. La estructura de las clases en una banda dada es la siguiente: de izquierda a derecha, en la parte superior se encuentran las clases uno, tres y cinco, y en la parte inferior, las clases dos, cuatro y seis.

Los parámetros empleados para la generación de las clases se resumen en la tabla 3.1. En cada renglón de dicha tabla se tienen seis números que constituyen los valores de abcisas para los cuales la función Rayleigh vale cero. En cada columna se tiene un valor de desvío estándar, que es del orden de 2^n , donde n varía desde cero hasta cinco. Para la generación de una banda se emplean los valores de un renglón de la tabla y un valor de una columna. De esta manera, para una banda dada, la diferencia entre medias consecutivas en el eje nivel de gris de un renglón dado es constante.

Los renglones de la tabla se enumeran desde uno hasta seis, y a los valores de desvío estándar se le asigna un índice que varía también desde uno hasta seis. Una banda sintética se almacena con un nombre que es la combinación de los índices asociados a la separación entre medias (IndiceMedia) y al desvío estándar (IndiceDesvio) mediante la siguiente expresión:

$$\text{Nombre} = \text{IndiceDesvio} * 10 + \text{IndiceMedia} \quad [3.4]$$

A continuación se presenta la tabla 3.1 que relaciona los nombres de las bandas generadas con los parámetros asociados a las clases que las componen:

Valores iniciales de abcisa	Separac. entre medias	$\sigma=1$	$\sigma=2$	$\sigma=4$	$\sigma=8$	$\sigma=16$	$\sigma=32$
16 48 80 112 144 176	32	11	21	31	41	51	61
72 88 104 120 136 152	16	12	22	32	42	52	62
100 108 116 124 132 140	8	13	23	33	43	53	63
114 118 122 126 130 134	4	14	24	34	44	54	64
121 123 125 127 129 131	2	15	25	35	45	55	65
124 125 126 127 128 129	1	16	26	36	46	56	66

Tabla 3.1

Entonces, sea por ejemplo, la banda de nombre 32. Esta banda contiene clases cuyo desvío estándar es 4 (puesto que el índice de desvío es 3) y cuya separación entre medias es 16 (puesto que el índice de separación entre medias es 2).

II.1.3 El simulador

A efectos de introducir mayor aleatoriedad en la distribución de las clases en una banda dada, se emplea un simulador que genera imágenes multiespectrales sintéticas a partir de las bandas generadas según los párrafos que anteceden, las segmenta según un criterio preestablecido y calcula un valor de precisión.

A grandes rasgos, el simulador trabaja de la siguiente manera:

- 1) Genera una imagen multiespectral sintética con clases distribuidas aleatoriamente.
- 2) Aplica sobre esta imagen un mismo algoritmo de segmentación especificado.
- 3) Calcula la precisión de dicho algoritmo y la acumula.
- 4) Repite los pasos anteriores tantas veces como iteraciones se hayan determinado con anterioridad. El número de iteraciones indica la cantidad de imágenes que se desean segmentar.
- 5) Por último, calcula la precisión promedio del algoritmo, dividiendo el acumulado de precisiones por el número de iteraciones realizadas.

A continuación se detallan, en primer lugar, las entradas al simulador, luego la generación de imágenes multiespectrales sintéticas, y por último las salidas del mismo.

II.1.3.1 Las entradas al simulador

Deben especificarse los siguientes datos:

- Cantidad de bandas
- Número de iteraciones
- Criterio de homogeneidad (el algoritmo de segmentación).
- Parámetros del criterio de homogeneidad

II.1.3.2 La generación de imágenes multiespectrales sintéticas

Con el objeto de efectuar pruebas con los algoritmos implantados, se utiliza un generador de imágenes multiespectrales sintéticas (GIMS), el cual emplea las bandas obtenidas en la etapa anterior con el generador de funciones de densidad Rayleigh. Para evitar confusiones, debido a que estas bandas se almacenan en un dispositivo físico, se las denominará "bandas almacenadas". A las bandas que sean salida del GIMS, debido a que sólo se almacenan en memoria RAM, se las denominará "bandas en memoria".

Para producir una banda en memoria, el GIMS necesita dos números aleatorios entre uno y seis, para establecer el nombre de una banda almacenada y otro número aleatorio entre uno y seis más para determinar el número de clase de dicha banda almacenada que extraerá para constituir la primera clase de la banda en memoria. El GIMS repite el paso anterior hasta completar seis clases de la banda en memoria.

Una imagen multiespectral sintética se genera de la manera explicada en el párrafo anterior con tantas bandas en memoria como cantidad de bandas se hayan especificado. Si la cantidad de bandas para cada imagen multiespectral es mayor que uno, y el operador solicita que las bandas estén decorrelacionadas, antes de pasar al paso siguiente, el GIMS aplica a cada imagen generada la descomposición en componentes principales.

II.1.3.3 Los procesos realizados sobre las imágenes multiespectrales sintéticas

Sobre cada imagen multiespectral sintética se siembran automáticamente los píxeles prototipo en las coordenadas del centro de la clase. Se inicia el algoritmo de segmentación por crecimiento de regiones sobre la imagen generada, aplicando el criterio de homogeneidad especificado con sus correspondientes parámetros. Al final de esta etapa se inicia un proceso de evaluación de la precisión del algoritmo de segmentación para esta imagen.

Los pasos anteriores se efectúan tantas veces como número de iteraciones de la simulación se hayan especificado.

A modo de resumen, se realizan los siguientes comentarios respecto de los algoritmos evaluados:

En cuanto al cálculo de la ventana óptima:

- Cuando se emplean ventanas promedio, éstas se calculan a partir de datos de todas las clases.
- Cuando se emplean ventanas por clase, cada una se obtiene a partir de los datos de la clase a la que está relacionada.

En cuanto al cálculo del umbral óptimo para el crecimiento de regiones:

- El algoritmo de diferencia de medias define los umbrales para el crecimiento con datos propios de cada clase.
- El algoritmo de diferencia de histogramas define los umbrales para el crecimiento a partir de datos propios de la clase y a partir de datos de las demás clases.

En cuanto al cálculo del umbral para la segmentación:

- En el algoritmo de diferencia de medias con ventana promedio los pixels son clasificados a la clase más cercana.
- En el algoritmo de diferencia de medias con ventana por clase, los umbrales se definen a partir de los crecimientos.
- En los algoritmos de diferencia de histogramas, ya sea con ventana promedio o con ventanas por clase, los pixels son clasificados en la clase más cercana, si su histograma asociado se intersecta con el de la clase.

II.1.3.4 Las salidas

Las salidas de la simulación son valores de precisión promedio y kappa promedio, calculados empleando las expresiones mencionadas en la parte III.1.1.

III.1.4 Las simulaciones realizadas

En primer lugar, se efectuaron segmentaciones sobre las bandas sintéticas almacenadas. Esto permitió observar detalladamente los crecimientos de regiones para cada clase y establecer parámetros para efectuar crecimientos suficientes, lo menos invasivos posibles. Un crecimiento es invasivo cuando toma pixels de sus clases vecinas. En segundo lugar, se efectuaron las simulaciones mencionadas.

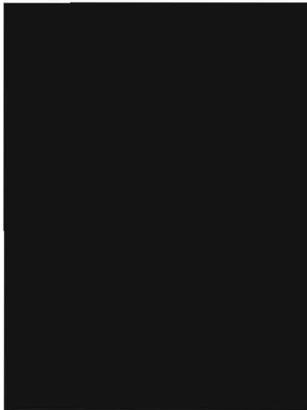
En todos los casos se sembraron los pixels prototipo en las posiciones que se muestran en la siguiente tabla (3.2):

Clase	Columna	Fila
1	32	64
2	32	192
3	96	64
4	96	192
5	160	64
6	160	192

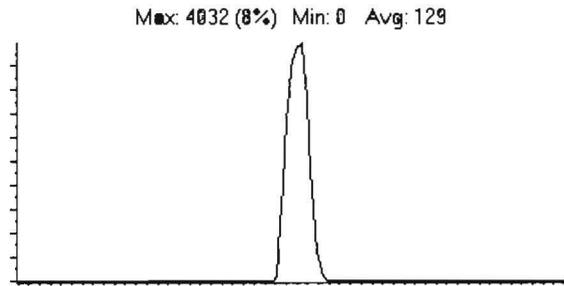
Tabla 3.2

Nota: Cada posición coincide con el centro geométrico de la clase respectiva.

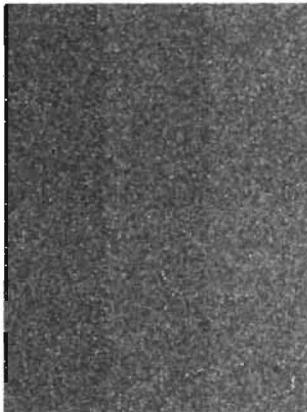
A modo ilustrativo se presentarán los resultados obtenidos por cada algoritmo para tres ejemplos. En primer lugar, se considerarán las bandas sintéticas almacenadas 35 (Imagen 3.1) y 64 (Imagen 3.2). En el caso 35 las clases tienen una separación entre medias de 2 y un desvío estándar de 4. En el caso 64 las clases tienen una separación entre medias de 4 y un desvío estándar de 32. Los histogramas que pueden apreciarse en las figuras 3.1 y 3.2 se corresponden con las imágenes 3.1 y 3.2 respectivamente. El tercer ejemplo es la imagen sintética denominada "ndi" (Imagen 3.3) cuyas clases tienen distribución Rayleigh con el mismo valor inicial de abcisa, pero con desvíos 1, 2, 4, 8, 16 y 32; su histograma es el de la figura 3.3.



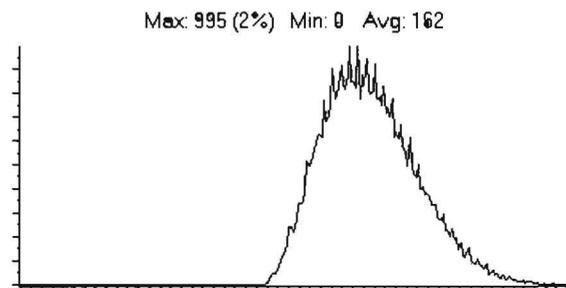
Caso 35
Imagen 3.1



Histograma de la imagen 3.1
Figura 3.1



Caso 64
Imagen 3.2



Histograma de la imagen 3.2
Figura 3.2

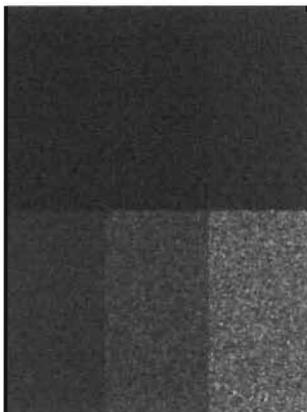
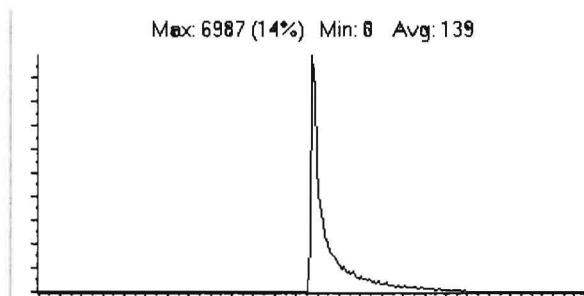


Imagen ndi
Imagen 3.3



Histograma de la imagen 3.3
Figura 3.3

III.1.4.1 Los resultados sobre las bandas sintéticas almacenadas

III.1.4.1.1 Algoritmo de segmentación por medias espectrales

Se muestran a continuación los resultados del algoritmo de diferencia de medias, empleando ventanas óptimas promedio y ventanas óptimas por clase. En ambos casos, los umbrales para los crecimientos se calcularon tomando la media más menos un desvío estándar.

III.1.4.1.1.1 Algoritmo de segmentación por medias espectrales con ventana óptima promedio

Los tamaños de ventana óptimos obtenidos en cada caso se presentan en la siguiente tabla (3.3):

Separación entre medias/Desvío estándar	32	16	8	4	2	1
1	25	3	11	13	7	37
2	5	3	3	15	3	11
4	9	7	11	3	5	23
8	5	7	59	23	5	11
16	19	9	19	9	61	33
32	9	9	37	59	13	35

Tabla 3.3

Los datos de la tabla 3.3 no tienen un comportamiento estadístico aparente, puesto que bandas con clases cuya separación entre medias es muy grande pueden tener tamaños de ventana muy grandes, y clases con separación entre medias muy pequeña pueden tener ventanas muy grandes también. Lo mismo sucede desde el punto de vista del desvío estándar.

Los valores de precisión y una gráfica que ilustra su comportamiento se presentan en la figura 3.4.

Los crecimientos de regiones y la segmentación con este algoritmo del caso 35, el caso 64 y la imagen ndi se muestran en las imágenes 3.4 y 3.5, 3.6 y 3.7 y 3.8 y 3.9 respectivamente.

