

03063



UNIVERSIDAD NACIONAL  
AUTÓNOMA DE  
MÉXICO

**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**

POSGRADO EN CIENCIA E INGENIERÍA DE LA COMPUTACIÓN

**“SEGMENTACION DE IMÁGENES  
MULTIESPECTRALES UTILIZANDO EL ALGORITMO  
ICM (ITERATED CONDITIONAL MODES)”**

**T E S I S**

**QUE PARA OBTENER EL GRADO DE:**

**MAESTRO EN INGENIERÍA  
(COMPUTACIÓN)**

**P R E S E N T A:**

**DAVID ESTRADA ESPINOSA**

**DIRECTOR DE TESIS: DR. BORIS ESCALANTE RAMÍREZ**

MÉXICO, D.F.

2005.

m 344628



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

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

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

# Agradecimientos

Posgrado en Ciencias e Ingeniería de la Computación:  
Gracias por el apoyo otorgado en la realización de este trabajo.

Dr. Boris Escalante Ramírez:  
Gracias por la dirección de este trabajo.

A mis amigos y compañeros del posgrado:  
Gracias por su apoyo y su amistad.

A mi familia:  
Gracias por su apoyo.

Autorizo a la Dirección General de Bibliotecas de la UNAM a difundir en formato electrónico e impreso el contenido de mi trabajo recepcional.

NOMBRE: David Estrada

FECHA: 25/05/05

FIRMA: 

# INDICE

ÍNDICE.....	1
RESUMEN.....	3
CAPÍTULO 1	
INTRODUCCIÓN.....	5
CAPÍTULO 2	
COMPONENTES PRINCIPALES, K-MEDIAS Y CAMPOS ALEATORIOS DE MARKOV.....	9
2.1 DESCOMPOSICIÓN EN COMPONENTES PRINCIPALES.....	9
2.1.1 REDUCCIÓN DE LA DIMENSIONALIDAD EN BASE A LA TRANSFORMACIÓN EN COMPONENTES PRINCIPALES.....	14
2.2 K-MEDIAS.....	15
2.3 CAMPOS ALEATORIOS DE MARKOV.....	17
2.3.1 MODELO DE ISING.....	17
2.3.2 DISTRIBUCIÓN GIBBS.....	19
2.3.3 SISTEMAS DE VECINOS.....	21
2.3.4 MODELO PARA LA SEGMENTACIÓN DE IMÁGENES.....	22
2.3.4.1 ALGORITMO ICM.....	23
2.4 EVALUACIÓN DE LA SEGMENTACIÓN.....	24
2.4.1. SEPARABILIDAD DE REGIONES.....	25
2.4.2. MATRIZ DE CONFUSIÓN.....	28
CAPÍTULO 3	
APLICACIÓN DE LOS ALGORITMOS.....	29
3.1 COMPONENTES PRINCIPALES.....	29
3.2 K-MEDIAS.....	34
3.3 SEGMENTACIÓN ICM.....	34
3.4 EVALUACIÓN.....	35

CAPÍTULO 4	
ANÁLISIS DE RESULTADOS.....	43
4.1 PARÁMETRO BETA.....	43
4.2 DIMENSIONALIDAD INTRÍNSECA.....	44
4.3 DIFERENCIAS ENTRE K-MEDIAS E ICM.....	47
4.4 TIEMPO DE PROCESAMIENTO.....	47
 CAPÍTULO 5	
CONCLUSIONES.....	51
 APÉNDICE 1	
PROBABILIDAD Y VARIABLES ALEATORIAS.....	53
 APÉNDICE 2	
ALGORITMOS EN MATLAB 6.5.....	59
 APÉNDICE 3	
MATRIZ DE LA IMAGEN 3.....	71
 BIBLIOGRAFÍA.....	77

# RESUMEN

La segmentación de una imagen multiespectral utilizando métodos contextuales (campos aleatorios de Markov) presentan considerables ventajas cualitativas y cuantitativas en la valoración de los resultados con respecto a los métodos estándar de segmentación de imágenes (K-medias, isodata, y otros).

En el presente trabajo se realiza la implementación un algoritmo segmentación bajo el algoritmo Iterated Conditional Modes (segmentación ICM) para imágenes multiespectrales encontrándose en las imágenes sintéticas de prueba a color con ruido mejoras hasta en un 57% con respecto a la segmentación K-medias.

También se realiza un análisis de la dimensionalidad intrínseca de las imágenes bajo la transformación en componentes principales concluyendo que esta sólo es de utilidad cuando las bandas no están fuertemente correlacionadas.

Los resultados obtenidos en el presente trabajo permiten recomendar su utilidad dado que la segmentación ICM converge en un número pequeño de iteraciones (6) y la complejidad del algoritmo es lineal, el cual resulta factible para su aplicación en imágenes de percepción remota.

# CAPÍTULO 1

## INTRODUCCIÓN

Una imagen multiespectral es la representación numérica de la distribución espacial de la respuesta espectral de los objetos que se encuentran en la escena, de acuerdo con un campo instantáneo de vista empleado para digitalizarla.

La segmentación de imágenes tiene su origen en numerosos estudios psicológicos que indican la preferencia de los humanos por agrupar regiones visuales en términos de proximidad, similitud y continuidad, para construir un conjunto de unidades *significativas*.

La segmentación de una imagen multiespectral es el proceso de particionar los píxeles de la imagen en número finito de regiones homogéneas que llamaremos etiquetas, basados en los valores espectrales de los píxeles. Si un píxel satisface cierto conjunto de criterios, el píxel es asignado a una región que corresponde a ese criterio.

La segmentación puede llevarse a cabo de manera puntual, es decir, segmentar un píxel en función únicamente de sus propiedades espectrales; o bien de manera contextual, donde un píxel se segmenta de acuerdo con el contexto espectral en el que se encuentra.

El método orientado a píxel considera sólo los valores espectrales del mismo sin considerar otro tipo de información para realizar la segmentación de la imagen. Ejemplos de este tipo de segmentación son los algoritmos K-medias e isodata.

Al contrario en el método contextual, la segmentación se realiza considerando los valores espectrales del píxel y los valores de los píxeles vecinos. Así se obtiene una mejor segmentación, tanto cuantitativa como cualitativa. En términos de probabilidad podemos decir que la probabilidad de que un píxel tenga la misma etiqueta que la de su vecino es mayor que la probabilidad de que sean distintas, y aún más, es muy improbable que un píxel tenga una

etiqueta diferente a la de todos sus vecinos e incluso es posible afirmar que ciertas configuraciones muy poco probables.

El mayor inconveniente de estos métodos consiste en que son iterativos ó recursivos, dificultando su implantación en sistemas que requieren un tiempo de respuesta rápido y reducen la posibilidad del paralelismo solamente al dominio espacial [Pr93]. También existe una pequeña dificultad a la hora de establecer el criterio por el cual una región pequeña es asimilada por otra grande, exigiendo un compromiso entre la eliminación de ruido y la pérdida de detalle. Otra dificultad más es la de determinar con precisión la frontera entre regiones cuando las variaciones de intensidad son graduales, produciendo a veces un efecto de ambigüedad.

La necesidad de incorporar la información contextual en el proceso de segmentación de imágenes surge dado que las segmentaciones no contextuales suelen ser segmentaciones ruidosas, entendiéndose que una segmentación ruidosa presenta mucha variabilidad en la distribución espacial de las etiquetas y, como consecuencia, no aparecen regiones homogéneas para una misma etiqueta. La fuente de este ruido pueden ser muy diversas [Ka94] y entre ellas se encuentran: una respuesta no uniforme del sensor, dispersión atmosférica en el caso de imágenes de percepción remota, errores en el procesamiento, la resolución del sensor y la naturaleza del método de segmentación, en consecuencia, es necesario algún tipo de procesamiento sobre la segmentación obtenida, de forma que reestructuren algunas de las etiquetas con el objetivo de conseguir regiones homogéneas que reflejen la estructura real de la escena. Con este objetivo se han desarrollado una amplia familia de técnicas de segmentación que se denominan segmentadores contextuales o métodos de corrección contextual; entre ellos, se encuentran la utilización de los campos aleatorios de Markov para la segmentación de una imagen [Ge84], [Ya99]

En este trabajo se realizará la implementación y la evaluación a través de la matriz de confusión de la segmentación de imágenes multiespectrales utilizando la aplicación de los campos aleatorios de Markov que incorporan la información relativa a los valores de los de los pixeles vecinos en el proceso de segmentación. Ya que en general se reportan en la literatura que a través de este tipo de segmentaciones se obtienen mejores resultados en comparación a los métodos estándar, además, del gran atractivo que han tenido recientemente por su empleo en el procesamiento digital de imágenes.



Los campos aleatorios de Markov son utilizados en procesamiento de imágenes para su segmentación [Lu03] [Mu01] [Ya99] [Wi02], restauración [Ge84] [Pr04], detección de bordes y el análisis del movimiento de secuencias de imágenes [Li95].

Estos modelos permiten representar adecuadamente la información a priori de una imagen de forma que la teoría de decisión bayesiana se puede aplicar, pero en la segmentación de una imagen hay que especificar primero un modelo de la imagen subyacente que sirva como información contextual del criterio de decisión (sistema de vecindad) y segundo un modelo de la imagen segmentada que sirva para describir la relación entre los valores espectrales de los píxeles y las etiquetas a discriminar (funciones de densidad de probabilidad de las etiquetas) [Ka93].

El modelo que se utiliza en este trabajo fue propuesto por Besag [Be74], [Be86], llamado ICM éste especifica el modelo contextual de la imagen subyacente que permite encontrar la probabilidad condicional de que el píxel considerado pertenezca a una región dado su sistema de vecindad de manera sencilla (interacción por parejas).

El procedimiento que se siguió en este trabajo para conseguir el objetivo planteado es el siguiente:

1. Descomposición en componentes principales.

La descomposición en componentes principales se realiza mediante la transformada de Karhunen-Loéwe con la finalidad de reducir la dimensionalidad de la imagen multiespectral y de esta manera reducir el tiempo computacional de cálculo del proceso de segmentación.

2. Aplicación del algoritmo K-medias.

A partir de considerar sólo las primeras componentes principales se procede a realizar la segmentación de la imagen aplicando el algoritmo de k medias (modelo a priori), que sirve de entrada al algoritmo ICM.

### 3. Aplicación del algoritmo ICM.

Se utiliza el algoritmo ICM para realizar la segmentación de la imagen teniendo como entrada la imagen original, la imagen segmentada por el algoritmo de K-medias y el sistema de vecindad utilizado.

### 4. Evaluación de los resultados.

El análisis cuantitativo de los resultados de la segmentación se realiza a partir de la matriz de confusión.

A continuación se describe la organización de los capítulos de esta tesis.

En el capítulo 2, se presentan los aspectos teóricos de la descomposición en componentes principales, la segmentación por el algoritmo K-medias y los campos aleatorios de Markov, presentando el algoritmo ICM en la segmentación de imágenes multiespectrales.

Por último, se presentan los métodos de evaluación cuantitativa de la segmentación.

En el capítulo 3, Se presentan los resultados de los algoritmos de componentes principales, K-medias, ICM y los resultados de los descriptores de la evaluación.

En el capítulo 4, Se realiza un análisis de los resultados en general, comenzando por la elección del valor del parámetro beta de la segmentación ICM, la dimensionalidad intrínseca de imágenes sintéticas en RGB con ruido gaussiano y finalmente la evaluación de la segmentación de ICM con respecto a K-medias utilizando el coeficiente kappa de la matriz de confusión.

En el capítulo 5, se presentan las conclusiones del trabajo.

# CAPÍTULO 2

## COMPONENTES PRINCIPALES, K-MEDIAS Y CAMPOS ALEATORIOS DE MARKOV

En éste capítulo, se presentan los aspectos teóricos de la descomposición en componentes principales, la segmentación por el algoritmo K-medias, los campos aleatorios de Markov, presentando el algoritmo ICM en la segmentación de imágenes multiespectrales y por último, se presentan los métodos de evaluación cuantitativa de la segmentación.

### 2.1 Descomposición en componentes principales

El objetivo del análisis de componentes principales es transformar la imagen multiespectral  $P$  en una nueva  $P'$ , en el que las bandas de la imagen estén descorrelacionadas (la matriz de covarianza en este espacio será diagonal). En otras palabras, se trata de encontrar un nuevo conjunto de ejes ortogonales en el que la varianza de las bandas sea máxima. El objetivo final es reducir la dimensionalidad de la imagen multiespectral una vez realizada la transformación al utilizar sólo las primeras componentes (las primeras bandas de la nueva imagen multiespectral  $P'$ ) [Ri99].

La transformación de componentes principales se plantea como la búsqueda de una transformación lineal de los valores de los píxeles,  $W$ , de la imagen multiespectral  $P$  en el espacio de coordenadas  $X$ , que da lugar a nuevas coordenadas  $Y$  en  $P'$ , tal que:

$$Y = W^T X \quad (2.1)$$

la ecuación anterior se puede escribir de manera explícita como:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_d \end{bmatrix} = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1d} \\ w_{21} & w_{22} & \cdots & w_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ w_{d1} & w_{d2} & \cdots & w_{dd} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_d \end{bmatrix} \quad (2.2)$$

para una imagen multiespectral de  $d$  bandas, de forma que su desarrollo es:



derivando respecto a  $W$ , se obtiene

$$(\sigma_x^2 - \lambda I)W = 0 \quad (2.9)$$

resolviendo el sistema de ecuaciones dado por (2.9), se encuentra la matriz,  $W$  que debe verificar que  $(\sigma_x^2 - \lambda I)W = 0$  con la condición de que  $\sigma_y^2 = W^T \sigma_x^2 W$  sea máxima y sujeta a la restricción de que  $W^T W = I$ .

Para que la ecuación (2.9) sea cierta sólo pueden ocurrir dos casos:

1. Que  $W = 0$  y en este caso la solución es trivial y no interesa.
2. Que  $\sigma_x^2 - \lambda I$  sea singular (no invertible), esto es, que:

$$|\sigma_x^2 - \lambda I| = 0 \quad (2.10)$$

La ecuación (2.10) es la ecuación característica de la matriz  $\sigma_x^2$  y su expresión es una ecuación polinomial de  $\lambda$ . Las soluciones a esta ecuación (los valores de  $\lambda$ ) se conocen como los autovalores de  $\sigma_x^2$ . Cuando se sustituyen en (2.9), se calculan los vectores asociados a cada valor de  $\lambda$ , que se conocen como los autovectores de  $\sigma_x^2$ .

Cada autovalor  $\lambda_i$ , es solución a una ecuación del sistema  $(\sigma_x^2 - \lambda_i I)W = 0$ . De tal forma que para cada ecuación, los parámetros de  $W$  asociados a la solución con  $\lambda_i$ , es un autovector  $\phi_i$ . Entonces podemos expresar la matriz de transformación como un vector de  $d$  vectores columna ( $W$  es un vector de autovectores).

$$W = [\phi_1, \phi_2, \dots, \phi_d] \quad (2.10)$$

Como  $\sigma_x^2$  es de orden  $d \times d$ , tendrá  $d$  autovalores asociados,  $\lambda_1, \lambda_2, \dots, \lambda_d$ , y como  $\sigma_x^2$  es simétrica, todos los autovalores serán reales.

Por lo tanto, como  $\sigma_x^2$  es simétrica, todos sus autovalores son reales y dado que es definida como positiva, sus autovalores están ordenados:

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d \quad (2.11)$$

La matriz de covarianza  $\sigma_y^2 = W^T \sigma_x^2 W$  es una matriz diagonal formada por los autovalores de  $\sigma_x^2$ :

$$\sigma_y^2 = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_d \end{bmatrix} \quad (2.12)$$

y los valores de la diagonal (autovalores de  $\sigma_x^2$ ) son las varianzas de los patrones en las respectivas coordenadas transformadas.

La matriz que contiene los coeficientes de la transformación,  $W$ , es la matriz de autovectores de  $\sigma_x^2$ , asumiendo que  $W$  es ortogonal.

Cada autovalor  $\lambda_i$  tiene asociado un autovector  $\phi_i$  y cada autovector define la dirección de un eje en el espacio transformado,  $P'$ . Dado que los autovalores están ordenados (por el valor de varianza en cada eje de  $P'$ ), y que cada autovalor tiene asociado un autovector, podemos establecer un orden entre las variables transformadas de forma que:

$Y^1$ : Primer eje en  $P'$  (primera componente principal).

La dirección de la máxima varianza de los patrones en  $P$  está determinada por este eje.

$Y^2$ : Segundo eje en  $P'$  (segunda componente principal).

La dirección de la máxima varianza en  $P$  entre todos los ejes ortogonales a  $Y^1$  está determinada por este eje.

y así sucesivamente para  $Y^3, \dots, Y^d$ .

Con estas consideraciones, el teorema fundamental del análisis de componentes principales se enuncia como sigue: Dado un conjunto de variables  $X^i$  ( $i = 1, 2, \dots, d$ ) con matriz de covarianza  $\sigma_x^2$ , no singular, siempre se puede derivar a partir de ellos un conjunto de variables descorrelacionadas  $Y^i$  ( $i = 1, 2, \dots, d$ ) mediante un conjunto de transformaciones lineales que corresponden a una rotación rígida cuya matriz de transformación  $W$  está formada por columnas, y los  $d$  autovectores de  $\sigma_x^2$ . La matriz de covarianza del nuevo conjunto de variables,  $\sigma_y^2$ , es diagonal, y contiene los autovalores de  $\sigma_x^2$ .

La transformación de componentes principales definida por (2.1) con las restricciones de diagonalidad (2.5) y (2.6) se conoce también como transformación de Karhunen-Loéwe ó de Hotelling.

A modo de resumen, y con una interpretación geométrica, los autovalores  $\lambda_i$  representan la varianza de las bandas en el espacio transformado y están relacionadas con las bandas en cada uno de los ejes de este espacio mientras que los autovectores  $\phi_i$  son vectores ortogonales que determinan la dirección de estos ejes.

La transformación por componentes principales es una transformación que preserva la varianza.

Si se define la varianza total de una imagen multispectral como la suma de las varianzas asociadas a cada atributo, como las varianzas individuales se encuentran en la diagonal de la matriz de covarianza, el cálculo de la varianza global se reduce al cálculo de la traza de la matriz de covarianza. Resulta evidente que si  $\sigma_Y^2$  es la matriz que contiene en su diagonal los autovalores  $\lambda_1, \lambda_2, \dots, \lambda_d$  de  $\sigma_X^2$ , entonces:

$$\text{tr}(\sigma_Y^2) = \sum_{i=1}^d \lambda_i$$

Además, si la transformación de componentes principales preserva la varianza global, entonces:

$$\text{tr}(\sigma_X^2) = \text{tr}(\sigma_Y^2) = \sum_{i=1}^d \lambda_i \quad (2.13)$$

El algoritmo para calcular  $W$  consiste en 4 pasos:

1. Calcular la matriz de covarianza global  $\sigma_X^2$ .
2. Calcular los autovalores de  $\sigma_X^2, \lambda_1, \lambda_2, \dots, \lambda_d$ .
3. Calcular los autovectores  $\phi_1, \phi_2, \dots, \phi_d$  asociados a  $\lambda_1, \lambda_2, \dots, \lambda_d$ .
4. Formar la matriz  $W = [\phi_1, \phi_2, \dots, \phi_d]$ .

Una vez formada la matriz de transformación  $W$  se procede a calcular los nuevos valores de las bandas  $Y$ , a partir de cada  $X$ . Como  $Y = W^T X$  y la matriz  $W$  es la matriz formada por los autovectores de  $\sigma_X^2$  sustituyendo en ec. (2.2) obtenemos:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_d \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} & \dots & \phi_{1d} \\ \phi_{21} & \phi_{22} & \dots & \phi_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{d1} & \phi_{d2} & \dots & \phi_{dd} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_d \end{bmatrix}$$

o lo que es lo mismo

$$\begin{aligned}
Y_1 &= \phi_{11}X_1 + \phi_{12}X_2 + \dots + \phi_{1d}X_d \\
Y_2 &= \phi_{21}X_1 + \phi_{22}X_2 + \dots + \phi_{2d}X_d \\
&\vdots \\
&\vdots \\
Y_d &= \phi_{d1}X_1 + \phi_{d2}X_2 + \dots + \phi_{dd}X_d
\end{aligned}$$

donde

$$\phi_1 = \begin{bmatrix} \phi_{11} \\ \phi_{12} \\ \vdots \\ \phi_{1d} \end{bmatrix}, \quad \phi_2 = \begin{bmatrix} \phi_{21} \\ \phi_{22} \\ \vdots \\ \phi_{2d} \end{bmatrix}, \quad \dots, \quad \phi_d = \begin{bmatrix} \phi_{d1} \\ \phi_{d2} \\ \vdots \\ \phi_{dd} \end{bmatrix}$$

### 2.1.1 Reducción de la dimensionalidad en base a la transformación en componentes principales

Este problema se conoce en la literatura como el cálculo de la *dimensionalidad intrínseca*, que puede definirse como el menor número de variables utilizado para representar con el mínimo error el conjunto complemento de variables.

El fundamento está en el orden los autovalores de una matriz de covarianza ec. (2.11) y que la transformación de componentes principales preserva la varianza global ec. (2.13). Se define  $\gamma_i$  como la contribución a la varianza global de un autovalor como el valor normalizado de un autovalor  $\lambda_i$ :

$$\gamma_i = \frac{\lambda_i}{\sum_{j=1}^d \lambda_j} \quad (2.14)$$

esto es, se asocia a cada autovalor ( y en consecuencia, a cada componente principal) un valor relativo de importancia  $\gamma_i$ , considerando su aportación porcentual a la varianza total.

Considerando que los autovectores están ordenados de manera decreciente, la mayor contribución a la varianza global será la de los primeros autovalores por lo que serán las primeras componentes principales las que aporten la mayor parte de la información (en términos de la varianza) mientras que las últimas se pueden considerar básicamente como ruido y pueden descartarse.



Por ejemplo, es bien conocido que la dimensionalidad intrínseca de los datos suministrados por los satélites Landsat 1 y 2 es aproximadamente 2. Esto significa que las primeras dos componentes principales pueden resumir la información contenida en el conjunto completo de variables originales. Es decir, la imagen que corresponde a la primera componente principal es la que asocia al mayor autovalor, o sea, la que presenta mayor varianza, que se traduce en una imagen con mayor contraste.

La manera de calcular la dimensionalidad intrínseca de una imagen multiespectral es utilizando una metodología empírica que se describe a continuación. Supongamos que disponemos de una imagen multiespectral de 6 bandas después de calcular su transformación en componentes principales, se procede a segmentar los píxeles utilizando las 6 componentes principales, para después segmentarlos utilizando 5 (descartando la componente 6), y así sucesivamente. Al realizar esta prueba se comprueba que no hay diferencia significativa cualitativa y cuantitativa al utilizar las 6 componentes principales o sólo utilizar las primeras 3, se dice, que la dimensionalidad intrínseca es 3.

## 2.2 K-Medias

Este es el método de segmentación más popular, también llamado de las medias móviles porque en cada iteración se calcula el valor de la media de los píxeles asignados a un cúmulo llamado centro. Inicialmente se seleccionan  $k$  centros de manera aleatoria y asignando los píxeles al cúmulo cuyo centro se encuentre más cercano. Cada cúmulo formará una clase al final del algoritmo, ya que comparten alguna correlación entre sus valores espectrales [St02].

El algoritmo puede plantearse en tres pasos:

1. Inicializar de manera aleatoria los centros de  $k$  cúmulos en el espacio multiespectral.
2. Asignación de los píxeles al cúmulo cuyo centro este más cercano, actualizando el valor de los centros y calculando las medias de los píxeles asignados al cúmulo. La medición de la distancia de un píxel al centro de un cúmulo se realiza empleando la métrica euclídeana, pero se pueden utilizar otras.

3. Convergencia. Los píxeles cambian de cúmulo asignado y en consecuencia, los centros de éstos. Si esto ocurre se repite el paso 2 hasta que no se modifiquen los centros. Cuando no haya modificaciones se considera que se ha encontrado una buena segmentación y el algoritmo termina.

El algoritmo es muy sencillo y eficiente ya que todos los cálculos resultan muy sencillos. El inconveniente que tiene, es que el resultado final depende del valor de  $k$  y la manera como se inicializan los centros, esto ocasiona que el resultado esté sesgado.

Ahora la convergencia es rápida para casos sencillos donde la imagen tiene etiquetas bien diferenciadas, de lo contrario el costo computacional es muy alto. De aquí que a veces se introducen criterios para la terminación anticipada del algoritmo: cuando el desplazamiento de las medias sea menor a una cierta distancia especificada, o bien, cuando el cambio del número píxeles de los cúmulos sea menor a una cota predeterminada.

Esto se justifica con el hecho de que experimentalmente los cúmulos de una segmentación dada se definen, en su mayor parte, en las primeras iteraciones y a partir de un cierto punto los cambios de los píxeles de cúmulos y los desplazamientos de las medias son muy pequeños.

Una vez que el espacio multiespectral ha sido segmentado de acuerdo al algoritmo de K-medias, es necesario generar la imagen en donde cada cúmulo represente una clase espectral. Esto se realiza de la siguiente manera: al terminar el proceso de segmentación, el valor de las medias de los cúmulos finales es conocida, de aquí que, en la última etapa, cada píxel de la imagen es etiquetado con el centroide cuya distancia a éste sea mínima. A partir de aquí se genera una imagen de salida, de una sola banda, donde los píxeles con la misma etiqueta tienen el mismo valor, es decir, en su representación visual se convierte en un mismo color para los píxeles asociados al cúmulo. La etiquetación y generación de la imagen de salida se realiza en un sólo paso, al establecer la relación del cúmulo con la clase espectral. La imagen de salida se le conoce como mapa temático dado que cada cúmulo representa una clase espectral.

Para hacer más eficiente el análisis de los cúmulos se puede aplicar primero la transformación en componentes principales a la imagen multiespectral y utilizando posteriormente sólo las primeras

bandas en el algoritmo de K-medias, esto reduce el costo computacional y mejora la separación espectral de las etiquetas.

## 2.3 Campos aleatorios de Markov

Alrededor de 1920 apareció un nuevo tipo de procesos estocásticos en la teoría probabilística inspirado en el modelo Ising llamado Campos Aleatorios de Markov (CAM). Esta se convirtió rápidamente en una herramienta ampliamente usada en una variedad de problemas no solamente en mecánica estadística donde tuvo su origen, su uso en el proceso de imágenes llegó a ser popular en el famoso artículo publicado por los hermanos S. Geman y D. Geman en 1984 [Ge84].

### 2.3.1 Modelo de Ising

El ferromagnetismo, que aparece en muchos metales ordinarios como el hierro y el níquel, es la presencia de magnetización espontánea incluso cuando no hay campo magnético externo. Se debe a que una fracción importante de los momentos magnéticos (ó espines) de los átomos se alinean en la misma dirección debido a la interacción entre los mismos.

El modelo de Ising es un modelo sencillo para el estudio analítico de la transición ferromagnética. Se parte de una red regular, que imita la red cristalina del hierro ó níquel, en cuyos sitios se colocan los átomos cuyo momento magnético o espín puede tomar los valores  $+1$  ó  $-1$ . Este alineamiento se produce únicamente a temperaturas bajas, por debajo de una temperatura característica llamada temperatura de Curie,  $T_C$ .

Procediendo como Ising, nosotros consideramos una red regular unidimensional como una secuencia  $0, 1, 2, \dots, n$  sobre una línea. En cada punto, hay un átomo que tiene un espín el cual puede estar hacia abajo ó hacia arriba en un momento dado, ver Figura 2.1.

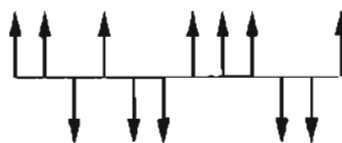


Figura 2.1 Una secuencia de espines, éstos pueden estar hacia abajo o hacia arriba.

Ahora, procederemos a definir la probabilidad de medir la configuración  $\omega = (\omega_0, \omega_1, \dots, \omega_n)$  sobre el conjunto de todas las posibles configuraciones  $\Omega$ . En este contexto el valor del espín de cada átomo está dado por la función

$$\delta_i(\omega) = \begin{cases} 1 & \text{Si } \omega_i \text{ está hacia arriba} \\ -1 & \text{Si } \omega_i \text{ está hacia abajo} \end{cases} \quad (2.15)$$

La energía asociada a esta configuración está dada por:

$$U(\omega) = -J \sum_{i,j} \delta_i(\omega) \delta_j(\omega) - mH \sum_i \delta_i(\omega) \quad (2.16)$$

En el primer término de la suma, Ising simplifica la interacción entre los espines de los átomos suponiendo que ésta sólo contribuye cuando los espines contiguos están alineados, es decir, apuntan en la misma dirección. La constante  $J$  es una propiedad del material. Si  $J > 0$  mantiene la contribución de los espines en la misma dirección (caso atractivo), en caso contrario, favorece a la contribución de los espines que apuntan en direcciones contrarias (caso repulsivo). El segundo término de la suma representa la influencia de un campo magnético externo  $H$  y  $m$  es una propiedad del material.

Entonces la probabilidad de medir la configuración  $\omega = (\omega_0, \omega_1, \dots, \omega_n)$  sobre  $\Omega$ , está dado por:

$$P(\omega) = \frac{e^{\left(\frac{1}{kT} U(\omega)\right)}}{Z} \quad (2.17)$$

donde  $T$  es la temperatura y  $k$  es una constante universal. La constante de normalización  $Z$  (función de partición) esta definida por

$$Z = \sum_{\omega \in \Omega} e^{\left(\frac{1}{kT} U(\omega)\right)} \quad (2.18)$$

La probabilidad definida por la ec. (2.17) es llamada distribución de Gibbs.

Este modelo puede extenderse a dos o más dimensiones naturalmente. En el caso de dos dimensiones colocamos a cada átomo en una red ortogonal, que está representado por dos coordenadas y cada átomo tiene al menos 4 vecinos excepto en las orillas. Si medimos la probabilidad de configuración de los espines de los átomos en la red, es inestable ya que se presenta una transición de fase. Como se precisa en [Ki80], considerando el caso atractivo y un campo externo  $h$ , la medición de  $P_h$  converge a  $P$  si  $h$  tiende a cero por la izquierda (a través de

valores negativos), pero converge a  $P^+ \neq P^-$  si  $h$  tiende a cero por derecha (a través de valores positivos).



Figura 2.2 Modelo del árbol de Cayley

Un caso especial es el modelo propuesto por Bethe [Ba90] llamado modelo del árbol de Cayley [Ki80], como una aproximación al modelo de Ising. En éste caso, los átomos se colocan sobre un árbol. La raíz tiene el nivel cero. A partir de la raíz, se colocan  $q$  ramas, (ver Figura 2.2). Si  $q=1$  tenemos el caso unidimensional de las cadenas de Markov. A una configuración sobre el árbol de  $n$  niveles se le asocia el estado en que se encuentra el espín del átomo. Y procediendo de la misma manera que el caso unidimensional del modelo de Ising se calcula la función de energía. Otra extensión del modelo de Ising a más de dos estados del espín por cada átomo es el modelo de Potts [Wu82], este puede utilizarse para imágenes de varios niveles de gris.

### 2.3.2 Distribución Gibbs

La manera más natural de definir CAM relacionado con los modelos de una imagen es definirlos en una red. Sin embargo, definiremos un CAM de manera más general sobre un grafo. Sea  $G = (S, E)$  un grafo donde  $S = \{s_1, s_2, \dots, s_N\}$  es el conjunto de vértices (ó sitios) y  $E$  es el conjunto de aristas

Def. 2.1. (Vecinos) Dos vértices  $s_i$  y  $s_j$  son vecinos si existe una arista  $e_{ij} \in E$  que los conecta. Al conjunto de vértices vecinos de  $s$  se denota por  $V_s$ .

Def. 2.2 (Sistema de vecindad)  $V = \{V_s / s \in S\}$  es un sistema de vecindad de  $G$  si

- (i)  $s \in V_s$
- (ii)  $s \in V_r \Leftrightarrow r \in V_s$

A cada sitio del grafo, se le asigna una etiqueta  $\lambda$  que pertenece a un conjunto finito de etiquetas  $\Lambda$ . Esta asignación es llamada la configuración  $\omega$  que tiene probabilidad  $P(\omega)$ . La restricción a un subconjunto  $T \subset S$  es denotada por  $\omega_T$  y  $\omega_s \in \Lambda$  denota la etiqueta dada al sitio  $s$ .

Def. 2.3 (Campo Aleatorio de Markov)  $\chi$  es un Campo Aleatorio de Markov con respecto a  $V$  si

(i) Para todo  $\omega \in \Omega$ :  $P(\chi = \omega) > 0$

(ii) Para cada  $s \in S$  y  $\omega \in \Omega$ :

$$P(\chi_s = \omega_s / \chi_r = \omega_r, r \neq s) = P(\chi_s = \omega_s / \chi_r = \omega_r, r \in V_s).$$

Def. 2.4 (Clique) Un subconjunto  $C \subseteq S$  es un clique si para cada par de diferentes sitios en  $C$  son vecinos.  $\zeta$  denota el conjunto de todos los cliques y su grado esta dado por  $\text{gdo}(\zeta) = \max_{C \in \zeta} |C|$ .

Sea  $v$  el potencial que asigna un número  $v_T(\omega)$  a cada configuración  $\omega_T$ . Entonces  $v$  define una energía  $U(\omega)$  sobre  $\Omega$  dada por:

$$U(\omega) = - \sum_T v_T(\omega) \quad (2.19)$$

Def. 2.5 La distribución de Gibbs es la probabilidad de medición  $\pi$  sobre  $\Omega$  y está dada por:

$$\pi(\omega) = \frac{e^{-U(\omega)}}{Z} \quad (2.20)$$

donde  $Z$  es la constante de normalización o función de partición dada por:

$$Z = \sum_{\omega} e^{-U(\omega)} \quad (2.21)$$

Si  $v_T(\omega) = 0$  siempre que  $T$  no es un clique entonces  $v$  es llamado los vecinos más cercanos del potencial de Gibbs.

Teorema 2.1 (Hammersley-Clifford)  $\chi$  es un CAM con respecto a un sistema de vecindad  $V$ , si y sólo si,  $\pi(\omega) = P(\chi = \omega)$  es una distribución Gibbs con los vecinos más cercanos al potencial de Gibbs  $v$ , esto es

$$\pi(\omega) = \frac{e^{-\sum_{C \in \zeta} v_C(\omega)}}{Z} \quad (2.22)$$

Por lo tanto, de las ecuaciones (2.20) y (2.22)

$$U(\omega) = \sum_{C \in \mathcal{C}} v_C(\omega) \quad (2.23)$$

La equivalencia entre los CAM y la distribución de Gibbs es que nos permite especificar los potenciales a partir de las características locales lo cual es muy difícil.

### 2.3.3 Sistemas de vecinos

Se considera al conjunto de sitios  $S$  como una rejilla  $L$  tal que  $\forall s \in S: s = (i, j)$  y definimos al sistema de vecindad de  $n$ -ésimo orden como:

$$V^n = \{V_{(i,j)}^n : (i, j) \in L\} \quad (2.24)$$

$$V_{(i,j)}^n = \{(k, l) \in L : (k - i)^2 + (l - j)^2 < n\} \quad (2.25)$$

Obviamente, los sitios en la fronteras de la rejilla tienen menos vecinos que los interiores (condiciones de frontera libre). Además,  $V^0 = S$  y para todo  $n \geq 0: V^n \subset V^{n+1}$ . La Figura 2.3 muestra un sistema de vecindad de primer orden y sus cliques dados por  $\{(i, j), \{(i, j), (i, j + 1)\}, \{(i, j), (i + 1, j)\}$ , también se muestra el sistema vecindad de segundo orden y sus correspondientes cliques.

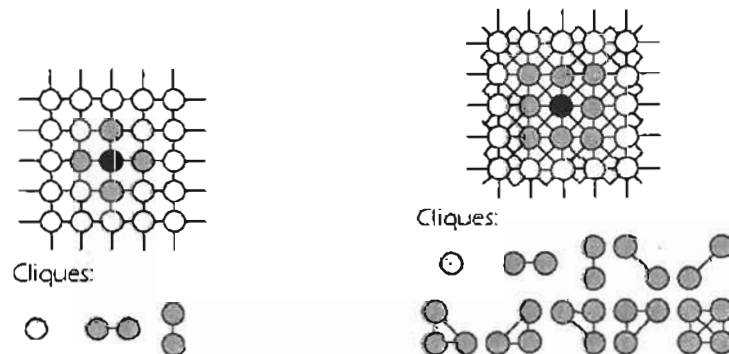


Figura 2.3 Sistemas de vecindad de primer (izquierda) y segundo orden (derecha) y sus correspondientes cliques.

Los sistemas de vecindad de un orden mayor a 2 son raramente utilizados ya que el calcular la función de energía puede ser complicado, ya que conlleva mucho tiempo de cómputo, además la elección depende de la resolución de la imagen y las regiones que se pretendan reconocer.

### 2.3.4 Modelo para la segmentación de imágenes

Los CAM son a priori muy apropiados para la segmentación ya que especifican claramente las propiedades locales de las regiones.