III.1.4.1.1.2 Algoritmo de segmentación por medias espectrales con ventana óptima por clase

Los tamaños de ventana óptimos obtenidos en cada caso con un umbral de estabilidad de 1% fueron los siguientes:

Separación entre medias/Desvío estándar	32	16	8	4	2	1
1	3	3	3	3	3	3
2	3	3	3	3	3	3
4	3	3	3	3	3	3
8	5	4	4	4	5	4
16	4	4	5	4	6	4
32	7	7	5	8	6	6

Tabla 3.4

De la tabla 3.4 se desprende que existe una relación directa entre el desvío estándar y el tamaño de ventana, puesto que a medida que aumenta el desvío estándar, el tamaño de ventana aumenta y viceversa, a medida que disminuye el desvío estándar, el tamaño de ventana disminuye. Cabe destacar que el desvío estándar es un dato inherente a la clase. Nota: Con valores más pequeños del umbral de estabilidad pueden aumentarse los tamaños de ventana.

Los valores de precisión se presentan en la figura 3.5. Los crecimientos de regiones y la segmentación con este algoritmo del caso 35, el caso 64 y la imagen ndi se muestran en las imágenes 3.10 y 3.11, 3.12 y 3.13 y 3.14 y 3.15 respectivamente.

III.1.4.1.2 Algoritmo de segmentación por histogramas conjuntos

Se muestran a continuación los resultados del algoritmo de segmentación por histogramas conjuntos, empleando ventanas óptimas promedio y ventanas óptimas por clase. En ambos casos, el factor a aplicar al umbral para los crecimientos fue del 75 %.

III.1.4.1.2.1 Algoritmo de segmentación por histogramas conjuntos con ventana óptima promedio

Los tamaños de ventana para cada caso fueron los siguientes:

Separación entre Medias/Desvío Estándar	32	16	8	4	2	1
1	3	3	3	3	5	23
2	3	3	3	3	5	13
4	3	3	5	5	9	15
8	3	3	7	9	9	21
16	3	5	11	17	23	3
32	3	3	19	21	43	25

Tabla 3.5

Algoritmo de Segmentación por Medias Espectrales
con ventana óptima promedio

Caso 35

Crecimiento de Regiones

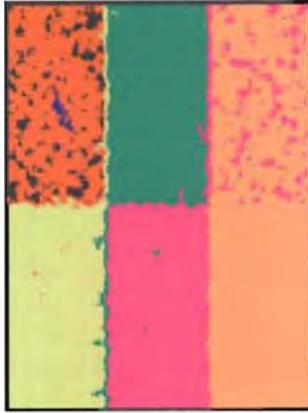


Imagen 3.4

Segmentación



Imagen 3.5

Caso 64

Crecimiento de Regiones



Imagen 3.6

Segmentación

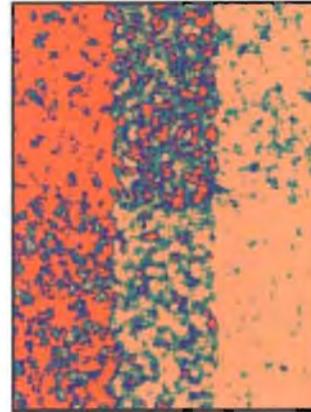


Imagen 3.7

Imagen ndi

Crecimiento de Regiones



Imagen 3.8

Segmentación



Imagen 3.9

Algoritmo de Segmentación por Medias Espectrales
con ventanas óptimas por clase

Caso 35

Crecimiento de Regiones



Imagen 3.10

Segmentación

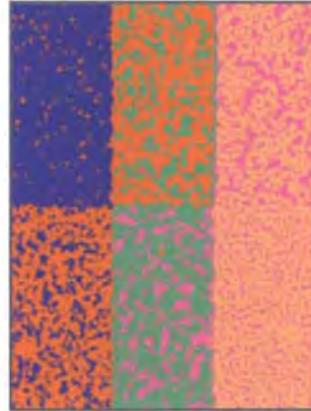


Imagen 3.11

Caso 64

Crecimiento de Regiones



Imagen 3.12

Segmentación

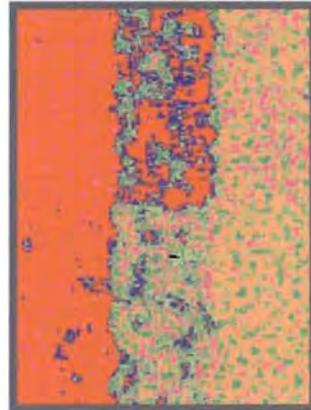


Imagen 3.13

Imagen ndi

Crecimiento de Regiones



Imagen 3.14

Segmentación



Imagen 3.15

Algoritmo de Segmentación por Medias Espectrales

Con ventana óptima promedio
(Umbral de acotación = 1 desvío estándar)

D/(Mi-M)	32	16	8	4	2	1	Promedio
1	85.997335	97.908432	93.567844	92.042717	98.002328	74.713170	90.371971
2	97.874630	97.908432	97.908465	90.771884	95.014015	57.071543	89.424828
4	95.677013	97.910483	95.860847	95.671581	69.358804	47.919928	83.733109
8	97.914681	97.650022	63.145673	86.933184	52.210701	16.666667	69.086822
16	88.960562	95.096504	55.431759	68.443957	16.666667	16.666667	56.877686
32	96.250127	94.373324	32.350013	32.949957	16.666667	16.666667	48.209459
Promedio	93.779058	96.807866	73.044100	77.802214	57.986530	38.284107	72.950646

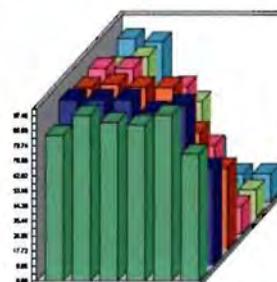


Figura 3.4

Con ventana óptima por clase
(Umbral de acotación = 1 desvío estándar; Umbral de Estabilidad = 1%)

D/(Mi-M)	32	16	8	4	2	1	Promedio
1	97.908399	97.908399	97.908399	97.929035	98.156253	86.790225	96.100119
2	97.908399	97.908399	97.904233	98.010210	96.211639	35.595745	87.256438
4	97.887198	97.885198	97.881398	95.169558	65.556396	34.593814	81.495593
8	97.603825	97.816460	97.343697	83.503678	46.501332	29.239413	75.334734
16	97.705260	94.783546	75.590608	49.096876	15.885246	22.663850	59.254231
32	96.134353	93.016010	50.838794	33.553559	15.219262	17.007650	50.961605
Promedio	97.524572	96.553002	86.244521	76.210486	56.221688	37.648450	75.067120

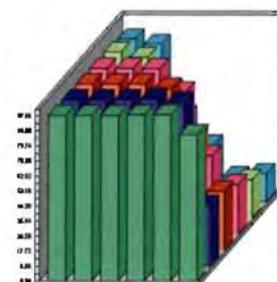


Figura 3.5

Unicamente en el algoritmo de diferencia de histogramas con ventana promedio se observa que, en general, la separación entre las medias de las clases tiene una relación inversa con el tamaño de ventana. A medida que disminuye la primera, el tamaño de ventana aumenta y viceversa, a medida que aumenta la separación entre las medias, el tamaño de la ventana disminuye.

Los valores de precisión se presentan en la figura 3.6. Los crecimientos de regiones y la segmentación con este algoritmo del caso 35, el caso 64 y la imagen ndi se muestran en las imágenes 3.16 y 3.17, 3.18 y 3.19 y 3.20 y 3.21 respectivamente.

III.1.4.1.2.2 Algoritmo de segmentación por histogramas conjuntos con ventana óptima por clase

Los tamaños de ventana para cada caso fueron los siguientes:

Sep. entre Medias/Desvío estándar	32	16	8	4	2	1
1	13	10	9	12	13	12
2	14	15	16	15	17	16
4	21	20	18	20	19	21
8	26	25	25	26	26	26
16	31	33	32	32	33	33
32	41	40	40	41	42	42

Tabla 3.6

Los valores de precisión se presentan en la figura 3.7.

Los crecimientos de regiones y la segmentación con este algoritmo del caso 35, el caso 64 y la imagen ndi se muestran en las imágenes 3.22 y 3.23, 3.24 y 3.25 y 3.26 y 3.27 respectivamente.

III.1.4.1.3 Conclusiones de las pruebas efectuadas con las bandas sintéticas almacenadas

- Cuando se calculan ventanas óptimas por clase, la precisión aumenta.
- Existe una relación directa entre el desvío estándar y el tamaño de ventana en algoritmos con ventana óptima por clase: a medida que aumenta el desvío estándar, el tamaño de ventana aumenta y viceversa, a medida que disminuye el desvío estándar, el tamaño de ventana disminuye. El desvío estándar es una propiedad estadística de la clase.
- Existe una relación directa entre la separación entre las medias y la precisión: a medida que aumenta la separación entre las medias, la precisión aumenta y viceversa, a medida que disminuye la separación entre las medias, la precisión disminuye.

Algoritmo de Segmentación por Histogramas
con ventana óptima promedio

Crecimiento de Regiones



Imagen 3.16

Caso 35

Segmentación



Imagen 3.17

Crecimiento de Regiones



Imagen 3.18

Caso 64

Segmentación



Imagen 3.19

Crecimiento de Regiones



Imagen 3.20

Imagen ndi

Segmentación

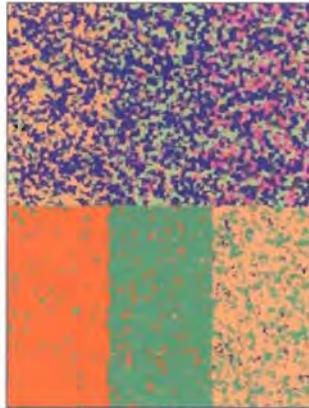


Imagen 3.21

Algoritmo de Segmentación por Histogramas
con ventanas óptimas por clase

Caso 35

Crecimiento de Regiones



Imagen 3.22

Segmentación

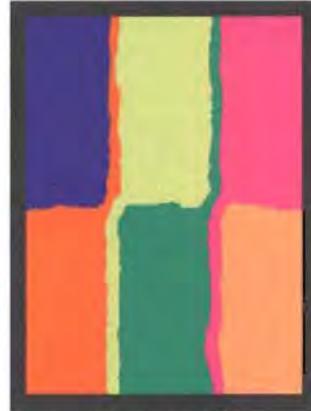


Imagen 3.23

Caso 64

Crecimiento de Regiones



Imagen 3.24

Segmentación



Imagen 3.25

Imagen ndi

Crecimiento de Regiones



Imagen 3.26

Segmentación



Imagen 3.27

Algoritmo de Segmentación por Histogramas



Con ventana óptima promedio
(Umbral de acotación = 0.75; Umbral de Estabilidad = 3%)

D/(Mi-Mj)	32	16	8	4	2	1	Promedio
1	99.991700	99.991684	100.000000	99.989682	99.937786	94.304054	99.035818
2	99.972967	99.981251	99.968867	99.908454	99.017248	93.550320	98.733185
4	99.925349	99.927432	99.970330	98.674399	95.301593	91.528355	97.554576
8	99.867119	99.361400	99.039880	84.491425	53.240411	49.339570	80.889967
16	98.434482	95.182998	94.556365	82.813425	66.306716	17.767845	75.843639
32	56.599041	35.570881	78.166493	67.024574	87.470684	54.500694	63.222061
Promedio	92.465110	88.335941	95.283656	88.816993	83.545740	66.831806	85.879874

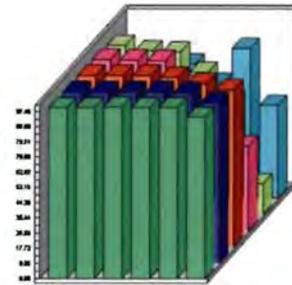


Figura 3.6

Con ventana óptima por clase
(Umbral de acotación = 0.75; Umbral de Estabilidad = 3%)

D/(Mi-Mj)	32	16	8	4	2	1	Promedio
1	99.977533	99.987917	99.995731	99.989239	99.903549	98.116266	99.661706
2	99.986758	99.979166	99.984721	99.929506	97.758728	91.288944	98.154637
4	100.000000	99.991867	99.938796	98.025143	91.611350	88.440379	96.334589
8	99.982301	99.894268	97.306566	86.667383	83.347513	69.656510	89.475757
16	99.790939	95.722046	83.354103	82.232798	83.484751	70.391652	85.829382
32	94.733989	75.642571	78.972286	90.618274	85.665370	62.136891	81.294897
Promedio	99.078587	95.202972	93.258700	92.910390	90.295210	80.005107	91.791828

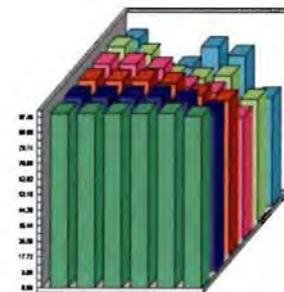


Figura 3.7

- Existe una relación inversa entre el desvío estándar y la precisión: a medida que aumenta el desvío estándar, la precisión disminuye y viceversa, a medida que disminuye el desvío estándar, la precisión aumenta.