Sea  $R = \{r_1, r_2, \dots, r_M\}$  un conjunto de sitios y  $F = \{F_r : r \in R\}$  el conjunto de pixeles de una imagen (u observaciones) sobre éstos sitios. El conjunto de todas las posibles observaciones  $f = \{f_{r_1}, f_{r_2}, \dots, f_{r_M}\}$  es denotado por  $\Phi$ . Además, dado otro conjunto de sitios  $S = \{s_1, s_2, \dots, s_N\}$  donde cada uno de estos sitios puede ser etiquetado en  $\Delta = \{0, 1, 2, \dots, L-1\}$ . El espacio de configuración  $\Omega$  es el conjunto de etiquetamiento discreto global  $\omega = (\omega_{s_1}, \omega_{s_2}, \dots, \omega_{s_N})$ , donde  $\omega_s \in \Delta$ . Los dos conjuntos de sitios  $S$  y  $R$  no necesariamente son disjuntos, pueden tener partes comunes ó estar referidos a un conjunto común de sitios.

La meta es modelizar las etiquetas y las observaciones con un campo aleatorio  $(\chi, F) \in \Omega \times \Phi$ . El campo  $\chi = \{\chi_s\}_{s \in S}$  es llamado el campo de etiquetado y  $F = \{F_r\}_{r \in R}$  es el llamado campo de observación.

El problema general de etiquetado es hallar el etiquetamiento  $\hat{\omega}$  que maximice la probabilidad a posteriori  $P(\omega/F)$ , este método es llamado el máximo a posteriori MAP.

A partir del teorema de Bayes, dado por:

$$P(\omega/F) = \frac{1}{P(F)} P(F/\omega)P(\omega) \quad (2.26)$$

Podemos observar que  $P(F)$  no depende del etiquetamiento  $\omega$  y haciendo la suposición que son independientes entonces

$$P(F/\omega) = \prod_{s \in S} P(f_s/\omega_s) \quad (2.27)$$

y considerando que  $P(\omega) = \pi(\omega)$  dado por la ecuación (2.20), se tiene que el etiquetamiento

$$\hat{\omega} = \arg \max_{\omega \in \Omega} \prod_{s \in S} P(f_s/\omega_s) \prod_{C \in \mathcal{C}} e^{-v_C(\omega_C)} \quad (2.28)$$

Si suponemos que  $P(f_s/\omega_s)$  es una función de distribución gaussiana multivariante, donde la etiqueta (clase)  $\lambda \in \Delta$ , es representada por la media  $\mu_\lambda$  y varianza  $\sigma_\lambda$ , y para el potencial  $v_C(\omega_C)$ , según el modelo por interacción de parejas se tiene que



$$v_c(\omega_c) = V_{(s,r)}(\omega_s, \omega_r) = \begin{cases} -\beta & \text{Si } \omega_s = \omega_r \\ 0 & \text{Si } \omega_s \neq \omega_r \end{cases} \quad (2.29)$$

Entonces si conocemos las probabilidades  $P(F/\omega)$  y  $P(\omega)$  definidas por la imagen (a partir de una pre-segmentación) podemos calcular  $\hat{\omega}$  que hace máxima la probabilidad a posteriori. Sin embargo, el número de configuraciones posibles para un conjunto de  $L$  etiquetas es  $L^N$ . Resulta evidente que incluso para imágenes pequeñas y para un número reducido de etiquetas no es posible calcular el valor óptimo.

Dado el gran esfuerzo computacional que acarrea el cálculo del MAP se han desarrollado algoritmos que intentan estimar el MAP. Uno de ellos es el recocido simulado (simulated annealing) que tiene por objetivo encontrar estimadores del MAP para todos los píxeles simultáneamente, por lo que es muy costoso computacionalmente. Este algoritmo está inspirado en la Termodinámica, concretamente en el fenómeno físico-químico del enfriamiento lento de las partículas que conduce a un estado de equilibrio. Este algoritmo fue utilizado por primera vez en la segmentación de imágenes por Geman y Geman [Ge84].

Otro algoritmo más atractivo, desde el punto de vista computacional, es el algoritmo Iterated Conditional Modes (ICM) que modela el proceso aleatorio espacial de etiquetado como un CAM

#### 2.3.4.1 Algoritmo ICM

El algoritmo ICM fue propuesto por Besag [Be86], para estimar las etiquetas utilizando un CAM como modelo a priori y que salva la tendencia de un CAM a degenerar en una imagen plana (de un solo color).

El algoritmo está basado en las siguientes hipótesis:

1. Las observaciones  $f_s$  son condicionalmente independientes con igual probabilidad  $P(f_s / \omega_s)$  que depende tan sólo de las  $\omega_s$ .
2. El verdadero valor de las etiquetas  $\hat{\omega}$  es una realización de un CAM con probabilidad  $P(\omega)$

El algoritmo para estimar las etiquetas por el método ICM necesita un vector de observaciones  $F$  y produce el vector de etiquetas  $\hat{\omega}$ , está dado por 4 pasos:

1. Elegir un modelo de CAM para las etiquetas  $\omega$ .

Besag sugiere utilizar un modelo de interacción por parejas de un solo parámetro. El uso de un sistema de segundo orden permite que los ocho vecinos de un pixel influyan en la asignación de la etiqueta para ese pixel. El valor de  $\beta$  es el único parámetro del algoritmo. Cuando mayor sea éste, mayor será la influencia de los pixeles vecinos y menor será el detalle. Besag sugiere utilizar el valor 1.5, mientras que otros autores sugieren tomar valor entre  $0.86 \leq \beta \leq 2.3$  y además es conveniente que varíe en cada iteración.

2. Inicializar  $\hat{\omega}$  eligiendo, para cada pixel, la etiqueta  $\omega_s$  que maximiza  $P(f_s / \omega_s)$ .

Escoger una segmentación previa asegura la convergencia más rápida del algoritmo, aunque ésta no es necesaria.

3. Para cada  $s$  se actualiza el valor  $\hat{\omega}$  por el  $\omega$  que maximiza  $P(\omega / F)$ .

$$\hat{\omega} = \arg \max_{\omega \in \Omega} \prod_{s \in S} P(f_s / \omega_s) \prod_{C \in \mathcal{C}} e^{-\nu_C(\omega_C)}$$

En este trabajo se considera  $P(f_s / \omega_s)$  como una probabilidad normal multivariable y  $P(\omega)$  viene determinada por el modelo escogido.

4. Repetir el paso anterior un número determinado de iteraciones  $n_{it}$ .

## 2.4 Evaluación de la segmentación

Un modelo de segmentación espectral es mejor entre mayor separabilidad espectral genere, la separabilidad espectral esta relacionada con:

1. La relación entre la señal espectral y el ruido.
2. La correlación entre las bandas.
3. La resolución espectral y espacial.
4. Definición y agudeza de los bordes.

La correlación entre las bandas y el contenido del ruido pueden ser prácticamente eliminados empleando la descomposición en componentes principales, mejorando de ésta manera la separabilidad espectral [Li02].

#### 2.4.1. Separabilidad de regiones

La separabilidad entre las etiquetas no depende únicamente de la distancia entre las medias de las etiquetas, sino también de las varianzas. Para una distancia  $|\mu_i - \mu_j|$  dada, las etiquetas pueden estar separadas para  $\sigma_i$  y  $\sigma_j$  suficientemente pequeñas; e inversamente, para  $\sigma_i$  y  $\sigma_j$  suficientemente grandes, las etiquetas pueden traslaparse aún para una diferencia  $|\mu_i - \mu_j|$  grande [De82] [Sw78].

Una cantidad que involucra la distancia entre las etiquetas y la varianza son las probabilidades  $P(F / \omega_i)$  y  $P(F / \omega_j)$  para las etiquetas  $\omega_i$  y  $\omega_j$ , para un valor espectral  $f$ , el cociente de estas da un parámetro para medir la separabilidad entre las etiquetas.

$$L_{ij}(f) = \frac{P(F / \omega_i)}{P(F / \omega_j)} \quad (2.30)$$

Por simplicidad se suele tomar el logaritmo de  $L_{ij}(F)$ , denotado por  $L_{ij}(F)'$ .

$$L_{ij}(F)' = \ln L_{ij}(F) = \ln (P(F / \omega_i)) - \ln (P(F / \omega_j)) \quad (2.31)$$

La cantidad  $L_{ij}(F)$  tiende a cero o a infinito cuando la separabilidad entre las etiquetas aumenta. Para todos los valores espectrales  $F$  se tiene una colección de cocientes  $L_{ij}(f)$ , por lo que un promedio ponderado de éstos proporciona una medida global de la separabilidad espectral de las etiquetas involucradas.

A partir de la ecuación (2.31) se define la divergencia como:

$$d_{ij} = E[L_{ij}(F)' / \omega_i] - E[L_{ji}(F)' / \omega_j] \quad (2.32)$$

donde

$$E[L_{ij}(F)' / \omega_i] = \int L_{ij}(F)' P(F / \omega_i) dF \quad (2.33)$$

por lo tanto,

$$d_{ij} = E[L_{ij}(F)' / \omega_i] - E[L_{ji}(F)' / \omega_j] = \int [P(F / \omega_i) - P(F / \omega_j)] \ln \left[ \frac{P(F / \omega_i)}{P(F / \omega_j)} \right] dF \quad (2.34)$$

A partir de las expresiones anteriores, se deducen las siguientes propiedades de la divergencia.

1. Positividad  $d_{ij} \geq 0$

Para cualesquiera dos etiquetas diferentes, la divergencia es mayor que cero.

2. Simetría  $d_{ij} = d_{ji}$

La divergencia es una medida simétrica.

3. Si  $P(F / \omega_i) = P(F / \omega_j)$ , para todo  $f$ , entonces  $d_{ij} = d_{ji} = 0$ .

4. Si  $F = (f^1, f^2, \dots, f^d)$  son estadísticamente independientes,  $P(F / \omega_i) = \prod_{k=1}^d P(f^k / \omega_i)$  donde  $d$  es el número de etiquetas, entonces,

$$d_{ij}(F) = \sum_{k=1}^d d_{ij}(f^k)$$

5.  $d_{ij}(f^1, f^2, \dots, f^d, f^{d+1}) \geq d_{ij}(f^1, f^2, \dots, f^d)$

La divergencia no disminuye con la adición de nuevas variables. Si las variables son estadísticamente independientes la separabilidad conjunta es la suma de las separabilidades que consideran sólo una variable.

El cálculo de la divergencia se simplifica si consideramos distribuciones de probabilidad normales, en este caso se simplifica a

$$d_{ij} = \frac{1}{2} \text{tr}[(\sigma_i^2 - \sigma_j^2)((\sigma_i^2)^{-1} - (\sigma_j^2)^{-1})] + \frac{1}{2} \text{tr}[(\sigma_i^2)^{-1} - (\sigma_j^2)^{-1}](\mu_i - \mu_j)(\mu_i - \mu_j)^T \quad (2.35)$$

donde  $\text{tr}$  es la traza de matriz correspondiente.

Al realizarse el cálculo para todos los pares de etiquetas, se obtiene una matriz, conocida como matriz de la divergencia, con las propiedades de simetría. Segmentando imágenes sintéticas, con un conjunto de distribuciones normales y aumentando la distancia espectral entre ellas, se obtiene una serie de resultados que muestran que la probabilidad de segmentar un pixel aumenta de manera exponencial al incrementar la distancia espectral.

La divergencia sin embargo muestra un comportamiento logarítmico, para subsanar esto es necesario modificar la divergencia para que tenga un comportamiento exponencial asintótico, a tal modificación se le conoce como la divergencia transformada, cuya expresión es la siguiente:

$$d_{ij}^T = 2\{1 - e^{-d_{ij}/B}\} \quad (2.36)$$

Por último, para obtener una medida de separabilidad entre K etiquetas conjuntamente, es necesario definir una medida de separabilidad global. Para ello se define la divergencia media entre K etiquetas, como:

$$\bar{d} = \sum_{i=1}^K \sum_{j=1}^K P(\omega_i)P(\omega_j)d_{ij} \quad (2.37)$$

y divergencia media transformada entre K etiquetas como:

$$\bar{d}^T = \sum_{i=1}^K \sum_{j=1}^K P(\omega_i)P(\omega_j)d_{ij}^T \quad (2.37)$$

donde  $P(\omega_i)$  es la probabilidad de que los pixeles de la imagen pertenezcan a la clase  $\omega_i$ .

Otra mejor caracterización de la separabilidad de etiquetas es la distancia de Jeffries-Matusita, también llamada Bhattacharyya [Ri99], que se define como:

$$J_{ij} = \int \left[ \sqrt{P(F/\omega_i)} - \sqrt{P(F/\omega_j)} \right]^2 dF \quad (2.38)$$

Esta expresión cuantifica la suma de las distancias al cuadrado entre las distribuciones de probabilidad para un par de etiquetas. De acuerdo, con esta definición, es posible afirmar que la divergencia es básicamente una suma de cocientes mientras que la distancia de Jeffries-Matusita es una suma de las distancias al cuadrado. Los cocientes tienen la particularidad de presentar valores altos cuando el denominador es pequeño, de aquí el comportamiento incorrecto de la divergencia en la caracterización de los resultados experimentales.

Si procedemos de la misma manera al considerar distribuciones normales, la distancia de Jeffries-Matusita se reduce a:

$$J_{ij} = 2\{1 - e^{-\kappa}\} \quad (2.39)$$

donde

$$\kappa = \frac{1}{8}(\mu_i - \mu_j)^t \left[ \frac{\sigma_i^2 + \sigma_j^2}{2} \right]^{-1} (\mu_i - \mu_j) + \frac{1}{2} \ln \left[ \frac{\sqrt{|\sigma_i^2 + \sigma_j^2|}}{\sqrt{|\sigma_i^2|} \sqrt{|\sigma_j^2|}} \right] \quad (2.40)$$

La distancia de Jeffries-Matusita media entre K etiquetas se define como

$$\bar{J} = \sum_{i=1}^K \sum_{j=1}^K P(\omega_i)P(\omega_j)J_{ij} \quad (2.41)$$

El intervalo de variación de la distancia de Jeffries-Matusita es de cero a dos, si el valor de  $J_{ij}$  tiende a dos entre dos etiquetas significa que esta segmentación es perfecta.

#### 2.4.2. Matriz de confusión

Esta es una matriz de igual dimensión al número de etiquetas de una segmentación dada. Para  $i \neq j$ , el elemento  $c_{ij}$  de esta matriz indica la proporción de pixeles clasificados de la clase  $i$  que fueron erróneamente clasificados en la clase  $j$ . Para  $i=1$ , el elemento de la matriz muestra la proporción de pixeles clasificados en la clase  $i$  con respecto al número de pixeles de dicha clase. La suma de los elementos a la largo de cada renglón produce el número total de pixeles de cada clase. En una segmentación perfecta, únicamente los elementos de la diagonal seran diferentes de cero e iguales al número de pixeles de la clase, es decir, se obtendría la matriz identidad.

A partir de la matriz de confusión es posible evaluar el coeficiente kappa ( $\hat{\kappa}$ ), indicador global de la calidad de la segmentación. El intervalo de variación de este coeficiente es  $-1 \leq \hat{\kappa} \leq 1$ , donde una segmentación perfecta daría un valor de 1 y una total errónea -1.

Este se define como:

$$\hat{\kappa} = \frac{d - q}{N - q} \quad (2.42)$$

donde  $d$  es la suma de los elementos de la diagonal en la matriz de confusión,  $N$  es la suma de todos los elementos de la matriz y  $q$  está dado por:

$$q = \frac{\sum_{i=1}^n \sum_{j=1}^n c_{ij}}{N} \quad (2.43)$$

donde  $n$  es el número de etiquetas.

Una manera de obtener la matriz de confusión es a partir de una segmentación supervisada [Co89].

# CAPÍTULO 3

## APLICACIÓN DE LOS ALGORITMOS

En este capítulo se presenta los resultados de los algoritmos de componentes principales, K-medias, ICM y los resultados de los descriptores de evaluación.

### 3.1 Componentes principales

El algoritmo de componentes principales mencionado en el capítulo anterior, se implementó en MatLab donde se cuenta con las rutinas para calcular los autovectores, simplificando considerablemente la programación. El código se muestra en el Apéndice 2.

El algoritmo de componentes principales se aplicó a la imagen de la Lena a color (Figura 3.1) los resultados se muestran en la Figura 3.2, donde se puede observar cualitativamente el aumento de contraste y una mejor definición de los objetos de la imagen en la primera componente principal.

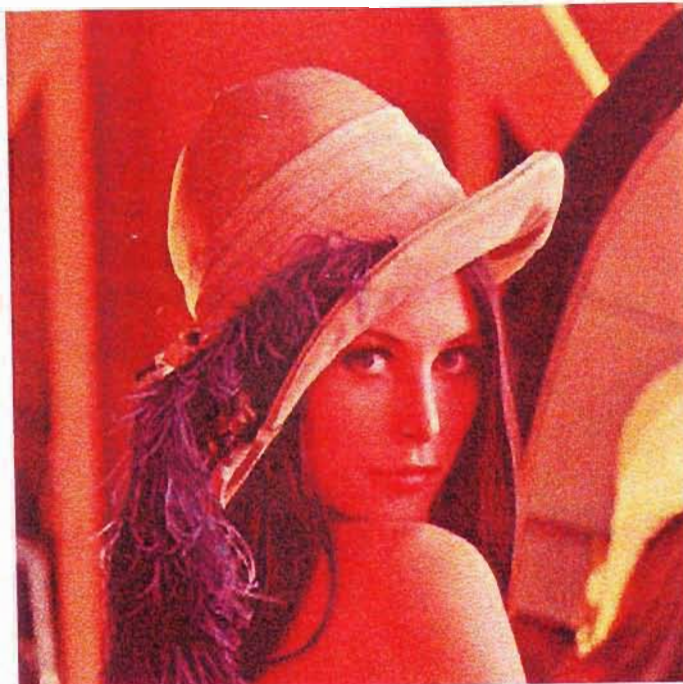


Figura 3.1. Imagen de la Lena a color







Bandas (RGB) de la imagen	Componentes de la Imagen
 <p data-bbox="489 700 551 733">Rojo</p>	 <p data-bbox="890 700 1121 733">Primera (90.89%)</p>
 <p data-bbox="482 1193 559 1226">Verde</p>	 <p data-bbox="890 1193 1121 1226">Segunda (8.24%)</p>
 <p data-bbox="489 1686 551 1719">Azul</p>	 <p data-bbox="897 1686 1113 1719">Tercera (0.87%)</p>

Figura 3.2 Imagen de la Lena a color descompuesta en sus componentes RGB y componentes principales donde se muestra la contribución a la varianza global de cada componente.



Los valores de las medias y la desviación estándar de la imagen a color y en componentes principales, la matriz de transformación lineal y la matriz de contribución de cada componente a la varianza global se muestra en la Tabla 3.1.

Lena RGB	
$\mu$ :	$\sigma^2$ :
(181.21, 100.05, 106.39)	$\begin{pmatrix} 244.03 & 230.44 & 119.39 \\ 230.44 & 276.94 & 165.90 \\ 119.39 & 165.90 & 117.84 \end{pmatrix}$
Matriz de transformación lineal	
w:	$\begin{pmatrix} -0.6101 & -0.7206 & 0.3295 \\ -0.6827 & 0.2670 & -0.6802 \\ -0.4021 & 0.6399 & 0.6548 \end{pmatrix}$
Lena Componentes Principales	
$\mu$	$\sigma^2$
(221.64, 35.79, 61.32)	$\begin{pmatrix} 580.59 & 0.01 & 0.01 \\ 0.01 & 52.6225 & -0.01 \\ 0.01 & -0.01 & 5.5934 \end{pmatrix}$
Matriz de contribución a la varianza global	
i:	$\begin{pmatrix} 0.9089 & 0 & 0 \\ 0 & 0.0824 & 0 \\ 0 & 0 & 0.0087 \end{pmatrix}$

Tabla 3.1 Valores de la media y la matriz de covarianza de la imagen Lena en RGB y en componentes principales, la matriz de la transformación lineal y la matriz de contribución a la varianza global.

Los resultados para a imagen de Pátzcuaro multiespectral de 6 bandas (Landsat) se muestran en la Figura 3.3.

En ésta podemos observar que hay una mayor correlación entre las bandas por lo que las bandas 1, 2 y 3 contribuyen considerablemente.

Los valores de las medias y la desviación estándar de la imagen a color y en componentes principales, la matriz de transformación lineal y la matriz de contribución de cada componente a la varianza global se muestra en la Tabla 3.2.

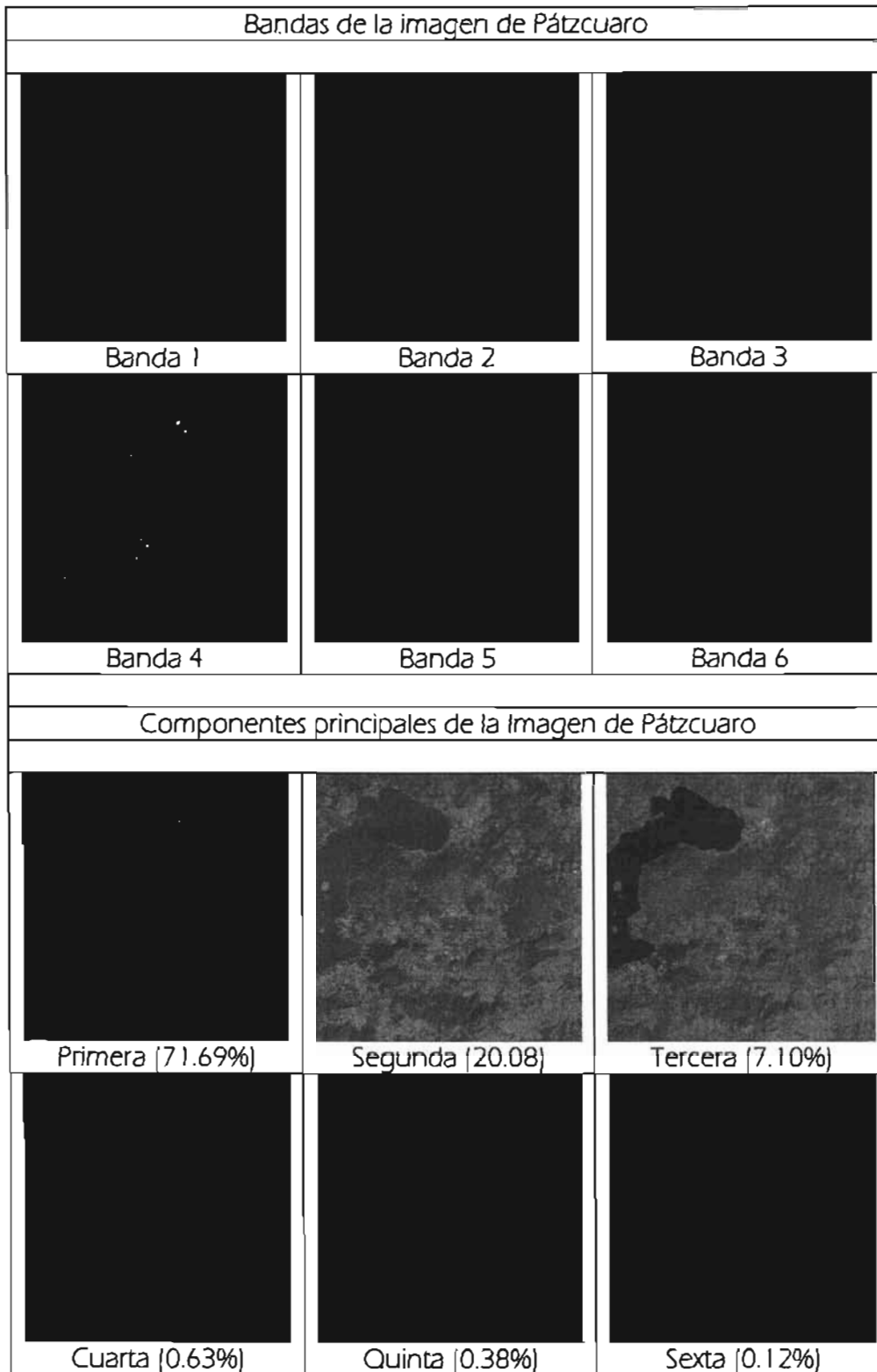


Figura 3.3 Imagen de Pátzcuaro y sus componentes principales donde se muestra la contribución a la varianza global de cada componente.

Pátzcuaro 6 bandas						
$\mu$ :	(65.50, 62.33, 71.65, 81.47, 77.19, 60.20)					
$\sigma^2$ :	$\begin{pmatrix} 102.62 & 127.26 & 135.58 & 63.41 & 39.56 & 45.19 \\ 127.25 & 171.20 & 187.22 & 57.64 & 43.22 & 58.36 \\ 135.58 & 187.22 & 251.48 & 126.85 & 181.97 & 179.51 \\ 63.41 & 57.64 & 126.85 & 440.08 & 436.23 & 317.24 \\ 39.55 & 43.22 & 181.97 & 436.23 & 672.40 & 522.15 \\ 45.19 & 58.36 & 179.51 & 317.24 & 522.15 & 432.47 \end{pmatrix}$					
Matriz de transformación lineal						
$W$ :	$\begin{pmatrix} 0.55 & -0.44 & -0.51 & -0.15 & -0.44 & 0.09 \\ -0.75 & -0.02 & -0.25 & -0.05 & -0.60 & 0.11 \\ 0.33 & 0.42 & 0.50 & 0.19 & -0.60 & 0.25 \\ -0.03 & -0.02 & 0.23 & -0.84 & 0.10 & 0.47 \\ 0.04 & 0.44 & -0.51 & 0.24 & 0.25 & 0.65 \\ -0.12 & -0.65 & 0.33 & 0.41 & 0.09 & 0.51 \end{pmatrix}$					
Pátzcuaro componentes principales						
$\mu$ :	(150.64, 76.98, 25.14, 13.44, 6.80, 7.07)					
$\sigma^2$ :	$\begin{pmatrix} 148.42 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 41.57 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 14.69 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 1.30 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.79 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.25 \end{pmatrix}$					
Matriz de contribución a la varianza global						
$l$ :	$\begin{pmatrix} 0.717 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.201 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.071 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.006 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.004 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.001 \end{pmatrix}$					

Tabla 3.2 Valores de la media y la matriz de covarianza de las bandas de la imagen Pátzcuaro y de las componentes principales, la matriz de la transformación lineal y la matriz de contribución a la varianza global.

### 3.2 K-medias

El algoritmo K-medias se implementó utilizando una variante del algoritmo expuesto en la sección 2.2 del capítulo anterior. Esta variante consiste en utilizar el histograma de la imagen para hacer la asignación al cúmulo que le correspondía a los píxeles que tienen el mismo valor espectral y no asignar a cada uno al leerlos de imagen, esto es posible dado que algoritmo K-medias sólo utiliza la información espectrales de los píxeles y por consecuente la información espacial es despreciable. De esta manera se redujo el tiempo de cómputo del algoritmo de K-medias, el código se muestra en apéndice 2.

El algoritmo de K-medias tiene como entrada la imagen y número de etiquetas. El número etiquetas que equivale al número de regiones a particionar la imagen se determinó de manera cualitativa y experimental, ya que es difícil de determinar ver Figura 3.4 y Figura 3.5, donde se selecciona una ventana de la imagen de Pátzcuaro y se realiza la segmentación bajo K-medias para un número de etiquetas entre 2 y 15, eligiendo 10 etiquetas ya que presentaba regiones homogéneas bien definidas.

El algoritmo de K-medias se aplicó a imágenes de prueba en niveles de gris, a color y multibanda, los resultados se muestran en las figuras 3.6.A, 3.6.B, 3.6.C y 3.6.D.

### 3.3 Segmentación ICM

Para la aplicación del algoritmo de segmentación ICM, se consideró la segmentación K-medias de las imágenes a partir de la cuál se pudo definir las probabilidades a priori (ecuación (2.20))

$$P(F / \omega) = \prod_{s \in S} P(f_s / \omega_s)$$

donde se consideró una distribución gaussiana multivariante, calculada a partir de las medias y la matriz de covarianza de las Etiquetas de la segmentación de K-medias. El valor del parámetro  $\beta$  varió de linealmente entre 0.86 y 2.3 en las 6 iteraciones que realiza el algoritmo y se utilizó un sistema de vecindad de conectividad 8.

El código del algoritmo empleado se muestra en el Apéndice 2.

Los resultados obtenidos se muestran en las figuras 3.6.A, 3.6.B, 3.6.C y 3.6.D. junto con los resultados de K-medias para realizar una comparación cualitativa.

Se puede observar una mejora de la segmentación de los objetos que conforman la escena en la imagen.

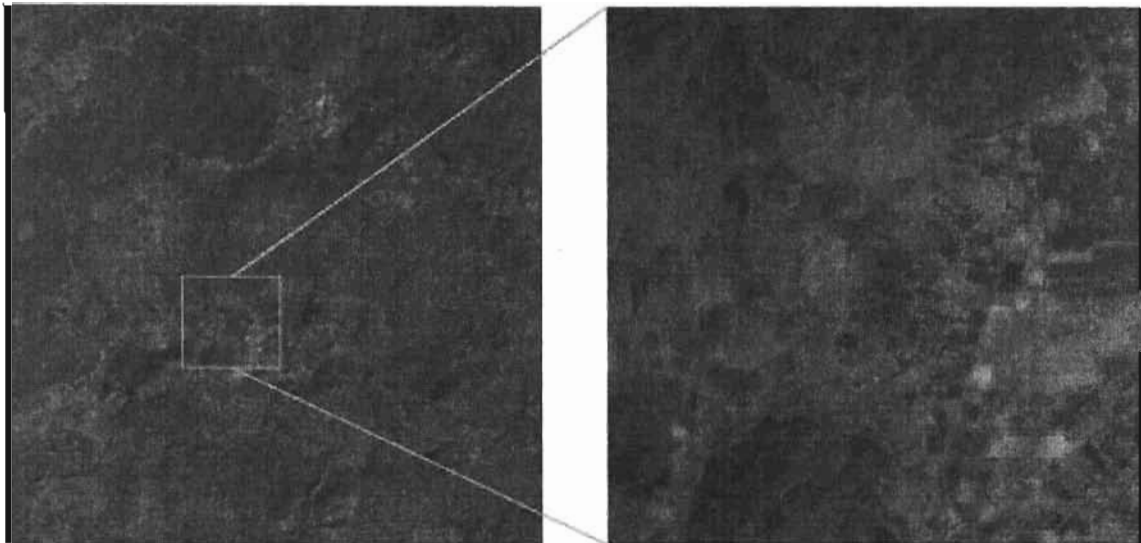


Figura 3.4. Ventana seleccionada sobre la imagen de Pátzcuaro (primera componente principal).

### 3.4. Evaluación

Los resultados de la evaluación utilizando las matrices de divergencia, Jeffries-Matusita y de confusión se muestran en las tablas 3.3 y 3.4 para las imágenes que se muestran.

Los descriptores de la matriz de divergencia y Jeffries-Matusita no podemos utilizarlos dado que dependen directamente de los valores de la media y la matriz de covarianza y resultan inadecuados para realizar la evaluación, ya que sus valores están en relación inversa con el error de segmentación.

Podemos ver que los valores medios de estos descriptores siempre son mayores para la segmentación K-medias que para ICM, aunque cualitativamente se muestre lo contrario (Tabla 3.3 y 3.4), por consiguiente no se utilizarán en la evaluación de los resultados posteriores.

Un buen descriptor es la matriz de confusión y el valor del coeficiente  $\hat{\kappa}$  para realizar la evaluación de los resultados según se muestra en las tablas 3.3 y 3.4, pero hay que señalar que se necesita de una segmentación supervisada para su obtención.

Segmentación de la ventana de la imagen de Pátzcuaro

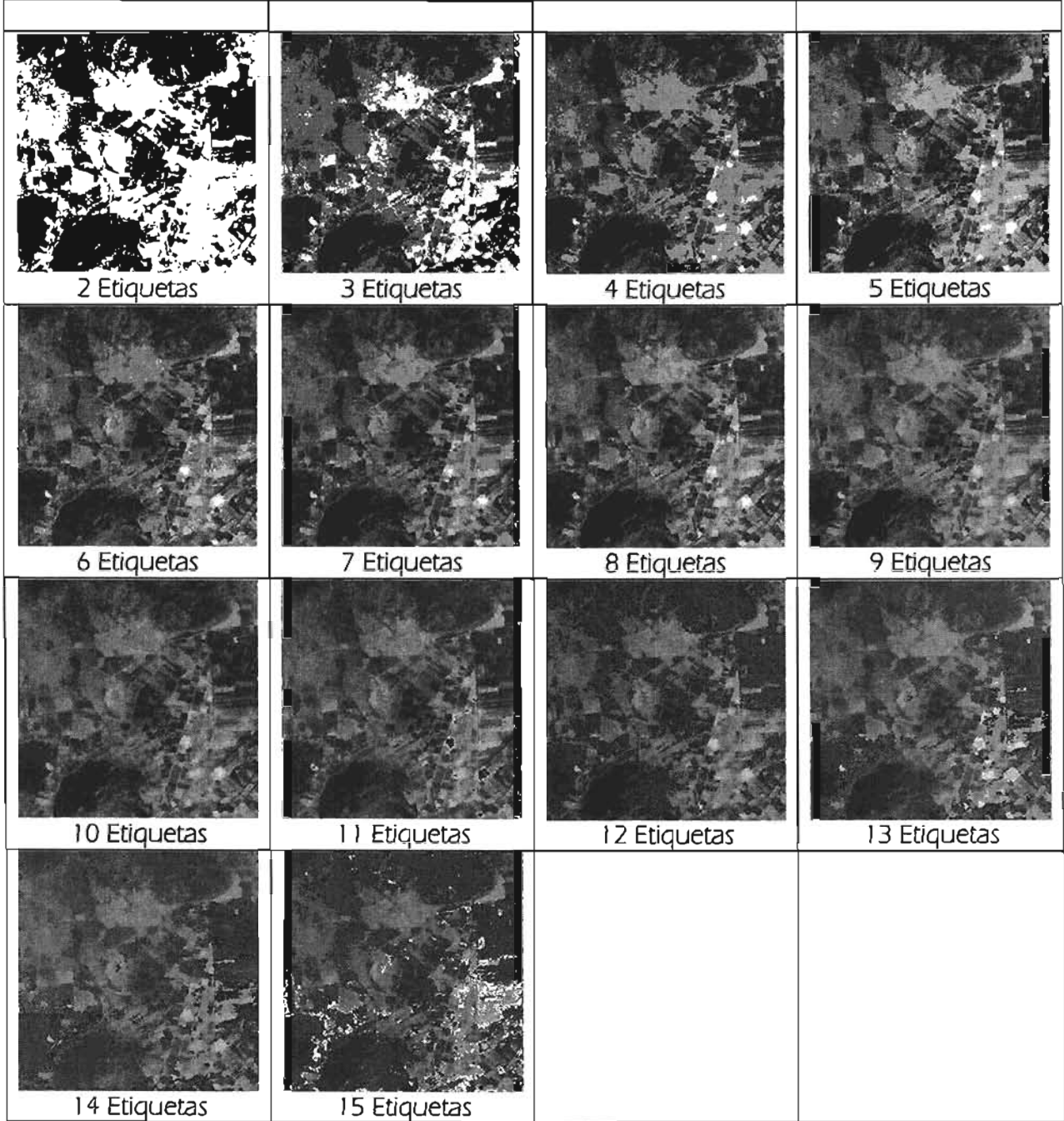


Figura 3.5 Segmentación de una ventana de la imagen de Pátzcuaro.

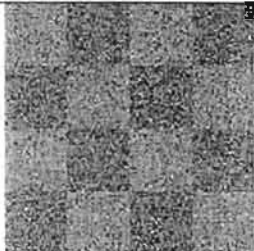
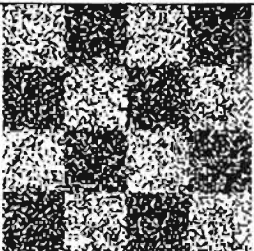
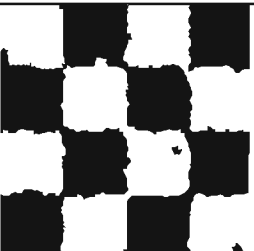




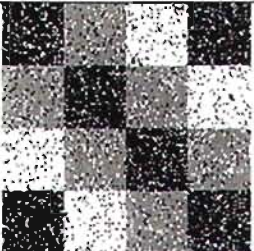
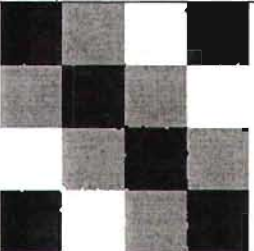






Imagen	Segmentación	
	K-medias	ICM
 Ruido (niveles de gris)	 2 Etiquetas	 $\beta=1..10$
 Paris (niveles de gris)	 5 Etiquetas	 $\beta=0.86..2.3$
 Ruido (RGB)	 3 Etiquetas	 $\beta=0.86..2.3$
 Lena (RGB)	 5 Etiquetas	 $\beta=0.02..0.12$
 Verduras (RGB)	 5 Etiquetas	 $\beta=0.86..2.3$

Figura 3.6.A Resultados de la segmentación bajo K-medias e ICM, se muestra el número de etiquetas y el valor del parámetro  $\beta$  en la segmentación ICM.