Las conclusiones anteriores se resumen a continuación:

Tamaño de ventana	Precisión
Promedio	Menor
Por clase	Mayor

Desvío Estándar	Tamaño de ventana
Aumenta	Aumenta
Disminuye	Disminuye

Separación entre medias	Precisión
Aumenta	Aumenta
Disminuye	Disminuye

Desvío Estándar	Precisión
Aumenta	Disminuye
Disminuye	Aumenta

El efecto de borde introducido por el tamaño de la ventana:

En los resultados obtenidos por cualesquiera de estos algoritmos puede apreciarse un efecto de borde introducido por el tamaño de la ventana. La razón de esto es la siguiente: dados los histogramas de dos clases A y B entre las cuales se encuentra el histograma de una tercera clase C, durante la clasificación, a medida que la ventana comienza a contener pixels de las clases A y B, las estadísticas obtenidas a partir de dicha ventana tienden a parecerse a las de la clase C, y no a una de las clases A o B, como corresponde.

A medida que la ventana se aproxima a la frontera entre las clases A y B, la dispersión de sus valores aumenta. Mientras los histogramas de las clases A, B y C no se traslapan el histograma de la ventana tiende a ser bimodal, pero a medida que tales histogramas se traslapan cada vez más, el histograma de la ventana pierde esta propiedad, y se hace unimodal.

En general, el efecto de borde se acentúa para tamaños de ventana grandes.

III.1.4.2 Los resultados de la simulación

En todas las simulaciones realizadas se efectuaron cien iteraciones. Para cada algoritmo se ha mantenido constante el umbral de estabilidad a efectos de comparar resultados entre bandas. Sin embargo, no se hacen comparaciones entre los resultados de

los algoritmos de segmentación por medias espectrales con los algoritmos de segmentación por histogramas conjuntos, puesto que los umbrales de estabilidad empleados son diferentes en cada caso.

III.1.4.2.1 Algoritmo de segmentación por medias espectrales

Se presentan las siguientes tablas comparativas:

Tabla comparativa de valores de precisión sobre pixel clasificado segmentando imágenes multispectrales por medias espectrales con ventanas óptimas promedio y ventanas óptimas por clase (Umbral de estabilidad = 1 %, Umbral de acotación = 1)

Tipo de imágenes	Precisión al calcular Ventana Optima Promedio	Precisión al calcular Ventana Optima por Clase
Imágenes de una banda	77.46	85.15
Imágenes de dos bandas no decorrelacionadas	90.74	94.60
Imágenes de dos bandas decorrelacionadas	89.30	95.14
Imágenes de tres bandas no decorrelacionadas	92.49	97.89
Imágenes de tres bandas decorrelacionadas	94.88	98.67

Tabla 3.7

Tabla comparativa de valores del coeficiente Kappa segmentando imágenes multispectrales por medias espectrales con ventanas óptimas promedio y ventanas óptimas por clase (Umbral de estabilidad = 1 %, Umbral de acotación = 1)

Tipo de imágenes	Coefficiente Kappa (en%) al calcular Ventana Optima Promedio	Coefficiente Kappa (en%) al calcular Ventana Optima Por Clase
Imágenes de una banda	72.66	82.11
Imágenes de dos bandas no decorrelacionadas	88.79	93.50
Imágenes de dos bandas decorrelacionadas	86.78	94.18
Imágenes de tres bandas no decorrelacionadas	90.95	97.47
Imágenes de tres bandas decorrelacionadas	93.71	98.41

Tabla 3.8

III.1.4.2.2 Algoritmo de segmentación por histogramas conjuntos

Se presentan las siguientes tablas comparativas:

Tabla comparativa de valores de precisión sobre pixel clasificado segmentando imágenes multiespectrales por histogramas conjuntos con ventanas óptimas promedio y ventanas óptimas por clase (Umbral de estabilidad = 10 %, Umbral de acotación = 0.75)

Tipo de imagen	Precisión al calcular Ventana Optima Promedio	Precisión al calcular Ventana Optima Por Clase
Imágenes de una banda	71.52	96.22
Imágenes de dos bandas no decorrelacionadas	80.56	97.35
Imágenes de dos bandas decorrelacionadas	80.65	96.76

Tabla 3.9

Tabla comparativa de valores del coeficiente Kappa segmentando imágenes multiespectrales por histogramas conjuntos con ventanas óptimas promedio y ventanas óptimas por clase (Umbral de estabilidad = 10 %, Umbral de acotación = 0.75)

Tipo de imagen	Coefficiente Kappa (en%) al calcular Ventana Optima Promedio	Coefficiente Kappa (en%) al calcular Ventana Optima Por Clase
Imágenes de una banda	70.08	95.43
Imágenes de dos bandas no decorrelacionadas	82.46	55.99
Imágenes de dos bandas decorrelacionadas	82.83	95.87

Tabla 3.10

Notas:

Debido a que el algoritmo de segmentación por histogramas conjuntos tiene un requerimiento de memoria muy elevado (la cantidad de bytes que se necesitan son $256^{\text{número de bandas}} \times \text{cantidad de clases} \times 4 \times 2$), no se han efectuado simulaciones para tres bandas.

Con un umbral más pequeño (3%) el algoritmo de segmentación por histogramas conjuntos con ventana óptima por clase aplicado sobre una sola banda produce mejores resultados que los presentados en la tabla anterior (la precisión es de 97.22% y kappa es igual a 96.42). Este mismo umbral de 3% sobre imágenes de dos bandas es demasiado pequeño, puesto que algunas clases quedan sin ventana óptima, ya que en ningún momento la función objetivo alcanza un valor tan pequeño. Sin embargo, si se emplea un umbral de estabilidad muy grande en segmentaciones sobre imágenes de una sola banda, los tamaños de ventana resultantes son demasiado pequeños y, en consecuencia, los crecimientos de algunas clases tienen área cero, y la segmentación no puede continuar, puesto que de tales clases no pueden obtenerse estadísticas.

III.1.4.2.3 Conclusiones de la simulación

- En ambos algoritmos se han obtenido buenos resultados al calcular ventanas óptimas por clase.
- La precisión aumenta cuando se aumenta el número de bandas.
- La precisión en la segmentación de imágenes multiespectrales por medias espectrales mejora cuando las bandas están decorrelacionadas.

III.2 LA IMAGEN SPOT

En primer lugar, se tomó la primera componente principal (Imagen 3.28) de la imagen SLP y se aplicó la segmentación por medias espectrales y por histogramas conjuntos. Ambos algoritmos calcularon ventanas óptimas por clase con un umbral de

Imagen SPOT
SLP

Primera Componente Principal

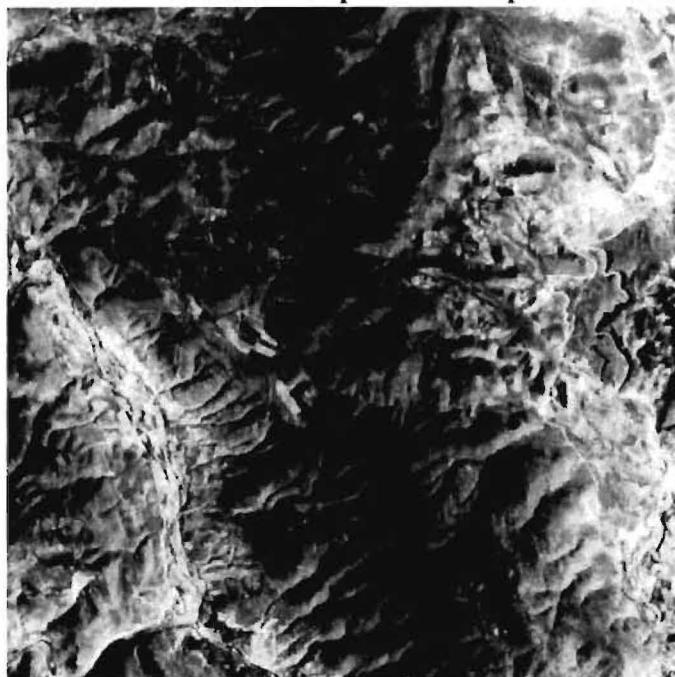


Imagen 3.28

Segunda Componente Principal

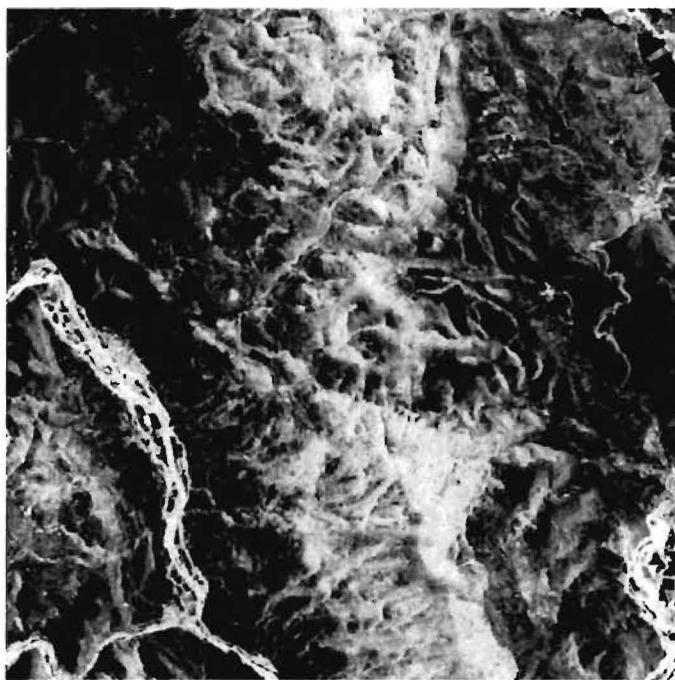


Imagen 3.29

estabilidad de 3%. En segundo lugar, se tomaron las dos primeras componentes principales de dicha imagen (la segunda componente es la imagen 3.29) y se aplicaron también tales segmentaciones calculando ventanas óptimas por clase. Los resultados se presentan a continuación.

III.2.1 Una banda con segmentación por medias espectrales con ventana óptima por clase

Las posiciones de los pixels sembrados y las ventanas óptimas de las clases fueron las siguientes:

Clase	Fila	Columna	Ventana Optima
1	267	95	3
2	171	181	3
3	190	158	7
4	163	303	3
5	200	257	3
6	86	444	3

Tabla 3.11

Debido a que el atributo relevante de este algoritmo es la “media” se hará un seguimiento de los valores que van tomando las medias de las clases a lo largo de cada una de las etapas. Los valores de la segunda columna, son las medias de los valores de los pixels pertenecientes a cada ventana óptima. En la tercera columna se tienen las medias de los valores de los pixels pertenecientes a cada región crecida. En la última columna se tienen las medias de los valores de los pixels de cada región segmentada.

Clase	Ventana Optima	Crecimiento	Segmentación
1	254	250	236
2	54	54	52
3	85	85	94
4	155	155	143
5	17	18	17
6	185	182	194

Tabla 3.12

Los umbrales tanto para el crecimiento como para la segmentación se obtienen sumando y restando un desvío estándar a las medias, por lo que los sucesivos valores del atributo desvío estándar se presentan a continuación:

Clase	Vent. Optima	Crecimiento	Segmentación
1	0	7	12
2	9	11	15
3	13	8	23
4	17	11	22
5	4	9	10
6	6	26	18

Tabla 3.13

Si se considera para cada clase que el dominio del atributo media resultante en la segmentación está dado también por la media de la región segmentada más menos un desvío estándar calculado sobre la región segmentada, puede observarse que para todas las clases los valores iniciales de media (obtenidos tanto en la etapa de cálculo de ventana óptima, como en la de crecimiento de regiones) se encuentran dentro de tal dominio.

El atributo área para cada una de las etapas y clases tuvo los siguientes valores:

Clase	Vent. Optima	Crecimiento	Segmentación
1	9	7	37579
2	9	424	34121
3	49	13341	52451
4	9	2829	52318
5	9	204	28465
6	9	35	51102

Tabla 3.14

La imagen 3.30 muestra los crecimientos de regiones y la imagen 3.31 la segmentación por medias espectrales de la componente principal de la imagen SPOT.