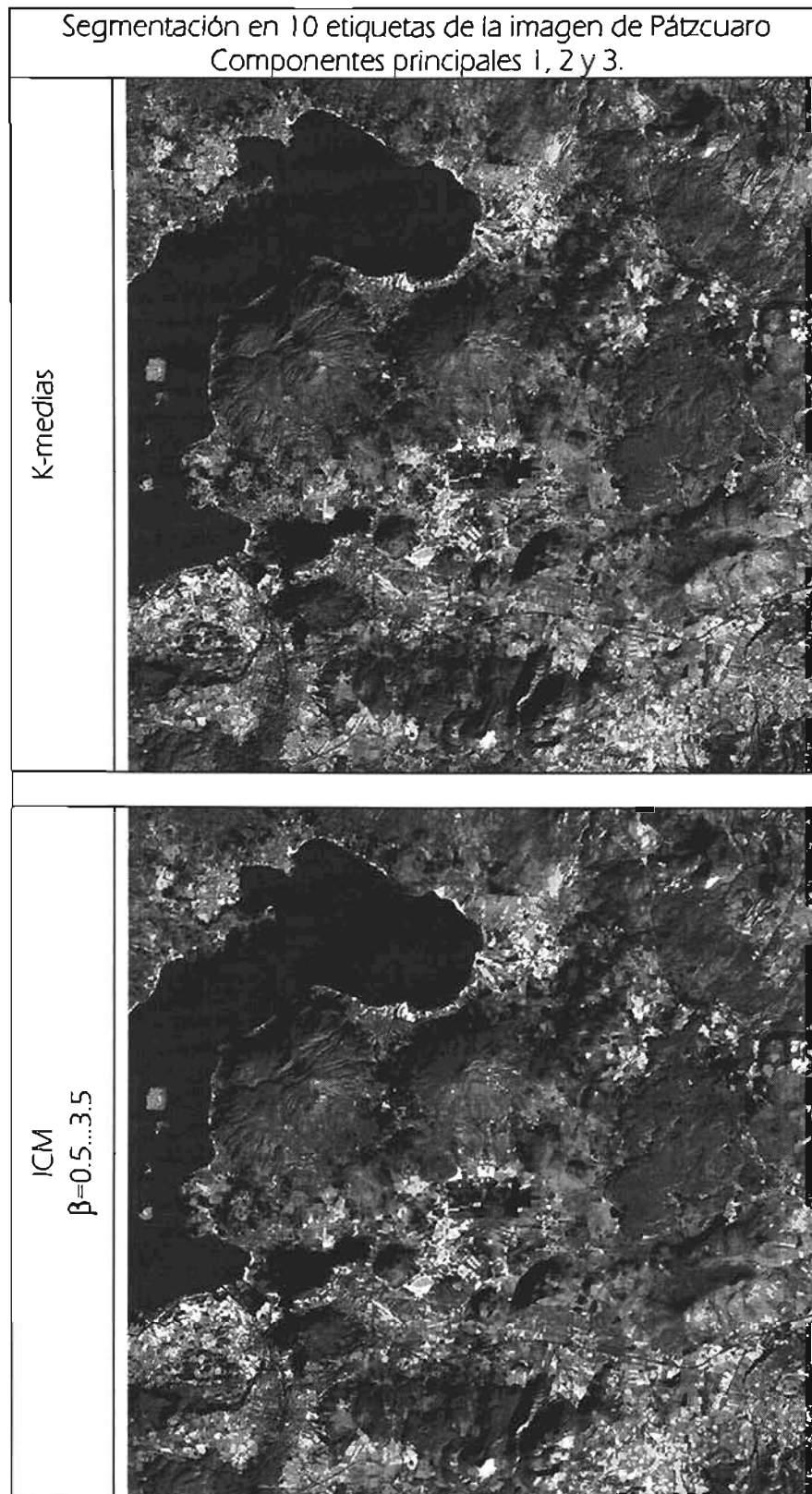


Figura 3.6.B Resultados de la segmentación bajo K-medias e ICM, se muestra el número de etiquetas y el valor del parámetro  $\beta$  en la segmentación ICM.



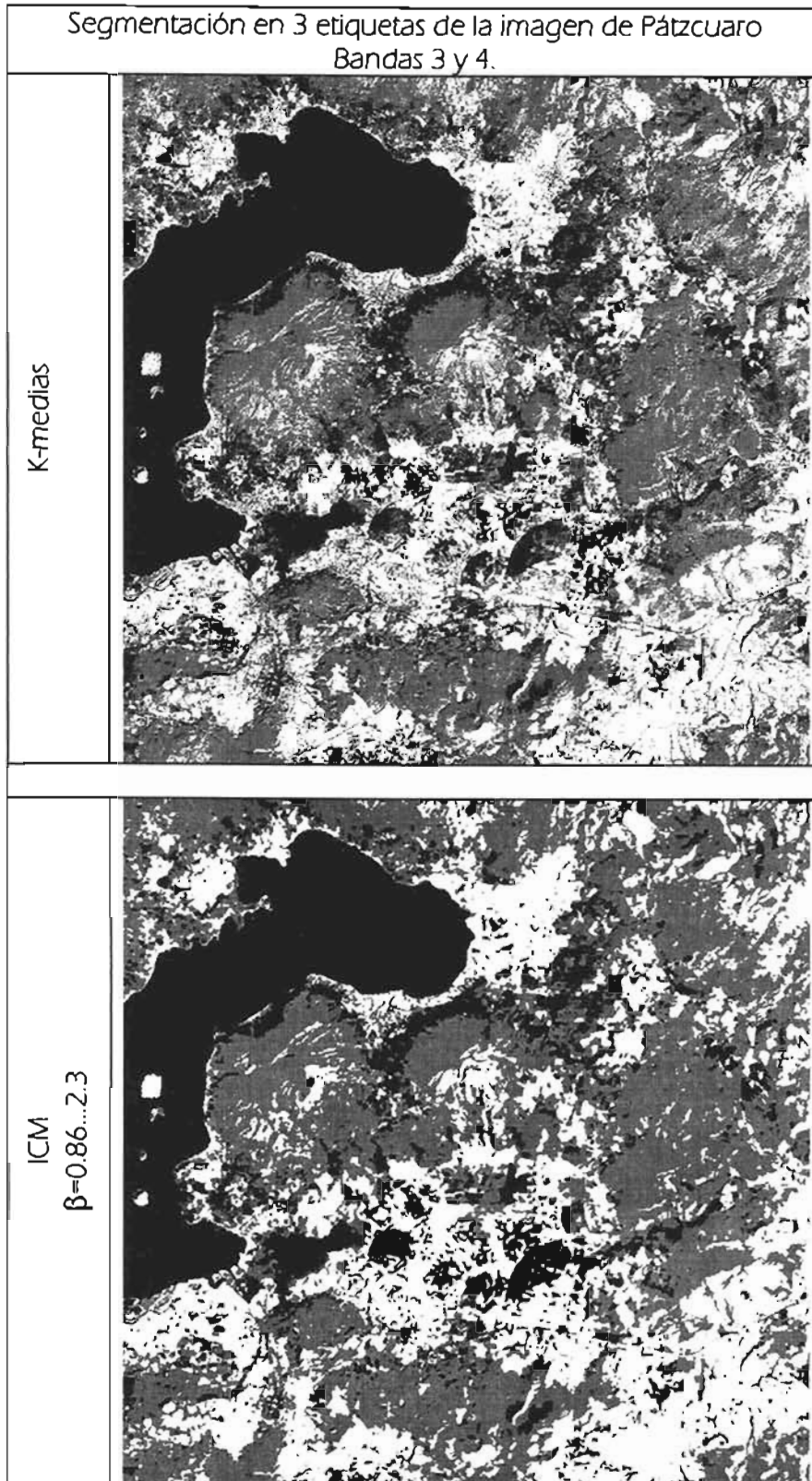


Figura 3.6 C Resultados de la segmentación bajo K-medias e ICM, se muestra el número de etiquetas y el valor del parámetro  $\beta$  en la segmentación ICM.

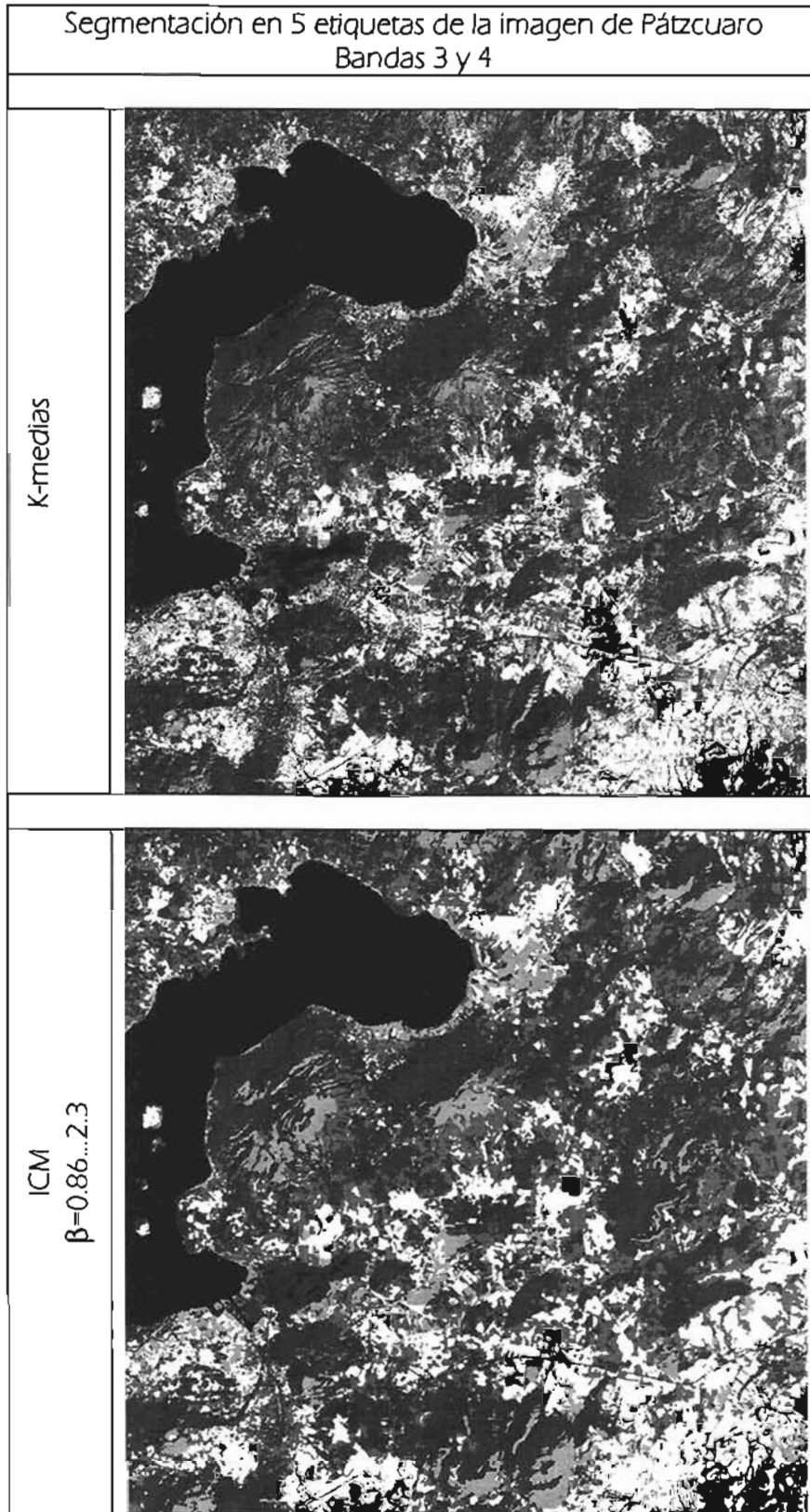


Figura 3.6.D Resultados de la segmentación bajo K-medias e ICM, se muestra el número de etiquetas y el valor del parámetro  $\beta$  en la segmentación ICM.

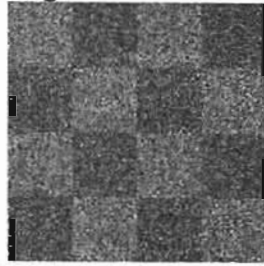
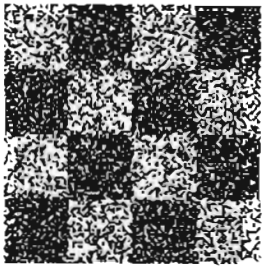

Descriptor	Segmentación	
	K-medias	ICM
Imagen:  Ruido (niveles en gris)		
Media y covarianza	$\mu = (95, 145)$ $\sigma^2 = (355, 341)$	$\mu = (103, 134)$ $\sigma^2 = (723, 742)$
Divergencia transformada	$\begin{pmatrix} 0.0 & 1.99 \\ 1.99 & 0.0 \end{pmatrix}$ $\bar{d} = 0.994$	$\begin{pmatrix} 0.0 & 1.48 \\ 1.48 & 0.0 \end{pmatrix}$ $\bar{d} = 0.737$
Jeffries-Matusita	$\begin{pmatrix} 0.0 & 1.19 \\ 1.19 & 0.0 \end{pmatrix}$ $\bar{J} = 0.589$	$\begin{pmatrix} 0.0 & 0.31 \\ 0.31 & 0.0 \end{pmatrix}$ $\bar{J} = 0.154$
Confusión	$\begin{pmatrix} 6149 & 2043 \\ 2646 & 5546 \end{pmatrix}$ $\hat{\kappa} = 0.428$	$\begin{pmatrix} 8056 & 136 \\ 459 & 7733 \end{pmatrix}$ $\hat{\kappa} = 0.927$

Tabla 3.3 Evaluación de la segmentación K-medias e ICM para imagen de ruido en niveles de gris, utilizando las matrices de divergencia transformada, Jeffries-Matusita y confusión.

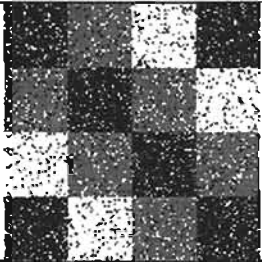
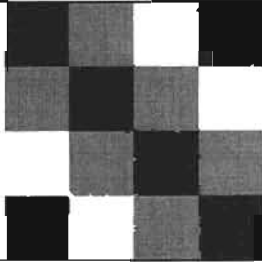
Descriptor	Segmentación	
	K-medias	ICM
Imagen: Ruido (RGB)		
Media y covarianza	$\mu = \begin{pmatrix} 211 & 39 & 44 \\ 54 & 43 & 220 \\ 58 & 217 & 55 \end{pmatrix}$ $\sigma^2 = \begin{pmatrix} 4514 & 740 & 705 \\ 740 & 3255 & 12 \\ 705 & 12 & 3634 \\ 5383 & 66 & 790 \\ 66 & 3606 & 553 \\ 790 & 553 & 3076 \\ 5880 & 881 & 93 \\ 881 & 3315 & 808 \\ 93 & 808 & 5520 \end{pmatrix}$	$\mu = \begin{pmatrix} 206 & 52 & 53 \\ 54 & 53 & 227 \\ 54 & 206 & 53 \end{pmatrix}$ $\sigma^2 = \begin{pmatrix} 5237 & 19 & 10 \\ 19 & 5295 & 71 \\ 10 & 71 & 5172 \\ 5433 & 59 & 111 \\ 59 & 5203 & 38 \\ 111 & 38 & 5208 \\ 5314 & 147 & 81 \\ 147 & 5179 & 0 \\ 81 & 0 & 5335 \end{pmatrix}$
Divergencia transformada	$\begin{pmatrix} 0.0 & 1.779 & 1.780 \\ 1.779 & 0.0 & 1.816 \\ 1.780 & 1.816 & 0.0 \end{pmatrix}$ $\bar{d}=1.187$	$\begin{pmatrix} 0.0 & 1.355 & 1.356 \\ 1.355 & 0.0 & 1.362 \\ 1.356 & 1.362 & 0.0 \end{pmatrix}$ $\bar{d}=0.891$
Jeffries-Matusita	$\begin{pmatrix} 0.004 & 1.330 & 1.332 \\ 1.330 & 0.004 & 1.340 \\ 1.332 & 1.340 & 0.003 \end{pmatrix}$ $\bar{J}=0.885$	$\begin{pmatrix} 0.003 & 1.164 & 1.164 \\ 1.164 & 0.003 & 1.167 \\ 1.164 & 1.167 & 0.003 \end{pmatrix}$ $\bar{J}=0.766$
Confusión	$\begin{pmatrix} 5075 & 511 & 558 \\ 441 & 5174 & 529 \\ 285 & 241 & 3570 \end{pmatrix}$ $\hat{\kappa}=0.763$	$\begin{pmatrix} 6120 & 18 & 6 \\ 10 & 6127 & 7 \\ 2 & 2 & 4092 \end{pmatrix}$ $\hat{\kappa}=0.996$

Tabla 3.4 Evaluación de la segmentación K-medias e ICM para imagen de ruido en RGB, utilizando las matrices de divergencia transformada, Jeffries-Matusita y confusión.

# CAPÍTULO 4

## ANÁLISIS DE LOS RESULTADOS

En este capítulo se presenta un análisis para la elección del valor del parámetro  $\beta$ , la dimensionalidad intrínseca de la imagen y por último sobre la disminución de error de segmentación de ICM con respecto a K-medias utilizando el coeficiente  $\hat{\kappa}$  de la matriz de confusión.

En este capítulo se utilizaron imágenes en RGB sintéticas a las cuales se les agregó ruido gaussiano en un 30%, éstas se seleccionaron porque se puede realizar su evaluación de manera más fácil y objetiva a través de la matriz de confusión ver Figura 4.2

### 4.1 Parámetro beta

El valor de  $\beta$  de la segmentación ICM se determinó por ensayo y error, el cuál varió entre un valor mínimo y máximo de manera lineal en 6 iteraciones.

Los resultados sobre la imagen de ruido en niveles de gris (ver Figura 4.1) se muestra en la tabla 4.2, donde se puede observar estos al utilizar un valor de  $\beta$  pequeño, uno grande, uno constante y variando este valor de manera lineal en un intervalo. Se muestra también el valor del coeficiente  $\hat{\kappa}$  que nos sirve de descriptor en la evaluación de los resultados.

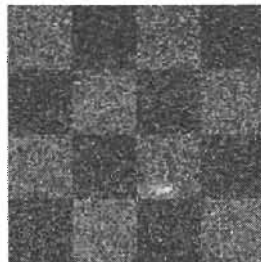


Figura 4.1 Imagen de ruido en niveles de gris.

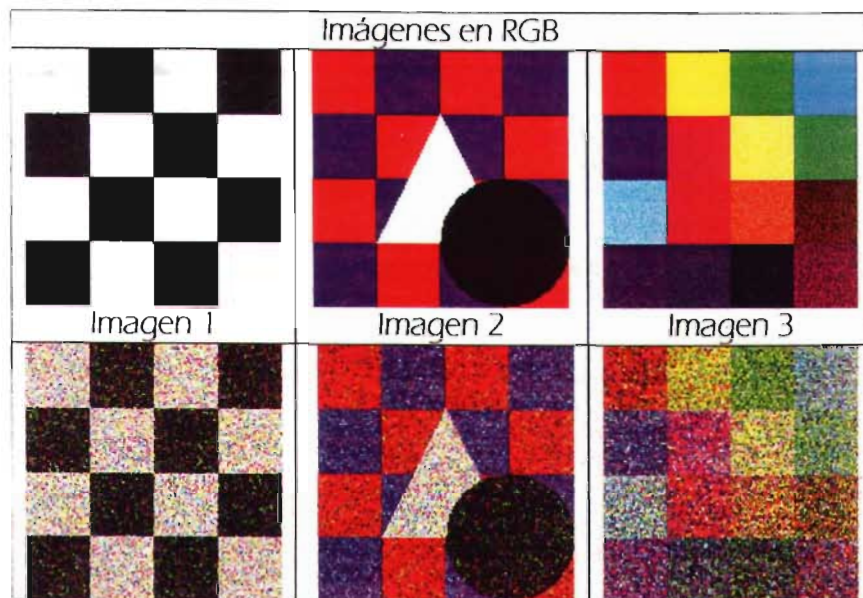


Figura 4.2 Imágenes en RGB sintéticas con ruido.