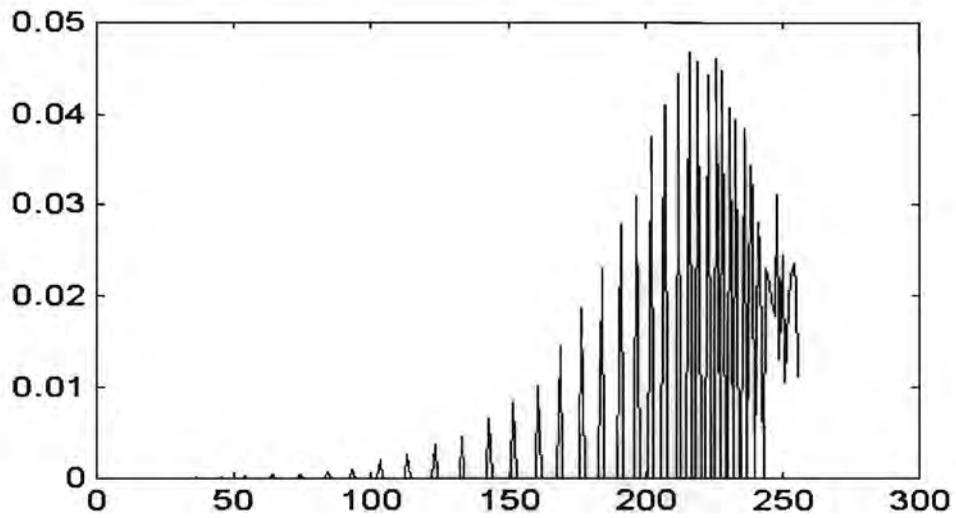
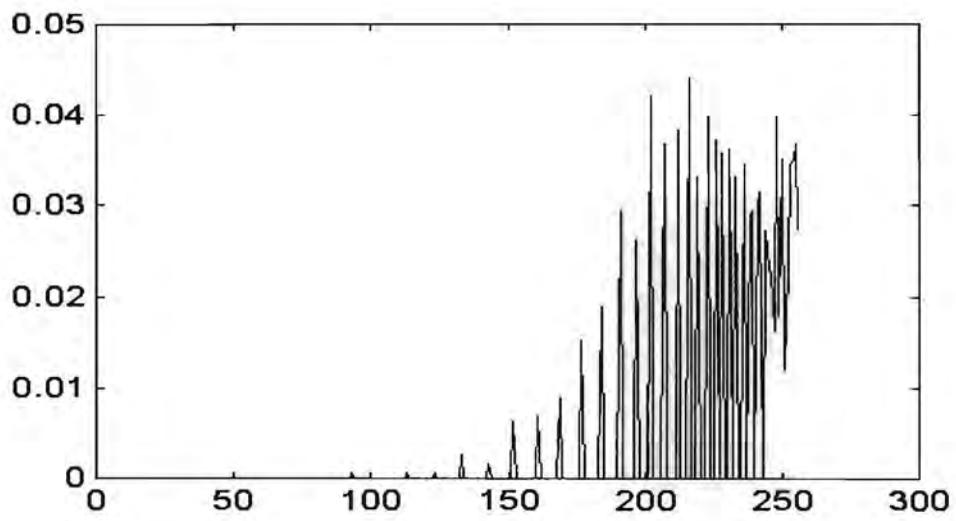
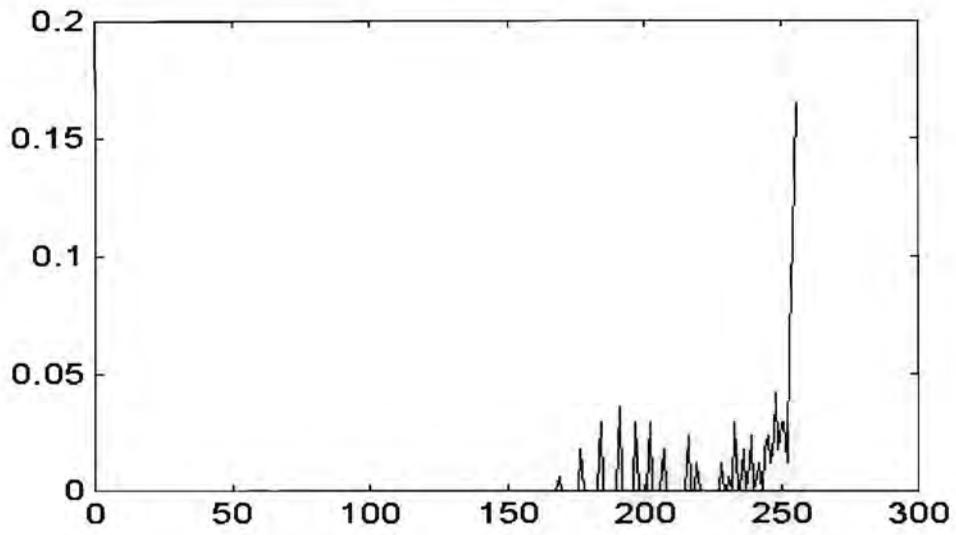
III.2.2 Una banda con segmentación por histogramas conjuntos con ventana óptima por clase

Las posiciones de los pixels sembrados y las ventanas óptimas de las clases fueron las siguientes:

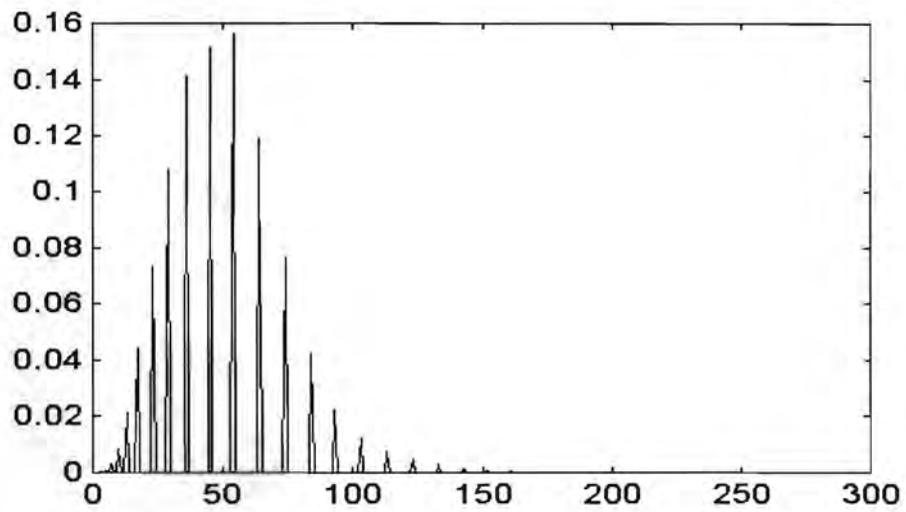
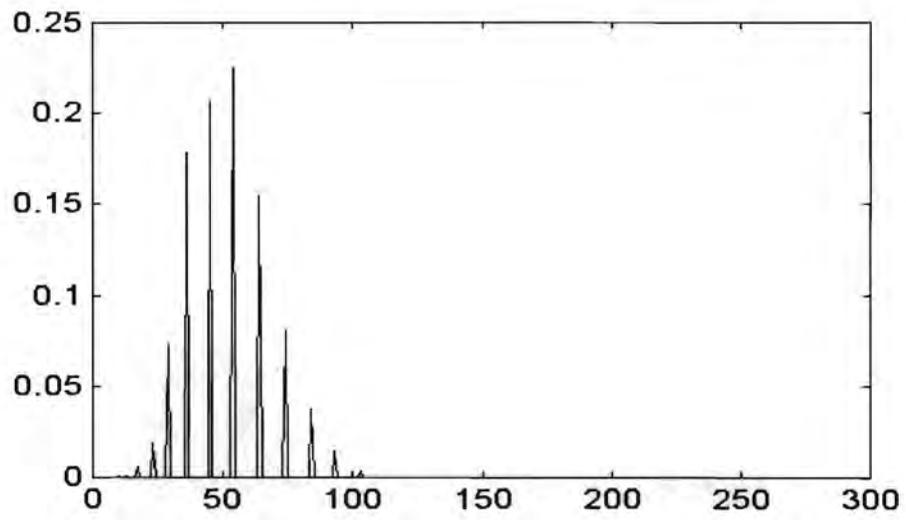
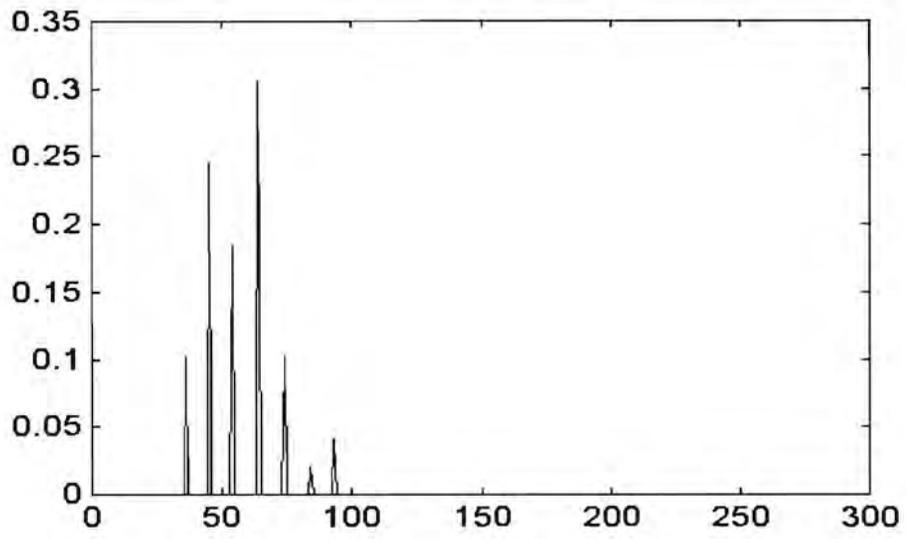
Clase	Fila	Columna	Vent. Optima
1	267	95	13
2	171	181	7
3	190	158	9
4	163	303	13
5	200	257	13
6	86	444	9

Tabla 3.15

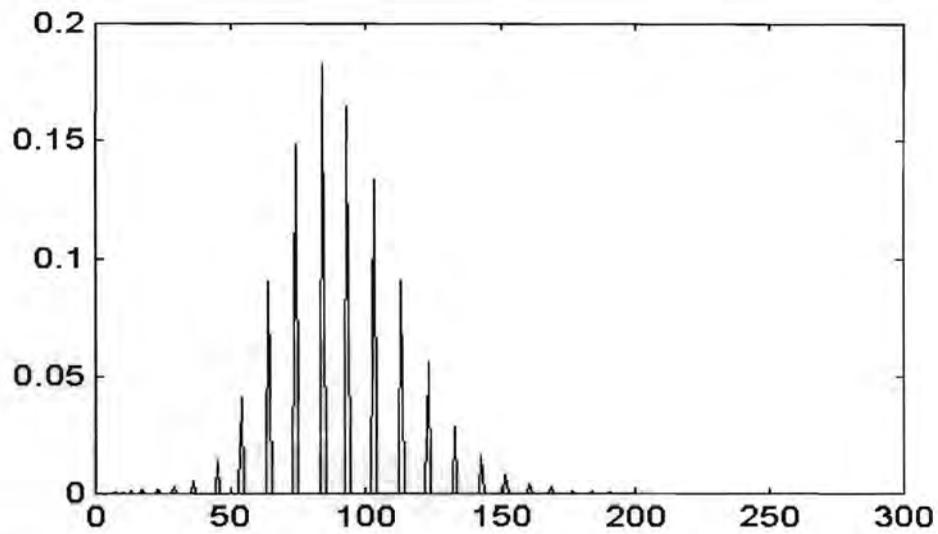
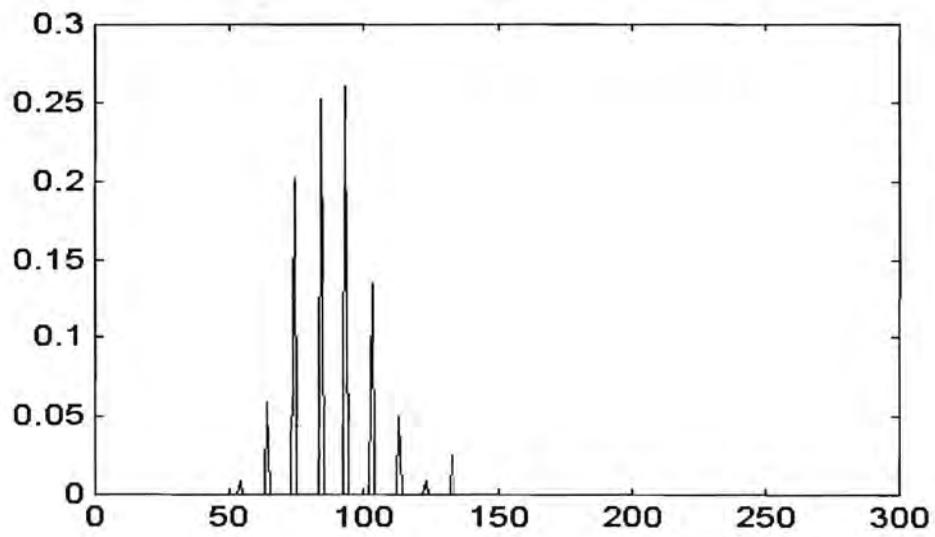
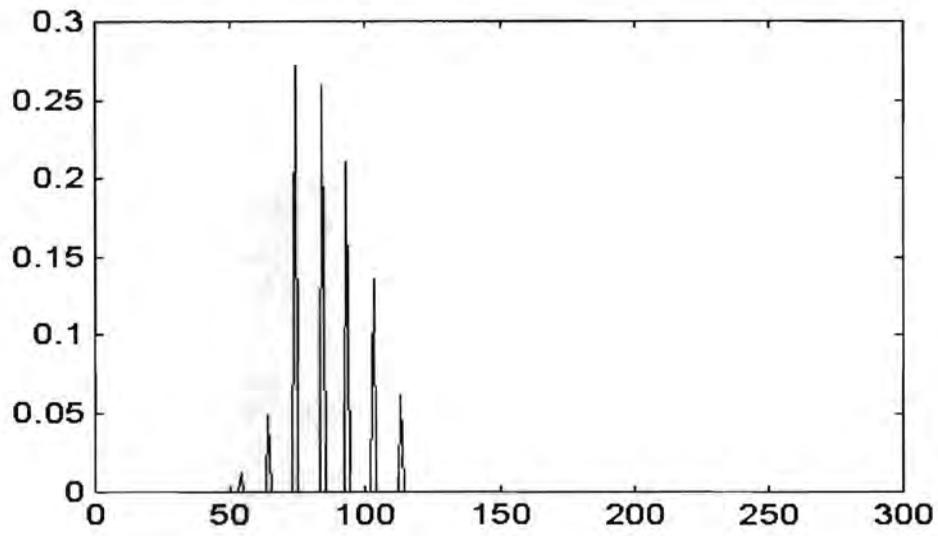
Clase 1