De los resultados obtenidos en la Tabla 4.1 se puede observar que el mejor de ellos, es elegir un valor de  $\beta$  que varíe de manera lineal en un rango ya que se tiene el valor del coeficiente  $\hat{\kappa}$  aumenta, lo que se traduce en una mejor segmentación.

#### 4.2 Dimensionalidad intrínseca

En esta sección se determina la dimensionalidad intrínseca de las imágenes de prueba utilizando la segmentación K-medias. Como se mencionó en el sección 2.1.1 la manera de calcular la dimensionalidad intrínseca de una imagen multiespectral de  $n$  bandas, es utilizando una metodología empírica que consiste en calcular su transformación en componentes principales, posteriormente se procede a segmentar los píxeles utilizando las  $n$  componentes principales, para después segmentarlos utilizando  $(n-1)$  (descartando la componente  $n$ ), y así sucesivamente. Al realizar esta prueba se evalúa de manera cualitativa y cuantitativa las diferencias, entonces, si la diferencia entre utilizar  $m$  y  $n$  componentes principales es mínima ó despreciable se dice que la dimensionalidad intrínseca es  $m$ .

La transformación en componentes principales de las imágenes de prueba se realizó obteniéndose los resultados mostrados en la Figura 4.3, en todos los casos la contribución de la primera componente es mayor.

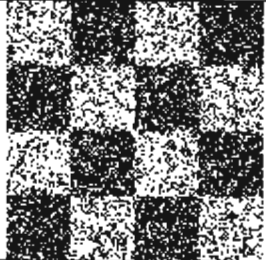
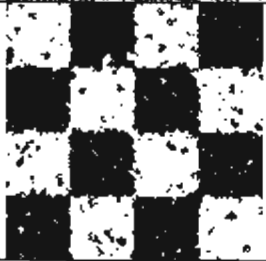
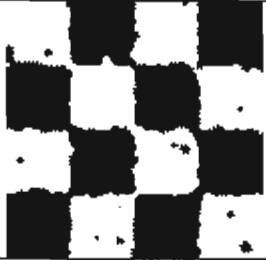
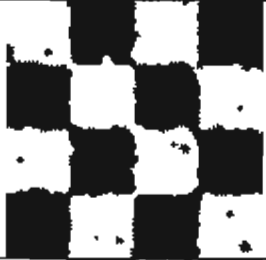
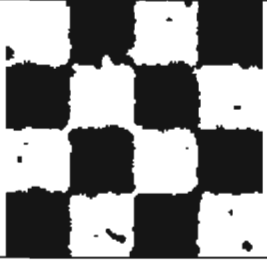
$\beta$	Segmentación ICM	$\hat{\kappa}$
0.1		0.4857
1		0.8734
10		0.9249
100		0.9248
1 a 10		0.9359

Tabla 4.1 Resultados de la segmentación ICM con diferentes valores de  $\beta$ .

Una vez obtenida la transformación en componentes principales se realizó la segmentación K-medias e ICM utilizando las tres componentes principales, después sólo dos primeras componentes principales y por último sólo la primera componente principal los resultados se muestran en las tablas 4.3 y 4.4, el valor del parámetro  $\beta$  varía entre 1 y 10.

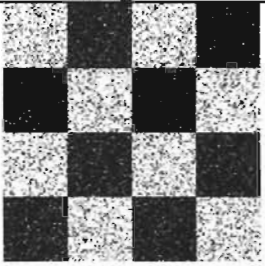
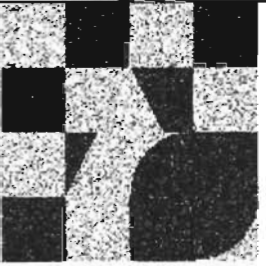
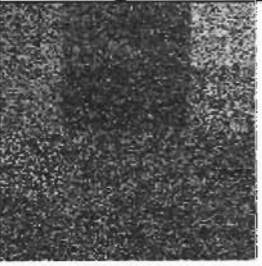
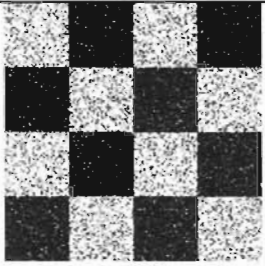


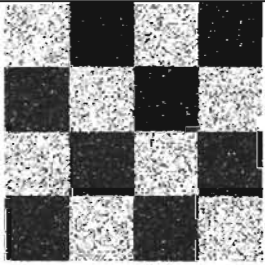
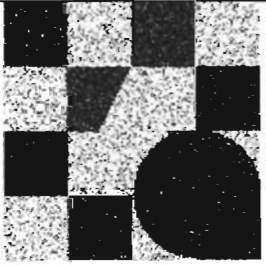
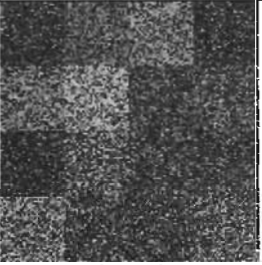
Componente principal	Imágenes		
	Imagen 1	Imagen 2	Imagen 3
Primera	 87.58%	 55.21%	 36.13%
Segunda	 6.30%	 32.68%	 34.49%
Tercera	 6.12%	 12.11%	 29.38%

Figura 4.3 Transformación en componentes principales de las imágenes a prueba donde se muestra su contribución a la covarianza global.

El análisis de los resultados se realizó utilizando la matriz de confusión. Estos resultados se muestran en la Tabla 4.3 y Tabla 4.4, la matriz de confusión para la segmentación en 16 etiquetas se muestra en el Apéndice 3.

Podemos observar en el caso de la imagen 1 en las tablas 4.3 y 4.4, que la variación del coeficiente  $\hat{\kappa}$  entre la segmentación utilizando 3 y 2 componentes principales es nula para el caso de ICM y para K-medias es de 0.0167 lo cual es despreciable, por lo que podemos decir que la dimensionalidad intrínseca es 2.



En el caso de las imágenes 2 y 3 resulta que la diferencia para el coeficiente  $\hat{\kappa}$  es significativa e incluso en el caso de sólo utilizar dos componentes principales en la imagen 2, se tiene que una región desaparece.

Por lo tanto la dimensionalidad intrínseca para las imágenes 2 y 3 es 3.

#### 4.3. Diferencias entre K-medias e ICM

Utilizando sólo las segmentaciones obtenidas utilizando las tres componentes principales podemos observar en la Tabla 4.2 las diferencias del coeficiente kappa entre la segmentación ICM y K-medias. Se puede observar que el coeficiente  $\hat{\kappa}$  es mayor para segmentación ICM en comparación con K-medias y por consecuente siempre hay una disminución del error.

El análisis de la disminución del error se realizó haciendo la diferencia del valor de los coeficientes  $\hat{\kappa}$  de las segmentaciones ICM y K-medias, y dividiendo entre el de K-medias, ver Tabla 4.2. La disminución del error está en el intervalo de 0.42% a 57.35%.

Imagen	$\hat{\kappa}$		$\hat{\kappa}_{ICM} - \hat{\kappa}_{K-medias}$	$(\hat{\kappa}_{ICM} - \hat{\kappa}_{K-medias}) / \hat{\kappa}_{ICM}$	Porcentaje (%)
	$\hat{\kappa}_{K-medias}$	$\hat{\kappa}_{ICM}$			
1	0.9958	1.0	0.0042	0.0042	0.42
2	0.9049	0.9944	0.0895	0.0900	9.00
3	0.1934	0.4535	0.2601	0.5735	57.35

Tabla 4.2 Disminución en error de la segmentación de K-medias respecto a ICM.

Hay que señalar que en el caso de la imagen 3 se tiene una disminución del error en un 57% lo que hace muy atractivo utilizar el algoritmo para imágenes a segmentar cuando se trata de un número considerable de regiones.

#### 4.4. Tiempos de procesamiento

El tiempo de procesamiento de los algoritmos de la transformación en componentes principales, K-medias e ICM para la imagen multiespectral de Pátzcuaro (6 bandas) que tiene un tamaño de 1500 por 1500 pixeles se muestran en la tabla 4.5.

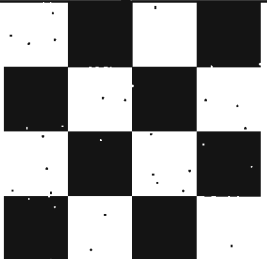
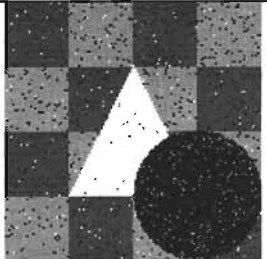
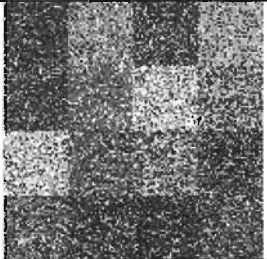
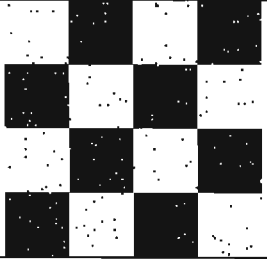
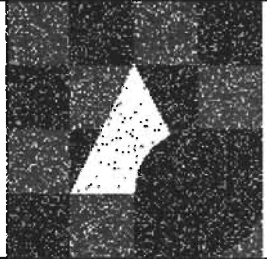
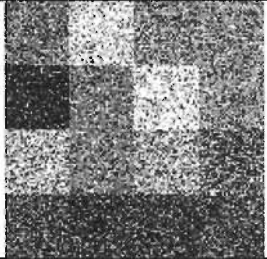
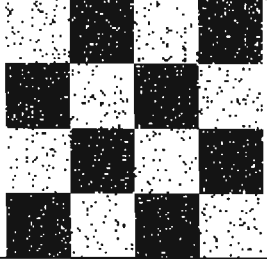
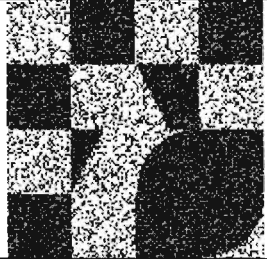
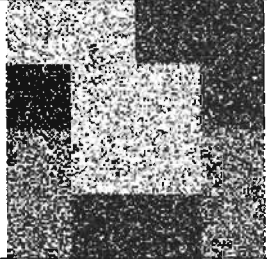
Componentes Principales	Imágenes		
	Imagen 1	Imagen 2	Imagen 3
1, 2, y 3			
$\hat{\kappa}$	0.9958	0.9049	0.1934
Matriz de confusión	$\begin{pmatrix} 8180 & 12 \\ 22 & 8170 \end{pmatrix}$	$\begin{pmatrix} 2892 & 156 & 158 & 2 \\ 302 & 5728 & 28 & 62 \\ 288 & 30 & 5160 & 55 \\ 0 & 9 & 7 & 1507 \end{pmatrix}$	
1 y 2			
$\hat{\kappa}$	0.9791	0.4591	0.0425
Matriz de confusión	$\begin{pmatrix} 8108 & 84 \\ 87 & 8105 \end{pmatrix}$	$\begin{pmatrix} 2527 & 35 & 620 & 26 \\ 35 & 4774 & 963 & 348 \\ 4390 & 41 & 1070 & 32 \\ 1 & 55 & 9 & 1458 \end{pmatrix}$	
1			
$\hat{\kappa}$	0.8948	0.2109	0.0152
Matriz de confusión	$\begin{pmatrix} 7738 & 454 \\ 408 & 7784 \end{pmatrix}$	$\begin{pmatrix} 2244 & 180 & 776 & 8 \\ 15 & 1468 & 256 & 4381 \\ 3887 & 291 & 1344 & 11 \\ 1 & 367 & 72 & 1083 \end{pmatrix}$	

Tabla 4.3 Segmentación K-medias de las componentes principales de las imágenes, utilizando tres, dos y una de ellas, se muestra debajo de cada imagen el valor del coeficiente kappa y la matriz de confusión que le corresponde

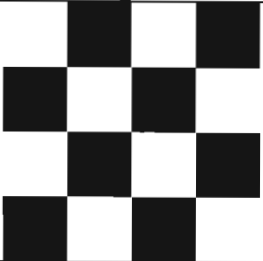
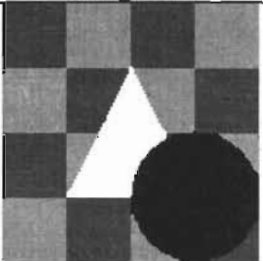
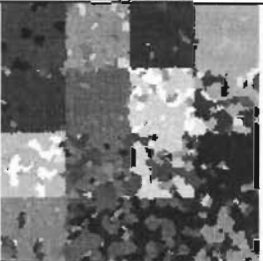
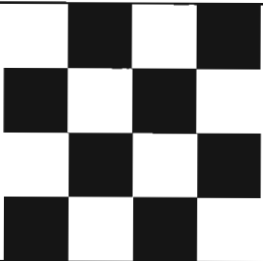

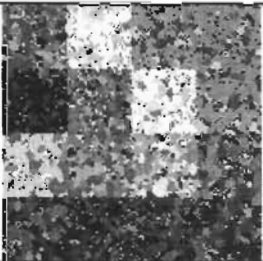
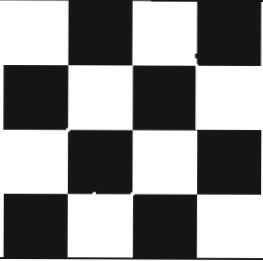
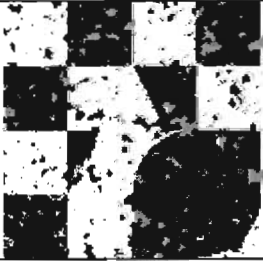
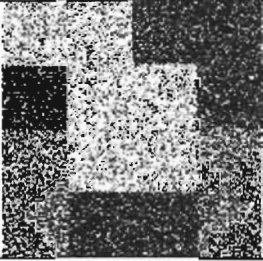
Componentes Principales	Imágenes		
	Imagen 1	Imagen 2	Imagen 3
1, 2, y 3			
$\hat{\kappa}$	1	0.9944	0.4535
Matriz de confusión	$\begin{pmatrix} 8192 & 0 \\ 0 & 8192 \end{pmatrix}$	$\begin{pmatrix} 3204 & 2 & 2 & 8 \\ 21 & 6097 & 0 & 2 \\ 28 & 0 & 5498 & 7 \\ 0 & 1 & 1 & 1521 \end{pmatrix}$	
1 y 2			
$\hat{\kappa}$	1	0.5453	0.0491
Matriz de confusión	$\begin{pmatrix} 8192 & 0 \\ 0 & 8192 \end{pmatrix}$	$\begin{pmatrix} 3198 & 4 & 1 & 5 \\ 13 & 6093 & 1 & 13 \\ 5514 & 4 & 1 & 14 \\ 0 & 1 & 0 & 1522 \end{pmatrix}$	
1			
$\hat{\kappa}$	0.9988	0.1846	0.0171
Matriz de confusión	$\begin{pmatrix} 8184 & 8 \\ 2 & 8190 \end{pmatrix}$	$\begin{pmatrix} 2817 & 73 & 315 & 3 \\ 14 & 508 & 87 & 5511 \\ 4736 & 98 & 687 & 12 \\ 2 & 190 & 21 & 1310 \end{pmatrix}$	

Tabla 4.4 Segmentación ICM de las componentes principales de las imágenes, utilizando tres, dos y una de ellas, se muestra debajo de cada imagen el valor del coeficiente kappa y la matriz de confusión que le corresponde.

Algoritmo	Tiempo de computo
Componentes principales	1 hora 33 minutos
K-medias	17 minutos
ICM	10 horas 47 minutos

Tabla 4.5 Tiempo de computo de los algoritmo de componentes principales, K-medias e ICM para la imagen de Pátzcuaro (tamaño 1500 por 1500 pixeles, 6 bandas) segmentada en 10 etiquetas.

# CAPÍTULO 5

## CONCLUSIONES

Las conclusiones del presente trabajo se enumeran a continuación:

1. El algoritmo ICM resulta atractivo para realizar la segmentación de imágenes multiespectrales dado que se obtienen buenos resultados de manera cualitativa (regiones homogéneas) y cuantitativa, mejorando los resultados de entrada a veces hasta en un 57% según el coeficiente kappa de la matriz de confusión.
2. Converge en un número pequeño de iteraciones (6), lo cual hace que sea factible su utilización en imágenes de percepción remota que son demasiado grandes.
3. La complejidad del algoritmo es lineal respecto a la imagen de entrada, ya que realiza un número constante de comparaciones de cada pixel con sus vecinos (orden de la vecindad) más otro número constante de operaciones para determinar la clase a la que se debe asignar.

Los inconvenientes son:

1. Determinación del valor del parámetro  $\beta$ , ya que se tiene que hacer por pruebas para encontrar su valor, aunque se puede facilitar su obtención a través de una ventana de entrenamiento.
2. Determinación del número óptimo de regiones a segmentar la imagen, este número es difícil de determinar ya que depende de propósito que se tenga al realizar la segmentación y por otra parte, el conocimiento del tipo imagen empleada.

Sobre la evaluación, a través de las matrices de divergencia y Jeffries-Matusita, resulta inadecuada por su dependencia directa con la covarianza de las medias, ya que sólo considera la separabilidad entre las mismas y no considera la información espacial de la segmentación.

Sobre la utilización de la transformación en componentes principales resulta atractiva cuando las bandas no están fuertemente correlacionadas, ya que se disminuye el tiempo de cómputo considerablemente al utilizar sólo la dimensionalidad intrínseca.

Sobre la evaluación de la segmentación a través de la matriz de confusión ésta resulta difícil si no se tiene conocimiento de la imagen a segmentar, si queremos garantizar que nuestros resultados sean objetivos.

Las aportaciones de este trabajo es el haber realizado la implementación de un algoritmo de segmentación de imágenes multiespectral utilizando los campos aleatorios de Markov y hacer la evaluación de los resultados empleando la matriz de confusión.