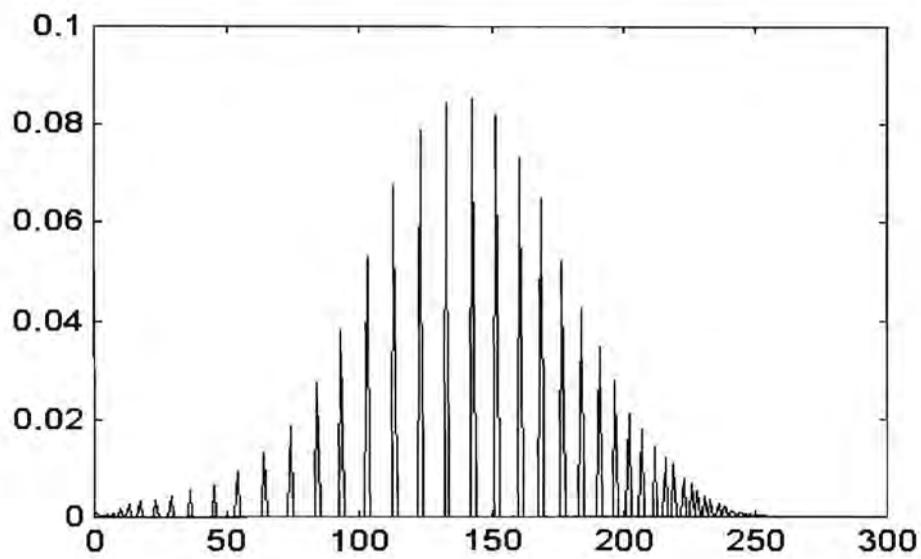
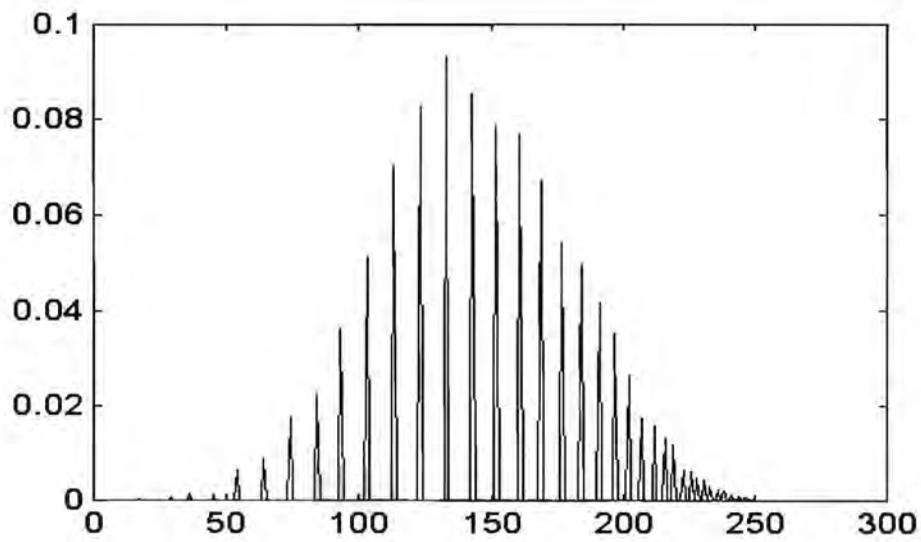
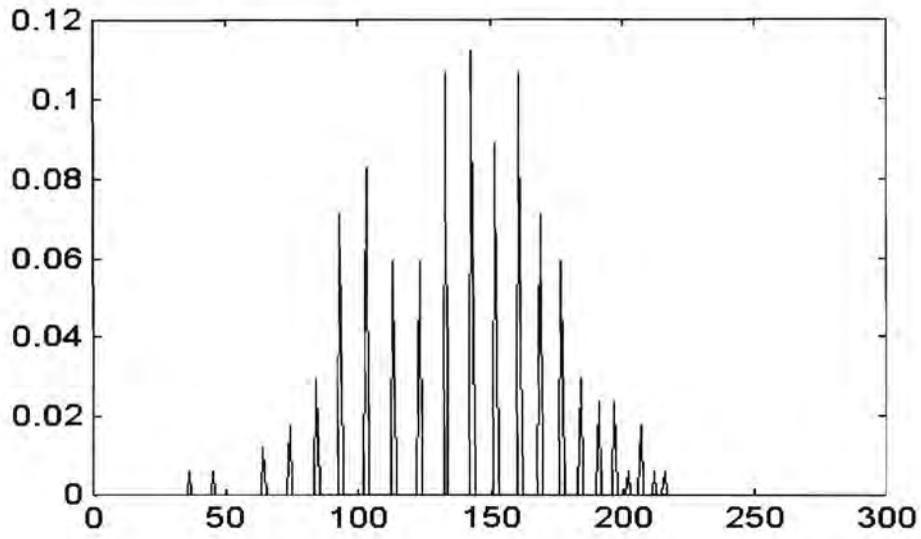
Clase 2



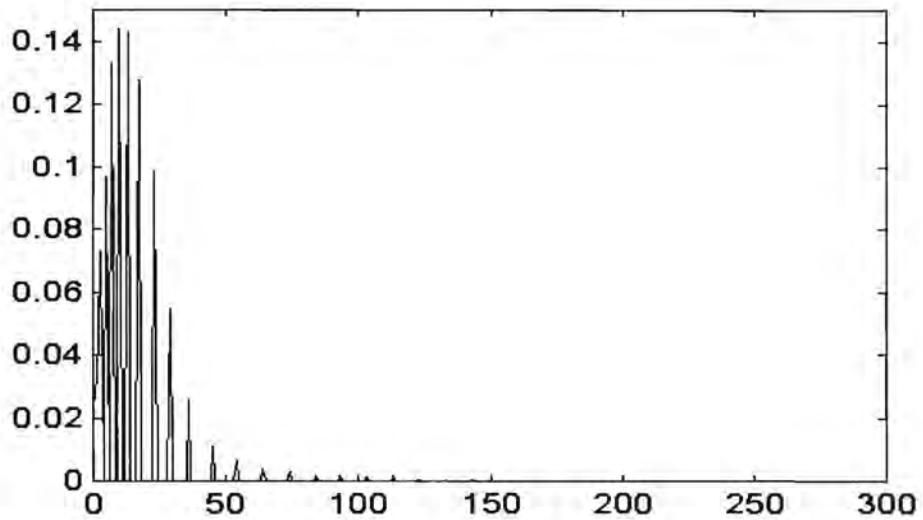
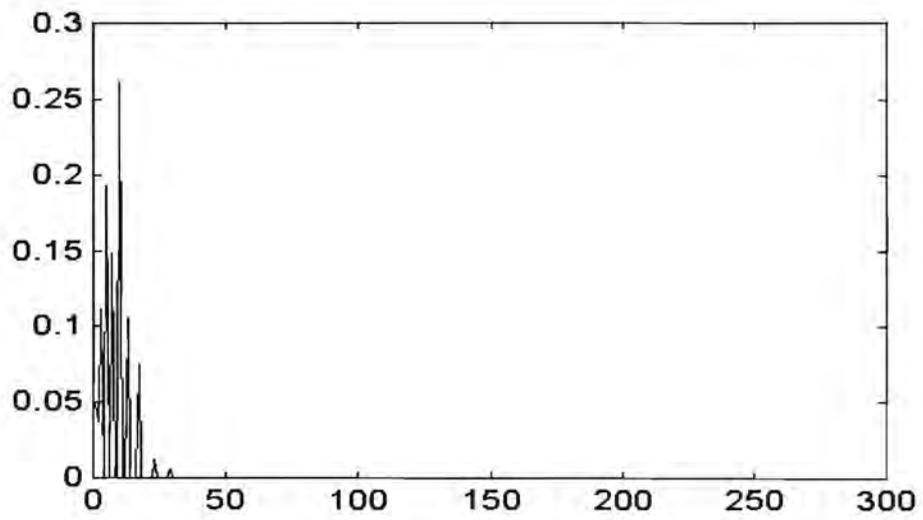
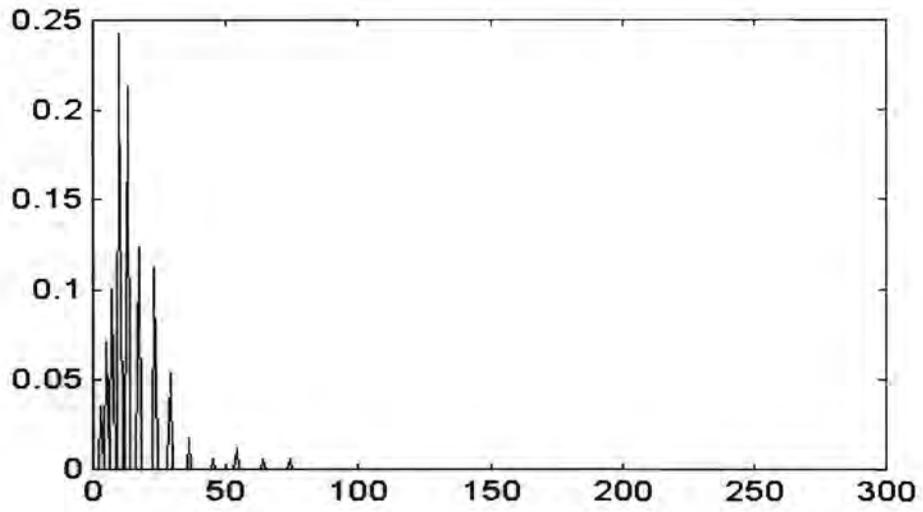
Clase 3



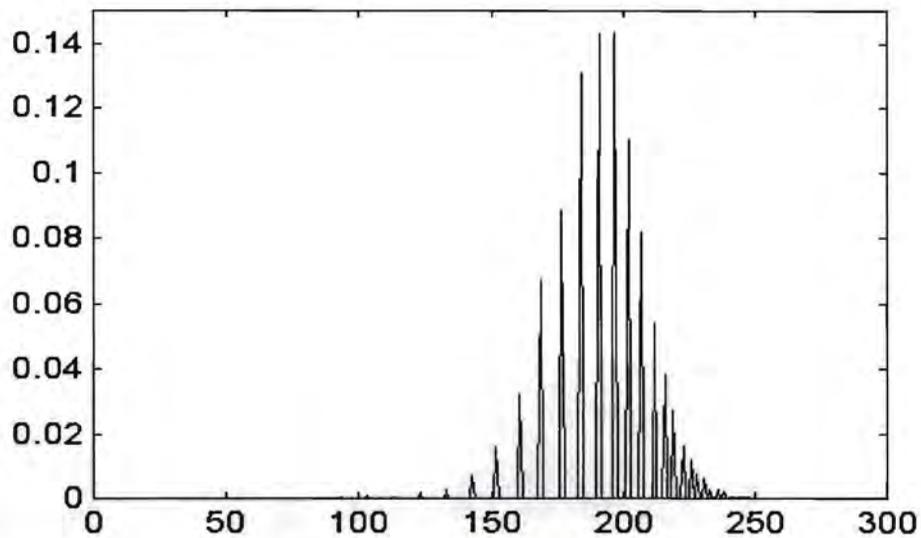
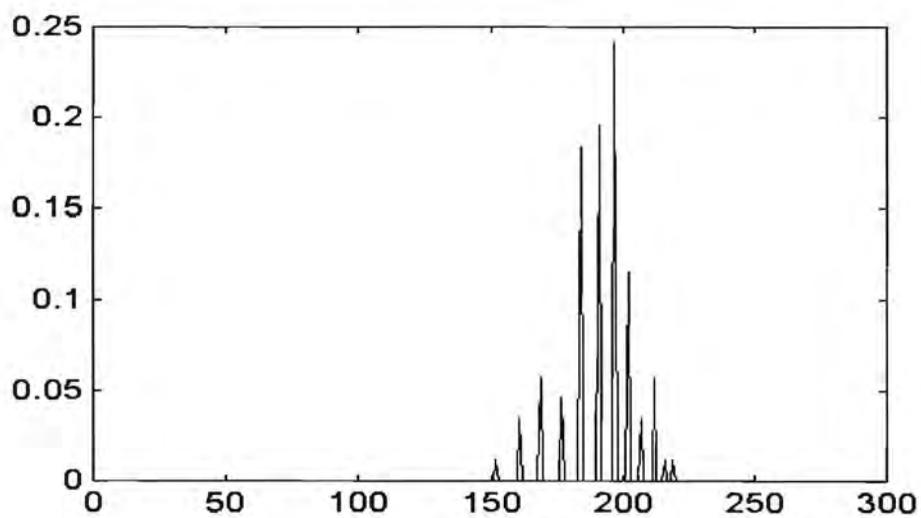
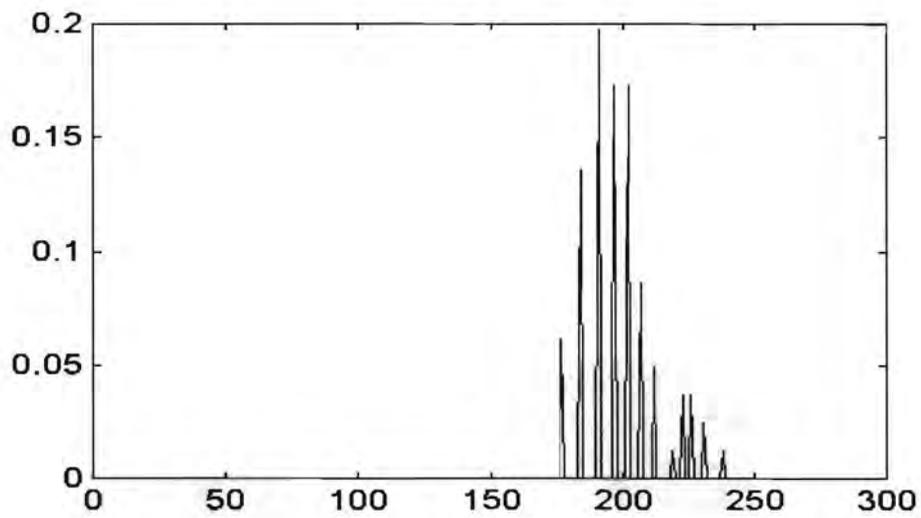
Clase 4



Class 5



Class 6



Primera Componente Principal de la imagen SPOT SLP

Segmentación en 6 clases por Medias Espectrales con ventanas óptimas por clase

Crecimiento de Regiones

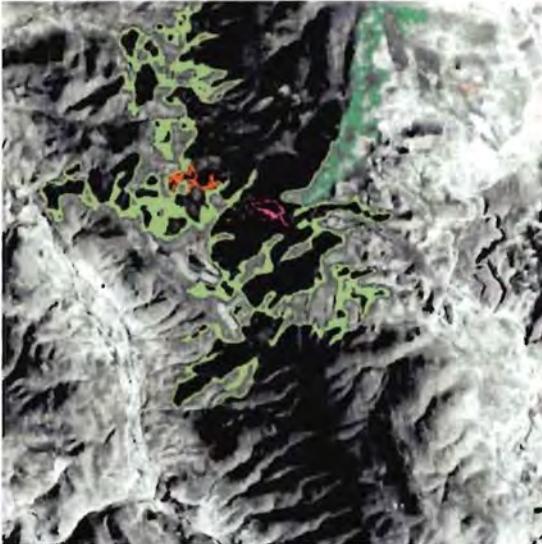


Imagen 3.30

Segmentación

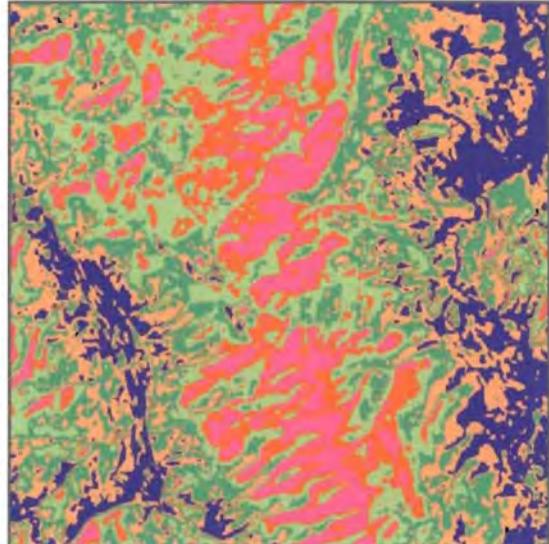


Imagen 3.31

Segmentación en 6 clases por Histogramas con ventanas óptimas por clase

Crecimiento de Regiones

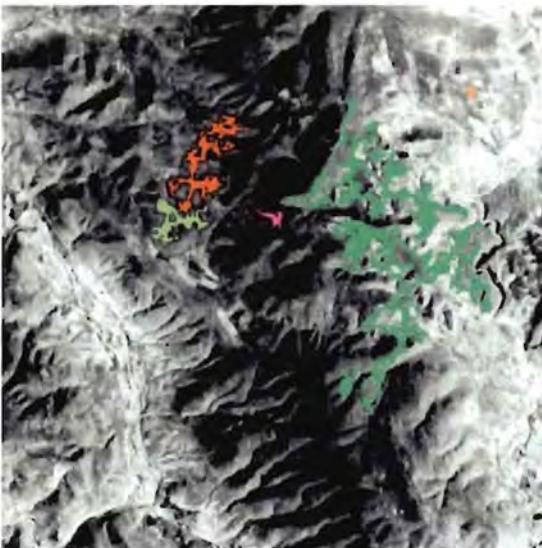


Imagen 3.32

Segmentación

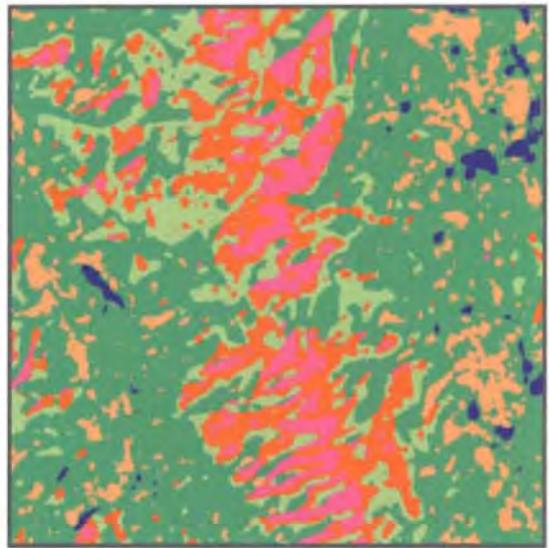


Imagen 3.33

Debido a que el atributo relevante de este algoritmo es el histograma se hará un seguimiento de los histogramas de las clases a lo largo de las etapas de cálculo de ventana óptima, de crecimiento y de segmentación en las figuras 3.8, 3.9, 3.10, 3.11, 3.12 y 3.13 correspondientes a las clases 1 al 6 respectivamente.

A efectos de comparar la correlación entre los histogramas de cada clase en cada etapa se emplea el coeficiente de correlación calculado mediante la siguiente expresión:

$$R_{xy} = \frac{\langle x, y \rangle}{\langle x, x \rangle + \langle y, y \rangle - \langle x, y \rangle} \quad [3.4]$$

Se calculan los siguientes coeficientes de correlación:

- 1) El coeficiente de correlación entre el histograma de la ventana óptima y el histograma del crecimiento denominado Rvc.
- 2) El coeficiente de correlación entre el histograma de la ventana óptima y el histograma de la segmentación denominado Rvs.
- 3) El coeficiente de correlación entre el histograma del crecimiento y el histograma de la segmentación denominado Rcs.

La tabla de coeficientes de correlación entre dichos histogramas es la siguiente:

Clase	Rvc	Rvs	Rcs
1	0.4060	0.2841	0.9453
2	0.8109	0.6531	0.8841
3	0.9593	0.8302	0.8642
4	0.9049	0.9112	0.9916
5	0.7059	0.8508	0.7384
6	0.9049	0.9112	0.9916

Tabla 3.16

Puede apreciarse que para todas las clases, el coeficiente Rcs es mayor que el coeficiente Rvs, es decir, que existe mayor correlación entre el histograma del crecimiento con el histograma de la segmentación, que entre el histograma de la ventana óptima y el histograma de la segmentación.

En las imágenes 3.32 y 3.33 pueden apreciarse el crecimiento de regiones y la segmentación respectivamente.

III.2.3 Dos bandas con segmentación por medias espectrales con ventana óptima por clase

Las posiciones de los pixels sembrados y las ventanas óptimas de las clases fueron las siguientes:

Clase	Fila	Columna	Ventana Optima
1	267	95	3
2	171	181	3
3	190	158	5
4	163	303	5
5	200	257	3
6	86	444	7

Tabla 3.17

El atributo relevante de este algoritmo es la Media espectral. Por ello se hará un seguimiento de los valores que van tomando las medias de las clases banda por banda a lo largo de las etapas de cálculo de ventana óptima, de crecimiento y de segmentación.

En la banda 1 las medias fueron las siguientes:

Clase	Vent. Optima	Crecimiento	Segmentación
1	254	244	235
2	54	62	78
3	87	86	109
4	159	155	138
5	17	24	39
6	197	196	197

Tabla 3.18

En la banda 2 se obtuvieron estas medias:

Clase	Vent. Optima	Crecimiento	Segmentación
1	217	190	201
2	66	69	41
3	120	114	110
4	204	202	210
5	193	190	203
6	147	143	106

Tabla 3.19

Como ya se ha mencionado, los umbrales tanto para el crecimiento como para la segmentación se obtienen sumando y restando un desvío estándar a la media, por lo que los sucesivos valores del atributo desvío se presentan a continuación también para cada banda.

Los valores de desvío estándar en la banda 1 fueron los siguientes:

Dos Primeras Componentes Principales de la imagen SPOT SLP

Segmentación en 6 clases por Medias Espectrales con ventanas óptimas por clase

Crecimiento de Regiones



Imagen 3.34

Segmentación

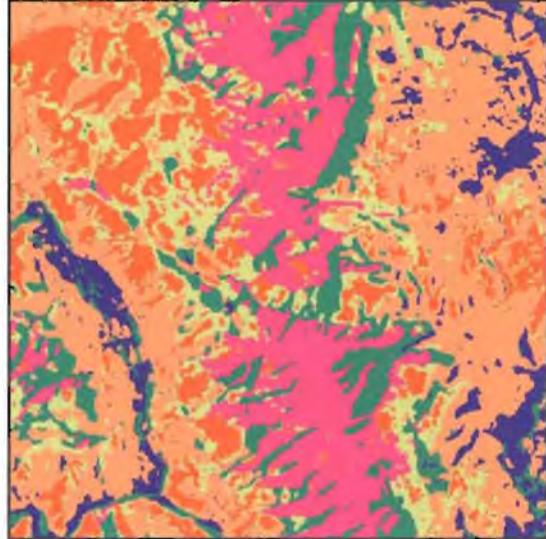


Imagen 3.35

Segmentación en 7 clases por Histogramas Conjuntos con ventanas óptimas por clase