Las perspectivas futuras son:

1. Se pueden obtener mejores resultados si se utiliza una segmentación supervisada de entrada (en lugar de K-medias), ya que los resultados de la segmentación ICM tiene una dependencia directa.
2. El algoritmo ICM se puede emplear sólo sobre los píxeles que presenten dispersos en la segmentación inicial, esto reduciría el tiempo de cómputo considerablemente.
3. Realizar la reducción de la dimensionalidad de la imagen multiespectral utilizando un análisis canónico, este es empleado en cuando las regiones están fuertemente correlacionadas

# APÉNDICE 1

## PROBABILIDAD Y VARIABLES ALEATORIAS

*La probabilidad es la misma guía de la vida.*  
Cicerón

La probabilidad de un evento cualquiera se define como la fracción obtenida al dividir el número de casos favorables al evento por el número total de casos igualmente probables, esta definición clásica es utilizada de manera general para resolver problemas físicos [Pa84].

Las matemáticas definen probabilidad de manera axiomática (axiomática de Kolmogorov) que permite operar y calcular la probabilidad para funciones que representan los eventos de un experimento.

Definición 1.1 Definimos espacio de muestreo  $\Omega$  como el conjunto de resultados posibles de un experimento.

Definición 1.2 Se llama evento aleatorio  $A$ , a cualquier subconjunto del espacio muestra.

Definición 1.3 El evento complementario  $A^c = \Omega - A$ , evento unión  $A \cup B$ , el evento intersección  $A \cap B$ . Si  $A \cap B = \emptyset$  entonces decimos que son mutuamente excluyentes.

Definición 1.4 La probabilidad  $P$  de un evento cualquiera  $A$  en  $\Omega$  está dada por:

1.  $P : A \rightarrow [0,1] \subset \mathfrak{R}$
2.  $P(\Omega) = 1$
3. Si  $A, B \in \Omega, A \cap B = \emptyset \Rightarrow P(A \cup B) = P(A) + P(B)$

Definición 1.5  $\mathcal{B}$  es un campo de Borel si los conjuntos  $\{A_i\}_{i=1}^n$  pertenecen al campo entonces

$$\bigcup_{i=1}^n A_i \text{ y } \bigcap_{i=1}^n A_i$$

también le pertenecen.

Dadas la definiciones de espacio muestra, campo de Borel y de probabilidad podemos definir un experimento.

Definición 1.7 Un experimento  $E=(\Omega, B, P)$  esta definido por:

1. Un espacio muestra  $\Omega$ .
2. Un campo de Borel  $B$  formado por subconjuntos de  $\Omega$  llamados eventos.
3. Un número  $P(A)$  asignado a cada evento  $A$  de  $\Omega$ .

Una variable aleatoria es una función que asigna a cada evento un número  $X(A)$  para cada evento  $A$  en  $\Omega$ .

Definición 1.8 Una variable aleatoria  $X$  es una función cuyo dominio es el espacio muestral  $\Omega$  que asigna un número  $X(A)$  a cada evento  $A$  de  $\Omega$  de un experimento  $E=(\Omega, B, P)$  tal que :

1. El conjunto  $\{X < x\}$  contiene un evento para cada número  $x$ , en otras palabras  $x$  es medible.
2. La probabilidad de los eventos  $\{X = +\infty\}$  y  $\{X = -\infty\}$  son iguales a cero.

Hay dos tipos de variable aleatoria: continua y discreta.

Definición 1.9 Dada una variable aleatoria  $X$ , la función  $F(x) = P\{X \leq x\}$  se le llama función de distribución de  $X$  definida para todo número  $x \in (-\infty, \infty)$ . Esta función asigna a cada número real  $x$  la probabilidad acumulada hasta ese valor.

A partir de la definición anterior, podemos ver que  $F(-\infty) = 0$  y  $F(\infty) = 1$ .

Definición 1.10 Dada una variable aleatoria  $X$ , se le llama función densidad de probabilidad a la derivada de la función de distribución  $F(x)$

$$f(x) = \frac{dF(x)}{dx}$$



Dada la función densidad de probabilidad podemos definir la función de distribución como:

$$F(x) = P\{X \leq x\} = \int_{-\infty}^x f(x) dx$$

La función de distribución  $F(x)$  es monótona por consiguiente  $f(x) > 0$  y

$$\int_{-\infty}^{\infty} f(x) dx = F(\infty) - F(-\infty) = 1$$

Definición 1.11 El valor esperado de una variable aleatoria  $X$  es la integral

$$E[X] = \int_{-\infty}^{\infty} xf(x) dx$$

donde  $f(x)$  es la función de densidad de  $X$ .

Definición 1.12 La varianza de una variable aleatoria con media  $\mu$  esta dada por:

$$\sigma^2 = E[(X - \mu)^2] = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx$$

La raíz cuadrada de  $\sigma^2$  es la desviación estándar.

Definición 1.13 Dada una variable aleatoria  $X$ , el momento  $m_k$  esta dado por:

$$m_k = E[x^k] = \int_{-\infty}^{\infty} x^k f(x) dx$$

Definición 1.14 La constante

$$\eta_k = E[(x - \mu)^k] = \int_{-\infty}^{\infty} (x - \mu)^k f(x) dx$$

es llamado el momento central.

Hasta aquí se ha definido los momentos centrales para variables aleatorias continuas, en el caso discreto se puede definir de la misma forma consultar la bibliografía recomendada [Pa 84].

Definición 2.15 La distribución conjunta de la función de variables aleatorias  $X$  y  $Y$  es definida por:

$$F_{xy}(x, y) = P\{X \leq x, Y \leq y\}$$

Definición 1.16 Sea  $F_{xy}(x, y)$  parcialmente diferenciable de segundo orden, la función densidad de probabilidad conjunta es dada por:

$$f_{xy}(x,y) = \frac{\partial^2 F_{xy}(x,y)}{\partial x \partial y}$$

Definición 1.17 La covarianza de dos variables aleatorias esta definida por:

$$\text{cov}_{xy}(x,y) = E[(X - \mu_x)(Y - \mu_y)]$$

y el cociente

$$r = \frac{E[(X - \mu_x)(Y - \mu_y)]}{\sqrt{E[(X - \mu_x)^2 (Y - \mu_y)^2]}} = \frac{\text{cov}_{xy}}{\sigma_x \sigma_y}$$

es llamado coeficiente de correlación.

Dos variables aleatorias  $X$  y  $Y$  son independientes si los eventos  $\{X < x\}$  y  $\{Y < y\}$  son independientes. Esto hace  $F_{xy}(x,y) = F_x(x) F_y(y)$  y  $f_{xy}(x,y) = f_x(x) f_y(y)$ . Por lo tanto,  $X$  y  $Y$  no están correlacionadas si  $E[XY] = E[X] E[Y]$  y son ortogonales si  $E[XY] = 0$ .

La independencia es una condición más fuerte que la no correlación.

La probabilidad es un número que satisface los axiomas y la manera como se asignan de la definición empírica enfoque de frecuencia de eventos, que nos dice que la probabilidad de un evento  $A$  es hallar por conteo del número total de resultados de un experimento  $N$  y el número de ocurrencias del evento  $A$  en el experimento  $N_A$  entonces:

$$P(A) = \frac{N_A}{N}$$

Esta es una buena definición en las aplicaciones físicas en un experimento y también puede aplicarse al tratamiento de imágenes a las cuales asignamos probabilidades con proposiciones tales como "La imagen  $X$  es la mejor descripción de las observaciones que  $Y$ ".

El concepto básico de la teoría de probabilidad Bayesiana es que todas las probabilidades son condicionales.

Def. 1.18 Dado un evento  $C$  con probabilidad diferente de cero, la probabilidad condicional de  $A$  dado  $C$  es definida por:

$$P(A/C) = \frac{P(A \cap C)}{P(C)}$$

Si A y C son mutuamente excluyentes entonces  $P(A/C)=0$  y si  $A \subset C$  entonces  $P(A \cap C) = P(A)$ .

Teorema 1.1 (Probabilidad completa) Dado n excluyentes eventos  $A_1, A_2, \dots, A_n$  cuya unión integran el espacio muestral, es decir,

$$A_i \cap A_j = \emptyset \quad \forall i \neq j, i=1,2,\dots,n$$

y

$$\bigcup_{i=1}^n A_i = \Omega$$

Entonces para cualquier evento B en espacio muestral  $\Omega$ :

$$P(B) = \sum_{i=1}^n P(B / A_i)P(A_i)$$

Teorema 1.2 (Bayes)

$$P(A / B) = \frac{P(B / A)P(A)}{P(B)}$$

La probabilidad  $P(A/B)$  es llamada probabilidad a posteriori y  $P(A)$  es la probabilidad a priori de A.

La distribución normal ó gaussiana es la distribución de probabilidad más usada en el procesamiento digital de imágenes.

Def. 1.19 (Distribución Normal) Una variable aleatoria es normalmente distribuida si su función densidad es una curva gaussiana

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\left(\frac{(x-\mu)^2}{2\sigma^2}\right)}$$

donde  $\mu$  es la media y  $\sigma$  es la desviación estándar.

En el caso n-dimensional, la función densidad de probabilidad esta expresada por

$$f(x_1, x_2, \dots, x_n) = f(\vec{x}) = \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} e^{-\left(\frac{1}{2}(\vec{x}-\vec{\mu})\Sigma^{-1}(\vec{x}-\vec{\mu})^T\right)}$$

donde  $\vec{\mu}$  es el vector de la media y  $\vec{\Sigma}$  es la matriz de covarianza dados por:

$$\Sigma = \begin{pmatrix} \sigma_{11} & \cdots & \sigma_{1n} \\ \vdots & \ddots & \vdots \\ \sigma_{n1} & \cdots & \sigma_{nn} \end{pmatrix}, \bar{\mu} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}$$

Si la variable aleatoria  $X_i$  no está correlacionada entonces su matriz de covarianza es diagonal y su función de densidad de probabilidad puede ser factorizada:

$$f(x_1, x_2, \dots, x_n) = f(x_1)f(x_2)\cdots f(x_n)$$

## Procesos estocásticos y cadena de Markov

Def. 1.20 Dado un experimento  $E = (\Omega, B, P)$ . Para cada evento  $\xi \in \Omega$ , se asigna la función  $X(t, \xi)$ .

La familia de éstas funciones se denomina proceso estocástico.

Observamos que si  $t$  es constante entonces  $X(t, \xi)$  es una variable aleatoria.

Def. 1.21 Un proceso es estrictamente estacionario si la probabilidad no cambia con el tiempo. En otras palabras, el proceso  $X(t)$  y  $X(t + \epsilon)$  tiene la misma estadística para todo  $\epsilon$ .

Def. 1.22 Un proceso estocástico  $X(t)$  es llamado proceso de Markov si para cada  $n$  y  $t_1 < t_2 < \dots < t_n$  se tiene

$$P\{X(t_n) \leq x_n / X(t_{n-1}), \dots, X(t_1)\} = P\{X(t_n) \leq x_n / X(t_{n-1})\}$$

Def. 1.23 Un proceso de Markov con un espacio discreto de estados es llamado Cadena de Markov. En otras palabras, sea  $\{X_i\} = X_1, X_2, \dots, X_n, \dots$  una secuencia de variables aleatorias que toman los valores  $a_1, a_2, \dots, a_n$ . Si el proceso de Markov satisface:

$$P\{X(t_n) = a_{i_n} / X_{n-1} = a_{i_{n-1}}, \dots, X_1 = a_{i_1}\} = P\{X_n = a_{i_n} / X_{n-1} = a_{i_{n-1}}\}$$

entonces  $\{X_i\}$  es una cadena de Markov.

# APÉNDICE 2

## ALGORITMOS EN MATLAB 6.5

### 2.1. Componentes Principales

Esta función recibe como argumento la matriz A de la imagen a calcular sus componentes principales y regresa en la matriz B las componentes principales de la imagen.

```
function[B] = tkI(A)
[nr, nc, nb]=size(A);
mu=zeros(1, nb);
sigma=zeros(1, nb);
for i=1:nb
    mu(i)=mean2(A(:,:,i)); % Calcula las medias de cada banda
    sigma(i)=std2(A(:,:,i)); % Calcula la desviación estándar
end;
mu, sigma
tam=nr*nc;
mcov=zeros(nb,nb);
for i=1:nb
    for j=i:nb % calculamos la matriz de covarianza
        mcov(i,j)=(sum(sum((A(:,:,j)-mu(j)).*(A(:,:,i)-mu(i)))))/tam;
    end;
end;
for j=1:nb-1
    for i=j+1:nb
        mcov(i,j)=mcov(j,i);
    end;
end;
[w,l]=eig(mcov) % Se calcula los eigenvalores a partir de la matriz de covarianza
```

```

w=-1*rot90(w',-1)
s=sum(sum(l));
l=rot90(l,2)/s
B=A;
for i=1:nr
    for j=1:nr
        B(i,j,:)=w*(squeeze(A(i,j,:))); % Se calculan las componentes principales
    end;
end;
B=uint8(abs(B));

```

## 2.2 K-medias e ICM

Este programa realiza la segmentación K-medias e ICM, hay que declarar el número de etiquetas, el número de iteraciones de ICM, el valor de beta máximo y mínimo, se necesita además el histograma de la imagen guardada en un archivo de texto y al finalizar da como resultados los valores de la matrices de divergencia y de Jeffries-Matussita, la segmentación K-medias e ICM.

```

clear all;
itera=100;
n=256;
filename_input='n16cr'; % Nombre de las imágenes de salida
k=16; % Numero de etiquetas
clc;
close all;
O=leer_imagen; % Llamada a la función que lee la imagen
A=double(O);
clear O;
[nr nc banda]=size(A)
ng=n;
means=(round((1:k)*ng/(k+1)))'*ones(1,banda) % Inicializa los valores de las medias
tam=11406; %Número de renglones del histograma
[fi, texto]= fopen('hist_n16cr_cp.txt', 'r'); %Nombre del histograma de la imagen

```

```

B=fscanf(fi, '%i %i %i %i', [4, tam]);
st=fclose(fi);
for j=1:itera % Esta parte corresponde a la segmentacion K-medias
pert=ones(1,tam);
centn=zeros(k, banda);
div=zeros(1,k);
for q=1:tam
    pixel=(B(1:banda,q));
    h=B(4,q);
    for t=1:k
        d(t)=sum((pixel-means(t,:)).^2);
    end;
    [v,p]=min(d);
    centn(p,:)=centn(p,:)+(pixel*h);
    div(p)=div(p)+h;
    pert(q)=p;
end;
for i=1:k
    if (div(i)~=0)
        centn(i,:)=(centn(i,:)/div(i));
    else
        centn(i,:)=randsrc(1,banda,[(1:n)]);
    end
    e(i)=sum((centn(i,:)-means(i,:)).^2);
end;
means=round(centn) % Redondea los valores de las medias
disp(e);
if e<=1
    break;
end;
end; % Aquí termina la parte de la segmentación K-medias
mcov=zeros(banda, banda, k); % Calcula la matriz de covarianza
for class=1:k

```

```

numclass=0;
nomb_arch=['class',int2str(class),'.txt'];
[fi, texto] = fopen(nomb_arch, 'w');
for i=1:tam
    if pert(i)==class
        for j=1:B(4,i)
            numclass=numclass+1;
            pixel=(B(1:banda,i));
            fprintf(fi, '%g %g %g \n', pixel);
        end;
    end;
end;
fclose(fi);
[fi, texto]= fopen(nomb_arch, 'r');
clase=fscanf(fi, '%i %i %i ', [3, numclass]);
fclose(fi);
clase=clase';
clase = clase - ones(numclass,1) * means(class,:);
mcov(:,:,class) = clase'*clase/numclass;
end;
mcov=round(mcov)
means=round(means);
for i=1:nr
    for j=1:nc
        for l=1:k
            for b=1:banda
                h(b)=((A(i,j,b) - means(l, b)).^2);
            end;
            d(l)=sum(h);
        end;
        [xmin, C(i,j)]=min(d);
    end;
end;

```



```

end;
max=zeros(1,5);
for i=1:k
    max(i)=mvnpdf(means(i,:), means(i,:), mcov(:,i));
end;
Ap=(ones(nr+2,nc+2,b))*123;
Ap(2:nr+1,2:nc+1,:)=A(:,,:);
Bp=ones(nr+2,nc+2)*123;
Bp(2:nr+1,2:nc+1)=C(:,,:);
imag=Bp;
bmin=1; % Valor mínimo de beta
bmax=10; % valor maximo de beta
itera=6;
pend=(bmax-bmin)/(itera-1);
ord=bmin-pend;
display('ya entree!!!');
for f=1:itera % Esta parte corresponde a la segmentación ICM
    display(f);
    b=pend*f+ord;
    display(b);
    %pact=0; pelec=0;
    for i=2:nr+1
        for j=2:nc+1
            G=squeeze(Ap(i,j,:));
            v=[imag(i,j-1) imag(i,j+1) imag(i-1,j) imag(i+1,j) imag(i-1,j+1) imag(i+1,j-1) imag(i-1,j-1)
            imag(i+1,j+1)]; % Sistema de vecindad
            % Sistema de vecindad de orden 2
            s=zeros(1,k);
            for tt=1:k
                for t=1:8
                    if v(t) == tt
                        s(tt)=s(tt)+1;
                    else

```

```

        s(tt)=s(tt)-1;
    end;
end;
end;
pp=exp(b*s);
for t=1:k
    Proba(t)=pp(t)*((mvnpdf(G, means{t,:}, mcov{:,t}))/max{t}); %sum(t);
end; %Calcula la probabilidad multivariable del pixel considerado
[xmax, clase]=max(Proba);
Bp(i,j)=clase;
end
end
imag=Bp;
end
Ap=imag(2:nr+1,2:nc+1);
means
mcov
ddiv_km=DIV(means, mcov,k) % Llamada a la función que obtiene la matriz de divergencia
djm_km=JM(means,mcov,k)% Lamada a la función que obtiene la matriz de Jeffries-Matusita
djm_pkm=0;
ddiv_pkm=0;
p_km=div/(nr*nc);
for i=1:k
    for j=1:k
        djm_pkm=djm_pkm+p_km(i)*p_km(j)*djm_km(i,j);
        ddiv_pkm=ddiv_pkm+p_km(i)*p_km(j)*ddiv_km(i,j);
    end
end
end
djm_pkm
ddiv_pkm
[means_icm, mcov_icm, n_icm] = DMCOVm (A, Ap, k)
%Llamada a la función que obtiene la medias y la matriz de covarianza
means_icm

```

```

mcov_icm
ddiv_icm=DIV(means_icm, mcov_icm,k)
djm_icm=JM(means_icm,mcov_icm,k)
djm_picm=0;
ddiv_picm=0;
p_icm=n_icm/(nr*nc);
for i=1:k
    for j=1:k
        djm_picm=djm_picm+p_icm(i)*p_icm(j)*djm_icm(i,j);
        ddiv_picm=ddiv_picm+p_icm(i)*p_icm(j)*ddiv_icm(i,j);
    end
end
djm_picm
ddiv_picm
A=uint8(A);
imshow (C, gray(k));
title('IMAGEN CLASIFICADA KMEANS', 'Color', 'black');
axis off;
imwrite (C, gray(k), [filename_input,num2str(k),'_km.bmp'], 'bmp');
figure;
imshow(Ap,gray(k));
title('IMAGEN CLASIFICADA MARKOV','Color','black')
axis off;
imwrite (Ap, gray(k),
filename_input,'_k',num2str(k),'b',num2str(bmin),'a',num2str(bmax),'_icm.bmp' , 'bmp');