Crecimiento de Regiones

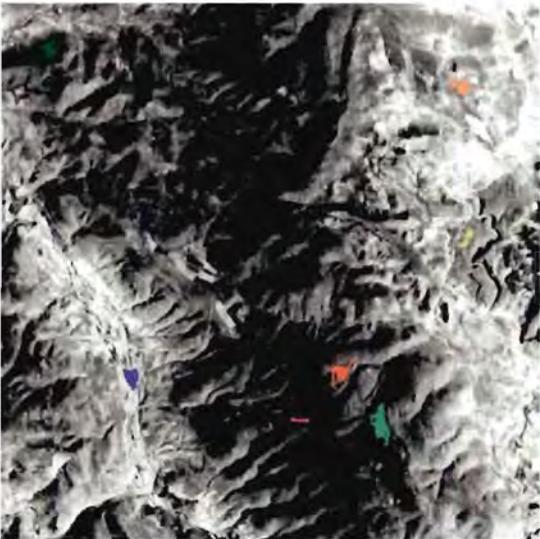


Imagen 3.36

Segmentación

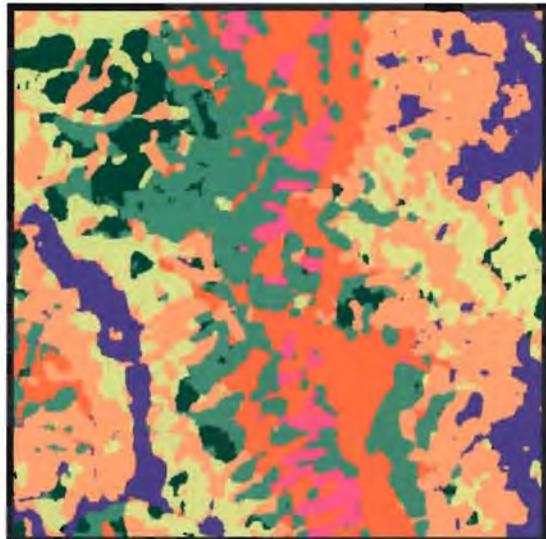


Imagen 3.37

Clase	Vent. Optima	Crecimiento	Segmentación
1	0	22	16
2	9	45	41
3	15	9	40
4	27	14	36
5	4	33	26
6	11	14	32

Tabla 3.20

En la banda 2 se presentaron los siguientes valores de desvío estándar:

Clase	Vent. Optima	Crecimiento	Segmentación
1	15	77	35
2	9	39	30
3	48	25	42
4	12	12	32
5	10	8	34
6	16	7	49

Tabla 3.21

Puede observarse que para todas las clases los valores iniciales de media (obtenidos tanto en la etapa de cálculo de ventana óptima, como en la de crecimiento de regiones) se encuentran dentro del dominio del atributo media resultante en la segmentación.

El crecimiento de regiones y la segmentación se presentan en las imágenes 3.34 y 3.35 respectivamente.

El atributo Area para cada una de las etapas y clases tuvo los siguientes valores:

Clase	Vent. Optima	Crecimiento	Segmentación
1	9	7	18028
2	9	19	39608
3	25	1683	40258
4	25	205	32298
5	9	15	47479
6	49	239	78365

Tabla 3.22

III.2.4 Dos bandas con segmentación por histogramas conjuntos con ventana óptima por clase

Los pixels fueron sembrados en las siguientes posiciones:

Clase	Fila	Columna
1	356	120
2	355	326
3	229	438
4	404	363
5	396	275
6	77	436
7	44	44

Tabla 3.23

El crecimiento de regiones y la respectiva segmentación se muestran en las imágenes 3.36 y 3.37.

III.3 LA IMAGEN SEASAT

A la imagen SEASAT (3.38) se le aplicó, en primer lugar, la segmentación por matrices de co-ocurrencia. En segundo lugar, se efectuó una segmentación por centros de masa de matrices de co-ocurrencia.

III.3.1 Segmentación por matrices de co-ocurrencia

Se han ingresado los pixels prototipo de las clases en las posiciones que a continuación se indican. El umbral de estabilidad fue de 3% y el factor a aplicar al umbral de 0.75. Los valores de ventana óptima resultantes para cada clase se incluyen en la siguiente tabla:

Clase	Fila	Columna	Vent. Optima
1	229	209	29
2	80	263	17
3	261	448	21
4	398	344	9
5	105	460	21
6	192	93	21
7	445	253	43

Tabla 3.24

El crecimiento de regiones y la imagen segmentada se muestran a continuación en las imágenes 3.39 y 3.40.

El atributo área a lo largo de las sucesivas etapas ha ido tomando para cada clase los siguientes valores:

Imagen SEASAT

Imagen SEASAT
Speckle eliminado con el Filtro Geométrico
[Lira, Frulla, Pastrana, Chavira, 1997]



Imagen 3.38

Clase	Vent. Óptima	Crecimiento	Segmentación
1	841	245	12512
2	289	4494	23282
3	441	339	11765
4	81	49	9371
5	441	326	44824
6	441	923	48144
7	1849	8926	59627

Tabla 3.25

A efectos de medir la correlación entre matrices de co-ocurrencia correspondientes a una misma clase a lo largo de las etapas de cálculo de ventana óptima, crecimiento de regiones y segmentación, se calcularon para cada clase los siguientes coeficientes:

- Coeficiente de correlación r_{vc} entre la matriz de co-ocurrencia de la ventana óptima y la matriz de co-ocurrencia de la región crecida.
- Coeficiente de correlación r_{vs} entre la matriz de co-ocurrencia de la ventana óptima y la matriz de co-ocurrencia de la región segmentada.
- Coeficiente de correlación r_{cs} entre la matriz de co-ocurrencia de la región crecida y la matriz de co-ocurrencia de la región segmentada.

Con este fin, las matrices de co-ocurrencia fueron reescaladas a 256 y almacenadas como imágenes raster. Para cada clase, las imágenes de las matrices de co-ocurrencia resultantes en cada una de las tres etapas constituyeron una imagen multispectral que fue entrada al proceso de descomposición en componentes principales. Este proceso genera una matriz de covarianza, de cuyos elementos se tomaron los datos para obtener los resultados que se presentan a continuación:

$$r_{ij} = \frac{\text{cov}(i, j)}{\sqrt{\text{cov}(i, i) \text{cov}(j, j)}} \quad [3.5]$$

Clase	r_{vc}	r_{vs}	r_{cs}
1	0.9527	0.9380	0.9850
2	0.9957	0.9928	0.9987
3	0.9905	0.9930	0.9857
4	0.9388	0.9248	0.9707
5	0.9795	0.9813	0.9964
6	0.9773	0.9609	0.9857
7	0.9915	0.9840	0.9916

Tabla 3.26

III.3.2 Segmentación por centros de masa de matrices de co-ocurrencia

Los pixels fueron sembrados en las mismas posiciones que en el ejemplo anterior. El algoritmo de segmentación por centros de masa de matrices de co-ocurrencia efectúa el cálculo de la ventana óptima de la misma manera que la segmentación por matrices de co-ocurrencia, por lo que los tamaños de ventana óptimos resultantes para cada clase también

coinciden con los del ejemplo anterior.

El crecimiento de regiones y la segmentación se realiza considerando el atributo centro de masa de la matriz de co-ocurrencia, por lo que, se hará un seguimiento de este atributo a lo largo de las tres etapas. Los umbrales para el crecimiento de regiones y para la segmentación se definen por medio del desvío estándar de la matriz de co-ocurrencia, por lo que también se mostrarán los valores que va tomando para cada eje de tal matriz. El factor a aplicar al umbral fue de 1, es decir, los umbrales se calcularon sumando más y menos un desvío estándar al centro de masa. El umbral de estabilidad fue de 3%.

Ya se ha señalado que la matriz de co-ocurrencia es simétrica respecto de su diagonal principal. En consecuencia, los centros de masa deberían encontrarse siempre sobre tal diagonal. Sin embargo, este algoritmo emplea una función de densidad que se obtiene reflejando la parte inferior izquierda de la matriz de co-ocurrencia sobre su diagonal principal y sumándola sobre la parte superior derecha. El dominio de esta nueva función de densidad se constituye por la parte superior derecha y la diagonal principal de la matriz de co-ocurrencia. De esta manera, los centros de masa pueden encontrarse en cualquier parte de este dominio, teniendo la peculiaridad de que nunca la fila del centro de masa puede ser superior que la columna del centro de masa.

A continuación se presentan los valores a lo largo de las etapas de cálculo de ventana óptima, crecimiento de regiones y segmentación de los atributos: fila del centro de masa de la matriz de co-ocurrencia, columna del centro de masa de la matriz de co-ocurrencia, desvío sobre la fila de la matriz de co-ocurrencia y desvío sobre la columna de la matriz de co-ocurrencia.

Fila del centro de masa de la matriz de co-ocurrencia:

Clase	Vent. Óptima	Crecimiento	Segmentación
1	4	3	3
2	3	3	3
3	5	7	7
4	3	3	3
5	3	4	4
6	4	4	4
7	6	6	7

Tabla 3.27

Columna del centro de masa de la matriz de co-ocurrencia:

Clase	Vent. Óptima	Crecimiento	Segmentación
1	4	4	4
2	3	3	3
3	6	13	9
4	4	4	4
5	4	4	4
6	5	5	5
7	8	8	9

Tabla 3.28

Desvío estándar de fila de la matriz de co-ocurrencia:

Clase	Vent. Optima	Crecimiento	Segmentación
1	1	2	2
2	1	1	1
3	3	5	4
4	2	2	2
5	2	1	1
6	2	2	3
7	4	5	4

Tabla 3.29

Desvío estándar de columna de la matriz de co-ocurrencia:

Clase	Vent. Optima	Crecimiento	Segmentación
1	3	3	2
2	2	2	2
3	4	7	5
4	1	1	1
5	1	2	3
6	2	2	3
7	7	7	6

Tabla 3.30

El crecimiento de regiones y de la segmentación se presentan a continuación en las imágenes 3.41 y 3.42.

Los valores del atributo área fueron los siguientes:

Clase	Vent. Optima	Crecimiento	Segmentación
1	841	246	36424
2	289	10871	15496
3	441	16	171
4	81	456	28798
5	441	380	27123
6	441	7561	66806
7	1849	7271	46082

Tabla 3.31

Los coeficientes de correlación para cada clase y etapa fueron los siguientes:

Imagen SEASAT

Segmentación en 7 clases por Matrices de Co-ocurrencia

Crecimiento de Regiones

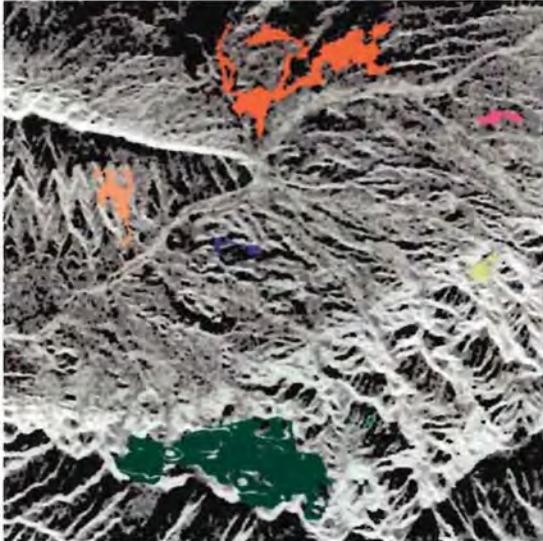


Imagen 3.39

Segmentación

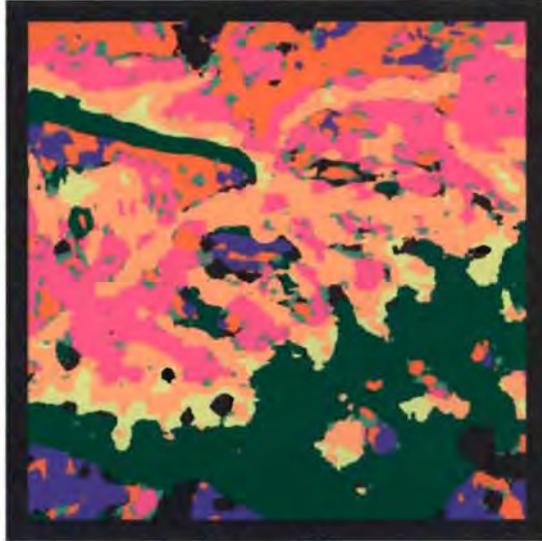


Imagen 3.40

Segmentación en 7 clases por Centros de Masa de Matrices de Co-ocurrencia

Crecimiento de Regiones



Imagen 3.41

Segmentación

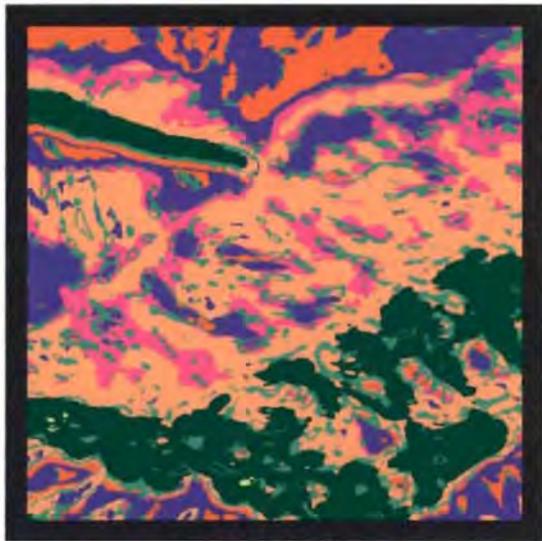


Imagen 3.42

Clase	r_{vc}	r_{vs}	r_{cs}
1	0.9527	0.9525	0.8386
2	0.9945	0.9756	0.9913
3	0.3354	0.8367	0.4361
4	0.9806	0.9494	0.9893
5	0.9829	0.8892	0.9554
6	0.9805	0.9572	0.9893
7	0.9845	0.8355	0.8975

Tabla 3.32

En este ejemplo pueden apreciarse claramente los siguientes aspectos:

- Cuando la región crecida de una clase tiene pocos pixels, como en la clase 3, la correlación entre las matrices de co-ocurrencia obtenidas a lo largo de las sucesivas etapas es muy baja. La definición de un umbral de suficiencia estadística en el crecimiento de regiones (o a fin de conservar el mismo nombre, un umbral de estabilidad para el crecimiento), es un área a explorar.
- La primera impresión visual de la imagen segmentada es que estos resultados son mejores que los del algoritmo anterior. En esta imagen aparece un efecto de borde que pareciera ser "agradable" a la vista, sin embargo, los resultados demuestran que este algoritmo no funciona tan bien como el anterior. La eliminación de este borde es un campo interesante de estudio. Ya se ha mencionado este efecto en la segmentación espectral.

III.4 Conclusiones y perspectivas

III.4.1 Conclusiones

- Los algoritmos implantados reportan buenos resultados al calcular ventanas óptimas por clase.
- Existe una relación directa entre el desvío estándar y el tamaño de ventana en algoritmos con ventana óptima por clase.
- La precisión aumenta cuando se aumenta el número de bandas.
- La precisión en la segmentación de imágenes multiespectrales por medias espectrales mejora cuando las bandas están decorrelacionadas.
- En los algoritmos implantados existe una relación directa entre la separación entre las medias y la precisión y una relación inversa entre el desvío estándar y la precisión.
- El algoritmo de segmentación por medias espectrales reporta buenos resultados en imágenes sintéticas y reales.
- El algoritmo de segmentación por histogramas conjuntos reporta buenos resultados en imágenes sintéticas y regulares en imágenes reales.

- El algoritmo de segmentación por matrices de co-ocurrencia da buenos resultados en imágenes texturales reales.

- El algoritmo de segmentación por centros de masa de matrices de co-ocurrencia es un camino de estudio interesante a seguir.

III.4.2 Caminos a seguir

- En cuanto al pixel prototipo:

Validar el pixel prototipo estableciendo si en su contexto existe energía suficiente como para obtener estadísticas válidas de la clase que representa.

- En cuanto al umbral de estabilidad:

Calcular este umbral a partir de los datos de la imagen. Que se pueda establecer cuál es la zona de la función objetivo de cada clase que contiene estadísticas válidas.

- En cuanto al factor de acotación a aplicar al umbral:

Eliminarlo cuando el umbral se obtiene a partir de los datos de la clase, aprovechando la propiedad de la Gaussiana de que considerando la media más menos un desvío estándar se tienen el 68% de los datos, considerando la media más menos dos desvíos se tienen el 95% de los datos y con tres desvíos se tienen el 100% de los datos.

Cuando el umbral se obtiene a partir de los datos de todas las clases, sería interesante averiguar si tal vez la apelación a un umbral de acotación es consecuencia de que con la ventana óptima se está teniendo la estadística relacionada a un pixel y su contexto y no a $v \times v$ pixels y sus respectivos contextos, es decir, se cuenta con un solo valor de la función criterio de homogeneidad insuficiente para determinar un rango de variación o umbrales. Es decir, un rango de variación se puede establecer cuando se cuenta con un conjunto de valores y con uno solo, puesto que se puede calcular, por ejemplo, la media del atributo en cuestión y su desvío estándar, y con ello definir el dominio del atributo criterio de homogeneidad.

Por otro lado, incluso puede que el atributo considerado criterio de homogeneidad no sea en realidad un atributo relevante para una imagen dada y, por ende, el factor de acotación no hace más que provocar una situación forzada incongruente con la realidad.

- En cuanto al atributo relevante:

Con base en los datos propios de la imagen y un conjunto de atributos dados, establecer automáticamente cuáles de tales atributos tienen los dominios correspondientes a las clases menos traslapados. En otros términos, automatizar la especificación del atributo relevante.

- En cuanto a los umbrales para el crecimiento de regiones:

Considerar la inclusión del cálculo de la ventana fin de la clase, de modo que con la ventana óptima se tenga una muestra con tantos pixels con sus respectivos contextos como elementos la componen. Esto quiere decir, que alrededor de cada pixel perteneciente a la ventana óptima debe poder considerarse un contexto o ventana del mismo tamaño que la ventana óptima cuyos elementos "sigan" perteneciendo a la clase del pixel prototipo. Esto lo puede garantizar el conocimiento de la ventana fin de la clase. Teniendo una muestra con $v \times v$ pixels cada uno con sus respectivos contextos es posible calcular un rango de variación de la función criterio de homogeneidad, por ejemplo, a través de su media y su desvío estándar. Se pueden tener, en consecuencia, umbrales calculados a partir de datos propios de la clase sin necesidad de tomar datos de otras clases.

- En cuanto al crecimiento de regiones:

Considerar un umbral de crecimiento suficiente (o umbral de estabilidad para el crecimiento de regiones), para detener el crecimiento cuando se han reunido pixels suficientes o para establecer si la región crecida no constituye una muestra estadística válida porque no cuenta con pixels suficientes. Contando con este umbral, en algunas circunstancias, puede que incluso se establezca que el crecimiento de regiones no es necesario.

- En cuanto a los umbrales para la segmentación:

En general, para cada etapa, emplear la misma metodología para el cálculo de los umbrales.

- En cuanto a la función potencial:

Las estadísticas resultantes de los crecimientos de regiones pueden servir de entrada a un algoritmo Markoviano como el Recocido Simulado. Un camino interesante a explorar es el empleo de la función potencial que utiliza el algoritmo segmentación por crecimiento de regiones en un algoritmo Markoviano y el estudio de una posible relación entre los clickees y la ventana óptima.

Referencias

- ACTON, S. T., 1996, On unsupervised segmentation of remotely sensed imagery using nonlinear regression, *Int. J., Remote Sensing*, Vol. 17, No 7, 1407 – 1415.
- ARAI, K., 1993, A classification method whit a spatial-spectral variability, *Int. J. Remote Sensing*, Vol. 14, No 4, 699 – 709.
- AYANS, J. & BIGING, G. S., 1996, An iterative classification approach for mapping natural resources from satellite imagery, *Int. J., Remote Sensing*, Vol. 17, No 5, 957 – 981.
- BEAULIEU, J.M. & GOLDBERG, M., 1989, Hierarchy in picture segmentation: a stepwise optimization approach. *I. E. E. E. Transactions on Pattern Analysis and Machine Intelligence*, 11, 150- 163.
- BEZDEK, J. C., 1987, *Pattern Recognition with Fuzzy Objective Function Algorithms*.
- CETIN, H. & LEVANDOWSKY, D. W, *Interactive Classification and Mapping of Multi-Dimensional Remotely Sensed Data using n-Dimensional Probability*
- CONGALTON, R. G., ODERWALD, R. G. & MEAD, R. A., 1983, Assessing Landsat classification accuracy using discrete multivariate analysis techniques. *Photogrammetric Engineering and Remote Sensing*, 49, 1671 – 1678.
- COHEN, J., 1960, A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20, 37 – 46.
- CHAVIRA, D., 1996, Evaluación de la complejidad (heterogeneidad) de una Imagen, Tesis de Licenciatura, Facultad de Ciencias Físico Matemáticas, Colegio Electrónica, BUAP.
- CHUVIECO, E., 1990, *Fundamentos de Teledetección*.
- DIKSHIT, O., 1996, Textural Classification for ecological research using ATM images, *Int. J. Remote Sensing*, Vol. 17, No. 5, 887 – 915.
- FU, K. S. & MUI, K. J., 1981, A Survey on Image Segmentation, *Pattern Recognition*, Vol. 13 pp. 3 - 16.
- GONG, P., MARCEAU, D. J. & HOWARTH P. J., 1992, A comparison of spatial feature extraction algorithms for land-use classification with SPOT HRV data. *Remote Sensing of Environment*, 40, 137 – 151.
- GURNEY, C. M. & TOWNSHEND, J. R. G., 1983, The use of contextual information in the classification of remotely sensed data. *Photogrammetric Engineering and Remote Sensing*, 49, 55 – 64.

HARALICK, R. M., 1973, SHANMUGAM, K.S., & DINSTEIN, I., 1973, Textural features for image classification, I.E.E.E. Transactions on Systems, Man and Cybernetics, 3, 610 – 621.

HARALICK, R. M., 1979, Statistical and structural approaches to texture. Proceedings of the I.E.E.E., 67, 786 - 804.

HSU, S., 1977, A texture–tone analysis for automated land–use mapping with panchromatic images. Proceedings of the annual meeting of ASPRS/ACSM, Washington D.C. (Falls Church, VA: American Congress on Surveying and Mapping).

HUDSON, W. D., & RAMM, C. W., 1987, Correct formulation of the Kappa coefficient of agreement. Photogrammetric Engineering and Remote Sensing, 53, 421 - 422.

JAIN, A. K., 1989, Fundamentals of Digital Image Processing, Prentice-Hall, Englewood Cliffs, New Jersey.

JENSEN, J. R., 1986, Introductory Digital Image Processing, A Remote Sensing Perspective (Englewood Cliffs, NJ: Prentice – Hall).

JENSEN, L., 1990, Knowledge-Based Classification of an Urban Area Using Texture and Context Information in Landsat-TM Imagery, Photogrammetric Engineering and Remote Sensing, Vol. 56, No. 6, pp. 899 – 904.

KAI, L. & MULLER, J. P., 1991, Segmenting Satellite Imagery: A Region Growing Scheme, International Geoscience and Remote Sensing Symposium, Helsinki University of Technology, Espoo, Finland, I.E.E.E., Vol 2., pp 1075 – 1078.

LIRA, J. & LANDGRAVE, R., 1992, Sistema de Análisis Digital de Imágenes, V. 2.0, Revista SELPER, Vol. 8.

LIRA, J., 1997, Procesamiento de Imágenes y Reconocimiento de Patrones, Notas de cursos, Maestría en Ingeniería Eléctrica, DEPEFI, UNAM, México, D.F.

LIRA, J., FRULLA, L., CHAVIRA, D., & PASTRANA, R., 1996, Morphology of Segmented Regions in SAR Images by Means of a Region Growing Algorithm. Proceedings of the First Latino-American Seminar on Radar Remote Sensing – Image Processing Techniques, Buenos Aires, Argentina.

LIRA, J. & FRULLA, 1998, An automated Region Growing Algorithm for Segmentation of Texture Regions in SAR Images, International Journal of Remote Sensing, aceptado para publicación.

OLSEN, E. R., RAMSEY R. & WINN D., 1993, A Modified Fractal Dimension as a Measure of Landscape Diversity, Photogrammetric Engineering & Remote Sensing, Vol. 59, pp. 1517 – 1520.

OPPENHEIM, A. V. & SCHAFER, R. W., 1975, Digital Signal Processing, Prentice-Hall International Editions.

PAPOULIS, A., 1991, Probability, Random Variables and Stockastic Processes, Mc. Graw Hill Book Co..

PASTRANA, R., 1996, Restauración de Imágenes Digitales: Eliminación de Ruido por Medio de un Filtro Geométrico, Tesis de Licenciatura, Facultad de Ciencias Físico Matemáticas Colegio Electrónica, BUAP.

RICHARDS, J. A., 1993, Remote Sensing Digital Image Analysis: An introduction (Berlin: Springer - Verlag).

RICHARDS, J. A., 1986, Remote Sensing Digital Image Analysis (Berlin, New York, London: Springer – Verlag).

ULABY, F., MOORE, R. & FUNG A. , 1986, Microwave Remote Sensing Active and Passive, Artech House, Washington.

WILSON, J. D., 1992, A comparison of procedures for classifying remotely-sensed data using simulated data sets, Int. J. Remote Sensing, Vol. 13, No 2, 365 – 386.

WOODCOCK, C. E., STRAHLER, A. H., & LOGAN, T. L., 1980, Stratification of forest vegetation for timber inventory using Landsat and collateral data. Proceedings of the 14th International Symposium on Remote Sensing of Environment, San Jose, Costa Rica, 23 – 30 April 1980, pp. 1769 – 1787.

WOODCOCK, C. & HARWARD, V. J., 1992, Int. J. Remote Sensing, Vol 13, No. 16, 3167 – 3187.

ZUCKER, S. W., 1976, Region Growing: Childhood and Adolescence, Computer Graphics and Image Processing, Vol. 5, pp. 382 – 399.