```

### 2.2.1 Función leer imagen

Esta función se encarga de leer la imagen de un archivo y cargarla a la matriz A.

```

function [A] = leer_imagen
resp=input('La imagen esta en RGB (S/N)?');
if (resp == 'S') | (resp == 's')

```

```

nomb_archivo = input ('¿Dame el nombre del archivo?');
info = imfinfo(nomb_archivo)
A = imread(nomb_archivo); %Lectura de imagen
A =(double(A))+1;
else
num_bandas = input ('¿Dame el numero de bandas?');
for i=1:num_bandas
    resp = input (['Nombre del archivo de la banda', int2str(i), '?']);
    A(:, :, i)=imread (resp);
end;
A=double(A)+1;

end;

```

### 2.2.2 Función divergencia

Esta función se realiza el cálculo de la matriz de divergencia a partir de los valores de las medias, la matriz de covarianza y el número de etiqueta.

```

function[D] = DIV(means,mcov,k)
D=zeros(k,k);
for i=1:k
    for j=1:k
        D(i,j)=0.5*trace((mcov(:, :, i)-mcov(:, :, j))*(mcov(:, :, i)^-1-mcov(:, :, j)^-1))+0.5*trace((mcov(:, :, i)^-1+mcov(:, :, j)^-1)*(means(i, :)-means(j, :))*(means(i, :)-means(j, :)));
        D(i,j)=2*(1-exp(-D(i,j)/8));
    end;
end;

```

### 2.2.2 Función Jeffries-Matusita

Esta función se realiza el cálculo de la matriz de Jeffries-Matusita a partir de los valores de las medias, la matriz de covarianza y el número de etiquetas.

```

function[J] = JM(means,mcov,k)
J=zeros(k,k);
for i=1:k
    for j=1:k
        kapa=((1/8)*(means(i,:)-means(j,:))*((mcov(:,i)+mcov(:,j))/2)^-1*(means(i,:)-
means(j,:)))+(1/2)*(det(mcov(:,i)+mcov(:,j))^(1/2)/(det(mcov(:,i))^(1/2)*det(mcov(:,j))^(1/2))));
        J(i,j)=sqrt(2*(1-exp(-kapa)));
    end;
end;

```

### 2.2.3 Función de obtención de las medias y la matriz de covarianza

Esta función se realiza el cálculo de las medias y matriz de covarianza de las regiones de una imagen cargada en la matriz A, y segmentada en la matriz B en k etiquetas.

```

function[m,mc,ne] = DMCOV (A, B, k)
m=zeros(1,k);
mc=zeros(1,k);
n=zeros(1,k);
[nr,nc]=size(B);
for i=1:k
    n(i)=size(find(B==i),1);
end;
ne=n;
s=zeros(1,k);
for i=1:nr
    for j=1:nc
        s(B(i,j))=s(B(i,j))+A(i,j);
    end
end
m=s./n;
s2=zeros(1,k);

```

```

for i=1:nr
    for j=1:nc
        s2(B(i,j))=s2(B(i,j))+(A(i,j)-m(B(i,j)))^2;
    end
end
mc=s2./n;

```

## 2.2.4 Obtención del histograma

Este programa calcula el histograma de una imagen cargada en la matriz A.

```

A=leer_imagen;
[nr nc nb]=size(A)
tam=nr*nc;
m=zeros(tam, nb);
cont=0;
for i=1:nr
    for j=1:nc
        cont=cont+1;
        m(cont,:)=squeeze(A(i,j,:));
    end;
end;
clear A;
display('ya termine');
n=256;
hist=zeros(n,n,n);
for i=1:tam
    hist(m(i,1),m(i,2), m(i,3))=hist(m(i,1),m(i,2), m(i,3))+1;
end
clear m;
nomb_arch=input('¿Dame el nombre del archivo?');
[fi, texto]= fopen(nomb_arch, 'w');
for i=1:n

```

```

display(i);
for j=1:n
    for k=1:n
        if hist(i,j,k) ~= 0
            fprintf(fi, '%g %g %g %g\n', i,j,k,hist(i,j,k));
        end
    end
end
end
clear hist;
st=fclose(fi);

```

### 2.3 Matriz de confusión y coeficiente kappa

Esta función calcula la matriz de covarianza a partir de la imagen segmentada (S), la imagen segmentada supervisadamente (O) y el número de etiquetas k.

```

function[m] = MC(k, O, S)
[nr, nc]=size(O);
m=zeros(k,k);
for i=1:nr
    for j=1:nc
        m(O(i,j),S(i,j))=m(O(i,j),S(i,j))+1;
    end
end
end

```

```

function [kp]=kappa(m)
[k,x]=size(m);
n=sum(sum(m));
d=0;
for i=1:k
    d=m(i,i)+d;
end
end

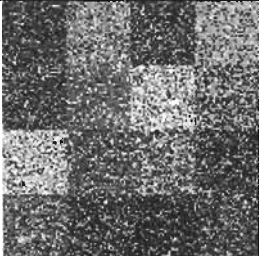
```

```
q=0;
for i=1:k
    h=0;
    for j=1:k
        h=m(i,j)+h;
    end
    g=0;
    for l=1:k
        g=m(l,i)+g;
    end
    q=h*g+q;
end
kp=(n*d-q)/(n*n-q);
```

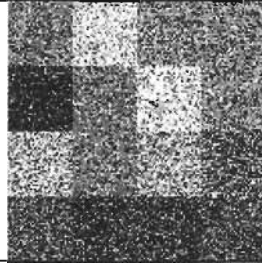


# APÉNDICE 3

## MATRIZ DE CONFUSIÓN PARA LA IMAGEN 3

Segmentación K-medias en 16 etiquetas de las componentes 1, 2 y 3 de la imagen 3																
																
$\hat{\kappa} = 0.1934$																
Matriz de confusión																
182	172	38	88	11	183	36	34	77	182	172	38	88	11	183	36	
258	103	113	136	76	42	99	10	43	258	103	113	136	76	42	99	
95	123	114	13	105	127	35	140	14	95	123	114	13	105	127	35	
79	27	15	469	6	8	70	0	140	79	27	15	469	6	8	70	
41	25	136	12	417	5	52	47	3	41	25	136	12	417	5	52	
63	169	20	10	6	475	8	79	33	63	169	20	10	6	475	8	
180	66	140	82	158	27	111	23	23	180	66	140	82	158	27	111	
4	27	16	0	55	94	9	493	2	4	27	16	0	55	94	9	
20	86	32	2	24	297	9	282	8	20	86	32	2	24	297	9	
25	43	57	5	193	42	33	236	6	25	43	57	5	193	42	33	
6	26	3	59	1	90	8	18	137	6	26	3	59	1	90	8	
7	6	20	75	71	2	59	3	9	7	6	20	75	71	2	59	
48	14	18	272	39	4	84	2	61	48	14	18	272	39	4	84	
26	77	12	28	11	220	20	96	65	26	77	12	28	11	220	20	
1	8	6	40	48	6	29	50	18	1	8	6	40	48	6	29	
182	172	38	88	11	183	36	34	77	182	172	38	88	11	183	36	

Segmentación K-medias en 16 etiquetas de las componentes 1 y 2 de la imagen 3

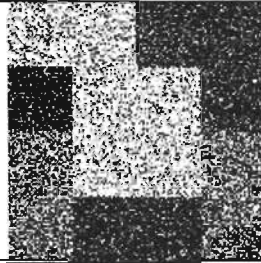


$\kappa = 0.0425$

Matriz de confusión

121	44	106	56	50	41	117	42	115	121	44	106	56	50	41	117
265	96	93	118	48	103	56	51	24	265	96	93	118	48	103	56
190	47	171	32	48	32	133	28	154	190	47	171	32	48	32	133
25	35	11	76	23	143	9	43	2	25	35	11	76	23	143	9
404	94	131	92	44	58	68	23	21	404	94	131	92	44	58	68
38	13	88	16	24	6	142	17	384	38	13	88	16	24	6	142
273	110	99	141	36	89	63	23	20	273	110	99	141	36	89	63
38	8	82	7	20	3	118	16	432	38	8	82	7	20	3	118
35	13	68	6	32	4	127	7	401	35	13	68	6	32	4	127
159	33	152	46	48	28	143	43	157	159	33	152	46	48	28	143
3	2	3	5	13	9	21	24	21	3	2	3	5	13	9	21
28	41	13	61	20	125	9	40	1	28	41	13	61	20	125	9
16	48	12	72	24	132	8	41	0	16	48	12	72	24	132	8
16	9	34	19	37	14	64	29	140	16	9	34	19	37	14	64
13	19	15	24	23	48	21	55	13	13	19	15	24	23	48	21
1	4	6	6	10	20	15	30	15	1	4	6	6	10	20	15

Segmentación K-medias en 16 etiquetas de la componentes 1 de la imagen 3

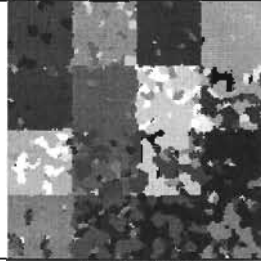


$\hat{\kappa}=0.0152$

Matriz de confusión

199	59	23	42	51	31	36	48	49	199	59	23	42	51	31	36
565	48	49	34	38	37	44	29	37	565	48	49	34	38	37	44
205	48	51	52	28	56	56	45	36	205	48	51	52	28	56	56
539	62	57	32	40	41	39	36	32	539	62	57	32	40	41	39
561	66	47	41	38	34	34	37	31	561	66	47	41	38	34	34
35	10	29	17	19	23	26	29	23	35	10	29	17	19	23	26
569	61	52	32	34	45	29	30	25	569	61	52	32	34	45	29
26	15	17	13	18	20	28	23	28	26	15	17	13	18	20	28
31	14	13	9	15	20	28	26	26	31	14	13	9	15	20	28
184	54	42	40	48	48	52	38	48	184	54	42	40	48	48	52
31	23	21	10	16	15	21	34	31	31	23	21	10	16	15	21
548	78	46	46	40	45	23	41	26	548	78	46	46	40	45	23
562	57	51	46	39	36	30	26	35	562	57	51	46	39	36	30
33	17	22	11	19	19	23	27	32	33	17	22	11	19	19	23
178	46	49	41	40	37	46	28	55	178	46	49	41	40	37	46
29	15	18	11	15	21	20	18	23	29	15	18	11	15	21	20

Segmentación ICM en 16 etiquetas de las componentes 1, 2 y 3 de la imagen 3



$\hat{\kappa}=0.4535$

Matriz de confusión

54	812	0	13	0	61	7	0	39	54	812	0	13	0	61	7
365	153	80	42	4	3	366	0	0	365	153	80	42	4	3	366
1	362	143	0	6	39	45	14	1	1	362	143	0	6	39	45
0	0	0	973	0	0	23	0	22	0	0	0	973	0	0	23
0	9	40	0	891	2	37	0	0	0	9	40	0	891	2	37
0	149	0	0	0	867	0	0	3	0	149	0	0	0	867	0
93	43	186	14	44	1	546	1	1	93	43	186	14	44	1	546
0	0	0	0	1	0	0	972	0	0	0	0	0	1	0	0
1	48	1	0	0	370	8	498	5	1	48	1	0	0	370	8
0	3	11	0	60	0	4	71	0	0	3	11	0	60	0	4
0	0	0	2	0	9	1	6	100	0	0	0	2	0	9	1
0	1	0	5	0	0	0	0	0	0	1	0	5	0	0	0
0	13	0	352	0	0	97	0	3	0	13	0	352	0	0	97
1	66	1	0	0	342	4	7	14	1	66	1	0	0	342	4
0	1	0	9	0	0	1	7	0	0	1	0	9	0	0	1
0	0	0	9	0	2	0	9	4	0	0	0	9	0	2	0

Segmentación ICM en 16 etiquetas de las componentes 1 y 2 de la imagen 3

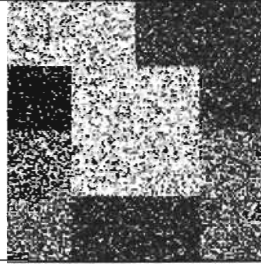


$\hat{\kappa}=0.0491$

Matriz de confusión

30	14	129	12	174	17	205	122	8	30	14	129	12	174	17	205
157	97	104	116	131	189	33	128	4	157	97	104	116	131	189	33
31	17	459	4	153	13	195	10	14	31	17	459	4	153	13	195
6	22	5	36	44	162	3	27	2	6	22	5	36	44	162	3
457	97	99	80	88	50	39	43	5	457	97	99	80	88	50	39
6	9	78	1	78	4	286	7	338	6	9	78	1	78	4	286
158	141	98	265	52	146	48	42	9	158	141	98	265	52	146	48
8	3	67	3	27	1	196	11	422	8	3	67	3	27	1	196
8	4	47	5	25	2	222	13	334	8	4	47	5	25	2	222
36	11	305	10	165	12	289	57	11	36	11	305	10	165	12	289
1	1	2	3	10	1	6	6	11	1	1	2	3	10	1	6
13	22	11	18	29	91	4	66	0	13	22	11	18	29	91	4
3	32	5	29	24	133	6	47	0	3	32	5	29	24	133	6
5	5	7	4	68	3	39	4	35	5	5	7	4	68	3	39
5	2	13	0	24	2	3	173	4	5	2	13	0	24	2	3
0	3	3	2	12	2	3	16	3	0	3	3	2	12	2	3

Segmentación ICM en 16 etiquetas de la componente 1 de la imagen 3



$\hat{\kappa}=0.0171$

Matriz de confusión

169	96	23	39	43	34	34	56	48	169	96	23	39	43	34	34
569	52	46	32	34	34	52	30	34	569	52	46	32	34	34	52
187	66	57	45	26	57	57	45	36	187	66	57	45	26	57	57
545	68	51	31	37	37	38	39	28	545	68	51	31	37	37	38
574	57	50	32	38	34	36	38	23	574	57	50	32	38	34	36
29	16	31	14	15	29	24	32	20	29	16	31	14	15	29	24
582	50	58	25	27	59	22	27	29	582	50	58	25	27	59	22
19	21	17	12	20	18	29	26	28	19	21	17	12	20	18	29
25	19	15	8	13	25	26	26	25	25	19	15	8	13	25	26
157	86	43	36	46	45	60	30	59	157	86	43	36	46	45	60
28	27	20	9	16	16	19	32	34	28	27	20	9	16	16	19
543	97	34	52	29	52	18	44	20	543	97	34	52	29	52	18
567	58	48	43	38	40	26	22	40	567	58	48	43	38	40	26
30	21	22	12	17	19	22	28	30	30	21	22	12	17	19	22
159	65	54	30	45	35	54	25	53	159	65	54	30	45	35	54
27	18	19	9	14	21	22	15	23	27	18	19	9	14	21	22

# BIBLIOGRAFÍA

[Ba 90] Baxter R. J. "Exactly Solved Models in Statistical Mechanics". Academic Press, 1990.

[Be 74] Besag J., "Spatial Interaction and the Statistical Analysis of Lattice Systems", Journal of the Royal Statistical Society, Ser. B, 36: 192-236, 1974.

[Be 86] Besag J., "On the statistical analysis of dirty pictures". Journal of the Royal Statistical Society. Serie B, 48(3):259-302, 1986.

[Co 95] Cortijo Bon F. J., "Un estudio comparativo de métodos de clasificación multibanda", Tesis Doctoral, Granada, Junio 1995.

[Co 89] Congalton Russell G., Green Kass, "Assessing the accuracy of remotely sensed data: principles and practices", Lewis publishers, 1998.

[De 82] Devliver P.A., Kittler J. V., "Pattern Recognition. A Statistical Approach", Prentice Hall-Englewood Cliffs, 1982.

[Ge 84] Geman D. and Geman S., "Stochastic Relaxation, Gibbs distributions and the bayesian restoration images", IEEE Transactions on Pattern and Machines Intelligence, 6, 721-741, 1984.

[Ka 93] Kato Z., Zerubia J., Berthod M., "Bayesian image classification using markov random field", in A. M.-D. G. Demoments, Editor, Maximum Entropy and Bayesian Methods, Kluwer Academic Publisher, 375-382, 1993.

[Ka 94] Kato Z., "Modelisations Markoviennes multiresolutions en vision par ordinateur. Application a la segmentation d'images SPOT" PhD. Thesis, University of Nice, 1994.

- [Ki 80] Kinderman R., Snell J. L. "Markov random fields and their applications", American Mathematics Society, 1:1-142, 1980.
- [Li 02] Lira Jorge, "Introducción al procesamiento digital de imágenes": IPN, Centro de Investigación en Computación, México, D. F. 2002.
- [Li 95] Li Stan Z., "Markov Random Field Modeling in Computer Vision", Springer Verlag 1995.
- [Lu 03] Luo J., Guo Cheng-en, "Perceptual grouping of segmented regions in color images", The Journal of the Pattern Recognition Society, 2781-2792, 36, 2003.
- [Mu 01] Murino V., "Reconstruction and segmentation of underwater acoustic images combining confidence information in MRF models", The Journal of the Pattern Recognition Society, 981-997, 34, 2001.
- [Pa 84] Papoulis A., "Probability random variables, and stochastic processes", Series in systems science. Mac-Graw\_hill Book Company, 1984.
- [Pr 93] Proesmans, M., Oosterlinck, A. "Parallel Segmentation Algorithms", Parallel Algorithms for Digital Image Processing, Computer Vision and Neural Networks ed. Pitas, I., John Wiley & Sons, pp. 123-146. 1993.
- [Pr 04] Provost J., Collet C., Rostaing P., Pérez P., Bouthemy P., "Hierarchical Markovian segmentation of multispectral images for the reconstruction of water depth maps", Computer Vision and Image Understanding, 155-174, 93, 2004.
- [Ri 99] Richards, John Alan, "Remote sensing digital image analysis : an introduction", Berlin, Springer Verlag, c1999
- [Sw 78] Swain P. H., Davis S. M: "Remote Sensing. The Quantitative Approach", Mc. Graw Hill 1978.



[St 02] Steven Phillips, "Reducing the computation time of the isodata and k-means unsupervised classification algorithms", Geoscience and Remote Sensing Symposium, 2002. IGARSS '02. 2002 IEEE International

[Wi 02] Wilson R., Li Chang-Tsun, "A class of discrete multiresolution random fields and its application to image segmentation", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 25, No. 1, January 2002.

[Wu 82] Wu F.Y. "The Potts model", Reviews of modern physics, 54(1):235-268, 1982.

[YA 99] Yamazaki Tatsuya, Gingras Denis, "Unsupervised multispectral Image classification using MRF models and VQ method", IEEE Transactions on Geoscience and Remote Sensing, Vol. 37, No. 2, march 1999.